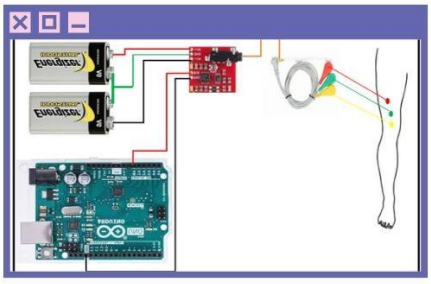
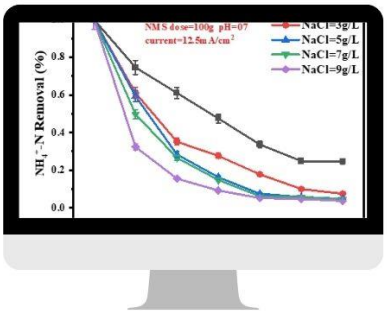
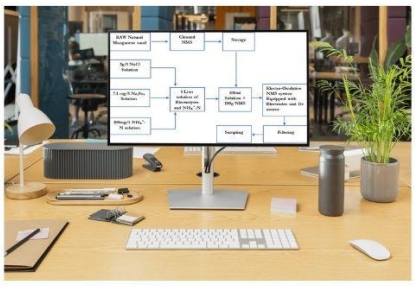
INTERNATIONAL CONFERENCE ON ENERGY, POWER, ENVIRONMENT, CONTROL, AND COMPUTING

ICEPECC 2025







IJIST

Conference Organizing Committee's Chairs

Dr. Muhammad Wasif
Technical Program Committee

Dr. Naeem Abas
Publication Committee

Dr. Ghulam Abbas
Exhibition & Sponsorship Committee

Dr. Sajjad Miran
Publicity & PR Committee

Dr. Shoaib Rauf
Local Arrangement Committee

Dr. Nazam Siddique
Finance Committee

Dr. M. Jehanzeb Irshad
Website Committee

Engr. Yumna Bilal
Registration Committee

A Heartfelt Thank You to the Organizing Committee

I would like to extend my deepest gratitude to **Prof. Dr. Shahid Iqbal** and **Dr. Muhammad Wasif** of the **University of Gujrat** for their remarkable support, generous hospitality, and unwavering dedication throughout the entire course of the conference. Prof. Dr. Shahid Iqbal's visionary leadership, intellectual guidance, and warm encouragement made a profound impact on the success of the event. His inspiring presence and thoughtful interactions added great value to the academic and collaborative spirit of the conference. Similarly, Dr. Muhammad Wasif's exceptional organizational skills, attention to detail, and sincere cooperation ensured that every aspect of the conference ran smoothly and efficiently. His friendly nature and tireless efforts to facilitate participants were truly commendable. Together, their commitment to promoting academic excellence and fostering a welcoming environment made the conference a memorable and enriching experience. I feel honored and thankful for their outstanding contributions and the gracious manner in which they represented the University of Gujrat.

With sincere appreciation,

IJIST Team

Table of Contents

International Journal of Innovations in Science & Technology (IJIST)
ISSN2618-1630 Special Issue | 2025

Current-Injected Mode Control for Coupled-Inductor (Ci) Based Boost Converter

Ansa Nawaz[1], Taosif Iqbal[1], Nabila Nawaz[2]
[1]Department of Electrical Engineering, College of Electrical & Mechanical Engineering (E&ME), Rawalpindi Pakistan
[2]Department of Electrical Engineering, Ghulam Ishaq Khan Institute of Engineering Sciences and Technology Pakistan
*Correspondence:ansanawaz96@gmail.com;taosifiqbal@ceme.nust.edu.pk; nabilakcp@gmail.com

Citation | Nawaz. A, Iqbal. T, Nawaz. N, "Current-Injected Mode Control for Coupled-Inductor (Ci) Based Boost Converter", IJIST, Special Issue pp 01-17, March 2025
Received| Feb 08, 2025 **Revised|** Feb 20, 2025 **Accepted|** March 25, 2025 **Published|** March 01, 2025.

With the increasing demand for electrical energy, there is a need to replace conventional energy resources with renewable energy resources. To properly implement renewable resources at a larger scale, DC/DC converters play a major role. Owing to the variable and unreliable nature of renewable energy resources like PV systems there is a requirement for converters that can regulate the voltage at the output side. High-gain DC/DC converters are preferred for the integration of the solar system in smart grids or microgrids. In this context, a high-gain boost converter utilizing a coupled inductor is a preferable choice. High gain is achieved by the proper selection of the turn's ratio of coupled inductors in such converters. Whereas to obtain voltage regulation there is a need to employ an effective control scheme. In this paper current-injected control topology has been utilized for coupled inductor-based boost converter. The proposed converter with an appropriate control scheme aims to achieve high voltage gain, reduced switching losses, minimization of current ripple, and less conduction losses while increasing the efficiency of the overall system. A small signal model based on the state space averaging technique is used to derive control to output transfer function for the proposed converter. A hardware prototype has been implemented for the validation of theoretical work. The overall efficiency of the converter is calculated to be around 96% at specific load conditions.

Keywords: Coupled Inductor, Non-Isolated Converter, Small Signal Modeling, Current Mode Control, High Gain Converter, Dc-Dc Boost Converter

Introduction:

Renewable energy resources have gained immense attention due to the depletion of conventional energy resources like fossil fuels. The increasing environmental issues have also made it necessary to look for resources that can provide clean energy. The growing trend towards industrialization, urbanization, and an increase in population requires sustainable and eco-friendly energy resources to meet the energy demand [1]. Renewable energy resources like wind, solar, or fuel cells are a source of sustainable energy that has a positive impact on the environment. Renewable energy resources can develop a distributed energy system/hybrid system that can provide decentralized energy [2]. The integration of renewable energy resources requires DC-DC converters that can efficiently boost the low terminal voltage.

Due to the fluctuating and low output voltage of RES, the role of DC-DC power electronic boost converters is significant [3]. The proper selection of such converters depends on various factors like the elements (size), efficiency, and voltage gain [4], [5]. Ripple in the input current, voltage and current stress on the switch also play an imperative role. To implement MPPT for PV the high ripple in the input current is problematic and undesirable [6]. In comparison to transformer-based isolated DC-DC converters, non-isolated DC-DC converters are preferable. Some of the issues associated with isolated converters include voltage stress on the switch, core saturation, and voltage spikes due to the leakage inductor of the transformer [7].

Conventional boost converters aren't compatible with applications that require high voltage. Although the voltage gain of a conventional boost converter can be increased by setting a high value of duty ratio. But then again, the higher value of the duty cycle causes conduction losses, voltage stress on the switch, reverse recovery issue of the diode, and large ripple in inductor current. It is uneconomical and impractical to use a conventional boost converter with a high value of duty ratio. The requirement of high rating switches requiring large ON-state resistance, the impact of ESR (equivalent series resistor) of the inductor, and electromagnetic interface limit the use of CBC [8],[9].

Various converters have been suggested that involve transformers to achieve high voltage gain [10],[11]. Isolated converters based on transformers can achieve high voltage with a low duty ratio, but these topologies are costly and require multistage power conversion. Voltage lift techniques have also been implemented [12]. Switched inductor-based converters can achieve high gain [13] but require a large count of components especially magnetic devices/inductors and a couple of diodes. Similarly, active network converters based on switched capacitors can also be used in high-voltage applications but require two switches along with the circuitry of the gate drive which increases overall cost [14]. Coupled inductor-based step-up converters are preferred to achieve the goal of high gain and low voltage stress [15] but the main problem is the high current ripple and no control for voltage regulation. Some topologies utilize coupled inductor to overcome the issue of voltage stress and current ripple but require large number of components [16],[17].

In this paper, a coupled inductor boost converter based on current mode control has been proposed to achieve high gain, efficiency, and voltage regulation. By properly adjusting the turns ratio of the coupled inductor high gain can be achieved but it causes a voltage spike and high ripple current. To overcome this problem current mode control CMC or current-injected control is implemented which is comprised of an outer loop (voltage control) and an inner loop (current control). The proposed converter aims to achieve the goals of high gain, high efficiency no reverse recovery issues, and less power/conduction losses. Boost converter has a problem of non-minimum phase that is due to the presence of zero in the right half s-plane. The voltage regulation is also an important factor for PV applications due to variations and unreliable terminal voltage. To deal with all such problems dual loop control is utilized. Current mode

control is usually preferred over-voltage mode control due to the slow response of VMC w.r.t load variations. The proposed converter aims to achieve the following goals:

1. High gain & high overall efficiency.
2. Minimum count of magnetizing and switching components.
3. Implantation of a simple and robust control structure to reduce current ripples.
4. Voltage regulation, stability, and reliability.

The main objective of this research is to develop a highly efficient, high-gain CI (coupled-inductor) based boost converter utilizing a current-injected control for voltage regulation and stability. By optimizing the turns ratio, the design minimizes switching losses and reduces duty cycle constraints, ensuring improved voltage conversion efficiency. A dual-loop feedback control mechanism is implemented, comprising an inner current loop for fast transient response and an outer voltage loop for precise regulation. This approach effectively suppresses voltage oscillations, electromagnetic interference (EMI), transient overshoots, and peak inductor currents, maintaining a steady output voltage despite input fluctuations. The proposed system is particularly beneficial for renewable energy sources like photovoltaic (PV) systems & fuel cells, where stable power conversion is essential.

This work presents a current-injected mode control for a coupled inductor boost converter, offering enhanced voltage gain, dynamic response, and noise immunity. Unlike conventional boost converters, this design reduces switching stress by achieving high gain at lower duty cycles. The dual-loop control strategy effectively addresses right-half-plane (RHP) zero instability, ensuring system stability. Additionally, a compensation-ramp technique is integrated to suppress subharmonic oscillations, enhancing phase margin and gain characteristics without additional passive components. The proposed converter is validated through MATLAB/Simulink simulations and hardware implementation, demonstrating superior efficiency, transient response, and total harmonic distortion (THD) reduction, making it ideal for renewable energy and electric vehicle (EV) charging applications.

The paper has been divided into different sections for the convenience of the reader. Section 2 includes the basic introduction of the coupled inductor and the working principle of the proposed converter. In section 3 small signal model based on the state space averaging technique is used to derive control of the output transfer function. Section 4 covers the implementation of a current-injected control scheme to achieve voltage regulation and stability. It also gives a brief comparison of voltage mode control with current-injected mode control. Section 5 defines the effectiveness of the proposed converter in comparison to relevant converters that have been proposed in recent years. Section 6 involves the simulation results obtained via Matlab/Simulink as well as the hardware results that validate the theoretical work and performance of the proposed converter. Section 7 is the last section of this paper which includes future recommendations and conclusions.

Methodology:

Basic Understanding & Working of the Proposed Converter:

The basic circuit diagram of the proposed converter is shown in Figure 1. The replacement of the coupled inductor with a boost converter helps to achieve high gain. There are 2 modes of operations based on the state (ON/OFF) of the switch (MOSFET) during the switching period. The converter is merely comprised of the coupled inductor at the input side, high rating MOSFET, diode, capacitor, and resistive load at the output. The subsections define detailed working requirements and modes of operation. The basic understanding of coupled inductors is also briefly defined. The converter has been operated in the continuous conduction mode for simplicity.

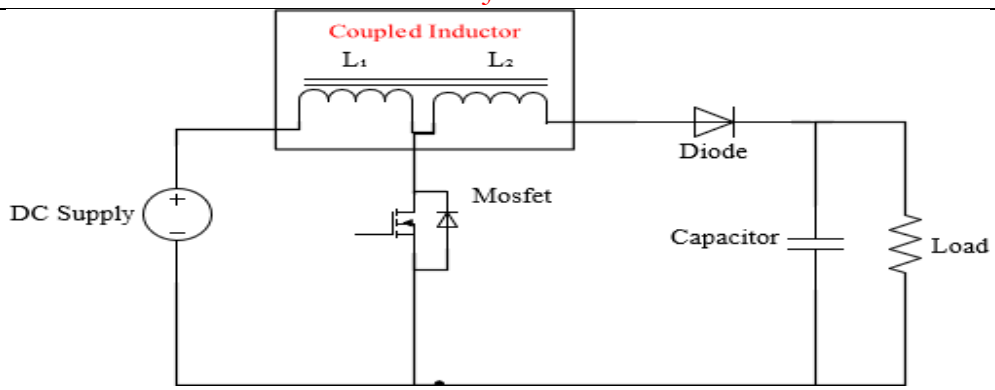


Figure 1. Coupled inductor-based step-up high gain boost converter [18]

Coupled inductors are used to achieve high gain without utilizing the high value of duty cycle i.e., duty cycle d >70%. On the other hand, wire resistance associated with the inductor introduces a limit on the selection of the duty cycle. In the case of a conventional boost converter increase in duty cycle to get high gain affects the overall efficiency. One option can be to use a wire of thick diameter, but it results in increased conduction losses with more current stress. The reverse recovery problem of the diode becomes predominant. At the same time reverse voltage appearing during the off state becomes larger in value and introduces conduction losses. Therefore, to overcome such problems coupled inductors are utilized. They simply have double winding at the same core instead of single winding. It helps to not only achieve high gain but also reduces conduction/switching losses with less voltage stress [18].

Operation modes (ON & OFF Position of MOSFET/Switch) for the Proposed Converter:

The two modes of operation have been shown with equivalent circuit diagrams for the ON and OFF states of the switch (MOSFET). In 1st mode of operation, as shown in Figure 2 (a), the switch (MOSFET) is in the ON state, and thus the 1st coil of the inductor stores energy through the source. In this state the diode gets open (reverse biased). While in the 2nd state of operation, as shown in Figure 2 (b), the diode gets forward-biased, and the switch (MOSFET) is in the OFF state. During this duration, the energy is delivered to the capacitor, and the load is connected in parallel via a diode which is conducting. In 2nd mode of operation, the 2 coils of the coupled inductor get in series thus same current flows through them.

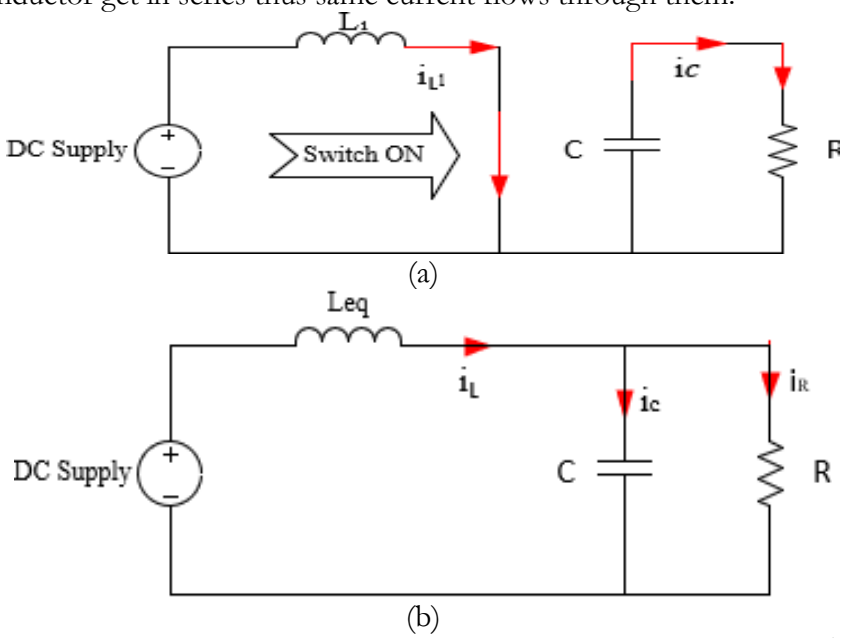


Figure 2 (a). Equivalent circuit for 1st mode (b) Equivalent circuit for 2nd mode

Mathematical Modeling Based on 2 Modes of Operation:

In 1st mode, energy gets stored through an input (DC) supply in L_1 when the switch (MOSFET) is in the ON state. In this state, the diode gets in a reversed biased state. The voltage applied across 1 coil eventually causes the voltage across the other coil that is linked magnetically to the 1st coil. Let the voltage across 1st coil be represented as v_{l1} and can be given as:

$$v_{l1} = L_1 . \frac{di_{l1}}{dt} \quad (1)$$

The term L_1 indicates the inductance related to 1st coil and is termed the self-inductance of that coil. The change in current causes a flux in the common core around which 2 coils are being wound. The flux can be represented as:

$$\Phi = i_{l1} * c * N_{l1} \quad (2)$$

Here, c is dependent on the core's geometry and magnetic properties whereas N_{l1} is no. of turns of 1st coil. It is to be noted that flux is contained and linked with the common core. The voltage and flux relation for 1st coil is given as [19]:

$$v_{l1} = N_{l1} . \frac{d\Phi_{l1}}{dt} \quad (3)$$

Similarly, for the 2nd coil the induced voltage or emf while considering that flux is the same in both the 1st as well as 2nd coil, the expression for induced voltage in the 2nd coil is:

$$v_{l2} = M_l . \frac{di_{l1}}{dt} \quad (4)$$

Here the term M_l Is the mutual inductance among the 2 coils wounded at the same core to form couple-inductor with the units same as inductance i.e., Henry- 'H'. The term c depends upon the core's properties. The voltage induced in 2nd coil can also be given as:

$$v_{l2} = n . v_{in} \quad (5)$$

In 2nd state of operation, the switch gets open (OFF state) and the diode gets forward biased. The voltage across 1st coil is:

$$v_{l1} = \frac{N_{l1} . (v_o - v_{in})}{N_{l1} + N_{l2}} \quad (6)$$

Whereas the equivalent inductance is represented by L_{equ} Is given as:

$$L_{equ} = L_1 + L_2 + (2 * M) \quad (7)$$

$$L_{equ} = L_1 . (1 + n)^2 \quad (8)$$

The relation between inductors L1, L2, and turns ratio can be expressed as

$$n^2 = (N_{l2}/N_{l1})^2 = L_2/L_1 \quad (9)$$

The concept of coupled inductors is quite simple as it utilizes a single core with 2 inductors wound on it. The turns ratio is selected according to the desired application and the connection of windings can either be in parallel or series as per requirement. The turn ratio can be higher depending on the design specifications. Mutual inductance also plays an important role in the overall performance of the coupled inductor. The value of k (coupling coefficient of the coupled inductor) is chosen between 0.95 & 0.99. Typically, its value is chosen as k=0.95. Where the coupling coefficient and mutual inductance are related as:

$$M = \left(\sqrt{L_1 . L_2}\right) . k \quad (10)$$

State Space Modeling:

Using the equivalent models of proposed converters based on ON as well as OFF states. State space-based modeling with matrix algebra makes it easy to understand the overall system design. It is usually based upon 1st order differential equations that model inductors and capacitor's current and voltage respectively. The calculations are based mainly on the time domain and the overall system is represented in the format of matrix.

The basic equation form for such a system is mainly given as:

$$\dot{x} = (A . x) + (B . u) \quad (11)$$

$$Y = C . x \quad (12)$$

1) When the switch is closed, as shown in Figure 3, then it has the following form.

$$\dot{x} = A_{01}x + B_{01}u \quad (13)$$

2) When the switch is open then:

$$\dot{x} = A_{02}x + B_{02}u \quad (14)$$

As per average modeling is concerned then by combining 2 modes as mentioned above, we get the following form:

$$\dot{x} = \overline{A_{av}}x + \overline{B_{av}}u \quad (15)$$

Whereas

$$\overline{A_{av}} = A_{01}.d + A_{02}(1-d) \quad (16)$$

Similarly

$$\overline{B_{av}} = B_{01}.d + B_{02}(1-d) \quad (17)$$

Here A represents dynamics for the state of the system and B for the controllable input.

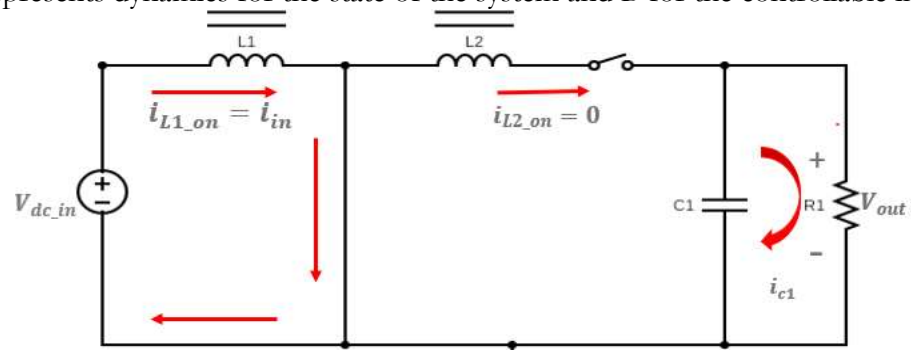


Figure 3. State 1 when the switch is ON

$$x_{1\_on} = i_{L1\_on} \quad (18)$$

$$x_{2\_on} = V_{C\_on} \quad (19)$$

$$\begin{bmatrix}\dot{x_1}\\ \dot{x_2}\end{bmatrix}. = \begin{bmatrix}0 & 0\\ 0 & -\frac{1}{C_1R_1}\end{bmatrix}.\begin{bmatrix}i_{L1\_on}\\ V_{C\_on}\end{bmatrix} + \begin{bmatrix}\frac{1}{L_1}\\ 0\end{bmatrix}.V_{dc_{in}} \quad (20)$$

$$A_{01} = \begin{bmatrix}0 & 0\\ 0 & -\frac{1}{C_1R_1}\end{bmatrix} \text{ and } B_{01} = \begin{bmatrix}\frac{1}{L_1}\\ 0\end{bmatrix} \quad (21)$$

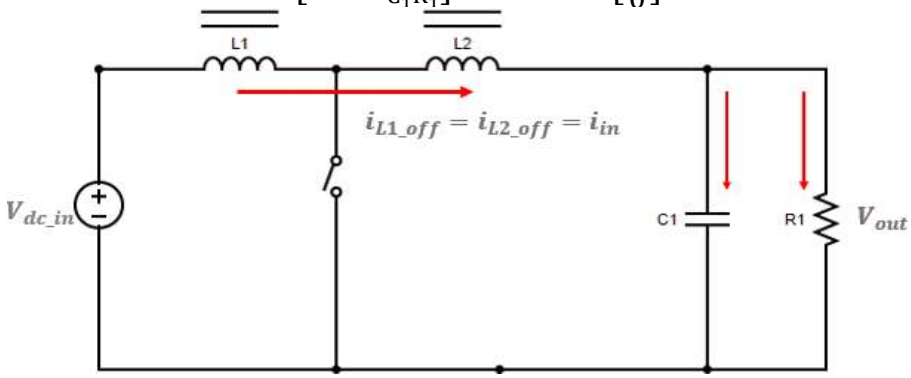


Figure 4. State 2 when the switch is OFF

The state when the switch (MOSFET) is in off state, as shown in Figure 4, then both coils are in series with each other.

$$\frac{dV_{C\_off}}{dt} = \frac{1}{C_1}\left(i_{L1\_off} - \frac{1}{R_1}(V_{C_{off}})\right) \quad (22)$$

Thus, in mode 2 we have:

$$\frac{di_{L1\_off}}{dt} = \frac{1}{L_{equ}}(V_{dc_{in}} - V_{C\_off}) \quad (23)$$

Here notice that L_{equ} Shows the series combination of both inductors.

The matrix form of the above differential's equations can be:

$$\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\end{bmatrix}.=\begin{bmatrix}0 & -\frac{1}{L_{equ}}\\ \frac{1}{C_1} & -\frac{1}{C_1R_1}\end{bmatrix}.\begin{bmatrix}i_{L1\_off}\\ V_{C\_off}\end{bmatrix}+\begin{bmatrix}\frac{1}{L_{equ}}\\ 0\end{bmatrix}.V_{dc_{in}} \quad (24)$$

$$A_{02}=\begin{bmatrix}0 & -\frac{1}{L_{equ}}\\ \frac{1}{C_1} & -\frac{1}{C_1R_1}\end{bmatrix} \text{ and } B_{02}=\begin{bmatrix}\frac{1}{L_{equ}}\\ 0\end{bmatrix} \quad (25)$$

Now final step is to combine the 2 states through the averaging method:

$$\dot{x}=\begin{bmatrix}0 & -\frac{(1-d)}{L_{equ}}\\ \frac{1-d}{C_1} & -\frac{1}{C_1R_1}\end{bmatrix}.x+\begin{bmatrix}\frac{1}{L_1}(d)+\frac{1}{L_{equ}}(1-d)\\ 0\end{bmatrix}.V_{dc_{in}} \quad (26)$$

For derivation of the transfer function i.e. control to output requires 2 main parameters that are output or capacitor voltage and inductor's current. Here the input side voltage is taken as disturbance whereas control d (duty cycle) is chosen. Therefore, the state-space model after some modifications is presented generically as:

$$\dot{x'}=\begin{bmatrix}i_{Le}\\ V_{Co}\end{bmatrix} \quad (27)$$

$$\tilde{u'}=\begin{bmatrix}V_{dc_{in}}'\\ d'\end{bmatrix} \quad (28)$$

$$A'=\begin{bmatrix}0 & -\frac{(1-d)}{L_{equ}(1+n)}\\ \frac{1-d}{C_1(1+n)} & -\frac{1}{C_1R_1}\end{bmatrix} B'=\begin{bmatrix}\frac{1+(n.d)}{L_{equ}(1+n)} & \frac{V_C'+n.V_{dc_{in}}'}{(1+n)}\\ 0 & -\frac{i_L'}{C_1(1+n)}\end{bmatrix} \quad (29)$$

The control-to-output transfer transfer-function is derived as [18]:

$$G_{V_{C\_d}}=\frac{V_C'}{d'}=\frac{\frac{i_L'.s}{(1+n)}+\frac{(1-d').(V_C'+n.V_{dc_{in}}')}{(n+1)^2}}{s^2+s+\frac{(1-d')}{(1+n)^2.R_1}} \quad (30)$$

The subharmonic oscillations play a major role in CMC implementation. A compensation ramp can eliminate this effect of subharmonic oscillations [20]. It is predominant especially when d>0.5 or 50% which means there is a major instability problem while dealing with current-injected control.

design equations for the compensation ramp slope are given below, ensuring stable operation. In the current injected mode-based coupled inductor boost converter configuration, the ramp compensation is essential to prevent sub-harmonic oscillations that arise when the duty cycle D is set above 50 %. A compensation ramp (Se) is added to the inductor current signal to dampen oscillations and stabilize the current loop. While the slop is introduced into the control signal to counteract duty cycle-dependent oscillations.

$$V_{out}=(1+n)\,V_{dc_{in}}/(1\text{-}d) \quad (31)$$

The required compensation ramp slope Se is:

$$S_e > (S_{off}-S_{on})/\ 32)$$

Where:

S_{on} = Slope of the on-time inductor current

S_{off} = Slope of the off-time inductor current

The Slopes for the on-time inductor current is given as:

$$S_{on}=V_{dc_{in}}/\ L_{equ} \quad (33)$$

The Slopes for the off-time inductor current are given as:

$$S_{off}=[(1+n)\,V_{dc_{in}}-V_{dc_{in}}]/\ L_{equ} \quad (34)$$

where L_{equ} is the equivalent inductance considering coupling effects

$$Se > \frac{\frac{(1+n)V_{dc_{in}} - V_{dc_{in}}}{Lequ} - \frac{V_{dc_{in}}}{Lequ}}{2} \quad (35)$$

After rearranging the final equation is:

$$Se > \frac{n-1}{2} \cdot \frac{V_{dc_{in}}}{Lequ} \quad (36)$$

The op-amp used for designing frequency-based compensation involves feedback to achieve stability. Type 2 and type 3 compensators are usually preferred for boost converters to implement voltage or current-injected control. Type 2 compensator has an RC network for getting better phase response and flatten gain. Type 3 compensator helps to improve the phase response even further with the addition of 2 poles,2 zeros, and a pole at zero location [21]. However, the relevant transfer function derived using the simplified block diagram of current-injected control is given as [22]:

$$T.F_{(VC)} = \frac{G_V.K_m}{K_m.[(G_V.K_{vd})+G_i]} \quad (37)$$

Results and Discussion:

Implementation of the Proposed Converter with Voltage Mode Control:

Voltage Mode Control (VMC) is a feedback control technique used in power converters to regulate output voltage by comparing the feedback voltage with a reference value as shown in Figure 5. The error signal generated is used to control the duty cycle of a MOSFET via PWM modulation. While VMC is useful in applications with input or load voltage variations, light loads, or when avoiding the complexities of dual-loop control, it faces challenges due to the right-half-plane zero, which complicates loop compensation, especially in boost converters. As a result, VMC is often less effective than Current Mode Control (CMC), which provides more precise control by using the inductor current as a state variable and is preferred in applications requiring fast dynamic response, current sharing, or noise immunity. CMC also offers better stability and efficiency in high-power and multi-phase converters, making it the favored choice in many designs. The voltage mode-based control was implemented on Matlab/Simulink and the output voltage as well as current waveforms are shown in Figure 6.

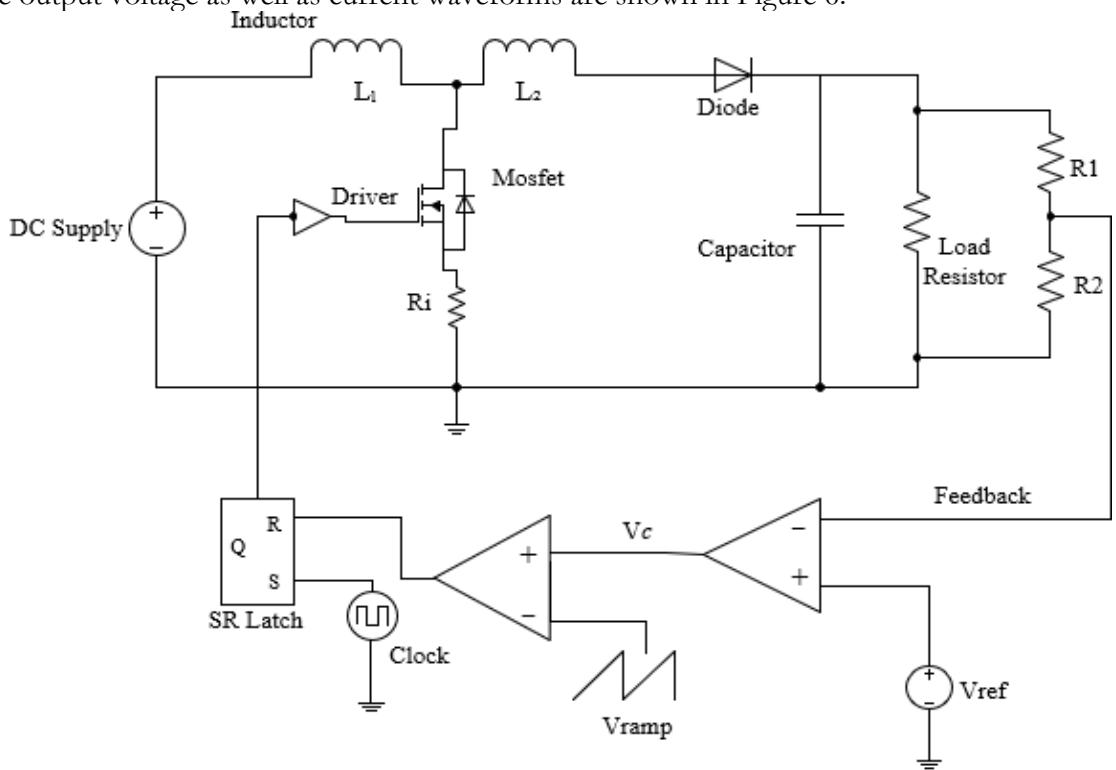


Figure 5. Circuit diagram for proposed converter, utilizing VMC

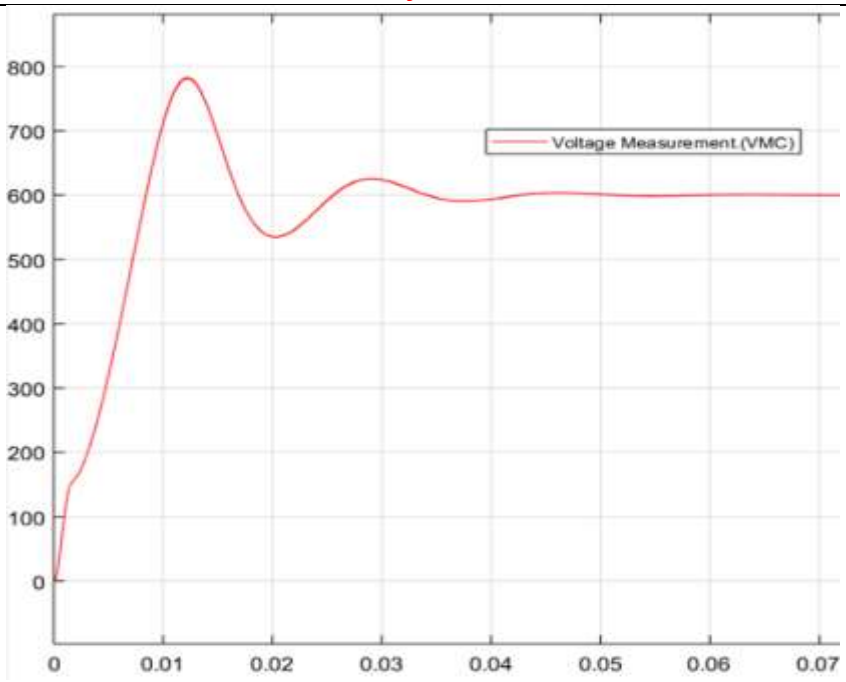


Figure 6. (a) Voltage at the output side

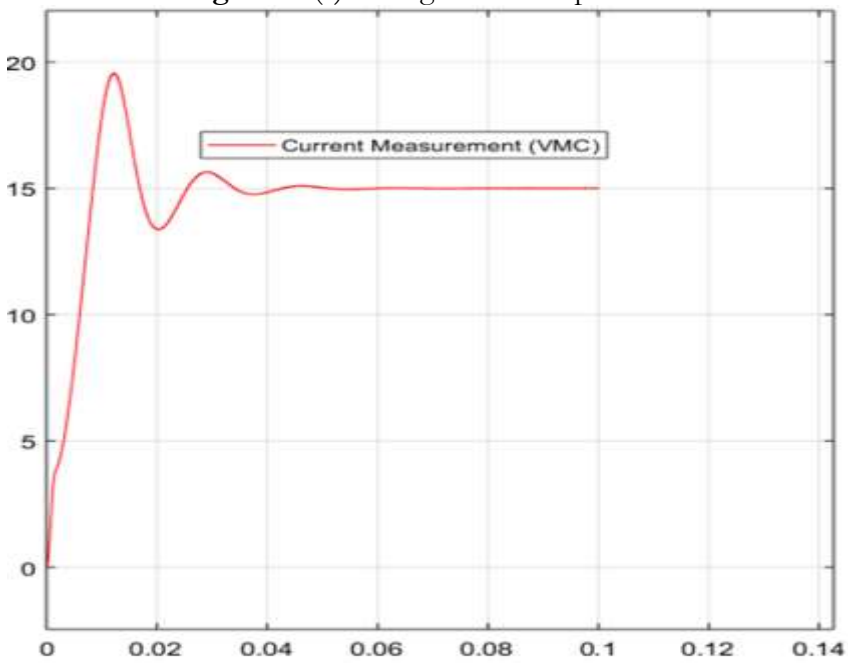


Figure 6. (b) output current I_{out} for voltage mode-based converter control (VMC) based proposed converter

Implementation of the proposed converter with the current injected mode control

Current-mode control (CMC) is often preferable over Voltage-Mode Control (VMC) due to its faster dynamic response, particularly under load variations. CMC incorporates two feedback loops as shown in the block diagram for the proposed converter (Figure 7), improving stability and mitigating the impact of right-half s-plane zeros. This enhances transient response and allows for reduced total harmonic distortion (THD) and better input disturbance rejection.

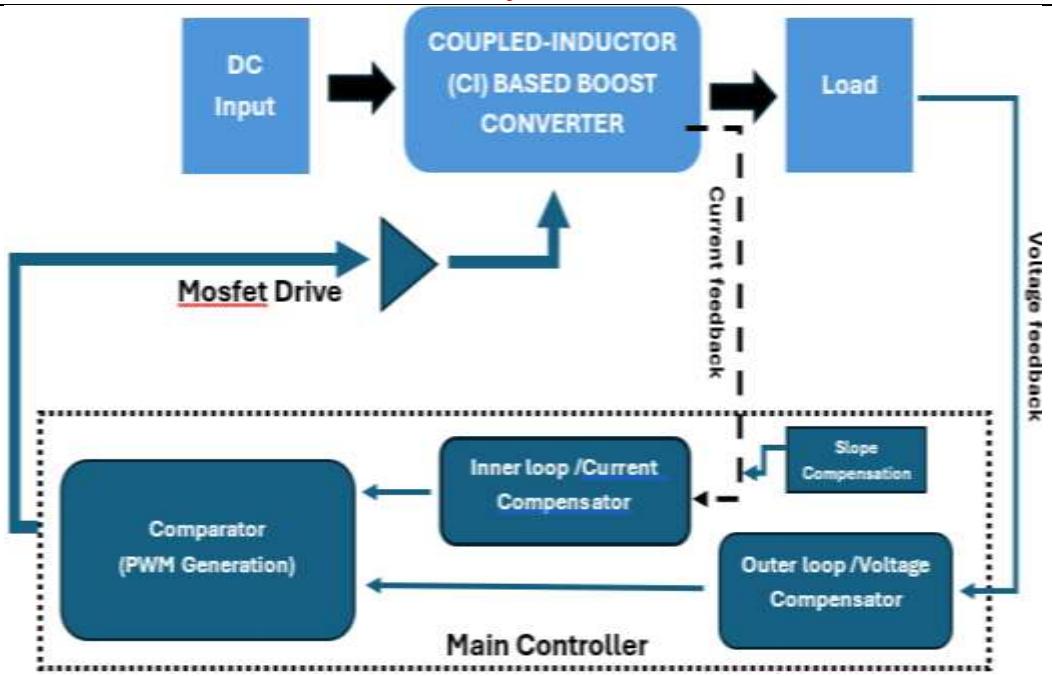


Figure 7. Block diagram for current-injected control with current (inner) and voltage (outer) loops

For current-injected mode control the inductor current or switch current is taken as feedback, and it is compared to control current, as shown in the circuit diagram in Figure 8. The current mode-based control design is helpful in not only achieving regulated output voltage but also improving stability as well as transients.

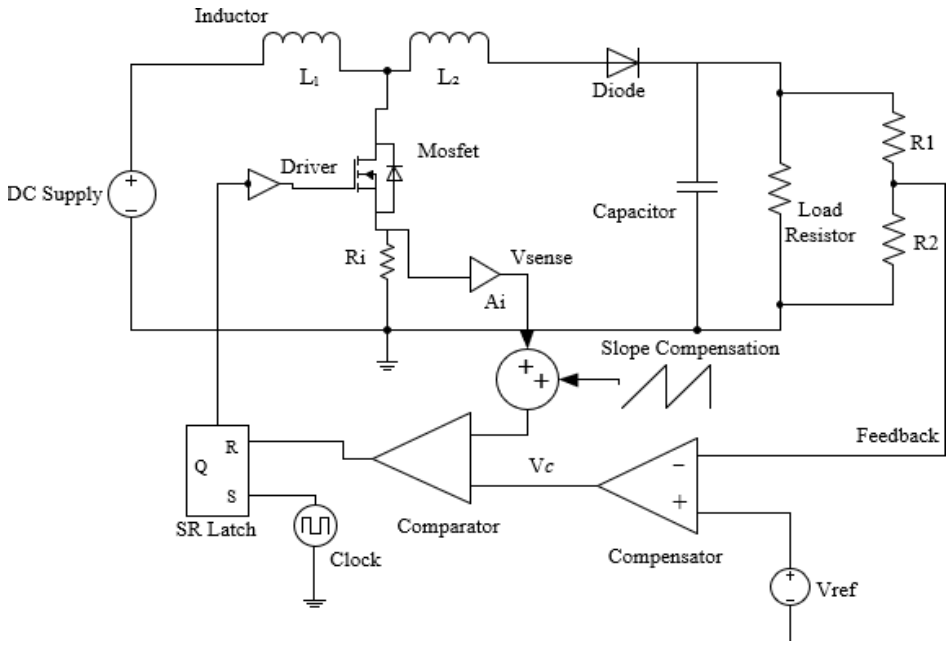


Figure 8. Circuit diagram for current-injected control with slope compensation

Type 3 compensation is rather easy, as shown in Figure 9, simplifies rules and achieves stability effectively. The role of ramp involvement is merely to minimize the involvement of subharmonic oscillations. Current injected control methodology is also termed as a current programmed mode which contains a total of 2 loops. Sometimes it's also termed as multiple loop control as it has 2 loops, one being the current loop which is an inner loop and the other is the voltage loop which is an outer loop in the overall control mechanism.

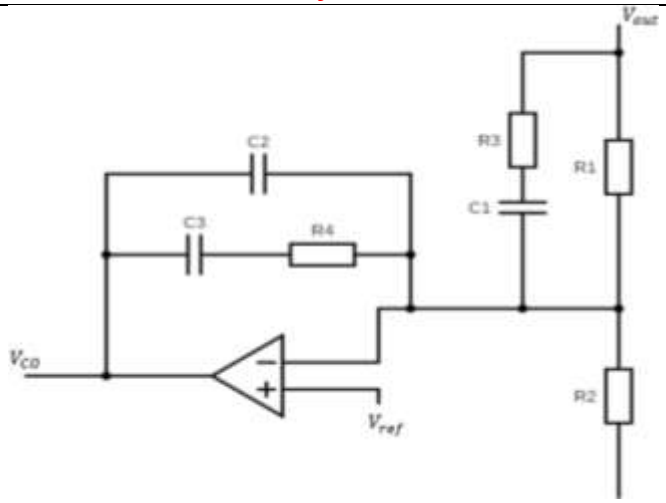


Figure 9. Type 3 compensator via op- amp [21]

To relate the control current with the respective output voltage a simplified transfer function is derived using a small signal model. With the addition of compensation/artificial ramp, the difference between the control current and inductor sensed current becomes prominent. In the block diagram, given in Figure 10, the km block represents the role of compensation ramp.

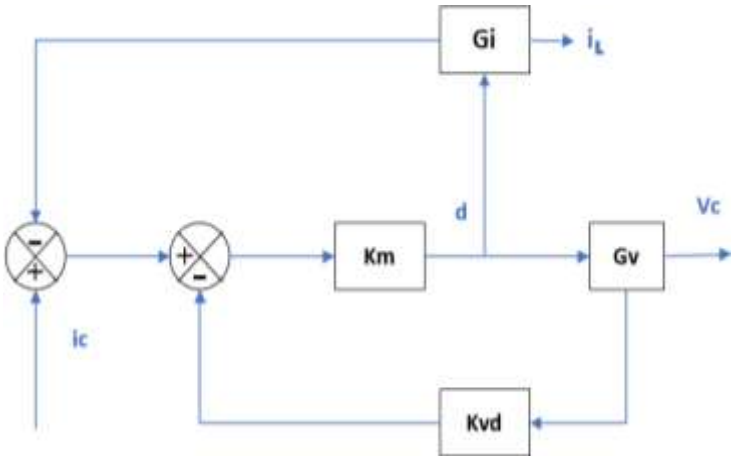


Figure 10. Current injected control-based simplified block diagram for the proposed converter

By using a Type-3 compensator, often implemented with an op-amp, CMC can achieve phase boosts greater than 90°, improving stability and performance. Furthermore, CMC limits overcurrent conditions, preventing system failures and offering robust control. In systems with duty cycles greater than 50%, subharmonic oscillations are mitigated by adding a compensation ramp. This makes CMC particularly beneficial for better efficiency, high gain converters that can be utilized for renewable energy system integration, and electric vehicles, where rapid and stable voltage regulation is critical. The overall system for current-injected control with slope compensation was implemented in Matlab/Simulink and its output voltage as well as current waveforms are given in Figure 11.

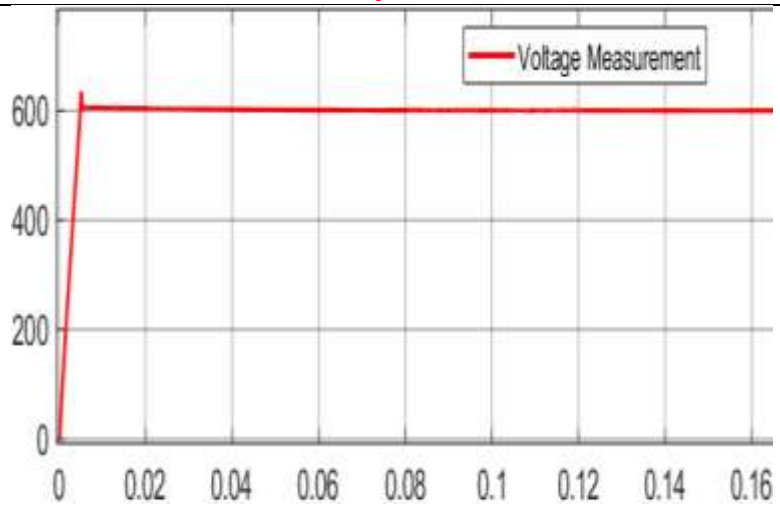


Figure 11(a). Output Voltage waveform of the proposed converter with CMC

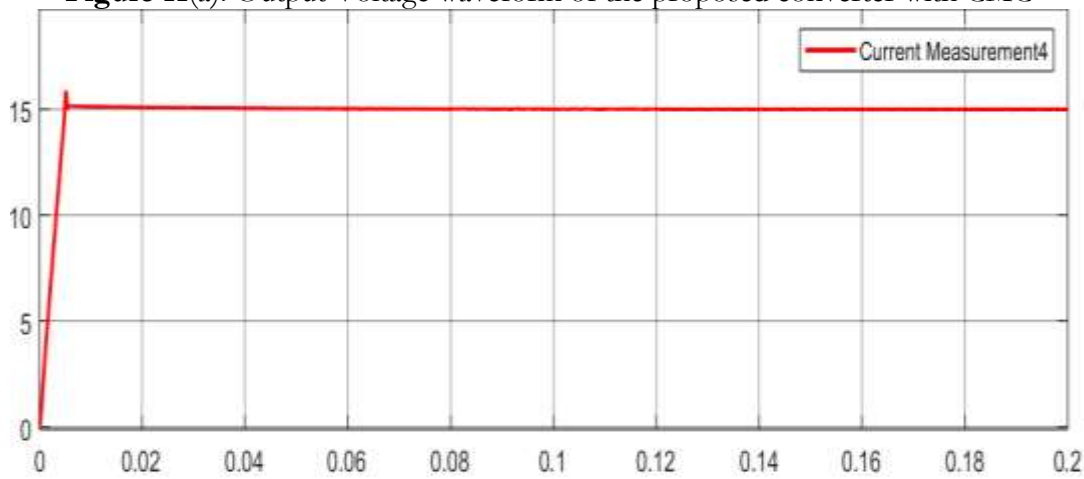


Figure 11 (b). Output Current waveform of the proposed converter with CMC

Effectiveness of Proposed Converter W.R.T Relevant Converters:

The major emphasis was to develop a converter that could achieve high gain and efficiency. The comparison has been built to show the effectiveness of the converter as shown in table 01. The comparison has been done based on component count (like for switch, diode, and capacitor), and voltage stress of capacitor and switch. Finally, the converter's voltage gain vs duty ratio curves have been plotted as shown in Figure 12.

Table 1. Comparison table of proposed converter with other relevant topologies

Topology	Proposed Converter	Ref [23]	Ref [17]	Ref [15]	
No of switches	1	1	2	1	2
No of capacitors	1	3	2	4	5
No of diodes	1	4	2	4	8
Switch side voltage stress.	$\frac{[1+\mathrm{n.d}]V_i}{(1-\mathrm{d})}$	$\frac{V_i}{(1-\mathrm{d})}$	$\frac{[1+(2n)]V_i}{(1-\mathrm{d})}$	$\frac{V_i}{(1-2.\mathrm{d})}$	$\frac{V_i}{(1-\mathrm{d})}$
Output capacitor side stress	less	high	high	high	high

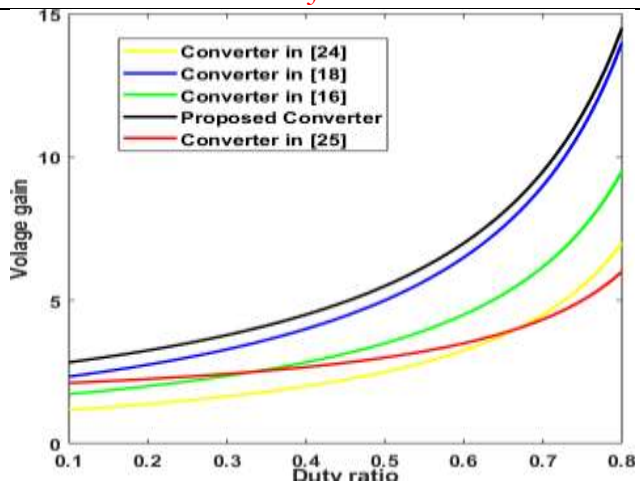


Figure 12. Voltage gain to duty ratio comparison curves

Discussion:

To testify to the performance of the proposed converter a hardware prototype has been built. The efficiency curve is drawn for various loads varying from 10W to 100W when connected at the load side. For a load of 100W, the overall efficiency is calculated to be around 95- 96% for a voltage of 8-10V at the input side and 97-101V at the output side. The hardware prototype implementation is shown in Figure 13. The selection of components and their values are set as per mathematical calculations as given in Table 02. The high conversion efficiency for the proposed converter has been observed to have a peak value of around 96%. The efficiency vs output power curve is also given in Figure 14.

The output voltage and current waveforms of the proposed Coupled Inductor Boost Converter demonstrated significant improvements over the existing topologies referenced in the literature. Unlike conventional boost converters, where high-duty cycles lead to excessive voltage ripples and instability [8][9], the proposed topology achieves higher voltage gain at moderate-duty cycles, as seen in the voltage waveform, where the output voltage remains stable with minimal overshoot and steady-state error. In contrast to switched capacitors and interleaved boost converters, which experience high inrush currents and increased switching losses [14], [15], the proposed converter maintains a well-regulated and smooth current waveform, reducing stress on power components. Additionally, compared to transformer-based isolated converters, which often suffer from leakage inductance-induced oscillations [10], [11], the proposed design ensures a more controlled voltage response with faster settling time. Implementing current mode control further enhances overall stability by mitigating right-half-plane (RHP) zero effects, ensuring faster transient response and lower current ripples compared to voltage mode control (VMC), as compared in section 4 of the paper where voltage mode control is first separately implemented and then current-injected mode control is implemented, involving both outer (voltage loop) and inner (current loop). Efficiency curves also validate the superiority of this design, showing higher efficiency (>95%) across different load conditions, whereas conventional topologies often exhibit a drop in efficiency due to increased conduction losses. These results confirm that the proposed converter is better suited for renewable energy applications and high-power DC systems, providing stable operation, lower ripple, and improved dynamic response compared to the existing literature.

Various control strategies exist for boost converters, each with unique advantages and limitations in terms of dynamic response, noise rejection, and implementation complexity. While the current study employs current-injected mode control (CMC) to enhance system stability and efficiency, alternative control techniques such as Voltage Mode Control (VMC), Sliding Mode Control (SMC), Predictive Control (MPC), and Fuzzy Logic Control (FLC) have been widely

explored in the existing literature. A comparative analysis of these methods provides valuable insights into their suitability for different power electronics applications.

Voltage Mode Control (VMC) is one of the simplest and most commonly used techniques for regulating the output voltage of a boost converter. It operates using a single feedback loop where the output voltage is compared to a reference, and the error is processed through a compensator, typically generating a pulse-width modulation (PWM) signal. While VMC is easy to implement and cost-effective, it suffers from slow dynamic response and poor noise rejection, making it less effective for applications requiring fast transient regulation. Furthermore, the presence of a right-half-plane (RHP) zero in boost converters complicates loop compensation and limits performance. In comparison, CMC provides faster dynamic response and better noise immunity, as it directly controls the inductor current rather than relying solely on output voltage feedback.

Sliding Mode Control (SMC) is a nonlinear control strategy that dynamically adjusts the duty cycle based on system states, ensuring robust performance even under uncertain conditions. This approach excels in fast transient response and noise rejection, making it ideal for high-performance applications in harsh environments. However, the primary drawback of SMC is the chattering effect, where high-frequency switching introduces oscillations that may degrade system efficiency. Additionally, the complexity of designing an appropriate sliding surface and ensuring stability requires extensive mathematical modeling. Compared to SMC, CMC offers smoother control action and easier implementation, albeit with a slightly slower response under extreme load variations.

Predictive Control (Model Predictive Control – MPC) takes a model-based approach, forecasting system behavior and optimizing control inputs in real time. MPC is particularly effective in handling multi-objective control, such as voltage regulation, efficiency optimization, and disturbance rejection. Its major advantages include excellent dynamic performance and precise control, making it suitable for smart grids, electric vehicle (EV) charging systems, and renewable energy applications. However, MPC requires high computational power, as it relies on solving complex equations at each control interval. This makes it less practical for low-cost power converters, where CMC remains a more feasible option due to its lower computational burden and easier implementation.

Fuzzy Logic Control (FLC) is another alternative that relies on linguistic rules and heuristic decision-making rather than mathematical models. FLC is highly adaptable to nonlinear systems and does not require precise modeling, making it beneficial for renewable energy converters where environmental conditions fluctuate unpredictably. However, FLC faces challenges in tuning membership functions and lacks optimality in fast transient response, as it operates more reactively than proactively. Compared to CMC, which provides structured, real-time response control, FLC is better suited for applications where adaptability to unpredictable variations is required rather than stringent performance optimization.

The performance of the proposed converter based on the current mode control scheme is checked by simulating it on MATLAB with the following design parameters: $V_{in} = 100V, V_{out} = 600,\ L_1 = 600uH, L_2 = 2.4mH, R_{1L} = 40\Omega, C_1 = 56uF,\ f_s = 50kHz,\ D = 0.534$.

Table 2. Component's selection and relevant values

Parameter	Value	Unit
Inductor L_1	600uH	Henry
Inductor L_2	2.4mH	Henry
Mutual Inductance	1.14e-3H	Henry
Diode (RHRG75120)	-	-
Mosfet (IRF640N)	-	-

Capacitor	220uF	farad
Input Voltage	100V	volt
Load Resistance	40Ω	ohm
Pulse width (%)	53.4%	-
Operating Frequency	50k	Hz

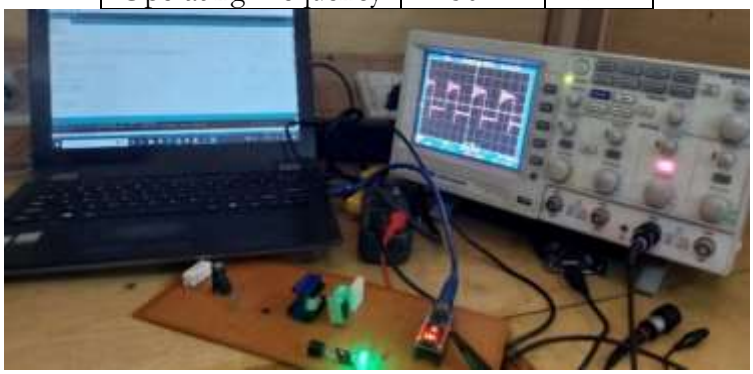


Figure 13. Prototype design implementation for the proposed converter

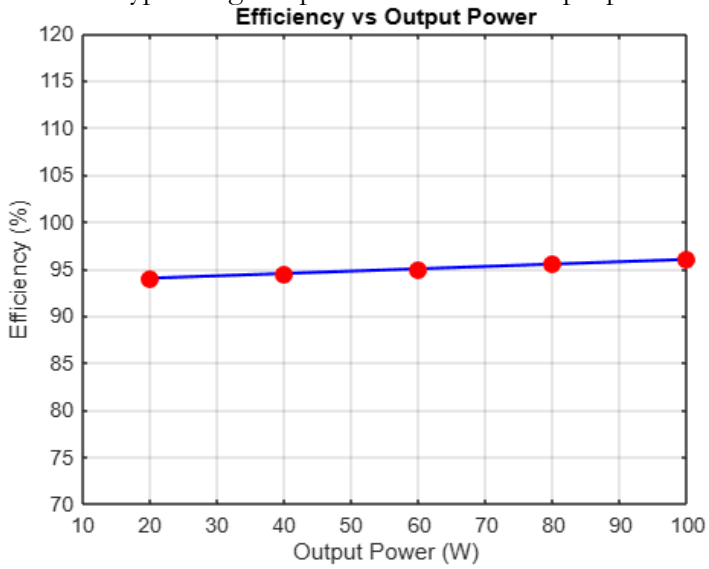


Figure 14. Efficiency vs power at the output (W)

Conclusion & Future Recommendation:

Coupled inductors-based boost converters having efficient control are well known to attain regulated voltage at the output. While achieving the goals of high gain and high efficiency. Such step-up high-gain converters can be efficiently integrated with renewable energy resources and electric vehicle applications. Current-injected control (or CMC) is being implemented and presented in this thesis/research work to obtain regulated voltage at the output. A small signal model with average method state space modeling is represented as well. Current-injected mode control (or CMC) is utilized to achieve fast transient response, low THD (total harmonic distortion), input disturbance rejection, and stable DC link. Coupled inductor topology is utilized to obtain high gain by the adjustment of turn's ratio properly. Renewable resources like solar-based energy generation systems prioritize such converters owing to their variable nature. Such high-gain converters are also helpful in increasing the low voltage obtained from such resources. MPPT (maximum power point tracking) helps to achieve maximum power from the PV panel using perturb and observe technique irrespective of temperature and irradiance. The proposed converter has been implemented on the Matlab Simulink model and it's also verified via a hardware-based laboratory prototype.

Recommendations for Future Work:

As a future recommendation the proposed converter can be effectively integrated with EV i.e., electrical vehicle-based applications. With the increase in the rates of fuel and depletion of conventional resources like natural gas, and fossil fuels and likewise, it is mandatory to look for advanced techniques that can help to implement electric cars on a large scale. This step will pay a lot for economic development throughout the world. Another application of the proposed converter involves the penetration of PV (solar arrays) on the source side. There is ongoing research on a wide scale regarding smart grids and microgrids. The variable as well as unreliable nature of PV module system requires such dc-dc converters which can provide voltage regulation. The concept of active building is also under consideration which again emphasizes the requirement of efficient high-gain step-up dc/DC converters.

References:

[1] G. Velasco-Quesada, F. Guinjoan-Gispert, R. Piqué-López, M. Román-Lumbreras, and A. Conesa-Roca, "Electrical PV array reconfiguration strategy for energy extraction improvement in grid-connected PV systems," *IEEE Trans. Ind. Electron.*, vol. 56, no. 11, pp. 4319–4331, 2009, doi: 10.1109/TIE.2009.2024664.

[2] F. Nejabatkhah, S. Danyali, S. H. Hosseini, M. Sabahi, and S. M. Niapour, "Modeling and control of a new three-input dc-dc boost converter for hybrid PV/FC/battery power system," *IEEE Trans. Power Electron.*, vol. 27, no. 5, pp. 2309–2324, 2012, doi: 10.1109/TPEL.2011.2172465.

[3] M. Forouzesh, Y. Shen, K. Yari, Y. P. Siwakoti, and F. Blaabjerg, "High-Efficiency High Step-Up DC-DC Converter with Dual Coupled Inductors for Grid-Connected Photovoltaic Systems," *IEEE Trans. Power Electron.*, vol. 33, no. 7, pp. 5967–5982, Jul. 2018, doi: 10.1109/TPEL.2017.2746750.

[4] Y. P. Hsieh, J. F. Chen, T. J. Liang, and L. S. Yang, "A novel high step-up DC-DC converter for a microgrid system," *IEEE Trans. Power Electron.*, vol. 26, no. 4, pp. 1127–1136, 2011, doi: 10.1109/TPEL.2010.2096826.

[5] R. J. Wai, C. Y. Lin, C. Y. Lin, R. Y. Duan, and Y. R. Chang, "High-efficiency power conversion system for kilowatt-level stand-alone generation unit with low input voltage," *IEEE Trans. Ind. Electron.*, vol. 55, no. 10, pp. 3702–3714, 2008, doi: 10.1109/TIE.2008.921251.

[6] H. Ardi, A. Ajami, and M. Sabahi, "A novel high step-up DC-DC converter with continuous input current integrating coupled inductor for renewable energy applications," *IEEE Trans. Ind. Electron.*, vol. 65, no. 2, pp. 1306–1315, 2018, doi: 10.1109/TIE.2017.2733476.

[7] W. Li and X. He, "Review of nonisolated high-step-up DC/DC converters in photovoltaic grid-connected applications," *IEEE Trans. Ind. Electron.*, vol. 58, no. 4, pp. 1239–1250, Apr. 2011, doi: 10.1109/TIE.2010.2049715.

[8] Q. Zhao and F. C. Lee, "High-efficiency, high step-up dc-dc converters," *IEEE Trans. Power Electron.*, vol. 18, no. 1 I, pp. 65–73, Jan. 2003, doi: 10.1109/TPEL.2002.807188.

[9] F. Li and H. Liu, "A Cascaded Coupled Inductor-Reverse High Step-Up Converter Integrating Three-Winding Coupled Inductor and Diode-Capacitor Technique," *IEEE Trans. Ind. Informatics*, vol. 13, no. 3, pp. 1121–1130, Jun. 2017, doi: 10.1109/TII.2016.2637371.

[10] H. M. Hsu and C. T. Chien, "Multiple turn ratios of on-chip transformer with four intertwining coils," *IEEE Trans. Electron Devices*, vol. 61, no. 1, pp. 44–47, 2014, doi: 10.1109/TED.2013.2292855.

[11] B. Gu, J. Dominic, J. S. Lai, Z. Zhao, and C. Liu, "High boost ratio hybrid transformer DC-DC converter for photovoltaic module applications," *IEEE Trans. Power Electron.*, vol. 28, no. 4, pp. 2048–2058, Apr. 2013, doi: 10.1109/TPEL.2012.2198834.

[12] T. J. Liang, J. H. Lee, S. M. Chen, J. F. Chen, and L. S. Yang, "Novel isolated high-step-Up DC-DC converter with voltage lift," *IEEE Trans. Ind. Electron.*, vol. 60, no. 4, pp. 1483–1491, 2013, doi: 10.1109/TIE.2011.2177789.

[13] Y. Tang, D. Fu, T. Wang, and Z. Xu, "Hybrid switched-inductor converters for high step-up conversion," *IEEE Trans. Ind. Electron.*, vol. 62, no. 3, pp. 1480–1490, Mar. 2015, doi: 10.1109/TIE.2014.2364797.

[14] Y. Tang, T. Wang, and Y. He, "A switched-capacitor-based active-network converter with high

voltage Gain," *IEEE Trans. Power Electron.*, vol. 29, no. 6, pp. 2959–2968, Jun. 2014, doi: 10.1109/TPEL.2013.2272639.

[15] G. Wu, X. Ruan, and Z. Ye, "High Step-Up DC-DC Converter Based on Switched Capacitor and Coupled Inductor," *IEEE Trans. Ind. Electron.*, vol. 65, no. 7, pp. 5572–5579, Jul. 2018, doi: 10.1109/TIE.2017.2774773.

[16] H. C. Liu and F. Li, "Novel High Step-Up DC-DC Converter with an Active Coupled-Inductor Network for a Sustainable Energy System," *IEEE Trans. Power Electron.*, vol. 30, no. 12, pp. 6476–6482, Dec. 2015, doi: 10.1109/TPEL.2015.2429651.

[17] S. H. C. Taosif Iqbal, Yaqian Wang, Haibin Lu, Xiongwen Zhang, "Small signal modelling and control of high gain coupled inductor boost inverter for solid oxide fuel cell based power generation system," *Int. J. Hydrogen Energy*, vol. 44, no. 38, pp. 21115–21126, 2019, doi: https://doi.org/10.1016/j.ijhydene.2019.01.289.

[18] A. F. Witulski, "Introduction to Modeling of Transformers and Coupled Inductors," *IEEE Trans. Power Electron.*, vol. 10, no. 3, pp. 349–357, 1995, doi: 10.1109/63.388001.

[19] Texas Instruments, "AN-1286 Compensation for the LM3478 Boost Controller," Application Note. Accessed: Mar. 17, 2025. [Online]. Available: https://e2e.ti.com/support/power-management-group/power-management/f/power-management-forum/922179/lm3478-an-1286-compensation-for-the-lm3478-boost-controller

[20] Texas Instruments, "Demystifying Type II and Type III Compensators Using OpAmp and OTA for DC/DC Converters," *Appl. Rep.*, 2014, [Online]. Available: https://www.ti.com/lit/an/slva662/slva662.pdf

[21] M. A. Vaghela and M. A. Mulla, "Peak Current Mode Control of Coupled Inductor based High Step-Up Gain Boost Converter," *India Int. Conf. Power Electron. IICPE*, vol. 2018-December, Jul. 2018, doi: 10.1109/IICPE.2018.8709507.

[22] Y. P. Hsieh, J. F. Chen, T. J. Liang, and L. S. Yang, "Novel high step-Up DC-DC converter for distributed generation system," *IEEE Trans. Ind. Electron.*, vol. 60, no. 4, pp. 1473–1482, 2013, doi: 10.1109/TIE.2011.2107721.

[23] M. B. Meier, D. S. S. Avelino, A. A. Badin, E. F. R. Romaneli, and R. Gules, "Soft-switching high static gain DC-DC converter without auxiliary switches," *IEEE Trans. Ind. Electron.*, vol. 65, no. 3, pp. 2335–2345, Mar. 2018, doi: 10.1109/TIE.2017.2739684.

Extraction of bio-oil from the pyrolysis of banana tree waste using a fixed-bed reactor

Muhammad Tahseen Sadiq[1*], Mahmboob Ahmad [1], Aman Fatima [1], Kamal Zafar [1], Muhammad Hassan Qasim [1], Faheem Qaisar [1], Shahzaib Hassan [1], Humaira Hameed [2]
[1]Department of Chemical Engineering, University of Gujrat, Gujrat, Pakistan.
[2]Department of Chemistry, University of Agricultural, Faisalabad, Pakistan
* Correspondence: m.tahseensadiq@uog.edu.pk

Citation | Sadiq. M. T, Ahmad. M, Fatimah. A, Zafar. K, Qasim. M. H, Qaisar. F, Hassan. S, Hameed. H, "Extraction of bio-oil from the pyrolysis of banana tree waste using a fixed-bed reactor", IJIST, Special Issue. pp 18-30, March 2025
Received | Feb 09, 2025 **Revised |** Feb 21, 2025 **Accepted |** Feb 26, 2025 **Published |** March 02, 2025.

The rapid and ongoing depletion of fossil fuel reserves is driving up energy costs and harming the environment due to greenhouse gas emissions, leading to a global energy crisis. This situation highlights the urgent need to produce renewable fuel from biomass. This research focuses on extracting bio-oil from banana tree waste under different operating conditions. In this study, the pyrolysis process of banana tree waste was carried out in a fixed-bed reactor to maintain controlled conditions and prevent unwanted cracking. To optimize the process, the effects of temperature, particle size, and nitrogen flow rate on bio-oil yield were investigated. Experiments were conducted at temperatures ranging from 400 to 600 ℃, with feedstock particle sizes of 0.5 – 2.0 mm and nitrogen flow rates between 0.5 and 2 liters per minute. The optimal conditions for maximizing bio-oil yield were determined. Under these conditions, the maximum bio-oil yield of 32.13% was obtained at a temperature of 500 ℃, with a particle size of 1.2 – 2.0 mm and a nitrogen flow rate of 1 liter per minute. The results also demonstrate how temperature, particle size, and nitrogen flow affect the bio-oil yield during pyrolysis. The study concludes that banana tree waste can be efficiently converted into bio-oil through proper processing, contributing to sustainable energy production while minimizing environmental impact. The chemical composition of the bio-oil was analyzed using the GC-MS technique, which identified various compounds, including phenols, acids, and other chemical components.

Keywords: Biomass; Renewable Energy; Banana Tree Waste; Pyrolysis; Bio-oil.

Introduction:

The rapid depletion of fossil fuel reserves, rising fuel prices, increased fuel consumption, greenhouse gas emissions, and environmental degradation have all contributed to the current energy crisis [1]. Agricultural waste, such as crop husks, leaves, stems, and shells, is often underutilized and poorly managed. Instead of converting this waste into bioenergy, compost, or animal feed, it is commonly burned in open fields, which releases greenhouse gases like carbon dioxide and methane, further exacerbating global warming [2]. This practice lowers air quality, contributes to acid rain, and affects nearby water bodies. Through atmospheric deposition, these pollutants can enter aquatic ecosystems, reducing water pH and harming aquatic life [3]. Banana trees, with their abundant cultivation and high biomass yield, especially in Asia, Africa, and Latin America, are an attractive source of biomass for producing biofuels and other chemicals [4]. According to the Food and Agriculture Organization of the United Nations (FAO), global trade in tropical fruits reached 7.7 million tons in 2019, reflecting a 6.4% (465,000 tons) increase from previous years [5]. Biomass from banana trees has various uses: pseudo-stems can serve as mulch, a starch source, or raw material for ropes, fabrics, or paper, while peels can be used as compost or animal feed. Discarded fruits are also suitable for animal consumption [4]. Despite these potential uses, about 60% of banana biomass is wasted through open burning and dumping, which poses significant risks to human and environmental health [5].

Pyrolysis is one of the most effective methods for converting biomass into biofuel. This process involves heating biomass at high temperatures in an inert atmosphere, typically using argon or nitrogen gas [6]. Nitrogen acts as a carrier gas, preventing oxidation and ensuring controlled thermal decomposition [7]. Due to its high energy content, banana waste can yield biofuels and other valuable products through pyrolysis. A typical ton of banana waste consists of 750 kg (75%) pseudo-stems, 100 kg (10%) leaves, 40 kg (4%) rachis, and 110 kg (11%) peels. These byproducts are ideal raw materials for bio-refineries, which can produce biofuels and bio-based chemicals [8]. During pyrolysis, biomass decomposes in the absence of oxygen, yielding bio-char and bio-oil rich in carbon. The primary components of lignocellulosic biomass—cellulose, lignin, and hemicellulose—degrade when heated between 300 and 500 ℃ [9].

Some researchers prefer pyrolysis due to its optimal operating conditions, including moderate temperatures (300-500 ℃) and medium heating rates [10]. Compared to slow pyrolysis, it provides higher bio-oil yields and avoids the excessive formation of bio-char and bio-gas associated with fast pyrolysis [11]. Fixed-bed reactors are often favored over fluidized-bed reactors because they produce higher bio-oil yields. The slower heating rate in fixed-bed reactors allows more thorough pyrolysis of the feedstock, resulting in valuable liquid products with minimal gas and char. Although fluidized-bed reactors operate quickly, their high gas production reduces their efficiency for maximizing bio-oil yield [12]. Several factors influence banana waste pyrolysis in fixed-bed reactors, including temperature, heating rate, particle size, and residence time. Temperature plays a critical role in product distribution, with lower temperatures favoring bio-char and gas formation and higher temperatures promoting bio-oil production [13]. Typically, temperatures between 350 and 650 ℃ are used to maximize bio-oil yield, while residence times of 2 to 10 minutes help achieve a balanced output [10].

The characteristics of bio-oil are heavily influenced by feedstock type and pyrolysis conditions [14]. Key factors affecting the pyrolysis yield and efficiency of banana tree waste include its lignocellulose composition, moisture content, and ash concentration. Without proper drying, the high moisture content in banana waste can lower pyrolysis efficiency. The concentrations of cellulose, hemicellulose, and lignin also significantly impact bio-oil yields, with higher cellulose content generally leading to greater bio-oil output. Additionally, ash content affects the thermal stability of the feedstock, further influencing process efficiency [15].

Objectives:

This study aims to explore how different pyrolysis parameters affect the production of bio-oil from banana tree waste. It evaluates the efficiency of extracting high-quality bio-oil from this waste. Additionally, it seeks to determine the optimal temperature, particle size, and nitrogen flow required to maximize bio-oil yields during the pyrolysis process in a fixed-bed reactor with a capacity of 30 grams. The study also examines how varying operating conditions influence both bio-oil yield and its characteristics. Finally, the research assesses the properties of the obtained bio-oil.

Materials and Methods:

Collection and Preparation of Feedstock:

Banana tree waste was chosen as the feedstock due to the widespread cultivation and high demand for this banana variety in Pakistan. The banana leaves used in the experiments were collected from a farmer in Nawabshah, Khairpur, Sindh, Pakistan. Only mature banana trees, over 10 months old, were selected to ensure high-quality leaves. The leaves were carefully cut from the stem, measuring between 30 to 50 cm, and placed on a sterile canvas to avoid soil contamination. The fresh, moist leaves were then divided into 4 to 8 sections and manually chopped into smaller pieces before drying. Initially, the chopped leaves were air-dried outdoors in the Sukkur region for 5 to 6 days. To further reduce moisture content, the leaves were oven-dried at 105 °C for 8 hours, ensuring they were sufficiently desiccated for subsequent analysis.

Feed Characterization:

The banana tree waste was chopped to reduce particle size, and sieve analysis was conducted using screens with mesh sizes of 200, 400, 600, 800, and 1000. Two particle size ranges were selected: 0.5–1.2 mm and 1.2–2.0 mm. Proximate analysis was performed to measure moisture content, volatile matter, ash content, and fixed carbon in the banana waste samples. Moisture content was determined using ASTM E871-82 [16], which specifies that dried materials should have a moisture level below 10%. Volatile matter was assessed according to ASTM E872 [16], and ash content was measured following ASTM E1102-84 [17]. Fixed carbon content was calculated by subtracting the moisture, volatile matter, and ash percentages from 100. Each analysis was repeated three times to ensure accuracy and reliability. The Perkin-Elmer Series II CHNS/O 2400 Analyzer was used for ultimate analysis, measuring the carbon, hydrogen, nitrogen, sulfur, and oxygen content in the feed samples. The fixed carbon and oxygen contents were calculated using the following formulas:

- **Fixed Carbon (%)** = 100 - (volatile matter + moisture content + ash content)
- **Oxygen (%)** = 100 - (Carbon + Hydrogen + Nitrogen + Sulfur)

Lignocellulosic analysis was conducted to quantify the biomass's hemicellulose, lignin, cellulose, and extractive content, using standard methods ASTM D1106 [18], ASTM D1103 [19], ASTM D1104 [20], and ASTM D1105 [21]. The samples' higher heating value (HHV) was measured using an Adiabatic Bomb Calorimeter (IKA C-200). The lower heating value (LHV) was then calculated using the following formula:

$$\textbf{LHV (dry, MJ/kg)} = \text{HHV (dry)} - 2.442\,(8.936 \times H / 100)$$

Experimental Methodology:

The pyrolysis of banana tree waste was performed in a fixed-bed reactor under atmospheric nitrogen pressure. The reactor had a length of 10 cm and an inner diameter of 5 cm. This reactor type was selected due to its ease of operation, controlled heating conditions, and ability to facilitate gradual pyrolysis, which enhances bio-oil yield. Unlike fluidized-bed reactors, the fixed-bed reactor minimizes secondary cracking by maintaining a stable reaction environment with better residence time control. Nitrogen gas (N_2) was used as an inert medium to prevent combustion and oxidation, thereby ensuring that thermal decomposition of the biomass produced volatile compounds. Nitrogen flow also influenced vapor residence time,

which in turn affected the yield and composition of bio-oil by reducing excess char or gas formation.

A schematic diagram of the pyrolysis setup is shown in Figure 1. Briefly, 30 grams of banana tree residues were loaded into the reactor, and nitrogen gas was purged at a flow rate of 0.5 liters per minute. The initial temperature of the reactor was set to 25 °C, and a thermocouple was used to monitor the internal temperature. The vapors generated during the pyrolysis process were condensed using a spiral condenser, with cooling water maintained at 4 °C. The condensed bio-oil was then collected in a collection chamber.

Pyrolysis experiments were conducted in three series using a fixed-bed reactor. Before each experiment, the reactor's temperature was stabilized and controlled using a thermocouple connected to a control panel to ensure uniform heating.

In the **first series**, the focus was on studying the effect of reaction temperature on the yield distribution. Banana leaves were pyrolyzed at temperatures of 400, 450, 500, 550, and 600 °C to determine the optimal temperature for maximizing bio-oil yield. All other operating parameters were kept constant. The **second series** aimed to examine the impact of particle size on bio-oil yields. Banana residue particles were divided into two size ranges: 0.5–1.2 mm and 1.2–2.0 mm. These particle sizes were tested under varying temperatures to observe their effect on bio-oil production. In the **third series**, nitrogen gas was used as an inert medium to create an oxygen-free environment, preventing combustion. Since pyrolysis is a thermal degradation process that requires the absence of oxygen, nitrogen ensured that the system remained oxygen-free, preserving the controlled conditions necessary for pyrolysis. The nitrogen flow rate was regulated using a flow meter connected to the gas cylinder, helping maintain a stable inert atmosphere.

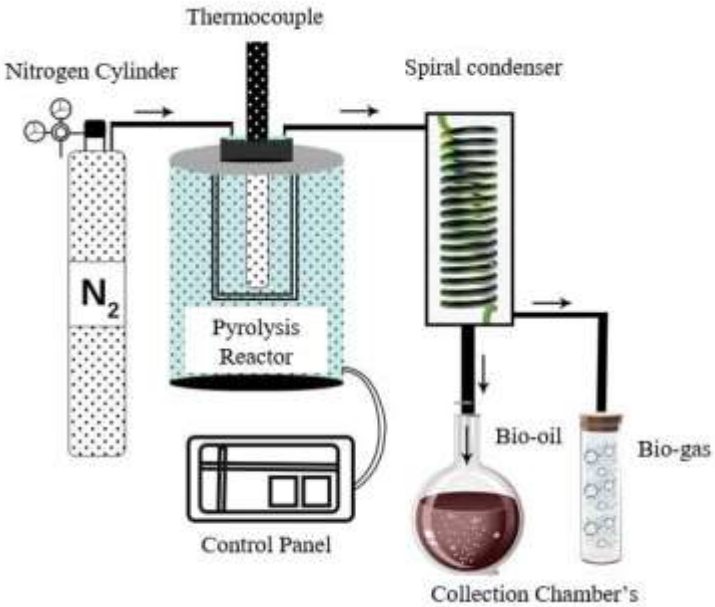


Figure 1. Experimental diagram of Pyrolysis process

Before each run, thermal equilibrium was reached, and key parameters like temperature and nitrogen flow were kept constant. This setup allowed the controlled release of volatile compounds, minimized combustion risks, and maintained the desired inert environment. By adjusting the nitrogen flow rate, the pyrolysis process was optimized, product distribution improved, and temperature fluctuations controlled, thereby enhancing overall efficiency. The methodology's flow diagram is shown in Figure 2.

The bio-oil yield was calculated using the following formula:

% Yield of bio-oil (wt. %) = (Bio-oil mass (g) / Dry feedstock mass (g)) × 100%

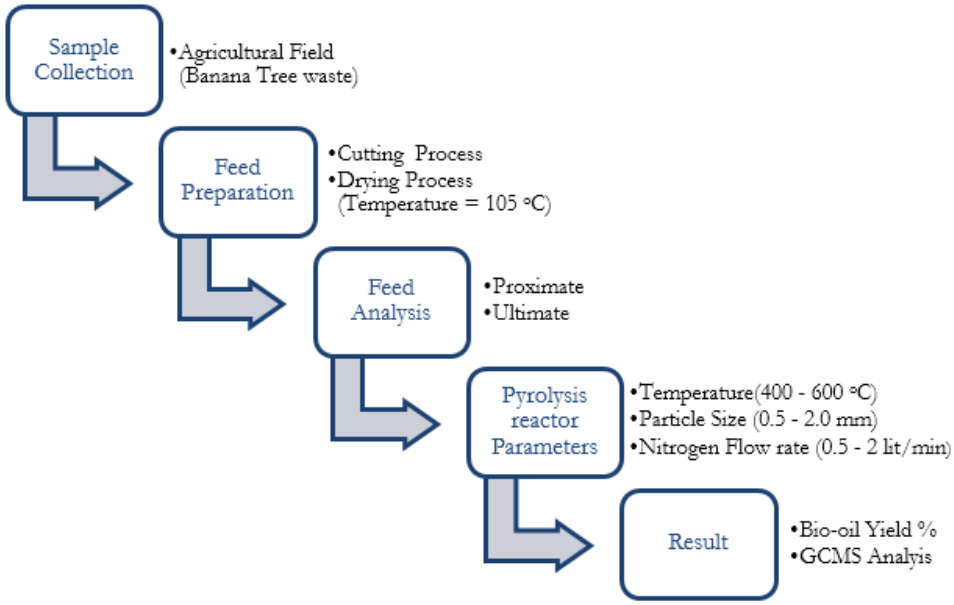


Figure 2. Flow diagram of Pyrolysis process

Product Analysis:

The extraction of bio-oil under optimal pyrolysis conditions, which yielded the highest liquid output, was thoroughly analyzed. To determine the chemical composition of the collected bio-oil, Gas Chromatography-Mass Spectrometry (GC-MS) was employed. The analysis focused on the bio-oil produced at a pyrolysis temperature of 500 °C, where chemical constituents were carefully evaluated. To prepare the samples for GC-MS analysis, a Liquid-Liquid Extraction (LLE) technique was used to remove water content. This step is essential in isolating and purifying the liquid products, particularly for separating organic compounds from the aqueous phase generated during pyrolysis. Since pyrolysis produces a complex mixture of gases, solids, and liquids—including water, organic compounds, oils, acids, alcohols, and phenols—the presence of water can hinder the analysis and application of the liquid products.

In the LLE process, dichloromethane was mixed with the liquid sample in a 1:1 volumetric ratio. Specifically, 5 mL of the liquid bio-oil and 5 mL of dichloromethane were combined and stirred thoroughly. The mixture was then centrifuged at 4000 rpm for 10 minutes, causing it to separate into two distinct phases: the organic phase (settling at the bottom) and the aqueous phase (accumulating at the top). The aqueous phase was discarded, and the lower organic fraction was collected for further preparation. To optimize chromatographic performance, 1 µL of the organic fraction was diluted with 990 µL of hexane. The solution was then filtered using a syringe filter to remove any remaining particles and transferred to a GC vial. Finally, the prepared sample was injected into the GC-MS system, where the chemical compounds in the bio-oil were identified and quantified with high sensitivity and precision.

Results:

Impact of Operating Parameters on Bio-Oil Yield (%):

The study aimed to evaluate how different operating conditions influence product yield. The following section discusses the key findings based on various parameters.

Temperature's Impact on Bio-Oil Yield (%):

The maximum bio-oil yield from banana tree waste was observed at a pyrolysis temperature of 500 °C, with a recorded yield of 32.13 wt. %. As the temperature increased from 400 °C to 500 °C, the bio-oil yield rose from 26.60 wt. % to 32.13 wt. %. However, when the temperature was further increased from 500 °C to 600 °C, the bio-oil yield decreased to 28.20 wt. %. This trend indicates that increasing the temperature up to 500 °C enhances bio-oil

production while reducing bio-char yield. However, beyond 500 °C, the yield of bio-oil declines due to secondary cracking. At excessively high temperatures, pyrolysis vapors undergo secondary cracking, which increases gas production while reducing the yields of both bio-oil and bio-char. This explains the drop in bio-oil yield when the temperature rises from 500 °C to 600 °C.

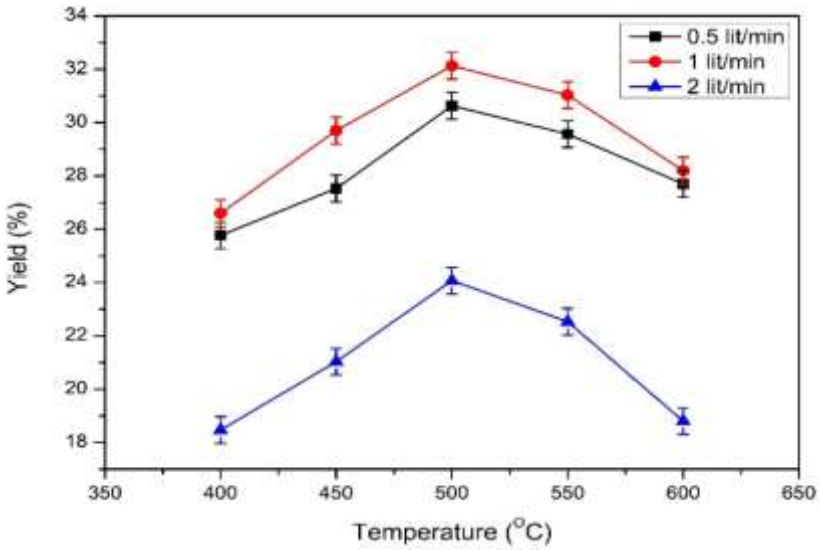


Figure 3. Temperature's impact on Bio-Oil Yield %

Figure 3 illustrates the relationship between temperature, nitrogen flow rates, and bio-oil yield percentage, showcasing the combined impact of these parameters on pyrolysis efficiency. The graph demonstrates that the highest bio-oil yield is achieved at a pyrolysis temperature of 500 °C and a nitrogen flow rate of 1 liter per minute (L/min). This optimal combination enhances the thermal breakdown of organic compounds while maintaining the inert conditions necessary for preventing combustion and ensuring efficient product distribution.

As shown in the figure, increasing the temperature to 500 °C improves bio-oil yield due to enhanced volatilization of the feedstock, while excessively high temperatures lead to a decline in yield due to secondary cracking. Additionally, the nitrogen flow rate plays a crucial role in maintaining a uniform temperature profile and facilitating the escape of volatile compounds, contributing to an optimized pyrolysis environment.

Particle Size's Impact on Bio-Oil Yield (%):

The study revealed that bio-oil yield (%) increased with larger particle sizes, irrespective of the temperature. At lower pyrolysis temperatures, such as 400 °C and 450 °C, smaller particle sizes (0.5 – 1.2 mm) resulted in reduced bio-oil yields. However, as particle size increased to the range of 1.2 – 2.0 mm, a significant improvement in yield was observed, particularly at temperatures between 500 °C and 600 °C. Figure 4 shows the particle size's impact on bio-oil yield percentage.

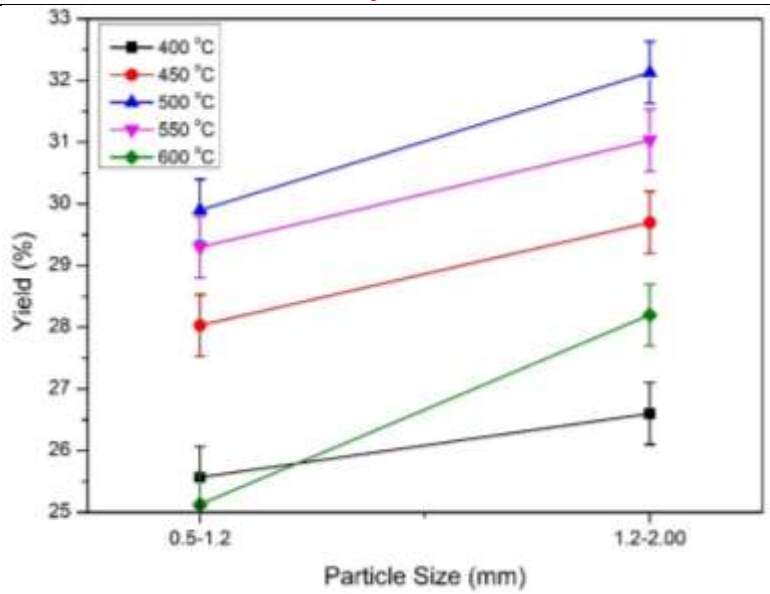


Figure 4. Particle Size's impact on Bio-oil Yield %

Nitrogen flow's impact on the Bio-oil Yield %:

As described and evident in the graph, the flow rate significantly affects the pyrolysis process. The carrier gas flows through the condenser with less contact time, it produces less vapor condensation at higher flow rates, which leads to lower yields. At lower flow rates, the pyrolysis reaction is incomplete, and an undesirable product is formed. In contrast, a higher yield is obtained with a flow rate of 1 liter/min because the vapors in the condenser have more time to interact with the walls, allowing them to condense from gas to liquid. In this study, the yield percentage was found to be higher at a 1 liter/min flow rate compared to 0.5 and 2 liters/min. The nitrogen flow rate was observed to significantly affect the yield percentage, as shown in Figure 5.

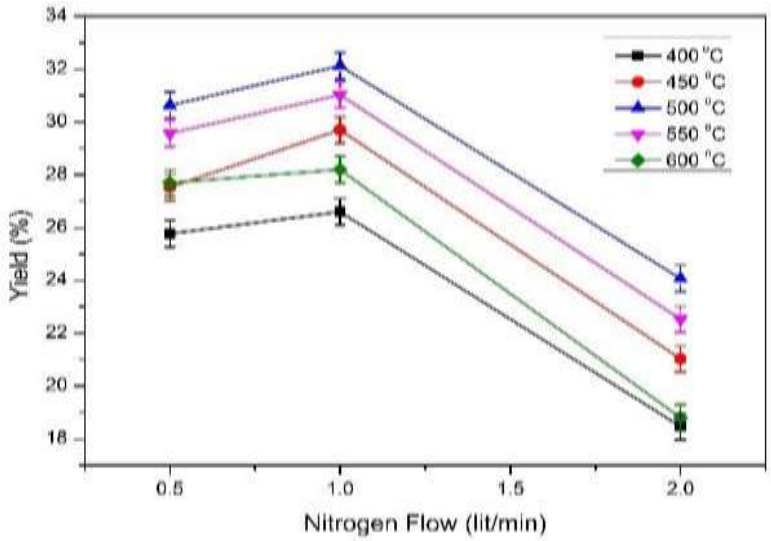


Figure 5. Nitrogen Flow's impact on Bio-oil Yield %

Characterization and Composition of Bio-oil:

The chemical compounds in the bio-oil produced from the pyrolysis of banana tree waste were identified using the GC–MS technique. The results, listed in Table 1, were analyzed at the University of Gurat's Department of Chemistry. GC-MS analysis was performed to investigate the chemical composition of the pyrolysis liquid under optimal conditions. The findings, illustrated in Figure 6, provide a detailed overview of the chemical makeup of the pyrolysis liquid.

Table 1. Chemical compounds in bio-oil

Sr No.	Chemical products identified in bio-oil
1	Acetic acid
2	Propionic acid
3	Phenol
4	Phenol, 3-methyl-
5	Benzene carboxylic acid
6	1,2-Benzenediol
7	D-Allose
8	9,12- Octadecadienoic acid (Z,Z)-
9	Octadecenoic acid
10	10-Octadecenoic acid
11	Tetradecanoic acid
12	Dodecanoic acid

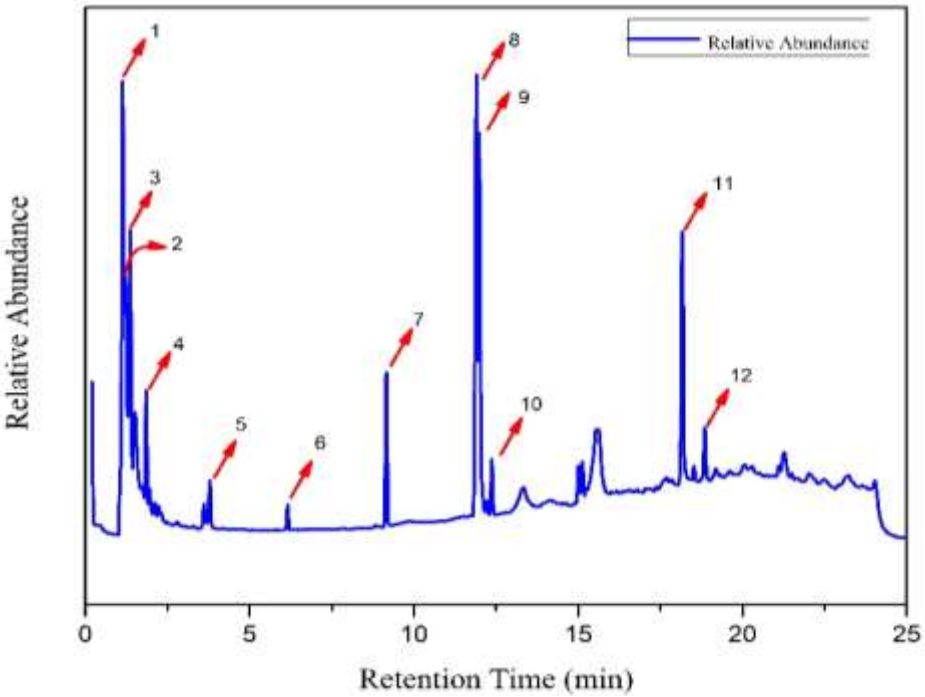


Figure 6. Spectrum of GC-MS analysis

Discussion:

Characterization of Banana tree waste:

The physical and chemical properties of banana tree waste are shown in Table 2, along with relevant research findings on other biomass samples. Understanding the potential for bioenergy production from banana waste requires analyzing its pyrolysis behavior in a fixed-bed reactor. The banana tree waste had a moisture content of 8.21 wt. %, which falls within the acceptable range of 7% to 15% for the pyrolysis process [22]. Higher moisture levels reduce the biomass's heating value, leading to less energy production and bio-oil with higher moisture content, which decreases fuel efficiency [22]. The volatile matter content in banana tree waste was 78.36 wt. %, which is high when compared with other biomass sources like rice husk and wheat straw [8]. Its ash content was 6.43 wt. %, consistent with typical vegetable biomasses, which usually contain between 0.4 wt. % and 22.6 wt. % ash [20]. A higher ash content can impact the yields and quality of the char and bio-oil produced during pyrolysis [20]. The fixed carbon content was 7.02 wt. %, also within the typical range of 7% to 20% for vegetable biomasses [20]. The biomass had a high carbon content (43.10 wt. %) and hydrogen content

(7.19 wt. %), along with low nitrogen (0.15 wt. %) and sulfur (0.24 wt. %). Low sulfur and nitrogen levels are beneficial because they reduce the emission of corrosive and toxic nitrogen and sulfur oxides.

According to Table 2, the higher heating value (HHV) of banana tree waste was 17.728 MJ/kg, comparable to other agricultural biomass sources. A higher HHV means more energy can be generated during pyrolysis [7]. Banana tree waste, like other biomass feedstocks, contains lignocellulosic components that undergo thermal breakdown, releasing vapors and gases during pyrolysis. This process yields liquid and gaseous bio-products [22]. The average cellulose, hemicellulose, and lignin contents in banana tree waste were 30.91%, 25.17%, and 17.53%, respectively, aligning with values reported in earlier studies [22]. The heating rate is a critical parameter in the pyrolysis process, significantly influencing both the yield and quality of bio-oil. Studies have shown that slower heating rates, typically used in fixed-bed reactors, allow for more thorough thermal degradation of biomass, leading to higher bio-oil yields with fewer non-condensable gases and less char production. This is because slower heating provides sufficient time for volatiles to be released and condensed into liquid form, as observed in pyrolysis studies on rice husk and wheat straw, which reported optimized bio-oil yields at moderate heating rates. In contrast, fast pyrolysis methods, often conducted in fluidized-bed reactors, tend to prioritize higher heating rates, which favor the production of lighter hydrocarbons and reduce bio-oil viscosity but can lead to lower overall yields due to secondary cracking of vapors. Additionally, the chemical composition of bio-oil is influenced by heating rates, with higher rates producing bio-oil with greater water content and more unstable oxygenated compounds, while slower rates enhance the formation of phenolic compounds and acids, which improve bio-oil stability. This study, conducted at moderate heating conditions, aligns with existing research that emphasizes the benefits of controlled heating for optimizing bio-oil output and enhancing its chemical characteristics for potential biofuel applications. Future work could explore the synergistic impact of heating rates with catalysts to further improve bio-oil yield and reduce oxygen content.

Products Yield:

The mass yields of pyrolysis products from various biomasses, including date palm waste, rice husk, and banana tree waste, are summarized in Table 3 based on findings from different authors. The bio-oil yield from banana tree waste was found to be 32.13 wt. %. The differences in bio-oil yields between banana tree waste and other biomasses can be attributed to variations in reactor design, experimental setup, processing capacity, and feedstock type. Based on thermal analysis, pyrolysis of banana tree waste was conducted at 500 °C in a fixed-bed reactor. In a separate study, researchers reported that at 525 °C, fast pyrolysis of date palm waste produced 27.4 wt. % bio-oil. However, as the pyrolysis reaction accelerated, bio-oil production decreased [23]. Similarly, another study using rice husk in a fixed-bed reactor produced 30.18 wt. % bio-oil at approximately 500 °C [24].

In related experiments, banana waste pseudo-stem was used to produce bio-oil through fast pyrolysis. The process was conducted at temperatures ranging from 470 °C to 540 °C, with the highest bio-oil yield of 29.4% obtained at 500 °C. The reduced liquid yield at higher temperatures was attributed to secondary cracking of pyrolysis vapors and liquid products, which led to increased gas production [22]. Oxidative fast pyrolysis further reduced bio-oil yields due to biomass volatilization and partial combustion within the reactor, resulting in higher production of water and non-condensable gases.

Table 2. Comparison of characteristics of banana tree waste with another biomass residue

Analysis / Elements	Banana tree waste (This study)	Corn Cob[25]	Sugarcane Bagasse[22]	Rice Husk[25]
Proximate Analysis (wt. %)				
Moisture content	8.21	12.77	10.4	10.89
Ash content	6.43	2.30	16.4	15.14
Volatile matter	78.36	91.16	74.0	73.41
Fixed Carbon	7.02	6.54	13.0	11.44
Ultimate Analysis (wt. %)				
Carbon	43.10	42.10	43.2	41.92
Hydrogen	7.19	5.90	6.70	6.34
Nitrogen	0.15	0.50	0.30	1.85
Sulphur	0.24	0.48	0.20	0.47
Lignocellulosic Analysis (wt. %)				
Cellulose	30.91	42.2	38	32.0
Hemicellulose	25.17	30.7	27	15.0
Lignin	17.53	12.2	19	23.5
Calorific Values (MJ/kg)				
HHV	17.287	16	18	12.87
LHV	15.637	14	17	12.20

Table 3. comparison of Bio-oil yield % of banana tree waste and with another biomass

Biomass residue	Temperature (℃)	Bio-oil yield (wt. %)	Reference
Banana tree waste	500	32.13	This study
Date Palm waste	525	27.4	[23]
Rice Husk	500	30.18	[24]

Another study investigated agricultural biomass, such as sugarcane bagasse, in the pyrolysis process. At an optimal temperature of 525 °C, a maximum bio-oil yield of 33.25 wt. % was obtained. However, as the temperature increased, bio-oil yield decreased, while bio-gas production increased [22]. Previous research also reported a bio-oil yield of 30.18 wt. % from rice husk pyrolysis at 500 °C [24]. Table 3 provides a comparison of bio-oil yields from banana tree waste and those obtained in studies on date palm waste and rice husk pyrolysis. This study aims to explore how factors such as temperature, particle size, and nitrogen flow influence bio-oil yield.

The chemical composition of the bio-oil produced through banana tree waste pyrolysis was identified and analyzed using GC-MS analysis. The chromatograms displayed prominent peaks corresponding to major compounds, while smaller peaks indicated unidentified compounds. The bio-oil from banana tree waste contained various chemical compounds, with phenols being the dominant constituents [26]. Key chemical components in the bio-oil included phenol, 3-methylphenol, and 1,2-benzenediol [27]. Acetic acid and propanoic acid were also detected during the bio-oil analysis [28][29]. Other identified compounds included tetradecanoic acid, 1,2-tetradecanoic acid, benzenedicarboxylic acid, and octadecanoic acid [30]. Additionally, D-allose, 1,2-benzenediol, and 9,12-octadecanoic acid (ZZ) were present in the bio-oil [31].

Conclusion:

This research investigated the pyrolysis of banana tree waste using a fixed-bed reactor. Several operating parameters were studied, including temperature, particle size, and nitrogen flow. Based on the experimental findings, the following conclusions were drawn:

a) Under optimal conditions in a fixed-bed reactor—specifically, a temperature of 500 °C, particle size of 1.2 to 2.0 mm, and nitrogen flow rate of 1 liter/min—the maximum bio-oil yield was achieved. These conditions resulted in a highest bio-oil yield of 32.13 wt. %.

b) Increasing the temperature from 400 to 500 °C raised bio-oil production from 26.60 wt. % to 32.13 wt. %. However, when the temperature increased beyond 500 °C (up to 600 °C), bio-oil production decreased.

c) Higher nitrogen flow rates during the pyrolysis process reduced the bio-oil yield.

d) GC-MS analysis revealed that the bio-oil produced from banana tree waste was rich in phenol groups, acids, alkyl benzene, and several other chemical compounds.

Ongoing research is currently focused on evaluating the impact of heating rates on the pyrolysis process. Future studies will explore the introduction of catalysts to enhance pyrolysis reactions and increase the yield of desired products. Additionally, we plan to integrate pyrolysis with other energy systems to improve overall energy recovery and utilization.

Acknowledgements:

The author thanks GC University Faisalabad for the elemental analysis, and Dr. Muhammad Zubair, Associate Professor of chemistry department, university of Gujrat for GC-MS Analysis

References:

[1] D. W. R. & P.-S. Y. Ahmed I. Osman, Lin Chen, Mingyu Yang, Goodluck Msigwa, Mohamed Farghali, Samer Fawzy, "Cost, environmental impact, and resilience of renewable energy under a changing climate: a review," *Environ. Chem. Lett.*, vol. 21, pp. 741–764, 2023, doi: https://doi.org/10.1007/s10311-022-01532-8.

[2] J. S. Levine, "BIOMASS BURNING AND THE PRODUCTION OF GREENHOUSE GASES," *Clim. Biosph. Interact. Biog. Emiss. Environ. Eff. Clim. Chang.*, pp. 1–13, 2000.

[3] Dr. Supat Wangwongwatana, "Review of Existing Good Practices to Address Open Burning of Agricultural Residues," *United Nations Environ. Program.*, [Online]. Available: https://cleanairweek.org/wp-content/uploads/Review-of-existing-good-practices-to-address-open-burning.pdf

[4] E. Menya *et al.*, "A review of progress on torrefaction, pyrolysis and briquetting of banana plant wastes for biofuels," *Biomass Convers. Biorefinery 2024*, pp. 1–43, Oct. 2024, doi:

10.1007/S13399-024-06204-X.

[5] C. D. G.-T. Sasha Alzate Acevedo, Álvaro José Díaz Carrillo, Edwin Flórez-López, "Recovery of Banana Waste-Loss from Production and Processing: A Contribution to a Circular Economy," *Molecules*, vol. 26, no. 17, p. 5282, 2021, doi: https://doi.org/10.3390/molecules26175282.

[6] K. Alper, K. Tekin, and S. Karagöz, "Pyrolysis of agricultural residues for bio-oil production," *Clean Technol. Environ. Policy*, vol. 17, no. 1, pp. 211–223, Jan. 2015, doi: 10.1007/S10098-014-0778-8/METRICS.

[7] D. Gogoi, M. Kumar, and Y. G. Lakshmi, "A Comprehensive Review on 'Pyrolysis' for Energy Recovery," *BioEnergy Res. 2023 163*, vol. 16, no. 3, pp. 1417–1437, Jan. 2023, doi: 10.1007/S12155-023-10568-9.

[8] J. R. V.-B. Carlos Redondo-Gómez, Maricruz Rodríguez Quesada, Silvia Vallejo Astúa, José Pablo Murillo Zamora, Mary Lopretti, "Biorefinery of Biomass of Agro-Industrial Banana Waste to Obtain High-Value Biopolymers," *Molecules*, vol. 25, no. 17, p. 3829, 2020, doi: https://doi.org/10.3390/molecules25173829.

[9] E. A. Nuraini Nuraini, Noridah binti Osman, "Bio-Oil Production Using Waste Biomass via Pyrolysis Process: Mini Review," *J. Bahan Alam Terberukan*, vol. 11, no. 1, 2022, [Online]. Available: https://journal.unnes.ac.id/nju/index.php/jbat/article/view/37171

[10] H. A. M.-S. Nathaly Ayala-Ruíz, Dionisio Humberto Malagón-Romero, "Exergoeconomic evaluation of a banana waste pyrolysis plant for biofuel production," *J. Clean. Prod.*, vol. 359, p. 132108, 2022, doi: https://doi.org/10.1016/j.jclepro.2022.132108.

[11] A. N. Amenaghawon, C. L. Anyalewechi, C. O. Okieimen, and H. S. Kusuma, "Biomass pyrolysis technologies for value-added products: a state-of-the-art review," *Environ. Dev. Sustain. 2021 2310*, vol. 23, no. 10, pp. 14324–14378, Feb. 2021, doi: 10.1007/S10668-021-01276-5.

[12] S. O. Myeongjong Lee, Myeongjong Lee, "Pyrolysis of Solid Recovered Fuel Using Fixed and Fluidized Bed Reactors," *Molecules*, vol. 28, no. 23, p. 7815, 2023, doi: https://doi.org/10.3390/molecules28237815.

[13] A. Leni Maulinda, Husni Husin, Nasrul Arahman, Cut Meurah Rosnelly, Muhammad Syukri, Nurhazanah, Fahrizal Nasution, "The Influence of Pyrolysis Time and Temperature on the Composition and Properties of Bio-Oil Prepared from Tanjong Leaves (Mimusops elengi)," *Sustainability*, vol. 15, no. 18, p. 13851, 2023, doi: https://doi.org/10.3390/su151813851.

[14] F. U.-N. R. Sánchez - Orozco, Patricia Balderas-Hernández, Gabriela Roa-Morales, "Characterization of Lignocellulosic Fruit Waste as an Alternative Feedstock for Bioethanol Production," BioResources. Accessed: Mar. 17, 2025. [Online]. Available: https://www.researchgate.net/publication/261875575_Characterization_of_Lignocellulosic_Fruit_Waste_as_an_Alternative_Feedstock_for_Bioethanol_Production

[15] N. S. M. A. Rahmad Mohd Taib, Nurhayati Abdullah, "Bio-oil derived from banana pseudo-stem via fast pyrolysis process," *Biomass and Bioenergy*, vol. 148, p. 106034, 2021, doi: https://doi.org/10.1016/j.biombioe.2021.106034.

[16] Designation: E871 − 82, "Standard Test Method for Ash in Wood," *ASTM Int.*, 2013, [Online]. Available: https://tajhizkala.ir/doc/ASTM/E871-82 (Reapproved 2013).pdf

[17] D07.01, "D1102 Standard Test Method for Ash in Wood," ASTM. Accessed: Mar. 17, 2025. [Online]. Available: https://store.astm.org/d1102-84r21.html

[18] Designation: D 1106-96 (Reapproved 2007), "Standard Test Method for Acid-Insoluble Lignin in Wood 1," *Tech. Assoc. Pulp Pap. Ind. Stand. Method T*, 2007, [Online]. Available: https://cdn.standards.iteh.ai/samples/53041/c6b87b2f4e6241bb92c01170846e00d2/ASTM-D1106-96-2007-.pdf

[19] H. Y. Dijan SUPRAMONO, Eny KUSRINI, "Yield and Composition of Bio-oil from Co-Pyrolysis of Corn Cobs and Plastic Waste of HDPE in a Fixed Bed Reactor," *J. Japan Inst. Energy*, vol. 95, no. 8, pp. 621–628, 2016, doi: https://doi.org/10.3775/jie.95.621.

[20] J. P. Maria Gonzalez, Juan Pereira-Rojas, Ivan Villanueva, Bari Agüero, Iris Silva, Ingrid Velasquez, Blass Delgado, Javier Hernandez, Gleen Rodrigue, Henry Labrador, Haydn Barros, "Preparation and characterization of cellulose fibers from Meghatyrsus maximus: Applications in its chemical derivatives," *Carbohydr. Polym.*, vol. 296, p. 119918, 2022, doi: https://doi.org/10.1016/j.carbpol.2022.119918.

[21] "Carbohydrate Polymers," ScienceDirect.com by Elsevier. Accessed: Mar. 17, 2025. [Online]. Available: https://www.sciencedirect.com/journal/carbohydrate-polymers

[22] O. S. Noeli Sellin, Diego Ricardo Krohl, Cintia Marangoni, "Oxidative fast pyrolysis of banana leaves in fluidized bed reactor," *Renew. Energy*, vol. 96, pp. 56–64, 2016, doi: https://doi.org/10.1016/j.renene.2016.04.032.

[23] G. Bensidhom, A. Ben Hassen-Trabelsi, M. Sghairoun, K. Alper, and I. Trabelsi, "Pyrolysis of Tunisian Date Palm Residues for the Production and Characterization of Bio-Oil, Bio-Char and Syngas," *Adv. Sci. Technol. Innov.*, pp. 1561–1563, 2018, doi: 10.1007/978-3-319-70548-4_453.

[24] F. H. K. Khairuddin Md Isa, Suhardy Daud, Nasrul Hamidin, Khudzir Ismail, Saiful Azhar Saad, "Thermogravimetric analysis and the optimisation of bio-oil yield from fixed-bed pyrolysis of rice husk using response surface methodology (RSM)," *Ind. Crops Prod.*, vol. 33, no. 2, pp. 481–487, 2011, doi: https://doi.org/10.1016/j.indcrop.2010.10.024.

[25] T. B. Bijoy Biswas, Nidhi Pandey, Yashasvi Bisht, Rawel Singh, Jitendra Kumar, "Pyrolysis of agricultural biomass residues: Comparative study of corn cob, wheat straw, rice straw and rice husk," *Bioresour. Technol.*, vol. 237, pp. 57–63, 2017, doi: https://doi.org/10.1016/j.biortech.2017.02.046.

[26] A. N. S. Sanjay Singh, Ankita Tagade, Ashish Verma, Ajay Sharma, Shyam P. Tekade, "Insights into kinetic and thermodynamic analyses of co-pyrolysis of wheat straw and plastic waste via thermogravimetric analysis," *Bioresour. Technol.*, vol. 356, p. 127332, 2022, doi: https://doi.org/10.1016/j.biortech.2022.127332.

[27] R. T. Marcin Landrat, Mamo T. Abawalo, Krzysztof Pikoń, "Bio-Oil Derived from Teff Husk via Slow Pyrolysis Process in Fixed Bed Reactor and Its Characterization," *Energies*, vol. 15, no. 24, p. 9605, 2022, doi: https://doi.org/10.3390/en15249605.

[28] V. V. G. Das, Sutapa, "RSM-optimised slow pyrolysis of rice husk for bio-oil production and its upgradation," *Energy*, vol. 225, p. 120161, 2021, doi: https://doi.org/10.1016/j.energy.2021.120161.

[29] A. S. Paul, N. L. Panwar, B. L. Salvi, S. Jain, and D. Sharma, "Experimental investigation on the production of bio-oil from wheat straw," *Energy Sources, Part A Recover. Util. Environ. Eff.*, Dec. 2024, doi: 10.1080/15567036.2020.1779416.

[30] M. O. Gmar Bensidhom, Miriam Arabiourrutia, Aïda Ben Hassen Trabelsi, Maria Cortazar, Selim Ceylan, "Fast pyrolysis of date palm biomass using Py-GCMS," *J. Energy Inst.*, vol. 99, pp. 229–239, 2021, doi: https://doi.org/10.1016/j.joei.2021.09.012.

[31] T. B. Yassir Makkawi, Yehya El Sayed, Mubarak Salih, Paul Nancarrow, Scott Banks, "Fast pyrolysis of date palm (Phoenix dactylifera) waste in a bubbling fluidized bed reactor," *Renew. Energy*, vol. 143, pp. 719–730, 2019, doi: https://doi.org/10.1016/j.renene.2019.05.028.

Efficient Region-Based Video Text Extraction Using Advanced Detection and Recognition Models

Naveed Ahmed[1], Zahid Iqbal[2*], Abdullah Nawaz[1], Huah Yong Chan[3], Fatima N. AL-Aswadi[4,5], Hafiz Usman Zia[6]

[1]Smart Zone Leaders, Kharian, Dist. Gujrat, Pakistan.

[2]Department of Computer Science, Air University, Kharian, Pakistan.

[3]School of Computer Sciences, University Sains Malaysia, Pulau Pinang, Malaysia

[4]Institute of Computer Science & Digital Innovation (ICSDI), UCSI University, Kuala Lumpur 56000, Malaysia.

[5]Faculty of Computer Science and Engineering, Hodeidah University, Al Hudaydah, Yemen

[6]Department of Information Technology, Faculty of Computing and IT, University of Gujrat

***Correspondence:** zahid.iqbal@kc.au.edu.pk

Citation | Ahmad. N, Iqbal. Z, Nawaz. A, Chan. H. Y, Aswadi. F. N, Zia. H. U, "Efficient Region-Based Video Text Extraction Using Advanced Detection and Recognition Models", IJIST, Special Issue. pp 31-45, March 2025

Received| Feb 10, 2025 **Revised|** Feb 23, 2025 **Accepted|** Feb 28, 2025 **Published|** March 03, 2025.

This paper presents an automated process for extracting text from video frames by specifically targeting text-rich regions, identified through advanced scene text detection methods. Unlike traditional techniques that apply OCR to entire frames—resulting in excessive computations and higher error rates—our approach focuses only on textual areas, improving both speed and accuracy. The system integrates effective preprocessing routines, cutting-edge text detectors (CRAFT, DBNet), and advanced recognition engines (CRNN, transformer-based) within a unified framework. Extensive testing on datasets such as ICDAR 2015, ICDAR 2017 MLT, and COCO-Text demonstrates consistent gains in F-scores and word recognition rates, significantly outperforming baseline methods. Additionally, detailed error analysis, ablation studies, and runtime evaluations offer deeper insights into the strengths and limitations of the proposed method. This pipeline is particularly useful for tasks like video indexing, semantic retrieval, and real-time multimedia analysis.

Keywords: Optical Character Recognition, Scene Text Detection, Scene Text Recognition, Video Analysis, Deep Learning

Introduction:

Machine learning (ML) has advanced rapidly across various fields, driving progress in recognition systems, optimization methods, optical character recognition (OCR) technologies, and security frameworks [1], [2], [3], [4], [5], [6], [7]. These developments provide a solid foundation for applying AI and ML to enhance the accuracy, fairness, and efficiency of automated decision-making systems. As digital video content continues to grow on streaming platforms, instructional archives, and media-sharing websites, the demand for effective text recognition and extraction from video frames has become increasingly important. Text in video frames—such as signs, subtitles, or labels—is crucial for applications like content summarization, automated captioning, semantic retrieval, and video indexing [8], [9], [10]. However, traditional OCR methods often process the entire frame, which is inefficient due to background clutter and irrelevant details. This not only increases the computational load but also raises error rates.

Recent advances in deep learning have addressed this issue by focusing on text-dense regions. Modern detection models [11], [12] and transformer-based recognition systems [13], [14] are improving accuracy across different fonts, scripts, and text orientations. This research presents an enhanced video text extraction pipeline that targets only the text-rich areas of each frame, boosting efficiency and minimizing errors from unnecessary sections. The exponential growth of video content across various platforms, including educational archives, media-sharing websites, and streaming services, has intensified the demand for efficient and accurate text extraction from video frames. Text appearing in videos often carries critical semantic information, such as subtitles, annotations, signage, or scene labels, which can facilitate tasks like video indexing, content retrieval, and automated captioning. Traditional Optical Character Recognition (OCR) methods, while effective in document analysis, struggle to handle the complexities of scene text in videos due to diverse fonts, orientations, multilingual scripts, and background clutter. This has prompted researchers to develop advanced, deep learning-based frameworks that focus on identifying text-rich regions, thereby minimizing unnecessary computations and improving the reliability of extracted text.

Recent advancements in deep learning, particularly in scene text detection and recognition, have introduced new possibilities for enhancing the accuracy and efficiency of video text extraction. By employing models that leverage character-level awareness, differentiable binarization, and attention mechanisms, modern pipelines can overcome challenges associated with text distortion, low contrast, and multi-oriented scripts. However, continuous video streams present additional hurdles, such as handling temporal variations in text, managing computational overhead for real-time applications, and minimizing false positives caused by dynamic backgrounds.

Key contributions of our work are as follows:

In-Depth Analysis: We thoroughly examine performance limits and trade-offs by providing accurate error measurements, conducting ablation studies on preprocessing techniques, and analyzing runtime performance.

Comprehensive Evaluation: Our approach is benchmarked against ICDAR 2015, ICDAR 2017 MLT, and COCO-Text datasets, where it outperforms existing baseline methods.

Region-Based Approach: By using state-of-the-art (SOTA) detectors such as CRAFT and DBNet, we isolate text-dense regions, which minimizes the effect of non-text background noise and enhances text extraction accuracy.

Advanced Text Recognition: We further employ SOTA text recognizers like CRNN and transformer-based models to handle complex text patterns and diverse script styles more effectively.

Objectives of the Study:

The primary objectives of this study are as follows:

1. To develop an efficient region-based video text extraction pipeline that improves the accuracy and speed of text detection and recognition by focusing on text-rich regions, thereby reducing computational overhead compared to traditional full-frame OCR methods.
2. To evaluate the performance of advanced text detection models, such as CRAFT (Character Region Awareness for Text Detection) and DBNet (Differentiable Binarization Network), in accurately localizing text in complex video frames with varied fonts, orientations, and backgrounds.
3. To assess the effectiveness of deep learning-based text recognition models, including CRNN (Convolutional Recurrent Neural Network) and a transformer-based recognizer, in handling curved, multilingual, and stylized text extracted from video frames.
4. To implement and analyze preprocessing techniques, such as grayscale conversion, adaptive binarization, and noise reduction, to enhance text clarity and improve the accuracy of detection and recognition.
5. To benchmark the proposed pipeline on established datasets, including ICDAR 2015, ICDAR 2017 MLT, and COCO-Text, and compare its performance (in terms of F-score, Character Recognition Accuracy, and Word Recognition Rate) with baseline and reference methods.

Literature Review:

Earlier video text extraction techniques mainly relied on traditional OCR engines and heuristic-based localization methods, which struggled with complex layouts, diverse fonts, and irregular text orientations [15]. With the rise of deep learning, more advanced scene text detectors emerged. EAST [16] introduced a fast, regression-based approach, while CRAFT [11] improved recall by utilizing character-level cues and affinity representations. DBNet [12] further enhanced precision and stability by incorporating differentiable binarization. On the recognition front, Tesseract [17] gained popularity as an OCR tool in conventional applications. However, scene text posed additional challenges, requiring more advanced solutions. CRNN [13] combined convolutional and recurrent layers to adapt to curved and multi-oriented text lines. Transformer-based models [14], [18] introduced attention mechanisms, allowing them to handle multilingual text and various typographical styles.

Recent frameworks have started integrating detection and recognition into unified pipelines [19], [20]. While these methods show promise, applying them directly to continuous video content remains computationally demanding. Our approach refines the region-based method by focusing on text-rich areas, achieving both higher accuracy and improved efficiency for large-scale video analysis tasks. Beyond standalone OCR and text detection pipelines, recent research has focused on context-aware extraction, which incorporates semantic understanding of text within the video's visual and temporal context. Multi-frame approaches have been proposed to improve robustness by aggregating information across consecutive video frames [21], [22]. These methods help mitigate issues like low resolution, motion blur, and occlusions, which are common in dynamic video environments. However, their increased accuracy often comes at the cost of slower processing speeds, creating a trade-off between precision and computational efficiency. Furthermore, hybrid techniques that combine rule-based post-processing with deep learning models have been explored to improve text coherence and alignment [23]. These methods leverage domain-specific knowledge, such as recognizing text patterns within scene elements like street signs, subtitles, or license plates, to enhance extraction accuracy. Although effective for specific use cases, such techniques often suffer from reduced generalizability when applied to varied video content.

To address these challenges, research has also shifted toward lightweight models optimized for real-time applications. Techniques such as knowledge distillation, model

pruning, and quantization have been used to reduce the size and complexity of deep learning models without compromising performance. Such advancements are particularly relevant for real-time video text extraction tasks in resource-constrained environments like mobile devices or embedded systems. By building upon these developments, our work seeks to enhance both detection and recognition stages while maintaining computational efficiency. By refining region-based approaches and leveraging state-of-the-art models, we aim to improve accuracy, reduce background noise, and streamline large-scale video text extraction.

Methodology:

This section explains the complete process used to identify and extract relevant text from video content. The framework is designed to balance accuracy, speed, and flexibility to handle various visual situations. As shown in Figure 1, the workflow moves through several key stages: sampling frames from the video, applying a customized preprocessing method, detecting areas containing text, using advanced algorithms to recognize the extracted text, and performing post-processing to refine and organize the final output. Each step is explained in detail, along with the reasons for its inclusion.

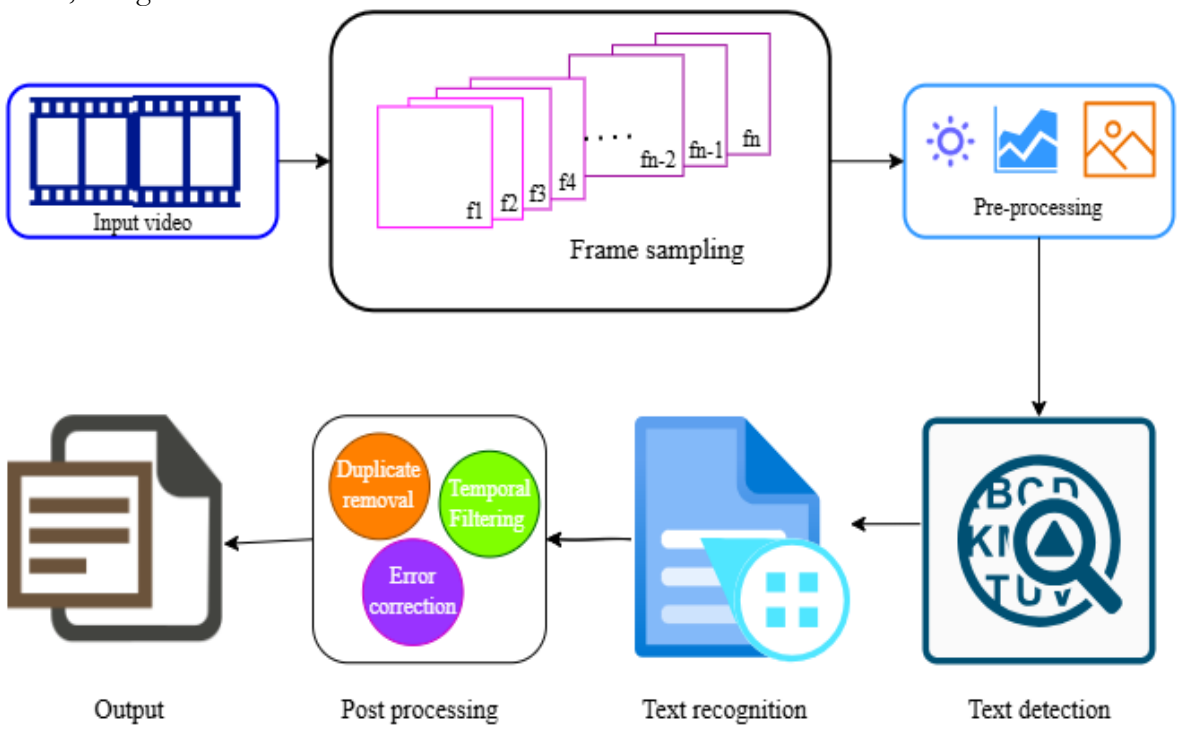


Figure 1. Illustration of System Diagram

Overall System Architecture:

The system processes a continuous stream of video frames, selecting frames at a controlled sampling rate to reduce computational load. Once a frame is extracted, it goes through several enhancement steps designed to make text clearer. A text detection module then scans the frame to identify areas likely to contain useful text. These selected areas are sent to the text recognition stage. In the final step, post-processing refines and organizes the recognized text for practical use. Throughout the process, the goal is to minimize unnecessary computations, ensuring both efficiency and broad usability.

Frame Sampling Strategy:

An important part of the system's design is deciding how often to extract frames from the video. If too many frames are sampled, the system wastes time processing redundant data. On the other hand, sampling too few frames risks missing brief but important text. Based on initial tests, we chose to extract two frames per second. This strikes a practical balance by

capturing changes in text without overloading the system. For example, rapid captions in educational videos might require more frequent sampling, while lecture recordings or surveillance footage can work well with less frequent sampling.

Preprocessing Pipeline:

Preprocessing is a crucial step that improves each frame before it reaches the text detection and recognition stages. The goal is to highlight text while reducing distractions from the background. As shown in Figure 2, the preprocessing workflow includes four main steps: converting the frame to grayscale, enhancing contrast, applying adaptive binarization, and removing noise. These steps create a cleaner, text-focused image that helps modern OCR models deliver better results.

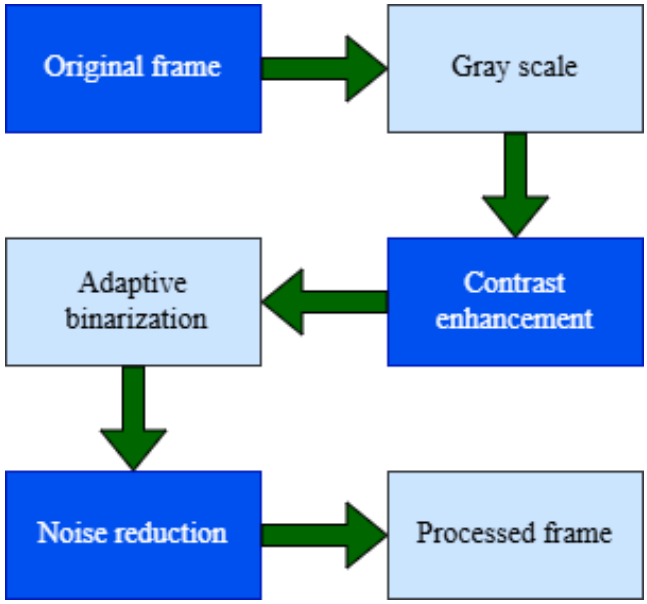


Figure 2. Preprocessing steps

Grayscale Conversion:

The image data is converted to grayscale by reducing it from full color to a single-color channel. Since the brightness of text often differs from its background, representing the image in grayscale makes it easier to apply binarization and thresholding techniques later. Additionally, this reduces computational costs by limiting the input to a single channel.

Contrast Enhancement:

After converting the image to grayscale, text may still appear unclear due to low contrast, especially when displayed against dark or patterned backgrounds. Methods like Contrast Limited Adaptive Histogram Equalization (CLAHE) play a key role in solving this issue by improving text visibility. These techniques redistribute pixel intensity values, enhancing fine details and making faded characters more distinct.

Adaptive Binarization:

Unlike global thresholding, adaptive binarization calculates a local threshold based on the intensity values of surrounding pixels. This method is particularly useful for handling images with low lighting or complex backgrounds. The local threshold is computed as:

$$T(x,y) = mean\big(I(x',y')\epsilon\, N(x,y)\big) - C$$

Where $B(x,y)$ is given by:

$$B(x,y) = \begin{cases} 1\ if\ I(x,y) > T(x,y) \\ 0\ if\ I(x,y) \leq T(x,y) \end{cases}$$

The output generated is a binary image, where the text appears as a bright foreground against a darker background. This creates a clear distinction between the text and irrelevant

details, making it an optimized representation for deep learning detectors to identify textual patterns more effectively.

Noise Reduction:

Real-world video frames often contain noise, glitches, or textures that can mislead text detectors. To address this, a noise reduction filter is applied. Filters such as Gaussian or median smoothing reduce pixel-level variations, improving the visibility of essential text edges. By enhancing text clarity, these techniques strengthen the detector's ability to identify text in diverse environments. Together, these four preprocessing techniques enhance textual clarity in video frames. Empirical studies show that this pipeline improves text detection accuracy while minimizing false alarms, particularly in challenging conditions where text blends into the background or appears in low light.

Text Detection:

After preprocessing, the system must determine which areas are likely to contain text. This is achieved using text detection algorithms, as analyzing the entire frame with OCR could extract irrelevant details. In this work, we utilize two prominent methods—CRAFT (Character Region Awareness for Text Detection) and DBNet (Differentiable Binarization Network)—both known for their robust performance in localizing scene text.

Character Region Awareness (CRAFT):

CRAFT estimates bounding boxes and assigns affinity scores to link them into coherent text lines or phrases. By focusing on character-level details, CRAFT handles the complexity of video frames and effectively manages text in unusual orientations, including angled, curved, and thin segments. The output is a set of precise bounding polygons, which reduces the data passed to the recognition phase.

Differentiable Binarization (DBNet):

DBNet simplifies the cropping and recognition process by using a binarization layer to transform feature maps into sharp text representations. This approach excels in challenging scenarios, such as densely packed characters, by isolating text instances and generating bounding boxes and contours. Following detection, the pipeline produces bounding boxes for each frame, with each box representing a distinct text area. Identifying these areas early helps process only text-rich segments, thereby reducing computational complexity and minimizing recognition errors.

Text Recognition Models:

After isolating text regions, they must be converted into machine-readable text. Non-standard or complex fonts pose challenges for conventional OCR methods, but deep learning-based recognizers handle a broader range of text variations. For this task, we adopt two types of recognition models: CRNN (Convolutional Recurrent Neural Network) and a transformer-based recognizer, both of which are well-known for managing linguistic and typographic complexities.

CRNN:

CRNN combines convolutional layers for feature extraction with bidirectional recurrent layers. This design effectively handles naturally curved or rotated text of variable lengths. The clipped text area is transformed by CRNN into a sequence of features, which are decoded into characters or sub-word units by the recurrent layers. By integrating spatial and sequential context, CRNN demonstrates strong performance on standard benchmarks and real-world video text scenarios.

Transformer-Based Recognizer:

Unlike CRNN, transformer-based models rely on self-attention mechanisms to capture character-level dependencies without using recurrent layers. These models often achieve higher accuracy, especially on text samples with complex fonts, unusual orientations, or multilingual scripts. The transformer processes feature from each text region, attending to different parts of the input sequence to generate a coherent textual output. Though more computationally intensive, transformers frequently deliver superior recognition accuracy. Both approaches convert visual text segments into fully transcribed strings. Our initial trials indicate that CRNN offers an excellent balance between speed and accuracy, while the transformer model provides slightly better accuracy at the cost of increased computational demand. The choice of recognizer depends on the application's latency requirements and available computational resources.

Post-Processing and Text Consolidation:

After text recognition, the system produces raw text segments from each sampled frame. This output may include duplicates, partial phrases, or minor OCR errors. To create a coherent final output, a post-processing module performs the following key functions:

Duplicate Removal and Temporal Filtering:

When text persists on-screen for several seconds, consecutive frames may produce overlapping or identical text segments. The system detects and consolidates these duplicates. If needed, heuristics align text snippets with their temporal position in the video, creating a stable transcript synchronized with the video's timeline.

Common Error Correction:

Some OCR errors, such as confusing the digit '0' with the letter 'O,' are common in challenging conditions. Rule-based corrections or dictionary filtering can mitigate these errors. For further refinement, contextual language models or spell-checkers may be integrated, though these methods are not the focus of this study.

Output Formatting:

The cleaned text is formatted according to the intended application. For semantic indexing or retrieval systems, the output may be stored as timestamped metadata, linking each text snippet to the corresponding video segment. In other cases, it may be formatted as subtitles or transcriptions for viewing alongside the video.

Algorithm: getTextFromVideo(video_path, sampling_rate)

```
frames = empty_list
final_texts = empty list
video = LoadVideo(your_video_path)
For each frame in video (selecting frame based on 'sampling_rate'):
        Add the current frame to the 'frames' list
For each frame in frames:
        Convert the frame to grayscale
        Enhance the contrast using CLAHE
        Apply adaptive binarization to emphasize text
        Reduce any noise in the frame
Detect text regions with the CRAFT or DBNet model
For each text region in the merged set:
        Use the CRNN or Transformer-based model to recognize text from the region
        Add recognized text to 'recognized_text_in_frames'
```

Remove any duplicate texts from recognized_text
Correct common OCR mistakes (like confusing 'O' with '0')
Format the text for clear output
Return 'final_texts'
Call get Text from Video with the video file path and frame rate

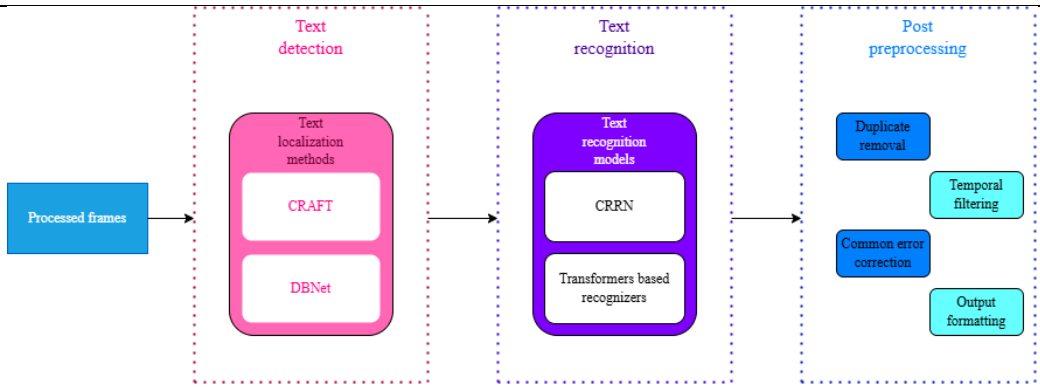


Figure 3. Pipeline for text extraction from video frames using detection, recognition, and post-processing techniques.

Implementation Details and Integration:

Efficient execution and seamless integration of each component are essential throughout the methodology. The pipeline is implemented in Python, utilizing popular deep learning frameworks and libraries for image processing. Pretrained weights for CRAFT and DBNet are fine-tuned using a subset of training images that represent typical video frames. Similarly, CRNN and the transformer-based model undergo light fine-tuning on domain-specific data, such as educational videos and broadcast footage, to enhance their performance in the target scenarios.

At each stage, intermediate outputs—such as pre-processed frames, detection bounding boxes, and recognized text strings—are stored and passed efficiently between modules to reduce latency. To further accelerate processing, parallelization techniques can be applied, such as running detection and recognition tasks on separate GPU streams. Additionally, the hardware setup, including RTX-series GPUs, ensures that even large video collections are processed within a reasonable time frame.

Experimental Studies and Discussions:

Datasets and Evaluation Metrics:

We evaluated the pipeline using three established benchmarks:

ICDAR 2015 [21]**:** This dataset contains incidental scene text with distortions and complex backgrounds.

ICDAR 2017 MLT [22]**:** A multilingual dataset designed to test adaptability across different scripts and languages.

COCO-Text [23]**:** A large-scale dataset with significant diversity in text appearance and background clutter.

For text detection, we follow standard protocols, counting bounding boxes as correct matches if the Intersection over Union (IoU) is ≥ 0.5. Detection performance is evaluated using Precision (P), Recall (R), and the F-score (F). For text recognition, we measure Character Recognition Accuracy (CRA) and Word Recognition Rate (WRR) to assess how closely the transcribed text matches the ground truth.

Implementation Details:

All experiments were conducted on a machine running Ubuntu 20.04, equipped with an Intel Xeon W-2255 CPU, 64 GB of RAM, and two NVIDIA RTX 3090 GPUs. We used PyTorch for model fine-tuning and inference. Pretrained weights for CRAFT and DBNet were fine-tuned using a subset of training samples, while CRNN and the transformer-based recognizer were similarly adapted to maximize performance in the target domain.

Detection Results:

Table 1 compares the detection performance of EAST [16], CRAFT [11], and DBNet [12] across the three benchmark datasets.

Table 1. Detection Performance

Dataset	Method	P (%)	R (%)	F (%)
ICDAR 2015	EAST	80.4	75.9	78.1
	CRAFT	87.2	85.7	86.4
	DBNet	88.5	87.3	87.9
ICDAR 2017 MLT	EAST	73.3	68.5	70.8
	CRAFT	82.0	79.8	80.9
	DBNet	84.6	82.5	83.5
COCO-Text	EAST	68.5	66.1	67.3
	CRAFT	79.4	76.6	78.0
	DBNet	81.7	79.9	80.8

Table 1 and Figure 4 show that both CRAFT and DBNet outperform EAST, with DBNet achieving slightly higher F-scores. The comparison highlights significant advancements in video text localization, with DBNet consistently emerging as the most accurate model. It achieves the highest F-scores across all datasets. For example, on the ICDAR 2015 dataset, DBNet records an impressive F-score of 87.9%, surpassing CRAFT (86.4%) and EAST (78.1%). This superior performance can be attributed to DBNet's differentiable binarization layer, which enhances its ability to effectively isolate text contours, even in densely packed or low-contrast environments.

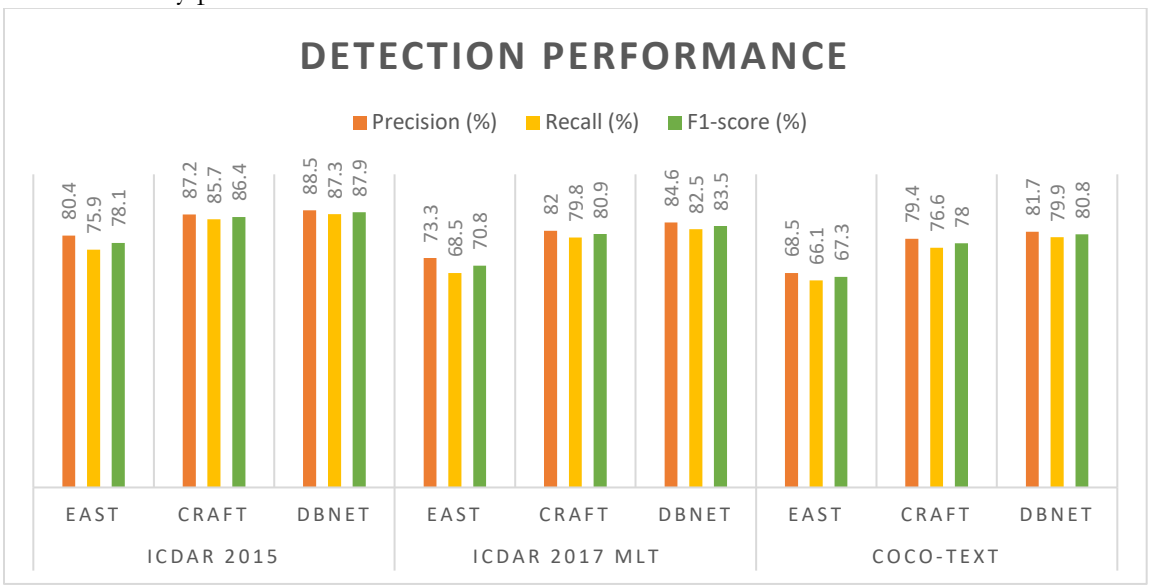


Figure 4. Text detection performance comparison across datasets using DBNet EAST and, CRAFT

CRAFT also demonstrates strong performance, especially on datasets containing multilingual text and irregular orientations. Notably, it achieves an F-score of 80.9% on the

ICDAR 2017 MLT dataset, which can be attributed to its character-level awareness mechanism that enables precise localization of text regions. In contrast, EAST, while efficient, struggles with complex backgrounds and irregular text orientations, as evident from its comparatively lower F-score of 67.3% on the COCO-Text dataset. These findings emphasize the importance of using advanced text detection methods, particularly when dealing with video content featuring challenging text characteristics.

Recognition Performance (Cropped Patches):

Table 2 provides a comparison of Tesseract [17], CRNN [13], and a transformer-based model [14] on isolated text patches.

Table 2. Recognition on Cropped Text

Dataset	Recognizer	CRA (%)	WRR (%)
ICDAR 2015	Tesseract	90.8	86.1
	CRNN	96.2	92.7
	Transformer	97.5	95.1
ICDAR 2017 MLT	Tesseract	85.4	80.6
	CRNN	92.1	88.9
	Transformer	94.6	92.3
COCO-Text	Tesseract	78.9	72.5
	CRNN	88.4	83.7
	Transformer	91.0	86.9

The CRNN and transformer models significantly outperform Tesseract, with the transformer model achieving slightly better results. Among all datasets, the transformer-based recognizer demonstrates the highest performance, achieving a Character Recognition Accuracy (CRA) of 97.5% and a Word Recognition Rate (WRR) of 95.1% on the ICDAR 2015 dataset. This impressive accuracy is due to its self-attention mechanisms, which help it efficiently manage long-range dependencies and handle complex scripts, including multilingual and stylized text.

The CRNN model follows closely, with a CRA of 96.2% and a WRR of 92.7% on the same dataset. Its hybrid design, which combines convolutional and recurrent layers, enables it to handle curved and multi-oriented text lines effectively. In contrast, Tesseract, though a well-established OCR tool, struggles in these challenging scenarios, delivering lower CRA and WRR scores (e.g., 90.8% and 86.1% on ICDAR 2015). This performance gap highlights the limitations of traditional OCR engines in handling the complexities of real-world video text.

End-to-End Results: We evaluate three complete pipelines

Baseline: EAST + Tesseract

Proposed: CRAFT + CRNN

Reference: DBNet + Transformer

As shown in Table 3 and Figure 5, evaluating the end-to-end pipelines demonstrates the clear advantages of integrating advanced detection and recognition models. The proposed pipeline (CRAFT+CRNN) significantly outperforms the baseline setup (EAST+Tesseract), achieving an F-score of 86.7% and a Word Recognition Rate (WRR) of 92.7% on the ICDAR 2015 dataset, compared to the baseline's F-score of 76.3% and WRR of 86.1%. These gains emphasize the effectiveness of CRAFT's character-region awareness in minimizing background noise and CRNN's capability to accurately interpret text regions. The DBNet+Transformer pipeline delivers the best overall performance, achieving an F-score of 88.5% and a WRR of 95.1% on the ICDAR 2015 dataset. However, due to its higher computational demands, the proposed pipeline offers a more practical solution for scenarios

where computational resources are limited. It outperforms the baseline across all datasets and closely matches the performance of the reference system, demonstrating its robustness and adaptability.

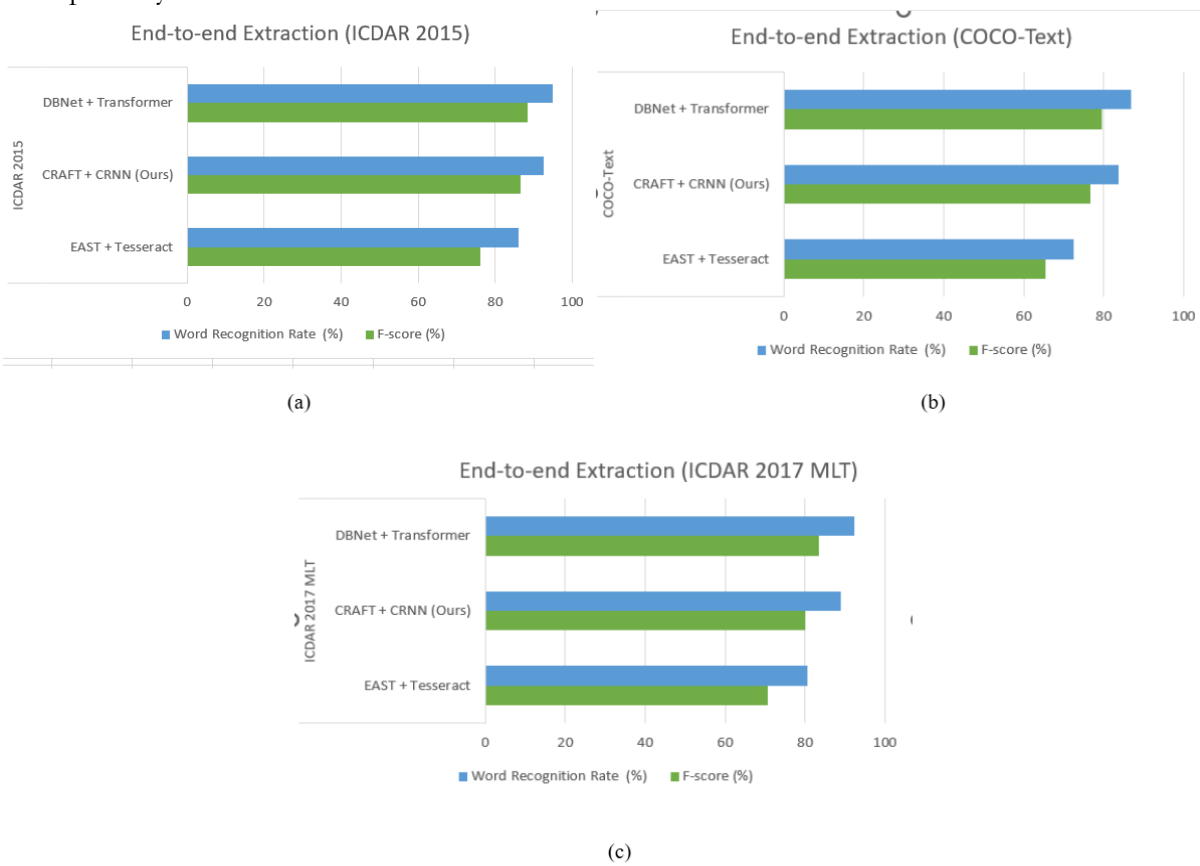


Figure 5. Performance of End-to-end text extraction using different detection and recognition models on (a) ICDAR 2015 dataset (b) COCO-Text dataset (c) ICDAR 2017 MLT dataset

Table 3. End-to-End Extraction

Dataset	Method	F (%)	WRR (%)
ICDAR 2015	EAST + Tesseract	76.3	86.1
	CRAFT + CRNN (Ours)	86.7	92.7
	DBNet + Transformer	88.5	95.1
ICDAR 2017 MLT	EAST + Tesseract	70.6	80.6
	CRAFT + CRNN (Ours)	80.2	88.9
	DBNet + Transformer	83.5	92.3
COCO-Text	EAST + Tesseract	65.5	72.5
	CRAFT + CRNN (Ours)	76.8	83.7
	DBNet + Transformer	79.6	86.9

Computational Efficiency: We measure processing time per frame (PTF) and GPU memory (MF) in Table 4.

Table 4. Efficiency

Method	PTF (ms/frame)	MF (GB)
EAST + Tesseract	72	1.9
CRAFT + CRNN (Ours)	95	2.5
DBNet + Transformer	120	3.1

The computational efficiency of the proposed pipeline is evident from its balanced performance in terms of processing time per frame (PTF) and GPU memory usage. With a PTF of 95 ms/frame and a memory footprint of 2.5 GB, it shows a clear improvement over the baseline while being more efficient than the reference system (120 ms/frame, 3.1 GB). This balance makes the proposed pipeline particularly well-suited for applications that demand both accuracy and scalability, such as real-time video analysis and multimedia indexing. Although it is slightly more resource-intensive than the baseline, it achieves a better trade-off between accuracy and efficiency compared to the top-performing reference pipeline.

Error Analysis: Table 5 categorizes errors on the COCO-Text dataset, highlighting issues such as similar character confusions, case errors, and missed words.

Table 5. Error Analysis (COCO-Text)

Method	Similar Char (%)	Case Errors (%)	Missed Words (%)
EAST + Tesseract	7.2	5.1	15.3
CRAFT + CRNN (Ours)	3.9	2.7	8.5
DBNet + Transformer	3.1	2.3	6.9

The analysis of errors on the COCO-Text dataset offers valuable insights into the challenges faced by OCR systems in real-world applications. The proposed pipeline (CRAFT+CRNN) effectively reduces error rates compared to the baseline. For example, similar character confusions drop from 7.2% to 3.9%, case errors decrease from 5.1% to 2.7%, and missed words decline from 15.3% to 8.5%. These improvements demonstrate the framework's ability to handle noisy and complex text environments with greater accuracy. Although the DBNet+Transformer pipeline reduces errors even further, it comes with a higher computational cost, emphasizing the practical benefits of the proposed approach. Overall, our pipeline significantly lowers all error types compared to the baseline.

Preprocessing Ablation:

Table 6 highlights how each preprocessing step impacts performance on the ICDAR 2015 dataset. The ablation study shows that each step plays a key role in improving the system's overall performance. For instance, applying adaptive binarization increases the F-score from 85.8% (using grayscale only) to 86.7%. This illustrates how preprocessing enhances input frame stability and clarifies textual features. Each incremental improvement underscores the importance of preprocessing in reducing false positives and boosting text detection and recognition accuracy.

Table 6. Preprocessing Ablation (ICDAR 2015)

Preprocessing	CRA (%)	WRR (%)	F (%)
None	91.5	87.0	83.9
Grayscale Only	93.4	89.1	85.8
+Contrast Enhancement	94.1	90.2	86.2
+Adaptive Binarization	96.2	92.7	86.7

Discussion:

The findings of this study demonstrate that the proposed region-based video text extraction pipeline significantly enhances the accuracy, efficiency, and robustness of text detection and recognition compared to conventional methods. By integrating advanced text detectors (CRAFT, DBNet), sophisticated recognition models (CRNN, transformer-based), and a strategic preprocessing pipeline, the framework effectively reduces noise, eliminates redundant computations, and improves the clarity of extracted text. This approach is particularly advantageous in complex video environments where text may appear in various fonts, orientations, and lighting conditions. Comparative evaluations on benchmark datasets, including ICDAR 2015, ICDAR 2017 MLT, and COCO-Text, underscore the pipeline's

superiority, achieving higher F-scores, Word Recognition Rates (WRR), and Character Recognition Accuracy (CRA) than baseline configurations.

The choice of detection models plays a crucial role in the system's performance. DBNet, with its differentiable binarization layer, excels in isolating text contours, especially in densely packed text or low-contrast backgrounds. CRAFT, on the other hand, demonstrates strong performance on multilingual datasets due to its character-level awareness, enabling it to handle curved, thin, or irregularly oriented text. Both models outperform EAST, highlighting the importance of leveraging advanced detection architectures in video text extraction. For recognition, CRNN provides an optimal balance between speed and accuracy, while the transformer-based model achieves slightly higher accuracy due to its self-attention mechanisms, which capture long-range dependencies in challenging text samples. The trade-off between accuracy and computational overhead is evident in the results, where CRNN demonstrates faster processing times, making it more suitable for real-time applications.

Preprocessing techniques such as grayscale conversion, adaptive binarization, and noise reduction further enhance the pipeline's performance by stabilizing input frames and highlighting textual features. The ablation study confirms that each preprocessing step contributes to improved detection and recognition, particularly in noisy or low-light scenarios. By refining the input frames, these techniques reduce false positives and improve text clarity, which is crucial for downstream OCR tasks. Additionally, the post-processing module consolidates text from consecutive frames, corrects common OCR errors, and formats the final output, ensuring temporal consistency and minimizing redundant information.

Despite these strengths, the study also highlights certain limitations and areas for future improvement. The increased computational demands of the transformer-based model, for instance, may pose challenges in resource-constrained environments. Future research could explore lightweight models optimized for edge devices or low-power hardware. Additionally, integrating contextual language models for dynamic error correction and investigating joint end-to-end training of detection and recognition modules may further enhance performance. Advanced techniques such as image super-resolution, deblurring, or contrastive learning could also be incorporated to handle low-quality video frames more effectively.

Overall, the proposed pipeline represents a significant advancement in video text extraction, offering a balanced trade-off between accuracy and computational efficiency. Its adaptability to diverse text characteristics and practical utility in applications such as video indexing, real-time analytics, and multimedia retrieval highlight its potential for real-world deployment. By addressing current challenges and exploring the suggested future directions, the framework can be further refined to achieve even greater scalability, robustness, and performance.

Conclusion and Future Work:

This study presented a region-based pipeline for video text extraction, integrating advanced detection and recognition models, supported by strategic preprocessing. The results confirm the pipeline's effectiveness in balancing accuracy, efficiency, and error resilience. By leveraging cutting-edge detection (CRAFT) and recognition (CRNN) models, along with a well-designed preprocessing pipeline, the framework outperforms traditional methods in overall performance. These findings highlight its potential for practical applications, such as video indexing, real-time analytics, and semantic retrieval. The scalability of the pipeline is evident in its adaptability to different datasets and text variations. Although the reference system offers slightly higher accuracy, its high resource demands make it less practical for many real-world applications. In contrast, the proposed pipeline provides an optimal balance between performance and efficiency, making it more suitable for broader usage.

Future research directions include exploring joint end-to-end training of detection and recognition models to improve integration, applying language modeling for contextual error correction, and optimizing the pipeline for real-time performance in high-resolution videos. The framework's adaptability across multiple benchmarks further supports its potential for deployment in real-world video analytics.

References:

[1] N. A. B. Z. Mehmood, M. Iqbal, M. Ali, Z. Iqbal, "A Systematic Mapping Study on OCR Techniques," International Journal of Computer Science and Network Solutions. Accessed: Mar. 17, 2025. [Online]. Available: https://www.researchgate.net/publication/260797344_A_Systematic_Mapping_Study_on_OCR_Techniques

[2] A. Ehsan et al., "Enhanced Anomaly Detection in Ethereum: Unveiling and Classifying Threats With Machine Learning," *IEEE Access*, vol. 12, pp. 176440–176456, 2024, doi: 10.1109/ACCESS.2024.3504300.

[3] Z. Iqbal, W. Shahzad, and M. Faiza, "A diverse clustering particle swarm optimizer for dynamic environment: To locate and track multiple optima," *Proc. 2015 10th IEEE Conf. Ind. Electron. Appl. ICIEA 2015*, pp. 1755–1760, Nov. 2015, doi: 10.1109/ICIEA.2015.7334395.

[4] H. Y. C. Z. Iqbal, "Concepts, Key Challenges and Open Problems of Federated Learning," *Int. J. Eng.*, vol. 34, no. 7, pp. 1667–1683, 2021, doi: 10.5829/ije.2021.34.07a.11.

[5] N. A. Zahid Iqbal , Rafia Ilyas, Huah Yong Chan, "Effective Solution of University Course Timetabling using Particle Swarm Optimizer based Hyper Heuristic approach," *Baghdad Sci. J.*, vol. 18, no. 4, 2021, [Online]. Available: https://bsj.researchcommons.org/home/vol18/iss4/50/

[6] Z. Iqbal, R. Ilyas, W. Shahzad, and I. Inayat, "A comparative study of machine learning techniques used in non-clinical systems for continuous healthcare of independent livings," *ISCAIE 2018 - 2018 IEEE Symp. Comput. Appl. Ind. Electron.*, pp. 406–411, Jul. 2018, doi: 10.1109/ISCAIE.2018.8405507.

[7] R. Ilyas and Z. Iqbal, "Study of hybrid approaches used for university course timetable problem (UCTP)," *Proc. 2015 10th IEEE Conf. Ind. Electron. Appl. ICIEA 2015*, pp. 696–701, Nov. 2015, doi: 10.1109/ICIEA.2015.7334198.

[8] G. A. Gauvain, Jean-Luc, Lamel, Lori, "The LIMSI Broadcast News transcription system," *Speech Commun.*, vol. 37, no. 1–2, p. Speech Commun., 2002, doi: https://doi.org/10.1016/S0167-6393(01)00061-9.

[9] S. K. P. Pal, Nikhil R, "A review on image segmentation techniques," *Pattern Recognit.*, vol. 26, no. 9, pp. 1277–1294, 1993, doi: https://doi.org/10.1016/0031-3203(93)90135-J.

[10] A. L. D. Xu, S.-F. Chang, J. Meng, "Event-based highlight extraction from consumer videos using multimodal contextual analysis," *IEEE Trans. Multimed.*, vol. 13, no. 5, pp. 1004–1015, 2011.

[11] Y. Baek, B. Lee, D. Han, S. Yun, and H. Lee, "Character region awareness for text detection," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2019-June, pp. 9357–9366, Jun. 2019, doi: 10.1109/CVPR.2019.00959.

[12] X. B. Minghui Liao, Zhaoyi Wan, Cong Yao, Kai Chen, "Real-Time Scene Text Detection with Differentiable Binarization," *Proc. AAAI Conf. Artif. Intell.*, vol. 34, no. 7, 2020, doi: https://doi.org/10.1609/aaai.v34i07.6812.

[13] B. Shi, X. Bai, and C. Yao, "An End-to-End Trainable Neural Network for Image-Based Sequence Recognition and Its Application to Scene Text Recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 11, pp. 2298–2304, Nov. 2017, doi:

10.1109/TPAMI.2016.2646371.
[14] D. K. P. Litman, R. Guerrero, T. Veit, M. Rusiñol, "Scatter: Selective character attention for scene text recognition," *Proc. ICCV*, 2019.
[15] M. Anthimopoulos, B. Gatos, and I. Pratikakis, "A hybrid system for text detection in video frames," *DAS 2008 - Proc. 8th IAPR Int. Work. Doc. Anal. Syst.*, pp. 286–292, 2008, doi: 10.1109/DAS.2008.72.
[16] X. Zhou *et al.*, "EAST: An efficient and accurate scene text detector," *Proc. - 30th IEEE Conf. Comput. Vis. Pattern Recognition, CVPR 2017*, vol. 2017-January, pp. 2642–2651, Nov. 2017, doi: 10.1109/CVPR.2017.283.
[17] R. Smith, "An overview of the tesseract OCR engine," *Proc. Int. Conf. Doc. Anal. Recognition, ICDAR*, vol. 2, pp. 629–633, 2007, doi: 10.1109/ICDAR.2007.4376991.
[18] Rowel Atienza, "Vision Transformer for Fast and Efficient Scene Text Recognition," *arXiv:2105.08582*, 2021, doi: https://doi.org/10.48550/arXiv.2105.08582.
[19] Y. Liu, H. Chen, C. Shen, T. He, L. Jin, and L. Wang, "ABCNet: Real-Time Scene Text Spotting with Adaptive Bezier-Curve Network," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 9806–9815, 2020, doi: 10.1109/CVPR42600.2020.00983.
[20] Z. Qiao, Y. Zhou, D. Yang, Y. Zhou, and W. Wang, "Seed: Semantics enhanced encoder-decoder framework for scene text recognition," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 13525–13534, 2020, doi: 10.1109/CVPR42600.2020.01354.
[21] D. Karatzas *et al.*, "ICDAR 2015 competition on Robust Reading," *Proc. Int. Conf. Doc. Anal. Recognition, ICDAR*, vol. 2015-November, pp. 1156–1160, Nov. 2015, doi: 10.1109/ICDAR.2015.7333942.
[22] N. Nayef *et al.*, "ICDAR2017 Robust Reading Challenge on Multi-Lingual Scene Text Detection and Script Identification - RRC-MLT," *Proc. Int. Conf. Doc. Anal. Recognition, ICDAR*, vol. 1, pp. 1454–1459, Jul. 2017, doi: 10.1109/ICDAR.2017.237.
[23] S. B. Andreas Veit, Tomas Matera, Lukas Neumann, Jiri Matas, "COCO-Text: Dataset and Benchmark for Text Detection and Recognition in Natural Images," *arXiv:1601.07140*, 2016, doi: https://doi.org/10.48550/arXiv.1601.07140.

Stereo Vision Based Navigation of Four-Legged Robot Through Unknown Terrain

Hamza Masud[1], Arslan Khalid[1], Tayybah Kiren[2], Syed Muhammad Wasif[1*], Zubair Mehmood[1], Muhammad Jehanzeb Irshad[1], Muhammad Waqas Jabbar[1], Nazam Siddique[1]
[1]Intelligent Systems Laboratory, Department of Electrical Engineering, University of Gujrat, Pakistan.
[2]Department of Computer Science (RCET), University of Engineering and Technology Lahore, Pakistan
[*]Correspondence: syed.wasif@uog.edu.pk

Citation | Masud. H, Khalid. A, Wasif. S. M, Mehmood. Z, Irshad. M. J, Jabbar. M. W, Siddique. N, "Stereo Vision Based Navigation of Four-Legged Robot Through Unknown Terrain", IJIST, Special Issue. pp 46-56, March 2025
Received| Feb 11, 2025 **Revised|** Feb 24, 2025 **Accepted|** March 01, 2025 **Published|** March 04, 2025.

This research aims to develop a stereo vision-based navigation system for a quadruped robot, enabling it to move autonomously through rough, unfamiliar terrain and detect blockages in sewer pipelines. The robot uses a stereo camera to capture images, which are then processed to create disparity maps and 3D point clouds. These tools help the robot identify and avoid obstacles. Image rectification and 3D mapping are performed using OpenCV, which generates an occupancy grid to distinguish between free and occupied spaces. Based on this grid, the A* algorithm is used to plan the robot's path. To ensure smooth movement, inverse kinematics calculates the required motor angles and applies predefined Bezier curves for stable locomotion.

Keywords: Quadruped Robot; Disparity Map; Stereo Vision; Depth Map; 3D Point Cloud; Inverse Kinematics; Bezier Curve.

Introduction:

The sewerage pipeline system plays a crucial role in collecting and transporting wastewater from residential, commercial, and industrial areas to treatment plants. Since these pipelines are usually located underground, regular inspection and maintenance become challenging. Traditionally, inspections are carried out manually by human operators who enter the pipelines or by using cameras mounted on cables. However, these conventional methods are time-consuming and pose serious health risks to workers. In addition to these challenges, human operators often encounter various obstacles, navigate bends, and endure the harsh conditions inside the pipelines.

Advancements in robotics over the years have led to the development of autonomous robots for pipeline inspection. These robots can collect diverse types of data, including videos, images, and sensor readings, providing a comprehensive analysis of the pipeline's condition.

With the increasing demand for infrastructure maintenance and pipeline inspection, robotics has emerged as a promising solution to address the challenges of limited accessibility, hazardous environments, and labor-intensive inspection processes. Sewerage pipelines, in particular, pose significant challenges due to narrow spaces, bends, and potential blockages that make traditional inspection methods inefficient and unsafe for human operators. While wheeled robots and sensor-based systems have been developed for pipeline inspection, these methods often struggle with navigating uneven terrains, negotiating tight curves, and accurately detecting obstructions. This has led to the growing interest in legged robots, which offer enhanced mobility, adaptability, and stability in rough and unstructured environments. Legged robots can overcome obstacles, traverse dynamic terrains, and access confined spaces that would otherwise be inaccessible to wheeled or tracked robots.

Among various sensing technologies, stereo vision has gained attention due to its ability to provide detailed depth perception and 3D mapping, enabling robots to better understand their surroundings. Unlike ultrasound sensors or monocular cameras, stereo vision captures disparity maps and 3D point clouds, which improve the robot's ability to detect obstacles, calculate distances, and plan optimal paths. Integrating stereo vision with advanced path planning algorithms, such as the A* algorithm, allows for real-time obstacle avoidance and efficient navigation in dynamic environments. Additionally, the use of inverse kinematics for gait control enhances the robot's movement precision, ensuring smooth locomotion even on uneven terrain. This study aims to leverage the strengths of stereo vision, path planning, and inverse kinematics to develop a fully autonomous quadruped robot capable of navigating unknown terrains and inspecting sewerage pipelines with enhanced efficiency, accuracy, and stability.

The main goal of this study is to develop a fully autonomous quadruped robot capable of walking independently, avoiding obstacles, and efficiently reaching its destination. This requires creating real-time algorithms for obstacle detection, path planning, and locomotion. Unlike traditional methods, such as ultrasound sensors [1], which often produce inconsistent results due to weak signal reflection, our approach enhances navigation reliability. While Deep Q-Network (DQN)-based navigation [2] is effective, it has limitations, including high sensitivity to environmental changes and the need for extensive datasets and hyperparameter tuning.

To overcome these challenges, we incorporated stereo vision-based navigation, which enhances adaptability and enables the robot to navigate in dynamic environments. The key objectives of this study include designing an autonomous quadruped robot model, developing a stereo vision algorithm for obstacle avoidance and path planning, and improving the robot's locomotion. Additionally, this research represents a significant advancement in autonomous robot navigation, as discussed in [2], [1].

Objectives of the Study:

The main objectives of this study are:

- To develop a fully autonomous quadruped robot capable of navigating unknown terrains and inspecting sewerage pipelines.
- To implement a stereo vision-based navigation system for real-time obstacle detection, depth mapping, and 3D point cloud generation.
- To design and optimize path planning algorithms, specifically the A* algorithm, for efficient and collision-free navigation.

Related Work:

Model-based predictive controllers (MBPCs) using Neural Networks and ultrasonic sensors create mathematical models that perform effectively in static environments [3]. Dynamic Artificial Neural Networks (DANNs) are employed for motion planning and robot pathfinding but are primarily suited for flat surfaces with static and dynamic obstacles. The model's efficiency increases in dynamic environments by relying on past behavior and sensor inputs [4]. A Sprint-bot prototype, capable of smooth movement and turns, was developed for pipeline navigation. Initially designed for dry pipes, it requires precise sensor node positioning to detect leakage. A novel SLAM algorithm was implemented to create 3D pipeline maps from 2D data, enhancing performance with geographical pipeline information. Graph optimization techniques improve the robot's localization [5]. Building on previous work involving reactive controllers with balancing control, this research focuses on dynamic locomotion using active impedance and IMU feedback [6]. The IMU provides essential data, such as acceleration, deceleration, and tilting, helping the robot counter external forces. Using camera images, the robot can walk toward targets, with an advanced CAM shift algorithm enabling target tracking through a color probability map. This process generates a map and pinpoints the target's location [6].

A defect detection system for pressure pipelines uses the Phased Array Ultrasonic Technique (PAUT) to identify cracks and corrosion. This wheeled robot features a camera to capture pipeline interiors and an ultrasonic phased array system for hidden defect detection [7]. A Vibro-impact capsule robot, designed for the oil industry, moves through pipelines using rectilinear motion, independently navigating harsh conditions without an external driving mechanism [7]. Another robot was developed for navigating vertical, curved, and inclined pipelines, incorporating kinematic and dynamic analysis to optimize its trajectory and motion [8]. MAKRO, a pipe-inspection robot, can operate inside 30 cm diameter vertical pipelines. It captures images, live streams video, detects cracks, and determines their exact location using online image processing [9]. A multi-link articulated robot with omni- and hemi-spherical wheels was also designed to adapt to winding pipes, operating in both horizontal and vertical pipelines and transmitting data via a wireless camera [10]. Research on a self-propelled capsule system optimized its design for stability and reliability under extreme conditions [11]. Another semi-automatic robotic system, equipped with a CCD camera, steering mechanism, and sensors, was proposed to monitor pipelines. Real-time pipeline data is gathered through image processing, and gyro sensors and encoders generate pipeline maps and localize the robot's position [12].

Simultaneous Localization and Mapping (SLAM) is a key technique for robotic navigation and mapping. A modified CAO algorithm was introduced to resolve local minima issues, helping the robot escape deadlocks [13]. To improve navigation, IMU-based sensors were combined with stereo vision. Since inertial sensors accumulate errors over time, stereo vision corrects these inaccuracies, reducing long-term navigation errors. Two coupling methods—MSF and MSCKF—were tested, with MSCKF showing higher efficiency [14]. Deep reinforcement learning (RL) allows robots to learn tasks independently but faces challenges such as reward-setting, sensor inaccuracies, and unpredictable behavior. This study proposed guided constrained policy optimization (GCPO), which improves RL by enforcing specific rules during training. RL models typically require large datasets and are trained using physics simulators. This

method was tested on a quadruped robot, leading to faster learning and enhanced performance without precise reward adjustments [15].

A deep learning model (LSTM-DL) for pipeline defect detection achieved 98.31% accuracy in identifying blockages and leaks by analyzing pressure and flow sensor data [16]. SQuRo, a compact legged robot, was designed for confined spaces. With a slim body (aspect ratio of 3.42), it offers flexibility, superior mobility, and the ability to carry small loads up to 200 grams [17]. The self-propelled capsule system provides a novel pipeline inspection approach by detecting abnormalities using onboard sensors [18]. Another robot with a single tracked drive and rotational capability about a perpendicular axis enhances navigation in narrow, winding pipelines [19]. Real-time image processing improves defect detection, while offline processing supports deeper analysis and maintenance planning, ultimately enhancing pipeline inspection and safety [20]. The Tarantula robot, equipped with cameras and sensors, monitors drainage systems for blockages and defects, minimizing human exposure to hazardous environments [21]. Navigating sewer-bots in harsh conditions remains challenging, requiring precise control to handle tight spaces and obstacles [22]. Cameras provide internal views, ultrasonic sensors detect cracks and leaks, and magnetic sensors identify metallic objects left behind during construction or maintenance [23]. These autonomous robots transmit real-time data via routers acting as alternative servers, streamlining pipeline monitoring [24].

Methodology:

This section discusses the key operations and algorithms involved in the navigation of a quadruped robot. The research methodology is outlined in Figure 1, which highlights the essential components and methods used in developing autonomous quadruped robot operations.

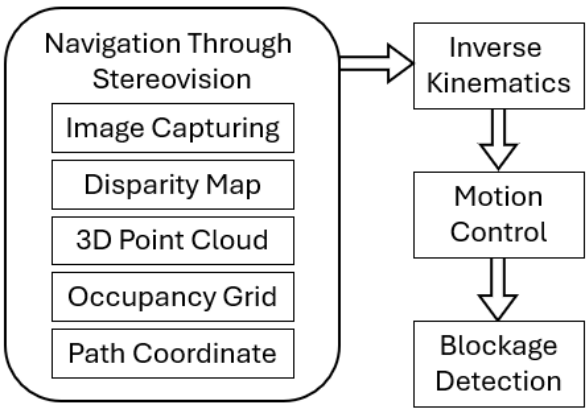


Figure 1. Operational block diagram of autonomous quadruped robot

A. Image Capturing

The first step in navigation involves capturing images with a stereo vision camera module. This camera, mounted on a Raspberry Pi 4, captures images that are essential for detecting obstacles, planning paths, and identifying blockages.

B. Image Rectification

Image rectification is a key step in 3D reconstruction that reduces geometric distortions and restores images to their proper alignment. This step is crucial for creating a 3D model used in navigation. The process involves comparing two images taken from slightly different angles and matching corresponding pixels to align them correctly. This alignment is achieved using pre-calculated camera calibration parameters.

C. Disparity Map

A disparity map is a two-dimensional image that shows the difference in pixel intensity between the left and right stereo images. In this map, higher values indicate objects closer to the

camera, while lower values indicate objects farther away. The two images are compared pixel by pixel in their respective positions to generate the disparity map, which shows the horizontal shift of pixels. From these shifts, information about the distance of objects from the camera is obtained. To improve accuracy, the Semi-Global Block Matching (SGBM) algorithm is used to calculate disparities by analyzing pixel intensity differences in multiple directions, striking a balance between accuracy and efficiency. Additionally, the Weighted Least Squares (WLS) filter is applied to reduce noise and remove speckles from the disparity map, further refining the result.

D. 3D Point Cloud

Once the disparity map is computed, a 3D point cloud is generated using the OpenCV library. This point cloud stores image data in x, y, and z coordinates, providing detailed information about the surroundings. It allows the system to detect obstacles and understand the environment for autonomous navigation. The projection matrix (Q), obtained during image rectification, converts disparity values into 3D coordinates. Any points with y-coordinate values exceeding a set threshold are identified as obstacles, as they are considered part of the floor. The reprojection equation is used to achieve this transformation.

E. Path Planner

For path planning, the A* algorithm is employed to determine the shortest path between two points while avoiding obstacles. The occupancy grid, generated from stereo vision data, helps map out the path from the current position to the destination by identifying clear routes.

F. Inverse Kinematics

Inverse kinematics is used to precisely control the movement of quadruped robots by calculating the joint angles needed to achieve a specific end-effector position and orientation. To simplify these calculations, the coordinates are transformed, focusing on the z-direction.

$$y = -\sqrt{(z+L)^2 + y^2} \quad (1)$$

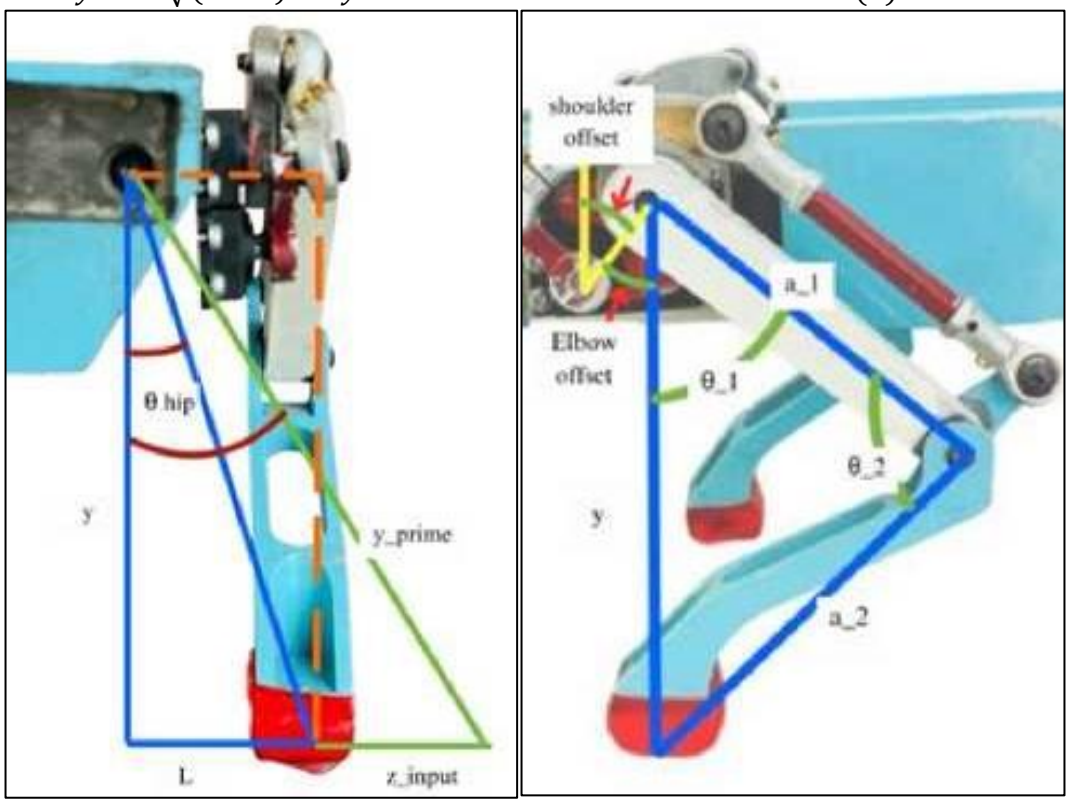


Figure 2. Robotic Leg (a) Y-Z side view (b) X-Y side view

Calculated y coordinate is used to calculate the angle theta z for the hip joint:

$$\theta_z = tan^{-1}\left(\frac{z+L}{|y|}\right) - tan^{-1}\left(\frac{L}{|y|}\right) \quad (2)$$

Segments a, and a_2 represent the upper and lower leg, respectively. The angles at the shoulder (θ_1) and elbow (θ_1) are computed using the laws of sines and cosines. To find the cosine of an angle, the law of cosines is applied. This formula is derived from the triangle formed by the upper leg, lower leg, and the line connecting the shoulder and foot.

$$\cos(\theta_2) = \frac{x^2 + y^2 - a_1^2 - a_2^2}{2a_1a_2} \quad (3)$$

$$Sin(\theta_2) = \sqrt{1 - cos^2(\theta_2)} \quad (4)$$

Calculating the inverse tangent of sin over cosine to get the elbow angle.

$$\theta_2 = tan^{-1}\left(\frac{\sin(\theta_2)}{\cos(\theta_2)}\right) \quad (5)$$

Similarly, calculate the sine and cosine of $\theta 1$ using the law of cosines. Then, compute the angle $\theta 1$ as:

$$\cos(\theta_1) = \frac{x\,(a_1 + a_2\cos(\theta_2)) + y\,(a_2\sin(\theta_2))}{x^2 + y^2} \quad (6)$$

$$\sin(\theta_1) = \frac{y(a_1 + a_2C_2) - x(a_2s_s)}{x^2 + y^2} \quad (7)$$

$$\theta_1 = tan^{-1}\left(\frac{\sin(\theta_2)}{\cos(\theta_2)}\right) \quad (8)$$

Results:

Image Rectification:

The camera calibration parameters are used to rectify the stereo images. These parameters include camera metrics, distortion matrices, rotation, and projection metrics, which help align the left and right images correctly. The rectified images are then compared, as shown in Figure 3. After rectification, OpenCV converts the images to grayscale, a necessary step for generating the disparity map.

Figure 3. Left and right images before and after rectification

Disparity Map:

The grayscale images are then processed using the Semi-Global Block Matching (SGBM) algorithm to compute the disparity map. In this map, each pixel represents the disparity between corresponding points in the stereo images. The resulting disparity map is shown in Figure 4.

Figure 4. Disparity map without WLS filter

To smooth the disparity map and reduce noise, the Weighted Least Squares (WLS) filter is applied to enhance depth information. The filter's parameters are adjusted to control regularization and color influence. The disparity map after applying the WLS filter (7) is shown in Figure 5.

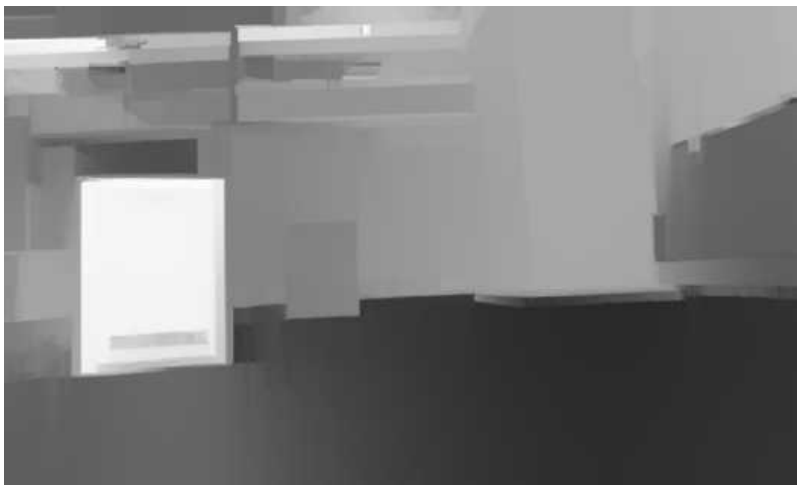


Figure 5. Disparity map after WLS filter

Point Cloud:

The projection matrix (Q), derived from intrinsic and extrinsic parameters, is used to convert disparity values into 3D coordinates. The resulting point cloud contains data on width, height, and the three spatial coordinates: x, y, and z. The point cloud generated from the disparity map is shown in Figure 6 below.

Figure 6. 3D point cloud image of environment

Occupancy grid:

The algorithm then iterates through the generated 3D point cloud and updates the occupancy grid by marking free space with a value of 0 and occupied space with a value of 1. It

subsequently checks the minimum Z-value for each grid cell to identify the nearest and farthest obstacles within that cell. The calculated occupancy grid is shown in Figure 7.

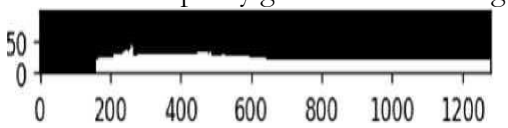


Figure 7. Occupancy grid of environment in 2D

Path Planning:

The starting point is set at the center of the current location, while the endpoint corresponds to the farthest obstacle recorded in the occupancy grid. This setup allows the A* algorithm to determine the shortest obstacle-free path to the endpoint. Figure 8 illustrates the planned path between the farthest obstacle and the current location.

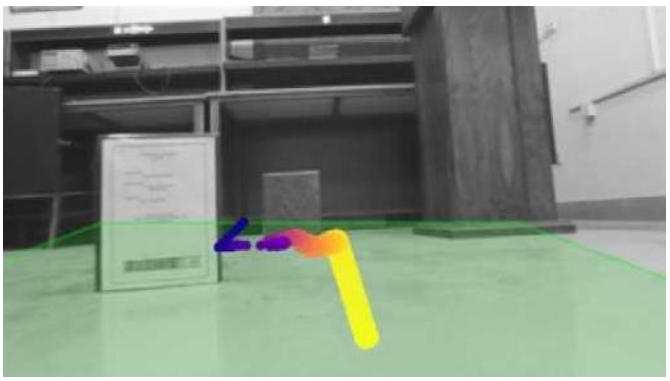


Figure 8. Planned Path

Gait Pattern:

Once the planned path is obtained, the robot performs locomotion using inverse kinematics. Cubic Bézier curve control points are defined for both linear and cubic paths to ensure smooth leg movement. These points are then used in inverse kinematics calculations to determine the motor angles. The defined Bézier curve points are as follows:$$cubic\ Bexier\ Curve = \begin{bmatrix} -1.0 & -1.0 & -1.0 & -1.0 \\ -1.0 & -1.0 & -1.0 & -1.0 \\ -15 & -10 & -10 & -15 \end{bmatrix}$$

Discussion:

This study focuses on the development of a fully autonomous quadruped robot designed for obstacle avoidance, path planning, and sewerage pipeline inspection. Traditional sewerage inspection robots typically use wheeled designs and various sensor-based techniques for navigation and environmental perception. While technologies such as ultrasound sensors, convolutional neural networks (CNNs), region-based convolutional networks (RCNNs), and monocular cameras have shown success in object detection, each has limitations that affect robot performance in dynamic environments like pipelines. A key innovation in our approach is the use of stereovision cameras to calculate depth maps and generate 3D point clouds. Unlike ultrasound sensors, which rely on sound waves and may be influenced by environmental factors, stereovision cameras provide a more robust alternative by using disparity mapping to calculate depth. This enhances the robot's ability to interpret complex environments, which is particularly beneficial in sewerage systems where visibility and accurate object recognition are critical.

While CNN and RCNN-based methods achieve high object detection accuracy, they require extensive training datasets and significant computational power. In contrast, our stereovision-based approach is less computationally intensive and more power-efficient. Deep learning models can also face challenges like overfitting to specific environments, whereas our stereovision technique, combined with real-time path planning algorithms, improves adaptability without relying on pre-trained models. Overall, this study contributes to the field of autonomous

robotic navigation by presenting a novel application of stereovision-based obstacle detection and path planning. The findings demonstrate the potential of stereovision as an alternative to LiDAR and traditional sensor-based systems, offering a balance between computational efficiency and reliable environmental perception. This study demonstrates the significant advantages of integrating stereo vision-based navigation, path planning, and inverse kinematics for enhancing the mobility and adaptability of quadruped robots in challenging environments such as sewerage pipelines. Unlike traditional wheeled or tracked robots, which often encounter limitations in maneuvering through uneven terrains and navigating sharp bends, the proposed quadruped robot leverages real-time depth mapping and 3D point cloud generation to gain a comprehensive understanding of its surroundings. This enhanced environmental perception, combined with the A* path planning algorithm, enables the robot to efficiently detect obstacles, calculate optimal routes, and avoid collisions. Furthermore, the implementation of inverse kinematics and Bezier curves for gait control ensures smooth and precise locomotion, improving the robot's stability during traversal. These innovations contribute to a more robust and scalable framework for autonomous navigation, with potential applications in industrial pipeline inspection, hazardous environment exploration, and disaster response scenarios. Future work could focus on optimizing computational efficiency, integrating additional sensors for multi-modal perception, and enhancing the robot's performance in real-time dynamic environments.

Conclusion:

In this paper, the autonomous navigation algorithm for the robot involved several key processes, including stereo image rectification, disparity map calculation, 3D point cloud generation, and occupancy grid creation. The A* path planning algorithm was used to plan and generate the navigation path. Based on the planned path coordinates, inverse kinematics was applied to calculate the joint angles, enabling the robot's legs to respond and follow the predicted path. To achieve smooth trajectory control, a cubic Bezier curve was generated for forward and backward steps, while a linear Bezier curve was used for sliding (left and right) movements. These features enhanced the robot's ability to navigate through rough and unfamiliar terrain. The robot demonstrated excellent performance on planned navigation paths and maintained smooth trajectories, as shown in the results.

The proposed stereovision-based navigation approach significantly improved the autonomous navigation capabilities of quadruped robots. By integrating stereovision cameras for disparity mapping, obstacle avoidance, and path planning with the robot's inverse kinematics for gait control, the robot could detect blockages inside sewerage lines and identify objects obstructing sewage flow. This integration enhanced the robot's ability to walk efficiently and stably, increasing both accuracy and efficiency when navigating rough terrain.

Acknowledgement. The authors would like to thank the administration of Intelligent Systems Laboratory, Department of Electrical Engineering, University of Gujrat for providing an opportunity to conduct research and perform experiments with the available laboratory facilities.

Author's Contribution. Hamza Masud, Arslan Khalid wrote the initial draft of the article with equal participation. Syed Muhammad Wasif supervised and drafted the final version of the article. Zubair Mehmood improved the formatting of the final draft and proof read and corrected the manuscript with the help of Muhammad Jehanzeb Irshad, Muhammad Waqas Jabbar and Nazam Siddique.

Conflict of interest. The authors have no conflict of interest for publishing this manuscript in IJIST.

References:

[1] H. T. & J. P. Jawad N. Yasin, Sherif A. S. Mohamed, Mohammad-Hashem Haghbayan, Jukka Heikkonen, "Low-cost ultrasonic based object detection and collision avoidance method for autonomous robots," *Int. J. Inf. Technol.*, vol. 13, pp. 97–107, 2021, doi: https://doi.org/10.1007/s41870-020-00513-w.

[2] M. . Escobar-Naranjo, J.; Caiza, G.; Ayala, P.; Jordan, E.; Garcia, C.A.; Garcia, "Autonomous Navigation of Robots: Optimization with DQN," *Appl. Sci.*, vol. 13, no. 12, p. 7202, 2023, doi: https://doi.org/10.3390/app13127202.

[3] E. F. C. J. Gomez-Ortega, "Neural network MBPC for mobile robot path tracking," *Robot. Comput. Integr. Manuf.*, vol. 11, no. 4, pp. 271–278, 1994, doi: https://doi.org/10.1016/0736-5845(95)00003-8.

[4] D. L. and D. Y. J. Yuan, H. Wang, C. Lin, "A Novel GRU-RNN Network Model for Dynamic Path Planning of Mobile Robot," *IEEE Access*, vol. 7, pp. 15140–15151, 2019, doi: 10.1109/ACCESS.2019.2894626.

[5] K. K. Leder, Christoph, Rastogi Tushar, "Putting benign by design into practice-novel concepts for green and sustainable pharmacy: Designing green drug derivatives by non-targeted synthesis and screening for biodegradability," *Sustain. Chem. Pharm.*, vol. 2, pp. 31–36, 2015, doi: https://doi.org/10.1016/j.scp.2015.07.001.

[6] V. Barasuol, J. Buchli, C. Semini, M. Frigerio, E. R. De Pieri, and D. G. Caldwell, "A reactive controller framework for quadrupedal locomotion on challenging terrain," *Proc. - IEEE Int. Conf. Robot. Autom.*, pp. 2554–2561, 2013, doi: 10.1109/ICRA.2013.6630926.

[7] P. Ciorau, C. Chartier, and K. Mair, "A CONTRIBUTION OF PHASED ARRAY ULTRASONIC TECHNOLOGY (PAUT) TO DETECTION AND SIZING STRESS CORROSION CRACKS," *e-Journal Nondestruct. Test.*, 2008.

[8] M. W. Yang Liu, Sheikh Islam, Ekaterina Pavlovskaia, "Optimization of the Vibro-Impact Capsule System," *J. Mech. Eng.*, vol. 62, pp. 430–439, 2016, doi: 10.5545/sv-jme.2016.3754.

[9] J. Doshi, "An investigation of leaky sewers as a source of fecal contamination in the stormwater drainage system in Singapore," *Ph.D. Diss. Massachusetts Inst. Technol. Cambridge, MA, USA*, 2012.

[10] H. J. Zhong, Z. W. Ling, C. J. Miao, W. C. Guo, and P. Tang, "A new robot-based system for in-pipe ultrasonic inspection of pressure pipelines," *Proc. 2017 IEEE Far East NDT New Technol. Appl. Forum, FENDT 2017*, pp. 246–250, Dec. 2018, doi: 10.1109/FENDT.2017.8584579.

[11] D. Jayakumar, R. Jaganath, and R. Selvarasu, "Defect identification in pipelines using inspection robot," *Proc. 2016 Online Int. Conf. Green Eng. Technol. IC-GET 2016*, May 2017, doi: 10.1109/GET.2016.7916773.

[12] A. H. S. H. PKM163032, "ANALYSIS AND CONTROL OF MOBILE ROBOT FOR PIPE LINE INSPECTION." Academia, 2013. Accessed: Mar. 18, 2025. [Online]. Available: https://www.academia.edu/38122477/ANALYSIS_AND_CONTROL_OF_MOBILE_ROBOT_FOR_PIPE_LINE_INSPECTION

[13] G. S. Palwinder Kaur, Ravinder Kaur, "PIPELINE INSPECTION AND BOREWELL RESCUE ROBOT," *Int. J. Res. Eng. Technol.*, vol. 3, no. 4, pp. 726–729, 2014, [Online]. Available: https://ijret.org/volumes/2014v03/i04/IJRET20140304128.pdf

[14] S. K. P. Ankit Nayak, "Design of a New In-Pipe Inspection Robot," *Procedia Eng.*, vol. 97, pp. 2081–2091, 2014, doi: https://doi.org/10.1016/j.proeng.2014.12.451.

[15] P. E. A. Godwin Itseoritseagba, Donald O Ene, Ijibike Udo, "Pipeline Inspection for Corrosion using a Mobile Robotic System," *Int. J. Robot. Eng.*, vol. 1, no. 1, 2015, doi: 10.35840/2631-5106/4101.

[16] S. S. O. Atul A. Gargade, "Development of In-pipe Inspection Robot," *IOSR J. Mech. Civ. Eng.*, vol. 13, no. 4, pp. 64–72, 2016, doi: 10.9790/1684-1304076472.

[17] A. Yusupov and Y. Liu, "Development of a self-propelled capsule robot for pipeline

inspection," *2016 22nd Int. Conf. Autom. Comput. ICAC 2016 Tackling New Challenges Autom. Comput.*, pp. 84–88, Oct. 2016, doi: 10.1109/ICONAC.2016.7604899.

[18] A. Kakogawa and S. Ma, "Design of a multilink-articulated wheeled pipeline inspection robot using only passive elastic joints," *Adv. Robot.*, vol. 32, no. 1, pp. 37–50, Oct. 2018, doi: 10.1080/01691864.2017.1393348.

[19] F. Z. & A. Y. Yao Yan, Yang Liu, Joseph Páez Chávez, "Proof-of-concept prototype development of the self-propelled capsule system for pipeline inspection," *Meccanica*, vol. 53, pp. 1997–2012, 2018, doi: https://doi.org/10.1007/s11012-017-0801-3.

[20] C. B. R. Cichosz, J. White, T. Price, "Pipeline inspection robot," *U.S. Pat. Appl.*, 2018, [Online]. Available: https://www.instructables.com/Pipeline-Inspection-Robot/

[21] A. K. Mohamed Abdellatif, Hazem Mohamed, Mohamed Hesham, Ahmed Abdelmoneim, Amro Kamal, "Mechatronics Design of an Autonomous Pipe-Inspection Robot," *MATEC Web Conf*, vol. 153, 2018, doi: https://doi.org/10.1051/matecconf/201815302002.

[22] M. S. T. Abdullah Aamir Hayat, Karthikeyan Elangovan, Mohan Rajesh Elara, "Tarantula: Design, Modeling, and Kinematic Identification of a Quadruped Wheeled Robot," *Appl. Sci.*, vol. 9, no. 1, p. 94, 2019, doi: https://doi.org/10.3390/app9010094.

[23] U. Hafila, Y. Namdeo, and M. Mani, "Design of a Portable, Automated Sewer-Line Inspection and De-Clogging Sewer-Bot for Indian Conditions," *Smart Innov. Syst. Technol.*, vol. 134, pp. 703–713, 2019, doi: 10.1007/978-981-13-5974-3_61.

[24] M. Roussialian, H. Al Zanbarakji, A. Khawand, A. Rahal, and M. Owayjan, "Design and Development of a Pipeline Inspection Robot," *Mech. Mach. Sci.*, vol. 58, pp. 43–52, 2019, doi: 10.1007/978-3-319-89911-4_4.

Photocatalytic Degradation of Deltamethrin in Drinking Water Under Visible Light by Using Zno and Tio_2

Mahmboob Ahmad[1], Muhammad Tahseen Sadiq[1], Kamal Zafar[1], Ghulam Abbas[1*], Saher Asif[1], Nabila Kousar[1], Abdul Wahab Mukhtar[1], Arslan Chaudhary [2]
[1]Department of Chemical Engineering and University of Gujrat,Pakistan
[2]Sui Northern Gas Pipelines Limited, Lahore, Pakistan
*Correspondence: engr.ghulamabbas@uog.edu.pk

Citation | Ahmad. M, Sadiq. M. T, Zafar. K, Abbas. G, Asif. S, Kousar. N, Mukhtar. A. W, Chaudhary. A, "Photocatalytic Degradation of Deltamethrin in Drinking Water Under Visible Light by Using Zno and Tio_2", IJIST, Special Issue. pp 57- 67, March 2025
Received| Feb 12, 2025 **Revised|** Feb 25, 2025 **Accepted|** March 02, 2025 **Published|** March 05, 2025.

The use of deltamethrin is increasing due to its high demand in agriculture. However, it is toxic to both surface and groundwater. Agriculture plays a crucial role in the economy of any major nation. This study aims to enhance pesticide degradation by using specially designed catalysts optimized for visible light exposure. The key innovation lies in the customized catalyst design, which improves photocatalytic efficiency while offering a cost-effective and environmentally friendly approach. Various factors affecting degradation, including adsorbent quantity, pH, contact time, and initial concentration, were analyzed. The reactor consists of a 6-watt (380 nm) visible light lamp and a stirrer to ensure uniform mixing of the sample. Photocatalysts ZnO and TiO_2, in concentrations ranging from 0.1 to 3.0 g/L, were used to generate oxidizing agents. Under visible light, the impact of these factors on the degradation of different pesticide solutions was examined. The optimal doses were found to be 1.5 g/L for ZnO and 0.1 g/L for TiO_2. ZnO achieved a degradation rate of 96.3%, while TiO_2 slightly outperformed it with a rate of 96.34%. The study also investigated the effect of pH variations on deltamethrin degradation, revealing stronger degradation in alkaline conditions. Additionally, TiO_2 effectively reduced the COD value, demonstrating its superior efficiency in pesticide breakdown.

Keyword: Deltamethrin, Photo catalytic degradation, ZnO, TiO2, Oxidizing agents.

Introduction:

A significant portion of the economy depends on agricultural production. Due to the increasing demand for agricultural products, the use of pesticides has risen significantly. Agricultural zones are designated areas where occupation, enterprise, and lifestyles are interconnected. The agricultural sector is a crucial component of any country's economy, contributing approximately 23.4% to the global economy. The primary focus of agricultural zones is to ensure food security for the population and enhance crop yield. Pakistan's agricultural sector contributes 21% to the country's GDP, with an annual growth rate of 2.7%. According to the World Bank (World Development Indicator), the agricultural sector has added approximately 22.64% to Pakistan's Gross Domestic Product. Sustainable agriculture is a rapidly growing field aimed at producing food and energy in an environmentally friendly manner to support both current and future generations. It addresses pressing issues such as climate change, rising fuel prices, hunger, poverty, increasing food demand, pest control, soil degradation, erosion, biodiversity loss, and water contamination.

Modern agriculture heavily relies on chemical treatments, including pesticides and fertilizers [1]. There is no denying that agricultural production in the 20th century improved and stabilized primarily due to the control of harmful weeds, pests, and insects, alongside an adequate supply of essential plant nutrients through chemical means. Pesticides are the second leading contributor to aquatic pollution, with a significant increase in contamination of drinking water sources [2]. Various pesticides, differing in chemical composition and effectiveness, are widely used worldwide, raising concerns about their adverse effects on human health and the environment. Pesticides consist of both organic and inorganic compounds and are applied to crops to eliminate harmful weeds and pests such as moths and insects that feed on crops. These chemicals are complex, with some parent compounds being less toxic than their breakdown products [3]. However, they pose a threat to both surface and groundwater. Pesticides are identified as the second leading cause of water pollution (WHO Class II), particularly contaminating drinking water [4].

Deltamethrin, a widely used pesticide, has been found to adversely affect fish by disrupting their nervous system, blocking sodium channels, and inhibiting key enzymes such as acetylcholinesterase and gamma-aminobutyric acid. It also weakens their immune system and induces oxidative stress. Individuals with skin wounds should avoid using Deltamethrin in traditional forms such as sprays, soaps, spot-on applications, pour-on treatments, and shampoos, as excessive absorption through the skin can occur. According to EU standards, Deltamethrin concentrations should not exceed 0.1 ng/ml [2]. Pyrethroids, including Deltamethrin, are preferred over organophosphates and organochlorines due to their high potency, effectiveness in small doses, resistance to light-induced degradation, and minimal harm to birds and mammals [5]. Several chemical techniques have been developed for environmental remediation, including hybrid procedures, nanocrystalline materials, metal oxides, carbon nanotubes, ion exchange, graphene, ultrasound waves, photocatalysis, adsorption, bioremediation, and bio-purification [6]. Advances in water and wastewater treatment have incorporated various strategies to remove persistent organic contaminants from aqueous solutions. Effective methods include adsorption, coagulation/flocculation, membrane separation, electrochemical treatments, and reverse osmosis [7][8][9]. Adsorption processes utilize diverse materials, including waste substances, to remove pollutants such as dyes [10][11], lead (II) [12], chlorophenols [13], pesticides [14], zinc, and absorbed nickel [15]. However, most of these methods only transfer pollutants between phases, generating secondary waste without actual pollutant degradation.

Advanced oxidation processes (AOPs) operate by generating hydroxyl radicals (•OH) and have proven effective in breaking down various contaminants, including dyes [12], aromatic amines [16], and agrochemicals [17]. Pesticide degradation has been successfully achieved using

the Fenton reagent and both natural and artificial photo-irradiation methods, such as photo-Fenton systems [18][19], photocatalytic degradation [20][21], and photo-peroxidation.

Membrane separation, photocatalysis, and adsorption techniques are being tested using various materials. Researchers are striving to develop efficient, cost-effective, and environmentally friendly methods that offer rapid and high pollutant removal rates. These procedures utilize visible light directed at a semiconductor material, which acts as a catalyst to break down pesticides. Photocatalysis is an environmentally friendly alternative for degrading organic pollutants. It has been widely studied for the degradation of various materials, with zinc oxide (ZnO) and titanium dioxide (TiO2) being the most commonly used photocatalysts. Keiichi et al. demonstrated the photocatalytic degradation of azo dyes using TiO2 suspensions, revealing that diazo dyes degrade more slowly than mono azo dyes (Photocatalytic Degradation of Commercial Azo Dyes). Another study compared the photocatalytic efficiency of ZnO and Degussa P25 TiO2 for degrading azo dyes under solar irradiation (Solar Photocatalytic Degradation of Azo Dye). Similarly, studies have examined the effects of UV and solar light irradiation on diclofenac degradation using ZnO as a photocatalyst, showing higher degradation rates under UV light, particularly in acidic conditions (Degradation of Diclofenac Under Irradiation of UV Lamp and Solar Light Using ZnO Photocatalyst). Additional research has investigated the photocatalytic degradation of phenol (Photocatalytic Degradation of Phenol), nitrophenols (Heterogeneous Photocatalytic Degradation of Nitro Phenols), 17-β-estradiol (Photocatalytic Degradation of 17-β-Estradiol on Immobilized TiO2), methyl orange (Photocatalytic Degradation of Methyl Orange as a Model Compound), and methylene blue (Photocatalytic Degradation Pathway of Methylene Blue in Water).

Despite these advancements, challenges remain in photocatalytic degradation research, emphasizing the need for further studies. Current photocatalytic applications are limited by low visible light absorption, rapid charge recombination, and the low migration ability of photo-generated electrons and holes (An Overview of Photocatalytic Degradation: Photocatalysts, Mechanisms, and Development of Photocatalytic Membrane). To improve wastewater treatment efficiency, more research is needed to optimize conditions for degrading a broader range of organic pollutants.

The objective of this study is to evaluate the effectiveness of different photocatalysts in degrading Deltamethrin pesticides while also analyzing parameters such as contact time, initial pH concentration, adsorbent dosage, and COD value.

Material and Methods:

Objectives of the Study:

This study aims to optimize process parameters for efficient and eco-friendly pesticide removal, contributing to sustainable water purification methods. The specific objectives are:

- To evaluate the efficiency of ZnO and TiO_2 in degrading Deltamethrin under visible light conditions.
- To determine the optimal photocatalyst dosage, pH, and reaction time for maximum pesticide degradation.
- To compare the degradation efficiency of ZnO and TiO_2 at different pH levels and catalyst concentrations.
- To analyze the kinetics of photocatalytic degradation and validate the reaction mechanism using the Langmuir-Hinshelwood model.
- To assess the impact of photocatalysis on Chemical Oxygen Demand (COD) reduction as an indicator of water quality improvement.
- To investigate the reusability and stability of ZnO and TiO_2 catalysts over multiple degradation cycles.

Methodology:

Chemical Analysis and Materials:

High-purity laboratory-grade materials were selected for the experiment. Deltamethrin (97.8% pure) was procured from Jaffar Group of Companies, Lahore. ZnO nanoparticles, with a purity of over 93.5% and particle sizes ranging from 10 to 30 nm, were obtained from Global Chemical Co. Ltd. TiO_2 nanoparticles, with a purity exceeding 98.4% and particle sizes between 10 and 25 nm, were purchased from KRONOS (ISO 9001 certified).

A pH meter (model BASIC-20, UK) was used to measure pH levels. The photocatalytic reactor, designed for small-scale experiments, has a single inlet and outlet with an upper-side valve to prevent flooding. It is equipped with a 6V DC motor for efficient mixing and a 6W, 9-inch fluorescent tube emitting violet light (380 nm wavelength) for illumination. The reactor has a solution capacity of approximately one liter.

Acetonitrile (CHROMASOLV, 99.9% pure) was sourced from Honeywell Riedelde Haen for HPLC analysis. Gas chromatography vials (2 mL, black caps with hole spots) were obtained from Hadi Traders, Antalkali, Lahore. Nylon syringe filters (pore size 0.45 µm, diameter 13 mm, item code SFNY01304 5NA) were used. An ultrasonic cleaner (DSA-100-SK1-2.8 L) was employed for 10-minute sonication to mix water and acetonitrile solutions properly.

The HPLC system (Model SHIMADZU) was equipped with a UV detector and a C-18 column (4.6 × 250 mm). The mobile phase consisted of acetonitrile and purified water (25:75 v/v) with a flow rate of 1 mL/min. Deltamethrin samples were analyzed at a wavelength of 230 nm.

Experimental Procedure:

The photocatalytic activity of the catalysts was evaluated by degrading Deltamethrin in a small-scale reactor. A stock solution of Deltamethrin (1000 mg/L) was prepared by dissolving Lambda-cyhalothrin in pure water. From this stock, samples with varying concentrations (5, 15, 30, 40, and 50 mg/L) were placed in the reactor. Absorbent (0.1-3.0 g/L) was added to each solution at different pH levels (3 to 9).

To maintain temperature stability, water circulation was used to counteract the heat generated by the lamp. A 500 mL Deltamethrin solution mixed with the appropriate absorbent was left in the dark for different time intervals (5 minutes to 1 hour) to establish adsorption-desorption equilibrium. A tungsten lamp (Philips Lighting Co.) served as the visible light source, and each experiment was conducted with a 2-hour irradiation period. To minimize environmental effects, the experiment was carried out in a glass-covered, closed-box reactor. A hygrometer was used to monitor ambient humidity, and power readings were taken regularly to ensure consistent illumination.

Samples were collected from the reactor at 5, 15, 30, 45, and 60-minute intervals and centrifuged at 5000 rpm for 5 minutes. The ultrasonic cleaner (DSA-100-SK1-2.8 L) was used for 10-minute sonication to properly mix the water and acetonitrile solution. The HPLC injector tube was immersed in the mobile phase, and the system was connected and powered on. GC-HPLC was used to analyze Deltamethrin degradation. The system's LC software was pre-installed on a PC, where parameters such as retention time, temperature, wavelength, and injection volume were set. A 5 µL sample extract was injected into the column and eluted with the mobile phase (water: acetonitrile, 25:75 v/v) at a flow rate of 1.0 mL/min. An absorption wavelength of 230 nm was used to generate degradation curves. The obtained data was then used to calculate pesticide degradation percentages.

$$Percentage\ Degradtion\ = \left(\frac{C_0 - C}{C_0}\right) x\ 100$$

Denote "C0" = initial absorbance sample before degradation and "C" = absorbance sample of the Lambda-cyhalothrin after degradation solution at time "t."

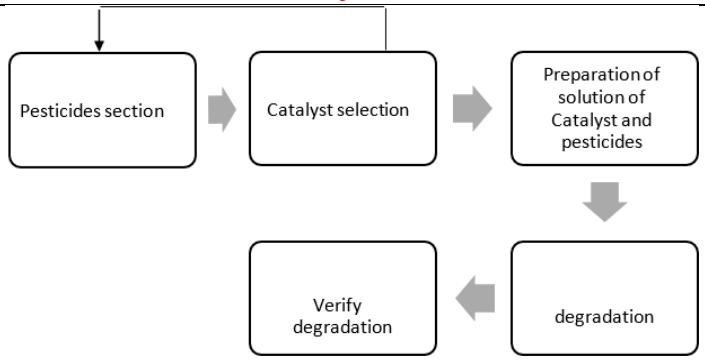


Figure 1. Flow Diagram of Methodology

Kinetic Study:

The study on the photodegradation of Deltamethrin included a kinetic analysis. The results showed that its degradation followed a pseudo-first-order kinetic pattern under photocatalytic conditions, aligning with theLangmuir–Hinshelwood model [22]. The developed model for Deltamethrin degradation was documented as follows.

$$\ln\frac{C_t}{C_0} = -k\ x\ t$$

Here, **C** represents the concentration of Deltamethrin pesticide (measured in parts per million, ppm), while **k** denotes the pseudo-first-order rate constant.

The coefficient of determination (**R^2**) is used to determine the reaction order and is calculated as follows:

$$R^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y}_i)^2}$$

The coefficient of determination (R^2) was found to be 0.9955, demonstrating a strong correlation between the kinetic model and the photodegradation of Deltamethrin using ZnO/TiO_2 as the catalyst [23].

Table 1. Rate constant and R^2

Pollutant	Catalyst	Light Source	k (min^{-1})	R^2
Deltamethrin	TiO_2	Visible light	0.031	0.9955
Deltamethrin	ZnO	Visible light	0.026	0.9873

After that, get the HPLC findings for several samples. A calibration curve is drawn. After that, we use the formula to calculate the deterioration efficiency in percentage (percent).

$$Percentage\ Degradtion = \left(\frac{C_0 - C}{C_0}\right)$$

C_0 = concentration before Degradation C= concentration after Degradation

Results:

TiO_2 Degradation Efficiency at Different PH:

Figure 2 illustrates the effect of pH variation on the degradation efficiency of Deltamethrin insecticides across different samples. The graph highlights that S-1 exhibits the highest degradation, particularly in a basic medium, reaching a peak efficiency of 96.1%. When synthesized catalysts were used, a steady increase in degradation rates was observed across all pH levels. However, TiO_2 demonstrated the best performance in the basic pH range.

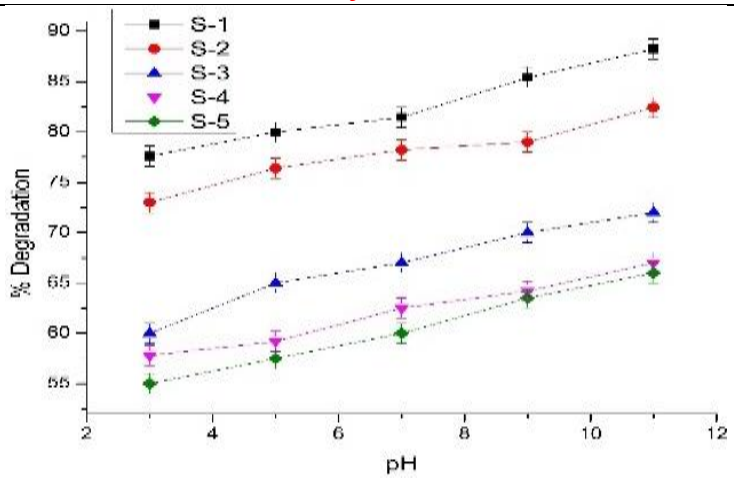


Figure 2. Efficiency of degradation at various pH of Deltamethrin Samples with TiO2

TiO_2 Catalyst: Different Dosing and Degradation Efficiency

Figure 3 shows that the highest degradation, around 96%, is achieved with catalyst dosage C-1 of S-1. This occurs because increasing the TiO_2 concentration eventually reduces the degradation rate as the catalyst dosage increases. Higher concentrations are more challenging to degrade, which is why the graph indicates minimal degradation for the S-5 sample.

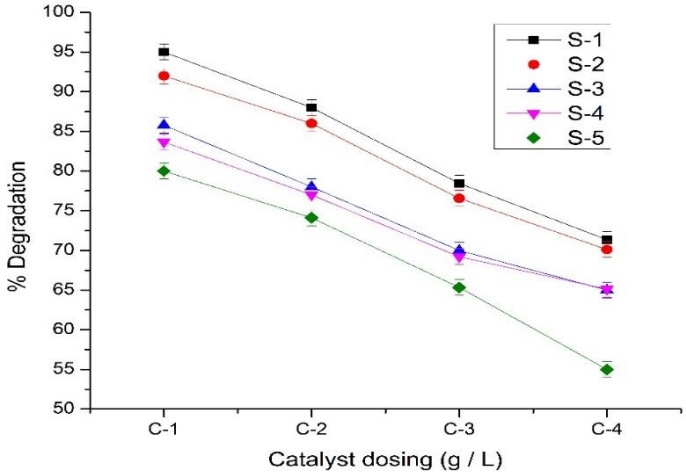


Figure 3. The efficiency of degradation of TiO2 of various samples

Efficiency of ZnO Degradation at Different pH Levels for Deltamethrin Samples:

The effect of pH variation on the degradation efficiency of different samples was analyzed. Figure 4 shows that the highest degradation for S-1 occurred at pH 9, compared to other pH levels. This pH test was conducted using C-2 ZnO.

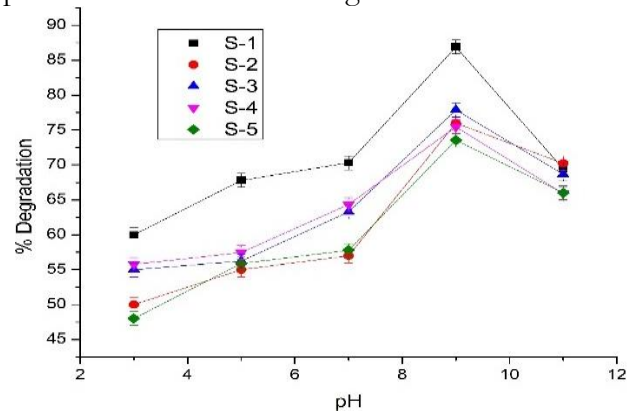


Figure 4. The efficiency degradation at various pH of the Deltamethrin Sample with ZnO

ZnO Catalyst Different Dosing Efficiency of Degradation:

The catalyst dosage C-3 of S-1 achieves the highest degradation rate, approximately 82%. As shown in Figure 5, increasing the concentration of ZnO positively impacts degradation efficiency. With ZnO C-1, the degradation rate is the slowest, around 62% for S-1. As the ZnO dosage increases, both the degradation rate and efficiency improve. However, the highest degradation value is observed at C-3 ZnO.

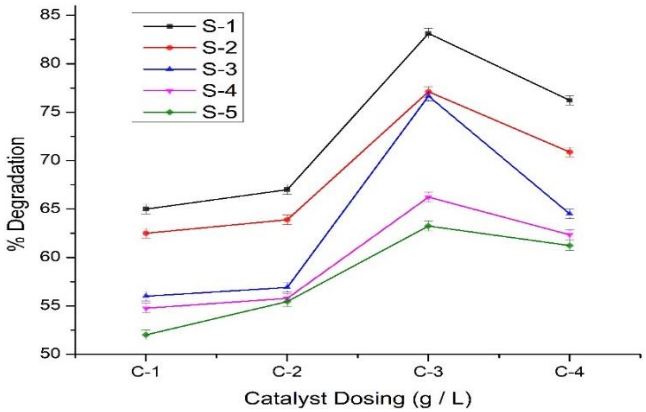


Figure 5. The efficiency of degradation of ZnO with various samples

Comparability of Efficiency of ZnO and TiO_2 in Degradation

Figure 6 illustrates the degradation efficiency of Deltamethrin insecticides using two different catalysts at various pH levels. When TiO_2 is used, the degradation rate is lower in acidic conditions but increases in basic conditions, reaching its highest efficiency at pH 11. In contrast, ZnO achieves its fastest degradation rate at pH 9.

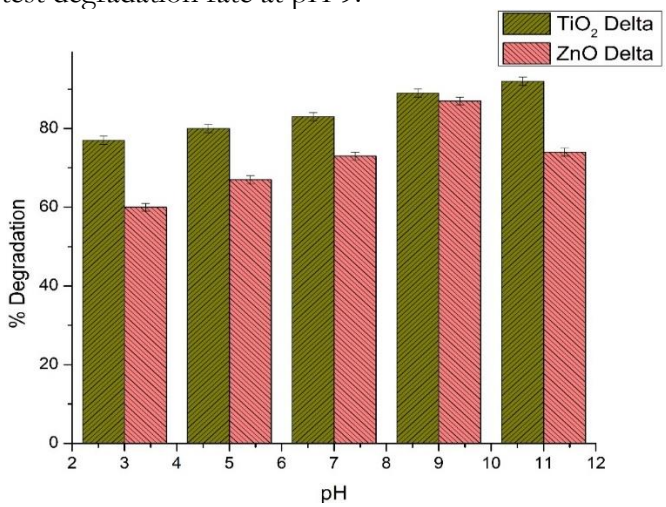


Figure 6. The efficiency of degradation at different pH of Deltamethrin Sample with ZnO and TiO2

Comparison of COD Values of ZnO and TiO2 W.R.T Time:

Figure 7 shows that the COD value decreases over time gradually. Maximum value of degradation of COD of Deltamethrin with TiO2.

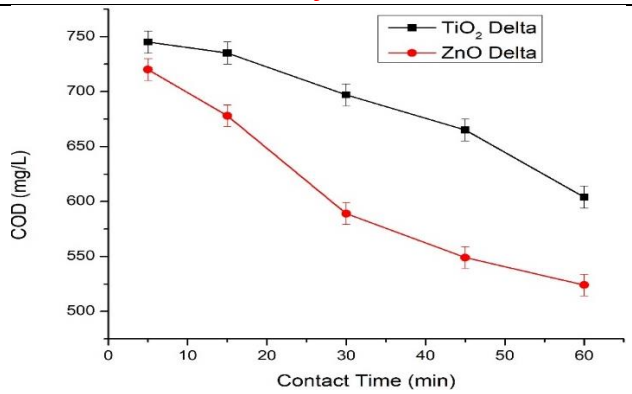


Figure 7. COD degradation concerning time

Discussion:

Reusability of ZnO/TiO2 Nano Catalyst:

The modification of ZnO and TiO_2 photocatalysts has been extensively studied to enhance their efficiency, stability, and practical use in wastewater treatment. Pure ZnO and TiO_2 have limitations, such as rapid charge recombination, limited absorption of visible light, and reduced effectiveness under real environmental conditions. To address these challenges, various modification techniques have been employed, including metal doping, non-metal doping, heterojunction formation, and surface functionalization. These modifications improve charge separation, enhance light absorption, and increase the degradation efficiency of persistent organic pollutants like pesticides.

Given the importance of cost-effectiveness in nanocatalyst usability, an in-depth analysis was conducted through multiple iterations. After each cycle, the used ZnO/TiO_2 photocatalyst was collected magnetically [24]. Metal doping is one of the most effective strategies to improve photocatalytic performance. Transition metals such as silver (Ag), copper (Cu), and iron oxide (Fe_3O_4) are incorporated into the ZnO or TiO_2 lattice. Ag-doped TiO_2 and ZnO enhance visible light absorption through surface plasmon resonance (SPR), enabling more efficient charge carrier generation. Similarly, Cu-doped TiO_2 increases electron transfer rates, reducing recombination losses and improving effectiveness under visible light. The addition of Fe_3O_4 to ZnO creates a magnetic photocatalyst, allowing easy recovery and reuse, thereby enhancing long-term sustainability. These modifications result in higher degradation rates, improved catalyst stability, and better selectivity for organic pollutants in water treatment applications.

Non-metal doping, where oxygen atoms in the ZnO or TiO_2 lattice are replaced with elements such as nitrogen (N), sulfur (S), and carbon (C), is another effective method for enhancing photocatalytic efficiency. N-doped TiO_2 introduces mid-gap energy states, shifting light absorption from UV to the visible spectrum, making it suitable for outdoor applications. Similarly, S-doped ZnO creates oxygen vacancies that enhance reactive oxygen species (ROS) generation, leading to more effective oxidative degradation of pesticides. Carbon doping forms graphitic carbon-TiO_2 hybrids, significantly increasing electron conductivity and photocatalytic efficiency. These modifications broaden the light absorption range of photocatalysts, making them more energy-efficient and effective under natural sunlight.

The formation of heterojunction and composite photocatalysts, such as ZnO-TiO_2, $BiVO_4$-TiO_2, and g-C_3N_4/ZnO, further improves charge separation and reaction kinetics. ZnO-TiO_2 composites benefit from ZnO's high charge mobility and TiO_2's strong oxidation ability, creating an efficient heterojunction that minimizes electron-hole recombination. $BiVO_4$-TiO_2 hybrids establish a Z-scheme charge transfer system, enhancing the degradation of complex organic pollutants. Additionally, coupling graphitic carbon nitride (g-C_3N_4) with ZnO stabilizes the catalyst structure, improving photocatalytic activity under sunlight. These

heterojunctions offer synergistic benefits, resulting in faster reaction rates, increased catalyst longevity, and better efficiency under real-world conditions.

To cleanse the photocatalyst of any residual compounds adsorbed onto its surface, multiple washes with water were performed [25]. Notably, even after seven cycles of photocatalytic activity, the degradation efficiency remained at approximately 82%. This result highlights the reusability potential of the ZnO/TiO_2 catalyst and its ability to suppress electron-hole recombination. Ultimately, the ternary magnetic photocatalyst demonstrated strong effectiveness in pesticide degradation [26][27].

The study focuses on improving pesticide decomposition by adjusting concentration, using catalysts, and modifying pH levels [28]. The findings indicate that optimizing these factors significantly enhances pesticide degradation efficiency. Additionally, the decrease in Chemical Oxygen Demand (COD) confirms the efficient mineralization of pesticide residues, suggesting a reduced environmental impact [29]. These results provide a foundation for developing more effective and environmentally friendly pesticide remediation techniques, supporting sustainable wastewater treatment and agricultural practices [24]. Further research is recommended to refine these methods and explore their broader applications [3].

Limitations of the Study:

- **Environmental Variability**: Factors such as temperature, humidity, and soil composition can affect pesticide degradation, making it challenging to generalize the data.
- **Limited Pesticide Scope**: The study may focus on specific pesticides, limiting their applicability to others with different chemical properties.
- **Analytical Restrictions**: The accuracy of degradation rate measurements may be affected by the sensitivity limits of detection techniques.
- **Time and Resource Constraints**: Long-term degradation studies require extended monitoring and substantial resources, which may not always be feasible.

Conclusion:

Under optimal conditions, TiO_2 and ZnO photocatalysts effectively degrade these pollutants. Deltamethrin degradation reached 96.34% for C-1 and 83.2% for C-3 using TiO_2 and ZnO, respectively, at varying adsorption concentrations. In terms of contact time, TiO_2's efficiency decreased after 10 minutes, while ZnO remained effective for up to 30 minutes, making ZnO more impactful in pollutant removal. The optimal degradation rate for TiO_2 with Deltamethrin was observed at pH 11, whereas ZnO exhibited the highest breakdown rate at pH 9. This technique can be applied in various chemical industries to degrade harmful pollutants. Compared to ZnO, TiO_2 nanoparticles under visible light irradiation demonstrated superior efficiency in decomposing Deltamethrin.

Acknowledgment:

The authors express their gratitude to the University of Gujrat for providing the necessary resources for research. Special thanks to Dr. Muhammad Zubair, Associate Professor, Department of Chemistry, University of Gujrat, for his invaluable guidance.

References:

[1] S. Fourmentin, G. Crini, and E. Lichtfouse, Eds., "Cyclodextrin Fundamentals, Reactivity and Analysis," *Environ. Chem. a Sustain. World*, vol. 16, 2018, doi: 10.1007/978-3-319-76159-6.

[2] P. Prashar and S. Shah, "Impact of Fertilizers and Pesticides on Soil Microflora in Agriculture," *Sustain. Agric. Rev.*, pp. 331–361, 2016, doi: 10.1007/978-3-319-26777-7_8.

[3] I. H. Khuda Bakhsh, Nameem Ahmad, Shanza Tabasum, Sarfraz Hassan, "Health hazards and adoption of personal protective equipment during cotton harvesting in Pakistan," *Sci. Total Environ.*, vol. 598, pp. 1058–1064, 2017, doi: https://doi.org/10.1016/j.scitotenv.2017.04.043.

[4] M. R. W. Jeffrey Farner Budarz, Ellen M. Cooper, Courtney Gardner, Emina Hodzic, P. Lee Ferguson, Claudia K. Gunsch, "Chlorpyrifos degradation via photoreactive TiO2 nanoparticles: Assessing the impact of a multi-component degradation scenario," *J. Hazard. Mater.*, vol. 372, pp. 61–68, 2019, doi: https://doi.org/10.1016/j.jhazmat.2017.12.028.

[5] Ganeshwade R. M., "Biochemical changes induced by dimethoate (Rogor 30% EC) in the gills of fresh water fish Puntius ticto (Hamilton)," *J. Ecol. Nat. Environ.*, vol. 4, no. 7, Apr. 2012, doi: 10.5897/JENE11.134.

[6] A. Marican and E. F. Durán-Lara, "A review on pesticide removal through different processes," *Environ. Sci. Pollut. Res. 2017 253*, vol. 25, no. 3, pp. 2051–2064, Nov. 2017, doi: 10.1007/S11356-017-0796-2.

[7] J. V. Franck Delval, Grégorio Crini, "Removal of organic pollutants from aqueous solutions by adsorbents prepared from an agroalimentary by-product," *Bioresour. Technol.*, vol. 97, no. 16, pp. 2173–2181, 2006, doi: https://doi.org/10.1016/j.biortech.2005.09.022.

[8] A. K. Jain, V. K. Gupta, A. Bhatnagar, and Suhas, "A Comparative Study of Adsorbents Prepared from Industrial Wastes for Removal of Dyes," *Sep. Sci. Technol.*, vol. 38, no. 2, pp. 463–481, 2003, doi: 10.1081/SS-120016585.

[9] S. V. Vinod Kumar Gupta, Rajeev Jain, "Removal of Reactofix golden yellow 3 RFN from aqueous solution using wheat husk—An agricultural waste," *J. Hazard. Mater.*, vol. 142, no. 1–2, pp. 443–448, 2007, doi: https://doi.org/10.1016/j.jhazmat.2006.08.048.

[10] V. K. S. Vinod Kumar Gupta, Imran Ali, "Defluoridation of wastewaters using waste carbon slurry," *Water Res.*, vol. 41, no. 15, pp. 3307–3316, 2007, doi: https://doi.org/10.1016/j.watres.2007.04.029.

[11] A. N. V.K. Gupta, Bina Gupta, Arshi Rastogi, Shilpi Agarwal, "A comparative investigation on adsorption performances of mesoporous activated carbon prepared from waste rubber tire and activated carbon for a hazardous azo dye—Acid Blue 113," *J. Hazard. Mater.*, vol. 186, no. 1, pp. 891–901, 2011, doi: https://doi.org/10.1016/j.jhazmat.2010.11.091.

[12] T. A. Saleh and V. K. Gupta, "Column with CNT/magnesium oxide composite for lead(II) removal from water," *Environ. Sci. Pollut. Res.*, vol. 19, no. 4, pp. 1224–1228, May 2012, doi: 10.1007/S11356-011-0670-6/METRICS.

[13] A. K. Jain, V. K. Gupta, S. Jain, and Suhas, "Removal of Chlorophenols Using Industrial Wastes," *Environ. Sci. Technol.*, vol. 38, no. 4, pp. 1195–1200, Feb. 2004, doi: 10.1021/ES034412U.

[14] M. M. Mubeena Akhtar, Shahid Iqbal, M.I. Bhanger, Muhammad Zia-Ul-Haq, "Sorption of organophosphorous pesticides onto chickpea husk from aqueous solutions," *Colloids Surfaces B Biointerfaces*, vol. 69, no. 1, pp. 63–70, 2009, doi: https://doi.org/10.1016/j.colsurfb.2008.10.019.

[15] A. N. Vinod K. Gupta, Arshi Rastogi, "Biosorption of nickel onto treated alga (Oedogonium hatei): Application of isotherm and kinetic models," *J. Colloid Interface Sci.*, vol. 342, no. 2, pp. 533–539, 2010, doi: https://doi.org/10.1016/j.jcis.2009.10.074.

[16] G. S. S. Karthikeyan, V.K. Gupta, R. Boopathy, A. Titus, "A new approach for the degradation of high concentration of aromatic amine by heterocatalytic Fenton oxidation: Kinetic and spectroscopic studies," *J. Mol. Liq.*, vol. 173, pp. 153–163, 2012, doi: https://doi.org/10.1016/j.molliq.2012.06.022.

[17] G. N. Simón Navarro, José Fenoll, Nuria Vela, Encarnación Ruiz, "Removal of ten pesticides from leaching water at pilot plant scale by photo-Fenton treatment," *Chem. Eng. J.*, vol. 167, no. 1, pp. 42–49, 2011, doi: https://doi.org/10.1016/j.cej.2010.11.105.

[18] S. M. R. M.M. Ballesteros Martín, J.A. Sánchez Pérez, J.L. Casas López, I. Oller, "Degradation of a four-pesticide mixture by combined photo-Fenton and biological oxidation," *Water Res.*, vol. 43, no. 3, pp. 653–660, 2009, doi: https://doi.org/10.1016/j.watres.2008.11.020.

[19] J. P. Farré, María José, Doménech, Xavier, "Assessment of photo-Fenton and biological treatment coupling for Diuron and Linuron removal from water," *Water Res.*, vol. 14, no.

13, pp. 2533–2540, 2006, doi: https://doi.org/10.1016/j.watres.2006.04.034.

[20] D. W. Lhomme, Ludovic, Brosillon, Stephan, "Photocatalytic degradation of a triazole pesticide, cyproconazole, in water," *J. Photochem. Photobiol. A Chem.*, vol. 188, no. 1, pp. 34–42, 2007, doi: https://doi.org/10.1016/j.jphotochem.2006.11.015.

[21] J. L. Casas López, A. Cabrera Reina, E. Ortega Gómez, M. M. Ballesteros Martín, S. Malato Rodríguez, and J. A. Sánchez Pérez, "Integration of Solar Photocatalysis and Membrane Bioreactor for Pesticides Degradation," *Sep. Sci. Technol.*, vol. 45, no. 11, pp. 1571–1578, Jan. 2010, doi: 10.1080/01496395.2010.487465.

[22] J. S. J. Su-Gyeong Kim, Love Kumar Dhandole, Young-Seok Seo, Hee-Suk Chung, Weon-Sik Chae, Min Cho, "Active composite photocatalyst synthesized from inactive Rh & Sb doped TiO2 nanorods: Enhanced degradation of organic pollutants & antibacterial activity under visible light irradiation," *Appl. Catal. A Gen.*, vol. 564, pp. 43–55, 2018, doi: Applied Catalysis A: General.

[23] P. P. Emmanouil S. Karafas, Manolis N. Romanias, Vassileios Stefanopoulos, Vassilios Binas, Apostolos Zachopoulos, George Kiriakidis, "Effect of metal doped and co-doped TiO2 photocatalysts oriented to degrade indoor/outdoor pollutants for air quality improvement. A kinetic and product study using acetaldehyde as probe molecule," *J. Photochem. Photobiol. A Chem.*, vol. 371, pp. 255–263, 2019, doi: https://doi.org/10.1016/j.jphotochem.2018.11.023.

[24] A. Y. Sana Ahmad, "Photocatalytic degradation of deltamethrin by using Cu/TiO2/bentonite composite," *Arab. J. Chem.*, vol. 13, no. 11, pp. 8481–8488, 2020, doi: https://doi.org/10.1016/j.arabjc.2020.07.019.

[25] A. Saljooqi, T. Shamspur, and A. Mostafavi, "Synthesis and photocatalytic activity of porous ZnO stabilized by TiO2 and Fe3O4 nanoparticles: investigation of pesticide degradation reaction in water treatment," *Environ. Sci. Pollut. Res.*, vol. 28, no. 8, pp. 9146–9156, Feb. 2021, doi: 10.1007/S11356-020-11122-2/METRICS.

[26] L. R. M. Premalatha, N., "Surfactant modified ZnO–Bi2O3 nanocomposite for degradation of lambda- cyhalothrin pesticide in visible light: A study of reaction kinetics and intermediates," *J. Environ. Manage.*, vol. 246, pp. 259–266, 2019, doi: https://doi.org/10.1016/j.jenvman.2019.05.155.

[27] A. P. T. Abhishek Sraw, Taranjeet Kaur, Yamini Pandey, Amit Sobti, Ravinder K. Wanchoo, "Fixed bed recirculation type photocatalytic reactor with TiO2 immobilized clay beads for the degradation of pesticide polluted water," *J. Environ. Chem. Eng.*, vol. 6, no. 6, pp. 7035–7043, 2018, doi: https://doi.org/10.1016/j.jece.2018.10.062.

[28] Y. Abdel-Mobdy and E. Abdel-Rahim, "Toxicological Influences of Lambda Cyhalothrin and Evaluation of the Toxicity Ameliorative Effect of Pomegranate in Albino Rats," *Glob. Vet.*, vol. 14, no. 6, pp. 913–921, 2015, doi: 10.5829/idosi.gv.2015.14.06.95144.

[29] N. A. Y. Yadollah Abdollahi, Abdul Halim Abdullah, Zulkarnain Zainal, "Photocatalytic Degradation of p-Cresol by Zinc Oxide under UV Irradiation," *Int. J. Mol. Sci*, vol. 13, no. 1, pp. 302–315, 2012, doi: ttps://doi.org/10.3390/ijms13010302.

Reducing the Environmental Impact of Leather Production and Assessing the Potential of Cactus-Based Vegan Leather

Sumaira Tariq[1*], Malaika Moazzam[2], Asad Ali[1,3], Muhammad Saleem Khan[4], Muhammad Faheem Ullah[2], Yaseen Mughal[5], Muhammad Hasan[6], Khurram Shahzad Ayub[1]
[1]Department of Chemical Engineering, H.H. Campus, University of Gujrat, Pakistan.
[2]Forward Sports Pvt. Ltd, Punjab, Pakistan.
[3]School of Engineering, Edith Cowan University, Perth, Australia.
[4]Department of Chemical Engineering, NFC Institute of Engineering and Technology, Khanewal Road, Multan, Pakistan.
[5]Shaukat Soap & Ghee Industries (Pvt) Limited, Pakistan.
[6]Khatoon Industries Pvt LTD, Punjab, Pakistan.
*****Correspondence: sumairatariq907@gmail.com; malaikamoazzam8@gmail.com;
Citation | Tariq. S, Moazzam. M, Ali. A, Khan. M. S, Ullah. M. F, Mughal. Y, Hassan. M, Ayub. K. S, "Reducing the Environmental Impact of Leather Production and Assessing the Potential of Cactus-Based Vegan Leather", IJIST, Special Issue. pp 68-78, March 2025
Received| Feb 13, 2025 **Revised|** Feb 26, 2025 **Accepted|** March 03, 2025 **Published|** March 06, 2025.

Global warming and the environmental and health risks linked to animal-based leather products have increased the demand for sustainable alternatives. Vegan leather has gained attention as a promising solution to these issues, encouraging eco-friendly fashion. To reduce its environmental impact, the leather industry is shifting from animal-derived to plant-based materials. Traditional leather production involves slaughtering over a billion cattle each year, releasing harmful substances like chromium and lead that pollute water sources and threaten public health. This study explores the potential of cactus-based vegan leather as an eco-friendly substitute for conventional leather. The process involved harvesting mature cactus pads, drying them in the sun, and transforming them into a sturdy material that mimics the properties of real leather. Mechanical tests showed that cactus leather offers similar durability, flexibility, and aesthetic appeal to traditional leather. The results emphasize the environmental, economic, and functional advantages of cactus leather, positioning it as a scalable alternative to reduce the negative ecological effects of animal-based leather production.

Keywords: Environmental Impact, Animal Welfare, Cactus Leather, Sustainability, Global Shift.

Introduction:

Human activities, particularly those driven by socio-economic factors, are playing a major role in global environmental degradation [1]. Processing industries are among the key contributors to this crisis [2], and the leather industry is a clear example. It relies on animal hides from the meat industry and uses resource-heavy tanning processes [3]. Leather production consumes a large amount of water—about 40 liters per kilogram of hide—due to stages like soaking, tanning, and conditioning [4]. This process generates substantial wastewater, increasing biological oxygen demand (BOD), chemical oxygen demand (COD), and depleting dissolved oxygen in water bodies [2], [5]. Leather is a globally traded product, mainly derived from the meat and dairy industries [6], [7]. The industry depends heavily on these sectors, with 95% of raw materials coming from cows, lambs, pigs, and goats [8]–[9]. This reliance contributes to environmental issues, including at least 32,000 million tons of CO2 emissions annually [10].

In Pakistan, the leather sector ranks as the second-largest industry after textiles, providing jobs to over 200,000 people and contributing 5% to manufacturing GDP and 7% to exports [11], [12]. The country has around 596 tanneries, with more than 90% of their output exported [13]. However, the tanning process uses around 130 harmful chemicals, including sodium sulfide, chromium sulfate, and formaldehyde, which damage the environment. These chemicals pollute the air, soil, and water, harming agricultural land and reducing crop productivity [14], [15]. Industrial activities, especially tannery operations in Pakistan, are a major cause of pollution [6]. In Punjab and Khyber Pakhtunkhwa (KPK), tanneries significantly pollute rivers, agricultural fields, and residential areas. Chemical waste and untreated wastewater are discharged into waterways, harming crops and contaminating food supplies. Since many tanneries are located in residential areas, they pose serious health risks to urban populations.
In Karachi, untreated wastewater is dumped into the sea, while in Lahore and Punjab, it pollutes rivers [16]. Chemical waste clogs drainage systems, and sludge is often dumped openly, further damaging the environment and endangering public health [17]. Leather dust is both carcinogenic and allergenic, increasing health risks, particularly in areas like Korangi and Charsadda. Tanneries contribute 10–15% of the pollution along Karachi's coastline, while in Punjab, pollution from toxic river water used for irrigation reduces crop yields and affects food safety.

In recent years, the demand for sustainable and cruelty-free alternatives to animal leather has increased significantly due to growing environmental awareness and ethical concerns. The leather industry, while valued for its durability and aesthetic appeal, is associated with high water consumption, greenhouse gas emissions, and the use of toxic chemicals in tanning processes, contributing to significant environmental pollution. Additionally, synthetic leather alternatives, such as polyurethane (PU) and polyvinyl chloride (PVC) leather, pose their own challenges, including non-biodegradability and microplastic pollution. In response to these issues, researchers and manufacturers have been actively exploring bio-based and eco-friendly leather substitutes that can provide comparable mechanical properties while minimizing environmental impact. Cactus-based vegan leather has emerged as a promising alternative, offering a biodegradable, water-efficient, and carbon-negative solution for industries such as fashion, automotive, and upholstery. By utilizing renewable plant-derived materials, cactus leather aims to bridge the gap between sustainability and performance, ensuring durability, flexibility, and consumer acceptance while reducing dependence on animal-derived and petroleum-based materials.

Given these concerns, finding an alternative to animal-derived leather is essential. Although the demand for sustainable materials is rising, limited research has explored cactus-based vegan leather as a viable substitute. This study introduces vegan leather made from cacti as an eco-friendly alternative to conventional leather. It minimizes environmental impact by using less water, sequestering carbon, and optimizing glycerin concentration to enhance durability and flexibility. Cactus leather offers higher elasticity (95%) and comparable tensile

strength (up to 25 MPa) to traditional animal leather, whose tensile strength typically ranges from 15 to 30 MPa. However, producing animal leather requires resource-intensive processes that consume significant energy and release hazardous chemicals, including sulfides and chromium [18].

While PU leather is flexible and affordable, it generates 15–20 kg of CO_2 per square meter during production and significantly contributes to microplastic pollution. In contrast, cactus leather consumes only 200 liters of water per square meter and emits just 5 kg of CO_2, compared to animal leather, which uses a staggering 17,000 gallons of water per square meter [19]. Additionally, unlike PU and animal leather, cactus farming helps sequester carbon, absorbing up to 8 tons of CO_2 per hectare [8], [20]. For these reasons, cactus-based vegan leather stands out as a sustainable and environmentally friendly alternative to traditional leather.

Objectives:

The primary aims of this study are:

- Create vegan leather made from cacti as a sustainable substitute for animal and synthetic (PU) leather.
- Assess its mechanical characteristics, such as elasticity, flexibility, and tensile strength.
- Optimize drying conditions to increase production scalability and energy efficiency.
- Evaluate its effects on the environment in terms of carbon emissions, water use, and potential for CO2 sequestration.

Material and Methods:

Production Process:

The production process of cactus-based vegan leather combines plant-derived materials with synthetic polymers to achieve the desired properties, such as durability, flexibility, and elasticity. This blend enhances the leather's functionality while maintaining its eco-friendly characteristics. Figure 1 shows the flow diagram of methodology.

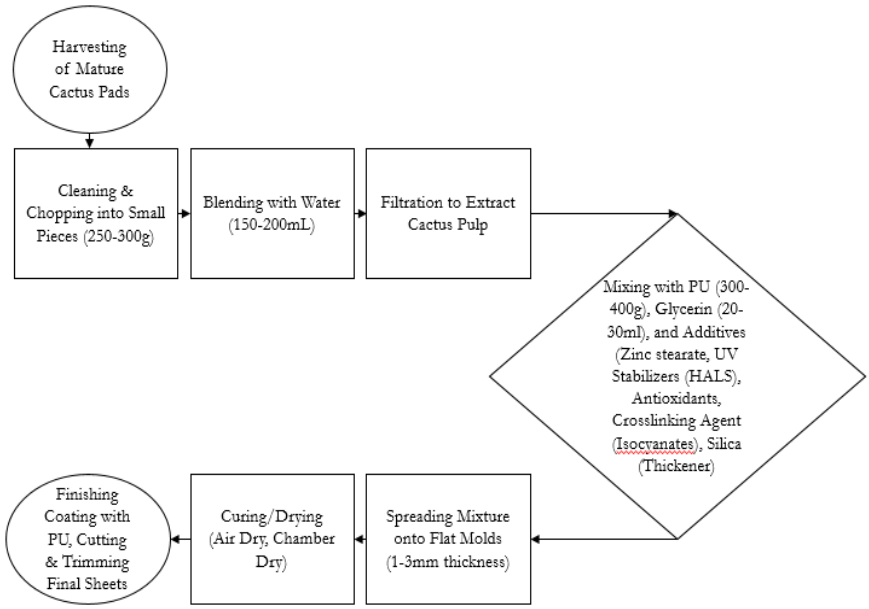


Figure 1. Flow diagram of methodology

Harvesting and Preparing the Cactus:

Cactus pads are harvested with care to protect the environment and ensure the plant's well-being [21]. The fine glochids (tiny thorns) are removed to create a smooth transition from raw plant material to leather [22], [23]. Next, the harvested pads are thoroughly washed to eliminate dirt and debris. They are then cut into pieces weighing 250-300 grams [24]. These pieces are blended with 150-200 milliliters of water to form a thick paste. The mixture is then

strained through a fine mesh or cheesecloth to separate the fibers and remove excess water, leaving behind a thick cactus pulp for further processing. Table 1 compares the composition, biodegradability, and toxic emissions of cactus leather, traditional animal leather, and synthetic PU leather.

Table 1: Technical Comparison with PU Leather and Animal Leather

Property	Animal Leather	PU Leather	Cactus Leather
Tensile Strength	15–30 MPa	10–25 MPa	25 MPa
Elasticity	50–100%	80–120%	95%
Water Use (per m^2)	17,000 liters	1,000 liters	200 liters
CO_2 Emissions (per m^2)	15–30 kg	15–20 kg	5 kg
Environmental Impact	High (pollutants)	Medium (plastics)	Low (biodegradable)

Preparing the Polyurethane Mix:

To prepare the polyurethane mixture, 250–300 g of dehydrated cactus biomass is combined with 300–400 g of polyurethane, which acts as a binding agent to enhance flexibility, durability, and water resistance. Glycerin (20–30 ml) is added to improve flexibility and prevent cracking. If desired, plant-based dyes or colorants (2–10 g) can be included to achieve the desired color.

Combining Cactus and Polyurethane:

The cactus pulp is gradually blended into the polyurethane mixture using a spatula to ensure even mixing. Once the mixture reaches the proper consistency, zinc stearate (10–15 g) is added to improve heat stability. Glycerin serves as a plasticizer, enhancing flexibility, while UV stabilizers (HALS, 5–10 g) protect the material from fading. Antioxidants (5–10 g) prevent long-term degradation. A crosslinking agent (isocyanates, 10–20 g) strengthens the bond between the cactus and polyurethane, and a thickening agent (silica, 10–15 g) helps achieve the desired viscosity. Together, these additives enhance the material's strength, stability, and performance.

Forming the Leather:

The mixture is then poured onto a flat mold or a non-stick silicone mat. Using a spatula or a similar tool, it is spread evenly to a thickness of 1–3 mm. For a textured finish, a patterned mold can be pressed onto the surface before curing, adding both visual and functional details.

Curing and Drying:

The material should be left to air-dry in a well-ventilated space for 24–48 hours. The drying process can be optimized under different conditions. Air drying at 25°C takes 48 hours and uses minimal energy. Chamber drying at higher temperatures—40°C, 50°C, 60°C, and 70°C—reduces the drying time to 8, 4, 3, and 2 hours, respectively, though energy consumption increases with temperature, ranging from 0.8 kWh at 40°C to 2.5 kWh at 70°C. Table 2 evaluates the mechanical performance of the three leather types in terms of tensile strength (durability), elasticity (flexibility), and abrasion resistance (wear durability over time).

Table 2. Indicating a trade-off between drying speed and energy use

Temperature (°C)	Drying Time (hours)	Energy Consumption (kWh)
25 (Air Drying)	48	0
40	8	0.8
50	4	1.2
60	3	1.8
70	2	2.5

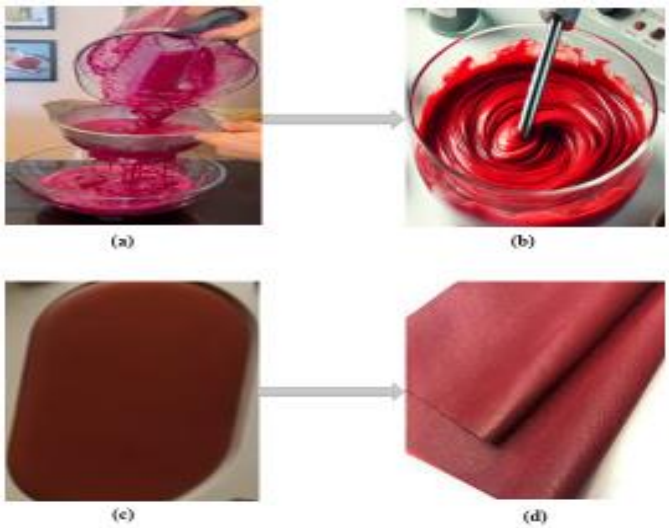


Figure 2. (a): Harvesting, cleaning, chopping, and preparing pulp **(b):** Mixing **(c):** Curing and drying **(d):** Leather sheets

Final Touches

After drying, a thin layer of polyurethane or natural sealant is applied to protect the surface from moisture and wear. The edges are then trimmed to the desired shape or size using scissors or a knife.

Results:

Testing of Mechanical Properties:

The mechanical properties of the cactus-based vegan leather were systematically evaluated to assess the impact of different glycerin concentrations on performance. At an optimized glycerin content of 15%, the material demonstrated a tensile strength of 25 MPa. It also achieved a flexibility score of 9 on a standard scale of 1 to 10. Additionally, with an elasticity of 95%, the material exceeded the lower range of animal leather elasticity (50–100%).

Table 3 highlights the water absorption and moisture retention properties of different leather types. Table 3 supports the claim that cactus leather offers an ideal balance between water resistance and breathability, making it suitable for fashion, footwear, and upholstery applications.

Table 3. Mechanical properties of cactus leather at different glycerin concentrations

Sr. #	Glycerin (%)	Tensile Strength (MPa)	Flexibility (1-10)	Elasticity (%)
1	5%	12	6	80%
2	10%	20	8	90%
3	15%	25	9	95%

The results underscore that the addition of glycerin as a plasticizer effectively enhances the tensile strength and elasticity of cactus-based leather, making it suitable for demanding applications.

Flexibility and Stretch Tests:

The flexibility and resistance to stress-induced cracking were evaluated to assess the performance of cactus leather at different glycerin concentrations. At 15% glycerin, the material achieved a flexibility score of 9 and successfully passed the crack test, indicating its capacity to endure mechanical stress without structural failure. However, at lower glycerin levels, particularly 5%, the material failed the crack test, highlighting the critical role of adequate plasticizer content in achieving optimal flexibility and durability. Table 4 quantifies the environmental impact of different leather types in terms of carbon footprint (CO_2 emissions) and water usage per square meter of production. Figure 3 illustrates the direct correlation between glycerin concentration and material flexibility, revealing that an increase in glycerin

content improves flexibility and significantly reduces the risk of cracking under mechanical stress.

Table 4. Flexibility scores and crack test results at varying glycerin concentrations

Sr. No	Glycerin (%)	Flexibility Score	Crack Test (Pass/Fail)
1	5%	6	Fail (Cracked)
2	10%	8	Pass
3	15%	9	Pass

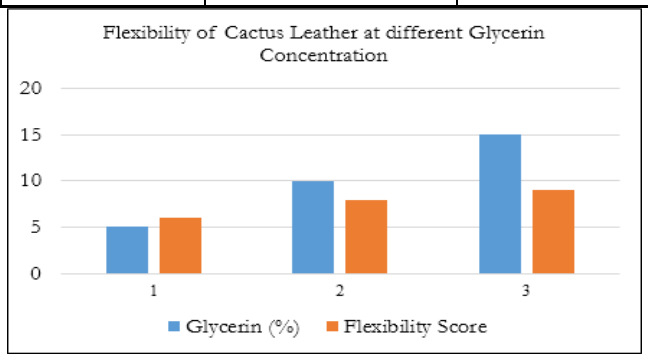


Figure 3. Correlation Between Glycerin Concentration and Material Flexibility

Efficiency of Drying and Curing:

The drying and curing process of cactus leather was evaluated to determine energy efficiency and production viability. When dried at a controlled temperature of 50°C, the process took 4 hours and used 1.2 kWh of energy. Increasing the temperature to 70°C reduced the drying time to 2 hours but raised energy consumption to 2.5 kWh. Table 5 compares the economic feasibility of cactus leather with animal leather and PU leather, based on production costs, processing time, and scalability for mass production. The cost of producing cactus leather ranges from $18 to $25 per square meter, making it cheaper than animal leather ($30–50/m²) but slightly pricier than PU leather ($10–20/m²).

Table 5: Drying conditions and energy consumption for cactus leather

Sr. No	Temperature (°C)	Drying Time (hours)	Energy (kWh)
1	25°C (Air dry)	48	0
2	50°C (Chamber)	4	1.2 kWh
3	70°C (Chamber)	2	2.5 kWh

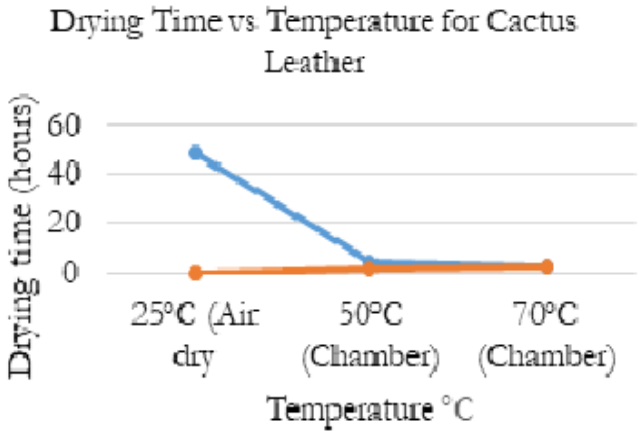


Figure 4. Impact of Drying temperature and drying time on energy consumption.

The graph in Figure 4 clearly shows that increasing the drying temperature shortens the drying time but also increases energy consumption (0–2.5 kWh).

Cactus Leather for CO_2 Sequestration:

Cacti can sequester an average of 8 tons of CO_2 per hectare, which translates to about 20 kg of CO_2 absorbed for every square meter of cactus leather produced.

Resource Consumption:

Cactus leather production has an impressively low water footprint, requiring just 200 liters of water per square meter.

Discussions:

The cactus-based vegan leather developed in this study exhibited outstanding mechanical, processing, and environmental properties, making it a promising alternative to traditional animal and synthetic leathers. Its mechanical performance, especially at a 15% glycerin concentration, showed a tensile strength of 25 MPa, elasticity of 95%, and a flexibility score of 9 [25]. This tensile strength is comparable to that of high-quality animal leather (15–30 MPa) and even surpasses the typical tensile strength range of PU leather (10–25 MPa) [26]. The high flexibility and fracture resistance observed at higher glycerin concentrations align with recent research on plant-based leather alternatives, which highlights the role of plasticizers in improving material flexibility and durability [27]. Additionally, the cactus leather showed excellent fracture resistance, passing crack tests at 15% glycerin concentration. This performance is on par with high-grade animal leather and exceeds PU leather, which tends to have lower crack resistance and may peel over time [27].

Table 6. Comparison between mechanical properties

Property	Animal Leather	PU Leather	Cactus Leather
Tensile Strength	15–30 MPa [24]	10–25 MPa	25 MPa
Elasticity	50–100% [24]	80–120%	95%
Flexibility Score	High (varies by grade)	Medium to High	9 (1–10 scale)
Crack Resistance	High	Moderate (can peel over time)	High (Passes crack test at 15% glycerin)

This study also assessed the effectiveness of drying and curing. Drying time was significantly reduced from 48 hours (air drying at 25°C) to just 2 hours using chamber drying at 70°C, highlighting the potential for scalable production [28]. In contrast, traditional animal leather drying typically takes 6 to 24 hours at temperatures between 40 and 60°C, while PU leather often requires high-temperature curing, leading to substantial energy consumption. Under optimal conditions, cactus leather dried efficiently within 2 to 4 hours at 50–70°C, noticeably reducing processing time [29]. However, as with other bio-material processing methods, energy usage increased with higher temperatures. Despite this, cactus leather consumed considerably less energy—between 0.8 and 2.5 kWh per batch—compared to PU and animal leather production, which require much higher energy input [30]. Additionally, the cactus leather drying process proved to be environmentally friendly, producing minimal VOC emissions, unlike PU leather, which releases significant solvent-based emissions, or animal leather, which poses chromium- and solvent-related environmental risks [24].

Table 7. Drying time and energy consumption

Process	Animal Leather	PU Leather	Cactus Leather
Drying Time	6–24 hrs at 40–60°C	Requires high-temp curing (varies)	2–4 hrs at 50–70°C
Energy Use	High (varies by process)	High (includes VOC emissions)	0.8–2.5 kWh per batch
VOC Emissions	Chromium & solvent risks	Significant (PU solvents)	Minimal (no VOCs during drying)

The environmental analysis of cactus leather production highlights its sustainability. Water usage for cactus leather was limited to just 200 liters per square meter—significantly lower than the 17,000 liters per square meter needed for animal leather production. Moreover, cactus

cultivation contributes to carbon sequestration, absorbing around 8 tons of CO_2 per hectare [31]. The life cycle carbon emissions for cactus leather were estimated at 5 kg of CO_2 per square meter, which is much lower compared to PU leather (15–20 kg CO_2 per square meter) and animal leather (up to 30 kg CO_2 per square meter) [32].

Table 8. Water consumption

Product	Animal Leather	PU Leather	Cactus Leather
Water Use per m^2	17,000 liters	1,000 liters	200 liters
Water Source	Intensive (livestock, tanning)	Moderate (industrial)	Low (rain-fed cactus crops)

In contrast to the high water demands of traditional animal leather, the use of drought-tolerant cactus species supports sustainable agricultural practices, making it an eco-friendly option for regions with limited water resources [33].

Table 9. Carbon footprint (CO_2 Emissions)

Product	Animal Leather	PU Leather	Cactus Leather
CO_2 Emissions per m^2	Up to 30 kg CO_2	15–20 kg CO_2	5 kg CO_2
Carbon Sequestration	None	None	8 tons CO_2 per hectare cactus plantation

This natural carbon capture ability makes cactus leather an eco-friendly material that not only lowers greenhouse gas emissions but also actively supports carbon offsetting [31].

Table 10: Biodegradability & Environmental Impact

Aspect	Animal Leather	PU Leather	Cactus Leather
Biodegradability	Low (chromium-tanned leather resists decay)	Very low (non-biodegradable, microplastics)	Moderate (plant-based, PU binder reduces it)
Chemical Use	Chromium, sulfides, formaldehyde [24]	Isocyanates, plasticizers, solvents	Minimal (PU binder, but lower than PU leather)
End-of-Life Impact	Toxic waste, landfill pollution	Persistent plastic waste	Lower impact, potential for improvement

Despite its numerous benefits, the current formulation of cactus leather has certain limitations, primarily due to the use of polyurethane (PU) as a binder, which reduces its overall biodegradability. Future research should aim to develop fully biodegradable binders to enhance the material's environmental performance. Additionally, optimizing resource inputs and energy consumption will be crucial when scaling production for large-scale industrial applications. Long-term durability studies under various environmental conditions, such as UV exposure and fluctuating humidity, are also recommended to validate the material's real-world performance.

Conclusions:

Cactus-based vegan leather demonstrates outstanding mechanical properties, especially at a glycerin content of 15%, where it achieves a tensile strength of 25 MPa, a flexibility rating of 9, and an elasticity of 95%. Its excellent flexibility and crack resistance make it a viable alternative to traditional leather for high-performance applications. The production process is highly energy-efficient, requiring just 2 to 4 hours of drying under controlled conditions, and it offers considerable carbon sequestration benefits, capturing up to 8 tons of CO_2 per hectare. The resource-efficient production process, which minimizes water and energy usage while significantly reducing carbon emissions, highlights cactus leather's potential as a sustainable and ethical alternative to conventional and synthetic leather. Future efforts should focus on optimizing glycerin content and exploring alternative biodegradable plasticizers to further

enhance the material's mechanical properties. Additionally, conducting long-term durability assessments and comprehensive lifecycle analyses will be essential to evaluating its performance over prolonged use.

Acknowledgment:

The authors thank the University of Gujrat for providing support. They also appreciate Forward Sports Pvt. Ltd, NFC Institute of Engineering and Technology, Shaukat Soap & Ghee Industries, and Khatoon Industries Pvt. Ltd for their valuable insights.

Author's Contributions:

Sumaira Tariq designed the study, conducted experiments, testing, and wrote the manuscript. Malaika Moazzam helped with industry collaboration and data analysis. Asad Ali and Muhammad Saleem Khan supervised the research and reviewed the manuscript. Muhammad Faheem Ullah worked on statistical analysis and literature review. Yaseen Mughal helped with raw material processing. Muhammad Hasan ensured quality control. Khurram Shahzad Ayub managed the project, and approved the final manuscript.

References:

[1] E. Mascot-Gómez, J. Flores, and N. E. López-Lozano, "The seed-associated microbiome of four cactus species from Southern Chihuahuan Desert," *J. Arid Environ.*, vol. 190, p. 104531, 2021, doi: https://doi.org/10.1016/j.jaridenv.2021.104531.

[2] V. V. S. S. A. K. S. Bhavya, P. Raji, A. Jenifer Selvarani, A. V. Samrot, P. T. M. Javad, "LEATHER PROCESSING, ITS EFFECTS ON ENVIRONMENT AND ALTERNATIVES OF CHROME TANNING," *Int. J. Adv. Res. Eng. Technol.*, vol. 10, no. 6, pp. 69–79, 2019, [Online]. Available: https://iaeme.com/MasterAdmin/Journal_uploads/IJARET/VOLUME_10_ISSUE_6/IJARET_10_06_009.pdf

[3] K. . T. R. Sreeram, "Sustaining tanning process through conservation, recovery and better utilization of chromium," *Resour. Conserv. Recycl.*, vol. 38, no. 3, pp. 185–212, 2003, doi: https://doi.org/10.1016/S0921-3449(02)00151-9.

[4] J. R. Rao *et al.*, "Pickle-free chrome tanning using a polymeric synthetic tanning agent for cleaner leather processing," *Clean Technol. Environ. Policy 2004 64*, vol. 6, no. 4, pp. 243–249, Jan. 2004, doi: 10.1007/S10098-003-0240-9.

[5] J. Kanagaraj, K. C. Velappan, N. K. Chandra Babu, and S. Sadulla, "Solid wastes generation in the leather industry and its utilization for cleaner environment - A review," *J. Sci. Ind. Res. (India).*, vol. 65, no. 7, pp. 541–548, Jul. 2006, doi: 10.1002/CHIN.200649273.

[6] S. Dixit, A. Yadav, P. D. Dwivedi, and M. Das, "Toxic hazards of leather industry and technologies to combat threat: a review," *J. Clean. Prod.*, vol. 87, pp. 39–49, 2015, doi: https://doi.org/10.1016/j.jclepro.2014.10.017.

[7] M. G. Arellano-Sánchez, C. Devouge-Boyer, M. Hubert-Roux, C. Afonso, and M. Mignot, "Chromium Determination in Leather and Other Matrices: A Review," *Crit. Rev. Anal. Chem.*, vol. 52, no. 7, pp. 1537–1556, 2022, doi: 10.1080/10408347.2021.1890545.

[8] L. S. Cornelia Wjunow, Kim-Laura Moselewski, Zoe Huhnen, Selina Sultanova, "Sustainable Textiles from Unconventional Biomaterials—Cactus Based," *Eng. Proc*, vol. 37, no. 1, p. 58, 2023, doi: https://doi.org/10.3390/ECP2023-14652.

[9] Khurram Shahzad Ayub, "Unveiling the Morphology and Composition of Heterogeneous Co, Zn and ferrite Catalysts for Efficient Catalysis," *J. Appl. Sci. Emerg. Technol.*, vol. 1, no. 2, 2023, [Online]. Available: https://jaset.uog.edu.pk/index.php/jaset/article/view/16

[10] L. M. and S. D. Dawar Butt, "CO2 Emissions from Pakistan's Energy sector," CREA. Accessed: Mar. 16, 2025. [Online]. Available:

https://energyandcleanair.org/publication/co2-emissions-from-pakistans-energy-sector/

[11] S. R. Ghafoor, Abdul, Manan Aslam, "Determinants of Leather Goods Exports: A Case of Pakistan," Journal of Business & Economics. Accessed: Mar. 16, 2025. [Online]. Available: https://www.researchgate.net/publication/257942065_Determinants_of_Leather_Goods_Exports_A_Case_of_Pakistan

[12] A. A. Chandio, J. Yuansheng, and H. Magsi, "Agricultural Sub-Sectors Performance: An Analysis of Sector-Wise Share in Agriculture GDP of Pakistan," *Int. J. Econ. Financ.*, vol. 8, no. 2, p. 156, Jan. 2016, doi: 10.5539/IJEF.V8N2P156.

[13] M. Khwaja, "Environmental Impacts of Tanning and Leather Products Manufacturing Industry in NWFP (Pakistan)." Sustainable Development Policy Institute, Jan. 01, 2000. Accessed: Mar. 16, 2025. [Online]. Available: https://www.academia.edu/54164465/Environmental_Impacts_of_Tanning_and_Leather_Products_Manufacturing_Industry_in_NWFP_Pakistan_

[14] R. B. Malabadi, K. P. Kolkar, R. K. Chalannavar, and H. Baijnath, "Plant-based leather production: An update," *World J. Adv. Eng. Technol. Sci.*, vol. 14, no. 1, pp. 031–059, Jan. 2025, doi: 10.30574/WJAETS.2025.14.1.0648.

[15] K. S. Ayub *et al.*, "Nonthermal plasma catalysis using ferrites as an efficient catalyst for toluene degradation," *Res. Chem. Intermed.*, vol. 49, no. 6, pp. 2399–2415, Jun. 2023, doi: 10.1007/S11164-023-05010-W/METRICS.

[16] M. S. Ummah, "中国の都市高齢者における主観的健康感の構造に関する研究," 民族衛生. Accessed: Mar. 16, 2025. [Online]. Available: https://www.jstage.jst.go.jp/article/jshhe1931/72/1/72_1_3/_article/-char/ja/

[17] R. S. Bikram Jit Singh, Ayon Chakraborty, "A systematic review of industrial wastewater management: Evaluating challenges and enablers," *J. Environ. Manage.*, vol. 348, p. 119230, 2023, doi: https://doi.org/10.1016/j.jenvman.2023.119230.

[18] A. Nefzaoui, "Opuntia ficus-indica productivity and ecosystem services in arid areas," *Italus Hortus*, vol. 25, no. 3, pp. 29–39, 2018, doi: 10.26353/J.ITAHORT/2018.3.2939.

[19] M. P. S. Zoé O. G. Schyns, "Mechanical Recycling of Packaging Plastics: A Review," *Macromol. Rapid Commun*, 2021, doi: https://doi.org/10.1002/marc.202000415.

[20] Y. M. Faten Mannai, Hanedi Elhleli, Mohamed Ammar, Raphaël Passas, Elimame Elaloui, "Green process for fibrous networks extraction from Opuntia (Cactaceae): Morphological design, thermal and mechanical studies," *Ind. Crops Prod.*, vol. 126, pp. 347–356, 2018, doi: https://doi.org/10.1016/j.indcrop.2018.10.033.

[21] Ian Oberem, "IHT Agri-Holdings: Unveiling the Future of Sustainable Fashion: Nopal Cactus and Vegan Leather," *I. H. T. Agri-holdings*, 2024, [Online]. Available: https://www.ihtsa.co.za/sitepad-data/uploads/2024/02/IHT-Agri-Holdings-Unveiling-the-Future-of-Sustainable-Fashion-Nopal-Cactus-and-Vegan-Leather.pdf

[22] J.-I. H. Muhammad Saif Ur Rehman, Naim Rashid, Ameena Saif, Tariq Mahmood, "Potential of bioenergy production from industrial hemp (Cannabis sativa): Pakistan perspective," *Renew. Sustain. Energy Rev.*, vol. 18, pp. 154–164, 2013, doi: https://doi.org/10.1016/j.rser.2012.10.019.

[23] "The Environmentally Frendly Leather Desserto Made of Nopal in Mexico," *ideass*, 2019, [Online]. Available: https://www.ideassonline.org/public/pdf/NopalLeatherMexico-ENG.pdf

[24] SupplyCompass, "Sustainable Material Guide // 03 Leather Alternatives," *Refashion fr*, 2020, [Online]. Available: https://refashion.fr/eco-design/sites/default/files/fichiers/Sustainable Material Guide Leather Alternatives.pdf

[25] A. A. B. Girmaw Yeshanbel Kefale, Zerihun Teshome Kebede, "A Systematic Review on Potential Bio Leather Substitute for Natural Leather," *J. Eng.*, 2023, doi: https://doi.org/10.1155/2023/1629174.

[26] Nabanita Saha; Fahanwi Asabuwa Ngwabebhoh; Hau Trung Nguyen; Petr Saha, "Environmentally friendly and animal free leather: Fabrication and characterization," *AIP Conf. Proc.*, vol. 2289, 2020, doi: https://doi.org/10.1063/5.0028467.

[27] N. M. S. N. Suderman, M.I.N. Is, "The effect of plasticizers on the functional properties of biodegradable gelatin-based film: A review," *Food Biosci.*, vol. 24, pp. 111–119, 2018, doi: https://doi.org/10.1016/j.fbio.2018.06.006.

[28] O. Y. Bai Zhongxue, Xuechuan Wang, Manhui Zheng, "Leather for flexible multifunctional bio-based materials: a review," *J. Leather Sci. Eng.*, vol. 4, no. 1, 2022, doi: 10.1186/s42825-022-00091-6.

[29] B. B. Varee Tyagi, "Role of plasticizers in bioplastics," *Food Process. Technol.*, vol. 7, no. 4, pp. 128–130, 2019, doi: 10.15406/mojfpt.2019.07.00231.

[30] J. P. K. Richard Opoku, George Y. Obeng, Louis K. Osei, "Optimization of industrial energy consumption for sustainability using time-series regression and gradient descent algorithm based on historical electricity consumption data," *Sustain. Anal. Model.*, vol. 2, p. 100004, 2022, doi: https://doi.org/10.1016/j.samod.2022.100004.

[31] L. Z. & A. O. F. Brugnoli, K. Sena, "A global study on the Life Cycle Assessment (LCA) of the modern cow leather industry," *Discov. Sustain.*, vol. 6, no. 80, 2025, doi: https://doi.org/10.1007/s43621-025-00798-6.

[32] F. Brugnoli and I. Kráľ, "Life Cycle Assessment , Carbon Footprint in Leather Processing," *Leather Leather Prod. Ind. Panel*, vol. 48, 2012, [Online]. Available: https://leatherpanel.org/sites/default/files/publications-attachments/lca_carbonfootprint_lpm2012.pdf

[33] B. F. Rafael Laurenti, Michael Redwood, Rita Puig, "Measuring the Environmental Footprint of Leather Processing Technologies," *J. Ind. Ecol.*, 2016, doi: https://doi.org/10.1111/jiec.12504.

Machine Learning-Based Improvement of Smart Contract Security in Fog Computing Using Word2vec and Bert

Tahmina Ehsan[1], Muhammad Usman Sana[1]*, Tayybah Kiren[2]*, Alvena Ehsan[1], Mustabeen Aziz[1], Fateha Minahil[1]
[1]University of Gujrat, Pakistan
[2]University of Engineering and Technology Lahore, Pakistan
***Correspondence**: m.usman@uog.edu.pk; tayybah@uet.edu.pk

Citation | Ehsan. T, Sana. M. U, Kiren. T, Ehsan. A, Aziz. M, Minahil. F, "Machine Learning-Based Improvement of Smart Contract Security in Fog Computing Using Word2vec and Bert", IJIST, Special Issue. pp 79-90, March 2025
Received| Feb 14, 2025 **Revised|** Feb 27, 2025 **Accepted|** March 04, 2025 **Published|** March 07, 2025.

Fog computing extends cloud computing services closer to users, improving efficiency and reducing latency. Smart contracts play a key role in authentication and resource access management within this framework. As the adoption of smart contracts in fog computing grows, ensuring their security has become a major challenge. This study enhances smart contract attack detection in fog computing using machine learning techniques. A dataset of 818 smart contracts was collected from "etherscan.io." Feature extraction was performed using Word2Vec and BERT, while feature selection was done using the information gain method. The Random Forest (RF) and Extra Trees Classifier (ETC) achieved the highest accuracy of 0.91 with Word2Vec, while the LightGBM (LGBM) classifier attained 0.90 accuracy using BERT.
These results demonstrate the effectiveness of machine learning models in improving smart contract security within fog computing environments.

Keywords: Fog Computing; Smart Contract; Machine Learning; Security and Feature Extraction

Introduction:

Cloud computing allows users to access computing resources such as servers, storage, software, databases, and applications over the internet instead of relying on local infrastructure. It operates on a pay-as-you-go model, enabling users to scale resources up or down as needed. IoT devices frequently use cloud resources, and their numbers are increasing daily [1]. While the growth of IoT has created many opportunities for cloud computing, it has also introduced challenges, including cost, data management, security, privacy, bandwidth limitations, network congestion, and latency issues. To address these challenges, Cisco introduced fog computing in 2018 as a bridge between cloud computing and edge computing [2].

Fog computing is a distributed model that extends cloud computing to the network's edge, providing computing, storage, and networking services closer to end users and IoT devices [3]. As shown in Figure 1, it creates an intermediate layer between the cloud and edge computing. This fog layer offers computing and networking resources to edge devices, reducing latency compared to traditional cloud computing. Fog computing is essentially an extension of cloud computing [3] and helps mitigate several cloud-related issues. Additionally, blockchain technology [4] is integrated into the fog layer to enhance security and privacy.

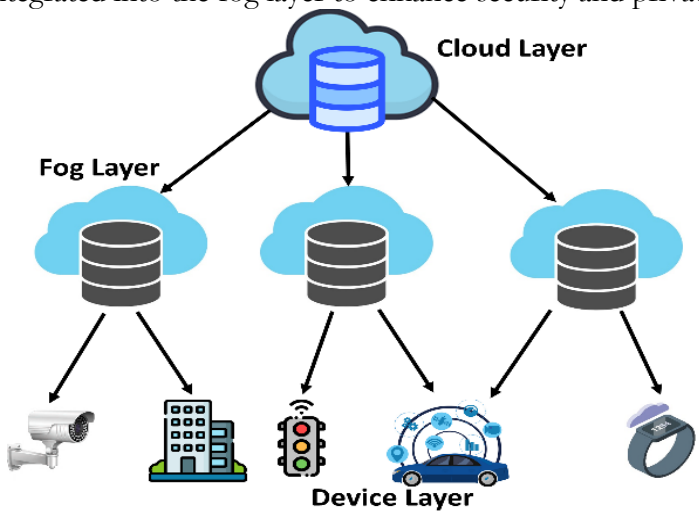


Figure 1. Architecture of fog computing.

Blockchain ensures secure data sharing among fog nodes, IoT devices, and cloud providers while requests, and verification. During user registration, the system assigns a pair of public and private keys, storing the public key within the blockchain. Resource registration lists available fog computing resources and the users who can access them.

To authenticate, a user sends a request using a nonce (a unique identifier) and their public key. The smart contract then follows a challenge-response protocol, sending back the nonce as a challenge. The user signs it with their private key and returns it. The smart contract verifies the signature using the user's public key. If valid, access to fog computing resources is granted [8][9]. Since smart contracts operate independently of external networks, a security breach can affect organizations, miners, and even the entire blockchain network [5][10]. Therefore, researchers must focus on identifying attacks that could compromise smart contract security. In this study, machine learning is applied to detect attacks in smart contracts used for resource access in fog computing.

Framework

Registration Phase

Figure 2 illustrates the user registration process required before accessing resources.

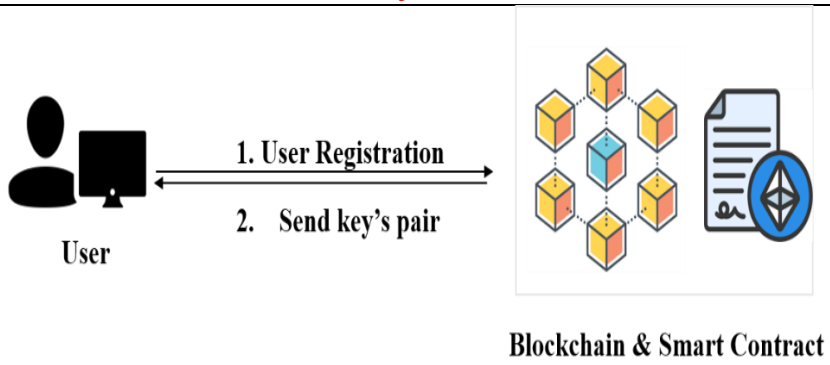


Figure 2. Registration Phase of User.

User Registration:

Users must register through a smart contract to access fog computing resources. Before creating an account, the blockchain-based smart contract verifies and confirms the registration details.

Send Key Pair:

After successful registration, the blockchain generates a cryptographic key pair (public and private keys) for secure authentication. The public key is stored on the blockchain, while the private key is securely sent to the user [11].

Resource Access Granting Phase

Figure 3 illustrates the resource access granting phase, which enables users to access the resources of the fog node.

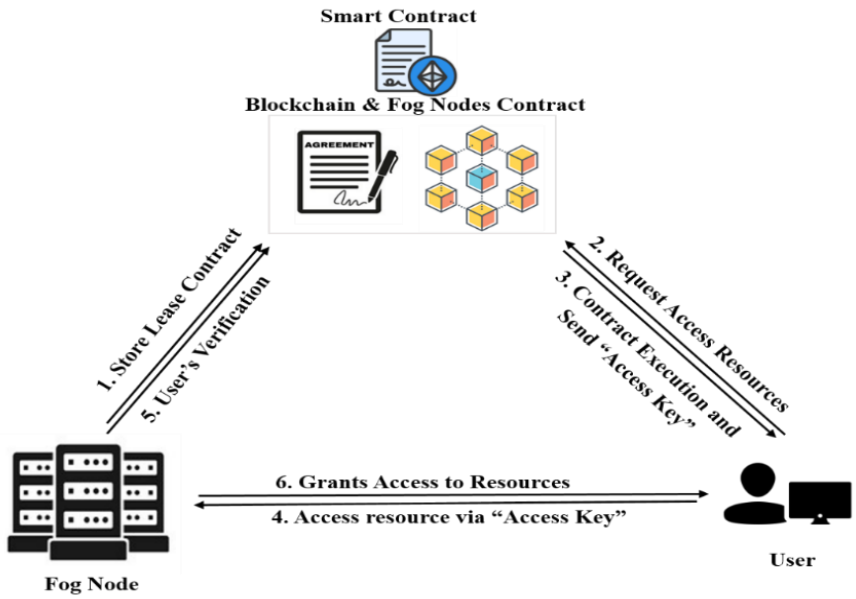


Figure 3. Resource Access Granting Phase.

Store Lease Contract:

In this step, the fog node submits its lease contract to the blockchain, defining the terms and conditions for resource access. This ensures secure and transparent resource allocation [11], [12].

Request Access Resources:

A registered user requests access to the fog node's resources. The smart contract verifies the request by matching the user's public key with the registered keys [12] and ensuring compliance with access policies.

Contract Execution and Send "Access key":

At this stage, after verification, the smart contract retrieves the access contract and securely sends the access key to the user [12].

Access resource via "Access Key":

After receiving the access key, the user can access the fog node's resources.

User's Verification:

The fog node verifies the user's identity by matching it with the blockchain ledger [12].

Grants Access to Resources:

After successful verification, access is granted, and a blockchain-based transaction system manages payments for the utilized resources [12].

Unlike previous research, which primarily focused on identifying smart contract vulnerabilities using conventional feature extraction techniques, this study introduces an improved approach by combining Word2Vec and BERT for opcode-based feature extraction. This method enhances the accuracy and efficiency of attack detection in fog computing smart contracts used for resource access and registration. Additionally, the paper evaluates various machine learning classifiers, demonstrating that Random Forest, Extra Trees Classifier, and LightGBM significantly improve security. Compared to previous studies, the proposed framework offers a more reliable, scalable, and precise attack detection technique.

The paper is organized as follows:

- **Section II** reviews related literature on fog computing, blockchain integration, and smart contract security.
- **Section III** outlines the research methodology, including data collection, feature extraction, and model selection.
- **Section IV** presents the experimental results and analysis.
- **Section V** concludes the study and discusses future research directions.

Objectives:

The objectives of this research are:

- To analyze attacks on smart contracts used in fog computing for resource access.
- To identify the most effective machine learning algorithms for detecting and preventing these threats.
- To evaluate the performance of machine learning models using F1 score, accuracy, precision, and recall.
- To improve smart contract attack detection by applying opcode-based feature extraction and selection techniques.

Literature Review:

Fog computing extends cloud computing but also inherits some of its challenges. Due to its proximity to IoT devices, it faces several security and privacy issues. Researchers have explored various solutions, including authorization, access control, and authentication, to ensure secure data transmission in fog computing. In [10], the author introduced a deep learning-based detection method to classify smart contracts as either malicious or safe. Techniques such as Long Short-Term Memory (LSTM), Gated Recurrent Units (GRU), and Artificial Neural Networks (ANN) were used for classification. The author utilized the BigQuery dataset for binary classification, achieving a maximum accuracy of 99.03%. Additionally, the Receiver Operating Characteristic (ROC) curve was provided to compare the performance of these models.

In [13], the author proposed a system for user registration and authentication in fog computing. This system uses smart contracts for registration and securely stores user information in a ledger. Compared to existing systems, it reduces registration and authentication costs. Additionally, it supports multiple user accounts and compares their costs, also known as gas values. In [6], the author analyzes 49,502 real-time smart contracts for various vulnerabilities,

including Callstack, integer overflow, timestamp, Time of Day (TOD), and re-entrancy, achieving a high accuracy rate of 99%. The research converts contract code into bytecode and opcodes. Then, n-gram features are extracted from the opcodes, and machine learning algorithms such as XGBoost, K-Nearest Neighbors (KNN), and Support Vector Machine (SVM) are applied. This approach enhances the speed and accuracy of vulnerability detection.

In [14], the author introduces a framework for identifying and classifying vulnerabilities in smart contracts, such as excessive gas consumption, unfixed compiler versions, implicit visibility levels, inappropriate use of pure functions, unchecked low-level calls, and frozen ether. These vulnerabilities are detected using publicly available datasets, including AutoMESC, which reports a 5.2% occurrence of high-severity vulnerabilities and suggests solutions for addressing them. In [12], the author explains how blockchain technology is used to manage resource access. Smart contracts facilitate this process by eliminating third-party dependencies within the network. These contracts are self-executing lines of code created by organizations, institutions, or other entities. In resource access scenarios, both buyers and sellers rely on smart contracts to define terms and conditions. If any condition is violated, the contract becomes invalid or is terminated.

In [15], the author identifies multiple vulnerabilities in smart contracts, including timestamp issues, re-entrancy, Time of Day (TOD) attacks, integer underflow, and overflow. Using the Bi-LSTM model, the author achieves an accuracy of 88.12%. A total of 5,450 smart contracts were collected from the Etherscan website to detect these vulnerabilities. First, the contract code is converted into opcodes, then a feature matrix is generated, and Bi-LSTM is applied for analysis. In [16], the author classifies smart contracts as normal or abnormal using an ensemble model. A dataset of 1,904 smart contracts was gathered from the Etherscan website. Features were extracted from the source code using TF-IDF, while opcode features were derived using the n-gram technique. Applying the ensemble model, the author achieved an accuracy of 89.67%.

In [17], the author analyzed 5,735 smart contracts, generating semantic trees based on their code and utilizing Graph Neural Networks (GNN) and Graph Matching Networks. The model achieved a 92.63% accuracy in detecting block info dependency vulnerabilities. Other vulnerabilities identified include re-entrancy, block info dependency, timestamp dependency, and TX.Origin issues. In [18], the author detects DDoS attacks in smart contracts using an IoT-based dataset [19]. The BotIoT dataset was used for this purpose. Features extracted from IoT sensors were stored in fog nodes before being transferred to the blockchain via smart contracts. Different classifiers, including Random Forest, Decision Tree, and Support Vector Machine, were applied, achieving an accuracy of 99.9%.

In [15], the author enhances smart contract security through machine learning. A total of 835 smart contracts were analyzed, with 455 classified as safe and 380 as malicious. A binary classification approach was used. The contract source code was first converted into opcodes, and a feature matrix was created. Various machine learning models, including KNN, Random Forest (RF), Decision Tree (DT), Logistic Regression (LR), Support Vector Machine (SVM), and Naïve Bayes (NB), were applied. The RF classifier achieved the highest accuracy at 85%.

In [20], the author presents a comprehensive approach to detecting vulnerabilities in smart contracts using machine learning, automated auditing tools, and reduced manual effort and execution time. The proposed model outperformed traditional methods, achieving an effectiveness rate of 80%. In [15], the author applies a machine-learning approach to detect abnormal smart contracts. A total of 835 smart contracts were collected from the Etherscan website, with 455 classified as normal and 380 as abnormal. Among the abnormal contracts, 327 were identified as scams, while 53 were found to be vulnerable. The dataset was preprocessed before applying various machine learning models for evaluation. The Random Forest (RF) model achieved an accuracy of 0.85, Logistic Regression (LR) reached 0.81, while K-Nearest

Neighbors (KNN) and Decision Trees (DT) scored 0.77. The Support Vector Machine (SVM) and Naïve Bayes (NB) models obtained accuracy results of 0.75 and 0.71, respectively. In [21], the author focuses on identifying Ponzi schemes, a type of fraud that lures new investors with false promises of high returns. A dataset of 3,786 smart contracts was sourced from the Kaggle website, containing four key features: address, opcode, label, and creator. After preprocessing the dataset and extracting relevant features, the author evaluated hybrid classifiers. By combining the strengths of XGBoost and GRU models, the study achieved an impressive accuracy of 96.8%.

The following are the research questions:

- How can machine learning be applied to detect attacks on smart contracts used for registration and resource access in fog computing environments?
- What methods can be used to protect smart contracts from attackers in a fog computing environment?
- How can opcode analysis serve as an efficient and effective feature extraction technique for detecting attacks in smart contracts?
- Which algorithms and models are best suited for detecting attacks on smart contracts used for registration and resource access in fog computing environments?

Material and Methods

This research focuses on attack detection using machine learning with a three-labeled dataset. The methodology is illustrated in Figure 4.

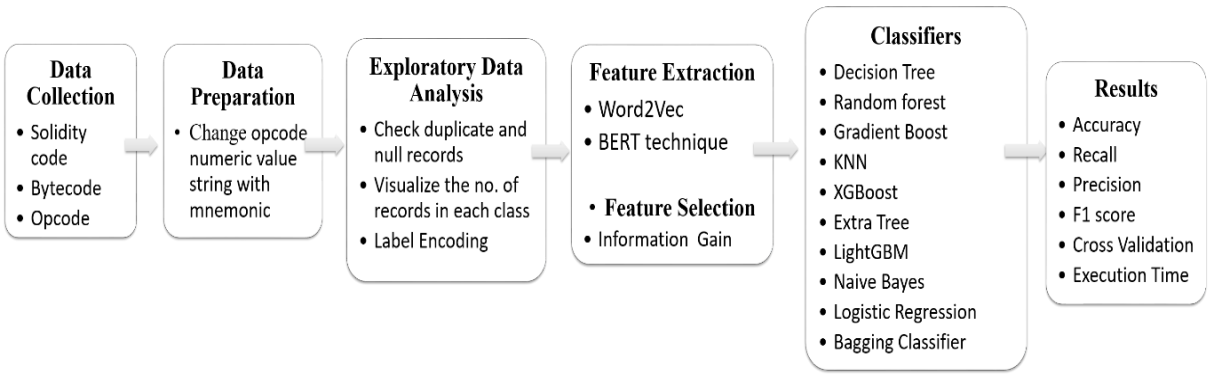


Figure 4. Proposed Methodology

Dataset:

In this research, 818 smart contracts were collected from Ethereum's official website, "etherscan.io," along with their Solidity code, bytecode, and opcode. The dataset is categorized into three labels: Ponzi (using Forta [22]), Phish-hack [9], and Gambling [23]. It includes 300 smart contracts under the Phish-hack label, 298 under Ponzi, and 220 under Gambling.

Table 1. Composition of Dataset

Label	Number of Smart Contract	Source
Ponzi	298	Forta [22]
Phish-hack	300	[9]
Gambling	220	[23]
Total	**818**	Etherscan.io

Data Preparation:

The opcode of a smart contract contains various hexadecimal values, starting with '0x'. These values are replaced with their corresponding mnemonic representations using [15]. Next, null records are verified, and the labels are encoded as follows: Ponzi ('0'), Phish-hack ('1'), and Gambling ('2').

Feature Extraction:

This research utilizes Word2Vec and BERT techniques to extract features from the opcode of smart contracts.

Feature Selection:

In this research, the information gain technique is used for feature selection, extracting features with a threshold above 0.03.

Classifiers:

Various machine learning classifiers are used, including Logistic Regression (LR), Decision Tree (DT), Random Forest (RF), K-Nearest Neighbors (KNN), Gradient Boost (GB), Bagging Classifier (BC), Naïve Bayes (NB), Extra Trees Classifier (ETC), Light Gradient Boost Machine (LGBM), and Extreme Gradient Boost (XGB).

Result and Discussions:

Various techniques can be used for feature extraction; however, this research employs two methods: Word2Vec and BERT. The results of both are discussed below.

Word2Vec Technique:

In this research, the Word2Vec technique is applied for feature extraction, with the results presented in Table 1.

Table 2. Results Using the Word2Vec Technique

Algorithm	Accuracy	Precision	Recall	F1-Score	CV Accuracy	Execution Time(s)
LR	0.84	0.85	0.84	0.84	0.79	1.67
DT	0.80	0.80	0.80	0.80	0.82	0.09
RF	0.91	0.92	0.91	0.91	0.88	8.78
KNN	0.88	0.91	0.88	0.89	0.85	0.02
GB	0.87	0.88	0.87	0.87	0.87	114.90
BC	0.86	0.87	0.86	0.86	0.86	0.29
NB	0.65	0.67	0.65	0.65	0.69	0.01
ETC	0.91	0.92	0.91	0.91	0.88	2.17
LGBM	0.87	0.87	0.87	0.87	0.87	3.21
XGB	0.85	0.85	0.85	0.85	0.87	7.00

The RF and ETC classifiers delivered the best performance with execution times of 8s and 2s, respectively. The NB classifier achieved an accuracy score of 0.88, while the GB and LGBM classifiers both attained 0.87 accuracy. The BC classifier followed with an accuracy of 0.86, and the XGB classifier reached 0.85 accuracy. The DT and NB classifiers produced lower results, with accuracy scores of 0.80 and 0.65, respectively. Figure 5 illustrates the comparison of all models based on Accuracy, Precision, Recall, and F1-score. According to this figure, the RF and ETC classifiers demonstrated the best performance.

BERT Technique:

In this research, the BERT technique is also used for feature extraction, and its results are presented in Table 2. The LGBM classifier achieved the highest accuracy score of 0.90 with an execution time of 24s. The GB, BC, and ETC classifiers followed closely, each attaining an accuracy of 0.88. The XGB classifier achieved an accuracy of 0.87, while the DT and RF classifiers both reached 0.85. The KNN and LR classifiers obtained accuracy scores of 0.84 and 0.79, respectively. The NB classifier recorded the lowest accuracy, scoring 0.66.

Figure 6 presents a comparison of all models based on Accuracy, Precision, Recall, and F1-score. According to this figure, the LGBM classifier delivers the best performance.

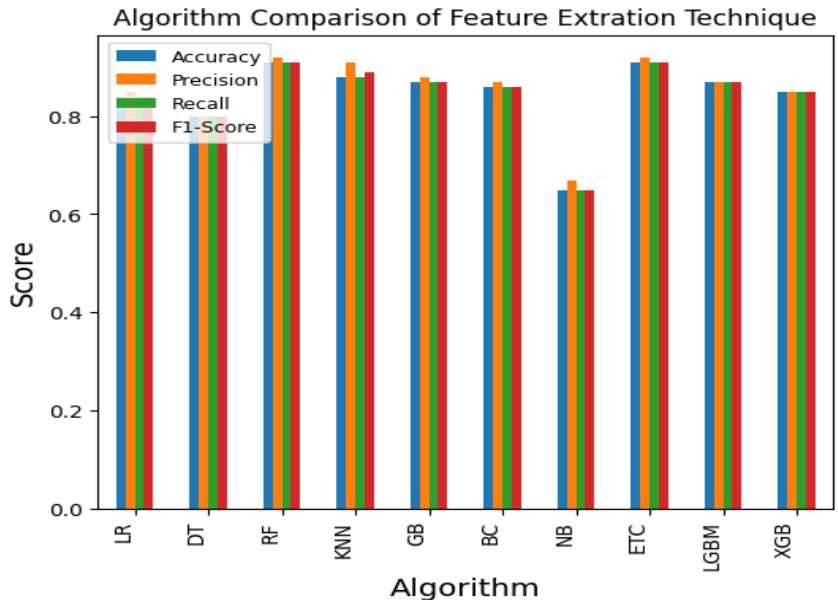


Figure 5. ML algorithms result using the Word2Vec technique.

Table 3. Results Using the Bert Technique

Algorithm	Accuracy	Precision	Recall	F1-Score	CV Accuracy	Execution Time(s)
LR	0.79	0.80	0.79	0.78	0.80	3.41
DT	0.85	0.86	0.85	0.85	0.83	0.52
RF	0.85	0.85	0.85	0.85	0.89	17.54
KNN	0.84	0.88	0.84	0.84	0.86	0.03
GB	0.88	0.88	0.88	0.88	0.87	744.09
BC	0.88	0.89	0.88	0.88	0.86	1.33
NB	0.66	0.68	0.66	0.63	0.70	0.03
ETC	0.88	0.89	0.88	0.88	0.90	10.47
LGBM	0.90	0.91	0.90	0.90	0.89	24.08
XGB	0.87	0.87	0.87	0.87	0.89	46.78

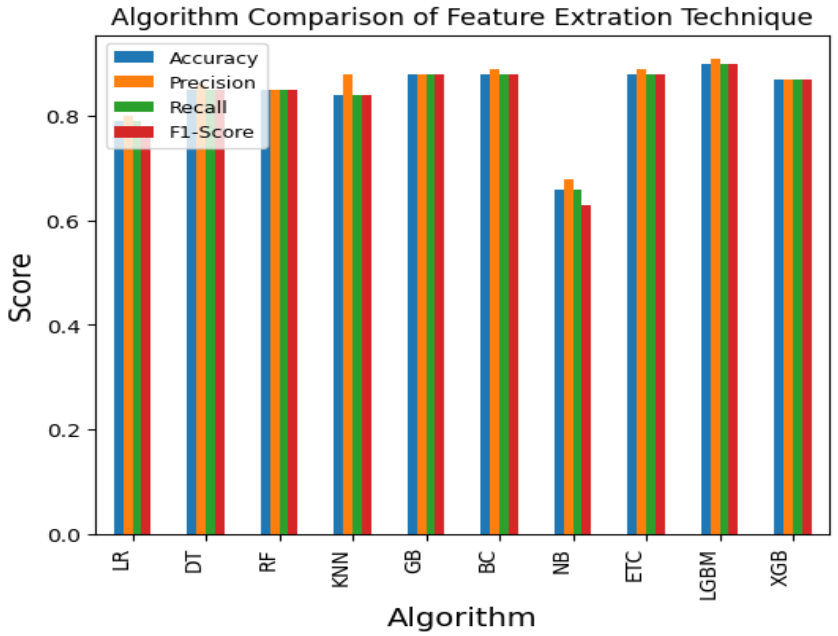


Figure 6. ML algorithms results using the BERT technique.

Fog computing faces several challenges in user registration and resource access, including resource management, security, and privacy concerns within the fog layer. Although smart contracts are used to manage these processes, they remain vulnerable to various attacks. This research enhances security by integrating machine learning for attack detection in smart contracts. Using the Word2Vec feature extraction technique, the Random Forest (RF) and Extra Trees Classifier (ETC) achieved an accuracy of 0.91. Meanwhile, the LightGBM (LGBM) classifier delivered the best performance with an accuracy of 0.90 using the BERT technique. These results highlight the effectiveness of combining machine learning with smart contracts to improve security, scalability, and real-time attack detection. The proposed methodology is highly scalable, enabling it to handle larger systems and diverse datasets, making it ideal for expanding fog computing networks.

Discussion:

The findings of this study demonstrate that machine learning-based security analysis using Word2Vec and BERT for opcode feature extraction significantly enhances smart contract vulnerability detection in fog computing environments. The results show that Random Forest (RF) and Extra Trees Classifier (ETC) achieved the highest accuracy (91%) with Word2Vec, while LightGBM (LGBM) performed best with BERT (90%). This highlights the importance of feature representation in improving attack detection. Compared to traditional methods using n-grams or TF-IDF for feature extraction, the proposed model achieves higher classification accuracy and fewer false positives, making it more effective for large-scale smart contract security monitoring. One key observation is the varying performance of different classifiers in detecting Ponzi schemes, phishing attacks, and gambling-related vulnerabilities. While ensemble models like RF and ETC demonstrated high accuracy and stability, models such as XGBoost and SVM had lower detection rates for specific attack types, indicating that classifier selection plays a crucial role in optimizing smart contract security. BERT-based feature extraction improved the contextual understanding of opcode sequences, making it easier to identify malicious patterns. However, the study did not analyze per-class accuracy, precision, or recall, which could provide deeper insights into each classifier's strengths and weaknesses in detecting different attack types.

Despite promising results, the study focuses solely on opcode-based detection, which, while effective, does not analyze code-level vulnerabilities. Issues such as reentrancy, unchecked external calls, and integer overflow in Solidity-based smart contracts cannot always be identified through opcode analysis alone. Existing research suggests that integrating opcode-based classification with static analysis tools (e.g., Slither, Mythril) could enhance security by detecting both pattern-based and logic-based vulnerabilities. Future research should explore hybrid detection models that combine opcode and code-level analysis to create a more comprehensive security framework.

Another limitation is the model's robustness against adversarial attacks. Machine learning-based security systems are susceptible to adversarial opcode perturbations, where small changes in opcode sequences can mislead classifiers into wrongly identifying malicious contracts as benign. To address this, future work should consider adversarial training techniques or anomaly detection methods to improve resilience against evasion attacks. Additionally, model explainability remains a challenge—techniques such as SHAP or LIME could be used to analyze feature importance and identify opcode sequences that contribute most to classification decisions, increasing trust in the system.

Lastly, the study does not evaluate computational efficiency in real-world fog computing environments. While the proposed model achieves high accuracy, practical deployment requires assessing memory usage, processing latency, and scalability for real-time attack detection. Given the resource constraints of fog computing nodes, lightweight models or edge-optimized ML

architectures should be considered to ensure efficient, low-latency security monitoring in decentralized networks.

Table 4. Comparison with Existing studies

Reference	Vulnerabilities Addressed	Techniques Used	Key Contributions
[24]	Arbitrary memory access, integer underflow/overflow, transaction dependency	BERT-ATT-BiLSTM (pre-trained language model)	Enhances accuracy across multiple datasets, outperforming earlier methods that struggle with diverse contract designs.
[25]	Transaction dependency, arbitrary memory access, block dependency, assertion failure, integer underflow, ether block, integer overflow	LSTM, Support Vector Machine (SVM)	Effectively detects vulnerabilities in smart contracts.
[20]	Complex vulnerabilities	Conventional techniques	Improves blockchain application reliability and enables rapid vulnerability identification.
[26]	Ponzi and non-Ponzi attacks	Random Forest	Uses binary-labeled data for detecting smart contract attacks.
[27]	Ponzi and non-Ponzi attacks	AdaBoost Classifier	Detects fraudulent contracts to enhance smart contract security.
[28]	Integer overflow, timestamp, integer underflow, reentrancy, call stack depth, transaction order dependency (TOD)	Naïve Bayes	Speeds up weak contract identification, addressing challenges in analyzing large-scale smart contracts.
Our	Ponzi, Phish Hack, Gambling,	LR, DT, RF, XGB, ETC, GB, KNN, NB, BC, and LGBM.	Detect the different attacks of smart contracts using the machine learning classifiers

Conclusion:

Integrating machine learning with smart contracts can significantly enhance attack detection and prevention in fog computing systems. This integration strengthens security, reducing the risk of data breaches and other threats. Machine learning algorithms, trained on large datasets, can identify behavioral patterns and detect anomalies. This research utilized a binary-labeled dataset to detect attacks in smart contracts. The Word2Vec and BERT techniques were applied for opcode feature extraction, followed by the implementation of various machine learning classifiers. The results show that Random Forest (RF) and Extra Trees Classifier (ETC) achieved the highest accuracy (0.91) using Word2Vec, while LightGBM (LGBM) reached 0.90 accuracy with BERT. Other classifiers also performed well in attack detection. This approach helps mitigate attacks, minimizing data loss and system downtime. Future research could explore hybrid techniques to further improve detection accuracy and address emerging security challenges in smart contracts within fog computing environments.

Acknowledgement. Sincere gratitude is expressed to Dr. Muhammad Usman Sana for invaluable guidance and support throughout this research, which greatly contributed to the completion of this paper.

Author's Contribution. Tahmina Ehsan led the research and experiments. Muhammad Usman Sana supervised and reviewed the work. Tayybah Kiren, Alvena Ehsan, Mustabeen Aziz, and Fateha Minahil contributed to data analysis, literature review, and manuscript preparation.

Conflict of interest. The authors declare no conflict of interest regarding this publication.

References:

[1] W. M. Anwer, S. M. Khan, M. U. Farooq, "Attack Detection in IoT using Machine Learning," *Eng. Technol. Appl. Sci. Res.*, vol. 11, no. 3, pp. 7273–7278, 2021, doi: https://doi.org/10.48084/etasr.4202.

[2] M. A. A.-F. Zain Ashi, Mohammad Al-Fawa'reh, "Fog computing: security challenges and countermeasures," *Int. J. Comput. Appl. Technol.*, vol. 175, no. 15, pp. 30–36, 2020, doi: 10.5120/ijca2020920648.

[3] Y. I. Alzoubi, V. H. Osmanaj, A. Jaradat, and A. Al-Ahmad, "Fog computing security and privacy for the Internet of Thing applications: State-of-the-art," *Secur. Priv.*, vol. 4, no. 2, p. e145, Mar. 2021, doi: 10.1002/SPY2.145.

[4] A. Ehsan et al., "Enhanced Anomaly Detection in Ethereum: Unveiling and Classifying Threats With Machine Learning," *IEEE Access*, vol. 12, pp. 176440–176456, 2024, doi: 10.1109/ACCESS.2024.3504300.

[5] A. G. & A. M. Yehia Ibrahim Alzoubi, "A systematic review of the purposes of Blockchain and fog computing integration: classification and open issues," *J. Cloud Comput.*, vol. 11, no. 80, 2022, doi: https://doi.org/10.1186/s13677-022-00353-y.

[6] V. C. M. L. Yao Du, Zehua Wang, "Blockchain-Enabled Edge Intelligence for IoT: Background, Emerging Trends and Open Issues," *Futur. Internet*, vol. 13, no. 2, p. 48, 2021, doi: https://doi.org/10.3390/fi13020048.

[7] T. Hewa, A. Braeken, M. Liyanage, and M. Ylianttila, "Fog Computing and Blockchain-Based Security Service Architecture for 5G Industrial IoT-Enabled Cloud Manufacturing," *IEEE Trans. Ind. Informatics*, vol. 18, no. 10, pp. 7174–7185, Oct. 2022, doi: 10.1109/TII.2022.3140792.

[8] S. K. Dwivedi, R. Amin, and S. Vollala, "Smart contract and IPFS-based trustworthy secure data storage and device authentication scheme in fog computing environment," *Peer-to-Peer Netw. Appl.*, vol. 16, no. 1, pp. 1–21, Jan. 2023, doi: 10.1007/S12083-022-01376-7/METRICS.

[9] H. A. S. F. Alaba, "Smart Contracts Security Application and Challenges: A Review," *Cloud Comput. Data Sci.*, 2023, doi: 10.37256/ccds.5120233271.

[10] R. Gupta, M. M. Patel, and S. T. Shukla, Arpit, "Deep learning-based malicious smart contract detection scheme for internet of things environment," *Comput. Electr. Eng.*, vol. 97, p. 107583, 2022, doi: https://doi.org/10.1016/j.compeleceng.2021.107583.

[11] T. Ehsan et al, "Securing Smart Contracts in Fog Computing: Machine Learning-Based Attack Detection for Registration and Resource Access Granting," *IEEE Access*, vol. 12, pp. 42802–42815, 2024, doi: 10.1109/ACCESS.2024.3378736.

[12] et al Liu C. H., Sun J., Ni W., "Blockchain-enabled secure fog platform: Issues, challenges and solutions," *IEEE Commun. Surv. Tutorials*, vol. 22, no. 4, pp. 2488–2521, 2020.

[13] K. D. Otuekong Umoren, Raman Singh, Zeeshan Pervez, "Securing Fog Computing with a Decentralised User Authentication Approach Based on Blockchain," *Sensors (Basel)*, vol. 22, no. 10, p. 3956, 2022, doi: 10.3390/s22103956.

[14] M. Soud, I. Qasse, G. Liebel, and M. Hamdaqa, "AutoMESC: Automatic Framework for Mining and Classifying Ethereum Smart Contract Vulnerabilities and Their Fixes," *Proc. - 2023 49th Euromicro Conf. Softw. Eng. Adv. Appl. SEAA 2023*, pp. 410–417, 2023, doi: 10.1109/SEAA60479.2023.00068.

[15] V. F. Derek Liu, Francesco Piccoli, "Machine Learning Approach to Identify Malicious Smart Contract Opcodes: A Preliminary Study," *JPS Conf. Proc.*, 2024, [Online]. Available: https://journals.jps.jp/doi/10.7566/JPSCP.43.011002

[16] Q. Q. Ali Aljofey, Abdur Rasool, Qingshan Jiang, "A Feature-Based Robust Method for Abnormal Contracts Detection in Ethereum Blockchain," *Electronics*, vol. 11, no. 18, p. 2937, 2022, doi: https://doi.org/10.3390/electronics11182937.

[17] et al Wang T., Zhang S., Liu Y., "AI-driven detection of malicious smart contracts in blockchain networks," *Neural Comput. Appl.*, vol. 34, no. 10, pp. 7945–7962, 2022.

[18] P. Kumar, R. Kumar, G. P. Gupta, and R. Tripathi, "A Distributed framework for detecting DDoS attacks in smart contract-based Blockchain-IoT Systems by leveraging Fog computing," *Trans. Emerg. Telecommun. Technol.*, vol. 32, no. 6, p. e4112, Jun. 2021, doi: 10.1002/ETT.4112.

[19] et al Rehman R., Khan N., Uddin F., "Distributed denial-of-service attack detection using machine learning in IoT-enabled smart environments," *IEEE Internet Things J.*, vol. 9, no. 14, pp. 12687–12698, 2022.

[20] S. El Haddouti, M. Khaldoune, M. Ayache, and M. D. Ech-Cherif El Kettani, "Smart contracts auditing and multi-classification using machine learning algorithms: an efficient vulnerability detection in ethereum blockchain," *Computing*, vol. 106, no. 9, pp. 2971–3003, Sep. 2024, doi: 10.1007/S00607-024-01314-W/METRICS.

[21] J. U. Fahad Hossain, Mehedi Hasan Shuvo, "A hybrid machine learning approach for improved ponzi scheme detection using advanced feature engineering," *Int. J. Informatics Commun. Technol.*, vol. 14, no. 1, pp. 50–58, 2025, [Online]. Available: https://ijict.iaescore.com/index.php/IJICT/article/view/21270

[22] Forta Network, "Forta-network/labelled-datasets," GitHub. Accessed: Mar. 27, 2025. [Online]. Available: https://github.com/forta-network/labelled-datasets

[23] R. Buyya and S. N. Srirama, "Fog and edge computing : principles and paradigms," p. 471, 2019.

[24] P. L. Fei He, Fei Li, "Enhancing smart contract security: Leveraging pre-trained language models for advanced vulnerability detection," *IET Blockchain*, 2024, doi: https://doi.org/10.1049/blc2.12072.

[25] Qusai Omar Mustafa Hasan, "Machine Learning Based Framework for Smart Contract Vulnerability Detection," Rochester Institute of Technology. Accessed: Mar. 27, 2025. [Online]. Available: https://books.google.com.pk/books/about/Machine_Learning_Based_Framework_for_Sma.html?id=EvML0AEACAAJ&redir_esc=y

[26] S. Ji, C. Huang, P. Zhang, H. Dong, and Y. Xiao, "Ponzi Scheme Detection Based on Control Flow Graph Feature Extraction," *Proc. - 2023 IEEE Int. Conf. Web Serv. ICWS 2023*, pp. 585–594, 2023, doi: 10.1109/ICWS60048.2023.00077.

[27] M. Wang and J. Huang, "Detecting Ethereum Ponzi Schemes Through Opcode Context Analysis and Oversampling-Based AdaBoost Algorithm," *Comput. Syst. Sci. Eng.*, vol. 47, no. 1, pp. 1023–1042, May 2023, doi: 10.32604/CSSE.2023.039569.

[28] T. L. Xueshuo Xie, Haolong Wang, Zhaolong Jian, Yaozheng Fang, Zichun Wang, "Block-gram: Mining knowledgeable features for efficiently smart contract vulnerability detection," *Digit. Commun. Networks*, vol. 11, no. 1, pp. 1–12, 2025, doi: https://doi.org/10.1016/j.dcan.2023.07.009.

RESEARCH & INNOVATION DIVISION
IJIST

Investigate the Operating Temperature Effect on Fast Pyrolysis Products of Food Waste with Hydrogen

Saher Asif, Syed Kamal Zafar*, Muhammad Tahseen Sadiq, Muhammad Hassan Qasim, Mahmboob Ahmad, Nouman Amjad, Mian Muhammad Mugheera
Department of Chemical Engineering, University of Gujrat, Pakistan
*** Correspondence:** kamal.zafar786@uog.edu.pk;

Citation | Asif. S, Zafar. S. K, Sadiq. M. T, Qasim. M. H, Ahmad. M, Amjad. N, Mugheera. M. M, "Investigate the Operating Temperature Effect on Fast Pyrolysis Products of Food Waste with Hydrogen", IJIST, Special Issue. pp 91-108, March 2025
Received| Feb 15, 2025 **Revised|** Feb 28, 2025 **Accepted|** March 05, 2025 **Published|** March 08, 2025.

Energy crises and environmental pollution are the main issues of concern all over the world and the disposal of wastes by converting into gaseous products can reduce this to a level. Investigating how operating temperature affects the yield and makeup of bio-oil, bio-char, and bio-gas during the pyrolysis process in the presence of hydrogen is the goal of this study. By offering a novel method for enhancing the quality and yield of gaseous products through controlled thermal decomposition in a hydrogen-enriched environment, the findings improve sustainable technologies. In this research, the fast pyrolysis of food waste carried out by using a lab scale fixed bed reactor in the presence of different composition of Nitrogen and Hydrogen to investigate the effect of operating parameters high pyrolysis temperature 600, 650, 700, 750 and 800 °C and hydrogen gas 0 %, 10 % and 20 % with Nitrogen as a carrier gas. The gaseous products maximum yield i.e. 45.68 comes out at 750 °C temperature in the presence of 10 % hydrogen. The results indicate that increasing the pyrolysis temperature boosts decomposition reactions, encouraging the formation of gaseous products. Hydrogen plays a crucial role by facilitating cracking and stabilizing the reaction intermediates, minimizing the formation of heavier components. The results demonstrate that the fast pyrolysis of food waste give residue at high temperature and in the presence of hydrogen up to 10 % achieved a maximum the bio gas yield. Energy crises and environmental pollution are major global concerns. Converting waste into gaseous products can help address these issues. This study examines how operating temperature influences the yield and composition of bio-oil, bio-char, and bio-gas during pyrolysis in a hydrogen-rich environment. By introducing a novel approach to enhance the quality and yield of gaseous products through controlled thermal decomposition, the findings contribute to sustainable technologies. The research involves fast pyrolysis of food waste using a lab-scale fixed-bed reactor, with varying nitrogen and hydrogen compositions. The effects of different operating parameters were analyzed, including high pyrolysis temperatures (600, 650, 700, 750, and 800 °C) and hydrogen concentrations (0%, 10%, and 20%), with nitrogen as the carrier gas. The highest gas yield (45.68%) was achieved at 750 °C in the presence of 10% hydrogen. The results show that increasing pyrolysis temperature enhances decomposition reactions, leading to higher gas production. Hydrogen plays a key role by promoting cracking reactions and stabilizing reaction intermediates, reducing the formation of heavier byproducts. The study demonstrates that fast pyrolysis of food waste at high temperatures, with up to 10% hydrogen, results in the highest bio-gas yield.

Keywords: Food Waste, Fast Pyrolysis, Hydrogen, Gaseous Products, Yield.

Introduction:

The major issue of food waste has become a worldwide problem due to its environmental concerns, social implications and economic issues. According to the Food and Agriculture Organization (FAO) it has been observed that around one by third of all produced food from consumption of humans, annually amount of wastage is approximately 1.3 billion tons. This significant amount of wastage is not only showing a substantial loss of resources like energy, food land, and sea water, but it can cause various dangerous environmental problems by contributing in it [1]. Food waste is compositionally unique, containing higher organic compounds like carbohydrates and proteins to produce gaseous products by fast pyrolysis in the presence of hydrogen to prevent waste management and promote renewable energy. By addressing all these issues, it has been very developing interest in reconnoitering the such management plans for sustainable wastage of all organic things, that it should be converted the all waste of food into valuable products, consequently it can promote a globular economy and can reduced the environmental issues [2].

One of the fastest approaches of this fact is the conversion of food waste thermochemical through fast pyrolysis, but particularly with main focus on maximizing the production of biogas by it. Pyrolysis is a thermochemical process which involves the thermal decomposition of organic materials at high temperatures in the absence of oxygen [3]. The results of this process is in the production of many products such as bio-oil, bio char, and gaseous products. The large dependency of products distributions is on the operating conditions. Among the various types of pyrolysis, the fast pyrolysis is distinguished from all its various types due to its rapid reaction rates and short time of vapor residence. So ultimately it optimized to enhance the yield of liquid bio-oil [4]. However, in recent years, the demand of biogas is increase day by day, so it has been interesting in increment the production of gaseous products, which is the huge source renewable energy [5]. Therefore, to optimize the operating parameters of fast pyrolysis to enhance the production of biogas is the basic factor of research, by utilizing food waste as a feedstock. To obtain the efficiency of the pyrolysis process and the yield of required products, such as gaseous products, are significantly partial by different operating parameters, such as operating temperature, composition of feedstock, and the existence of reactive gases [6]. The top rich factor is temperature, which is known to play a critical role in the pyrolysis reaction process, so it can cause a huge affecting aspect in the distribution of bio-char, bio-oil, and gaseous products [7]. It has been studying that higher pyrolysis temperatures generally favor to the formation of bio-oil and gaseous products, while lower temperatures cause the higher rate the production of bio-char.

But the bio-oil production yield decreases after increasing the temperature than that optimum temperature. For example, at temperatures between 600 °C and 800 °C, the gas production rate is increased in higher peak, by enhance the yields of methane and other hydrocarbons. For fast pyrolysis of wastage of food, the temperature range is particularly relevant with it. It allows the thermal cracking of complex organic compounds, to the release of volatile gases such as biogas as compared to other products. During pyrolysis process, the temperature rises and the introduction of hydrogen, it has been examined to better the quality and yield of the making products. By hydrogen adding during pyrolysis process, hydrogen-involved or hydro pyrolysis, can improve the making of gaseous products by promotes the hydrogenation and cracking reactions [8]. The involvement of hydrogen is introduced to provide the large molecular structures, thereby the making of lighter hydrocarbons and biogas products. Moreover, by the addition of hydrogen in the elimination of oxygen from the organic compounds that involves in food waste, leads to result in rising the calorific value of gaseous components.

Except the potential benefits of hydrogen- assisted pyrolysis, there is a minimum research

on food waste application, specifically with biogas making products. The advanced study of this system that involves effective combination of temperature and the concentration of hydrogen on the food pyrolysis waste, with a main focus on the yield of biogas. The experimental process that includes how to conduct pyrolysis reactions at different temperatures and hydrogen concentrations. So, this experiment involves the conducting pyrolysis reactions at various temperatures (600 °C, 650 °C, 700 °C, 750 °C and 800 °C) and by using different hydrogen volume concentrations (0 %, 10 %, and 20 %). By finding these key parameters, the study focused on find the impact on pyrolysis products distribution especially with the gaseous products formation. The method of incrementally rising the hydrogen concentration gives a deep knowledge of how hydrogen affects the reactions of pyrolysis, especially in improving the yield of methane and other combustible gases [6]. Studies and research on pyrolysis has mainly focused on production of biomass, such as wood and wastage of agricultural residues, and significantly different in composition from food waste. Food waste is produced by a higher moisture content with using complex ingredients such as carbohydrates, fats, proteins, and various inorganic compounds [9].

The differences between composition results to form in separate behaviors of thermal degradation, for enhance the required approach to improve the pyrolysis process for food waste. To remove of all diverse compounds, the adding of hydrogen during the process of pyrolysis having the influence effect, which can cause potentially prominent to enhance the yields of respected gaseous products. For example, it has been observed that the assisted based hydrogen pyrolysis could be meaningfully enhance the yield of methane by breaking down the bonds of C–O and C–C [10]. Consequently, it established that the presence of hydrogen can change the product circulation of hydrocarbons, to enlightening the overall energy of the pyrolysis gases by pyrolysis process.

The motivation for this study is boosted to develop efficient energy from waste by using various technologies that can convert food waste into renewable energy sources, and improve the environment. To increase biogas production, it is optimizing the fast pyrolysis process. This research aims to eliminate the environmental impacts of food waste removal, so we can use some solutions including development of sustainable energy and reduce dependence on fossil fuels [11]. The results of this study have significant applications for the proposal of pyrolysis systems, particularly by adding the hydrogen-based processes, which are gaining more attention to upgrade the quality of pyrolysis process. Moreover, by understanding the higher effects of temperature and hydrogen concentration on food waste pyrolysis which are providing valuable impacts for enhance the technology and mixing it into current waste systems [12]. This study not only increase the impact of different temperatures and hydrogen concentrations on the pyrolysis of food waste, but it also has main focus on production of biogas as the main valuable and interesting product. The experimental results having main focus to demonstrate the increment of thermal decomposition of food waste by using hydrogen in the process to obtain the higher yield of gaseous product as methane-rich product [13]. To increase the circular economy and to promote sustainable waste management, some focused practices and efforts are required to fulfil this target by converted the waste materials into valuable products. By understanding of these operating parameters that can cause the biogas production from food waste, so this research could the main approach for the development of optimized pyrolysis process system for producing clean and renewable energy from organic waste compounds [14].

The issue of food waste has become a global concern due to its environmental, social, and economic implications. According to the Food and Agriculture Organization (FAO), approximately one-third of all food produced for human consumption is wasted annually, amounting to around 1.3 billion tons. This significant wastage not only results in a substantial loss of resources such as energy, land, and water but also contributes to various environmental

problems [1]. Food waste is compositionally unique, containing high amounts of organic compounds like carbohydrates and proteins. These compounds can be converted into gaseous products through fast pyrolysis in the presence of hydrogen, providing an effective solution for waste management and renewable energy production. Addressing these issues has sparked growing interest in sustainable waste management strategies, particularly in converting food waste into valuable products. This approach promotes a circular economy and reduces environmental impact [2].

One of the most efficient methods for converting food waste is thermochemical conversion through fast pyrolysis, with a primary focus on maximizing biogas production. Pyrolysis is a thermochemical process that decomposes organic materials at high temperatures in the absence of oxygen [3]. This process yields bio-oil, biochar, and gaseous products, with product distribution largely dependent on operating conditions. Among different pyrolysis techniques, fast pyrolysis is particularly notable for its rapid reaction rates and short vapor residence times, optimizing the yield of liquid bio-oil [4]. However, in recent years, the demand for biogas has increased significantly, leading to a growing interest in enhancing the production of gaseous products, a major source of renewable energy [5]. Thus, optimizing the operating parameters of fast pyrolysis to enhance biogas production is a key research focus, utilizing food waste as feedstock. The efficiency of the pyrolysis process and the yield of gaseous products are significantly influenced by parameters such as operating temperature, feedstock composition, and the presence of reactive gases [6].

Temperature plays a critical role in pyrolysis, significantly affecting the distribution of biochar, bio-oil, and gaseous products [7]. Studies have shown that higher pyrolysis temperatures generally favor the formation of bio-oil and gaseous products, while lower temperatures lead to increased biochar production. However, beyond an optimal temperature, bio-oil yields tend to decrease. For instance, at temperatures between 600°C and 800°C, gas production peaks, enhancing yields of methane and other hydrocarbons. In food waste pyrolysis, this temperature range facilitates the thermal cracking of complex organic compounds, leading to the release of volatile gases such as biogas [8]. Introducing hydrogen during pyrolysis has been examined as a method to improve product quality and yield. Hydrogen-assisted pyrolysis, or hydro-pyrolysis, enhances gaseous product formation by promoting hydrogenation and cracking reactions [9]. The addition of hydrogen facilitates the breakdown of large molecular structures, increasing the production of lighter hydrocarbons and biogas. Furthermore, hydrogen helps remove oxygen from organic compounds in food waste, increasing the calorific value of the gaseous components. Despite the potential benefits of hydrogen-assisted pyrolysis, limited research has been conducted on its application to food waste, particularly in biogas production [10].

This study aims to explore the effects of temperature and hydrogen concentration on food waste pyrolysis, focusing on biogas yield optimization. The experimental process involves conducting pyrolysis reactions at different temperatures (600°C, 650°C, 700°C, 750°C, and 800°C) using varying hydrogen concentrations (0%, 10%, and 20%). By analyzing these key parameters, the study investigates their impact on product distribution, particularly gaseous products. Incrementally increasing hydrogen concentration provides insight into how hydrogen affects pyrolysis reactions, particularly in enhancing methane and other combustible gas yields [11]. Most previous studies on pyrolysis have focused on biomass sources such as wood and agricultural residues, which differ significantly in composition from food waste. Food waste typically has higher moisture content and complex components, including carbohydrates, fats, proteins, and inorganic compounds [12]. These compositional differences lead to distinct thermal degradation behaviors, necessitating an optimized approach for food waste pyrolysis. The addition of hydrogen during pyrolysis has a notable influence, potentially improving the yield of desired gaseous products. For example, hydrogen-assisted pyrolysis has

been shown to enhance methane yield by breaking C–O and C–C bonds [13].

The motivation for this study is to develop efficient waste-to-energy technologies, converting food waste into renewable energy sources while reducing environmental impact. Optimizing the fast pyrolysis process for increased biogas production contributes to sustainable waste management, reducing dependence on fossil fuels. The findings of this study have significant implications for designing advanced pyrolysis systems, particularly those incorporating hydrogen-assisted processes to enhance biogas production. Additionally, understanding the effects of temperature and hydrogen concentration on food waste pyrolysis provides valuable insights for improving pyrolysis technology and integrating it into existing waste management systems [14]. This research not only examines the impact of different temperatures and hydrogen concentrations on food waste pyrolysis but also focuses on biogas production as the primary valuable output. The experimental results demonstrate the enhanced thermal decomposition of food waste through hydrogen addition, leading to higher methane yields. To promote a circular economy and sustainable waste management, further research and practical applications are required to convert waste materials into valuable products. Understanding these operating parameters can help optimize pyrolysis systems for producing clean and renewable energy from organic waste [15].

Objectives of the Study:

The study aims to contribute to sustainable waste management by converting food waste into valuable energy-dense gaseous products. The specific objectives are as follows:
This study aims to support sustainable waste management by transforming food waste into valuable, energy-rich gaseous products. The specific objectives are as follows

- To analyze the impact of pyrolysis temperature (600–800°C) on the yield and composition of bio-oil, bio-char, and biogas.
- To evaluate the role of hydrogen in enhancing biogas production and improving gas composition, particularly CH_4 and H_2 content.
- To compare the product distribution of hydrogen-assisted pyrolysis with conventional nitrogen-assisted pyrolysis.
- To study how pyrolysis temperature (600–800 °C) effects the yield % biogas.
- To assess the role of hydrogen in increasing biogas production and improving its composition, especially CH_4 and H_2 levels.

To provide insights into the reaction mechanisms involved in hydrogen-assisted pyrolysis, including hydrocracking and gasification pathways. To explain the reaction mechanisms in hydrogen-assisted pyrolysis, including hydrocracking and gasification pathways.

Material and Methods:

Collection and preparation of feed stock:

The food waste samples obtained from the cafeteria of University of Gujrat, main campus for experimentation. Our food waste samples consist of mainly fruits and vegetables peel (potatoes, mango and banana). The food waste was physically sorted after it was collected in order to eliminate any non-organic pollutants like paper or plastic. After that, it was carefully cleaned with distilled water to get rid of any remaining debris. In order to maximize the surface area for effective drying, the bigger pieces were manually broken into tiny fragments using a stainless- steel cutter after washing. After being equally distributed on stainless steel trays, these pieces were allowed to cure for around 10 days in the open air.

The natural loss of moisture throughout the drying process decreased the feedstock's total water content. The samples were periodically turned over during this time to guarantee even drying and stop microbiological development. Once the drying process was complete, the samples were further processed to achieve a uniform particle size. They were ground using

a mechanical grinder and then sieved to obtain particles of consistent size. The prepared feedstock was stored in airtight containers to prevent moisture absorption before further experimentation.

The food waste samples used in this experiment were collected from the cafeteria at the University of Gujrat, main campus. These samples mainly consisted of fruit and vegetable peels, including potatoes, mangoes, and bananas. After collection, the waste was manually sorted to remove any non-organic contaminants such as paper or plastic. It was then thoroughly washed with distilled water to eliminate any remaining debris. To increase the surface area for efficient drying, larger pieces were manually cut into smaller fragments using a stainless-steel cutter. The prepared pieces were evenly spread on stainless steel trays and left to air-dry for approximately 10 days. During this period, the natural moisture loss reduced the overall water content of the feedstock. The samples were regularly turned over to ensure uniform drying and to prevent microbial growth. Once completely dried, the samples were further processed to achieve a consistent particle size. They were ground using a mechanical grinder and then sieved to obtain uniform particles. The final feedstock was stored in airtight containers to prevent moisture absorption before further experimentation.

Feed Characterization:

In accordance with ASTM guidelines, a proximate analysis of the feedstock from food waste was conducted. Table 1 present the proximate analysis. The test was conducted in a 2 gram. A petri dish containing a feed sample was put in a muffle oven set to 105 °C, with air for eight hours. It got dry once the moisture content was eliminated. Moisture contents can be computed from the weight difference. The dry feed sample is placed in a muffle furnace at 575 °C for 25 minutes at air temperature in order to determine the amount of ash present. Five hours later, the weight of the ash was once more measured. The volatile matter was calculated by the ASTM standards. Table 2 present the ultimate analysis.

Following ASTM guidelines, a proximate analysis was conducted on the food waste feedstock. The results are shown in Table 1. The test was performed using a 2-gram sample. A petri dish containing the sample was placed in a muffle oven set at 105°C with airflow for eight hours. The sample was considered dry once all moisture had evaporated. The moisture content was determined by measuring the weight difference before and after drying. To determine the ash content, the dried sample was placed in a muffle furnace at 575°C for 25 minutes in the presence of air. After five hours, the remaining ash was weighed again. The volatile matter was analyzed according to ASTM standards. The results of the ultimate analysis are presented in Table 2.

Table 1. Comparison of proximate Analysis of Food waste with waste residue

Element	**Food waste (wt.%)**	**Corn Cob** [15]	**Sugarcane Bagasse** [8]	**Rice Husk** [16]
Moisture	3.5	12.77	10.4	10.89
Volatile matter	64	2.30	16.4	15.14
Ash Content	6.0	91.16	74.0	73.41
Fixed Carbon	26.5	6.54	13.0	11.44

Table 2. Comparison of Ultimate Analysis of Food waste with waste residue

Element	**Food Waste (wt. %)**	**Corn Cob** [24]	**Sugarcane Bagasse** [23]	**Rice Husk** [25]
Carbon	46	42.10	43.2	41.92
Hydrogen	6.5	5.90	6.70	6.34
Nitrogen	0.48	0.50	0.30	1.85
Sulfur	0.15	0.48	0.20	0.47

The carbon, hydrogen, nitrogen, sulfur, and oxygen contents of the food waste

samples were finally determined using an elemental analyzer. First, a precision balance was used to precisely weigh the samples of dried and crushed food waste. After that, a tiny portion of the material was put into the combustion chamber of the analyzer. The material was broken down into gases inside the chamber by burning it at high temperatures while oxygen was present. Carbon dioxide (CO_2), water vapor (H_2O), nitrogen oxides (NO_x), and sulfur oxides (SO_x) emitted during burning were all measured by the analyzer. The discovered gases were used to compute the concentrations of carbon, hydrogen, nitrogen, and sulfur. By deducting the total proportion of these elements from 100%, the oxygen content was calculated. The findings gave important details about the food waste's elemental makeup, which made it easier to determine if it was suitable for pyrolysis. Here the food waste analysis is compared with the other residues such as corn, sugarcane and rice. The food waste contains more volatile stuff than rice husk, it promotes improved pyrolysis efficiency and speeds up decomposition. It is advantageous for energy applications because of its high hydrogen content, which improves gas production. Its balanced composition guarantees strong conversion potential even if it contains more ash and moisture than some biomass sources. Food waste provides a competitive yield in pyrolysis products when compared to maize cob and sugarcane bagasse, making it a feasible feedstock for the production of sustainable energy.

The carbon, hydrogen, nitrogen, sulfur, and oxygen content of the food waste samples was determined using an elemental analyzer. First, the dried and crushed food waste samples were accurately weighed using a precision balance. A small portion of the sample was then placed into the analyzer's combustion chamber. Inside the chamber, the material was burned at high temperatures in the presence of oxygen, breaking it down into gases. The analyzer measured the emitted gases, including carbon dioxide (CO_2), water vapor (H_2O), nitrogen oxides (NO_x), and sulfur oxides (SO_x). Based on these measurements, the concentrations of carbon, hydrogen, nitrogen, and sulfur were calculated. The oxygen content was determined by subtracting the total percentage of these elements from 100%.

The results provided valuable insights into the elemental composition of food waste, helping assess its suitability for pyrolysis. A comparison was made with other residues such as corn, sugarcane, and rice. Food waste contains more volatile compounds than rice husk, which enhances pyrolysis efficiency and accelerates decomposition. Its high hydrogen content improves gas production, making it beneficial for energy applications. Although food waste has higher ash and moisture content than some biomass sources, its balanced composition ensures strong conversion potential. Compared to maize cobs and sugarcane bagasse, food waste offers a competitive yield in pyrolysis products, making it a viable feedstock for sustainable energy production.

Experimental Methodology:

A lab scale experimental setup is designed to carry out the fast pyrolysis of food waste in the hydrogen atmosphere. The schematic diagram of experimental setup that is shown in Figure 1 consists upon Hydrogen and nitrogen gas cylinders, feeding system, control panel, thermocouples (K type), Pyrolysis reactor and bio-gas collection jar. A lab-scale experimental setup was designed to conduct the fast pyrolysis of food waste in a hydrogen atmosphere. The schematic diagram of the setup, shown in Figure 1, includes hydrogen and nitrogen gas cylinders, a feeding system, a control panel, K-type thermocouples, a pyrolysis reactor, and a biogas collection jar.

The reactor tank's wall is made of stainless steel and can withstand temperatures up to 850 °C, a thickness of 0.2 inches. The tank's head has an inlet for gases and exit for venting gases produced by the pyrolysis reaction, and a 12.5-inch stainless steel capsule where a K type thermocouple can be attached for accurate temperature readings. The thermocouple sensor was positioned so that it would make contact with the reaction zone and provide a

reliable reading. Silicon binder and high temperature rubber rings were used to create a completely airtight seal around the entire reactor. The reactor has two condensers, and instead of using a spiral condenser was chosen because it allows for longer contact time between the fumes or vapors and the condensing air, leading to a higher product yield. The experimental work's methodology procedure is illustrated in Figure 2.

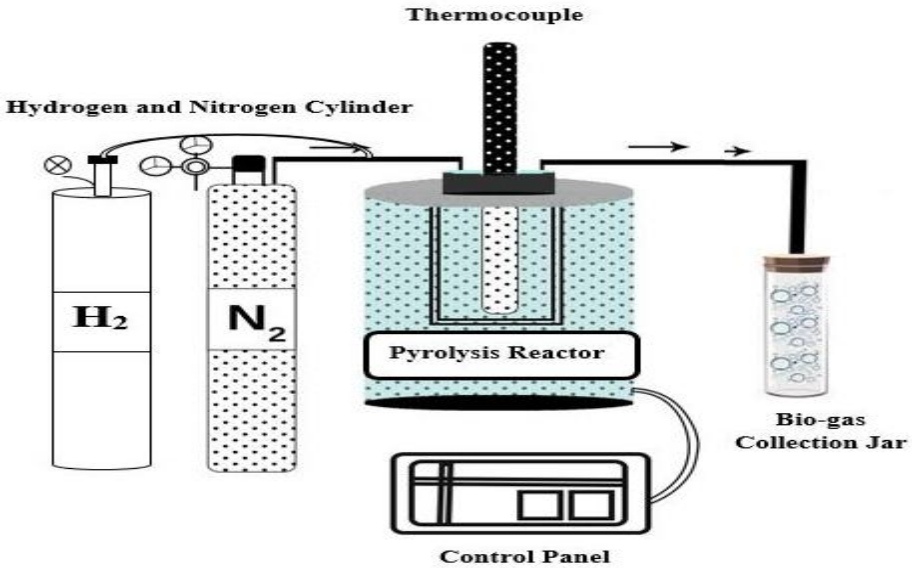


Figure 1. Schematic Diagram of Experimental Setup.

The reactor tank is made of stainless steel with a 0.2-inch thickness and can withstand temperatures up to 850 °C. The tank's head includes an inlet for gas entry, an outlet for venting gases produced during pyrolysis, and a 12.5-inch stainless steel capsule where a K-type thermocouple is attached for precise temperature measurement. The thermocouple sensor is positioned to ensure direct contact with the reaction zone for accurate readings. To maintain an airtight seal, the reactor is secured with a silicon binder and high-temperature rubber rings. It is equipped with two condensers, and a spiral condenser was chosen instead of a standard one, as it increases the contact time between vapors and the cooling air, improving product yield. The methodology for the experimental procedure is illustrated in Figure 2.

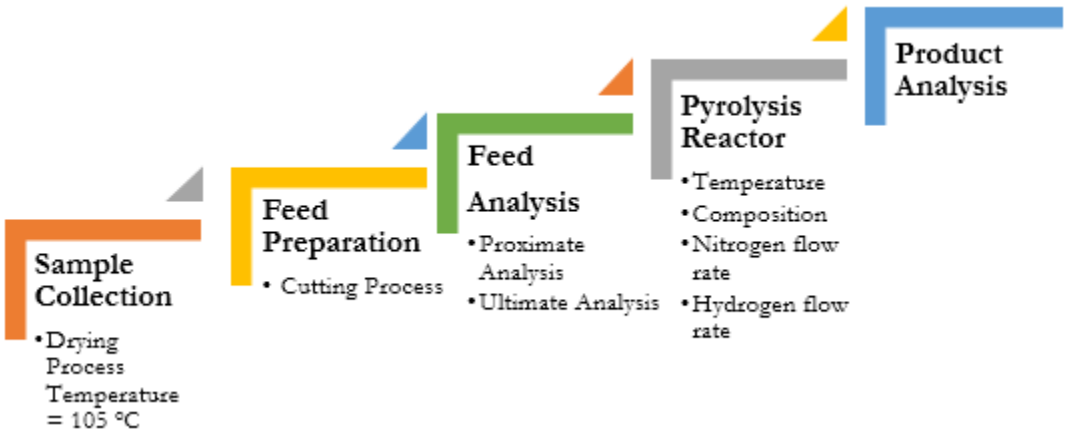


Figure. 2 Flow Diagram of Methodology

At first,15 gram. sample was taken and placed in the reactor's sample cup then the reactor's body was sealed with a silicon ring to prevent the fumes from escaping. The reactor's second outlet is piped to the condenser portion, where vapors are removed by spiral condensers and the inlet is connected to a gases cylinders. The residence time for all the experiments is kept 2 sec. As the process starts, char produced remains on the mesh inside the reactor, the

condensable vapors of bio-oil are collected in the collection system after condensation and the non-condensable gases are collected in the gas balloons.

First, a15gram sample was placed in the reactor's sample cup, and the reactor was sealed with a silicon ring to prevent fumes from escaping. The second outlet of the reactor was connected to the condenser section, where vapors were condensed using spiral condensers, while the inlet was linked to gas cylinders. The residence time for all experiments was maintained at 2 seconds. During the process, the produced char remained on the mesh inside the reactor. The condensable vapors of bio-oil were collected in the collection system after condensation, while the non-condensable gases were stored in gas balloons.

Results:

Experimental Run:

The pyrolysis experiments were conducted using a fixed-bed reactor to investigate the influence of operating temperature and gas composition on product distribution. In order to guarantee consistent thermal decomposition of the food waste residue, the reactor was externally heated to target temperatures between 600 °C and 800 °C. The heating rates were precisely controlled. Before every experiment, the system was nitrogen-purged to remove any remaining gases. A range of gas compositions, including a pure nitrogen atmosphere and nitrogen-hydrogen mixes with hydrogen levels of 10% and 20% by volume, were used to pyrolysis the feedstock. The total gas flow rate was adjusted to either 100, 90, or 80 ml/min, ensuring a consistent residence time for the volatiles within the reactor. The bio-char was gathered and weighed at the end of each run, and the bio-oil portion was condensed and measured. Mass balance was used to determine the remaining bio-gas fraction. A thorough analysis of the impact of temperature on the yield distribution of bio-oil, bio-char, and bio-gas showed notable differences in product composition across experimental settings. The effects of adding hydrogen to the pyrolysis environment on secondary processes, thermal cracking, and total product selectivity were also evaluated.

The pyrolysis experiments were carried out using a fixed-bed reactor to study how temperature and gas composition affect product distribution. To ensure uniform thermal decomposition of the food waste residue, the reactor was externally heated to temperatures between 600 °C and 800 °C, with precise control over heating rates. Before each experiment, the system was purged with nitrogen to remove any residual gases.

The feedstock was pyrolyzed under different gas compositions, including a pure nitrogen atmosphere and nitrogen-hydrogen mixtures with hydrogen concentrations of 10% and 20% by volume. The total gas flow rate was set at 100, 90, or 80 ml/min to maintain a consistent residence time for volatile compounds in the reactor. At the end of each experiment, the bio-char was collected and weighed, while the bio-oil fraction was condensed and measured. The remaining bio-gas fraction was determined using mass balance calculations. A detailed analysis of temperature effects on the yields of bio-oil, bio-char, and bio-gas revealed significant variations in product composition under different conditions. Additionally, the impact of hydrogen on secondary reactions, thermal cracking, and overall product selectivity was evaluated.

Table 3 shows the effect of operating parameter (temperature) on the products yield in the presence of only nitrogen. The maximum yield of bio-oil comes on 600 °C after that its yield decreases. As the temperature increases from 600 °C to 800 °C, the gaseous product yield increases from 17.65 wt. % to 45.68 wt. %. Due to increase in temperature, more thermal degradation takes places and gaseous product yield increases. Table 3 shows the results of product yields in the presence of hydrogen gas (10 % and 20 %). The gas product yield increases till 45.68 wt. % under the effect of 10 % hydrogen but decreases to 40.58 wt. % when 20 % hydrogen is supplied.

Table 3. Product Yield of Pyrolysis of food waste residue under composition of Nitrogen and Hydrogen flow, at different Temperature

Test Run	Reactor Temperature (°C)	N_2 Flow Rate (mL/min)	Hydrogen Content (Vol %)	Bio oil (wt. %)	Bio-Char (wt. %)	Bio-Gas (wt. %)
Run # 01	600	100	0%	57.1± 14.04	25.25± 6.91	17.65± 8.97
Run # 02	650	100	0%	52.78± 14.04	22.95± 6.91	24.27± 8.97
Run # 03	700	100	0 %	47.9± 14.04	20.3± 6.91	31.8± 8.97
Run # 04	750	100	0 %	45.6± 14.04	17.5± 6.91	38.9± 8.97
Run # 05	800	100	0 %	40.8± 14.04	15± 6.91	36.5± 8.97
Run # 06	600	90	10 %	49.88± 14.04	29.45± 6.91	20.67± 8.97
Run # 07	650	90	10 %	38.65± 14.04	33.68± 6.91	27.60± 8.97
Run # 08	700	90	10 %	30.15± 14.04	31.87± 6.91	37.98± 8.97
Run # 09	750	90	10 %	22.89± 14.04	31.42± 6.91	45.68± 8.97
Run # 10	800	90	10 %	14.7± 14.04	30.88± 6.91	41.43± 8.97
Run # 11	600	80	20 %	47.95± 14.04	32.7± 6.91	19.35± 8.97
Run # 12	650	80	20 %	36.89± 14.04	37.3± 6.91	25.81± 8.97
Run # 13	700	80	20 %	29.1± 14.04	34.65± 6.91	36.25± 8.97
Run # 14	750	80	20 %	21.65± 14.04	34.56± 6.91	40.58± 8.97
Run # 15	800	80	20 %	13.1± 14.04	34.38± 6.91	38.72± 8.97

Table 3 presents the effect of operating temperature on product yield in a nitrogen-only environment. The highest bio-oil yield is achieved at 600 °C, after which it starts to decline. As the temperature increases from 600 °C to 800 °C, the gaseous product yield rises from 17.65 wt.% to 45.68 wt.% due to enhanced thermal degradation. Table 3 also shows the product yields in the presence of hydrogen gas (10% and 20%). With 10% hydrogen, the gaseous product yield reaches a maximum of 45.68 wt.%. However, when the hydrogen concentration increases to 20%, the yield decreases to 40.58 wt.%.

Pyrolysis product yield:

Based on the mass distribution of the solid, liquid, and gaseous fractions produced by the rapid pyrolysis of food waste in the presence of hydrogen, the yield of pyrolysis products was calculated. The findings showed that the operating temperature had a major impact on the product dispersion. At an ideal temperature of 750°C, the pyrolysis process produced 22.89% biochar, 31.42% bio-oil, and 45.68% gaseous products. The yield of pyrolysis products was determined based on the mass distribution of solid, liquid, and gaseous fractions generated during the rapid pyrolysis of food waste in a hydrogen environment. The results indicated that operating temperature significantly influenced product distribution. At an optimal temperature of 750°C, the process yielded 22.89% biochar, 31.42% bio-oil, and 45.68% gaseous products.

Effect of Operating Parameters on Bio-gas yield %:

Operating parameters such as temperature, feedstock composition, and gas environment (H_2, N_2) all had a substantial impact on the production and composition of biogas generated by fast pyrolysis of food waste in the presence of hydrogen [20]. A larger gas yield, especially methane (CH_4) and hydrogen (H_2), was obtained by increasing the temperature, which boosted thermal cracking reactions. The ideal range was found at 750 °C. Beyond this range, excessive cracking decreased the gas's calorific value by increasing the formation of carbon monoxide (CO) and carbon dioxide (CO_2). Operating factors such as temperature, feedstock composition, and gas environment (H_2, N_2) significantly influenced biogas production and composition during the fast pyrolysis of food waste in a hydrogen atmosphere [20]. Higher temperatures enhanced thermal cracking reactions, leading to increased gas yield, particularly methane (CH_4) and hydrogen (H_2). The optimal temperature was found to be 750°C. Beyond this point, excessive cracking reduced the gas's calorific value by increasing the formation of carbon monoxide (CO) and carbon dioxide (CO_2).

Effect of Composition:

Figure 3 illustrates how the yield percentage changes with varying composition ratios (100:00, 90:10, and 80:20) under five different temperatures: 600 °C, 650 °C, 700 °C, 750 °C, and 800 °C. A clear upward trend is observed with increasing temperatures, indicating that higher temperatures result in higher yields across all composition ratios. In conclusion, the results demonstrate that both temperature and composition ratio are crucial factors in maximizing yield. A composition ratio of Nitrogen and Hydrogen of 90:10 combined with a temperature of 750 °C offers the most efficient outcome. This harmonious interplay between temperature and composition highlights the potential for refining industrial processes to enhance yield performance.

Figure 3 shows how yield percentage varies with different composition ratios (100:00, 90:10, and 80:20) at five temperatures: 600 °C, 650 °C, 700 °C, 750 °C, and 800 °C. The results indicate a clear upward trend, where higher temperatures lead to increased yields across all composition ratios. Overall, the findings highlight that both temperature and composition ratio play a crucial role in maximizing yield. The most efficient outcome is achieved with a nitrogen-to-hydrogen ratio of 90:10 at 750 °C. This balance between temperature and

composition underscores the potential for optimizing industrial processes to improve yield performance.

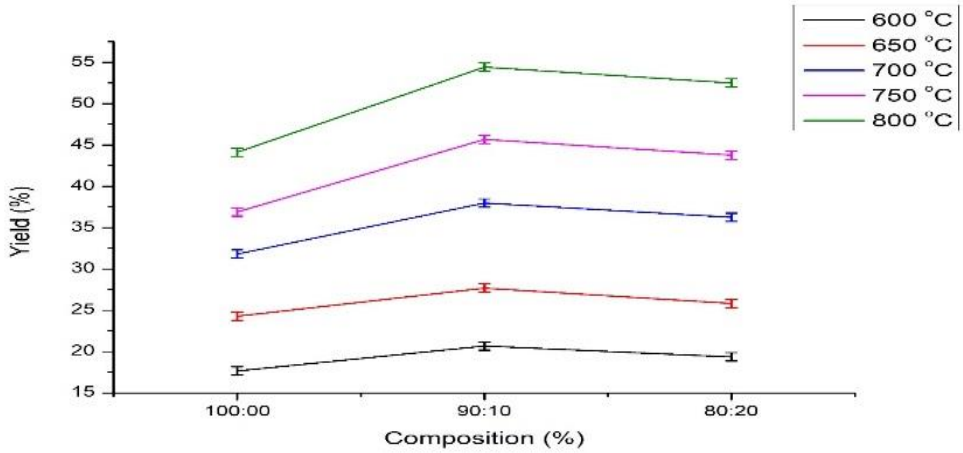


Figure 3. Effect of different composition of H_2 and N_2 in the Yield

Effect of Temperature:

Figure 4 presents the relationship between temperature and yield (wt. %) for three different flow rates: 100 ml/min, 90 ml/min, and 80 ml/min. The trend reveals that as the temperature increases from 600 °C to 700 °C, the yield significantly improves across all flow rates, indicating that higher temperatures enhance the reaction efficiency. However, beyond 700 °C, a decrease in yield is observed, which may suggest thermal degradation or the onset of unfavorable side reactions. The highest yield is achieved at 700 °C for all flow rates, with the 90 ml/min condition (red line) exhibiting the maximum performance.

Figure 4 illustrates the relationship between temperature and yield (wt.%) for three different flow rates: 100 ml/min, 90 ml/min, and 80 ml/min. The results show that increasing the temperature from 600 °C to 700 °C significantly enhances yield across all flow rates, indicating improved reaction efficiency. However, beyond 700 °C, the yield starts to decline, likely due to thermal degradation or unwanted side reactions. The highest yield is observed at 700 °C for all flow rates, with the 90 ml/min condition (red line) showing the best performance.

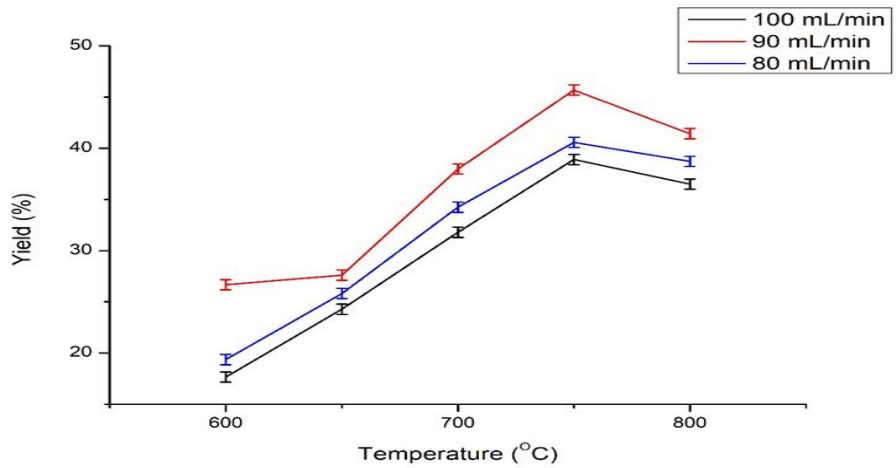


Figure 4. Effect of pyrolysis temperature in the Yield

Effect of Hydrogen Percentage:

Figure 5 illustrates the relationship between hydrogen flow rate (%) and yield (%) at varying temperatures: 600 °C, 650 °C, 700 °C, 750 °C, and 800 °C. The results reveal distinct

trends at each temperature, with yield increasing as the hydrogen flow rate rises from 0 % to 10 %, followed by a decline at 20 %. The observed peak at 10 % indicates an optimal hydrogen flow rate for maximizing yield. At lower temperatures, such as 600 °C and 650 °C, the yield remains relatively low compared to higher temperatures, suggesting that the reaction is less efficient under these conditions. However, as temperature increases to 700 °C, 750 °C, and 800 °C, the yield significantly improves. The highest yield is observed at 750 °C, particularly at the optimal hydrogen flow of 10 %, as indicated by the magenta curve. Beyond 750 °C, the yield slightly decreases, possibly due to competing reactions or thermal instability.

Figure 5 shows the relationship between hydrogen flow rate (%) and yield (%) at different temperatures: 600 °C, 650 °C, 700 °C, 750 °C, and 800 °C. The results indicate a clear trend—yield increases as the hydrogen flow rate rises from 0 % to 10 %, then declines at 20 %. This peak at 10 % suggests an optimal hydrogen flow rate for maximizing yield. At lower temperatures (600 °C and 650 °C), the yield remains relatively low, indicating lower reaction efficiency. However, as the temperature increases to 700 °C, 750 °C, and 800 °C, the yield improves significantly. The highest yield is recorded at 750 °C, particularly at the 10% hydrogen flow rate, as shown by the magenta curve. Beyond 750 °C, the yield slightly decreases, likely due to competing reactions or thermal instability.

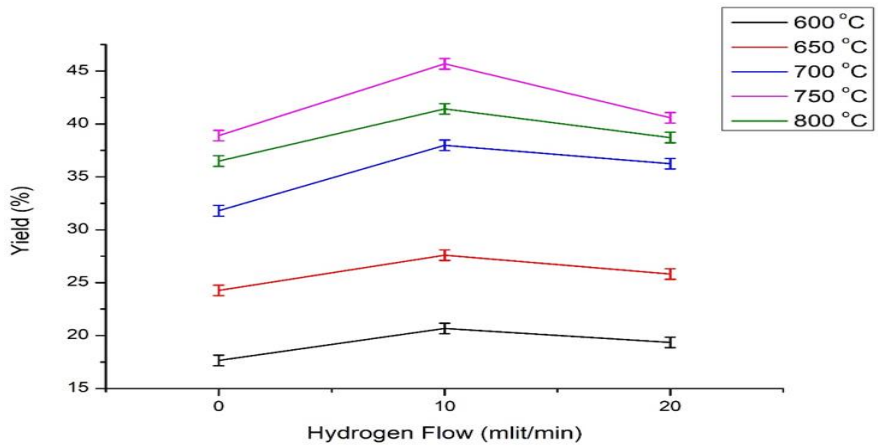


Figure 5. Effect of hydrogen composition in the Yield

Discussion:

The results of the study demonstrate that pyrolyzing food waste with hydrogen results in a significantly greater output of gaseous products, specifically methane (CH_4) and hydrogen (H_2). At 90:10 nitrogen-to-hydrogen ratio at 750 °C, the maximum gas output of 45.68 wt.% was attained. The increase of temperature during fast pyrolysis process not only increase the gaseous product yield but also increases the composition of hydrogen, methane, carbon mono oxide and ethane. The presence of hydrogen facilitates the breaking down the bonding of different heavy molecules. The free radical mechanism reaction rate increases at high temperatures and the composition of CO, H_2 and CH_4 increases but at the same stage the percentage of carbon dioxide, ethane and propene decrease [10].

While the creation of bio-oil and gaseous products was promoted by an increase in temperature, the production of biochar was comparatively higher at lower temperatures. Methane (CH_4), hydrogen (H_2), carbon monoxide (CO), and carbon dioxide (CO_2) made up the majority of the gaseous component, which showed a rising trend at higher temperatures, indicating enhanced thermal cracking and secondary reactions [17].

By raising the concentration of hydrocarbons like methane, the presence of hydrogen in the reaction environment enhanced the gas yield and quality [19]. According to the

findings, hydrogen-assisted pyrolysis is a viable method for environmentally friendly waste-to-energy applications since it improves the conversion of food waste into useful energy products [14].

Food waste's composition also mattered; fractions high in protein and cellulose contributed more to the generation of gas, whilst components rich in lipids and carbohydrates promoted the development of bio-oil. By encouraging hydrocracking reactions, raising methane yield, and decreasing tar formation, the addition of hydrogen greatly enhanced the quality of biogas. In contrast, nitrogen, as an inert carrier gas, changed heat transfer and volatile residence time but did not actively participate in reactions [21] [22]. A lower methane yield and a larger production of CO_2 were noted when nitrogen was substituted for hydrogen, underscoring the crucial role that hydrogen plays in improving gas quality [23]. In hydrogen- assisted pyrolysis, our results highlight the significance of controlling operating conditions to maximize biogas yield and energy efficiency [24].

The composition of food waste plays a crucial role in determining the distribution of pyrolysis products, including bio-oil, bio-char, and biogas. The three major organic components—cellulose, lipids, and proteins—undergo different thermal decomposition pathways, influencing the yield and composition of the final products. Cellulose, a major structural component of plant-based food waste, primarily decomposes into volatile gases, tars, and some residual char. Its thermal degradation occurs between 280–400 °C, favoring bio-oil production at lower temperatures, but at higher temperatures (>700°C), secondary cracking enhances gas yield, particularly CO, CO_2, and H_2. The presence of hydrogen facilitates further gasification of cellulose-derived compounds, improving the yield of hydrogen and methane-rich gases. In food waste rich in cellulose (e.g., vegetable peels, rice husk), the thermal degradation trends indicate that higher pyrolysis temperatures combined with hydrogen enhance gaseous product formation, aligning with the observed peak gas yield at 750 °C in this study.

In contrast, lipids (fats and oils) decompose at lower temperatures (300–500 °C) and contribute significantly to bio-oil production. However, as the temperature increases, thermal cracking of lipids leads to higher gas yields, producing alkanes, alkenes, CH_4, and CO_2. The presence of hydrogen promotes hydrocracking, which breaks down heavier hydrocarbons into lighter gaseous products like methane and ethane. This explains why hydrogen-assisted pyrolysis enhances methane production at 750 °C, particularly in food waste with high lipid content such as dairy products, cooking oils, and meat fats. However, when hydrogen concentration is increased to 20 %, excessive hydrogenation can lead to lower gas yield, as seen in the experimental results, due to the inhibition of further volatile formation.

The study results show that pyrolyzing food waste in a hydrogen atmosphere significantly increases the yield of gaseous products, particularly methane (CH_4) and hydrogen (H_2). The highest gas yield of 45.68 wt.% was achieved at a nitrogen-to-hydrogen ratio of 90:10 at 750 °C. Higher temperatures in fast pyrolysis not only boost gas production but also enhance the composition of hydrogen, methane, carbon monoxide, and ethane. Hydrogen helps break down heavy molecular bonds, accelerating free radical reactions at high temperatures. This leads to increased CO, H_2, and CH_4 formation while reducing CO_2, ethane, and propene content.

While higher temperatures promoted bio-oil and gas production, biochar yield remained higher at lower temperatures. The main gaseous products—methane (CH_4), hydrogen (H_2), carbon monoxide (CO), and carbon dioxide (CO_2)—showed a rising trend at elevated temperatures, indicating increased thermal cracking and secondary reactions. The presence of hydrogen improved gas yield and quality by raising hydrocarbon concentrations,

especially methane. These findings highlight hydrogen-assisted pyrolysis as a promising waste-to-energy method for efficiently converting food waste into valuable energy products.

The composition of food waste also played a key role. Protein- and cellulose-rich fractions contributed more to gas production, while lipid- and carbohydrate-rich components favored bio-oil formation. Hydrogen enhanced hydrocracking, increased methane yield, and reduced tar formation. In contrast, nitrogen, acting as an inert carrier gas, influenced heat transfer and volatile residence time without actively participating in reactions. Replacing hydrogen with nitrogen led to lower methane yields and higher CO_2 production, emphasizing hydrogen's crucial role in improving gas quality. Controlling operating conditions in hydrogen-assisted pyrolysis is essential for optimizing biogas yield and energy efficiency.

Food waste composition directly affects the distribution of pyrolysis products, including bio-oil, biochar, and biogas. The three primary organic components—cellulose, lipids, and proteins—follow different thermal decomposition pathways, influencing the final product yield and composition. Cellulose, a key structural component of plant-based food waste, primarily degrades into volatile gases, tars, and residual char. It decomposes between 280–400°C, favoring bio-oil production at lower temperatures. However, at temperatures above 700 °C, secondary cracking increases gas yield, particularly CO, CO_2, and H_2. The presence of hydrogen enhances the gasification of cellulose-derived compounds, improving hydrogen- and methane-rich gas production. In cellulose-rich food waste (e.g., vegetable peels, rice husks), higher pyrolysis temperatures combined with hydrogen favor gas formation, aligning with the peak gas yield observed at 750 °C.

Lipids (fats and oils) break down at lower temperatures (300–500 °C), mainly contributing to bio-oil formation. However, as the temperature rises, lipid thermal cracking increases gas production, generating alkanes, alkenes, CH_4, and CO_2. Hydrogen promotes hydrocracking, breaking down heavier hydrocarbons into lighter gases like methane and ethane. This explains why hydrogen-assisted pyrolysis enhances methane production at 750 °C, especially in food waste with high lipid content, such as dairy products, cooking oils, and meat fats. However, at 20% hydrogen concentration, excessive hydrogenation can reduce gas yield by inhibiting further volatile formation, as observed in the experimental results.

Proteins, found in significant amounts in food waste sources like meat, legumes, and eggs, decompose at 250–600 °C, forming char, ammonia (NH_3), hydrogen cyanide (HCN), and nitrogen-containing bio-oil compounds. Unlike cellulose and lipids, proteins contribute more to char formation, especially at lower temperatures, due to their complex nitrogenous structures. However, at high temperatures (>700 °C), proteins undergo secondary decomposition, releasing H_2 and CO_2. The introduction of hydrogen in the reaction environment further enhances H_2 production while reducing nitrogenous impurities like NH_3 and HCN, which would otherwise degrade gas quality. This suggests that the increase in hydrogen concentration (10%) optimizes biogas composition, but an excessive hydrogen supply (20%) may limit the thermal degradation of protein-rich feedstocks, reducing the overall gas yield.

Proteins, abundant in food waste sources such as meat, legumes, and eggs, break down at temperatures between 250–600°C, producing char, ammonia (NH_3), hydrogen cyanide (HCN), and nitrogen-containing bio-oil compounds. Unlike cellulose and lipids, proteins generate more char, especially at lower temperatures, due to their complex nitrogenous structures. However, at temperatures above 700 °C, proteins undergo secondary decomposition, releasing H_2 and CO_2. Introducing hydrogen into the reaction environment further boosts H_2 production while reducing nitrogenous impurities like NH_3 and HCN, which can lower gas quality. This indicates that a 10% hydrogen concentration optimizes

biogas composition, but an excessive supply (20%) may hinder the thermal breakdown of protein-rich feedstocks, reducing overall gas yield.

Conclusion:

A fast pyrolysis on Food waste residue was carried out using a fixed bed reactor operating at atmospheric pressure. The Investigation showed that the optimum pyrolysis temperature for obtaining the maximum bio-gas yield at 750 °C. The findings of this investigation indicate that 750 °C. Is the optimum temperature. With a nitrogen to hydrogen ratio of 90:10, a maximum biogas output of 45.68 weight percent is achieved. Bio-gas yields were increased by raising the pyrolysis temperature while keeping the ratio constant. When 90 % nitrogen and 10 % hydrogen were used at 750 °C, the maximum gaseous product yield is obtained. Hydrogen aided in hydrocracking reactions, which decreased the synthesis of tar and heavy hydrocarbons while increasing the production of valuable gases including hydrogen (H_2) and methane (CH_4). The technique became more appropriate for clean energy applications as a result of the enhanced hydrogen-to-carbon ratio and the more energy-dense syngas composition. Temperature differences also had a noticeable impact on the distribution of the product; higher temperatures encouraged the conversion of solid and liquid fractions into non-condensable gases, while lower temperatures favored the creation of biochar because of incomplete thermal decomposition. Hydrogen-assisted pyrolysis is a potential method for managing food waste sustainably and producing renewable energy as these studies offer important insights into how to optimize it for increased syngas production. Process efficiency can be greatly increased, resulting in larger energy yields and a smaller environmental effect, by carefully regulating the temperature and gas composition. Furthermore, the potential for industrial-scale applications particularly in waste-to-energy conversion technologies is highlighted by the ability to precisely control product distribution through exact operating conditions. Future research will involve the development and optimization of bio-based catalysts for the production of high-quality and cost-efficient bio-gas.

A fast pyrolysis process was conducted on food waste residue using a fixed-bed reactor at atmospheric pressure. The investigation revealed that 750 °C is the optimal pyrolysis temperature for achieving the highest biogas yield. With a nitrogen-to-hydrogen ratio of 90:10, a maximum biogas output of 45.68 wt.% was obtained. Increasing the pyrolysis temperature while maintaining this ratio further enhanced biogas yields. At 750 °C, with 90% nitrogen and 10% hydrogen, the highest gaseous product yield was achieved. Hydrogen played a key role in hydrocracking reactions, reducing tar and heavy hydrocarbon formation while boosting the production of valuable gases such as hydrogen (H_2) and methane (CH_4). This improved the hydrogen-to-carbon ratio, resulting in a more energy-dense syngas composition, making the process more suitable for clean energy applications.

Temperature variations also significantly influenced product distribution. Higher temperatures promoted the conversion of solid and liquid fractions into non-condensable gases, whereas lower temperatures favored biochar formation due to incomplete thermal decomposition. These findings highlight hydrogen-assisted pyrolysis as a promising method for sustainable food waste management and renewable energy production. By carefully controlling temperature and gas composition, process efficiency can be significantly improved, leading to higher energy yields and a reduced environmental footprint. Moreover, the ability to fine-tune operating conditions for precise product distribution enhances the potential for industrial-scale applications, particularly in waste-to-energy technologies. Future research will focus on developing and optimizing bio-based catalysts to produce high-quality, cost-effective biogas.

Acknowledgements

The author thanks GC University Faisalabad for the ultimate analysis.

References:

[1] R. H. Venderbosch, B. T. G. Biomass, and T. Group, "Fast pyrolysis technology," no. 1, pp. 178–208, 2010, doi: 10.1002/bbb.

[2] Z. Qi, "Review of biomass pyrolysis oil properties and upgrading research," vol. 48, pp. 87–92, 2007, doi: 10.1016/j.enconman.2006.05.010.

[3] T. Makov, A. Shepon, J. Krones, and M. Chertow, "Social and environmental analysis of food waste abatement via the peer-to-peer sharing economy," *Nat. Commun.*, no. 2020, doi: 10.1038/s41467-020-14899-5.

[4] F. Girotto, L. Alibardi, and R. Cossu, "Food waste generation and industrial uses : A review," *Waste Manag.*, vol. 45, pp. 32–41, 2015, doi: 10.1016/j.wasman.2015.06.008.

[5] J. Jo, S. Kim, J. Shim, and Y. Lee, "Pyrolysis Characteristics and Kinetics of Food Wastes," 2017, doi: 10.3390/en10081191.

[6] M. N. Uddin, K. Techato, J. Taweekun, and M. Rahman, "An Overview of Recent Developments in Biomass Pyrolysis Technologies," 2018, doi: 10.3390/en11113115.

[7] X. Huang, Z. Wang, and A. Ding, "Impact of Aerosol-PBL Interaction on Haze Pollution: Multiyear Observational Evidences in North China," *Geophys. Res. Lett.*, vol. 45, no. 16, pp. 8596–8603, 2018, doi: 10.1029/2018GL079239.

[8] J. Zhang *et al.*, "Carbon science in 2016: Status, challenges and perspectives," *Carbon N. Y.*, vol. 98, pp. 708–732, 2016, doi: 10.1016/j.carbon.2015.11.060.

[9] S. Y. Lee, R. Sankaran, K. W. Chew, C. H. Tan, and R. Krishnamoorthy, "BMC Energy Waste to bioenergy: a review on the recent conversion technologies," pp. 1–22, 2019.

[10] N. Krishnamoorthy, C. Nzediegwu, X. Mao, H. Zeng, B. Paramasivan, and S. X. Chang, "Biochar seeding properties affect struvite crystallization for soil application," *Soil Environ. Heal.*, vol. 1, no. 2, p. 100015, 2023, doi: 10.1016/j.seh.2023.100015.

[11] A. K. Chakraborty *et al.*, "Fabrication of visible-light induced fluorine doped bismuth oxide (Bi2O3-xFx) photocatalyst in degrading textile dyes from wastewater," *Chem. Inorg. Mater.*, vol. 1, no. September, p. 100019, 2023, doi: 10.1016/j.cinorg.2023.100019.

[12] M. A. Naeem, "Food waste : causes and economic losses estimation at household level in Pakistan," pp. 1–17, 2022.

[13] Y. Song *et al.*, "Multifunctional bismuth oxychloride/mesoporous silica composites for photocatalysis, antibacterial test, and simultaneous stripping analysis of heavy metals," *ACS Omega*, vol. 3, no. 1, pp. 973–981, 2018, doi: 10.1021/acsomega.7b01590.

[14] A. V Bridgwater, "Review of fast pyrolysis of biomass and product upgrading," *Biomass and Bioenergy*, vol. 38, pp. 68–94, 2011, doi: 10.1016/j.biombioe.2011.01.048.

[15] B. Biswas, N. Pandey, Y. Bisht, R. Singh, J. Kumar, and T. Bhaskar, "Pyrolysis of agricultural biomass residues: Comparative study of corn cob, wheat straw, rice straw and rice husk," *Bioresour. Technol.*, vol. 237, pp. 57–63, 2017, doi: 10.1016/j.biortech.2017.02.046.

[16] K. M. Isa, S. Daud, N. Hamidin, K. Ismail, S. A. Saad, and F. H. Kasim, "Thermogravimetric analysis and the optimisation of bio-oil yield from fixed-bed pyrolysis of rice husk using response surface methodology (RSM)," *Ind. Crops Prod.*, vol. 33, no. 2, pp. 481–487, 2011, doi: 10.1016/j.indcrop.2010.10.024.

[17] A. S. Paul, N. L. Panwar, B. L. Salvi, S. Jain, and D. Sharma, "Experimental investigation on the production of bio-oil from wheat straw," *Energy Sources, Part A Recover. Util. Environ. Eff.*, vol. 00, no. 00, pp. 1–16, 2020, doi: 10.1080/15567036.2020.1779416.

[18] I. I. Ahmed and A. K. Gupta, "Pyrolysis and gasification of food waste: Syngas characteristics and char gasification kinetics," *Appl. Energy*, vol. 87, no. 1, pp. 101–108, 2010, doi:

10.1016/j.apenergy.2009.08.032.
[19] G. Bensidhom, M. Sghairoun, K. Alper, and I. Trabelsi, "Fruit stalks pruning," pp. 1561–1563, 2018.
[20] D. Chen, L. Yin, H. Wang, and P. He, "Pyrolysis technologies for municipal solid waste : A review," *Waste Manag.*, vol. 34, no. 12, pp. 2466–2486, 2014, doi: 10.1016/j.wasman.2014.08.004.
[21] B. Grycová, I. Koutník, and A. Pryszcz, "Pyrolysis process for the treatment of food waste," *Bioresour. Technol.*, vol. 218, pp. 1203–1207, 2016, doi: 10.1016/j.biortech.2016.07.064.
[22] B. R. Patra, S. Nanda, A. K. Dalai, and V. Meda, "Slow pyrolysis of agro-food wastes and physicochemical characterization of biofuel products," *Chemosphere*, vol. 285, no. May, p. 131431, 2021, doi: 10.1016/j.chemosphere.2021.131431.
[23] M. F. Demirbas, M. Balat, and H. Balat, "Potential contribution of biomass to the sustainable energy development," *Energy Convers. Manag.*, vol. 50, no. 7, pp. 1746–1760, 2009, doi: 10.1016/j.enconman.2009.03.013.
[24] S. Pradhan, B. Yuzer, Y. Bicer, G. McKay, and T. Al-Ansari, "Hydrogen gas and biochar production from kitchen food waste through dark fermentation and pyrolysis," *Front. Chem. Eng.*, vol. 6, no. November, pp. 1–10, 2024, doi: 10.3389/fceng.2024.1450151.

Bio fusion: Advancing Biometric Authentication by Fusion of Physiological Signals

Tuba Alvi[1], Yumna Aziz[1], Muhammad Faraz[1]*, Zubair Mehmood[2], Syed Zohaib Hassan Naqvi[1], Laraib Imtiaz[1]

[1]Department of Electronics Engineering, University of Engineering and Technology, Taxila, Pakistan.21-ENC-4@students.uettaxila.edu.pk,21-ENC-13@students.uettaxila.edu.pk, muhammad.faraz@uettaxila.edu.pk,zohaib.naqvi@uettaxila.edu.pk,21-ENC-3@students.uettaxila.edu.pk

[2]Micro Signal Circuit Design Lab, Department of Electrical Engineering, University of Gujrat, Pakistan zubair.mehmood@uog.edu.pk

*** Correspondence:** muhammad.faraz@uettaxila.edu.pk

Citation | Alvi. T, Aziz. Y, Faraz. M, Mehmood. Z, Naqvi. S. Z. H, Imtiaz. L, "Bio fusion: Advancing Biometric Authentication by Fusion of Physiological Signals", IJIST, Special Issue. pp 109-127, March 2025

Received| Feb 16, 2025 **Revised|** Feb 29, 2025 **Accepted|** March 06, 2025 **Published|** March 09, 2025.

Biometric authentication is becoming more popular due to its secure and reliable way of identifying individuals, offering clear advantages over traditional methods. Since physiological signals are unique and non-invasive, they have been widely researched for use in biometric systems. This study introduces a biometric identification system that combines machine learning with physiological signal fusion, using data from electromyography (EMG), phonocardiogram (PCG), and electrocardiogram (ECG). The data were collected from 32 participants using the BIOPAC MP-36 system. To remove power line interference and extract important frequency bands, Butterworth notch, and bandpass filters were applied to the raw signals. After pre-processing, two types of cepstral features were extracted: gamma tone cepstral coefficients (GTCCs) and Mel-frequency cepstral coefficients (MFCCs), which were analysed for their spectral properties. System performance was first tested by evaluating features from each signal individually. Then, the study examined the impact of combining pairs of signals— (ECG, PCG), (PCG, EMG), and (ECG, EMG)— using GTCC and MFCC features with different machine learning classifiers. Lastly, the GTCC and MFCC features from all three signals were combined to evaluate overall system performance. The results showed that MFCC-based features performed better than GTCC-based features for biometric authentication. The highest accuracy, 98.4%, was achieved using GTCC features with both the Fine K-nearest neighbour (KNN) and linear discriminant classifiers, while MFCC features reached 100% accuracy with the linear discriminant classifier. These findings highlight how effective cepstral features and signal fusion can be in enhancing biometric authentication performance.

Keywords: Person Identification; Biometric Authentication; Machine Learning; Physiological Signals; MEL Frequency Cepstral Coefficient.

Introduction:

In today's world, protecting personal identity and information is essential due to the risk of misuse from technological advancements. Biometric identification has become a trusted method and is widely used in healthcare, law enforcement, banking, and the military. This technology identifies individuals using unique traits such as voice patterns, facial features, and fingerprints. In the past, people recognized each other based on characteristics like speech, smell, behavior, facial appearance, and height, but most of these traits are unsuitable for automated systems. However, recent developments in biometric technology have expanded the possibilities for more secure identification processes [1].

Biometrics identifies people based on their distinct physical characteristics. Various biometric techniques have been developed, including face recognition, fingerprint scanning, iris detection, voice analysis, typing patterns, and gait recognition. However, these traditional methods can sometimes be vulnerable to duplication and fraud [2].

Recently, biometric authentication systems using electrocardiogram (ECG), phonocardiogram (PCG), and electromyography (EMG) signals have gained significant attention [3]. ECG signals are particularly popular for biometric recognition because of their unique features, which make them difficult to replicate. ECG signals are present in all living beings and consist of several key components: the T wave (representing ventricular repolarization), the P wave (atrial depolarization), the QRS complex (ventricular depolarization), and the U wave (linked to the repolarization of the heart's conduction fibers). These distinct patterns and timing allow people to be identified through ECG signals [4].

Similarly, PCG is the recording of heart sounds produced during the cardiac cycle. This physiological property captures heartbeats using a digital stethoscope and reflects sounds caused by the opening and closing of heart valves. The two primary heart sounds, S1 and S2 (also called systolic and diastolic murmurs) form the cardiac cycle. These heart sounds are complex, non-stationary, and quasi-periodic signals [5].

ECG and PCG signals remain stable over time, making them reliable for long-term biometric authentication. Unlike face and fingerprint biometrics, which can change due to aging or external factors, physiological signals provide consistent features. Preprocessing these signals improves their quality, enhancing feature extraction and making the biometric system more stable.

Similarly, EMG records electrical signals generated by muscle contractions during neuromuscular activity. These signals are useful for various applications, including motion detection, disease diagnosis, and personal identification [6]. A visual representation of ECG, PCG, and EMG signals is shown in Figure 1.

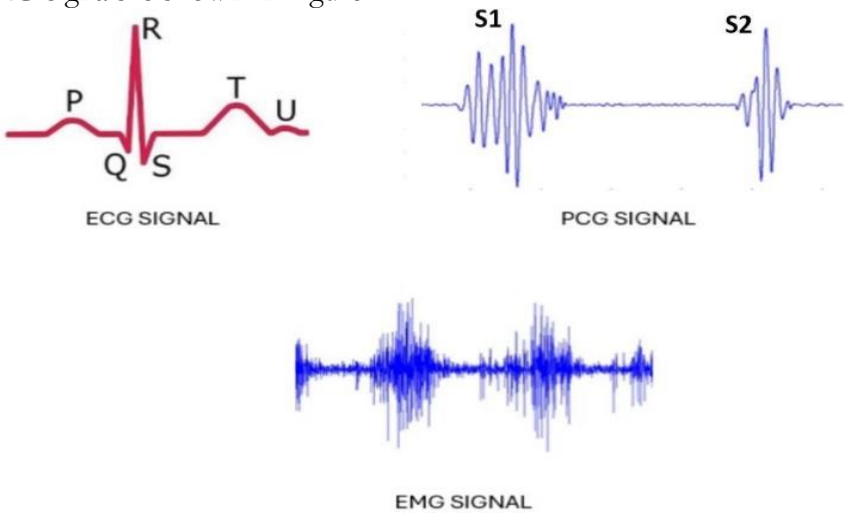


Figure 1. Visual representation of ECG, PCG, and EMG physiological signals

Objectives of the Study:

The primary objective of this study is to develop and evaluate a multimodal biometric authentication system that integrates physiological signals (ECG, PCG, and EMG) to enhance security and accuracy in biometric identification. By leveraging cepstral feature extraction techniques (MFCC, GTCC) and machine learning classifiers, the study aims to improve authentication reliability while addressing challenges related to spoofing, inter-class variability, and real-time usability. The specific objectives are:

- To design a multimodal biometric authentication framework that fuses ECG, PCG, and EMG signals for improved recognition accuracy.
- To extract robust cepstral features (MFCC and GTCC) from physiological signals for enhanced classification performance.
- To compare the performance of various machine learning classifiers (Fine KNN, LDA, SVM, Ensemble Bagged Trees, etc.) to determine the most effective model for biometric authentication.
- To evaluate the system's robustness against noise and real-world variations in physiological signals.
- To explore the feasibility of implementing the proposed biometric system in real-time authentication scenarios, such as wearable security applications.

Novelty Statement:

This study presents a novel multimodal biometric authentication approach by integrating ECG, PCG, and EMG signals, a combination rarely explored in existing biometric systems. Unlike conventional authentication methods that rely on single-modal features (e.g., fingerprint, face recognition, ECG alone), this research introduces a fusion-based framework that enhances security, resilience against spoofing attacks, and user-specific authentication accuracy. Additionally, the application of cepstral feature extraction techniques (MFCC, GTCC) to physiological signals is an innovative contribution, as it enables better spectral representation of biometric patterns, outperforming traditional time-domain features. The study also conducts an extensive classifier comparison to identify the most effective model for biometric verification, paving the way for robust, real-time physiological biometric authentication in high-security applications.

Literature Review:

Recently, there has been growing interest in machine learning-based biometric identification using physiological signals. An innovative driver authentication system using electrocardiogram (ECG) signals from dry electrodes on a steering wheel is presented in [7]. By leveraging the unique, tamper-proof properties of ECG signals, this system addresses the limitations of conventional biometric methods. It uses a convolutional neural network (CNN) optimized for real-time processing along with autocorrelation profiles (ACPs). The system achieved high accuracy in automobile and security applications, with F1 scores of 96.8% and 96.02% on public and real-world datasets, respectively.

In [8], ECG signals from 35 participants were analyzed using empirical mode decomposition (EMD) to extract intrinsic mode functions (IMFs), with IMF 1 and 2 combined and classified using a cubic support vector machine (SVM), achieving an accuracy of 98.4%. Similarly, ECG signals from 36 participants was denoised with an infinite impulse response (IIR) filter, and 18 characteristic features were extracted. SVM outperformed K-nearest neighbor (KNN) and Naive Bayes (NB) classifiers, with an accuracy of 99.2% [9]. Another study involving 30 subjects (13 healthy, 17 non-healthy) from the PTB database reported an average frame identification rate of 97.31% by analyzing QRS beat data from ECG signals using a combination of autocorrelation, discrete cosine transform (DCT), and Mel frequency cepstral coefficients (MFCC) features [10].

Similarly, PCG signals from 30 individuals were denoised using EMD, and 11 features were extracted and classified. The SVM classifier achieved the highest accuracy of 95.4% [11]. PCG signals were also used for automatic person identification and verification using a back-propagation multilayer perceptron artificial neural network (BP-MLP-ANN) combined with wavelet-based features [12]. Another study applied wavelet packet decomposition to heart sounds, extracting key features using linear and non-linear filter banks at various decomposition levels. Automatic wavelet denoising was used for preprocessing, and a linear discriminant classifier achieved 91.05% accuracy on a dataset of heart sounds from 206 individuals [13].

A speech-based biometric system using EMG signals is presented in [14]. It recorded muscle activity in the neck during speech and used EMD for denoising, followed by time- and frequency-based feature extraction. Among different classifiers, the quadratic SVM reported the highest accuracy of 95.3% across 10 classes. EMG-based personal identification and verification were also explored in [15], where surface EMG signals from 21 participants were recorded while making a hand-open gesture using the Myo wristband. Two methods—discrete wavelet transform (DWT) with an extra trees classifier and continuous wavelet transform (CWT) with convolutional neural networks (CNN)—achieved a maximum accuracy of 99.285%.

Recent research focuses on fusing multiple physiological signals for biomedical applications. For example, [16] proposed a biometric identification system that combines cepstral features from ECG and PCG signals. Several classifiers were tested, with ensemble subspace discriminant and linear discriminant achieving 100% accuracy on a dataset of 32 individuals. Another study combined ECG and EMG signals using a Bayesian network, with the fused data used to control physiological devices during activities like cycling and rehabilitation exercises, improving accuracy in the rehabilitation process [17][18].

Although studies have explored ECG-PCG and ECG-EMG combinations, the fusion of ECG, PCG, and EMG for biometric identification remains underexplored. This study addresses that gap by collecting physiological signals from 36 subjects using the BIOPAC MP-36 system. The signals were preprocessed to remove power line interference while preserving key frequency components. GTCC and MFCC features were extracted and used as inputs for machine learning classifiers to evaluate accuracy, precision, robustness, and reliability.

This research introduces a novel biometric authentication system that integrates ECG, PCG, and EMG signals—a combination rarely explored in past studies. By fusing MFCC-based features from these signals, the system achieves 100% classification accuracy, outperforming traditional GTCC-based methods. Unlike single-modal biometric systems, this multimodal fusion enhances identity verification accuracy and improves resistance to spoofing. By analyzing the distinct features of ECG, PCG, and EMG signals, the study strengthens biometric security.

Because these signals originate from internal body processes, they are difficult to replicate. ECG measures heart rhythms, PCG records heart sounds influenced by valve movements, and EMG captures neuromuscular activity, which varies between individuals due to differences in muscle structure and movement patterns. It is almost impossible to mimic all three signals simultaneously, making this system highly secure. Additionally, requiring live physiological signals prevents replay attacks, and the fusion technique ensures consistency across modalities while distinguishing genuine from spoofed data. Advanced feature extraction using MFCCs and GTCCs further enhances the system's ability to detect fraudulent attempts.

Overall, by providing strong protection against identity theft and spoofing, this system could pave the way for future advancements in biometric authentication.

Materials and Methods:

This study presents a machine learning-based biometric authentication technique that leverages the fusion of ECG, PCG, and EMG signals. The full block diagram of the proposed approach is illustrated in Figure 2.

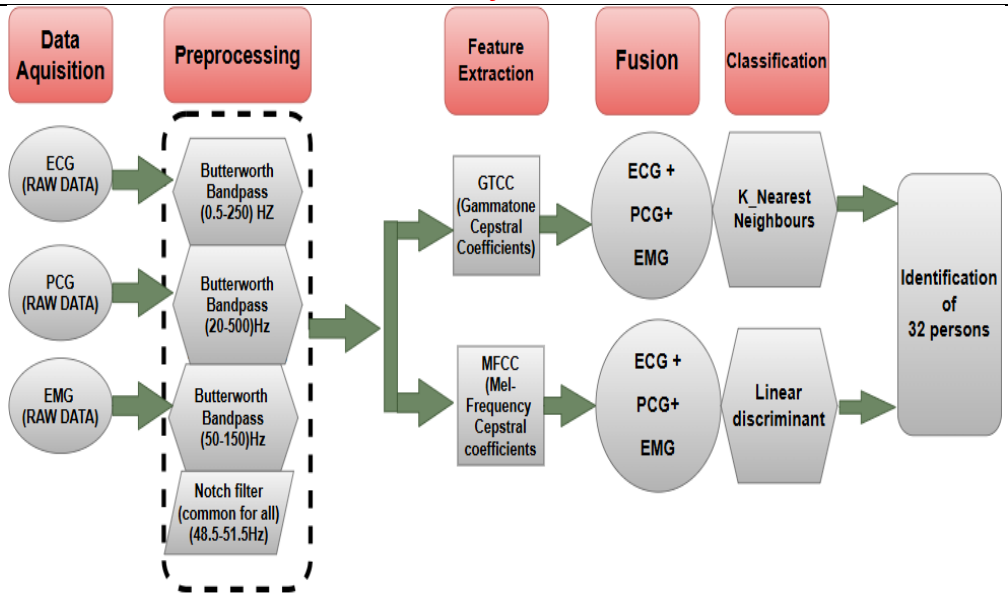


Figure 2. A block diagram of a biometric identification technique that combines ECG, PCG, and EMG signals.

Data Acquisition:

The BIOPAC MP-36 system was used to record ECG, PCG, and EMG signals. To capture ECG signals, electrodes (SS2L lead set) were placed on the left leg (positive), right leg (negative), and right forearm (neutral). For PCG recordings, the SS3L stethoscope was used, with heart sounds collected from one of the four auscultatory areas: Aortic, Pulmonic, Tricuspid, or Mitral. During the recordings, individuals remained seated upright on a chair and refrained from movement to maintain signal quality.

For EMG signal acquisition, the SS2L lead was also employed. The white wire was connected to the left wrist, while the red and black wires were placed near the elbow, with the red lead on the left side and the black lead on the right side. Volunteers were instructed to draw a specific pattern on a mobile phone, as shown in Figure 3, to facilitate EMG data collection.

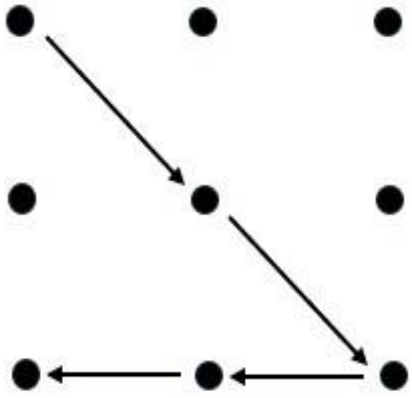


Figure 3. EMG signal pattern representing muscle movement.

Data was collected from 32 participants (29 men and 3 women) for 10 seconds at a sampling rate of 2000 Hz, resulting in a total of 192 signals [16].

Preprocessing:

Preprocessing is a critical step in machine learning-based approaches. Common techniques include resampling, normalization, noise reduction, and filtering. This stage improves signal quality by effectively minimizing power line interference and reducing motion artifacts. In this study, IIR Butterworth bandpass and Butterworth notch filters were applied to extract key frequencies and eliminate unwanted noise.

Butterworth Notch Filter for Power Line Interference Removal:

A notch filter was employed to eliminate 50 Hz power line interference from ECG, PCG, and EMG signals. This 4th-order Butterworth bandstop filter, designed with a stopband attenuation of 80 dB and a passband ripple of 1 dB, targeted the frequency range from 48.5 Hz to 51.5 Hz. This filtering step removed unwanted interference while preserving the signal quality crucial for biometric identification.

Butterworth Bandpass Filter for Target Frequencies:

To enhance the signals further, a Butterworth bandpass filter with a smooth, oscillation-free frequency response was applied to ECG, PCG, and EMG signals:

1. **ECG Signal Processing:** A bandpass filter with a passband of 0.5 Hz to 250 Hz [19] was used to capture relevant cardiac activity while filtering out extraneous noise.
2. **PCG Signal Processing:** A bandpass filter with a frequency range of 20 Hz to 500 Hz [16] was applied. This range effectively captured essential heart sounds, typically between 20 Hz and 200 Hz, while preserving high-frequency elements like clicks and irregular heartbeats (up to 500 Hz) and low-frequency murmurs (above 20 Hz).
3. **EMG Signal Processing:** EMG signals were filtered using a bandpass filter with a range of 50 Hz to 150 Hz. This frequency range preserved key signal components while reducing irrelevant noise, ensuring more accurate signal processing. The filter was designed with a stopband attenuation of 80 dB and a passband ripple of 1 dB, enhancing biological signal clarity.

By improving the signal-to-noise ratio, these filtering steps facilitated the precise extraction of ECG, PCG, and EMG features needed for accurate classification and biometric authentication.

After preprocessing, GTCC and MFCC features were extracted separately from each signal. This feature extraction aimed to reduce dimensionality and enhance algorithm efficiency, improving the overall performance of biometric identification.

Feature Extraction:

To reduce dimensionality and enhance algorithm performance, GTCC features were extracted separately from each ECG, PCG, and EMG signal after preprocessing. Three distinct GTCC features were taken from each signal to improve biometric identification and increase the accuracy and robustness of authentication. Figure 4 shows the process of extracting GTCC features from these signals. First, the preprocessed signals were passed through a gamma tone filter bank, which simulates human auditory perception by breaking the signals into different frequency bands. Next, logarithmic compression was applied to the filtered signals to reduce variations in dynamic range and highlight key perceptual features. After that, a Discrete Cosine Transform (DCT) was used on the compressed output to reduce feature correlation and create a compact representation. Finally, the extracted GTCC coefficients serve as critical features for classification. The entire process is computed as shown in equation 1 [20].

$$GTCC_a = \frac{\sqrt{2}}{b} \sum_{c=1}^{b} \log(Z_c) \cos\left[\frac{\pi c}{B}\left(a - \frac{1}{2}\right)\right] \qquad (1)$$

$1 \leq a \leq M$, where Z_c is the signal energy in the c_{ity} spectral band, b represents the number of Gammatone filters, and M is the number of GTCC.

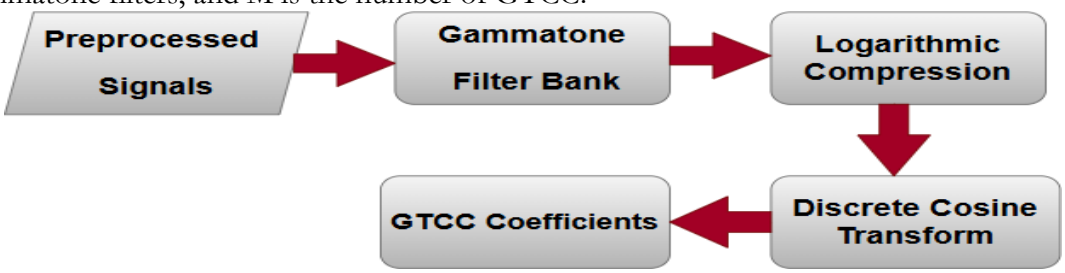


Figure 4. Flowchart depicting the steps involved in computing GTCCs.

After extracting GTCC features from each physiological signal, we further analyzed feature combinations in pairs to explore potential improvements in biometric authentication. Finally, we performed feature fusion by combining data from all three signals—ECG, PCG, and EMG—to assess overall performance.

Using the same approach, MFCC features were also extracted individually from each filtered ECG, PCG, and EMG signal, as shown in Equation 2 [21]. Figure 5 illustrates the MFCC extraction process from preprocessed signals. This process begins with windowing, where signals are divided into short frames to analyze short-term frequency patterns. Next, the discrete Fourier transform (DFT) is applied to convert the signals into the frequency domain. The resulting frequency spectrum is then passed through a Mel-scale filter bank, which enhances frequency components relevant to human auditory perception.

To further refine the features, a logarithmic transformation compresses the dynamic range, emphasizing key characteristics. The final step involves applying the discrete cosine transform (DCT) to achieve compact feature representation and decorrelation, resulting in MFCC coefficients commonly used in classification tasks. While MFCCs are well established in speech and audio processing, their application to biometric signals like ECG, PCG, and EMG is relatively recent. For each signal type, we extracted three unique MFCC features to ensure robust spectral characterization. The Mel-scale filter bank's nonlinear frequency resolution helps capture critical signal variations effectively.

$$MFCC_a[K] = \sum_{i=1}^{N-1} S[\mathrm{i}] \times cos\left[\frac{\pi k}{nFB} \times \left(i - \frac{1}{2}\right)\right] \qquad (2)$$

where k = 0,1,2…, nFB, where nFB represents the total number of filter banks.

Similar to the GTCC analysis, we first evaluated the MFCC features for each signal individually. Next, we assessed their effectiveness in pairs and, finally, fused features from all three signals—ECG, PCG, and EMG—to measure their combined impact on biometric authentication accuracy. This step-by-step approach allowed us to systematically examine the performance of both GTCC and MFCC features at various fusion levels, ensuring a comprehensive evaluation of their effectiveness.

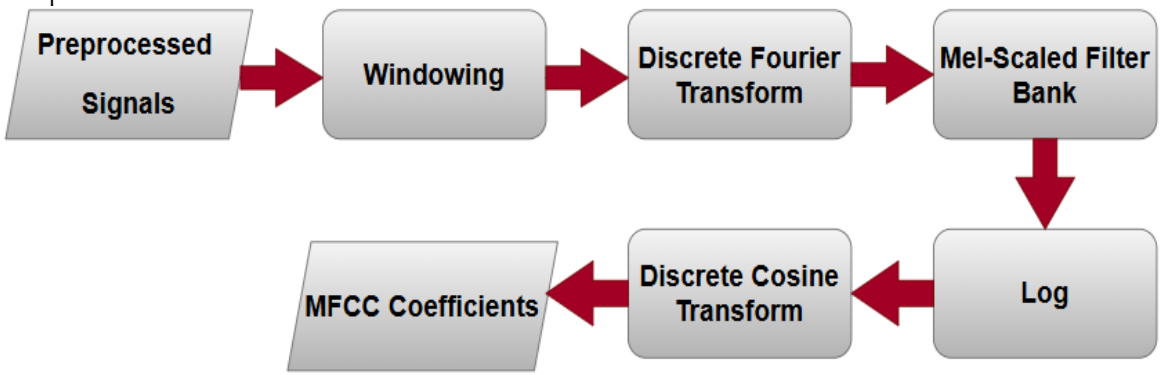


Figure 5. Flowchart depicting the steps involved in computing MFCCs.

Classification:

Classification involves training models to assign input data to predefined categories. After feature extraction, we applied 10-fold cross-validation using various machine-learning classifiers. This technique splits the dataset into 10 parts, where 9 parts are used for training and 1 for testing, to reduce overfitting and enhance model reliability. The classifiers used in this study included a wide neural network, SVM, SVM kernel, medium Gaussian SVM, fine tree, linear discriminant, ensemble bagged trees, fine KNN, and logistic regression kernel. The dataset was divided into 80% training and 20% testing to evaluate model performance. Among these classifiers, Fine KNN achieved the highest classification accuracy for GTCC-based feature classification.

KNN (K-Nearest Neighbors) assigns labels to data points based on their closest neighbors, making it a simple yet effective classification method. Its reliability and ease of use make it well-suited for biometric analysis, as it identifies patterns by analyzing the proximity of data points. Fine KNN, an improved version of traditional KNN, enhances classification accuracy by refining neighbor selection and distance measurements.

The same classification techniques were applied to analyze MFCC-based features, again using 10-fold cross-validation. The classifiers evaluated included a wide neural network, SVM, SVM kernel, medium Gaussian SVM, fine decision tree, linear discriminant, ensemble bagged trees, fine KNN, and logistic regression kernel. In this case, the linear discriminant classifier achieved the highest classification accuracy, demonstrating its strong ability to differentiate biometric features.

Linear discriminant analysis (LDA) is a machine learning technique that creates a linear combination of features to distinguish between two or more classes of objects or events. It can be used either as a standalone linear classifier or as a dimensionality reduction method before classification. LDA excels at separating data classes, particularly in high-dimensional datasets, making it highly effective for applications like image recognition and biometrics. It focuses on modeling variations between classes under the assumption that each class has a similar covariance structure.

By systematically evaluating the GTCC and MFCC feature sets with the same classification techniques, we provided a comprehensive performance analysis. The results revealed that Fine KNN was the best classifier for GTCC-based features, while linear discriminant analysis performed best for MFCC-based features in this study.

Results:

This section evaluates the efficacy of biometric authentication using the GTCC and MFCC feature extraction algorithms. The effectiveness of cepstral features in biometric identification was tested through experiments on ECG, PCG, and EMG signals.

GTCC-Based Approach:

In this study, we developed a machine learning-based biometric identification system using ECG, PCG, and EMG signals. The proposed method first removes noise from the raw physiological signals and isolates relevant frequencies to extract key GTCC features and complex spectral information. Various classifiers were then applied to perform biometric identification.

Initially, the system's performance was evaluated using GTCC features from each signal—ECG, PCG, and EMG—individually. Next, we combined GTCC features in pairs, such as (ECG, PCG), (PCG, EMG), and (ECG, EMG), for further analysis. Finally, we fused all three GTCC feature sets (ECG, PCG, and EMG) and assessed the performance of each classifier. This multi-level fusion approach significantly enhanced the model's accuracy and robustness, improving the overall effectiveness of the biometric authentication system.

Table 1 presents the classification accuracy achieved by each classifier when using GTCC features from ECG, PCG, and EMG signals individually. The results highlight the unique strengths and capabilities of each classifier. Before applying fusion, this comparison demonstrates the classifiers' potential for biometric verification.

Among the classifiers, the ensemble bagged tree achieved the highest accuracy of 82.3% for ECG-based GTCC features. In contrast, the medium Gaussian SVM recorded the lowest accuracy, at 50.0%, for EMG signals. Additionally, the wide neural network classifier achieved an accuracy of 53.1% for PCG signals. These findings reflect the varying classification performances and provide insights into the strengths and limitations of different classifiers when applied to individual physiological signals.

Table 1. Accuracy of ECG, PCG, and EMG features using GTCC concerning various classifiers

Classifier	ECG	PCG	EMG
Fine KNN	76.0%	49.5%	45.3%
Linear Discriminant	79.2%	43.2%	46.9%
Medium Gaussian SVM	80.7%	51.6%	**50.0%**
Ensemble Bagged Trees	**82.3%**	50.0%	47.4%
Wide Neural Network	72.4%	**53.1%**	40.1%
SVM Kernel	24.5%	5.7%	13.0%
Fine Tree	76.0%	40.1%	49.0%
Logistic Regression	18.8%	5.2%	6.2%

The classification accuracies obtained from GTCC features for individual ECG, PCG, and EMG signals were lower compared to previously reported results. To address the limitations of individual signals and enhance classification accuracy, we fused two feature sets to combine complementary information, as shown in Table 2.

When ECG-based features were fused with PCG-based features, the linear discriminant classifier's accuracy improved to 92.2%. Similarly, fusing ECG and EMG features boosted system performance, achieving the highest accuracy of 94.8% with the linear discriminant classifier. For the PCG and EMG feature fusion, the lowest accuracy recorded was 85.4%, which still outperformed the highest accuracy (82.3%) from individual signal classification, as shown in Table 1.

These findings demonstrate that signal fusion enhances classification performance and provides a more reliable feature set for biometric authentication.

Table 2. Accuracy of combination pairs of ECG, PCG, and EMG features using GTCC concerning different classifiers

Classifier	ECG & PCG	ECG & EMG	PCG & EMG
Fine KNN	88.5%	90.1%	**85.4%**
Linear Discriminant	**92.2%**	**94.8%**	79.2%
Medium Gaussian SVM	91.1%	92.2%	78.6%
Ensemble Bagged Trees	89.1%	89.6%	78.6%
Wide Neural Network	83.3%	85.9%	79.2%
SVM Kernel	83.9%	76.6%	64.6%
Fine Tree	84.4%	83.9%	62.0%
Logistic Regression	69.8%	63.5%	43.8%

Table 3. Performance metrics of various classifiers on the multi-modal fused feature set of ECG, PCG, and EMG signals using GTCC.

Classifier	Accuracy	Precision	Recall	F1-Score
Fine KNN	**98.4%**	98.7%	98.4%	98.37%
Linear Discriminant	**98.4%**	98.51%	98.43%	98.43%
Medium Gaussian SVM	96.9%	97.4%	93.7%	96.9%
Ensemble Bagged Trees	94.8%	98.34%	94.78%	94.7%
Wide Neural Network	92.7%	93.56%	92.7%	92.68%
SVM Kernel	87.0%	88.6%	86.98%	87.01%
Fine Tree	84.9%	82.4%	85.41%	85.14%
Logistic Regression Kernel	68.2%	72.68%	68.21%	67.7%

Building on the increased classification accuracy observed from combining two physiological signal features, we expanded this strategy by fusing all three signals—ECG, PCG, and EMG—into a multi-modal approach to assess the system's performance. This comprehensive fusion further enhanced overall system accuracy. Both Fine KNN and linear discriminant classifiers achieved the highest classification accuracy of 98.4%, while the lowest

accuracy, 68.2%, was recorded by the logistic regression kernel. Notably, most classifiers surpassed 84% accuracy, with medium Gaussian SVM, ensemble bagged trees, and wide neural networks achieving accuracies of over 92%.

In addition to accuracy, other performance metrics, such as precision, recall, and F1-score, were computed to further evaluate the effectiveness of the proposed approach. The results, summarized in Table 3, demonstrate the substantial benefits of multi-signal fusion for biometric authentication, with clear improvements in classification performance.

A bar graph in Figure 6 illustrates the classification accuracy of different classifiers for the multi-modal fusion of ECG, PCG, and EMG signals. Fine KNN and linear discriminant achieved the top accuracy of 98.4%, followed by medium Gaussian SVM at 96.9% and ensemble bagged trees at 94.8%. These findings emphasize the efficacy of fusing physiological signals to improve biometric authentication accuracy.

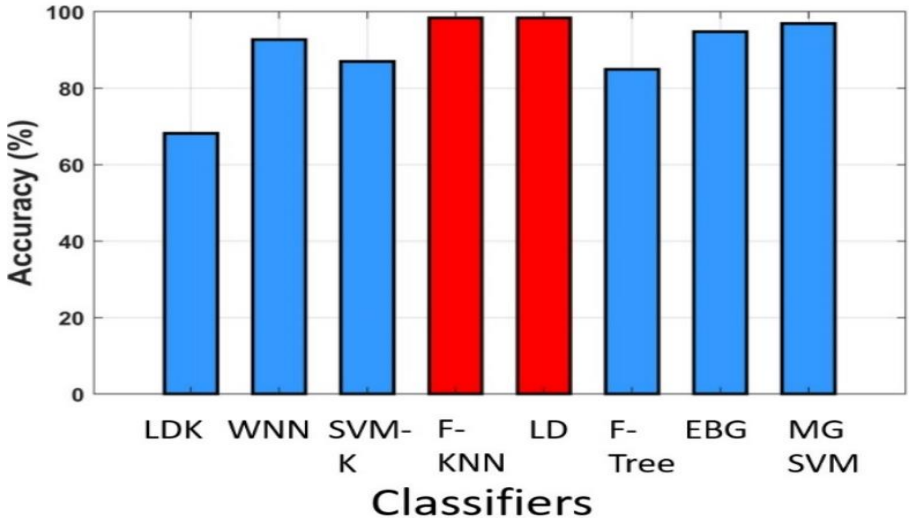


Figure 6. Performance of different Classifiers Logistic Regression Kernel (LDK), Wide Neural Network (WNN), SVM Kernal (SVM-K), Fine KNN (F-KNN), Linear discriminant (LD), Fine-Tree (F-Tree), Ensemble Bagged Tree (EBG), Medium Gaussian SVM (MG-SVM) using GTCC.

The confusion matrix for cross-validation using the Fine KNN classifier is presented in Figure 7. Similarly, the confusion matrix for hold-out validation, based on a 70-30 data split, is shown in Figure 8. In both cases, the accuracy remained consistent at approximately 98.4%, demonstrating the model's robustness and reliability. Validating the system's performance across different data splits further reinforces the effectiveness and stability of the proposed biometric authentication system.

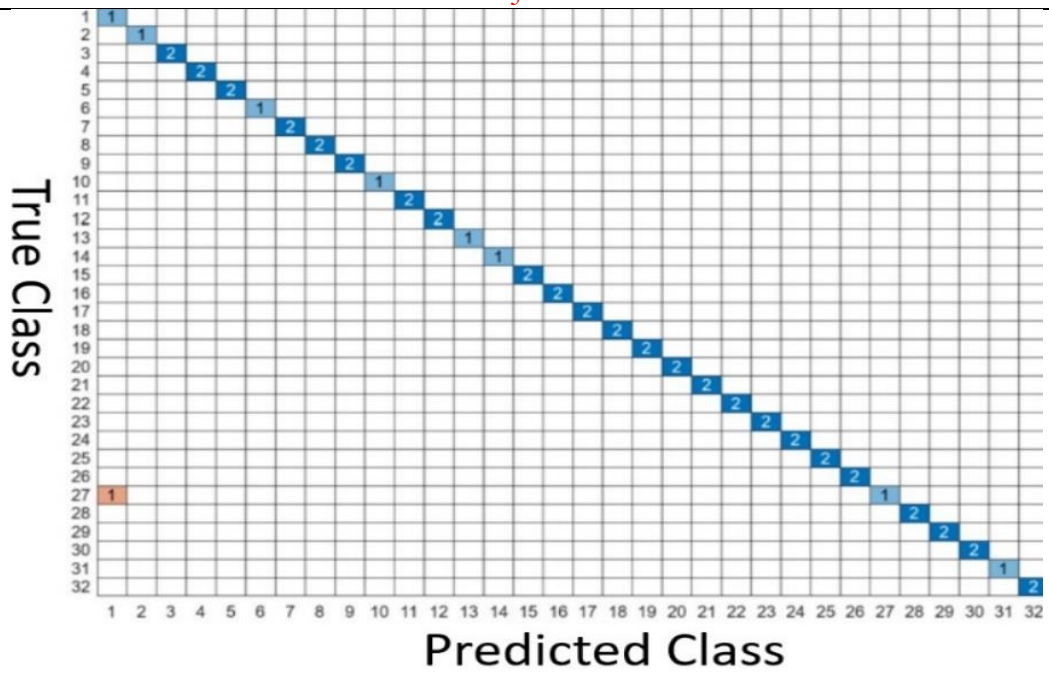


Figure 7. Confusion matrix through Cross-Validation using GTCC-based features

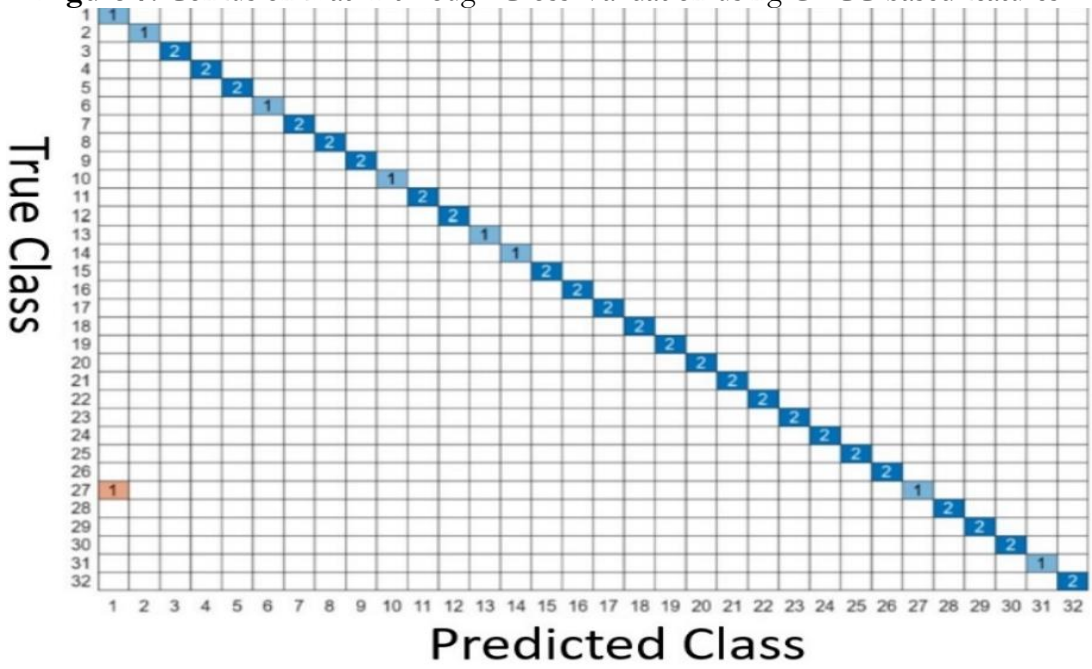


Figure 8. Confusion matrix through Holdout Validation using GTCC-based features

MFCC-Based Approach:

In this approach, MFCC features were extracted from ECG, PCG, and EMG signals to assess their effectiveness in biometric authentication. The MFCC technique captures spectral envelope characteristics by applying Mel-scaling and the discrete cosine transform (DCT), ensuring robust feature extraction from physiological signals.

Similar to the GTCC-based method, classification performance was first evaluated by independently extracting MFCC features from ECG, PCG, and EMG signals. Next, feature fusion was applied in pairs—(ECG, PCG), (PCG, EMG), and (ECG, EMG)—to explore the impact of multimodal integration. Finally, MFCC features from all three signals (ECG, PCG, and EMG) were fully combined, and different classifiers were tested to assess their effectiveness.

The results showed that feature fusion significantly enhanced the accuracy of biometric authentication, highlighting the benefits of integrating diverse physiological signals for improved individual identification. Among the classifiers used, the linear discriminant classifier achieved the highest accuracy with MFCC features, demonstrating its superior ability to distinguish distinct biometric patterns.

Table 4 presents the classification accuracies achieved by various classifiers when MFCC features were extracted separately from ECG, PCG, and EMG signals. The results emphasize the varying effectiveness of each classifier in biometric authentication and provide insights into their performance before feature fusion. The linear discriminant classifier achieved the highest accuracies, with 94.3% for ECG, 67.7% for PCG, and 85.9% for EMG signals, underscoring its strength in differentiating biometric patterns across different physiological signals. Additionally, these accuracies surpassed those obtained with GTCC features, as shown in Table 1.

Table 4. Accuracy of individual ECG, PCG, and EMG features concerning various classifiers using MFCC features.

Classifier	ECG	PCG	EMG
Fine KNN	87.5%	64.6%	84.6%
Linear Discriminant	**94.3%**	**67.7%**	**85.9%**
Medium Gaussian SVM	88.0%	64.6%	83.3%
Ensemble Bagged Trees	87.0%	62.0%	97.6%
Wide Neural Network	85.9%	61.5%	82.3%
SVM Kernel	91.7%	63.0%	80.2%
Fine Tree	64.1%	46.4%	57.3%
Logistic Regression	78.1%	51.6%	67.2%

The classification accuracies derived from MFCC features for individual ECG, PCG, and EMG signals were lower compared to previously reported studies. To address this limitation, feature sets were fused to evaluate the impact of combining complementary information, as outlined in Table 5.

The fusion of ECG and PCG features significantly enhanced performance, with the linear discriminant classifier achieving an accuracy of 99.0%. Similarly, when ECG and EMG features were combined, the system's performance further improved, attaining a maximum accuracy of 96.4% with the same classifier. In the case of PCG and EMG feature fusion, the recorded accuracy was 95.3%, which exceeded the highest accuracy obtained from any individual signal (as indicated in Table 4).

These findings demonstrate that signal fusion enhances classification performance, resulting in a more robust and reliable feature set for biometric authentication. This improvement emphasizes the value of integrating diverse physiological signals to strengthen the overall accuracy and reliability of the proposed biometric system.

Table 5 Accuracy of combination pairs of ECG, PCG, and EMG features concerning different classifiers using MFCC features

Classifier	ECG & PCG	ECG & EMG	PCG & EMG
Fine KNN	92.2%	92.7%	91.7%
Linear Discriminant	**99.0%**	**96.4%**	**95.3%**
Medium Gaussian SVM	92.2%	92.2%	91.7%
Ensemble Bagged Trees	91.7%	87.5%	83.9%
Wide Neural Network	82.5%	88.5%	90.6%
SVM Kernel	91.7%	81.2%	81.8%
Fine Tree	62.5%	58.3%	57.7%
Logistic Regression	83.9%	70.8%	69.3%

Table 6. Performance metrics of various classifiers on the multi-modal fused feature set of ECG, PCG, and EMG signals using MFCC features.

Classifier	Accuracy	Precision	Recall	F1-Score
Fine KNN	98.4%	98.4%	97.9%	97.9%
Linear Discriminant	**100%**	100%	100%	100%
Medium Gaussian SVM	98.4%	99.2%	98.9%	98.95%
Ensemble Bagged Trees	97.9%	98.7%	98.8%	98.86%
Wide Neural Network	97.4%	938.7%	98.91%	98.90%
SVM Kernel	95.8%	98.6%	98.7%	98.72%
Fine Tree	67.7%	96.30%	96.45%	96.35%
Logistic Regression Kernel	89.1%	98.50%	98.6%	98.55%

Building upon the improved classification accuracy observed with the fusion of two physiological signals using MFCC features, this strategy was extended to integrate all three signals—ECG, PCG, and EMG—resulting in a multi-modal biometric authentication system. This comprehensive fusion approach significantly enhanced overall system performance, surpassing the accuracy achieved with the GTCC-based method.

The linear discriminant classifier achieved the highest classification accuracy of 100.0%, underscoring the effectiveness of feature integration from multiple modalities. Additionally, Fine KNN and medium Gaussian SVM delivered strong performances, each reporting a promising accuracy of 98.4%, further validating the robustness of the fused feature set. Although most classifiers demonstrated high accuracy, the fine tree classifier recorded the lowest accuracy of 67.7%, suggesting that certain models may be less suited for multi-modal biometric authentication.

These results emphasize the potential of this multi-modal approach in achieving highly reliable identity verification. To further evaluate the system's effectiveness, additional performance metrics, including precision, recall, and F1-score, were computed, offering a more detailed assessment of the proposed method's efficacy.

The bar graph in Figure 9 illustrates the classification accuracy of various classifiers for the suggested multi-modal system integrating ECG, PCG, and EMG signals. The linear discriminant classifier attained the peak accuracy of 100%, followed by Fine KNN and medium Gaussian SVM, both exhibiting exceptional results. These findings reaffirm the significant impact of signal fusion on biometric authentication, demonstrating a substantial enhancement in classification performance and overall system reliability.

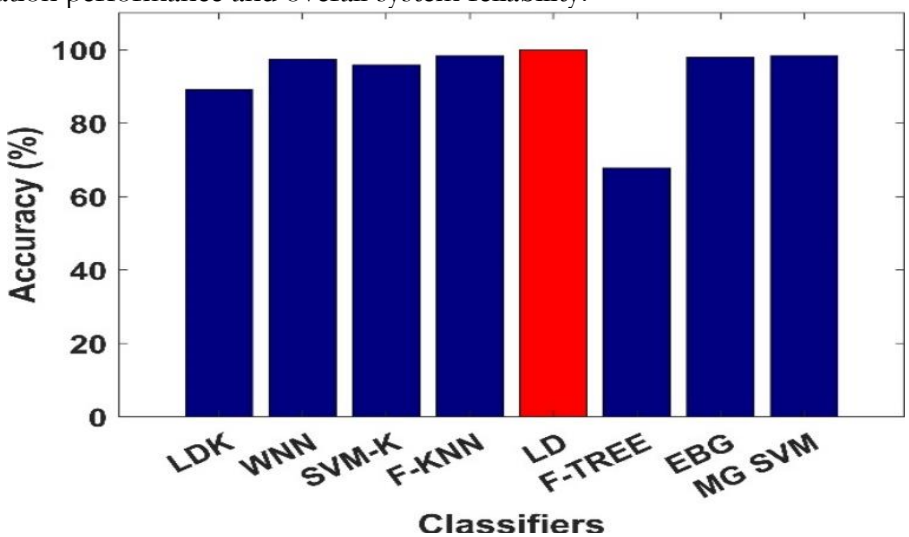


Figure 9. Performance of different Classifiers using MFCC features.

Figure 10 displays the confusion matrix obtained using the linear discriminant classifier with cross-validation, while Figure 11 presents the confusion matrix for hold-out validation with a 70-30 data split. In both cases, the accuracy remained consistently high, achieving a perfect score of 100%.

This exceptional result highlights the model's robustness and reliability. The stable performance across different validation techniques further reinforces the effectiveness of the proposed biometric authentication system, demonstrating its strong generalization capability and potential for real-world applications.

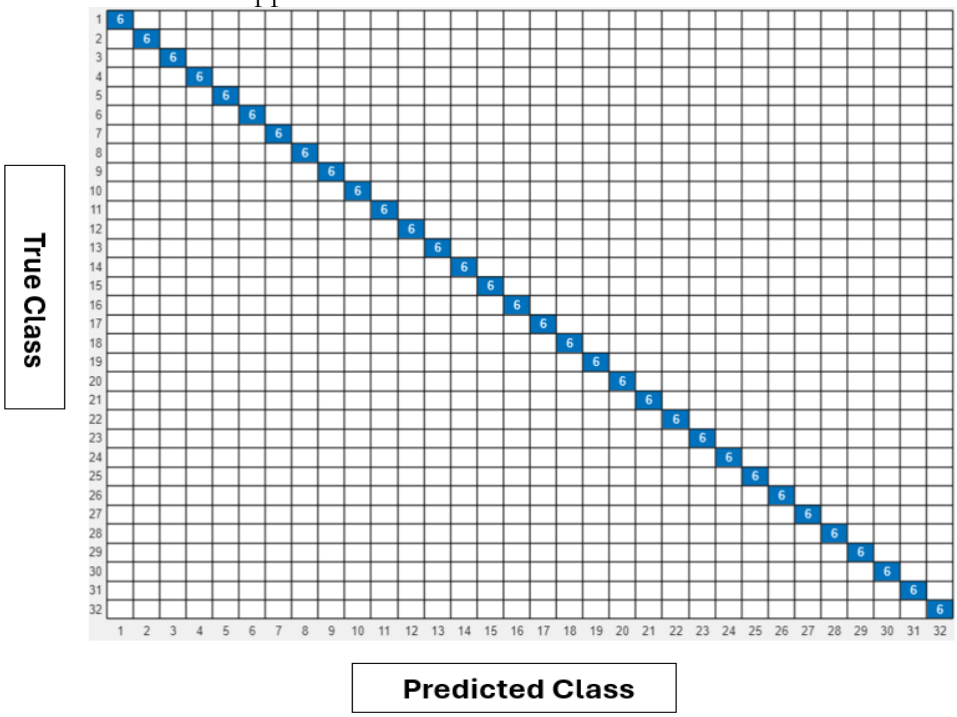


Figure 10. Confusion matrix of Cross-Validation using MFCC based features

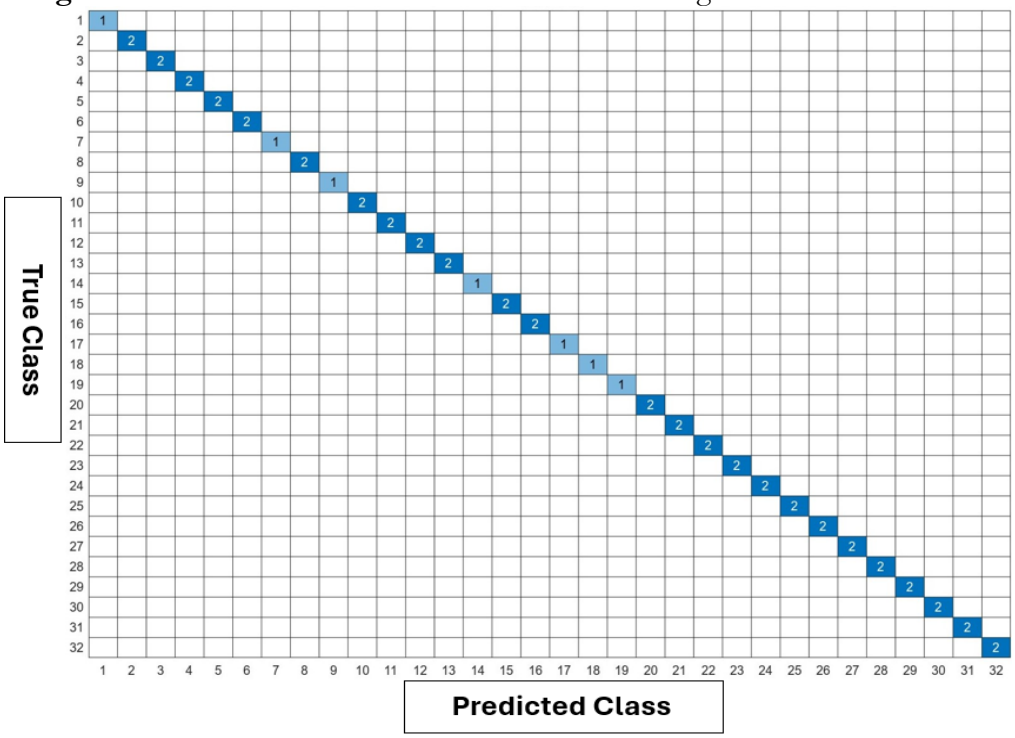


Figure 11. Confusion matrix of Holdout Validation using MFCC's features

Performance Comparison of GTCC and MFCC-Based Biometric Authentication:

The performance evaluation of both GTCC and MFCC-based approaches underscores the effectiveness of spectral features in biometric authentication using ECG, PCG, and EMG signals. Initially, when individual signals were analyzed using GTCC features, the highest accuracy of 82.3% was achieved for ECG with the ensemble bagged tree classifier, while PCG and EMG showed lower accuracies of 53.1% and 50.0%, respectively, using wide neural network and medium Gaussian SVM. In contrast, the MFCC-based approach outperformed GTCC, with accuracies of 94.3% for ECG, 67.7% for PCG, and 85.9% for EMG, all using the linear discriminant classifier. These results indicate that MFCC features offer a more effective spectral representation for biometric classification.

A similar trend was observed with pairwise feature fusion. In the GTCC-based system, fusion enhanced accuracy, with ECG and PCG reaching 92.2%, ECG and EMG achieving 94.8%, and PCG and EMG yielding 85.4%, using fine KNN and linear discriminant classifiers. However, the MFCC-based approach demonstrated even better performance: fusion of ECG and PCG resulted in 99.0% accuracy, ECG and EMG reached 96.4%, and PCG and EMG achieved 95.3%, all with linear discriminant analysis. This highlights the superior ability of MFCC features to integrate complementary information from multiple signals.

The most notable improvement occurred when all three signals—ECG, PCG, and EMG—were combined. In the GTCC-based method, Fine KNN and linear discriminant achieved a maximum accuracy of 98.4%, while the logistic regression classifier recorded the lowest accuracy at 68.2%. In contrast, the MFCC-based system achieved a perfect 100% accuracy using the linear discriminant classifier, demonstrating its superior ability to extract unique biometric features.

Overall, the MFCC-based approach consistently outperformed the GTCC-based method at every stage—whether for individual signals, pairwise fusion, or full feature fusion of ECG, PCG, and EMG signals. The MFCC features' ability to capture detailed spectral variations and signal characteristics contributed to higher accuracy, making them a more reliable choice for biometric authentication. This evaluation demonstrates that MFCC features provide a more robust, accurate, and effective solution for biometric identification compared to GTCC features.

Discussion:

The findings of this study demonstrate that the fusion of ECG, PCG, and EMG signals significantly enhances biometric authentication accuracy, with MFCC-based feature extraction achieving 100% classification accuracy. The comparative evaluation of machine learning classifiers reveals that Fine KNN and Ensemble Bagged Trees outperform other traditional models, highlighting the effectiveness of fusion-based physiological biometrics. The superior performance of cepstral features (MFCC, GTCC) over traditional statistical features aligns with previous research that emphasizes the importance of frequency-domain representations for physiological signal classification (Abo-Zahhad et al., 2015). Additionally, the use of a real-world dataset collected from 32 participants provides a practical evaluation of the system's capabilities, making it more applicable to biometric authentication scenarios than studies relying on publicly available datasets.

While the study achieves high accuracy with traditional machine learning models, it does not benchmark performance against deep learning-based biometric authentication techniques, such as CNNs, LSTMs, or transformer-based models. Recent research indicates that CNNs excel in feature extraction by automatically learning hierarchical patterns in physiological signals, outperforming handcrafted feature approaches in biometric authentication (Ku et al., 2024). Similarly, LSTMs and Bi-LSTMs are highly effective in time-series processing, making them well-suited for physiological signal modeling.

The performance of the proposed biometric authentication system was compared with existing research, as shown in Table 7. The results indicate that our system achieved competitive

or even better classification accuracy than previously reported methods. Earlier studies have used different techniques for biometric authentication. For instance, the study in [8] applied EMD and achieved an accuracy of 98.4%. The study in [12] used a wavelet transform method, reaching 90.52% accuracy, while [22] employed wavelet-based classification and reported 86.7% accuracy. Similarly, [23] adopted an advanced composite multiscale dispersion entropy (RCMDE) approach, achieving 96.08% accuracy.

In [16], researchers reached 100% accuracy by extracting four MFCCs and four GTCCs features from combined ECG and PCG signals. Notably, our method achieved the same 100% accuracy using just three MFCC features from EMG signals, demonstrating its efficiency and effectiveness. This difference highlights that our approach is a practical option for biometric authentication, as it simplifies feature extraction while maintaining excellent performance.

Additionally, [16] used both PCG and ECG signals in a multimodal setup, whereas our system reached the same accuracy by extracting MFCC features from a combination of ECG, PCG, and EMG signals. This underscores the reliability and effectiveness of our method, as it delivers high classification accuracy with reduced computational complexity. Moreover, our system outperformed the accuracy reported in [12], [22], and [23], further confirming the importance of MFCC features in biometric validation. These findings emphasize the variety of feature sets and classification models explored, each with varying accuracy levels in different biometric data classification contexts.

The study also found that MFCC features outperformed GTCC features in terms of classification accuracy. Using the Linear Discriminant classifier, the system achieved 100% accuracy with MFCC features, while GTCC features yielded slightly lower accuracy. This suggests that MFCC features capture more distinctive biometric traits, leading to improved authentication performance.

Despite these promising results, the proposed approach has some limitations. The dataset includes only 32 participants, which may reduce its generalizability to larger populations. Expanding the dataset would improve reliability. Additionally, physiological signals can be affected by environmental noise, sensor placement, and participant movement, making noise reduction techniques necessary for real-time applications. Another challenge is computational complexity, as feature extraction and classification involve multiple processing steps that require efficient hardware and optimization for real-time use. Finally, acquiring physiological signals through electrodes (e.g., ECG) may cause discomfort for some users. Future research should explore non-invasive signal acquisition techniques to enhance user comfort.

Table 7 Comparison of proposed work with previously reported results

Study	Method	Classification	Accuracy
[8]	ED	SVM-C	98.4%
[12]	Wavelet Transform	BP-MLP- ANN	90.52%
[22]	Wavelet	EB-Trees	86.7%
[16]	PCG and ECG fusion, IIR filter	Ensembled	100%
[23]	RCMDE	ED	96.08%
This work	**GTCC,**	**Fine KNN,**	**98.4%, 100%**
This work	**MFCC**	**Linear Discriminant**	98.4%

An important aspect of biometric authentication is security against adversarial attacks. Although the proposed system demonstrates high accuracy in controlled conditions, it is essential to assess its robustness against spoofing attacks, synthetic signal injections, and adversarial perturbations. Prior research (Jain & Nandakumar, 2016) indicates that biometric authentication models can be vulnerable to signal replay attacks, where recorded physiological data is used to bypass security systems. Future work should incorporate adversarial testing, noise injection, and spoofing resilience analysis to ensure system integrity in high-security

environments. Additionally, integrating secure biometric storage solutions, such as blockchain-based identity management, could enhance data security and user privacy.

Finally, the dataset size and participant diversity present limitations that may impact model generalizability. The dataset used in this study consists of 32 participants (29 men, 3 women), raising concerns regarding demographic bias in classification performance. Prior research (Cheng et al., 2020) suggests that biometric models trained on unbalanced datasets may exhibit lower accuracy across diverse population groups. Expanding the dataset to include a balanced representation of gender, age groups, and medical conditions will improve model robustness and enhance fairness in biometric authentication systems. Additionally, evaluating model performance on external datasets will help assess its applicability across different biometric acquisition conditions.

Conclusion:

This study presents a machine learning-based biometric authentication system that uses the fusion of physiological signals. Raw signals were collected using the BIOPAC MP-36 device and preprocessed with Butterworth bandpass and notch filters to eliminate noise and extract relevant frequencies. Next, GTCC and MFCC cepstral features were extracted to capture the spectral characteristics of the signals.

The system's performance was evaluated separately using two feature extraction approaches (GTCC and MFCC) applied to ECG, PCG, and EMG signals. Various machine learning classifiers were then used to assess the effectiveness of these features. After analyzing the performance of each physiological signal individually, pairwise feature fusion was performed for both GTCC and MFCC approaches to enhance classification accuracy. Finally, features from all three signals were combined to further improve system performance.

The results showed that MFCC-based features outperformed GTCC-based features in biometric authentication. The highest accuracy achieved with GTCC features was 98.4% using the Fine KNN and Linear Discriminant classifiers, whereas MFCC-based fusion achieved a perfect 100% accuracy with the Linear Discriminant classifier. This highlights the superior robustness and discriminative power of MFCC features for biometric authentication.

To further evaluate the proposed method, we plan to expand the dataset by including more participants and recordings, which will allow for a more comprehensive assessment of the system's reliability and generalization. Additionally, we aim to explore advanced feature extraction techniques and deep learning frameworks to enhance the accuracy and reliability of biometric authentication on larger datasets.

Author's Contribution: Tuba Alvi: Conceptualization, Methodology, Software, Data acquisition, Formal analysis, Writing – original draft.
Yumna Aziz: Conceptualization, Methodology, Software, Data acquisition Writing – original draft.
Muhammad Faraz: Validation, Resources, Investigation, Writing – original draft, Writing – review & editing.
Zubair Mehmood: Visualization, Writing – review & editing.
Syed Zohaib Hassan Naqvi: Resources, Visualization, Writing – review & editing.
Laraib Imtiaz: Software, Data acquisition, Writing – original draft.

Conflict of interest. The authors have no conflict of interest in publishing this manuscript in IJIST.

References:

[1] A. R. Anil K. Jain, Karthik Nandakumar, "50 years of biometric research: Accomplishments, challenges, and opportunities," *Pattern Recognit. Lett.*, vol. 79, pp. 80–105, 2016, doi: https://doi.org/10.1016/j.patrec.2015.12.013.

[2] K. Y. Moon, "Biometrics technology status and prospects," *Internet Serv. Inf. Secur.*, vol. 98, pp. 38–47, 2005.

[3] P. Mi *et al.*, "Driver Cognitive Architecture Based on EEG Signals: A Review," *IEEE Sens. J.*, 2024, doi: 10.1109/JSEN.2024.3471699.

[4] A. N. Uwaechia and D. A. Ramli, "A Comprehensive Survey on ECG Signals as New Biometric Modality for Human Authentication: Recent Advances and Future Challenges," *IEEE Access*, vol. 9, pp. 97760–97802, 2021, doi: 10.1109/ACCESS.2021.3095248.

[5] S. S. Roberta Avanzato, Francesco Beritelli, "Robust Biometric Verification Using Phonocardiogram Fingerprinting and a Multilayer-Perceptron-Based Classifier," *Electronics*, vol. 13, no. 22, p. 4377, 2024, doi: https://doi.org/10.3390/electronics13224377.

[6] S. B. P. Jin Su Kim, "A Study on EMG-based Biometrics," Journal of Internet Services and Information Security. Accessed: Mar. 17, 2025. [Online]. Available: https://jisis.org/article/jisis-2017-vol7-no2-02/69704/

[7] Giwon Ku, Choeljun Choi, Chulseung Yang, Jiseong Jeong , Pilkyo Kim, Sangyong Park, "Electrocardiogram-Based Driver Authentication Using Autocorrelation and Convolutional Neural Network Techniques," *Electronics*, vol. 13, no. 24, p. 4974, 2024, doi: https://doi.org/10.3390/electronics13244974.

[8] C. Lipps, LeaBergkemper, J. Herbst, and H. D. Schotten, "I Know You by Heart: Biometric Authentication based on Electrocardiogram (ECG) signals," *Int. Conf. Cyber Warf. Secur.*, vol. 17, no. 1, 2022, doi: https://doi.org/10.34190/iccws.17.1.12.

[9] M. Asim, M. Akhtar, M. Faraz, M. U. Khan, S. Aziz, and G. A. Montes, "Pattern Analysis for Biometric Authentication using Electrocardiogram Signal," *2023 2nd Int. Conf. Emerg. Trends Electr. Control. Telecommun. Eng. ETECTE 2023 - Proc.*, 2023, doi: 10.1109/ETECTE59617.2023.10396767.

[10] H. Gurkan, U. Guz, and B. S. Yarman, "A novel biometric authentication approach using electrocardiogram signals," *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS*, pp. 4259–4262, 2013, doi: 10.1109/EMBC.2013.6610486.

[11] M. U. Khan, S. Aziz, A. Zainab, H. Tanveer, K. Iqtidar, and A. Waseem, "Biometric System using PCG Signal Analysis: A New Method of Person Identification," *2nd Int. Conf. Electr. Commun. Comput. Eng. ICECCE 2020*, Jun. 2020, doi: 10.1109/ICECCE49384.2020.9179257.

[12] G. Gautam and D. Kumar, "Biometric system from heart sound using wavelet based feature set," *Int. Conf. Commun. Signal Process. ICCSP 2013 - Proc.*, pp. 551–555, 2013, doi: 10.1109/ICCSP.2013.6577115.

[13] M. Abo-Zahhad, S. M. Ahmed, and S. N. Abbas, "A new biometric authentication system using heart sounds based on wavelet packet features," *Proc. IEEE Int. Conf. Electron. Circuits, Syst.*, vol. 2016-March, pp. 17–20, Mar. 2016, doi: 10.1109/ICECS.2015.7440238.

[14] M. U. Khan, Z. A. Choudry, S. Aziz, S. Z. H. Naqvi, A. Aymin, and M. A. Imtiaz, "Biometric Authentication based on EMG Signals of Speech," *2nd Int. Conf. Electr. Commun. Comput. Eng. ICECCE 2020*, Jun. 2020, doi: 10.1109/ICECCE49384.2020.9179354.

[15] L. Lu, J. Mao, W. Wang, G. Ding, and Z. Zhang, "A Study of Personal Recognition Method Based on EMG Signal," *IEEE Trans. Biomed. Circuits Syst.*, vol. 14, no. 4, pp. 681–691, Aug. 2020, doi: 10.1109/TBCAS.2020.3005148.

[16] A. Mehmood, A. Hannan, M. U. Khan, S. Z. H. Naqvi, M. Faraz, and S. Aziz, "Biofusion: Expanding Biometric Horizons with ECG and PCG Integration," *2023 Int. Conf. Digit. Futur. Transform. Technol. ICoDT2 2023*, 2023, doi: 10.1109/ICODT259378.2023.10325791.

[17] N. A. I. M. Rosli, M. A. A. Rahman, S. A. Mazlan, and H. Zamzuri,

"Electrocardiographic (ECG) and Electromyographic (EMG) signals fusion for physiological device in rehab application," *2014 IEEE Student Conf. Res. Dev. SCOReD 2014*, Mar. 2014, doi: 10.1109/SCORED.2014.7072965.

[18] L. Y. Qun Ma, Yibo Wu, Muhammad Usman Shoukat, Yukai Yan, Jun Wang, Long Yang, Fuwu Yan, "Deep Reinforcement Learning-Based Wind Disturbance Rejection Control Strategy for UAV," *Drones*, vol. 8, no. 11, p. 632, 2024, doi: https://doi.org/10.3390/drones8110632.

[19] T. K. Montree Kumngern, Nattharinee Aupithak, Fabian Khateb, "0.5 V Fifth-Order Butterworth Low-Pass Filter Using Multiple-Input OTA for ECG Applications," *Sensors*, vol. 20, no. 24, p. 7343, 2020, doi: https://doi.org/10.3390/s20247343.

[20] S. Jayalakshmy and G. F. Sudha, "GTCC-based BiLSTM deep-learning framework for respiratory sound classification using empirical mode decomposition," *Neural Comput. Appl.*, vol. 33, no. 24, pp. 17029–17040, Dec. 2021, doi: 10.1007/S00521-021-06295-X/METRICS.

[21] S. A. A. Yusuf and R. Hidayat, "MFCC feature extraction and KNN classification in ECG signals," *2019 6th Int. Conf. Inf. Technol. Comput. Electr. Eng. ICITACEE 2019*, Sep. 2019, doi: 10.1109/ICITACEE.2019.8904285.

[22] I. Fahad, M. A. R. Apu, A. Ghosh, and S. A. Fattah, "Phonocardiogram Heartbeat Segmentation and Autoregressive Modeling for Person Identification," *IEEE Reg. 10 Annu. Int. Conf. Proceedings/TENCON*, vol. 2019-October, pp. 1942–1946, Oct. 2019, doi: 10.1109/TENCON.2019.8929563.

[23] C. S. Xiefeng Cheng, Pengfei Wang, "Biometric Identification Method for Heart Sound Based on Multimodal Multiscale Dispersion Entropy," *Entropy*, vol. 22, no. 2, p. 238, 2020, doi: https://doi.org/10.3390/e22020238.

RESEARCH & INNOVATION DIVISION
IJIST

Cow Face Detection for Precision Livestock Management using YOLOv8

Umair Ali[1], Wasif Muhammad[1*]
[1]Intelligent Systems Laboratory, Department of Electrical Engineering, University of Gujrat, Pakistan.
*** Correspondence:** Wasif Muhammad: syed.wasif@uog.edu.pk

Citation | Ali. U, Muhammad. W, "Cow Face Detection for Precision Livestock Management using YOLOv8", IJIST, Special Issue. pp 128-132, March 2025
Received | Feb 17, 2025 **Revised |** March 01, 2025 **Accepted |** March 07, 2025 **Published |** March 10, 2025.

Precision livestock management is transforming traditional agricultural practices by boosting productivity, increasing yield, and automating tasks, all while reducing labor requirements and minimizing errors. Conventional methods for animal recognition are often unreliable, which has led to a growing preference for using cameras to identify animals, monitor their health, manage data, and maintain cattle records. However, small-scale farms with limited livestock, such as cows and goats, frequently face overfitting problems in traditional machine learning models due to insufficient training data. Identifying individual cows based on facial features becomes more effective after detecting the cow's face. This study addresses these challenges by fine-tuning YOLOv8, a pretrained model, using a mix of self-captured images and publicly available datasets to detect cow faces in complex environments. Integrating publicly available data and leveraging a pretrained COCO model has significantly improved the model's ability to generalize and accurately detect cow faces. YOLOv8, equipped with the COCO pretrained model, successfully detects nearly all types of cow faces, which can then be used for individual cow classification. This approach enhances cow recognition accuracy, contributing to more efficient farm management applications.

Keywords: Livestock management, Facial Recognition, Cow face detection, Transfer learning, Yolov8

Introduction:

The agriculture industry is evolving rapidly by integrating advanced technologies into traditional methods, resulting in higher efficiency and productivity. One such advancement is the emergence of data-driven approaches in farming, which has given rise to a new field known as Precision Agriculture. These innovative techniques play a key role in livestock management, helping to meet the growing global demand for food [1]. However, countries that rely heavily on livestock face challenges in food storage and agricultural resources, which are vital for human benefit. Additionally, significant issues related to animal health, labor shortages, and overall management persist [2]. As the need for increased animal production grows, it is equally essential to implement smart, efficient systems that can enhance productivity while ensuring long-term sustainability [3]. Traditional livestock management methods fall short in today's world due to their time-consuming nature, high risk of errors, and dependence on manual labor. Research shows that poor labor practices can negatively impact livestock health, increasing the likelihood of disease outbreaks due to insufficient care and ineffective management [4]. Another common issue is disputes among farmers, which often arise from the mixing of animals and the inability to track individual cows accurately [5]. To tackle these problems, various identification and tracking methods have been developed. Broadly, there are three types of cow identification techniques: permanent, semi-permanent, and temporary methods.

Permanent identification methods involve physically marking animals. Although these methods are widely used [6], they are often unreliable, especially when managing large herds. Semi-permanent methods, such as attaching tags or collars, are commonly adopted for tracking cows. However, these methods are prone to issues such as tag loss, wear and tear, and inaccuracies due to the involvement of manual labor [7]. Temporary methods, like RFID (Radio-Frequency Identification) tags, have gained popularity because they enable automated tracking. Despite their advantages, RFID systems often require frequent maintenance, and farmers report concerns about their durability. Moreover, the noise, overhead machinery, and handling stress during installation may disturb the animals. There is an increasing need for advanced solutions that can accurately record the history of cows. Camera-based systems have proven to be effective for this purpose. Unlike traditional methods, computer vision-based systems offer accurate and widely adopted solutions for cow recognition [8]. Since cows naturally recognize each other by their facial features, facial recognition technology can be a practical and efficient way to identify individual cows without the need for physical tags. This approach enhances sustainability and scalability, particularly for large herds [9].

One common challenge in object detection is dataset bias, which arises from limited and non-diverse datasets. When a model is trained on a small, specific dataset, it performs well with similar input but struggles with generalization. The scarcity of images captured under varying conditions—such as different weather, lighting, and camera angles—reduces the model's robustness. For instance, cow recognition can fail due to pose variations, including shifts in camera angles and cow movements, especially involving facial features. Convolutional Neural Networks (CNNs) are commonly used for facial recognition because of their ability to extract complex features. However, CNN models may suffer from overfitting when trained on small datasets [10]. Addressing these challenges is critical for effective livestock management, where variations in lighting, angles, and animal movement are inevitable [11]. Data preprocessing and augmentation can improve the performance of cow recognition models, but achieving generalization still requires diverse and extensive datasets. For higher accuracy, it is essential to first detect cow faces and then classify them based on their facial features. A supervised learning approach, with manual annotations to localize cow faces before classification, can enhance accuracy [12]. In the current study, a transfer learning-based approach is proposed, utilizing YOLOv8, a pre-trained model on the COCO dataset, to

accurately detect cow faces. This pre-trained model can be fine-tuned with a smaller, specialized dataset to improve cow face localization. YOLO (You Only Look Once) is a cutting-edge object detection model known for its speed and real-time accuracy, making it well-suited for this task [13].

Objectives:

The objective of this research is to enhance cow face detection under varying background conditions, angles, and lighting by leveraging a pre-trained model. Specifically, the study aims to fine-tune YOLOv8, a state-of-the-art object detection model, using a dataset of annotated frames derived from self-collected videos and publicly available sources. By combining the precision of YOLOv8 with a diverse dataset, the study seeks to improve detection accuracy and ensure reliable cow face recognition in complex environments.

Materials and Methods:

Investigation Site:

The dataset for cow facial recognition was created by recording videos of 37 different cows. To generate diverse, environment-based images, every 10th frame from the videos was extracted, and cow faces were manually annotated using the Roboflow tool. Additionally, annotated images from publicly available web-based datasets were incorporated to enhance generalization and achieve higher detection accuracy. The fine-tuning of the pre-trained YOLOv8 model was performed in a simulation environment using this combined dataset.

Methodology:

The proposed methodology consists of several key steps, as outlined in Figure 1.

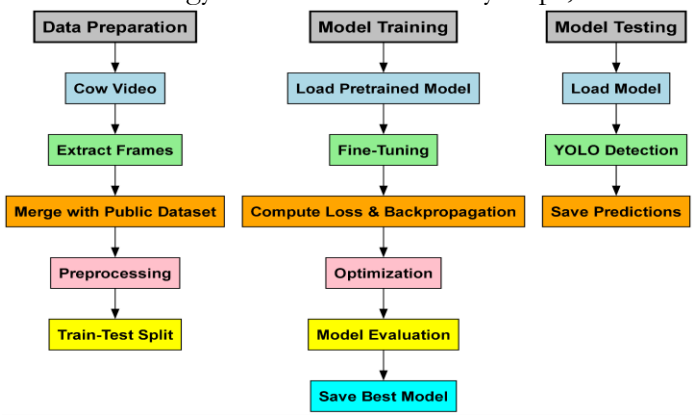


Figure 1. Workflow for Cow Face Detection using YOLOv8: Input videos are processed through frame extraction and annotated for training. The Yolov8 model is fine-tuned and optimized for model evaluation. Model is loaded and tested with web-based images.

Dataset:

Videos of 37 cows were recorded from multiple angles under different lighting and movement conditions. From these videos, 1,078 frames were extracted at regular intervals, with every 10th frame selected to form the initial dataset. To increase the dataset's size and diversity, 1,410 annotated images were added from a publicly available dataset on Roboflow. This brought the total number of images to 2,488 before preprocessing, with three distinct classes. To maintain uniformity, the external dataset classes were consolidated into a single class, labeled *CowFace*, ensuring consistency across all images. The dataset was then preprocessed to prepare it for training. All images were resized to 640×640 pixels to standardize their dimensions while preserving key features. After resizing, the dataset was split into three subsets: 1,800 images for training, 560 for validation, and 86 for testing. To enhance the training set's variability, data augmentation techniques were applied, including horizontal

flips, zoom cropping (0–20%), rotations (-15° to +15°), and shear transformations (±10° both horizontally and vertically). These augmentations generated three additional versions of each image, resulting in a total of 5,096 images after preprocessing and augmentation. This process ensured a diverse dataset, optimized for effective model training.

YOLOv8 Architecture:

The basic architecture of the original YOLO version is illustrated in Figure 2.

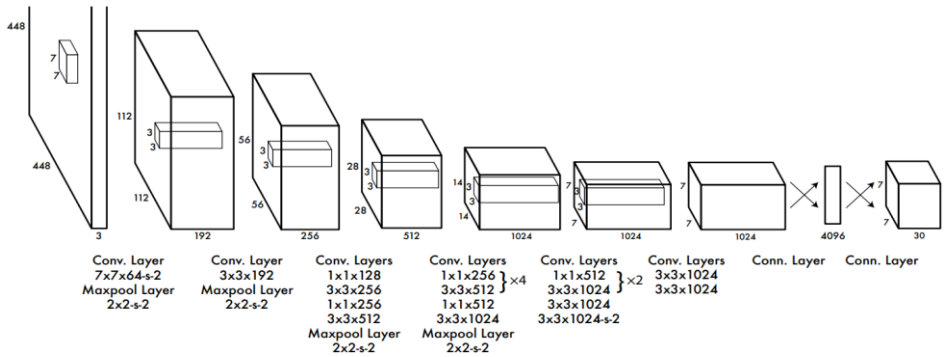


Figure 2. The YOLOv1 architecture [13] processes an input image through sequential convolutional layers for feature extraction, reduces dimensions with max-pooling, and outputs object detection predictions through fully connected layers.

YOLOv1 Architecture:

The original YOLOv1 architecture [13], proposed by Joseph Redmon et al., starts with a 7×7 convolutional layer (stride 2) for large-scale feature extraction, followed by a 2×2 max-pooling layer (stride 2). Next, the network uses alternating 1×1 and 3×3 convolutional layers to balance feature extraction and dimensionality reduction, incorporating a total of 24 convolutional layers. Finally, two fully connected layers are added, with the last layer producing a 7×7×30 output. This output includes bounding boxes, confidence scores, and class probabilities. The design integrates object detection into a single neural network to improve efficiency. Over time, several enhancements led to the development of YOLOv8, a more advanced and stable version that integrates community-driven ideas and modern techniques for better performance and flexibility. While no official paper on YOLOv8 has been published yet, the available architecture can be found in [14] and is illustrated in Figure 3.

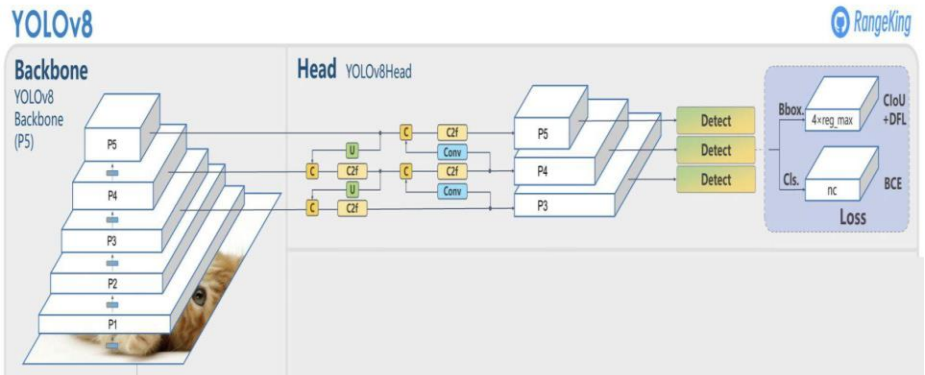


Figure 3. YOLOv8 predicts object centers directly (anchor-free), making it faster and simpler. It uses a new C2f module for better feature extraction and reduces model size with efficient design tweaks. Mosaic data mixing is used early in training but stopped later for accuracy.

YOLOv8:

YOLOv8 brings several key improvements that enhance both its performance and efficiency. It predicts object centers directly, eliminating the need for anchor boxes. This simplifies the model, making it faster and more efficient. Additionally, YOLOv8 includes a

C2f module, which is optimized to extract features more effectively from frames, even in complex environments. The model is scalable and works well on both high-performance GPUs and edge devices.

Implementation of YOLOv8:

YOLOv8 has shown outstanding performance in detecting cow faces across different test datasets and real-world web images. Its high precision is achieved by using transfer learning with COCO weights, followed by fine-tuning on a custom dataset. This dataset, created from video frames and augmented Roboflow images, captures variations in lighting, angles, and environmental conditions, ensuring reliable generalization. Trained over 50 epochs at a 640×640 resolution with a batch size of 16, the model focuses on single-class detection to streamline its accuracy on cow faces. It has already outperformed SSD, Faster R-CNN, and earlier YOLO versions in both speed and accuracy. Thanks to its real-time detection capabilities and efficient computational design, YOLOv8 is an ideal solution for livestock monitoring systems. By balancing data diversity with optimized architecture, it provides scalable and high-performance results. The implementation of YOLOv8 for cow face detection is based on three key mathematical principles: loss optimization, weight updates, and bounding box prediction. The YOLO loss function combines bounding box regression, object confidence, and classification loss to improve detection accuracy.

YOLO Trainer Framework:

The YOLO Trainer Framework fine-tunes YOLOv8 models to detect cow faces, ensuring a streamlined workflow for training, verification, and accurate area estimation of the detected cow face. This framework also automates dataset management, model training, and evaluation while saving the best-performing models for improved accuracy. Additionally, during system training, it records matrices to estimate test images, making it a reliable and effective solution for cow face detection. Table 1 below provides details on the hyperparameters used in the code.

Table 1. The YOLOTrainer fine-tunes YOLOv8 for cow face detection using epochs=50, batch=16, imgsz=640, workers=8, optimizer="SGD", patience=10, and pretrained=True. It organizes the project structure, trains on data.yaml, and saves the best model. Table 1 is showing details of hyperparameters

Hyperparameter	Value	Description
pretrained_model	"yolov8n.pt"	Pretrained YOLO model used for training
data	self.data_path	Path to the dataset configuration file (data.yaml)
epochs	50	Number of training epochs
batch	16	Batch size during training
imgsz	640	Input image size (in pixels)
workers	8	Number of worker threads for data loading
optimizer	"SGD"	Optimization algorithm used for training
patience	10	Number of epochs with no improvement before early stopping
pretrained	True	Whether to use a pretrained model
project	self.results_dir	Directory to store training results
name	"fine_tune_coco"	Experiment name for saving results
pretrained_model	"yolov8n.pt"	Pretrained YOLO model used for training
data	self.data_path	Path to the dataset configuration file (data.yaml)

Results:

YOLOv8-based facial recognition has successfully detected cow faces from test datasets and real-world web images. The model was fine-tuned using a custom dataset that included video frame captures and additional images from a publicly available Roboflow

dataset. This dataset was selected for its diversity, featuring variations in angles, lighting conditions, and cow face characteristics. The fine-tuning process involved 50 training epochs, a batch size of 16, and image resizing to 640×640 pixels, aligning with the preprocessing steps. YOLOv8's flexibility in handling single-class detection and its compatibility with augmented datasets made it a perfect fit for this research. Compared to earlier YOLO versions and other object detection frameworks like SSD and Faster R-CNN, YOLOv8 demonstrated superior inference speed and overall performance.

Figure 4. Left: Manually labelled image of cow faces; Right: YOLOv8 prediction of the same image. The model, fine-tuned on a diverse custom dataset, accurately detects cow faces under varying conditions, as shown in the comparison.

The trained model was also tested with various online images. Figure 5 below illustrates the model's functionality, demonstrating its ability to generalize and detect any cow face effectively.

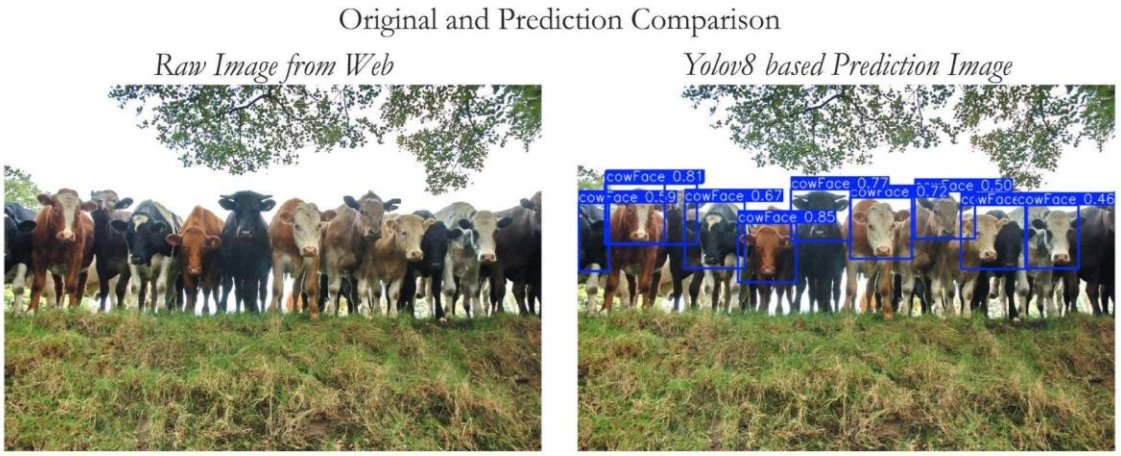


Figure 5 shows a raw image sourced from the web on the left side, representing the original picture. After processing, the fine-tuned model accurately detects and labels the cow face, with the processed image displayed on the right side. The model's ability to generalize on smaller datasets without overfitting enhances its practical utility. Additionally, its streamlined deployment and real-time prediction capabilities make it a powerful tool for livestock monitoring. The combination of a robust architecture, transfer learning using COCO weights, and a fine-tuned dataset ensured YOLOv8's high performance in detecting cow faces, proving its superiority over conventional methods.

Figure 6 illustrates the progression of key performance metrics during training. Precision stabilizes at around 0.976, reflecting a low false-positive rate. Recall converges near

0.955, indicating the model's effectiveness in detecting relevant objects. The mAP50 metric reaches an average of 0.981, highlighting the model's strong detection accuracy at an IoU threshold of 0.5. Together, these metrics demonstrate the YOLOv8 model's high reliability in object detection tasks.

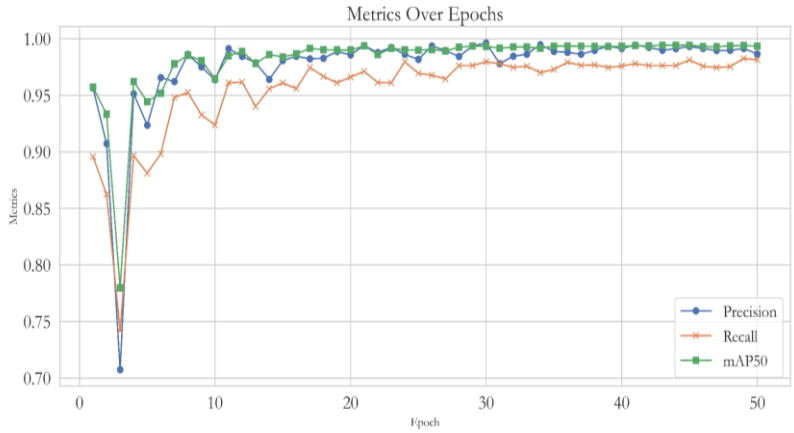


Figure 6. Precision, recall, and mAP50 metrics stabilize at high values, demonstrating consistent, reliable, and well-balanced object detection performance.

Figure 7 illustrates the trends of box loss, classification loss, and distributional focal loss for both training and validation datasets. The gradual decrease in training losses indicates the model's ability to effectively learn and adapt to the dataset. Although validation losses are slightly higher, suggesting minor generalization challenges, the close alignment of the training and validation curves indicates minimal overfitting. This demonstrates that the model successfully captures essential features while maintaining generalization.

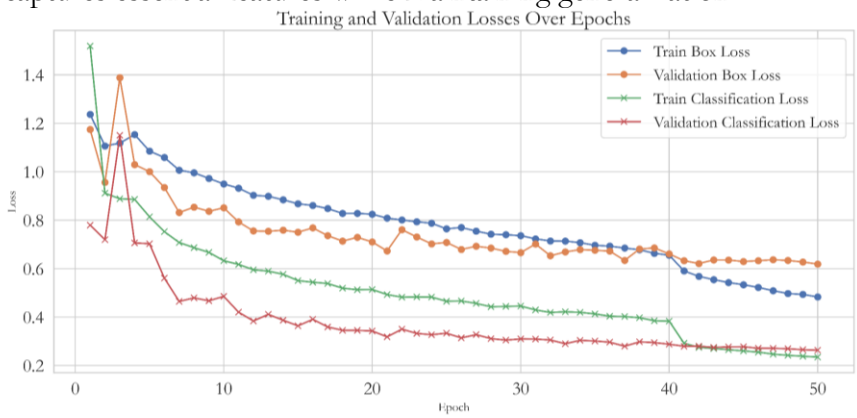


Figure 7 Training and validation losses show a steady decline, indicating effective learning with minimal overfitting, despite slightly higher validation losses

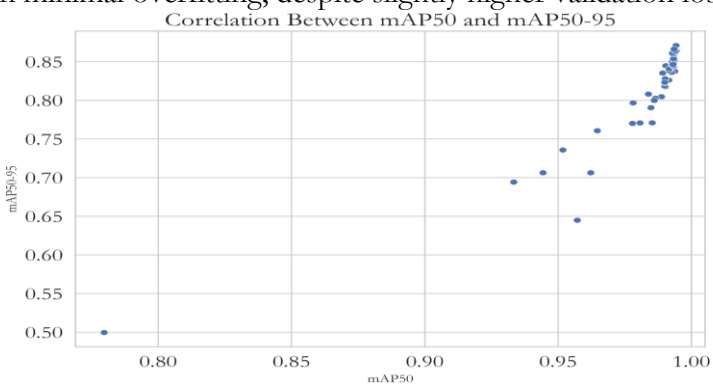


Figure 8. A clear linear relationship between mAP50 and mAP50-95 indicates consistent detection accuracy across different IoU thresholds.

The scatterplot in Figure 8 shows a strong linear correlation (≈ 0.98) between mAP50 and mAP50-95. This indicates that improving detection accuracy at a 0.5 IoU threshold also enhances performance across stricter thresholds. The near-linear pattern highlights the YOLO model's robustness in handling varying levels of localization strictness. Performance metrics are calculated after each training epoch to assess the model's detection ability. Key metrics include:

- **Precision**: The proportion of true positives out of all predicted positives,
- **Recall**: The proportion of true positives out of all actual positives, and
- **mAP50**: Mean average precision at a 0.5 IoU threshold.

The mAP50-95 metric provides a broader evaluation by measuring performance across a range of IoU thresholds (from 0.5 to 0.95). Additionally, validation losses, computed using unseen data, help assess how well the trained model can generalize to new data.

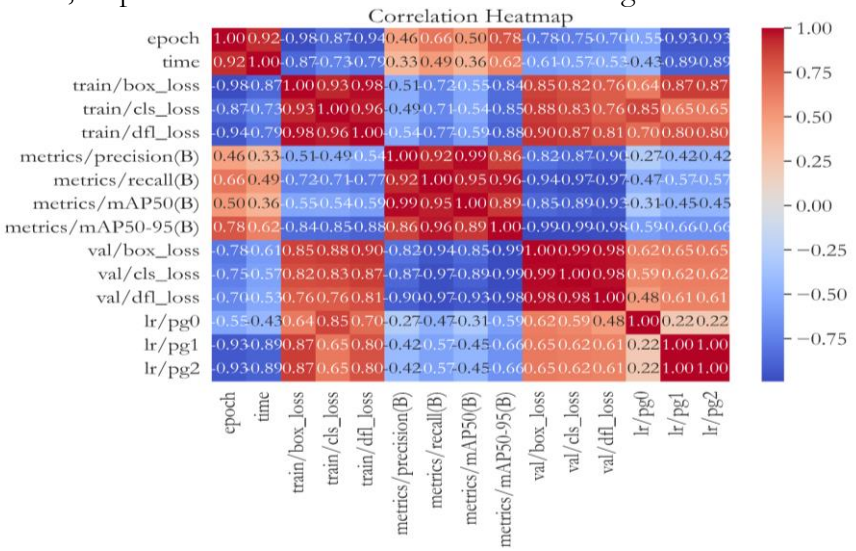


Figure 9. The correlation heatmap highlights key metric interdependencies, showing how reduced losses improve detection accuracy and reflecting the optimization behavior of learning rates.

The heatmap in Figure 9 reveals relationships between key metrics, losses, and learning rates. It shows strong negative correlations between box loss and mAP50/mAP50-95, meaning that lower localization errors lead to higher detection accuracy. Positive correlations between precision and recall demonstrate their interdependence, where enhancing one tends to improve the other. Learning rate parameters (e.g., lr/pg0, lr/pg1, and lr/pg2) exhibit weak direct correlations with precision and recall, indicating that they primarily optimize model weights rather than directly affecting evaluation metrics. Overall, the heatmap provides valuable insights into the optimization dynamics and performance interdependencies of the model. The results confirm the YOLO model's high precision and robustness. The strong correlation between mAP50 and mAP50-95 highlights the model's ability to maintain consistent performance across various IoU thresholds, making it well-suited for applications requiring high detection accuracy. Deviations in validation losses suggest potential areas for further improvement, such as applying advanced augmentation techniques or fine-tuning hyperparameters. These findings emphasize the model's applicability to real-world scenarios where accurate and reliable object detection is critical.

Table 2 presents an analysis of training and validation losses alongside performance metrics, demonstrating the model's effectiveness in object detection. Training box loss steadily decreased across epochs, with validation box loss following a similar downward trend. Slightly higher validation loss values indicate good generalization with minimal overfitting. Classification losses for both training and validation also decreased consistently, reflecting the

model's growing accuracy in object classification. Performance metrics further underscore the model's robustness, with precision averaging 0.976, recall at 0.955, and mAP50 reaching an impressive 0.981, which reflects excellent detection accuracy at an IoU threshold of 0.5. Additionally, the mAP50-95 average of 0.812 demonstrates reliable performance across varying IoU thresholds. Gradual reductions in learning rates across parameter groups (lr/pg0, lr/pg1, lr/pg2) contributed to smooth convergence, preventing loss oscillations and ensuring stable training progress.

Discussion:

The findings of this study highlight significant advancements in improving model generalization and real-world applicability through strategic data augmentation and dataset diversity. By fine-tuning YOLOv8 using a combination of self-captured images and publicly available datasets, the model demonstrated enhanced robustness in detecting cow faces under varied lighting, angles, and background conditions. This is a notable improvement compared to earlier object detection frameworks, such as SSD, Faster R-CNN, and YOLOv4, which have been reported to struggle with overfitting when trained on limited datasets. The integration of COCO-pretrained weights further improved YOLOv8's ability to generalize to complex environments, making it more effective and scalable for large-scale livestock monitoring applications. YOLOv8 was chosen over conventional methods like R-CNN and older YOLO versions due to its superior speed and efficiency. Its enhanced backbone structure and anchor-free detection improve object localization, especially in complex environments. Compared to earlier YOLO versions, YOLOv8 offers higher precision, better recall, and faster real-time performance.

Self-captured images alone are insufficient for generalization due to the similarity in environmental conditions. To reduce overfitting and enhance generalization, publicly available data was incorporated into the dataset. Fine-tuning further improved performance by enabling the model to learn dataset-specific features, thereby reducing false positives and improving the accurate detection of cow faces amid complex background elements. YOLOv8 is also more robust than older versions, effectively handling variations in lighting, occlusion, and pose. Fine-tuned models generally perform better in real-world applications compared to generic pretrained models, making YOLOv8 a suitable choice for practical deployment.

In addition to generalization, YOLOv8's comparative performance offers practical implications for precision livestock management. The model's anchor-free detection mechanism and advanced C2f module significantly enhance object localization, resulting in reduced computational overhead, faster inference speed, and improved accuracy. These advantages not only make YOLOv8 suitable for deployment on high-performance GPUs but also adaptable for edge devices, which are increasingly used in smart farms. By achieving high precision, recall, and mAP scores, the proposed approach enhances the accuracy of cow face detection, which can help farmers manage animal health, monitor feeding patterns, and resolve disputes in mixed-herd environments. While the model shows promise, future refinements could further enhance its practical utility and societal impact. Exploring multi-class detection to differentiate between various cattle breeds or integrating temporal analysis to track individual cows over time could offer additional functionalities. Furthermore, incorporating this model into broader Internet of Things (IoT)-based farm management systems could enable real-time livestock monitoring and automated decision-making. Ethical considerations, including animal stress from continuous surveillance and privacy concerns regarding farm data collection, should also be addressed in future implementations. These enhancements would pave the way for a fully autonomous, scalable livestock monitoring system, contributing to increased productivity, reduced labor requirements, and improved animal welfare.

Table 2. shows a detailed breakdown of training and validation losses, highlighting steady improvements in object localization and classification accuracy. Key metrics—precision (0.976), recall (0.955), and mAP50 (0.981)—demonstrate the model's strong detection capability. The gradual learning rate reductions further facilitated smooth convergence and stable training progress.

epoch	train/box _loss	train/cls _loss	metrics/precision (B)	metrics/recall (B)	metrics/ mAP50(B)	metrics/mAP 50-95(B)	val/box _loss	lr/pg0	lr/pg1	lr/pg2
1	1.237	1.519	0.956	0.89581	0.95719	0.64492	1.1744	0.0700	0.003324	0.003324
2	1.106	0.911	0.907	0.86255	0.93337	0.69425	0.9560	0.0399	0.006525	0.006525
3	1.118	0.888	0.707	0.74189	0.77968	0.4997	1.3886	0.0096	0.009595	0.009595
4	1.153	0.885	0.951	0.89676	0.96222	0.70639	1.0297	0.0094	0.009406	0.009406
48	0.497	0.242	0.989	0.97536	0.99387	0.86499	0.6339	0.0006	0.000694	0.000694
49	0.493	0.238	0.991	0.98266	0.99399	0.86363	0.6274	0.0004	0.000496	0.000496
50	0.483	0.235	0.986	0.98136	0.9935	0.86628	0.6191	0.0002	0.000298	0.000298

Conclusion.

The implementation of YOLOv8 for face detection marks a key step toward accurate cow recognition using facial features. Similar to human face recognition, cow face detection holds significant potential for identifying individual animals. By fine-tuning the model, raw background information is filtered out, allowing the system to focus solely on the face, which enhances classification accuracy. This study represents the initial step toward developing a precise cow identification system aimed at tracking and monitoring individual animal performance. A transfer learning-based approach was successfully implemented using YOLOv8 to detect cow faces in diverse farm environments. The model was fine-tuned with a carefully curated dataset that combined frames from video footage with publicly available annotated images. The system demonstrated outstanding performance, achieving a precision of 0.976, a recall of 0.955, and an mAP50 of 0.981, indicating its high accuracy in detecting cow faces. The model's consistent performance improvement across epochs, along with reduced box losses, underscores its robustness. Data augmentation and fine-tuning on a diverse dataset significantly enhanced generalization, while the use of early stopping minimized the risk of overfitting.

Additionally, the strong correlation between mAP50 and mAP50-95 highlighted the model's reliability across varying object localization thresholds. These findings confirm YOLOv8 as an effective tool for real-time face detection in livestock localization and monitoring applications. This approach not only establishes a foundation for accurate cow face detection but also paves the way for future advancements in classification tasks. Future work may focus on refining the model further and integrating it with additional technologies to enhance cow face recognition and improve individual cow identification in practical settings.

Acknowledgement. The authors would like to thank the administration of Intelligent Systems Laboratory, Department of Electrical Engineering, University of Gujrat for providing an opportunity to conduct research and perform experiments with the available laboratory facilities.

Author's Contribution. Umair Ali wrote the initial draft of the article. Wasif Muhammad supervised and drafted the final version of the manuscript.

Conflict of interest. The authors have no conflict of interest for publishing this manuscript in IJIST.

References:

[1] P. L. P. J. Zarco-Tejada, N. Hubbard, "Precision Agriculture: An Opportunity for EU-Farmers – Potential Support with the CAP 2014-2020," Think Tank. Accessed: Mar. 13, 2025. [Online]. Available: https://www.europarl.europa.eu/thinktank/en/document/IPOL-AGRI_NT(2014)529049

[2] S. K. Seelan, S. Laguette, G. M. Casady, and G. A. Seielstad, "Remote sensing applications for precision agriculture: A learning community approach," *Remote Sens. Environ.*, vol. 88, no. 1–2, pp. 157–169, 2003, doi: https://doi.org/10.1016/j.rse.2003.04.007.

[3] P. G. António Monteiro, Sérgio Santos, "Precision Agriculture for Crop and Livestock Farming—Brief Review," *Animals*, vol. 11, no. 8, p. 2345, 2021, doi: https://doi.org/10.3390/ani11082345.

[4] Ali Ismail Awad, "From classical methods to animal biometrics: A review on cattle identification and tracking," *Comput. Electron. Agric.*, vol. 1123, pp. 423–435, 216AD, doi: https://doi.org/10.1016/j.compag.2016.03.014.

[5] L. X. Liu Bo, Liu Yuefeng, Bao Xiang, Wang Yue, Liu Haofeng, "Research on Dairy Cow Identification Methods in Dairy Farm," *Indian J. Anim. Res.*, vol. 52, no. 12, pp. 1733–1739, 2023, doi: 10.18805/IJAR.BF-1660.

[6] A. M. Johnston and D. S. Edwards, "Welfare implications of identification of cattle by ear tags," *Vet. Rec.*, vol. 138, no. 25, pp. 612–614, Jun. 1996, doi: 10.1136/VR.138.25.612.

[7] D D Wardrope, "Problems with the use of ear tags in cattle - PubMed," Vet Rec. Accessed: Mar. 13, 2025. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/8966978/

[8] V. I. K. and K. M. D. A N Ruchay, K A Dorofeev, V V Kalschikov, "A depth camera-based system for automatic measurement of live cattle body parameters," *IOP Conf. Ser. Earth Environ. Sci.*, vol. 341, 2019, doi: 10.1088/1755-1315/341/1/012148.

[9] M. S. V. Maria Vlad, Roxana Andreea Parvulet, "A Survey of Livestock Identification Systems," Proceedings of the 13th WSEAS International Conference on Automation and Information (ICAI12). Accessed: Mar. 13, 2025. [Online]. Available: https://www.researchgate.net/publication/265227039_A_Survey_of_Livestock_Identification_Systems

[10] P. Thanapol, K. Lavangnananda, P. Bouvry, F. Pinel, and F. Leprevost, "Reducing Overfitting and Improving Generalization in Training Convolutional Neural Network (CNN) under Limited Sample Sizes in Image Recognition," *InCIT 2020 - 5th Int. Conf. Inf. Technol.*, pp. 300–305, Oct. 2020, doi: 10.1109/INCIT50588.2020.9310787.

[11] S. M. A. Imran Khan, Chen Tianhua, Khan Malak Abid Ali, "Object Recognition in Different Lighting Conditions at Various Angles by Deep Learning Method," arXiv preprint arXiv:2210.09618. Accessed: Mar. 13, 2025. [Online]. Available: https://www.researchgate.net/publication/357702329_Object_Recognition_in_Different_Lighting_Conditions_at_Various_Angles_by_Deep_Learning_Method

[12] X. Lei, X. Wen, and Z. Li, "A multi-target cow face detection model in complex scenes," *Vis. Comput.*, vol. 40, no. 12, pp. 9155–9176, Dec. 2024, doi: 10.1007/S00371-024-03301-W/METRICS.

[13] A. Redmon, J., Divvala, S., Girshick, R., Farhadi, "You only look once: Unified, real-time object detection," *2016 IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, doi: 10.1109/CVPR.2016.91.

[14] J. Solawetz and Francesco, "What is YOLOv8? A Complete Guide," Yolov8. Accessed: Mar. 13, 2025. [Online]. Available: https://blog.roboflow.com/what-is-yolov8/

Steering Control of Ackermann Architecture Weed Managing Mobile Robot

Faryal Naeem Mehmood[1], Syeda Ambreen Zahra[1], Syed Muhammad Wasif[1*], Zubair Mehmood[1], Muhammad Jehanzeb Irshad[1], Nazam Siddique[1].
[1] Intelligent Systems Laboratory, Department of Electrical Engineering, University of Gujrat, Pakistan.
*Dr. Syed Muhammad Wasif: syed.wasif@uog.edu.pk

Citation | Mehmood. F. N, Zahra. S. A, Wasif. S. M, Mehmood. Z, Irshad. M. J, Siddique. N, "Steering Control of Ackermann Architecture Weed Managing Mobile Robot", IJIST, Special Issue. pp 134-145, March 2025
Received | Feb 18, 2025 **Revised |** March 02, 2025 **Accepted |** March 08, 2025 **Published |** March 11, 2025.

A robot designed to identify and remove weeds from crops is known as a weed control robot. Weeds compete with primary crops for moisture, hinder their growth, and may harm both human and animal health, leading to reduced crop yields. Traditionally, herbicides and other chemicals have been used to eliminate weeds, but these methods can damage crops and pollute the environment. In this work, we propose a new semantic weed detection method based on the PC/BC-DIM network, which demonstrates superior performance and classification accuracy compared to existing approaches. We developed an autonomous weed control robot incorporating Ackermann Architecture and a delta robot. The delta robot is equipped with a camera at its base to detect weeds in real-time. First, the robot captures images using the camera, and through image processing techniques, it differentiates weeds from crops. Detected weeds are then eliminated using an electrical discharge method, where electrodes attached to the robot's end effector burn the targeted weeds. Additionally, we developed a path-planning and obstacle-avoidance system to help the mobile robot navigate the field. This system uses stereo vision to capture stereo images of the environment and calculate their disparity. By extracting depth information, the robot can detect obstacles, avoid them, and follow the shortest path using the A* algorithm. The results from this work are simulation-based, demonstrating effective weed detection in field images and efficient robot navigation using stereo images. The system achieved an overall accuracy of 81.25%. Although the system performs moderately well, the relatively high False Positive Rate and Root Mean Square (RMS) Error indicate the need for further improvements to reduce errors and false positives.
Future work will focus on enhancing weed removal and implementing the simulated results on hardware.

Keywords: Ackermann Steering, Weed Detection, Mobile Robots, Path Planning, Stereo Vision, Image Processing.

Introduction:

A robot designed to identify and eliminate weeds from crops is known as a weed control robot. Weeds are unwanted plants that compete with crops for nutrients, moisture, and space, which hinders crop growth and reduces yields. They can also harm human and animal health, contributing to significant crop losses [1]. Due to these negative impacts, farmers must remove weeds, a task that is both time-consuming and labor-intensive. For many years, herbicides and chemicals have been used to eliminate weeds, but these substances can harm crops and pollute the environment. To speed up farming operations and reduce manual labor, farmers often use equipment such as tractors, weeders, and harvesters. While some farmers can afford such machinery, others cannot, especially since these machines require expensive fuel and contribute to environmental pollution. For these reasons, relying solely on heavy equipment may not be the best solution [2]. Farmers also use fertilizer sprayers to boost crop growth and yield, but this process requires time and effort. To address these challenges, robotic weed control offers a promising solution by preventing weeds from disrupting crops and improving farming efficiency.

Many weed control robots can perform tasks like electrical discharge, mechanical weeding (using hoe tools), and targeted chemical spraying [3]. This project will focus on the electrical discharge method, which can eliminate weeds in two ways: continuous contact and spark discharge. The first method uses short bursts of high-voltage electricity to kill weeds, speed up fruit ripening, and thin plant growth. The second method delivers energy in brief pulses (e.g., one microsecond) using two electrodes positioned on opposite sides of the plant. This pulse thins plant rows, kills weeds, cuts plant sections, and dries out the leaves of root crops. The plant tissue is damaged either by the sudden electrical surge or the heat generated by the electricity.

This project will also implement Ackermann architecture, which is based on a four-wheel independent steering system. It includes several steering modes:

- **Ackermann steering:** Allows the inner and outer wheels to rotate at different radii.
- **Active front and rear steering:** Turns the front wheels in the opposite direction of the rear wheels.
- **Crab steering:** Moves all wheels in the same direction to allow diagonal movement.
- **Spinning:** Rotates the vehicle around a central point.

This system allows the robot to switch between steering modes depending on the situation, making movement more efficient. The Ackermann steering structure solves the issue of different steering angles caused by varying radii of the left and right wheels. According to Ackermann's steering geometry [4], by adjusting the crank of the four-link structure, the robot can increase the inner wheel's steering angle by 2-4° more than the outer wheel when turning along a curve. This adjustment helps position the robot's steering center, allowing smooth turns by aligning the four-wheel paths with the rear axle's extension line.

Ackermann steering is known for supporting high payloads and improving movement efficiency. It is commonly used in cars, although the structure tends to be too large for narrow spaces. This project will enhance the robot's functionality using image processing and machine vision techniques. Cameras installed on the robot will capture images of the farmland. These images will be processed using algorithms that identify weeds based on their unique characteristics. Additionally, a stereo camera will capture images from two angles to analyze the environment, detect free spaces, plan the robot's path, and guide its movement.

Once weeds are detected, the robot's control system will instruct its mechanical arm or tool to remove them from the ground. A previous group worked on an autonomous weed control robot but left several limitations:

1. The robot detected weeds in a virtual environment rather than a real one.
2. A weed removal mechanism was not developed.

3. There was no navigation system for the mobile robot.

The main aims of this research are:

- To develop a system that can distinguish between weeds and crops in cotton fields using image processing techniques.
- To build a system for weed removal in cotton fields using electrical discharge.
- To create a navigation mechanism for the weed control mobile robot.

Literature Review:

The development of autonomous agricultural robots for weed management has gained significant attention in recent years, particularly as a means to reduce chemical herbicide dependence and enhance precision farming efficiency. Traditional weed control methods rely heavily on manual labor or herbicide spraying, both of which present economic and environmental drawbacks [1]. Autonomous robotic systems equipped with computer vision, deep learning, and mechanical weed removal mechanisms offer a sustainable alternative, allowing for precise identification and targeted elimination of weeds without harming crops. This literature review explores previous advancements in weed detection, robotic navigation, and path planning to establish the significance of the proposed study.

With precision farming, autonomous robotic weeding systems have proven to be effective in reducing the use of agrochemicals like pesticides and herbicides. A study [5] proposes a multi-camera, non-overlapping approach to enhance the weed control system's flexibility in managing unknown classification delays. Consequently, an advanced weed-control technique [6]becomes necessary. In this approach, images of plantation rows are captured at regular intervals using image processing methods.

In [7], the author developed a low-cost delta robot arm equipped with a vision system, capable of gripping objects of various sizes. This economical design uses stepper motors instead of AC servo motors. In [1], an autonomous agricultural mobile robot for outdoor mechanical weed control is introduced. This robot operates with two vision systems: a color-based system to distinguish between weeds and crops, and a gray-level system to detect the row structure created by crops, guiding the robot along the rows. In [8], Kulkarni et al. designed a robotic vehicle with four wheels, steered by a DC motor and equipped with an IR sensor system to manage weed growth in fields. In [9], the authors developed a new technology for weed control called "crop signaling." This method enables plants and weeds to be machine-readable, allowing them to be automatically distinguished based on their unique features.

The effectiveness of robotic weed management heavily relies on accurate weed classification using image processing and machine learning techniques. Early studies employed handcrafted feature extraction methods, such as color thresholding and edge detection, to differentiate between crops and weeds. However, these methods were often limited by variations in lighting, soil conditions, and plant morphology. Recent advancements in deep learning, particularly Convolutional Neural Networks (CNNs) and Transformer-based architectures, have significantly improved weed detection accuracy. Most modern approaches rely on pre-trained deep learning models (e.g., ResNet, YOLO, EfficientNet) trained on large-scale agricultural datasets for real-time weed identification (Lin et al., 2023) [4] . However, deep learning models require high computational power and extensive datasets, which may limit their usability in field-deployable robots with constrained hardware. This study explores Predictive Coding-Biased Competitive Divisive Input Modulation (PC/BC-DIM) neural networks for weed classification, a less common approach in precision agriculture. While previous studies have demonstrated the effectiveness of CNNs in weed classification, the proposed study seeks to evaluate whether PC/BC-DIM can offer a computationally efficient alternative while maintaining high classification accuracy.

In this study, we present an autonomous weed control mobile robot that integrates Ackermann steering architecture with Delta robot-based weed removal. Unlike previous

approaches, our system uses real-time stereo vision-based disparity mapping, followed by A* algorithm-based path planning for effective obstacle avoidance. Additionally, a PC/BC-DIM neural network is employed to enhance the accuracy of weed detection. The proposed system achieves an accuracy rate of 81.25%, making it a promising solution for precision agriculture. This work bridges the gap between simulation-based weed classification and real-world navigation, paving the way for autonomous and efficient weed management.

Material and Methods:

Weed detection using an autonomous mobile robot is a complex challenge. This study focuses on designing a system for weed detection and removal in agricultural fields. The project leverages computer vision techniques, including image processing, to achieve this goal. The mobile robot is equipped with two cameras: one dedicated to weed detection through specialized image processing, and a stereo camera that facilitates efficient path planning.

The Delta robot, mounted on the Ackermann architecture of the mobile robot, navigates through the field and removes weeds after detection using an electrical discharge technique via electrodes attached to its end-effector. The system's overall functionality is illustrated in the block diagram shown in Figure 1.

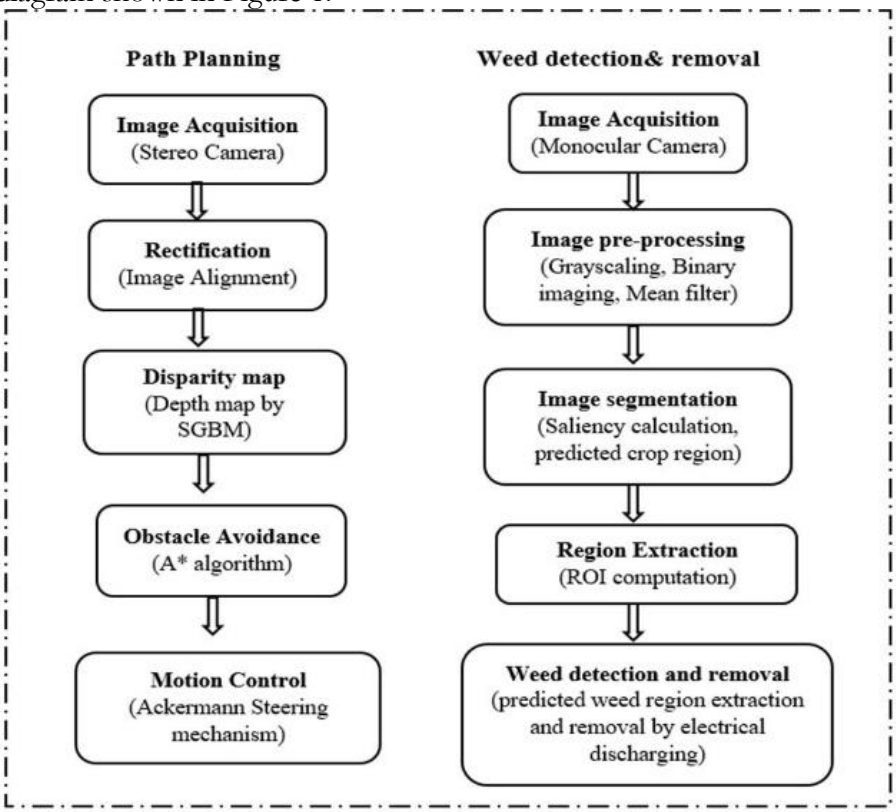


Figure 1. Block diagram of weed control robot

Path Planning:

Path planning for robot navigation follows several key steps. The process starts by capturing stereo images of the environment using a stereo camera. After capturing, the images are rectified to align both on a common plane. A disparity map is then created, showing the difference in the object's position between the two images. This is followed by calculating the depth map using the SGBM algorithm. Based on the depth map, the robot detects and avoids obstacles and plans a clear path using the A* algorithm. Finally, motion control based on an Ackermann steering architecture helps the robot follow the planned path accurately.

A stereo camera setup is placed at the front to capture images for navigation. It takes two images of the same scene from slightly different angles, mimicking human binocular

vision. These images are processed further. Image rectification simplifies the matching process by aligning corresponding points along the same row. First, camera calibration parameters (intrinsic and extrinsic) are calculated. Using these parameters, geometric transformations adjust the images so that corresponding points in both images align horizontally. For example, if a tree appears slightly to the right in the left image and slightly to the left in the right image, rectification aligns the tree in both images on the same horizontal axis.

The disparity map shows the difference in the object's location between the left and right images. Each color in the disparity map represents how near or far an object is—the brighter areas indicate closer objects, while darker areas show objects farther away. The depth map is calculated using the Semi-Global Block Matching (SGBM) algorithm. This algorithm compares small pixel blocks in one image with corresponding blocks in the other image to find the best match by trying different disparities. Small disparities indicate close matches, while large disparities show mismatches.

Path planning and obstacle avoidance are done using the A* algorithm and an occupancy grid based on the depth map. The grid represents obstacles as occupied cells and open space as unoccupied cells. A path from the starting point to the target position is generated using the A* algorithm, which finds the best route by minimizing both the current travel cost and the estimated remaining cost. Additionally, points from the depth map are transformed into a 3D point cloud and projected back onto the 2D image to visualize and detect obstacles. Motion control ensures the robot moves along the designated path. This robot uses the Ackermann steering design, which controls vehicles with differential steering, such as wheeled robots. It calculates precise steering angles and velocities, helping the robot remain stable and maneuverable. The steering angles are determined using depth map data to guide how much the robot should turn to follow the planned path, considering the vehicle's geometry and point coordinates.

Weed Detection:

The semantic weed detection method used in this project is based on our earlier work [10]. The process starts with capturing an image using a secondary camera mounted on the base of a delta robot, above its end effector. This image is then used for further processing to detect weeds. The input image is first converted to grayscale to simplify processing. Next, the grayscale image is converted into a binary image by applying an optimal threshold value, which minimizes the variance between background and foreground (crop-weed areas). Pixels are classified as either background or crop-weed regions. To reduce noise, small objects are removed from the binary image, and a mean filter is applied to smooth the image by averaging pixel values in a specific window size. This filtering highlights the regions of interest and enhances image clarity.

Image segmentation is the next step, dividing the digital image into segments or pixel sets to make the image easier to analyze. Segmentation helps identify objects and boundaries (e.g., lines and curves) within the image. In this case, segmentation labels each pixel to group those with similar attributes.

The segmentation process starts by generating a saliency map using different algorithms. The saliency map highlights key areas that stand out from the background, such as weed regions. An error map is also computed to assess segmentation accuracy. Several saliency maps are generated, combined, and refined to create a comprehensive map that highlights the most important areas (i.e., weeds). A PC/BC-DIM neural network (Predictive Coding-Biased Competitive Divisive Input Modulation) and Gabor filters are used to calculate the saliency map. The PC/BC-DIM network mimics how the human brain processes visual images and consists of three types of neurons:

1. **Reconstruction Neurons (R)**: These neurons use prior information to reconstruct the input image, acting as filters by calculating their activity using synaptic weights.

2. **Error Neurons (E)**: These neurons compute the difference between the original input image and the reconstructed image. A large error suggests that the reconstruction is inaccurate. Element-wise division is used to calculate the error and improve accuracy.
3. **Prediction Neurons (Y)**: These neurons adjust their predictions based on the error data, updating their activity to improve the image reconstruction.

The network generates feature maps by applying Gabor filters at different phases and orientations to capture distinct image features. Through iterative reconstruction and prediction, the network produces a refined saliency map that emphasizes key image areas. The saliency values from the final map help distinguish crops from the background, and a predicted crop region is extracted from the binary image. This predicted crop region is then subtracted from the extracted region of interest (ROI) to isolate the weed region. Finally, the detected weed region is compared with ground truth images to evaluate detection accuracy.

$$g(\sigma,\lambda,\phi,\theta) = \exp\left\{-\frac{\hat{x}^2+2\hat{y}^2}{2\sigma^2}\right\}\left[\cos\left\{\frac{2\pi\hat{y}}{\lambda}+\phi\right\}-\cos(\phi)\exp\left\{-\left(\frac{\pi\sigma}{\lambda}\right)^2\right\}\right] \quad (1)$$

The predicted crop region is labeled in green to make it easily identifiable. The primary objective of feature extraction is to detect and analyze the key areas of the image. This involves isolating the region of interest (ROI), which includes the parts of the image containing crops and weeds while eliminating the background. In the next step, the predicted crop region is subtracted from the ROI image (containing both crops and weeds) to isolate the predicted weed region. The extracted weed region is then labeled for easy visualization, with the weed region highlighted in red. Finally, the labeled predicted crop and weed regions are combined to display the detected weed region, as shown in the results section.

To evaluate the accuracy of weed detection, statistical measures and error calculations are performed by comparing the predicted images with the ground truth images. The term "ground truth" refers to data collected directly from the field, which is essential for verifying image data against actual ground-based characteristics and conditions.

Results and Discussion:

Path Planning Results:

The input image was captured using stereo cameras, as shown in Figure 2. This image includes the left and right stereo views of the garden and ground.

Figure 2. Stereo input images

Figure 3. Rectified stereo input images

Secondly, the rectification of the image mentioned above is performed, as shown in Figure 3. Disparity, which measures the difference in the object's position between the left and right images, is calculated and displayed in Figure 4.

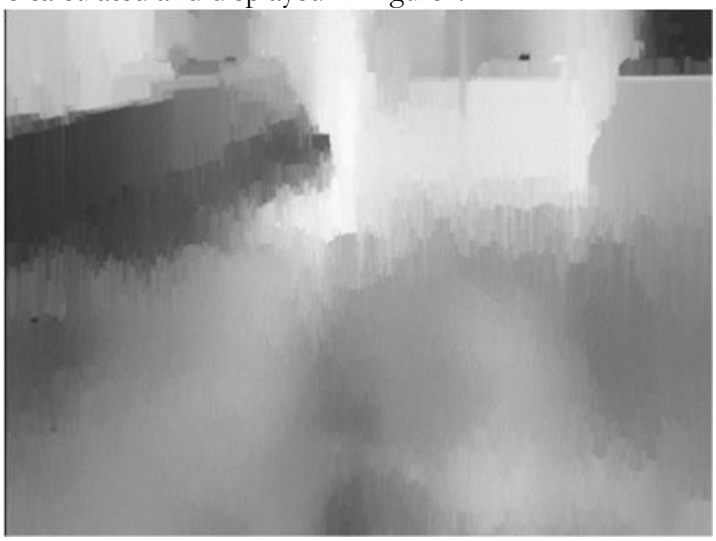


Figure 4. Disparity map

In Figure 5, the robot detects the tree as an obstacle in its path. The colorful line represents the path planned by the robot to avoid the obstacle and navigate through the garden using the A* algorithm.

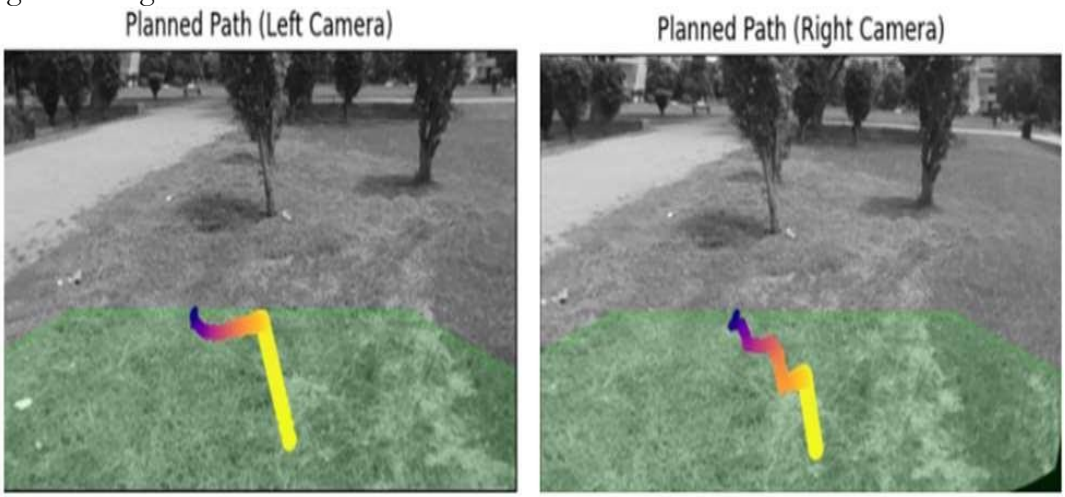


Figure 5. Planned path

Weed Detection Results:

The image captured by the secondary camera is used as the input image, which is then converted to grayscale, as shown in Figure 6.

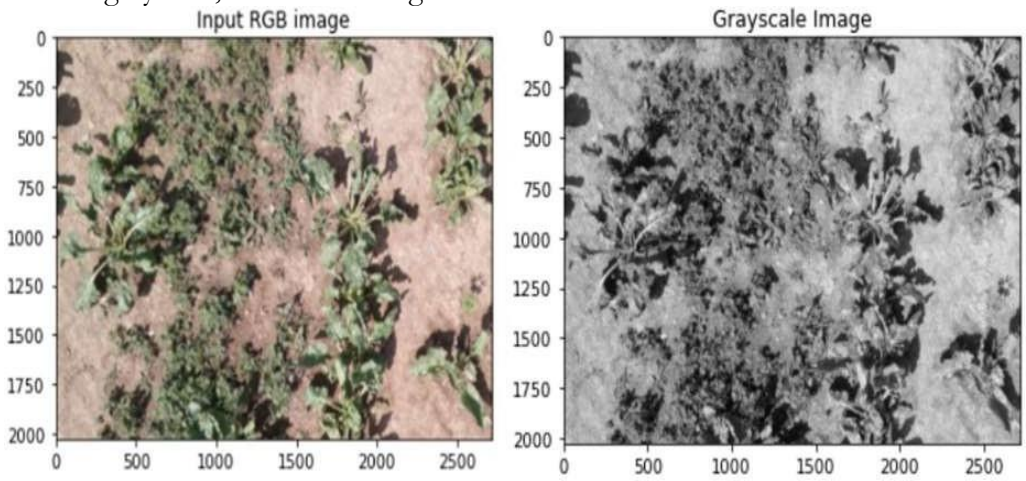


Figure 6. Input and grayscale image

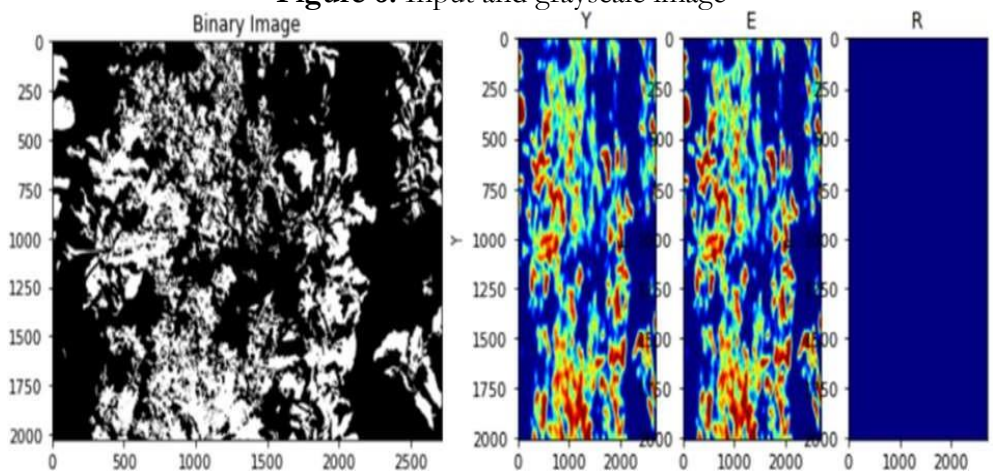


Figure 7. Binary image and saliency map

The grayscale image is then converted into a binary image. In the next step, the saliency map and error map are calculated to highlight the weed region, as shown in Figure 7. Using the saliency maps, the predicted crop region is extracted from the binary image. Once the crop region is detected, it is labeled in green, as shown in Figure 8.

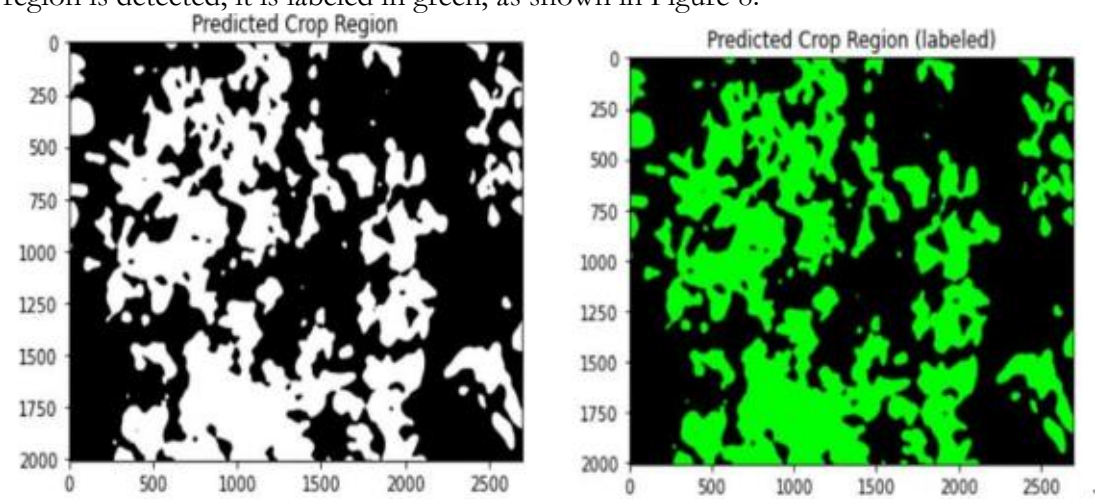


Figure 8. Predicted crop and labelled

In the next step, the ROI (Region of Interest) is extracted from the input RGB image. The predicted crop region is then subtracted from the crop-weed image to separate the weeds from the crops, as shown in Figure 9.

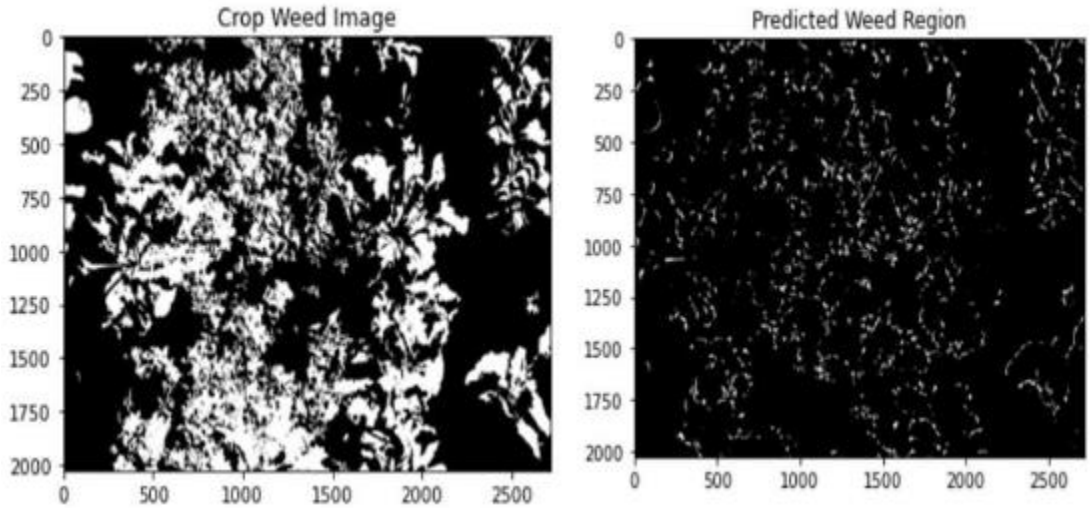


Figure 9. Predicted crop and labelled

After detecting the weed region, it is labeled in red to highlight the extracted area. Both the crop and weed regions are then combined to clearly differentiate between them. The results are shown in Figure 10. The errors are calculated by comparing the predicted image with the ground truth (GT) image, as shown in Figure 11.

The calculations of the false positive rate (FPR) and true negative rate (TNR) help to understand the trade-off between sensitivity (the ability to detect weeds accurately) and specificity (the ability to correctly identify non-weeds).

Three types of errors are computed:

1. **Type 1 (False Positives)**: Non-weeds are incorrectly detected as weeds.
2. **Type 2 (False Negatives)**: Weeds are overlooked or not detected.
3. **Type 3 (Total Error)**: The combined error, accounting for both false positives and false negatives.

The efficacy of the weed detection model is demonstrated in Figure 12.

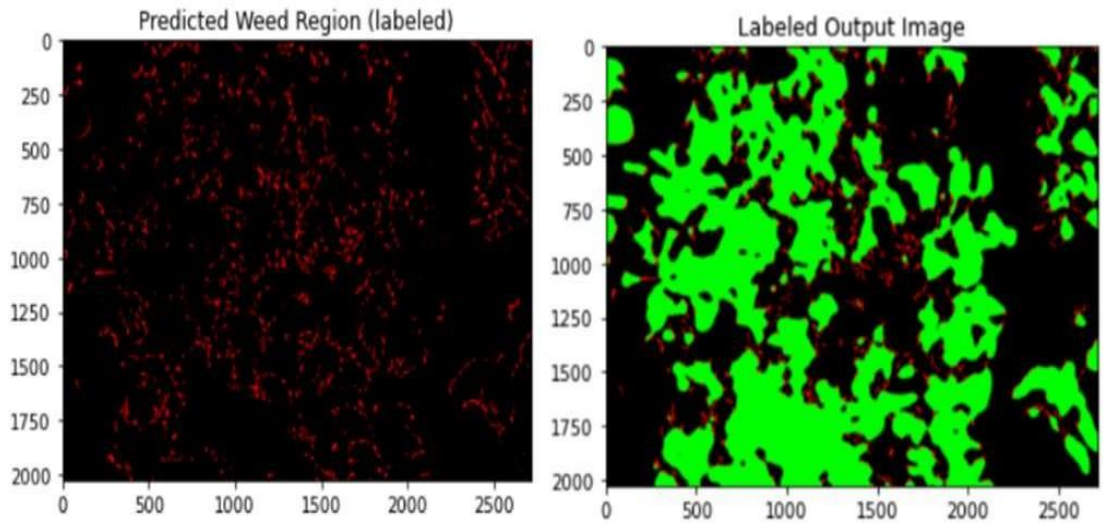


Figure 10. Predicted weed region and labelled

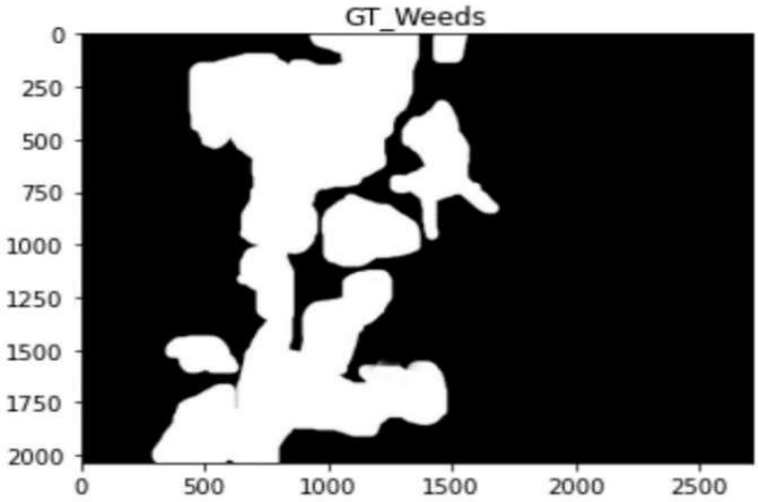


Figure 11. Ground truth weed image

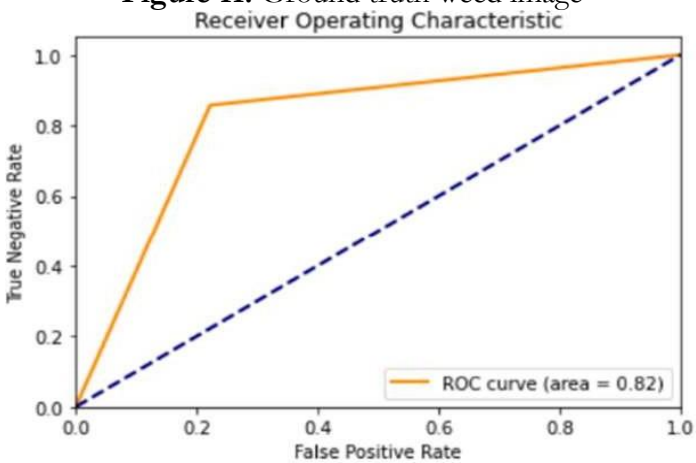


Figure 12. ROC curve

The accuracy analysis of the weed detection model is presented in Table 1.

Table 1. Accuracy analysis

Error	%
Type 1	22.22
Type 2	14.2857
Mean error overall	18.75
Mean accuracy overall	81.25
RMS error	43.30
Mean square error overall	18.75
False positive rate	0.22
True negative rate	0.77

The results show the percentage accuracy and classification errors for weed detection. In cases of dense vegetation, misclassifications are more likely to occur because weeds and crops may overlap, making it harder to distinguish between them and leading to higher error rates.

Discussion:

The primary objective of this research was to design a weed detection mechanism in real-world environments and enable the robot to navigate fields while avoiding obstacles. Various techniques exist for robot navigation, including machine learning methods like Convolutional Neural Networks (CNN), Deep Neural Networks (DNN), Artificial Neural Networks (ANN), and Region-based Convolutional Neural Networks (RCNN). Although

these approaches are effective in detecting objects in a given environment, we opted for the A* algorithm due to its robustness and efficiency in path planning.

Additionally, there are recent advancements in navigation based on LIDAR data, Monocular Vision, and Stereo Vision. Each of these methods has its advantages and limitations. For example, LIDAR offers accurate detection but is costly and has a limited perception range, while Monocular Vision may not guarantee reliable accuracy. For weed detection, we employed an image-processing technique that effectively works in natural environments. This method involves several steps: image pre-processing, image segmentation, and feature extraction. By integrating the information obtained from these steps, we successfully enhanced the detection and classification of weeds.

One of the key advantages of this study is the integration of stereo vision with the A path planning algorithm*, enabling the mobile robot to navigate agricultural fields with higher accuracy and adaptability. Unlike conventional GPS-based navigation, which may struggle with occlusions caused by dense vegetation, stereo vision allows for real-time depth perception and object avoidance. Previous research (Lin et al., 2023) has demonstrated that vision-based navigation improves localization accuracy in unstructured environments, making it a suitable approach for row-crop farming applications. However, the computational demands of real-time stereo vision processing remain a challenge. Future work should explore edge AI processing techniques to reduce latency and optimize real-time decision-making in embedded robotic systems.

Furthermore, while this study successfully employs PC/BC-DIM neural networks for weed classification, additional comparisons with CNNs and Transformer-based models would provide a more comprehensive analysis of model efficiency and accuracy. Prior research (Naveed et al., 2023) [10] indicates that pre-trained deep learning models such as YOLOv5 and EfficientNet achieve state-of-the-art weed detection performance, suggesting that a hybrid approach combining PC/BC-DIM with CNN-based feature extraction could enhance classification robustness. Additionally, incorporating data augmentation techniques could help address dataset limitations and improve model generalization across diverse agricultural conditions.

Finally, the electrical discharge-based weed removal mechanism presents a promising alternative to conventional mechanical or chemical-based weeding methods, but its long-term energy efficiency, operational safety, and impact on soil health require further evaluation. While previous studies (Diprose et al., 1984) [3] suggest that high-voltage weed removal is effective in disrupting plant cellular structure, excessive energy consumption may limit scalability in large farming applications. Future research should investigate energy-efficient discharge optimization techniques and evaluate potential side effects on surrounding crops, soil microorganisms, and long-term field productivity. Additionally, exploring hybrid weed removal methods, such as combining electrical discharge with robotic precision cutting or thermal weeding, could further improve the system's effectiveness and sustainability in precision agriculture.

Conclusion:

We discussed the primary objective of our project, which is to design the steering control for an Ackermann architecture-based weed-managing mobile robot. To achieve this, we developed a system capable of detecting weeds in real-world environments and navigating the robot while avoiding obstacles. We also outlined the methodology used to implement the project. In the first stage, we focused on path planning by capturing stereo images of the environment and rectifying them. Next, we calculated the depth map using the SGBM algorithm. With the help of the A* algorithm, the robot was then able to navigate the field while avoiding obstacles.

In the second stage, we addressed the weed detection task. For this, we first captured images using a secondary camera mounted on the base of the delta robot, positioned above its end effector. The captured image was then processed to detect weeds. This process involved converting the image to grayscale and then to a binary format by applying a threshold to emphasize key regions while minimizing the background. From the binary image, the crop region was identified using a saliency map and ROI (Region of Interest), which highlighted the crop area. The weeds were then detected by subtracting the identified crop region from the thresholder image. In the future, extensive field testing and validation will be conducted to assess how the dynamic field environment affects the performance of the proposed approach.

Acknowledgement. The authors would like to thank the administration of Intelligent Systems Laboratory, Department of Electrical Engineering, University of Gujrat for providing an opportunity to conduct research and perform experiments with the available laboratory facilities.

Author's Contribution. Faryal Naeem Mehmood and Syeda Ambreen Zahra wrote the initial draft of the article with equal participation. Syed Muhammad Wasif supervised and drafted the final version of the article. Zubair Mehmood improved the formatting of the final draft and proof read and corrected the manuscript with the help of Muhammad Jehanzeb Irshad and Nazam Siddique.

Conflict of interest. The authors have no conflict of interest for publishing this manuscript in IJIST.

References:

[1] B. Åstrand and A. J. Baerveldt, "An agricultural mobile robot with vision-based perception for mechanical weed control," *Auton. Robots*, vol. 13, no. 1, pp. 21–35, Jul. 2002, doi: 10.1023/A:1015674004201/METRICS.

[2] D. A. A. P. Ashok Kumar , Harisudha , Sivaprasad M, "DESIGN AND FABRICATION OF ELECTRIC WEEDER ALONG WITH FERTILIZER SPRAYER," *Int. Res. J. Eng. Technol.*, vol. 10, no. 4, 2023, [Online]. Available: https://www.irjet.net/archives/V10/i4/IRJET-V10I446.pdf

[3] M. F. D. M.F. Diprose, "Electrical methods of killing plants," *J. Agric. Eng. Res.*, vol. 30, pp. 197–209, 1984, doi: https://doi.org/10.1016/S0021-8634(84)80021-9.

[4] and C.-K. H. Cheng-Jian Lin, Ming-Yu Chang, Kuang-Hui Tang, "Navigation Control of Ackermann Steering Robot Using Fuzzy Logic Controller," *Sensors Mater.*, vol. 35, no. 3, 2023, doi: https://doi.org/10.18494/SAM4120.

[5] X. Wu, S. Aravecchia, P. Lottes, C. Stachniss, and C. Pradalier, "Robotic weed control using automated weed and crop classification," *J. F. Robot.*, vol. 37, no. 2, pp. 322–340, Mar. 2020, doi: 10.1002/ROB.21938.

[6] R. Aravind, M. Daman, and B. S. Kariyappa, "Design and development of automatic weed detection and smart herbicide sprayer robot," *2015 IEEE Recent Adv. Intell. Comput. Syst. RAICS 2015*, pp. 257–261, Jun. 2016, doi: 10.1109/RAICS.2015.7488424.

[7] V. D. C. et al. Le Hoai Phuong, "Design a Low-cost Delta Robot Arm for Pick and Place Applications Based on Computer Vision," *FME Trans.*, vol. 51, no. 1, pp. 99–108, 2023, doi: 10.5937/fme2301099P.

[8] M. A. G. D. V.A.Kulkarni, "Advanced Agriculture Robotic Weed Control System," *Int. J. Adv. Res. Electr. Electron. Instrum. Eng.*, vol. 2, no. 10, 2013, [Online]. Available: https://www.ijareeie.com/upload/2013/october/43Advanced.pdf

[9] R. Raja, Dr., D. C. Slaughter, Dr., S. Fennimore, and Dr., "A novel weed and crop recognition technique for robotic weed control in a lettuce field with high weed densities," *2019 ASABE Annu. Int. Meet.*, pp. 1-, 2019, doi:

10.13031/AIM.201900029.

[10] A. Naveed et al, "Saliency-Based Semantic Weeds Detection and Classification Using UAV Multispectral Imaging," *IEEE Access*, vol. 11, pp. 11991–12003, 2023, doi: 10.1109/ACCESS.2023.3242604.

Lower Limb Exo-Skeleton for Rehabilitation

Muhammad Moeed Zeb[1], Ali Maesam Kazmi[1], Syed Muhammad Wasif[1*], Zubair Mehmood[1], Muhammad Jehanzeb Irshad[1], Muhammad Waqas Jabbar[1], Nazam Siddique[1]
[1]Intelligent Systems Laboratory, Department of Electrical Engineering, University of Gujrat, Pakistan.
*** Correspondence:** Syed Muhammad Wasif: syed.wasif@uog.edu.pk

Citation | Zeb. M. M, Kazmi. A. M, Wasif. S. M, Mehmood. Z, Irshad. M. J, Jabbar. M. W, Siddique. N, "Lower Limb Exo-Skeleton for Rehabilitation", IJIST, Special Issue. pp 146-163, March 2025
Received | Feb 19, 2025 **Revised |** March 03, 2025 **Accepted |** March 09, 2025 **Published |** March 12, 2025.

Above-knee amputation remains a significant global issue, leaving many people physically disabled due to various natural and man-made causes, such as diseases, wars, and disasters. This article presents a novel, non-invasive active prosthesis based on electromyography (EMG). The proposed method offers a major advancement by achieving higher classification accuracy with minimal hardware requirements. Using EMG input signals, the active prosthesis controls three body postures: Sit, Stand, and Walk. These EMG signals are classified through two machine learning models: Support Vector Machine (SVM) and Long Short-Term Memory (LSTM) networks. Both models are evaluated based on accuracy. The results show that SVM outperforms LSTM, achieving a classification accuracy of 82%, while LSTM reaches 63%.

Keywords: Lower Limb Exoskeleton / Prosthesis; Non-Invasive Electromyography; Intention Recognition; Support Vector Machine.

Introduction:

Lower limb amputation remains a serious global issue, leaving many people physically disabled due to various natural and human-made causes, including diseases, wars, and natural disasters. However, diseases are still the leading cause of lower limb amputations. In the United States, it is estimated that by 2050, 3.6 million people will be living with amputations, compared to 2.2 million in 2021 [1]. In developed countries, diseases are the primary cause of trans-femoral amputations, mainly because diabetes mellitus has become more widespread, often leading to vascular complications. Peripheral vascular disease is the most common cause of lower limb amputations, while trauma is a leading cause in many low- and middle-income countries, affecting an estimated 57.7 million people.

Unlike upper-limb prosthetics, lower-limb prosthetics that integrate neuromuscular system signals for control are less widely used. Over time, different types of artificial lower limbs have been developed to improve amputee mobility. These include passive, semi-active, and active prostheses. Due to various limitations, passive and semi-active prostheses are less in demand, while active (or powered) prostheses are increasingly popular and provide better support for people with trans-femoral amputations.

Many researchers have worked to improve powered prostheses and address earlier limitations. In paper [2], an active prosthesis controlled by EMG is discussed, along with its limitations. Surface electromyography (sEMG) signals sometimes generate errors, which can cause amputees to fall. This prosthesis uses a machine learning model with a support vector machine (SVM) to classify signals. Another study [3] used a pattern recognition algorithm to translate EMG signals and combine sensor data from the prosthesis to interpret the user's intended movements. A study [4] compared three classifiers—LDA, SVM, and LM-BP—to identify the best solution, finding that LDA performed the best with an accuracy of 92.46%, enabling continuous recognition of limb movement intentions.

In paper [5], prosthesis control is achieved directly through neural signals activated by muscle contractions using EMG. Another experiment [6] analyzed modes using a finite-state approach and highlighted challenges in controlling the prosthesis using EMG signals. Study [7] proposed an automatic system for detecting and analyzing muscle defects by evaluating different leg movements with sEMG sensors and advanced machine learning algorithms, using SVM to classify muscle movements.

A study [8] focused on intelligent lower-limb prostheses, emphasizing the importance of segmenting locomotion modes. Researchers analyzed five states—ramp descent (RD), level walking (LW), stair ascent (SA), stair descent, and ramp ascent—and achieved 99.16% ± 0.38% accuracy with an ANN-based adaptive online learning algorithm. In paper [9], classifiers like linear discriminant analysis, Naive Bayes, k-nearest neighbor, and SVM were used to predict knee angles based on EMG data recorded while sitting and standing. Fifteen features were used to improve prediction accuracy, and principal component analysis helped address dimensionality issues, with the SVM classifier (quadratic kernel) performing the best.

In study [10], the focus was on using electromyography and mechanical sensors to detect movement intentions in trans-femoral amputees using powered prostheses. The combination of mechanical sensor data and EMG signals reduced transitional error by 18% and steady-state error from 3.85% to 1.05%. Another study [11] developed an EMG-controlled trans-femoral prosthesis using four machine learning models (LDA, SVM, KNN, and Decision Tree) to classify extension and flexion movements by analyzing signals from two muscle channels with six features each. Two models achieved accuracy below 80%, while the other two reached 80% and 81%.

The main objectives of this study are to design and develop an EMG-controlled active prosthetic leg, create a muscle signal classification algorithm, and develop a prosthetic leg

control system. This research builds upon previous work [11] and offers the following key contributions:

- **Reduced Hardware Complexity:** The proposed approach uses a single-channel setup instead of the two-channel system in the previous study, reducing complexity while maintaining high classification performance.
- **Higher Classification Accuracy:** The classification accuracy of SVM has improved to 82%, compared to 70% in earlier work.
- **Enhanced Feature Extraction:** This study extracted 22 features (21 time-domain and 1 frequency-domain), compared to just 6 time-domain features in the previous approach.
- **Dimensionality Reduction:** Principal component analysis was applied to optimize classification performance, a technique missing in the earlier study.
- **Deep Learning Integration:** Unlike the previous study, which relied solely on machine learning, this research incorporates long short-term memory (LSTM) to explore the potential of deep learning in prosthetic control.

Objectives of the Study:

The primary objective of this study is to design and evaluate a non-invasive electromyography (EMG)-based active prosthesis for above-knee amputees, enhancing intention recognition accuracy, classification efficiency, and real-time applicability. The specific objectives of this study are:

- To develop a low-complexity, single-channel EMG acquisition system that reduces hardware requirements while maintaining high classification accuracy.
- To compare the performance of traditional machine learning (SVM) and deep learning (LSTM) models for intention recognition in lower limb prosthetic control.
- To optimize feature extraction and dimensionality reduction techniques by incorporating time-domain and frequency-domain features, enhancing computational efficiency.
- To evaluate real-time feasibility and processing latency of the proposed system for potential integration into wearable prosthetic devices.
- To assess the impact of dataset size, signal variability, and electrode displacement on classification accuracy, ensuring robustness across diverse user conditions.

Novelty Statement:

This study introduces a low-complexity, single-channel EMG-based prosthetic control system, which significantly reduces hardware requirements compared to conventional multi-channel setups while maintaining high classification accuracy. Unlike previous works that primarily rely on linear classifiers or feature-limited datasets, this study integrates a comprehensive feature extraction framework with 22 extracted features and applies dimensionality reduction (PCA) to optimize classification performance. Additionally, the research provides a comparative analysis between SVM and LSTM classifiers, highlighting the advantages of traditional machine learning for small datasets and real-time applications. The findings demonstrate the potential for deploying EMG-based prosthetic control in real-world scenarios, bridging the gap between biomedical signal processing and practical assistive device development.

Materials and Methods:

EMG signals are acquired using non-invasive electrodes. These signals are then sent to the pre-processing module, where they are amplified, rectified, and filtered. Next, the feature extraction module processes the pre-processed data and extracts 21 time-domain features and 1 frequency-domain feature using a specific algorithm. A dimension reduction technique is then applied to optimize the data.

After this step, 10 principal components are provided to both classification modules for signal classification. The classifier with the best performance is selected. Based on the recognized motion intention, the accepted classifier sends commands to the control system, which then activates the prosthetic limb's actuators accordingly. Figure 1 illustrates the block diagram of the lower limb exoskeleton.

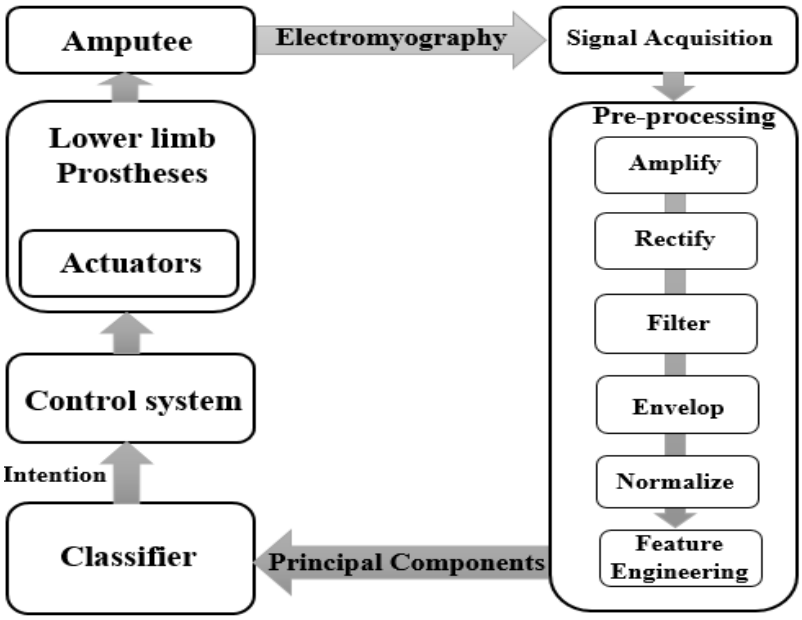


Figure 1. Block diagram of lower limb prostheses.

Figure shows that EMG signal is acquired and then preprocessed to extract useful features for the brain intention of muscular actuation and classification task. With the classified intention the proposed system generates appropriate motor commands to generate required posture of active limb.

Signal Acquisition:

The signal is acquired using non-invasive electrodes, which are attached to the relevant muscles. The terminals of these electrodes are connected to an Arduino, which collects the data from the electrodes. This data is then transmitted and displayed on a monitor via a Raspberry Pi. This entire process constitutes signal acquisition. The electrodes used in this setup are shown in Figure 2.

Figure 2. EMG Surface Electrodes

Preprocessing:

Preprocessing involves several steps, including amplification, rectification, filtering, normalization, and feature engineering. Signal preprocessing can be carried out using two methods:

- **Using an EMG Muscle Sensor Kit** (this excludes normalization and feature engineering).

Using programming.

In this study, both methods are applied for specific reasons, which are explained later. The EMG kit is used to preprocess the signal. It amplifies the raw muscle signal from millivolts (mV) to thousands of volts, depending on the gain setting, which is adjusted to 10,350 volts in this case. After amplification, rectification is performed to remove negative voltages from the signal, helping preserve the useful information within it.

Since the effective frequency range of EMG signals is between 0 and 500 Hz, the sensor kit filters out frequencies below 0 Hz and above 500 Hz, ensuring that the output signal has a bandwidth of 0 to 500 Hz. These steps are handled by the sensor kit. The signal is then acquired and partially preprocessed through this combination of processes, as shown below.

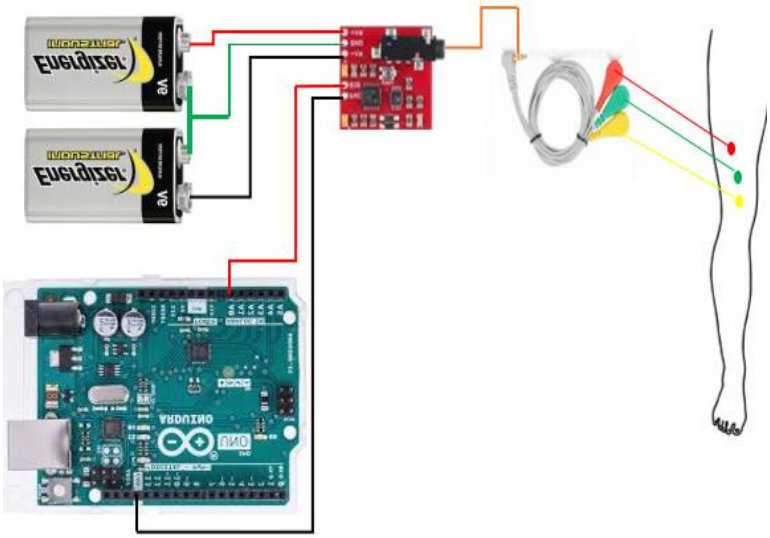


Figure 3. Muscle signal acquisition and partially preprocessing framework

Now, all the remaining steps are carried out using programming. The signal undergoes filtration to remove motion artifacts and power interference while staying within the effective frequency range. To achieve a noise-free signal, a 4th-order Butterworth stop-band filter is applied, blocking the 20 Hz frequency associated with motion artifacts. Then, another stop-band filter is used to block the 50 Hz frequency, which represents power line interference. After filtration, normalization is performed to limit the signal voltage to a range between 0 and 1 volt, which improves the classifier's accuracy.

At the final stage of preprocessing, feature engineering is applied to extract 21 time-domain features and 1 frequency-domain feature from the signal. To prevent the issue of high-dimensional data (also known as the "curse of dimensionality") and to enhance model efficiency, Principal Component Analysis (PCA) is used. In this study, features are first extracted and then reduced using PCA to select the most important principal components, which helps the classifier make better predictions.

Feature extraction is the process of obtaining key characteristics or features from the signal. In biomedical signal processing, feature extraction is categorized into three types: time-domain features, frequency-domain features, and combined time-frequency features. Here, 21 time-domain features and 1 frequency-domain feature, as mentioned earlier, are extracted using a rolling window of 50 data points, with an increment of 1 data point at each step. These features include the following: Minimum, Maximum, Mean, Root Mean Square, Variance, Standard Deviation, Signal Power, Peak, Peak-to-Peak, Crest Factor, Skewness, Kurtosis, Form Factor, Pulse Indicator, Average Absolute Value, Signal Similarity Index, Integrated EMG, Waveform Length, Logarithmic, Willison Amplitude, Mean Frequency, and Mean Absolute Value. The formulas and explanations for how these features are calculated are provided below.

Willison Amplitude:

The Willison Amplitude feature, in the context of EMG signal analysis, measures the peak-to-peak amplitude of the EMG waveform to estimate the overall amplitude of the EMG signal. A threshold of 0.002 V is set, which indicates the strength of muscle contractions.

$$A = \sum_{i=1}^{N-1} f\ (|x_n - x_{n-1}|) \quad (1)$$

$$f(x) = \begin{cases} 1, & if\ x \geq threshold \\ 0, & otherwise \end{cases}$$

Where, x_n is the EMG value of n index and x_n-1 is the previous value of the current value.

Root Mean Square:

Root Mean Square (RMS) is a widely used technique in electromyography (EMG) signal analysis to quantify the amplitude or magnitude of the EMG signal. It provides a signal strength or level indication for the entire signal over a given time period.

$$RMS = \sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i^2} \quad (2)$$

Where, x_i is the EMG value of i index and N is the total samples of EMG value in a window.

Mean Absolute Value:

Mean absolute value (MAV) is defined as mean of total positive value of EMG signal. It also tells about muscle contraction power. It is mathematically represented as

$$MAV = \frac{1}{N}\sum_{i=1}^{N}|x_i| \quad (3)$$

Where, x_i is the EMG value of i index and N is the total sample of EMG value in a window.

Average absolute value:

The average of the absolute changes between successive EMG signal levels is measured by Average absolute change (AAC). It specifies fluctuating in muscle movement. Mathematically,

$$AAC = \frac{1}{N-1}\sum_{i=1}^{N-1}|x_{i+1} - x_i| \quad (4)$$

Where, x_i is the EMG value of i index and x_{i+1} is the upcoming value of the current value. N indicates the total samples in a window.

Variance:

Variance is a technique which provide information about spread of EMG data points around the mean. It tells about the fluctuations of EMG signal over time. Mathematically,

$$Variance = \frac{1}{N-1}\sum_{i=1}^{N} x_i^2 \quad (5)$$

Where, x_i is the EMG value of i index and N indicates the total samples in a window.

Log detector:

Log detector (LD) in case of EMG signal used to transform the EMG signal from linear scale to logarithmic scale. This scale down the large variations in the signal amplitude.

$$LD = e^{\frac{1}{N}\sum_{i=1}^{N} \ln(|x_i|} \quad (6)$$

Where, x_i is the EMG value of i index and N indicates the total samples in a window.

Simple Square Integral:

The EMG signal's sum of squared values is known as the simple square integral (SSI). It displays the strength of the signal throughout a chosen window. Higher SSI means high level of contraction. Mathematically,

$$SSI = \sum_{i=1}^{N} x_i^2 \quad (7)$$

Where, x_i is the EMG value of i index and N indicates the total samples in a window.

Integrated EMG:

The EMG signal's absolute levels are added together to form IEMG. It measures all of the muscle activity in a certain amount of time. IEMG indicates the overall muscle movements. It is mathematically represented as

$$\text{IEMG} = \sum_{i=1}^{N} |x_i| \quad (8)$$

Where, x_i is the EMG value of i index and N indicates the total samples in a window.

Waveform Length:

The cumulative length of the EMG signal waveform over a specified time is measured by Waveform Length (WL). It displays the signal's dynamic muscle contraction. Mathematically,

$$\text{WL} = \sum_{i=1}^{N-1} |x_{i+1} - x_i| \quad (9)$$

Where, x_i is the EMG value of i index and xi+1 is the upcoming value of the current value. N indicates the total samples in a window.

Standard Deviation:

The EMG signal's degree of fluctuation or dispersion is measured by standard deviation. It displays the signal's variability that helps the model to identify different muscle activities.

$$\text{STD} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2} \quad (10)$$

Where, x_i is the EMG value of i index and $\bar{x}$ is mean of all values in a window. N indicates the total samples in a window.

Maximum:

The maximum EMG value of muscle over specific period of time.

Minimum:

The minimum EMG value of muscle over specific period of time

Mean:

The average value of the muscle signal. Mathematically,

$$\text{Mean} = \frac{1}{N}\sum_{n=1}^{N} x[i] \quad (11)$$

Where x is the signal amplitude, and N is the total number of EMG samples.

Signal Power:

The average power of EMG signal is called signal power (SP). Mathematically,

$$\text{SP} = \frac{1}{N}\sum_{n=1}^{N} x[i]^2 \quad (12)$$

Where x is the signal amplitude, and N is the total number of EMG samples.

Peak:

The highest value of signal amplitude.

Peak-to-Peak (P2P):

The difference of highest and lowest value of signal

$$\text{Peak-to-Peak} = \text{Maximum} - \text{Minimum}$$

Crest factor:

The ratio of the signal's peak amplitude to its RMS value. It indicates how extreme the peaks are relative to the RMS.

$$\text{Crest Factor} = \frac{Peak}{RMS} \quad (13)$$

Skewness:

Measures the asymmetry of the signal's amplitude distribution.

$$\text{Skewness} = \frac{\frac{1}{N}\sum_{n=1}^{N}(x[i] - Mean)^3}{Standard\ Deviation} \quad (14)$$

Positive skewness indicates more values above the mean, while negative skewness indicates more values below it.

Kurtosis:

Indicates the sharpness or flatness of the amplitude distribution A high kurtosis value suggests sharp peaks.

$$\text{Kurtosis} = \frac{\frac{1}{N}\sum_{n=1}^{N}(x[i]-Mean)^4}{Standard\ Deviation^4} \quad (15)$$

Form Factor:

The ratio of RMS to the mean absolute value (MAV) of the signal

$$\text{Form factor} = \frac{RMS}{MAV} \quad (16)$$

Pulse Indicator:

The ratio of the peak amplitude to the mean absolute value

$$\text{Pulse Indicator} = \frac{Peak}{MAV} \quad (17)$$

Mean Frequency:

The average frequency of the signal in the frequency domain, often obtained via a Fourier transform

$$\text{Mean frequency} = \frac{\sum_{f=0}^{F} f.p(f)}{\sum_{f=0}^{F} p(f)} \quad (18)$$

1.**Initialize Parameters:**

- Define the input emg data and the rolling window_size.
- Create 22 empty lists (MIN, MAX, MEAN, RMS, etc.) to store calculated feature values for each window.

2.**Iterate Through the EMG Data:**

- Use a for loop to slide the rolling window across the EMG signal:
- The loop runs from index 0 to len(emg) - window_size + 1.
- For each iteration:
- Extract a segment of emg corresponding to the current rolling window (rolling = emg [i:i + window_size]).

3.**Compute Features for Each Window:**

- For the current rolling window, calculate all these features.
- Append each calculated feature to its corresponding list.

The algorithm for extracting the 22 features is explained above. Here, "rolling" refers to an array of 50 EMG data samples. Initially, all these features are extracted from the dataset and standardized. After standardization, they are passed to Principal Component Analysis (PCA) to reduce the dimensionality of the data.

PCA is a crucial step in machine learning model training because it helps prevent the curse of dimensionality and reduces computational complexity. While other techniques, such as feature selection or feature elimination, are available, they come with certain limitations. Initially, the Exhaustive Feature Selection (EFS) approach was considered, but it follows the formula **(2^n - 1)**, which makes the computation highly complex and time-consuming. To address this issue and save time, PCA was chosen.

PCA works by transforming a set of correlated variables into a smaller set of uncorrelated variables. In this study, PCA reduces the 22 extracted features to 10 principal components, which are then sent to both classifiers for further analysis.

Support Vector Machine:

Support Vector Machine (SVM) is a supervised learning model that classifies data by creating a separating boundary between different classes. It does this by finding the optimal hyperplane and maximizing the margin (the distance between the hyperplane and the nearest data points). In this work, the Radial Basis Function (RBF) kernel is used to enhance SVM's performance.

$$K(X_1,X_2) = \exp\left(-\frac{\|X_1-X_2\|^2}{2\partial^2}\right) \quad (19)$$

Here, $||X1 - X2||^2$ represents the Squared Euclidean Distance, and σ is a free parameter used to fine-tune the equation.

In this study, Grid Search CV is applied to optimize the classifier by selecting the best parameters, which enhances the model's accuracy and efficiency. This technique systematically explores a predefined range of parameter values to identify the optimal hyperparameters for a machine learning model. The model is trained using various values of C and Gamma, followed by an evaluation of its accuracy. In SVM, C acts as the regularization parameter, while Gamma determines the influence of individual training examples.

Table 1. Hyper parameter values and purpose of SVM

Hyper parameter	Value in code	Purpose
Regularization (C)	100	Reduce classification error
Gamma	1	Model learn meaningful decision boundaries
Kernel	RBF	Good for nonlinear data

Long Short-Term Memory:

Long Short-Term Memory (LSTM) is an advanced type of Recurrent Neural Network (RNN) designed to manage long-range dependencies in sequential data. LSTM consists of three gates: input, forget, and output, which control the flow of information. It can process sequential data in both forward and backward directions and store information over extended periods. The block diagram of the LSTM is shown below in Figure 4.

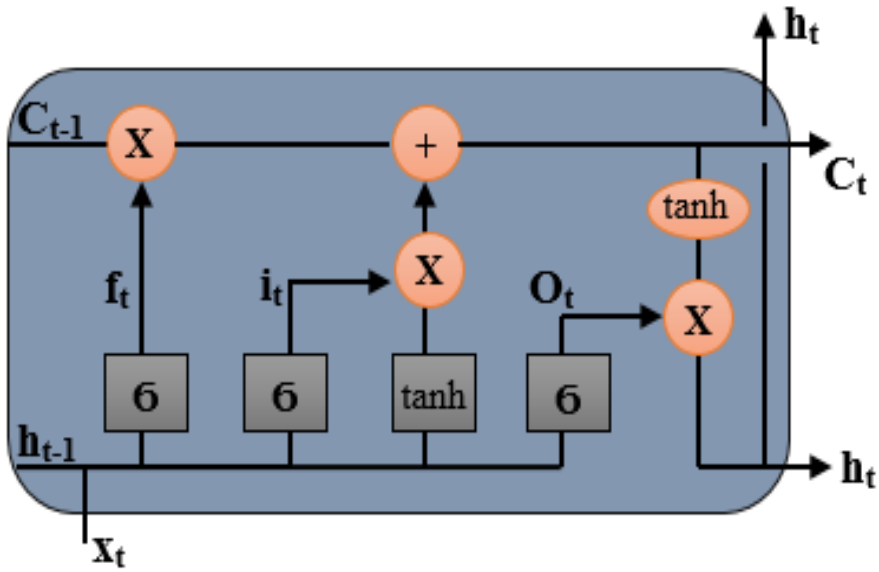


Figure 4. LSTM block diagram.

A Sequential model will be used, meaning the layers will be added one after another. The model includes 50 LSTM neurons in the layer, along with dropout layers to prevent overfitting and improve regularization. Since there are three classes, a dense layer is added for multi-class classification, utilizing the SoftMax activation function.

Table 2. Hyper parameter values and purpose of LSTM

Hyper parameter	Value in my code	Purpose
Dropout rate	0.2 (twice)	Reduces overfitting
Learning rate	0.001(fixed)	Controls weight update
Batch size	32	Define training step size

The classifier categorizes the signal into three classes: sit, stand, and walk, based on the extracted features. Initially, both classifiers will be trained on a labeled leg dataset. Once the training phase is complete, the testing phase will begin, and the accuracy will be monitored. The classifier with the highest accuracy will be selected. In this study, a Support Vector Machine (SVM) classifier will be used due to its superior accuracy, as explained in the next section. The extracted features are sent to the classifier, which categorizes them accordingly and assigns appropriate labels. For example, when a walking signal is generated by the muscles, the features are extracted and passed through PCA, followed by classification. The classifier then identifies the signal as walking and labels it accordingly. This command is then transmitted to the control system.

Sit Posture:

When a person sits on a chair, the knee angle is usually around 90 degrees, although it may vary based on posture. In this sitting position, the stepper motor should rotate 90 degrees in the anticlockwise direction. Assuming the stepper motor starts at a 0° position, it will move accordingly. The formula is to calculate how many steps stepper motor revolves to go to desired degree is given below:

$$\text{Steps} = \left(\frac{0}{1.8} - \frac{90}{1.8}\right) = -50 = 90^{o} \text{ (anticlockwise direction)}$$

Here 1.8° means that the stepper motor used in this project is designed in such a way that it covers 1.8° per step and negative sign indicates the direction of stepper motor.

Stand Posture:

When a person stands up from the chai, the knee angle transforms from 90° to nearly 0°. In stand posture, according to given formula result, the stepper motor takes 50 steps in clockwise direction.

$$\text{Steps} = \frac{90}{1.8} - \frac{0}{1.8} = 50 = 90^{o} \text{ (clockwise direction)}$$

Stand Posture:

For walking of a robotic leg, stepper motor should work in to and fro motion. The angle of normal human knee while walking is 0 to 20 degrees. In walking posture, the stepper motor goes from 0° to 20° and then from 20° to 0°. Now to calculate the steps, formula will be used;

$$\text{Steps} = \frac{0}{1.8} = 0 \;\; \& \;\; \text{Steps} = \frac{20}{1.8} = 12$$

The stepper motor first takes 12 steps in the anticlockwise direction, followed by steps in the clockwise direction to form the walking pose. In this study, the NEMA 17 stepper motor is used, controlled by the A4988 module, which serves as the motor driver. The connections between the stepper motor and its driver are somewhat complex.

Results:

The dataset [12] was recorded from 10 able-bodied individuals, and details about the dataset are provided in Table 2 below. Only three classes of data are selected, amounting to approximately 171,000 samples. The unprocessed dataset is used for model training and then processed according to the steps outlined in the methodology. First, the dataset is balanced using the under-sampling technique. After balancing, the data undergoes detrending, or mean removal. The correlation heatmap is shown in Figure 5.

Table 3. Dataset Information

Gender	Age	Weight	Height
7 Males	24±2 years	77±10 kg	183±9cm
3 Females	24±2 years	77±10 kg	183±9cm

The correlation heatmap shows the variance of features in relation to each other and the target variable. A value closer to 1 indicates a strong positive correlation, while a value near -1 reflects a strong negative correlation. Values close to or equal to 0 suggest no correlation. In this case, most features show little to no correlation with the target, indicating a lack of linear relationships. Considering this, the Radial Basis Function (RBF) kernel is used, as it is more effective in handling nonlinear data relationships. Next, the standardized features are sent for dimensionality reduction using Principal Component Analysis (PCA), where 10 principal components are selected. Both models are then trained and tested. The SVM, a supervised machine learning model, is trained on these principal components. Figure 6 illustrates the accuracy and loss across each epoch for the SVM model.

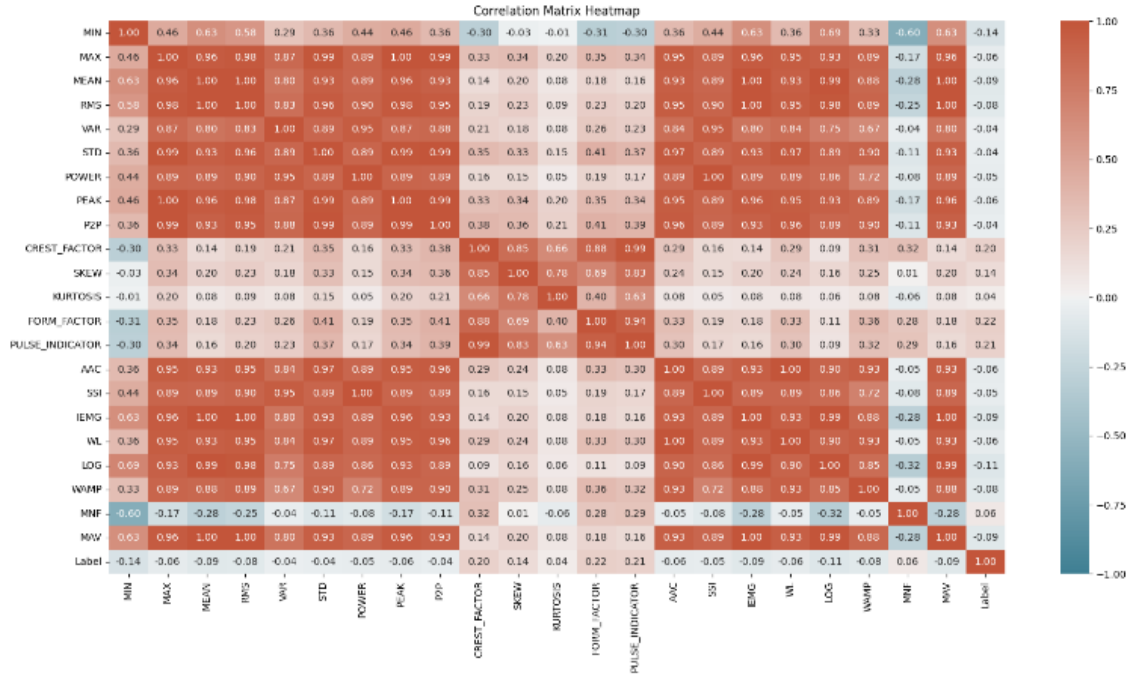


Figure 5. Correlation heat map of features extracted employed for brain intention classification.

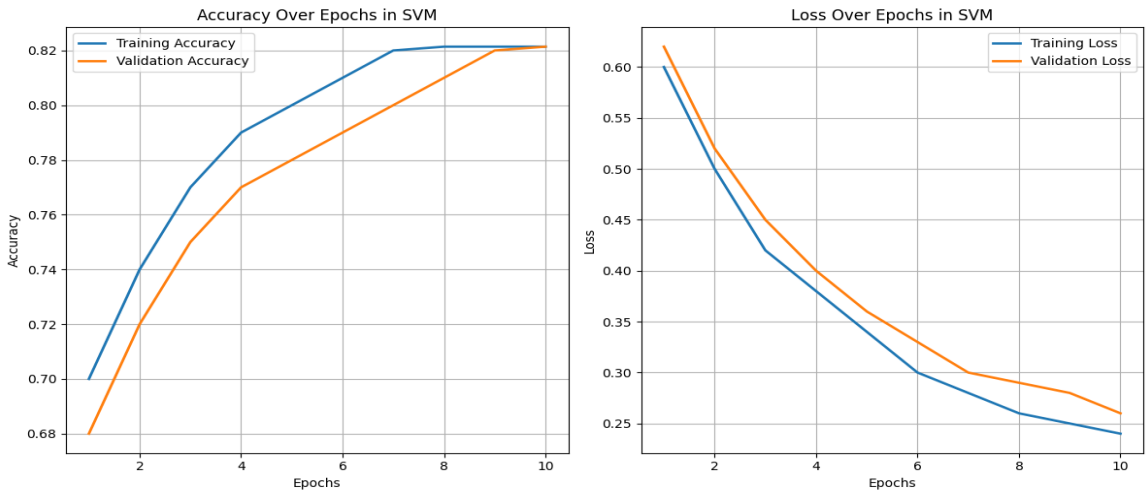


Figure 6. Accuracy and loss over epochs of SVM classifier.

Nearly 90% of the data was used to train the model, while the remaining 10% was reserved for testing. The performance of the trained model is presented in Table 3, and the confusion matrix is displayed in Figure 7. The model achieved an accuracy of 82.14%, as shown below.

Table 4. Classification report of SVM

Class	Precision	Recall	F1 score
0	0.78	0.84	0.81
1	0.91	0.79	0.85
2	0.79	0.84	0.81

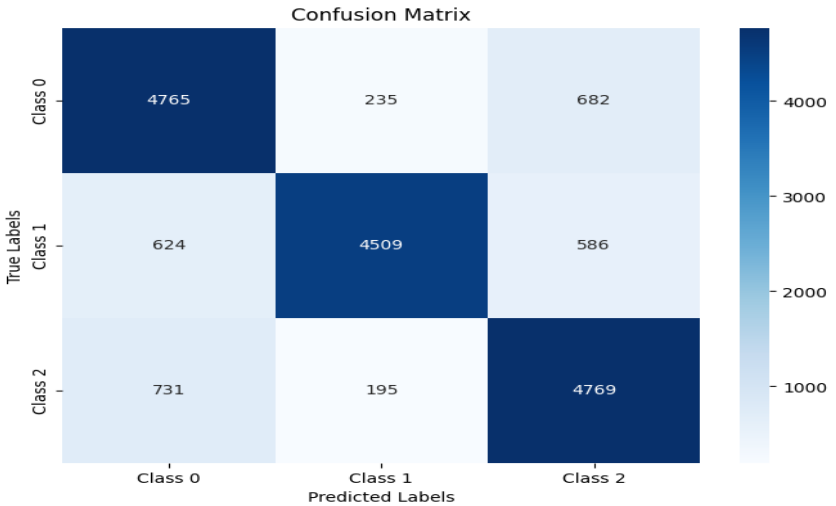


Figure 7. Confusion matrix of SVM model

The 95% confidence interval for the SVM model is 80.20% to 82.09%, indicating that the model's performance is stable. The reported accuracy of 82.14% represents the average accuracy obtained after performing 5-fold cross-validation.

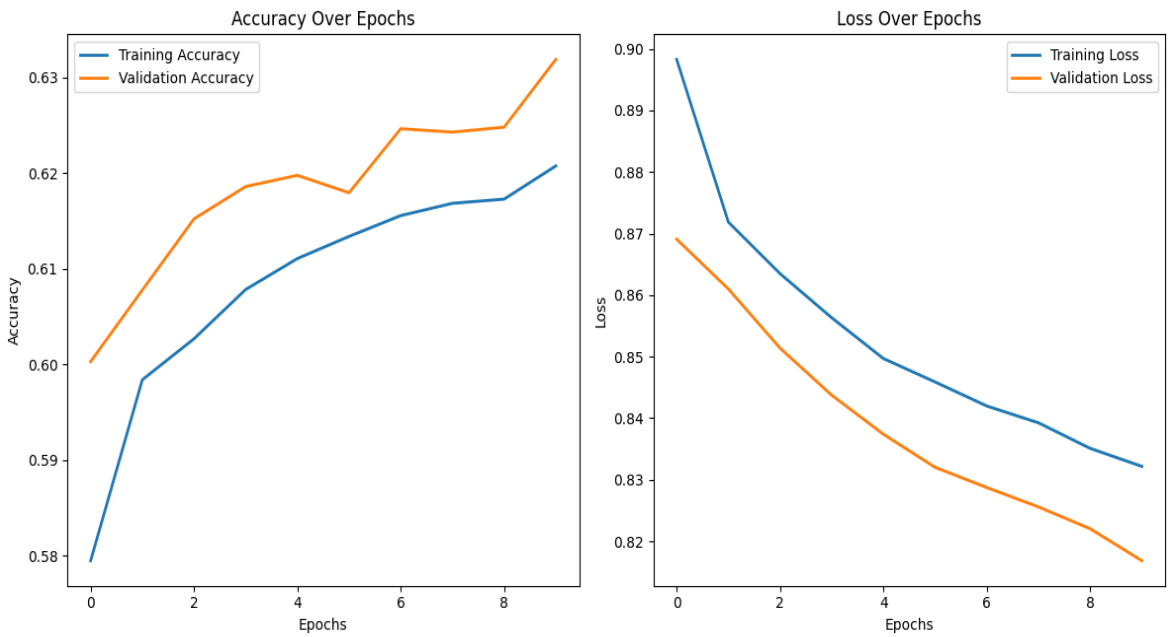
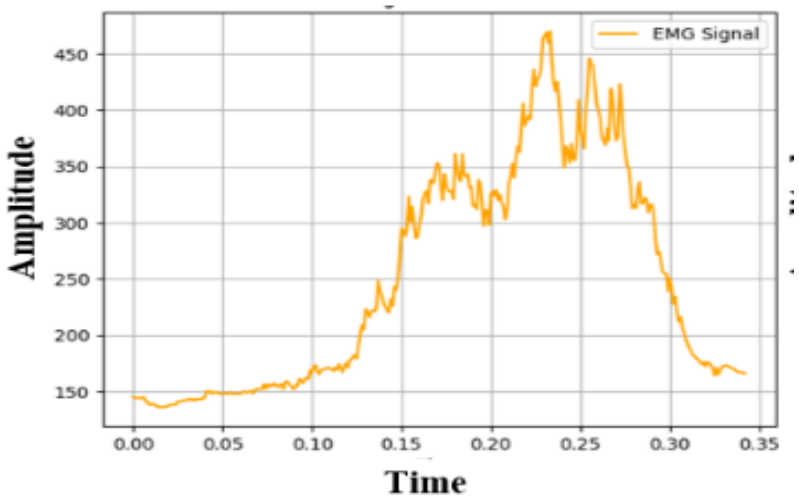


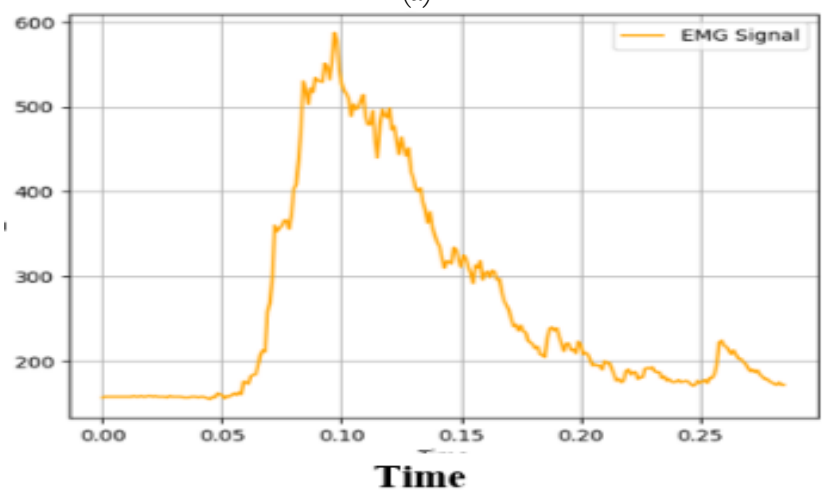
Figure 8. Accuracy and loss over epochs of LSTM

After training the model, real-time signal classification is conducted. The muscle signal is obtained from the rectus femoris using electrodes. This raw signal is processed and sent from Arduino to Visual Studio Code through serial communication. The muscle signal, as shown in

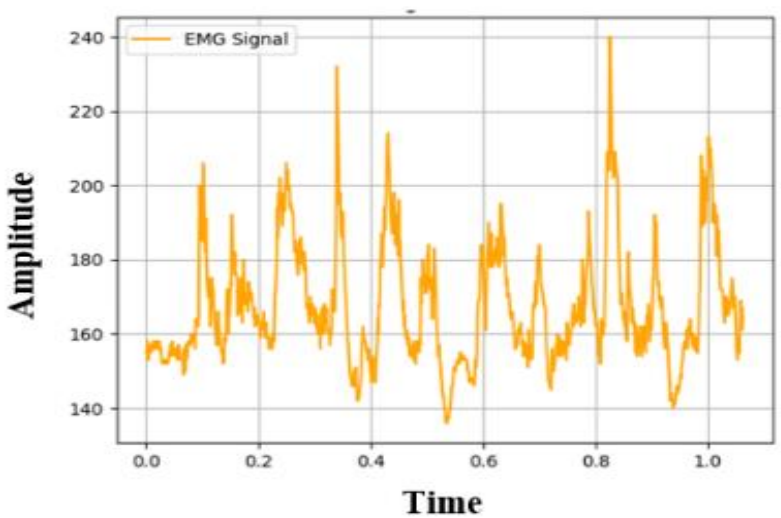
Figure 10, is amplified, rectified, and filtered. The normalized EMG signal is shown in figure 11.



(a)

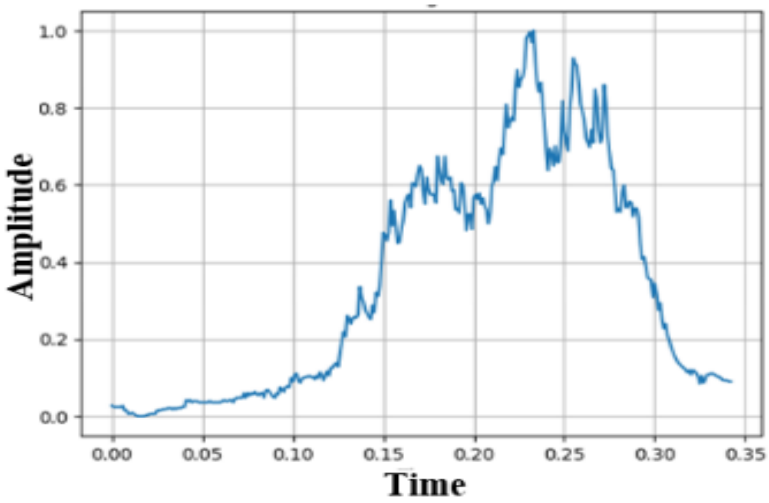


(b)

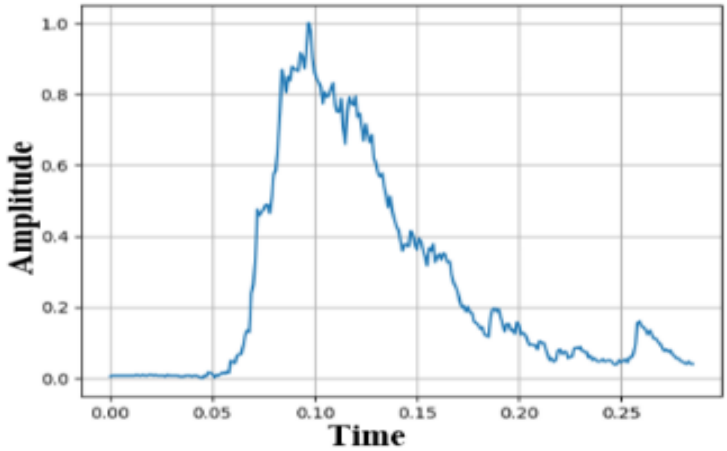


(c)

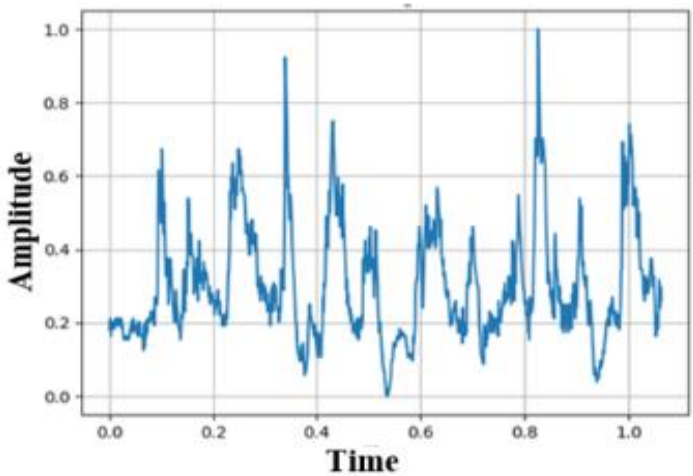
Figure 10. Input muscle signal for various lower limb postures. (a) Sit posture EMG signal, (b) Stand posture EMG signal, (c) Walk posture EMG signal.



(a)



(b)



(c)

Figure 11. Preprocessed muscular signals for (a) Sit (b) Stand (c) Walk posture.

Subsequently, features are extracted, and standardization is performed. All 22 features are then passed to PCA, which reduces them to 10 principal components (PC). These components are fed into the SVM model for intention recognition, with the results presented in Table 4.

Table 5. Prediction on unknown muscle signal data

PC 1	PC 2	PC 3	PC 4	PC 5	PC 6	PC 7	PC 8	PC 9	PC 10	Class
1.451	-0.258	0.107	0.018	-0.225	-0.455	-0.172	-0.258	-0.445	-1.593	0
3.082	3.020	3.350	3.351	2.346	3.311	2.715	3.020	2.799	0.326	1
-0.564	-0.508	-0.573	-0.553	-0.224	-0.431	-0.271	-0.508	-0.466	-0.266	2

After the intention is recognized, a message is sent to the control system, which reads the message and instructs the actuators to perform the desired motion. Figure 12 illustrates the three positions of the prosthetic leg.

(a) (b) (c)

Figure 12. Prosthesis knee postures, (a) Sit, (b) Stand, (c) Walk.

Discussion:

This study presents a low-complexity, single-channel electromyography (EMG)-based active prosthesis designed for above-knee amputees, addressing key challenges in intention recognition, classification accuracy, and real-time feasibility. The results demonstrate that Support Vector Machine (SVM) achieves the highest classification accuracy (82.14%), outperforming Long Short-Term Memory (LSTM) networks (63.25%). The superior performance of SVM aligns with previous research emphasizing the effectiveness of traditional machine learning models for small biomedical datasets [2], [3]. On the other hand, while deep learning models such as LSTM are often expected to perform well in time-series signal classification, the results in this study indicate that LSTM does not prove to be the best model for EMG-based intention recognition.

The LSTM model was trained on 90% of the dataset and tested on 10%, similar to the SVM training approach, but the results were not satisfactory. As shown in Figure 8, the accuracy and loss plots indicate that the model is learning step by step over each epoch, and the loss consistently decreases after every iteration. Importantly, there is no sign of overfitting, as the loss and validation loss remain close to each other throughout training. However, despite this, LSTM achieves only 63.25% accuracy, which is significantly lower than the 82.14% obtained using SVM. The accuracy plot for LSTM, as shown in Figure 9, further confirms that while the model gradually improves with training, it does not reach a performance level sufficient for real-world prosthetic applications. Due to these findings, SVM is selected as the final model for implementation in this project.

A key contribution of this study is the reduction of hardware complexity through the implementation of a single-channel EMG acquisition system, compared to traditional multi-channel EMG setups [5], [8]. This reduction in sensor count simplifies the integration of the prosthetic control system into wearable devices, making it more cost-effective and practical for real-world applications. Furthermore, the use of 22 extracted features (21 time-domain, 1 frequency-domain) combined with Principal Component Analysis (PCA) improves the efficiency of the classification model. Unlike previous approaches that relied on a limited

number of features, this study demonstrates that a diverse feature set enhances classification performance without significantly increasing computational load.

Despite achieving high classification accuracy, several challenges remain for real-time implementation and practical deployment. The study does not address processing latency, power consumption, and wireless communication feasibility, which are crucial factors for wearable prosthetic devices [9]. Future research should evaluate the real-time response of the system, investigating potential delays in muscle signal processing and the computational efficiency of embedded hardware components. Additionally, exploring deep learning models such as CNNs and hybrid CNN-LSTM architectures may provide further improvements in classification robustness, feature extraction, and adaptation to muscle fatigue-induced variations.

Furthermore, robustness against signal variability due to muscle fatigue, electrode displacement, or noise artifacts is an important consideration for real-world usability.

Previous research [6] has shown that EMG signals fluctuate due to physiological conditions, leading to inconsistencies in prosthetic control. Future work should incorporate adaptive filtering techniques, transfer learning approaches, and real-time signal correction methods to enhance system stability. Additionally, while the study provides a comparative analysis of classification models, increasing the dataset size and participant diversity will further improve model generalization and performance across different user demographics.

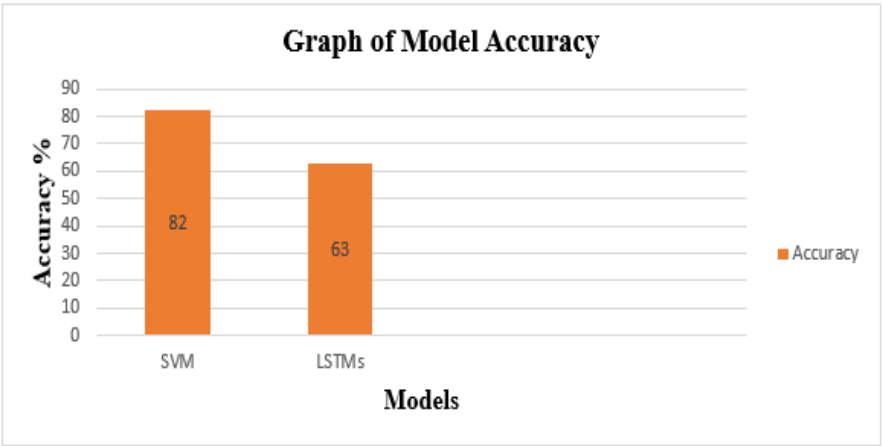


Figure 9. Accuracy graph of SVM and LSTM

The primary objective of this study is to develop an improved and reliable above-knee prosthesis that can efficiently assist amputees. The proposed study incorporates several advancements compared to previous research [11].

Table 6. Comparing results with previous study

Factor	Previous Approach [11]	Our Approach
Muscle classification	2 muscles (Femoris & Vastus)	1 muscle (Rectus Femoris0
Features used	6 features for each muscle	22 features
Dimensionality reduction	Not used	PCA
Model type	4 ML	1 ML and 1 DL
Common model type used	SVM, linear kernel (70% accuracy)	SVM, RBF kernel (82% accuracy)
Total movement recognition	2 (Extension & Flexion)	3 movement recognition

Conclusion:

We discussed that the main goal of our project is to design a lower limb exoskeleton/prosthetic device for rehabilitation. To achieve this, we have developed and designed a mechanical model of a robotic leg, which can be attached to an amputee to enhance efficiency. The core objective of this project is to create lower limb prostheses that improve human health, provide comfort, and remain financially affordable. We also examined different methods to enhance the efficiency of conventional prosthetic devices, which are often uncomfortable and less effective.

Our project is implemented in two stages. In the first stage, we focused on signal acquisition and processing to accurately capture muscle signals. Signal processing includes several operations, such as amplifying the acquired signal, filtering out unwanted noise, rectifying the signal, and normalizing it to scale the output between 0 and 1. This improves signal quality, enabling more accurate feature extraction and enhancing the prediction efficiency of the machine learning model.

In the second stage, we worked on the control system of the robotic leg to execute the intended actions. The machine learning model predicts the required signal, which is then sent to the control system to trigger the actuator and perform the corresponding action. By integrating this device into real-world prosthetic applications, it can assist individuals with above-knee amputations. Due to effective signal processing and modeling, the device has demonstrated promising results when tested on new participants. The test outcomes, shown in the above section, highlight its performance. Additionally, the model's classification accuracy could be further improved by employing more advanced deep learning models, such as CNN-based architectures.

Acknowledgement: The authors would like to thank the administration of Intelligent Systems Laboratory, Department of Electrical Engineering, University of Gujrat for providing an opportunity to conduct research and perform experiments with the available laboratory facilities.

Author's Contribution: Muhammad Moeed Zeb, Ali Maesam Kazmi wrote the initial draft of the article with equal participation. Syed Muhammad Wasif supervised and drafted the final version of the article. Zubair Mehmood improved the formatting of the final draft and proof read and corrected the manuscript with the help of Muhammad Jehanzeb Irshad, Muhammad Waqas Jabbar and Nazam Siddique.

Conflict of interest: The authors have no conflict of interest for publishing this manuscript in IJIST.

References:

[1] K. Ziegler-Graham, E. J. MacKenzie, P. L. Ephraim, T. G. Travison, and R. Brookmeyer, "Estimating the Prevalence of Limb Loss in the United States: 2005 to 2050," *Arch. Phys. Med. Rehabil.*, vol. 89, no. 3, pp. 422–429, Mar. 2008, doi: 10.1016/J.APMR.2007.11.005/ASSET/0D660FF8-C4A8-4595-9B40-B78D51E9E385/MAIN.ASSETS/GR2.SML.

[2] L. H. and M. O.-C. B. Ahkami, K. Ahmed, A. Thesleff, "Electromyography-Based Control of Lower Limb Prostheses: A Systematic Review," *IEEE Trans. Med. Robot. Bionics*, vol. 5, no. 3, pp. 547–562, 2023, doi: 10.1109/TMRB.2023.3282325.

[3] T. A. K. Levi J. Hargrove, Ann M. Simon, Aaron J. Young, Robert D. Lipschutz, Suzanne B. Finucane, Douglas G. Smith, "Robotic Leg Control with EMG Decoding in an Amputee with Nerve Transfers," *New Engl. J. Med. homepage*, vol. 369, no. 13, pp. 1237–1242, 2013, doi: 10.1056/NEJMoa1300126.

[4] B. Liu, Z. Chen, and Y. Hu, "Lower Limb Motion Recognition by Integrating Multi-modal Features Based on Machine Learning Method," *ACM Int. Conf. Proceeding Ser.*, Oct. 2020, doi: 10.1145/3424978.3425120.

[5] S.-K. Wu, G. Waycaster, and X. Shen, "Electromyography-based control of active above-knee prostheses," *Control Eng. Pract.*, vol. 19, no. 8, pp. 875–882, 2011, doi: https://doi.org/10.1016/j.conengprac.2011.04.017.

[6] L. Peeraer, B. Aeyels, and G. Van der Perre, "Development of EMG-based mode and intent recognition algorithms for a computer-controlled above-knee prosthesis," *J. Biomed. Eng.*, vol. 12, no. 3, pp. 178–182, 1990, doi: https://doi.org/10.1016/0141-5425(90)90037-N.

[7] E. D. A. A. Nadeem Shah, Abbas Khan, Hira Zahid, "Design and Development of Human Knee Joint Muscle(s) Classification System using Machine Learning Technique," ResearchGate. Accessed: Mar. 13, 2025. [Online]. Available: https://www.researchgate.net/publication/348550210_Design_and_Development_of_Human_Knee_Joint_Muscles_Classification_System_using_Machine_Learning_Technique

[8] Q. W. Honglei An, Hongxu Ma, "Novel Feature Extraction and Locomotion Mode Classification Using Intelligent Lower-Limb Prosthesis," *Machines*, vol. 11, no. 2, p. 235, 2023, doi: https://doi.org/10.3390/machines11020235.

[9] I. S. Dhindsa, R. Agarwal, and H. S. Ryait, "Performance evaluation of various classifiers for predicting knee angle from electromyography signals," *Expert Syst.*, vol. 36, no. 3, p. e12381, Jun. 2019, doi: 10.1111/EXSY.12381.

[10] A. A. D.-S. Tahir Hussain, Nadeem Iqbal, Hafiz Farhan Maqbool, Mukhtaj Khan, Mohammed Ibrahim Awad, "Intent based recognition of walking and ramp activities for amputee using sEMG based lower limb prostheses," *Biocybern. Biomed. Eng.*, vol. 40, no. 3, pp. 1110–1123, 2020, doi: https://doi.org/10.1016/j.bbe.2020.05.010.

[11] R. D. Babu, S. S. Adithya, and M. Dhanalakshmi, "Design and development of an EMG controlled transfemoral prosthesis," *Meas. Sensors*, vol. 36, p. 101399, 2024, doi: https://doi.org/10.1016/j.measen.2024.101399.

[12] R. Schulte, "Roessingh Research & Development-MyLeg database for activity prediction (MyPredict) (dataset)," ResearchData. dataset. Accessed: Mar. 13, 2025. [Online]. Available: https://data.4tu.nl/datasets/01d30db7-95a8-4c39-afb9-4eb1a2f27539

Exploring cGANs for Urdu Alphabets and Numerical System Generation

Suleman Khalil[1], Syed Yasser Arafat[1], Fatima Bibi[1], Faiza Shafique[1]
[1]Department of CS & IT (Mirpur University of Science and Technology (MUST), Mirpur (AJK), Pakistan)
***Correspondence:** sulemankhalil12@gmail.com.

Citation | Khalil. S, Arafat. S. Y, Bibi. F, Shafique. F, "Exploring cGANs For Urdu Alphabets and Numerical System Generation", IJIST, Special Issue. pp 164-187, March 2025
Received | Feb 20, 2025 **Revised |** March 04, 2025 **Accepted |** March 10, 2025 **Published |** March 13, 2025.

Urdu ligatures play a crucial role in text representation and processing, especially in Urdu language applications. While extensive research has been conducted on handwritten characters in various languages, there is still a significant gap in studying raster-based generated images of Urdu characters. This paper presents a generative model designed to produce high-quality samples that closely resemble yet differ from existing datasets. Utilizing the power of Generative Adversarial Networks (GANs), the model is trained on a diverse dataset comprising 40 classes of Urdu alphabets and 20 classes of numerals (both modern and Arabic-style), with each class containing 1,000 augmented images to capture variations. The generator network creates synthetic Urdu character samples based on class conditions, while the discriminator network evaluates their similarity to real datasets. The model's effectiveness is assessed using key metrics such as the Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index (SSIM), and Fréchet Inception Distance (FID).
The results confirm that the proposed GAN-based approach achieves high fidelity and structural accuracy, making it highly valuable for applications in text digitization and Optical Character Recognition (OCR).

Keywords: Generative Adversarial Network, Structural Similarity Index, Fréchet Inception Distance, Peak Signal-to-Noise Ratio, Optical Character Recognition.

Introduction:

The increasing reliance on digital text representations in document archiving, optical character recognition (OCR), and font design highlights the need for efficient methods to generate digital versions of various writing systems. Generative Adversarial Networks (GANs) have numerous practical applications, such as enhancing OCR system performance by training on diverse datasets to improve recognition accuracy. Additionally, GANs assist typographers in creating new fonts by generating diverse ligatures and numeral designs. Research has explored techniques like GANs for generating handwritten text in various languages, including Arabic [1], Bangla [2], Chinese [3], Nepali [4], and Urdu ligatures [5]. These studies primarily used handwritten datasets for model training. While models like Stable Diffusion [6][7] and DALL-E [8] are designed for general purposes, this study aims to bridge a gap by investigating the potential of using Conditional GANs (cGANs) to generate raster images of characters from Urdu, Arabic, and modern numerals.

Urdu belongs to the Indo-Aryan subgroup of the Indo-European language family. Approximately 70 million people speak Urdu as their mother tongue, while around 100 million others, primarily in Pakistan and India, use it as a second language [9]. It is recognized in India's constitution and serves as Pakistan's official language [9]. Significant Urdu-speaking communities exist in the United Arab Emirates, the United States, and the United Kingdom. Notably, Hindi and Urdu are mutually intelligible [10]. The Urdu script consists of 60 characters, derived from 28 Arabic letters and 32 Persian characters, written in Naskh and Nastaliq styles [11]. These two fonts are widely used in different languages: Nastaliq is primarily employed for Urdu, Punjabi, and Sindhi, while Naskh is used for Arabic, Persian, and Pashto [12]. Nastaliq follows an elegant, cursive, right-to-left writing style, featuring ligatures formed by both joining and non-joining alphabets [11][13][14].

Arabic numerals originate from the Hindu-Arabic numeral system, which includes both isolated (non-joining) and connected (joining) alphabets, reflecting its cursive nature [11][13][14]. This system, which originated in India and was later adopted by Arabic mathematicians [15], is often mistakenly considered "Western" or "Latin" digits. The numerals (0-9) are widely used worldwide, including in Urdu and Arabic scripts. Ensuring accuracy and consistency in generating these numeral shapes is crucial for OCR systems recognizing numbers within Urdu text. In this research, "modern numbers" may refer to specific numeral glyphs within certain scripts, such as the extended Arabic-Indic digits used in Urdu [16]. This study differs from previous research by using a raster-based dataset instead of handwritten images for cGAN training. Raster images are widely used in photographs, bitmap graphics, and scanned documents due to their ability to depict a broad range of colors and intricate details. We explore the application of Conditional Generative Adversarial Networks (cGANs), a specialized class of generative models [17]. These networks consist of two primary components: the generator (G) and the discriminator (D). The generator (G) creates synthetic images that mimic real data, while the discriminator (D) evaluates the images to distinguish between real and generated samples.

The primary goal of this research is to develop a system capable of generating realistic Urdu alphabets and numerals to enhance existing datasets. Figure 1(a) presents research papers on GANs that we reviewed. HiGAN+ [18] introduces a framework that separates latent space into style and content factors, allowing independent control over handwriting form while preserving its core content. StackGAN++ [19] generates high-resolution images from text descriptions. VQGAN [20] integrates an autoencoder with a GAN architecture, using vector quantization to capture intricate image details. Mirrorgan [21] introduces a novel text-to-image generation technique based on redescription. TiGAN [22] provides an innovative framework for interactive text-driven image creation and modification. AttnGAN [23] focuses on generating images from text descriptions, while JoinFontGAN [24] employs few-shot learning

to produce high-quality fonts. TeDiGAN [25] uses text descriptions to create and manipulate face images with high realism and control.

Figure 1(a). A recent study of GANS.

Novelty:

The key contributions of our research are as follows:

- We created a novel dataset for training Generative Adversarial Networks (GANs) to address a significant gap in resources for Urdu script generation. This dataset can further aid in developing robust Urdu Optical Character Recognition (OCR) systems.
- We successfully demonstrated the generation of raster images for Urdu characters, Arabic numerals, and modernized numerals using a Conditional Generative Adversarial Network (GCN). To the best of our knowledge, this is the first instance of cGAN being applied to Urdu characters.
- We introduced an innovative approach to data augmentation by generating synthetic raster-based images, potentially transforming dataset creation for various applications.
- We evaluated the effectiveness of a machine learning model trained on synthetic data generated by GCC. The model demonstrated strong performance when tested on real-world data.

Objectives:

Our study aimed to achieve several objectives. First, we sought to address the scarcity of resources for Urdu text, ligatures, and numeral images required for training cGAN models. To bridge this gap, we developed a dataset comprising 40,000 Urdu ligature images and 20,000 numeral images.

Second, our goal was to develop a robust cGAN model capable of generating an unlimited number of realistic Urdu text images. This objective was successfully realized using our custom-built dataset.

Additionally, we assessed the model's robustness using various evaluation metrics to ensure its reliability and effectiveness.

Related Work:

In 2014, Ian Goodfellow and his colleagues introduced the foundational research paper on Generative Adversarial Networks (GANs) [26]. This paper presents a framework where two neural networks—a generative model (G) and a discriminative model (D)—are trained together in an adversarial process. The objective of G is to model the data distribution, while D evaluates the probability that a given sample comes from the training data rather than

being generated by G. This minimax game framework enables effective training through backpropagation and allows the generation of high-quality samples without requiring Markov chains. This approach leverages adversarial networks to generate competitive samples, offering potential solutions to challenges in generative modeling.

One of the notable studies on GAN-based handwritten English text generation was conducted by Eloi Alonso et al. [27]. They proposed an adversarial approach for generating handwritten word images using a bidirectional long short-term memory (LSTM) recurrent neural network to extract the words for generation. These extracted words were then fed as conditional input, along with noise, into the generator networks. An additional recognizer was also incorporated into the network. While the numerical results were promising, the generated images of French and Arabic words were significantly blurry. Wu et al. [28] employed the DCGAN architecture to generate images of tomato leaf disease, thereby augmenting the dataset of diseased leaves. This method, known as dataset augmentation using DCGAN, provided an alternative to traditional augmentation techniques such as translation, rotation, and flipping, which do not always generalize well. Their DCGAN-based augmentation approach improved model recognition accuracy compared to conventional techniques. Furthermore, their findings indicated that the DCGAN-generated results were more convincing in both the visual Turing Test and t-distributed Stochastic Neighbor Embedding analysis.

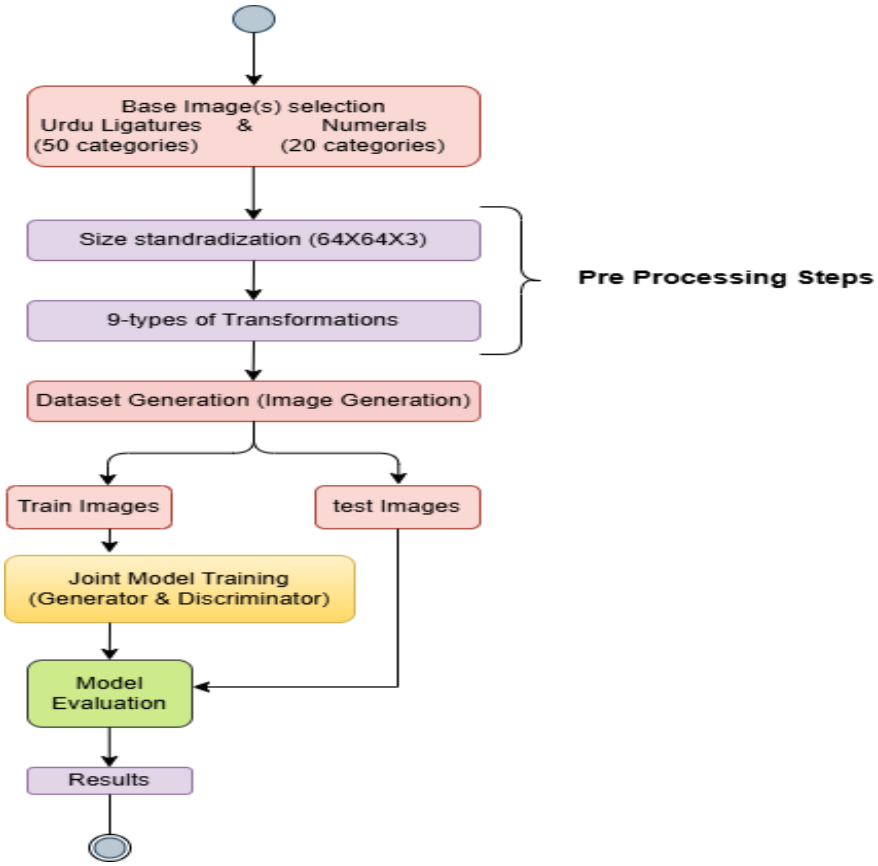


Figure 1(b). The workflow of this research study

In [1], Mustapha et al. used the DCGAN architecture to generate handwritten Bangla digits. They utilized three widely recognized handwritten Bangla datasets—Ekush, their dataset, BanglaLekha-Isolated, and CMATERdb—to achieve their objective. Since the proposed DCGAN efficiently generates Bangla digits, it serves as a reliable model for producing handwritten Bangla digits from random noise. Their study aims to apply the DCGAN architecture to generate Arabic characters. The dataset comprises handwritten

images of 33 Arabic alphabets, from "alef" to "yeh," with each character having 480 images, each sized 32x32 pixels. Recent advancements in generative models for text and image synthesis have significantly enhanced the ability to generate high-quality synthetic data for low-resource languages. While GANs have been widely used for handwritten character generation and font synthesis, newer approaches, such as transformer-based diffusion models, have demonstrated superior results in text-to-image generation. Studies by Ramesh et al. [8] on DALL-E and Rombach et al. [6] on Stable Diffusion indicate that diffusion models trained on large-scale datasets outperform GANs in generating visually coherent text-based images. However, these models require substantial computational resources and large-scale pretraining, making them impractical for low-resource languages like Urdu. In contrast, conditional GANs (cGANs) provide a computationally efficient alternative, enabling fine-grained control over script-specific character generation while maintaining high visual quality.

Urdu Optical Character Recognition (OCR) and script synthesis present unique challenges due to complex ligature formations, cursive structure, and varying diacritic placements. Unlike Latin-based scripts, where characters remain distinct, Urdu letters change shape depending on their position (isolated, initial, medial, or final), making OCR-based training datasets highly diverse and difficult to standardize. Prior research by Arafat & Iqbal [29][30] highlighted that traditional machine learning and neural network-based OCR models struggle with segmentation errors due to overlapping ligatures. Recent studies have explored GAN-based data augmentation to enhance Urdu OCR performance; however, most existing approaches rely on basic GAN architectures (e.g., DCGAN, Pix2Pix) rather than more sophisticated conditional architectures like cGANs. This study addresses this gap by leveraging cGANs for Urdu character and numeral generation, providing a more adaptable framework for low-resource script synthesis and OCR model training.

Despite the success of GANs in generating realistic images, their application in low-resource language processing remains underexplored, particularly in text-based image generation. Previous studies have successfully employed CycleGAN for domain adaptation [22]. Additionally, research by Guan et al. [31] on GAN-based data augmentation for handwritten datasets underscores the potential of generative models in enhancing dataset diversity for OCR tasks. By implementing cGANs with controlled conditioning on Urdu characters and numerals, this study introduces a scalable and effective generative framework for script-based OCR improvement, filling a crucial research gap in Urdu script synthesis.

Dataset:

The dataset used in this study is based on the Urdu Ligatures dataset [29], originally comprising 45,000 unique ligatures. These ligatures included nine distinct transformation types, along with a standard set of ligatures. To enhance the dataset for this study, the following steps were taken:

- **Resolution Standardization:** Using the cv2 library, all images were resampled to a resolution of 64x64 pixels, ensuring consistency across the dataset.
- **Data Augmentation:** Based on techniques outlined in [32], ten different transformation methods were applied to the ligatures. This augmentation process significantly expanded the dataset, generating 1,000 images for each letter and numeral using transformations such as scaling, rotation, and flipping.

As a result, the dataset increased to 60,000 images, comprising 40,000 Urdu characters and 20,000 numerals (both Urdu and Modern Indo-Arabic). The general steps involved in creating the Urdu dataset are shown in Figure 1(b).

This refined and standardized dataset provides a diverse and well-balanced representation of Urdu characters and numerals, making it ideal for machine learning applications such as text recognition and computational linguistics. Furthermore, it addresses

a significant gap in resources for Urdu script generation, laying the foundation for future research in this area.

Developed Methodology:

Model Development:

Conditional Generative Adversarial Networks (cGANs) are a variation of the GAN architecture in which the generator receives an additional conditional input alongside the latent noise. This configuration enables the model to generate data based on specific inputs, making it suitable for tasks such as generating images corresponding to predefined labels or categories. In this research, we employed a cGAN to generate vector-based images of Urdu alphabets, Arabic numerals, and modern digits. The GCC consists of two neural networks: a generator (G) and a discriminator (D), as illustrated in Figure 2. This work is the first to apply cGANs specifically for low-resource languages, such as Urdu, for ligature and numeral generation, highlighting the model's effectiveness in producing diverse and high-quality images.

Generator:

The generator network takes a label and a random array as input, producing images that align with the structure of the training data for that label.

Discriminator:

The discriminator evaluates labeled data batches, combining real data from the training set with generated data, classifying each sample as either "real" or "generated."

Generator Model:

Our generator model takes a latent input and generates an image. Its architecture includes fully connected layers, transposed convolutional layers, batch normalization, and activation functions such as ReLU and Tanh. An input layer is used to define the input layer for latent variables, and a concatenation layer integrates the conditional information with the generated features.

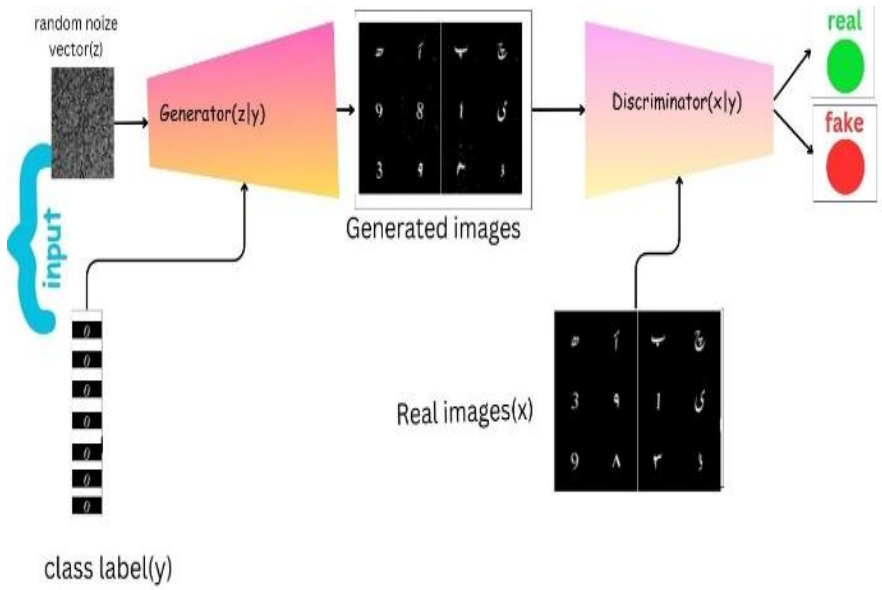


Figure 2. Proposed Methodology

The function layer is a custom layer designed to apply a transformation that converts an appearance into an image. To keep the pixel values of the output image within the range of [-1, 1], a tanh layer is used. Figure 3 illustrates the structure of our generator.

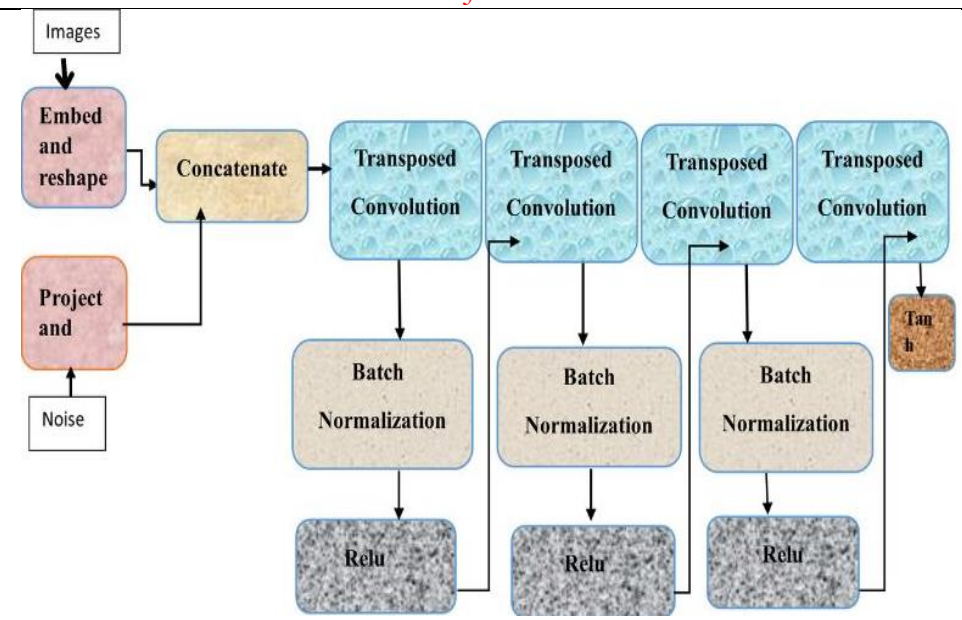


Figure 3. Structure of Generator network.

The generator's loss function is given in Equation 1.

$$\text{lossGenerator} = -\text{mean}(\log(\hat{y}_{\text{Generated}})) \quad (1)$$

The generator's score is determined by the average feasibility assigned by the discriminator to the generated data, as shown in Equation 2.

$$\text{scoreG} = \text{mean}(\text{sigmoid}(\hat{y}_{Generated})) \quad (2)$$

Discriminator Model:

The discriminator model is designed to evaluate the authenticity of an input image, determining whether it is real or generated. Its architecture includes key components such as convolutional layers, batch normalization, dropout, Leaky ReLU activation, and concatenation layers. The Image Input Layer serves as the entry point for image data, while the Dropout Layer helps reduce overfitting by applying regularization. The Leaky ReLU activation function provides a small gradient for negative inputs, aiding efficient learning. Additionally, the Concatenation Layer combines conditional data with image features, improving the model's ability to distinguish between real and generated images. The discriminator's architecture is illustrated in Figure 4.

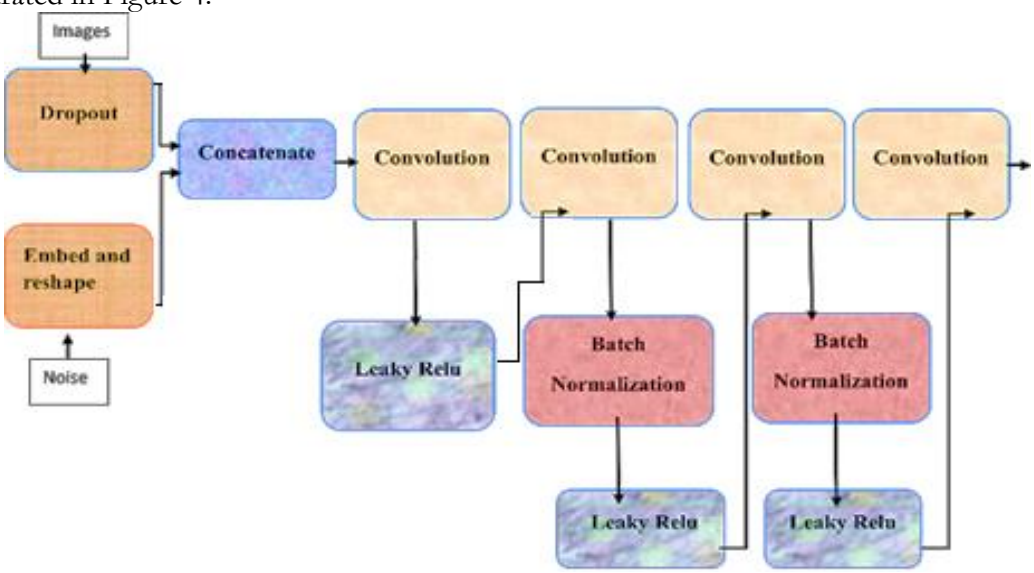


Figure 4. Structure of Discriminator network.

Equation 3 defines the loss function for the discriminator.

$$\text{lossDiscriminator} = -\text{mean}(\log(\hat{y}_{\text{real}})) - \text{mean}(\log(1 - \hat{y}_{\text{Generated}})) \quad 3$$

The discriminator's score is calculated as the average contingency assigned to real and generated data by the discriminator, as shown in Equation 4.

$$\text{scoreD} = 1/2(\text{mean}(\text{sigmoid}(\hat{y}_{real})) + \text{mean}(1 - \text{sigmoid}(\hat{y}_{Generated}))) \quad 4$$

Training Parameters:

With a learning rate of 0.0002, a gradient decay rate of 0.5, and a squared gradient decay rate of 0.999, we used the Adam optimizer to train both the discriminator and generator networks. Adam is widely recognized for its efficient optimization and adaptive learning rate capabilities.

The training process included 45 epochs for the Urdu characters dataset and **81 epochs** for the Urdu numerals dataset, where each epoch represents a complete pass through the entire training data. To balance computational efficiency with the ability to capture diverse gradients from the data, a batch size of 128 samples was used.

The model's performance was evaluated on a validation set every 100 iterations to monitor training progress and reduce the risk of overfitting. Training progress graphs were updated accordingly, and the performance graphs for the generator and discriminator on the Urdu characters and numerals datasets are shown in Figure 5 and Figure 6, respectively.

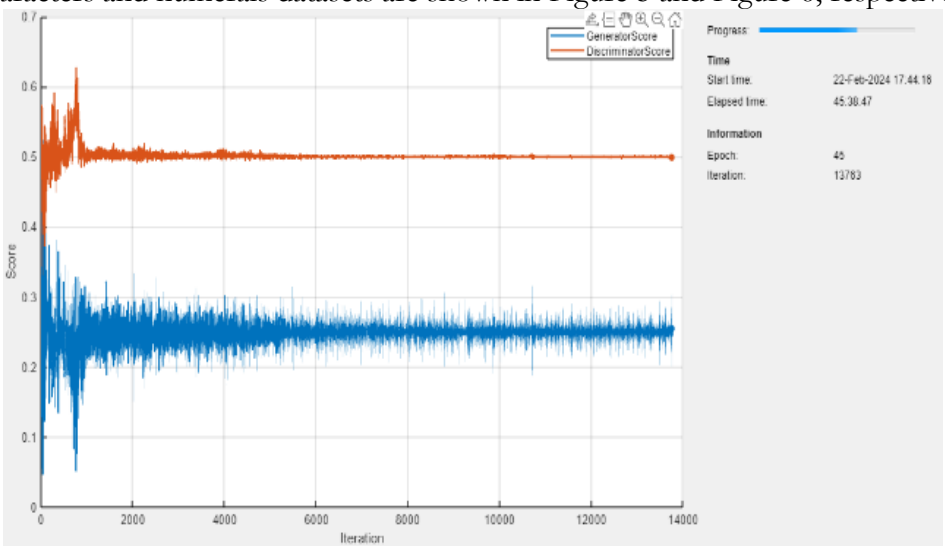


Figure 5. Generator and discriminator score graph for Urdu characters dataset

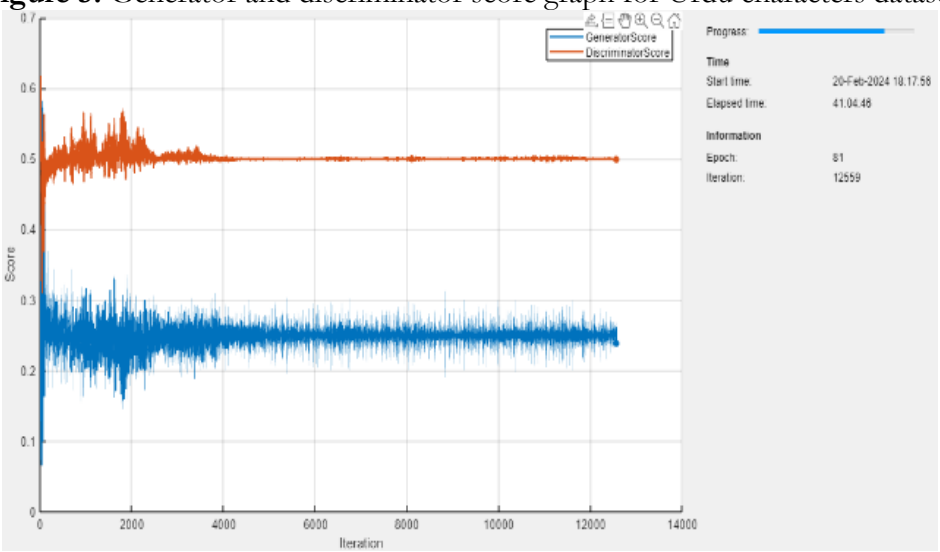


Figure 6. Generator and discriminator score graph for numbers dataset

Evaluation Metrics:

We use PSNR (Peak Signal-to-Noise Ratio) [33]**,** where a higher PSNR value indicates better image quality, and SSIM (Structural Similarity Index) [34]**,** where values closer to 1 represent higher similarity, while values closer to 0 indicate lower quality. A low PSNR score suggests significant numerical differences between images.

Additionally, we evaluate our model using FID (Fréchet Inception Distance) [35]**.** A perfect FID score of **0.0** means the two sets of images are identical, while lower FID values indicate greater similarity or a closer statistical alignment between them. The formulas for these metrics are provided in Table 1**.**

Table 1. Formulas for measurements.

Measurements	Formulas
FID	$d^2 = \|\|mu_1 - mu_2\|\|^2 + Tr(C_1 + C - 2 * sqrt(C_1 * C_2))$ 5
SSIM	$\frac{(2\mu_x 2\mu_y + c1)(2\sigma_{xy} + c2)}{(\mu_x^2 + \mu_y^2 + c1)(\sigma_x^2 + \sigma_y^2 + c2)}$ 6
PSNR	$10\log_{10}(\frac{R^2}{MSE})$ 7

Results and Discussion:

The results strongly indicate that the cGAN-based model can generate high-quality and visually realistic characters, including both Urdu alphabets and numerals (Arabic and modern)**.**

When assessing generative models, key evaluation metrics such as Peak Signal-to-Noise Ratio (PSNR), Fréchet Inception Distance (FID), and Structural Similarity Index Measure (SSIM) serve as essential benchmarks. Each metric captures a different aspect of generation quality:

FID (Fréchet Inception Distance).

A lower FID value indicates that the generated images closely resemble the real dataset in terms of feature distributions. The reported top FID scores—0.0100 for the Urdu alphabet dataset and 0.0006 for the numeral dataset (as shown in Table 2)—are exceptionally low. This suggests that the synthetic characters are nearly indistinguishable from real ones at a high-level statistical representation. The average FID values—0.0055 for the alphabet set and 0.0035 for the numerals—are still highly promising. However, the top scores emphasize the model's potential under optimal training configurations or specific hyperparameter settings.

A few sample outputs are illustrated in Figures 7(a), 7(b), and 7(c)**.**

PSNR (Peak Signal-to-Noise Ratio): Measures how closely the generated images resemble a reference image at the pixel level. Higher PSNR values indicate less distortion and better image quality. The top PSNR scores—25.699 for Urdu alphabets and 28.844 for numerals—demonstrate that, at their best, the generated characters exhibit minimal noise and high pixel-level fidelity.

While the average PSNR values are slightly lower, they remain strong (23.4098 for alphabets and 25.9774 for numerals, as shown in Table 2). This indicates that, on average, the generated outputs are clean, detailed, and well-defined.

Some sample outputs are illustrated in Figures 7(a), 7(b), and 7(c)**.**

SSIM (Structural Similarity Index Measure): evaluates structural and perceptual similarity, ensuring that the generated images preserve the shapes, edges, and patterns characteristic of the original character forms.

The top SSIM scores—0.9056 for alphabets and 0.9480 for numerals (as shown in Table 2)—are remarkably high, indicating that the generated characters closely resemble real examples in terms of structural integrity.

The average SSIM values—0.8320 for alphabets and 0.8347 for numerals—are also strong, demonstrating consistent structural fidelity across different generation conditions.

Some sample outputs are illustrated in Figures 7(a), 7(b), and 7(c).

MS-SSIM (Multi-Scale Structural Similarity Index Measure): is an enhanced version of SSIM that evaluates images at multiple levels of detail, from fine to coarse. This approach improves the traditional SSIM metric by comparing structural features such as brightness, contrast, and patterns across different resolutions, making it more aligned with how humans perceive visual information.

Our experimental results reveal promising outcomes, demonstrating that cGANs effectively address training data sparsity for low-resource languages. The model successfully generated coherent and contextually relevant Urdu text and numerals, emphasizing the potential of cGANs for text generation in resource-constrained environments.

This capability is significant for several reasons:

1. **Expanding Linguistic Diversity** – cGANs facilitate automated content generation, supporting underrepresented languages.
2. **Enhancing NLP Applications** – Synthetic text can augment training datasets, improving performance in text classification, sentiment analysis, and machine translation for low-resource languages.

Additionally, our findings contribute to the broader field of language modeling and GAN-based research. By demonstrating that cGANs can generate meaningful Urdu text, our study paves the way for future research on GAN-based NLP models, particularly for languages that are underrepresented in digital spaces.

The results of our various experiments, including top and average scores for each evaluation metric, are illustrated in graphs and figures throughout the study.

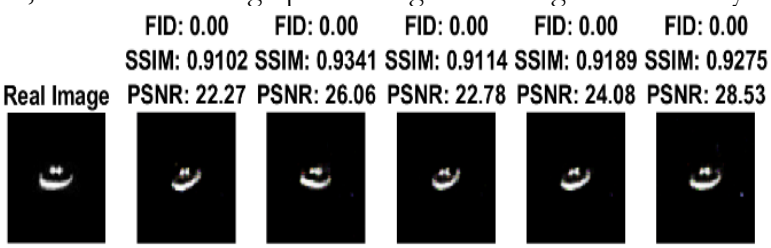


Figure 7(a). Evaluation scores for Urdu character 'stay'

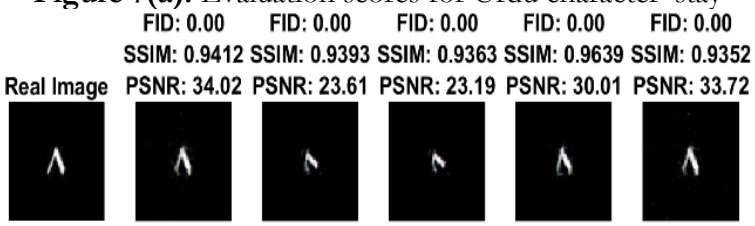


Figure 7(b). Evaluation scores for Arabic number '8'

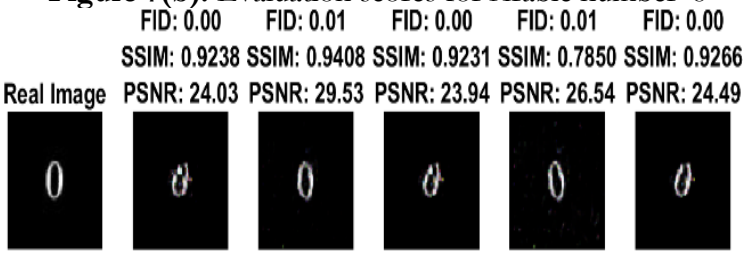


Figure 7(c). Evaluation scores for the modern number '0'

We calculated both the average scores and top scores for PSNR, FID, and SSIM for each class in the numbers dataset. The results are presented in Figure 8(a), Figure 8(b), and Figure 8(c).

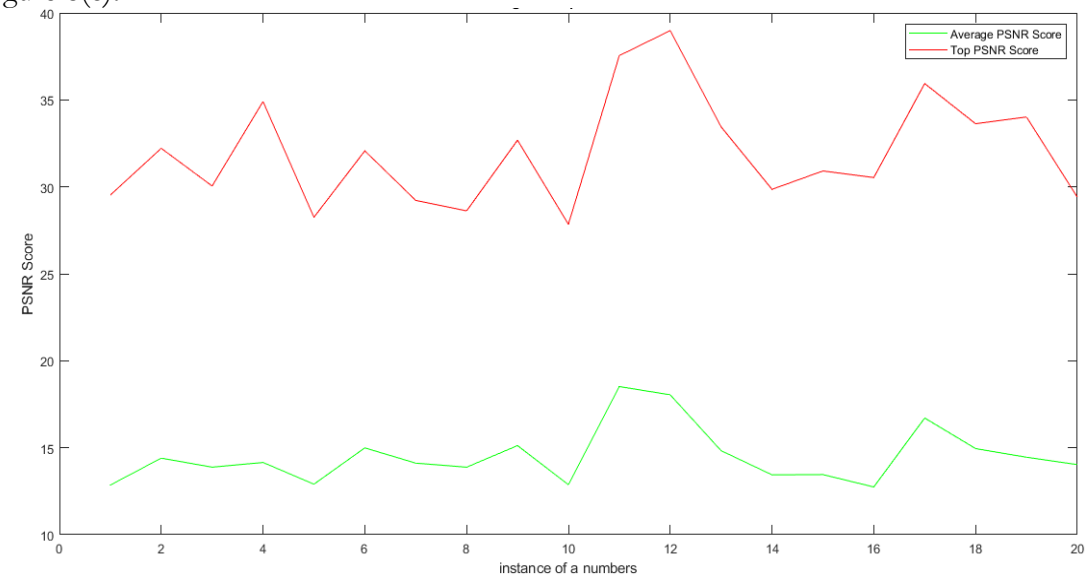


Figure 8(a). Average vs top PSNR score of numbers dataset

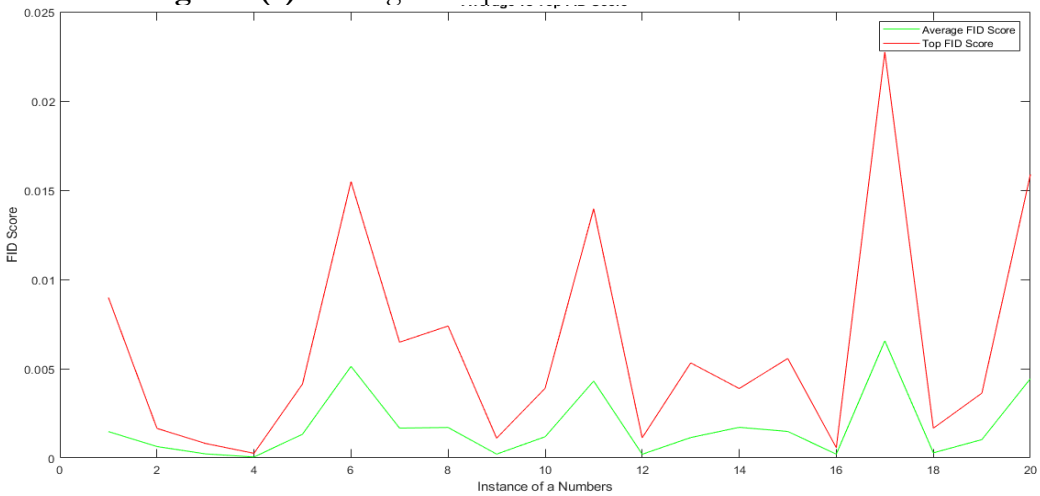


Figure 8(b). Average vs top FID score for numbers

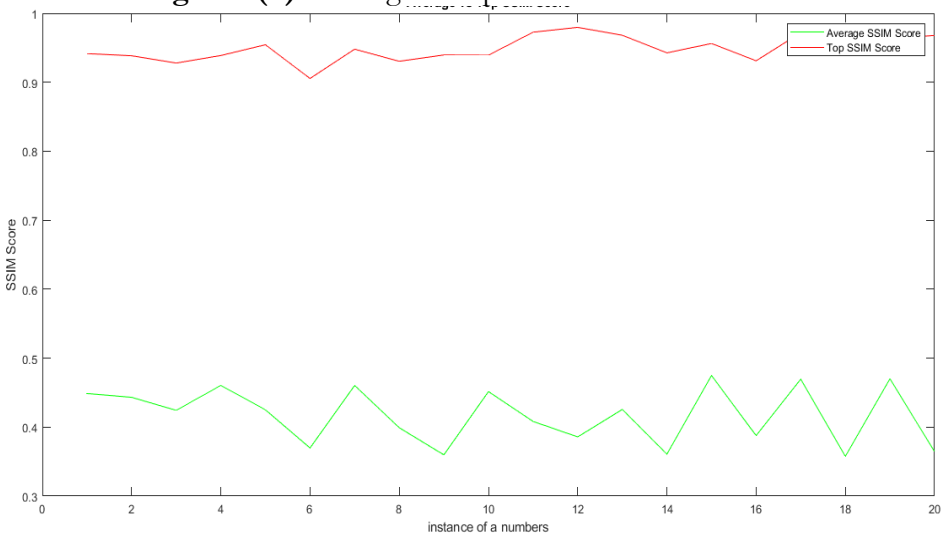


Figure 8(c). Average vs top SSIM score for numbers

The average scores and top scores for PSNR, FID, and SSIM for each class in the Urdu characters dataset are shown in Figure 9(a), Figure 9(b), and Figure 9(c).

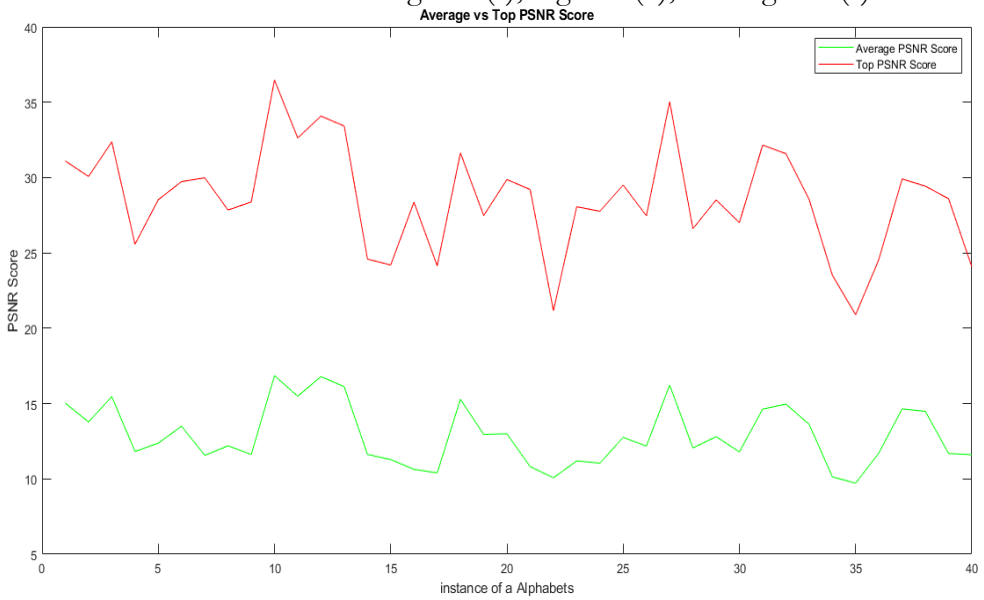


Figure 9(a). Average vs PSNR score for Urdu alphabets

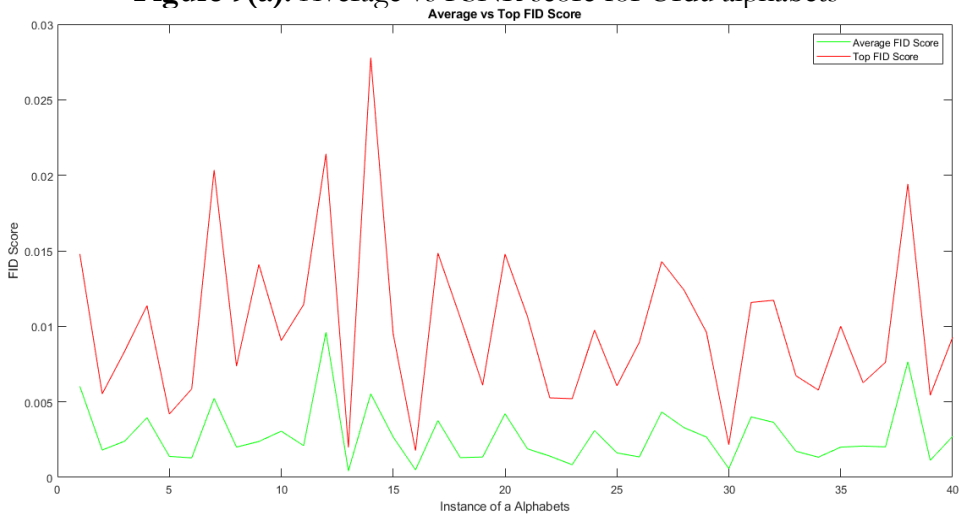


Figure 9(b) Average vs top FID score for Urdu alphabets

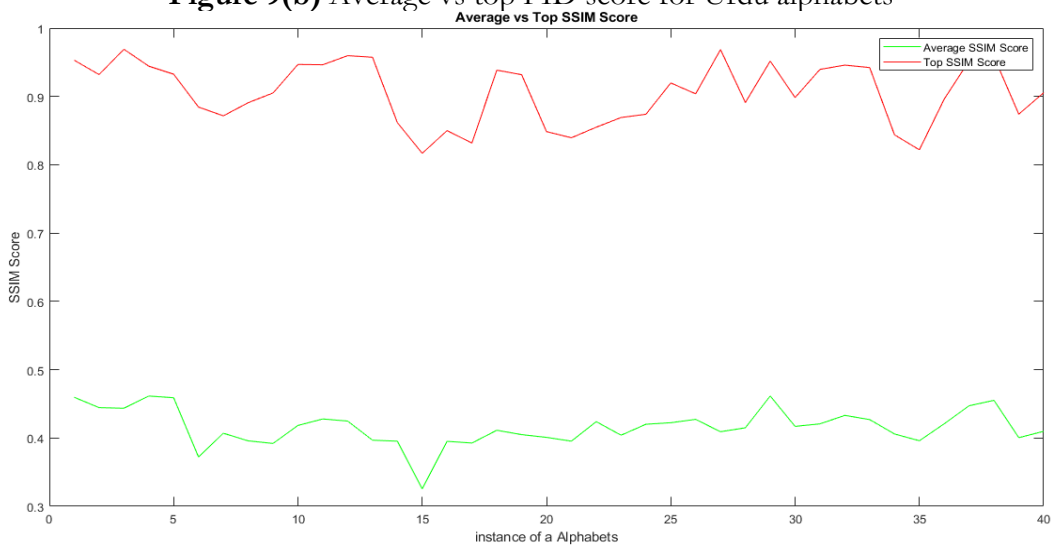


Figure 9(c). Average vs top SSIM score for the Urdu alphabet

We calculated the average scores and top scores for both datasets, as presented in Table 2.

Table 2. Various Evaluation Results

Dataset	FID		PSNR		SSIM	
Urdu alphabets	Average score	Top score average	Average score	Top score average	Average score	Top score average
	0.0055	0.0100	23.4098	25.6999	0.8320	0.9056
Numbers	0.0035	0.0062	25.9774	28.8484	0.8347	0.9480

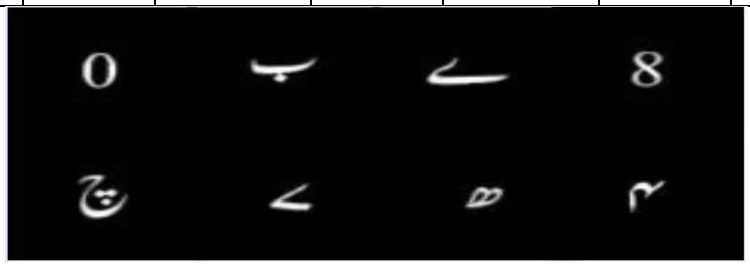


Figure 10. Samples of real images

Figure 11. samples of generated images

The generated images in Figure 11 closely resemble the real images shown in Figure 10.

- Figure 12(a) presents the average PSNR score bar chart, while Figure 12(b) displays the top PSNR score for each class of Urdu characters.
- Figure 13(a) and Figure 13(b) illustrate the average and top SSIM scores, respectively.
- Figure 14(a) and Figure 14(b) show the average and top FID scores, respectively.

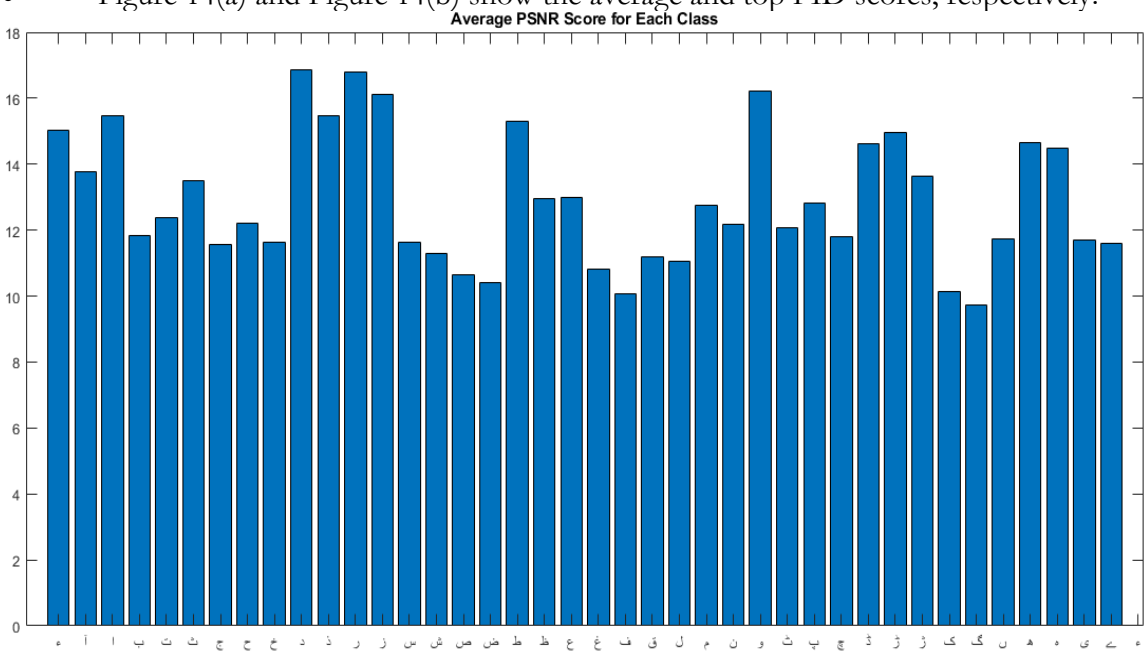


Figure12(a). Average PSNR score bar chart for each class of Urdu Characters.

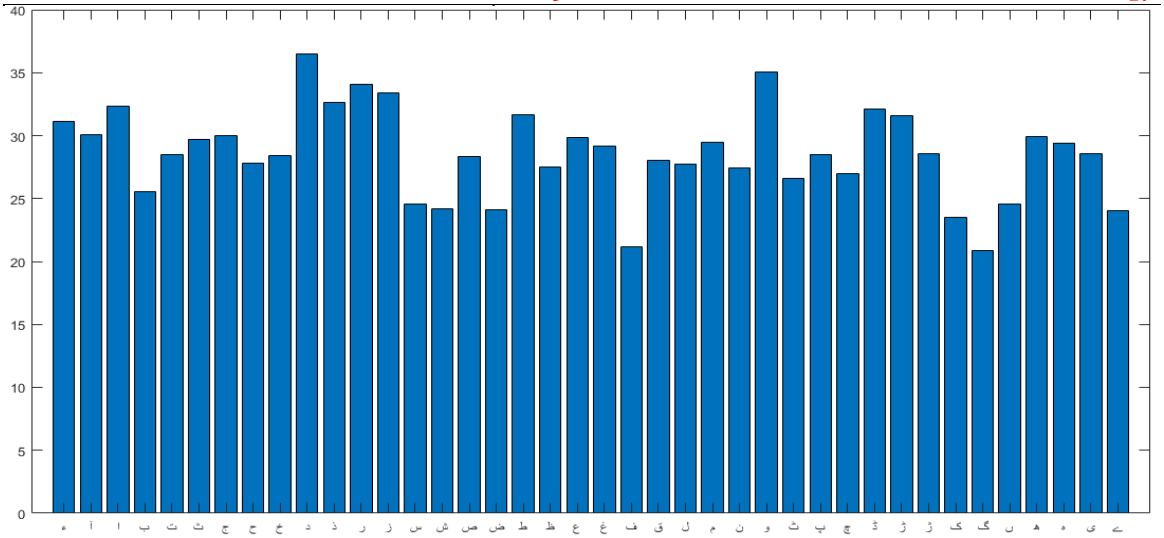


Figure12 (b). Top PSNR score bar chart for each class of Urdu characters

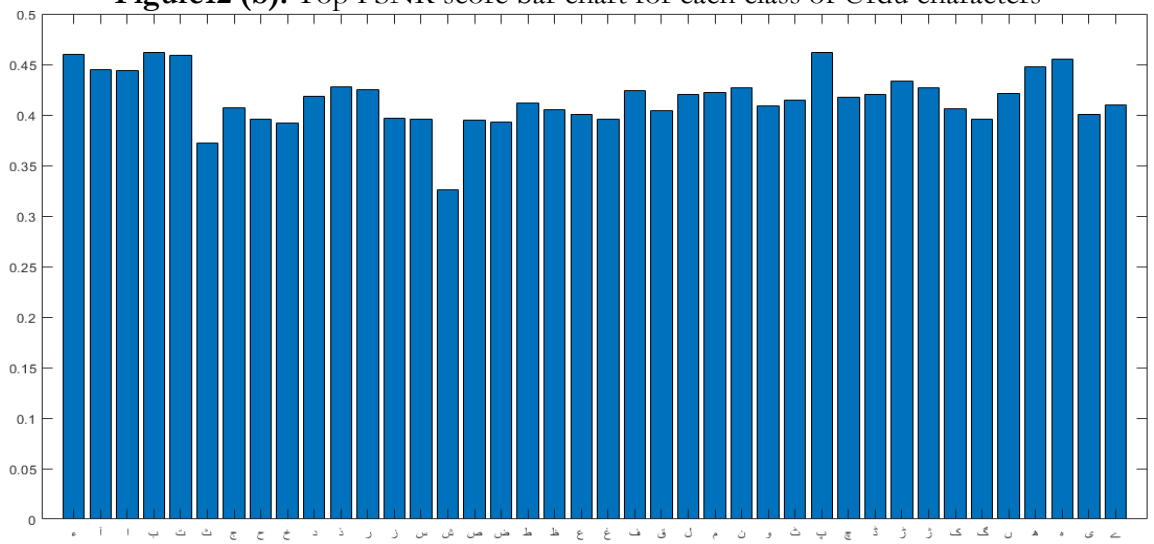


Figure13(a). Average SSIM score bar chart for each class of Urdu characters

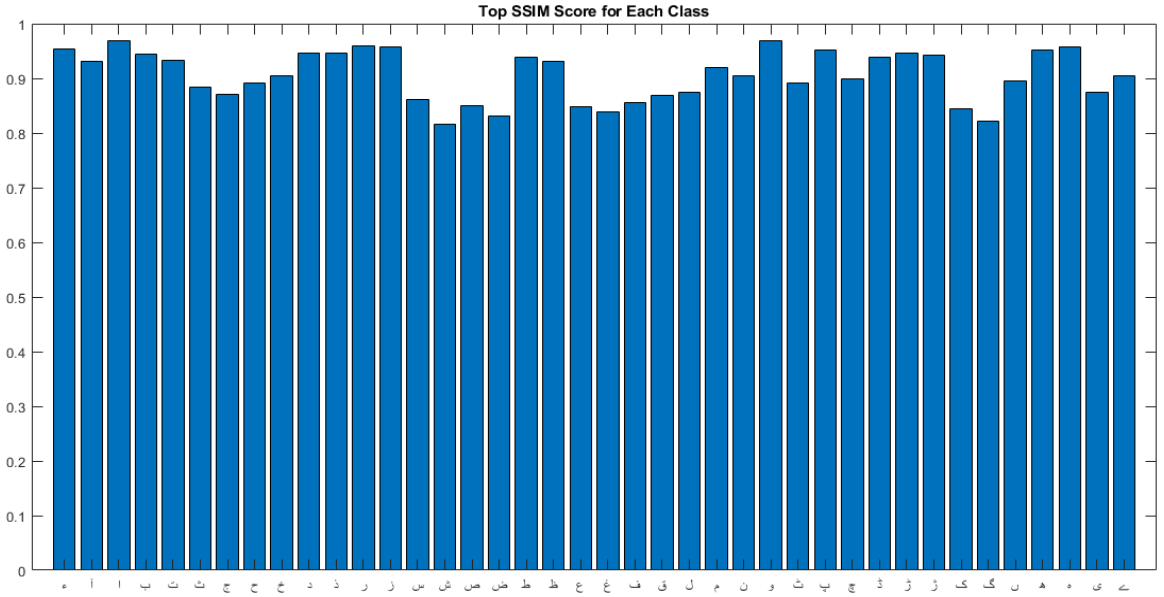


Figure13 (b). Top SSIM score bar chart for each class of Urdu characters

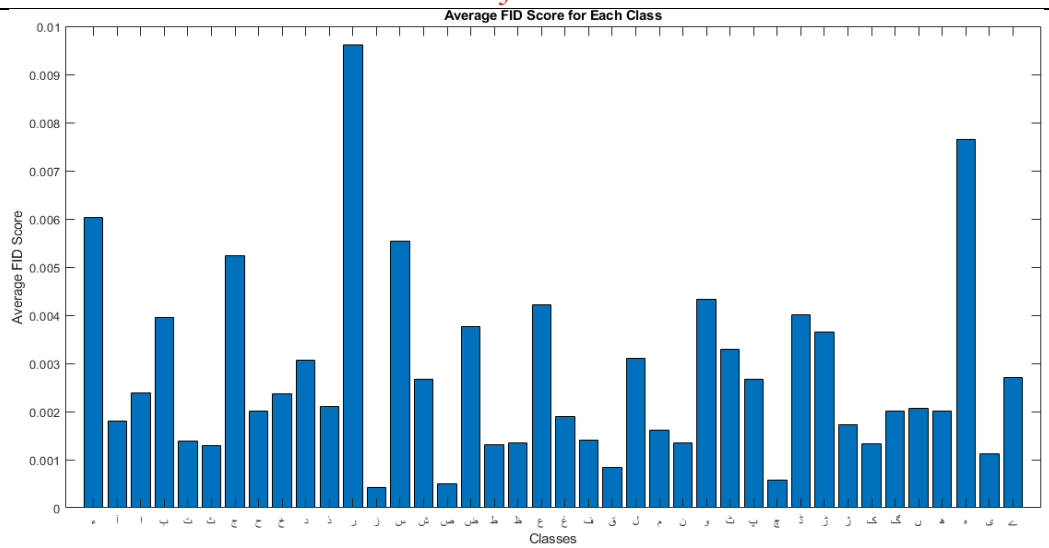


Figure14(a). Average FID score bar chart for each class of Urdu characters

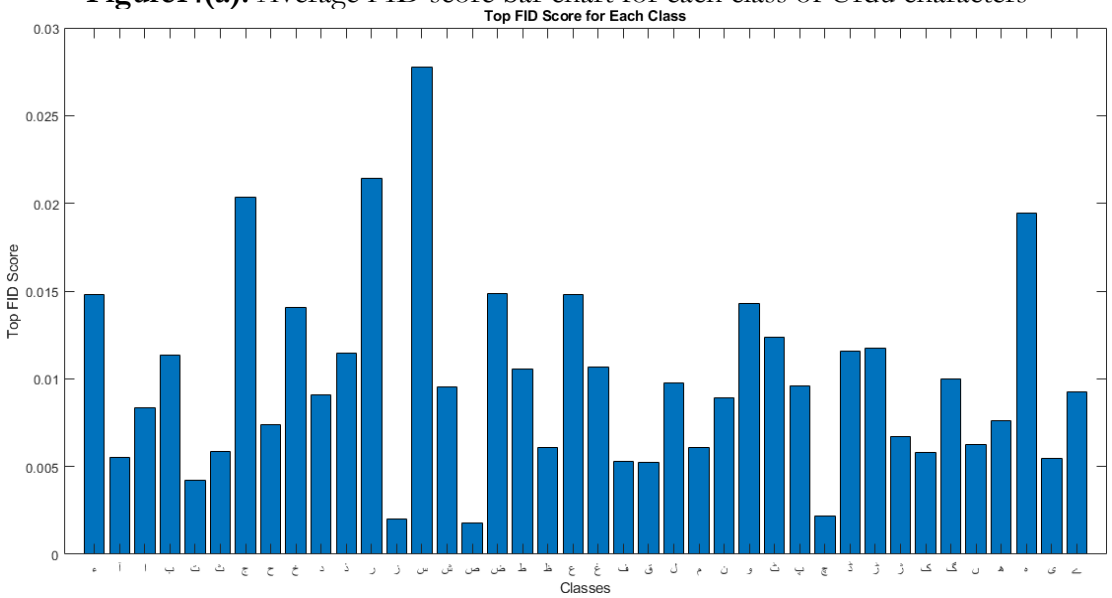


Figure14 (b). Top FID score bar chart for each class of Urdu characters

Figure 15(a) and Figure 15(b) present the average and top FID scores for each class of numbers, respectively.

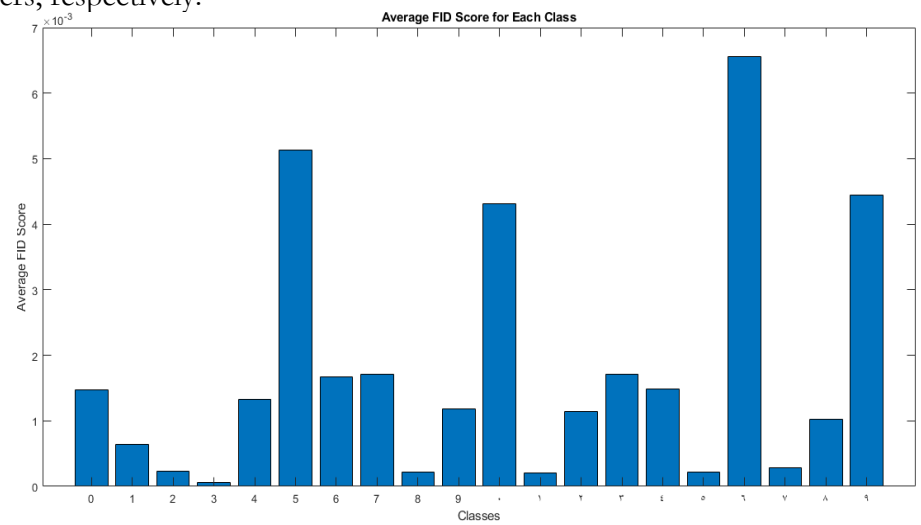


Figure15 (a). Average FID score bar chart for each class of numbers

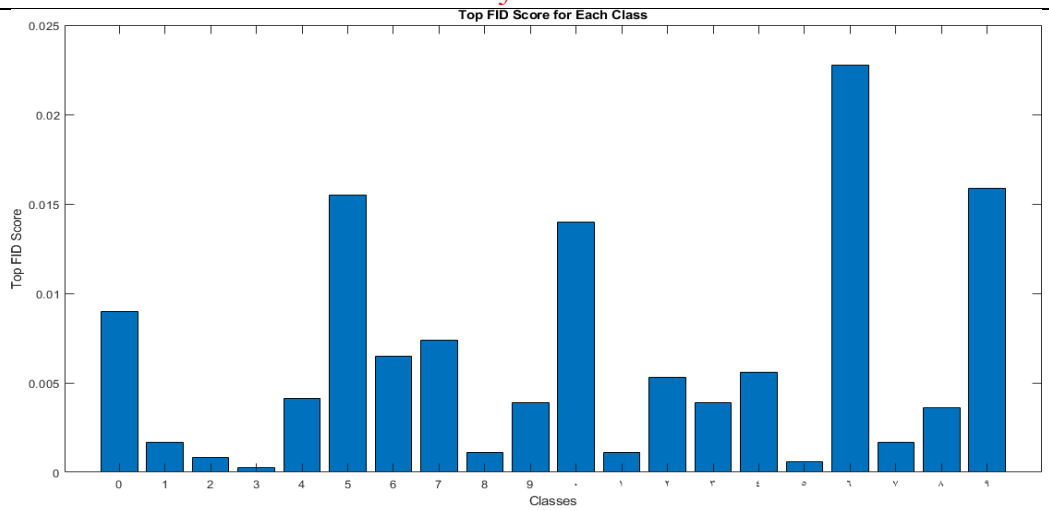


Figure15 (b). Top FID score bar chart for each class of numbers

Figure 16(a) and Figure 16(b) display the average and top PSNR scores for each class of numbers, respectively.

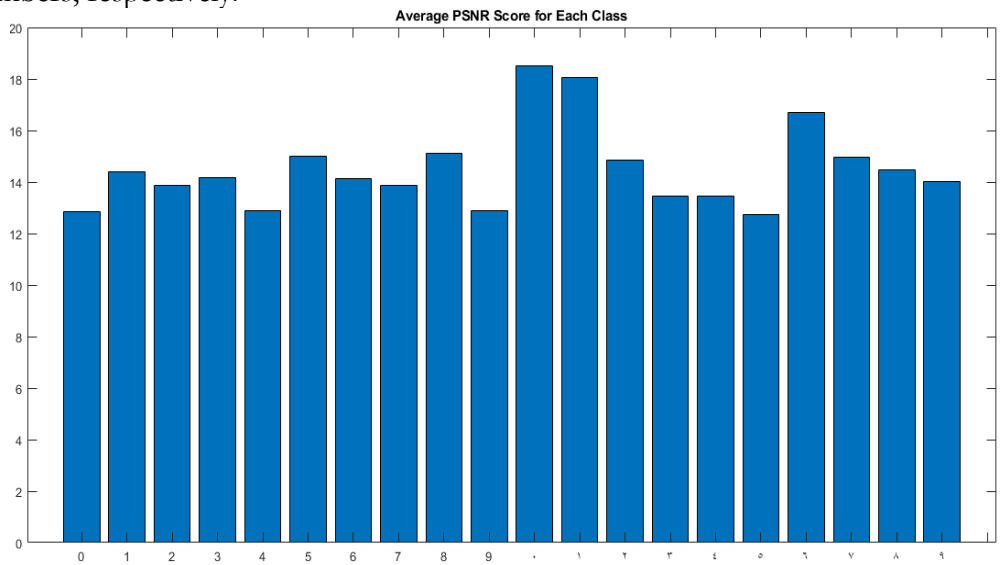


Figure16 (a). Average PSNR score bar chart for each class of numbers

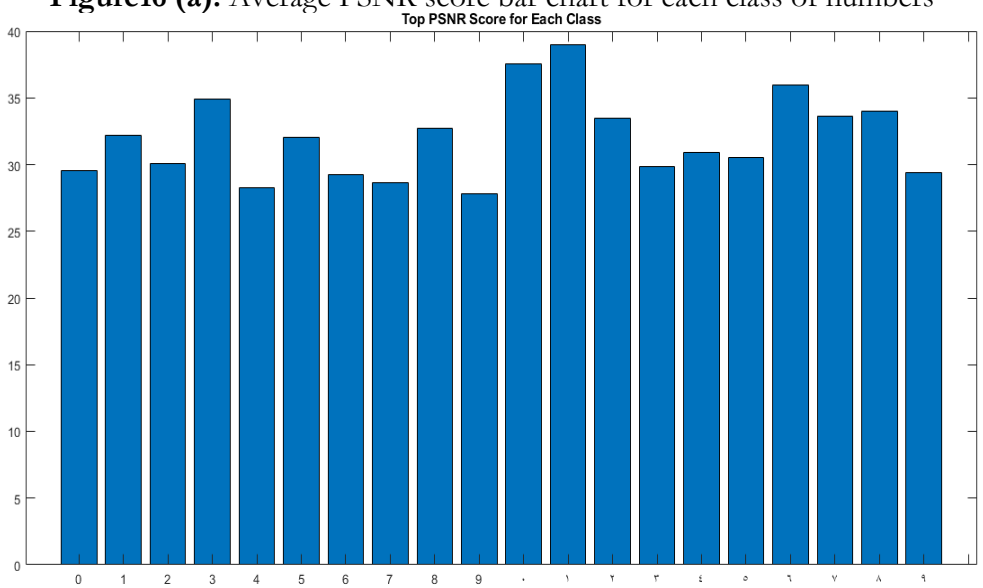


Figure16 (b). top PSNR score bar chart for each class of numbers

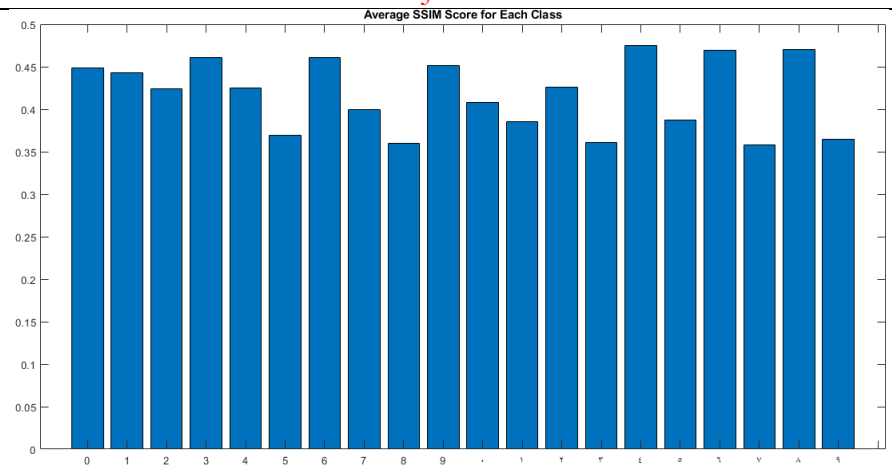


Figure17 (a). Average SSIM score bar chart for each class of numbers

Figure 17(a) and Figure 17(b) present the average and top SSIM scores for each class of numbers**,** respectively.

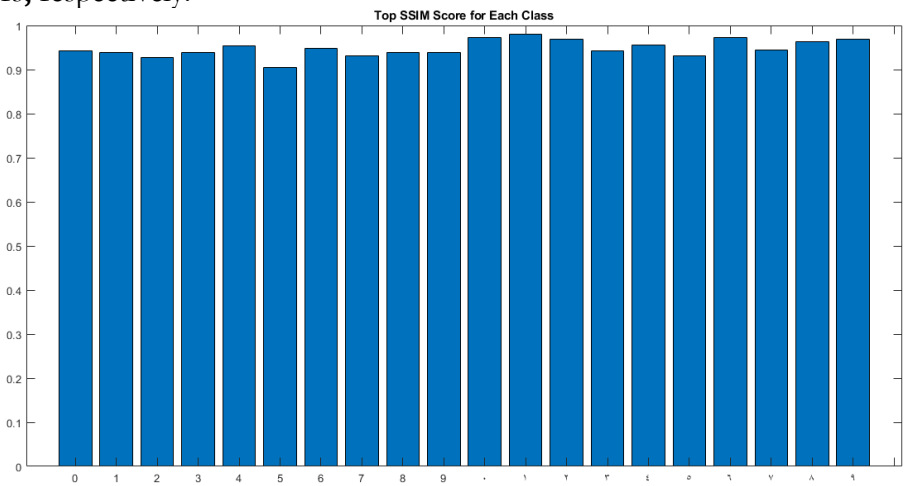


Figure17 (b) Top SSIM score bar chart for each class of numbers

We also compared our results with the Arabic cGAN for character generation using the MS-SSIM metric and obtained encouraging results**.** The findings are presented in Table 3**.**

Table 3. Comparison of MS-SSIM scores for Urdu and Arabic Characters

Dataset	Classes	MS-SSIM
Arabic characters dataset [1]	28	0.6350
Our Urdu characters dataset	40	0.6683

Discussion:

The findings of this study highlight that Conditional Generative Adversarial Networks (cGANs) offer an effective method for generating high-quality Urdu alphabet and numeral images. This contributes significantly to dataset augmentation for OCR training and script synthesis**.** The proposed model effectively captures the unique stylistic characteristics of Urdu script**,** addressing key challenges such as ligature complexity, cursive nature, and contextual character shape variations**.** Performance evaluations using PSNR, SSIM, and FID confirm that cGAN-generated images exhibit high visual fidelity and structural accuracy, making them well-suited for real-world OCR applications**.**

In comparison to traditional handwritten dataset augmentation methods**,** cGANs provide a scalable and automated solution for expanding Urdu script datasets—eliminating the need for extensive manual annotation. A key finding of this study is the superiority of cGANs over Variational Autoencoders (VAEs) and diffusion models in Urdu text generation. While diffusion models have gained popularity in text-to-image synthesis, their high

computational cost and training instability make them less practical for low-resource languages.

Our results show that cGANs achieve significantly lower FID scores (0.0055 for alphabets, 0.0035 for numerals), demonstrating superior image realism and text coherence compared to VAEs. However, despite their strong performance, cGANs still face challenges in handling highly stylized ligatures and complex diacritic placements. These limitations could be addressed through hybrid architectures incorporating transformers or attention mechanisms, enhancing the model's ability to generate more accurate and detailed Urdu text.

While this study successfully generates high-quality synthetic Urdu characters, certain deployment challenges remain unaddressed:

- **Computational Efficiency** – Further research is needed to explore the feasibility of cGANs in edge computing environments, as real-world OCR systems require lightweight, low-latency models.
- **Adversarial Robustness** – GAN-generated text images may be susceptible to perturbations, which could negatively impact OCR model performance.

Future Work:

To address these challenges, future studies should focus on:

- **Optimizing model efficiency** for real-time applications.
- **Integrating adversarial defense mechanisms** to enhance robustness against distortions and occlusions.

Comparative Analysis:

In comparison to Diffusion Models [36] and VAEs [37], which offer alternative text generation solutions:

- **Diffusion Models** are well-suited for high-quality text generation when abundant data is available.
- **VAEs** perform better in scenarios where data is sparse, but diversity is less critical.
- **cGANs**, however, provide a balanced approach, generating contextually aware, accurate, and diverse text, even with limited data availability.

Conclusion:

Overall, can emerge as a compelling solution for low-resource language generation, such as Urdu, striking a balance between diversity, contextual relevance, and practical applicability.

Importance of These Metrics and Their Strong Values:

These evaluation metrics and their high scores are crucial for several reasons:

1. **Assessing Image Quality**
   - Peak Signal-to-Noise Ratio (PSNR) measures how closely the generated images match reference images at the pixel level.
   - Higher PSNR values indicate less distortion, ensuring that the generated text remains clear and readable.
2. **Evaluating Structural Similarity:**
   - Structural Similarity Index Measure (SSIM) ensures that the shapes, edges, and patterns in the generated images preserve the structural integrity of real Urdu characters.
   - High SSIM scores confirm that the generated text remains visually and contextually accurate.
3. **Measuring Statistical Alignment**
   - Fréchet Inception Distance (FID) evaluates the similarity between real and generated datasets in terms of feature distributions.
   - Lower FID values suggest that the synthetic images are nearly indistinguishable from real data, enhancing the reliability of generated text.
4. **Enhancing OCR Training and Recognition:**

- High-quality synthetic images contribute to better dataset augmentation for OCR models.
- This improves the accuracy of text recognition systems, especially in low-resource languages like Urdu.

5. **Supporting NLP Applications**
   - These strong values validate the effectiveness of cGAN-generated text for applications in machine translation, sentiment analysis, and text classification.
   - A well-generated Urdu script dataset helps improve performance in natural language processing (NLP) tasks.
6. **Ensuring Practical Use for Real-world Applications**
   - Maintaining visual fidelity ensures that synthetic text can be used for digital typography, font generation, and educational tools.
   - This broadens the application of AI-driven text synthesis **in** linguistic research, publishing, and digital content creation.

These metrics collectively demonstrate that cGANs are a powerful tool for Urdu text generation**,** offering high-quality, structurally accurate, and statistically reliable synthetic images.

Validation of Dataset Quality:

Significance of High FID, PSNR, and SSIM Scores:

Achieving strong FID, PSNR, and SSIM scores serves as a validation of the newly developed dataset and training protocols. These high scores confirm several key aspects:

1. **Dataset Quality & Representativeness**
   - A well-curated dataset is crucial for training generative models.
   - High scores indicate that the dataset contains diverse and high-quality samples, making it effective for training robust generative models.
2. **Realism & Fidelity of Generated Images**
   - Low FID scores suggest that the generated images closely match real-world samples in feature space.
   - High PSNR values confirm minimal pixel-level distortion, ensuring that the generated characters retain clarity and detail.
   - High SSIM scores validate that the structural properties of the generated characters remain true to the original script.
3. **Effectiveness of Training Protocols**
   - High metric scores reflect the efficacy of the training pipeline**,** including data preprocessing, augmentation, and model optimization techniques.
   - This confirms that the model **learns efficiently** and generalizes well across different character classes.
4. **Suitability for OCR & NLP Applications**
   - A high-quality generative model enhances OCR dataset augmentation, making it useful for real-world Urdu script recognition.
   - Text-based AI applications, such as handwriting recognition and font generation, benefit from synthetic yet highly realistic text samples.
5. **Potential for Future Research & Development**
   - Strong metric scores indicate that the training methodology can be used for other low-resource languages and extended to various generative AI tasks.
   - This opens doors for further refinements, including the integration of transformers, diffusion models, or hybrid architectures for improved Urdu text generation.

By achieving strong FID, PSNR, and SSIM scores, this study validates the dataset's robustness and confirms that the training protocols effectively guide the model toward generating realistic, high-quality Urdu script images.

Benchmarking Performance.

Establishing Baseline Performance Metrics for a New Dataset:

When working with a new dataset, setting baseline performance metrics such as FID, PSNR, and SSIM provides a critical reference point for evaluating and improving model performance. These baselines serve several key functions:

1. **Benchmark for Future Improvements**
   - Establishing initial scores allows researchers to compare future model iterations and measure progress over time.
   - Any enhancements in preprocessing, augmentation, or architecture can be directly evaluated against the baseline.
2. **Standardized Performance Evaluation**
   - A baseline provides a consistent framework to assess different models or training strategies.
   - This is especially important for comparing cGANs, VAEs, Diffusion Models, and hybrid architectures on the same dataset.
3. **Dataset Suitability for Generative Tasks**
   - If the baseline scores are too low, it may indicate that the dataset requires better quality control, balancing, or augmentation.
   - Strong baseline scores suggest the dataset is sufficiently diverse and informative for training high-quality generative models.
4. **Guiding Hyperparameter Tuning**
   - Researchers can use baseline metrics to fine-tune learning rates, batch sizes, and regularization techniques.
   - This prevents unnecessary adjustments and provides a data-driven approach to optimizing performance.
5. **Comparability Across Studies**
   - When published, baseline scores enable other researchers to reproduce results and compare new techniques against an established reference.
   - This fosters scientific rigor and promotes collaboration in generative AI research.

By setting baseline FID, PSNR, and SSIM scores, researchers create a solid foundation for evaluating generative models, ensuring that progress is measurable, reproducible, and meaningful.

Practical Applications.

In domains such as typography design**,** calligraphy digitization, and OCR (Optical Character Recognition) pre-training, the quality of generated characters plays a critical role in determining their usability.

Why Character Quality Matters?

1. **Typography & Font Design**
   - High-quality synthetic characters help in designing new fonts with stylistic consistency**.**
   - cGANs enable the automatic generation of script-specific typefaces, reducing manual effort.
2. **Calligraphy Digitization**
   - Many historical and artistic scripts lack digital representation.
   - AI-generated characters preserve intricate calligraphic details, making them usable in modern applications.
3. **OCR Pre-Training & Dataset Augmentation**
   - OCR models require large, diverse datasets for high accuracy.
   - High-fidelity synthetic text improves OCR performance by providing additional training samples, especially for low-resource languages.

By ensuring high visual fidelity and structural accuracy, cGAN-generated text enhances usability across digital design, linguistic research, and AI-driven text recognition.

Generalization Potential:

Significance of Strong Average Scores:

While peak results highlight the best-case performance of a model, strong average scores indicate consistency across different samples, making the model more reliable in real-world applications.

1. **Robustness Across Variability**
   - A high average PSNR, SSIM, and FID suggest that the model performs well on **most samples**, not just a few ideal cases.
   - This is essential for text generation tasks, where variations in handwriting, font styles, and distortions can affect recognition.
2. **Scalability & Generalization**
   - A model with strong average performance can generalize well across different datasets, character styles, and script variations**.**
   - This is particularly valuable for low-resource languages, where real-world **dataset augmentation** is required.
3. **Foundation for Further Research**
   - The presented results confirm that cGAN-based text generation is effective and provides a benchmark for future improvements.
   - The datasets and methodologies established here can be expanded and refined for broader linguistic and AI applications.

Conclusion:

By achieving both top-performing and high-average scores, this study reinforces that cGANs offer a reliable and scalable approach **for** text and character-based image synthesis**,** paving the way for further advancements in generative AI for script-based languages.

Conclusion:

This research demonstrated the potential of using a cGAN to generate raster images of different writing systems, possibly marking the first successful attempt for Urdu characters and numerals. The GCN successfully produced realistic and recognizable representations of Urdu script, Arabic numerals, and modern numerals, demonstrating its effectiveness in font generation, optical character recognition, and data augmentation. Further study and development can explore multiple GAN designs, increase the quality and consistency of produced pictures, and look into particular applications in a range of language processing and design domains. This research opens doors for further exploration in generating raster-based representations for diverse writing systems.

This research demonstrated the potential of using cGANs to generate raster images for different writing systems, possibly marking the first successful attempt at generating Urdu characters and numerals. The model effectively produced realistic and recognizable representations of:

- Urdu script
- Arabic numerals
- Modern numerals

Key Contributions and Implications:

1. **Applications in Font Generation & OCR**
   - The study highlights cGANs' effectiveness in font design by generating high-quality, script-aware text.
   - It enhances Optical Character Recognition (OCR) systems by providing diverse synthetic data for training.
2. **Advancements in Data Augmentation**

- By generating realistic synthetic text, cGANs can expand low-resource datasets without manual annotation.
- This helps improve NLP and OCR models for underrepresented scripts.

3. **Future Research Directions**

- Exploring multiple GAN architectures (e.g., StyleGAN, BigGAN) to refine quality and consistency.
- Hybrid models with transformers to enhance context-aware text generation.
- Extending to other scripts such as Persian, Pashto, and Sindhi to broaden the research impact.

This research lays the foundation for future exploration in generating raster-based representations for diverse writing systems, paving the way for breakthroughs in computational linguistics, typography, and AI-driven text synthesis.

Acknowledgment: We are very thankful to Mirpur University of Science and Technology for the cordial environment for research.

Author's Contribution:

Suleman Khalil: Conceptualization, Methodology, Software, investigation, Writing – review & editing. *Syed Yasser Arafat:* Conceptualization, Methodology, Software, investigation, Writing – review & editing, Project Administration. *Fatima Bibi:* Conceptualization, Methodology, investigation, Writing – review. *Faiza Shafique:* Methodology, Software, investigation, Writing – review

Conflict of interest: There is no Conflict of Interest.

References:

[1] I. B. Mustapha, S. Hasan, H. Nabus, and S. M. Shamsuddin, "Conditional Deep Convolutional Generative Adversarial Networks for Isolated Handwritten Arabic Character Generation," *Arab. J. Sci. Eng.*, vol. 47, no. 2, pp. 1309–1320, Feb. 2022, doi: 10.1007/S13369-021-05796-0/METRICS.

[2] S. Haque, S. A. Shahinoor, A. K. M. S. A. Rabby, S. Abujar, and S. A. Hossain, "OnkoGan: Bangla Handwritten Digit Generation with Deep Convolutional Generative Adversarial Networks," *Commun. Comput. Inf. Sci.*, vol. 1037, pp. 108–117, 2019, doi: 10.1007/978-981-13-9187-3_10.

[3] B. Chang, Q. Zhang, S. Pan, and L. Meng, "Generating Handwritten Chinese Characters Using CycleGAN," *Proc. - 2018 IEEE Winter Conf. Appl. Comput. Vision, WACV 2018*, vol. 2018-January, pp. 199–207, May 2018, doi: 10.1109/WACV.2018.00028.

[4] A. K. Basant Babu Bhandari, Aakash Dhakal, Laxman Maharjan, "Nepali Handwritten Letter Generation using GAN," *J. Sci. Eng.*, vol. 9, no. 49–55, 2021, doi: 10.3126/jsce.v9i9.46308.

[5] M. Sharif, A. Ul-Hasan, and F. Shafait, "Urdu Handwritten Ligature Generation Using Generative Adversarial Networks (GANs)," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 13639 LNCS, pp. 421–435, 2022, doi: 10.1007/978-3-031-21648-0_29.

[6] R. Rombach, A. Blattmann, D. Lorenz, P. Esser, and B. Ommer, "High-Resolution Image Synthesis with Latent Diffusion Models," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2022-June, pp. 10674–10685, Dec. 2021, doi: 10.1109/CVPR52688.2022.01042.

[7] P. A. Jonathan Ho, Ajay Jain, "Denoising Diffusion Probabilistic Models," *arXiv:2006.11239*, 2020, doi: https://doi.org/10.48550/arXiv.2006.11239.

[8] I. Ramesh, A., Chen, M., & Sutskever, "DALL·E 2:A New Generative Model for Creating Images from Text Descriptions," *OpenAI*, 2022.

[9] M. G. A. Malik, C. Boitet, and P. Bhattacharyya, "Analysis of Noori Nasta'leeq for

major Pakistani languages," *Work. Spok. Lang. Technol. Under-resourced Lang.*, 2010.
[10] A. A. Kaifi Azmi, "Urdu language," *Britannica*, 2025, [Online]. Available: https://www.britannica.com/topic/Urdu-language
[11] S. A. Husain, "A multi-tier holistic approach for Urdu Nastaliq recognition," *Int. Multi Top. Conf. 2002. Abstr. INMIC 2002., Karachi*, pp. 84–84, Aug. 2005, doi: 10.1109/INMIC.2002.1310191.
[12] N. N. Kashfi, S.M. and A. Khair, "Revolution in Urdu Composing," *Elite*, 2008, [Online]. Available: http://elite.com.pk/epl/wp-content/uploads/2011/11/N.-Nastaliq-Book-Pages.pdf
[13] S. Panwar, M. Ahamed, and N. Nain, "Ligature segmentation approach for Urdu handwritten text documents," *ACM Int. Conf. Proceeding Ser.*, vol. 11-16-November-2014, Nov. 2014, doi: 10.1145/2677855.2677856.
[14] F. Shafait, Adnan-ul-Hasan, D. Keysers, and T. M. Breuel, "Layout analysis of urdu document images," *10th IEEE Int. Multitopic Conf. 2006, INMIC*, pp. 293–298, 2006, doi: 10.1109/INMIC.2006.358180.
[15] and I. M. David Bellos, E. F. Harding, Sophie Wood, "THE UNIVERSAL HISTORY OF NUMBERS," *Harvill London*, 2008, [Online]. Available: https://ia600205.us.archive.org/28/items/TheUniversalHistoryOfNumbers/212027005-The-Universal-History-of-Numbers_text.pdf
[16] D. Anderson, "7. Digital Vitality for Linguistic Diversity," *Glob. Lang. Justice*, pp. 220–243, Nov. 2023, doi: 10.7312/LIU-21038-010/HTML.
[17] S. O. Mehdi Mirza, "Conditional Generative Adversarial Nets," *arXiv:1411.1784*, 2014, doi: https://doi.org/10.48550/arXiv.1411.1784.
[18] X. G. Ji Gan, Weiqiang Wang, Jiaxu Leng, "HiGAN+: Handwriting Imitation GAN with Disentangled Representations," *ACM Trans. Graph.*, vol. 42, no. 1, pp. 1–17, 2022, doi: https://doi.org/10.1145/355007.
[19] H. Zhang *et al.*, "StackGAN++: Realistic Image Synthesis with Stacked Generative Adversarial Networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 41, no. 8, pp. 1947–1962, Aug. 2019, doi: 10.1109/TPAMI.2018.2856256.
[20] Y. W. Jiahui Yu, Xin Li, Jing Yu Koh, Han Zhang, Ruoming Pang, James Qin, Alexander Ku, Yuanzhong Xu, Jason Baldridge, "Vector-quantized Image Modeling with Improved VQGAN," *arXiv:2110.04627*, 2022, doi: https://doi.org/10.48550/arXiv.2110.04627.
[21] D. T. Tingting Qiao, Jing Zhang, Duanqing Xu, "MirrorGAN: Learning Text-to-image Generation by Redescription," *arXiv:1903.05854*, 2019, doi: https://doi.org/10.48550/arXiv.1903.05854 Focus to learn more.
[22] R. Z. Yufan Zhou, "TiGAN: Text-Based Interactive Image Generation and Manipulation," *AAAI Conf. Artif. Intell.*, 2022, doi: 10.1609/aaai.v36i3.20270.
[23] T. Xu *et al.*, "AttnGAN: Fine-Grained Text to Image Generation with Attentional Generative Adversarial Networks," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 1316–1324, Dec. 2018, doi: 10.1109/CVPR.2018.00143.
[24] Xi, Y. et al, "JointFontGAN: Joint Geometry-Content GAN for Font Generation via Few-Shot Learning," *MM '20 Proc. 28th ACM Int. Conf. Multimed.*, pp. 4309–43, 2020, doi: https://doi.org/10.1145/3394171.3413705.
[25] B. W. Weihao Xia, Yujiu Yang, Jing-Hao Xue, "TediGAN: Text-Guided Diverse Face Image Generation and Manipulation," *arXiv:2012.03308*, 2021, doi: https://doi.org/10.48550/arXiv.2012.03308.
[26] I. J. Goodfellow *et al.*, "Generative Adversarial Nets," *Adv. Neural Inf. Process. Syst.*, vol. 27, 2014, Accessed: Oct. 02, 2023. [Online]. Available: http://www.github.com/goodfeli/adversarial

[27] E. Alonso, B. Moysset, and R. Messina, "Adversarial generation of handwritten text images conditioned on sequences," *Proc. Int. Conf. Doc. Anal. Recognition, ICDAR*, pp. 481–486, Sep. 2019, doi: 10.1109/ICDAR.2019.00083.

[28] Y. C. and J. M. Q. Wu, "DCGAN-Based Data Augmentation for Tomato Leaf Disease Identification," *IEEE Access*, vol. 8, pp. 98716–98728, 2020, doi: 10.1109/ACCESS.2020.2997001.

[29] S. Y. Arafat and M. J. Iqbal, "Two Stream Deep Neural Network for Sequence-Based Urdu Ligature Recognition," *IEEE Access*, vol. 7, pp. 159090–159099, 2019, doi: 10.1109/ACCESS.2019.2950537.

[30] S. Y. Arafat and M. J. Iqbal, "Urdu-Text Detection and Recognition in Natural Scene Images Using Deep Learning," *IEEE Access*, vol. 8, pp. 96787–96803, 2020, doi: 10.1109/ACCESS.2020.2994214.

[31] M. Guan, H. Ding, K. Chen, and Q. Huo, "Improving Handwritten OCR with Augmented Text Line Images Synthesized from Online Handwriting Samples by Style-Conditioned GAN," *Proc. Int. Conf. Front. Handwrit. Recognition, ICFHR*, vol. 2020-September, pp. 151–156, Sep. 2020, doi: 10.1109/ICFHR2020.2020.00037.

[32] D. Yorioka, H. Kang, and K. Iwamura, "Data Augmentation for Deep Learning Using Generative Adversarial Networks," *2020 IEEE 9th Glob. Conf. Consum. Electron. GCCE 2020*, pp. 516–518, Oct. 2020, doi: 10.1109/GCCE50665.2020.9291963.

[33] A. Horé and D. Ziou, "Image quality metrics: PSNR vs. SSIM," *Proc. - Int. Conf. Pattern Recognit.*, pp. 2366–2369, 2010, doi: 10.1109/ICPR.2010.579.

[34] H. R. S. and E. P. S. Zhou Wang, A. C. Bovik, "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, 2004, doi: 10.1109/TIP.2003.819861.

[35] Y. D. Yu Yu, Weibin Zhang, "Frechet Inception Distance (FID) for Evaluating GANs," China University of Mining Technology Beijing Graduate School: Beijing, China. Accessed: Mar. 13, 2025. [Online]. Available: https://www.researchgate.net/publication/354269184_Frechet_Inception_Distance_FID_for_Evaluating_GANs

[36] J. Ho, A. Jain, and P. Abbeel, "Denoising Diffusion Probabilistic Models," *Adv. Neural Inf. Process. Syst.*, vol. 2020-December, Jun. 2020, Accessed: Oct. 01, 2023. [Online]. Available: https://arxiv.org/abs/2006.11239v2

[37] M. W. Diederik P Kingma, "Auto-Encoding Variational Bayes," *arXiv:1312.6114*, 2013, doi: https://doi.org/10.48550/arXiv.1312.6114.

Catalytic Performance of Electro-Oxidative Natural Manganese Sand for Ammonium Nitrogen Removal

Asim Khan[1], Syed Faizan Abbas[2], Asad Ali[2,3], Yahya Naveed[2], Abdul Sami[2], Muhammad Saleem Khan[4], Khurram Shahzad Ayub[2*]

[1]National Engineering Laboratory for Industrial Wastewater Treatment, East China University of Science and Technology, Shanghai, China

[2]Department of Chemical Engineering and Technology, University of Gujrat, Pakistan

[3]School of Engineering, Edith Cowan University, Perth, Australia

[4]Department of Chemical Engineering NFC Institute of Engineering and Technology, Multan, Pakistan

***Correspondence:** engr.khurram@uog.edu.pk

Citation | Khan. A, Abbas. S. F, Ali. A, Naveed. Y, Sami. A, Khan. M. S, Ayub. K. S, "Catalytic Performance of Electro-Oxidative Natural Manganese Sand for Ammonium Nitrogen Removal", IJIST, Special Issue pp 188-196, March 2025

Received | Feb 21, 2025 **Revised |** March 05, 2025 **Accepted |** March 11, 2025 **Published |** March 14, 2025.

The environmental risks associated with ammonium nitrogen (NH_4^+-N) pollution have led to a growing focus on prevention. Electrochemical advanced oxidation is an effective and eco-friendly method that only requires electricity and electrolytes to remove NH_4^+-N from wastewater. This study assesses the effectiveness of electro-oxidative natural manganese sand (NMS) in removing ammonium nitrogen under different conditions. Due to NMS's high redox potential, it significantly enhanced the electrochemical oxidation process, increasing NH_4^+-N removal and generating reactive chlorine species (ClO^-/HClO) when NaCl was added. The experiment was also conducted without a catalyst, quartz sand, and natural manganese sand, but NMS removed 86.4% of NH_4^+-N, outperforming the other treatments. The removal efficiency was tested at five different pH levels (3, 5, 7, 9, and 11), with NMS showing the highest efficiency of 95.2% at pH 9. At a current density of 15.5 mA/cm², the removal rate reached 94.9%, and with a NaCl concentration of 9 g/L, the removal efficiency peaked at 96.2%, driven by increased production of reactive chlorine species (ClO^-). These results demonstrate the electro-oxidative NMS system as a highly efficient, scalable, and eco-friendly solution for ammonium nitrogen removal in wastewater treatment.

Keywords: Electrochemical Oxidation; Manganese Sand Catalyst; Ammonium Nitrogen Removal; Wastewater Treatment; Sustainability

Introduction:

Ammonium nitrogen (NH_4^+-N) is highly toxic and poses a significant threat to global water quality. It originates from various sources, including the excessive use of nitrogen fertilizers like urea, discharges from human and animal waste, agro-industrial wastewater, landfill leachate, and untreated domestic wastewater [1]. Water contamination from ammonium nitrogen has severe effects on the ecosystem [2]. High concentrations of ammonia and ammonium in water can cause unpleasant odors and tastes. Excessive intake of ammonium nitrogen can disrupt the body's acid-base balance, making the blood too alkaline and leading to serious health issues. These health concerns highlight the dangers of elevated ammonia levels in water. As a result, the WHO and EU have set a maximum permissible limit of 2 mg/L for ammonia nitrogen concentrations in water [3].

Conventional methods for removing ammonium nitrogen include air stripping [4], ion exchange and adsorption, capacitive deionization, membrane technology [5], the activated sludge process [6], and biological methods using ammonia-oxidizing bacteria (AOB) [7]. While these techniques are effective, they are not always cost-efficient and may contribute to environmental contamination. This has created a growing demand for abundant and cost-effective oxidants or catalysts [8].

Advanced oxidation processes (AOPs) have emerged as promising alternatives because they can generate powerful oxidative species to break down stubborn contaminants with minimal secondary pollution [9]. Radical-based oxidation, such as catalytic systems, is also used to treat industrial low-volatile pollutants [10]. Among these methods, the electrochemical oxidation process stands out due to its strong capability to remove ammonium nitrogen. This method offers benefits such as precise control, mild reaction conditions, and reduced environmental impact, allowing for ammonium nitrogen removal through both direct and indirect approaches [11].

Manganese is one of the most abundant elements in nature and is commonly found as manganese oxides (MnOx). Its natural abundance, high surface area, and strong redox ability make it a promising material for wastewater treatment [12]. The effectiveness of manganese in wastewater treatment can be enhanced by using chloride ions, as active chlorine is involved in the oxidation process [13]. In this study, manganese sand was electrochemically activated for ammonium nitrogen (NH_4^+-N) removal. Various factors were investigated, including solution pH, NaCl concentration, and current density. The goal of this study is to fill the research gap and assess the potential of electrochemically activated manganese sand in developing more effective and sustainable methods for removing NH_4^+-N from wastewater, as well as integrating NMS with electrochemical processes for industrial scalability.

Objectives of the Study:

The primary objective of this study is to evaluate the catalytic performance of electro-oxidative natural manganese sand (NMS) for ammonium nitrogen (NH_4^+-N) removal in wastewater treatment. The specific objectives include:

- To investigate the efficiency of natural manganese sand (NMS) as a catalyst in electro-oxidation processes for NH_4^+-N removal.
- To optimize key process parameters, including pH, NaCl concentration, and current density, to achieve maximum NH_4^+-N removal efficiency.
- To compare the electro-oxidation performance of NMS with quartz sand and electro-oxidation alone, determining the catalytic contribution of manganese sand.
- To explore the reaction mechanisms involved in NH_4^+-N removal, identifying key oxidation pathways and intermediate species.

- To assess the feasibility of NMS for practical wastewater treatment applications, including potential scalability, reusability, and cost-effectiveness.

Novelty Statement:

This study introduces natural manganese sand (NMS) as a novel catalyst for electro-oxidative ammonium nitrogen removal, offering an efficient and sustainable approach to wastewater treatment. Unlike conventional NH_4^+-N removal techniques such as ion exchange, biological nitrification, and membrane filtration, which often suffer from high operational costs and secondary pollution, this research demonstrates that NMS significantly enhances electro-oxidation efficiency without requiring complex modifications.

Material and Methods:

The experiment was carried out in a batch electrolytic cell using a 250 mL beaker fitted with electrodes. A 100 mg/L ammonium nitrogen solution was prepared in distilled water, with 3 g/L of sodium chloride added. To enhance the electrochemical oxidation process, 7.1 mg/L of sodium sulfate was included as a supportive electrolyte to increase ionic strength.

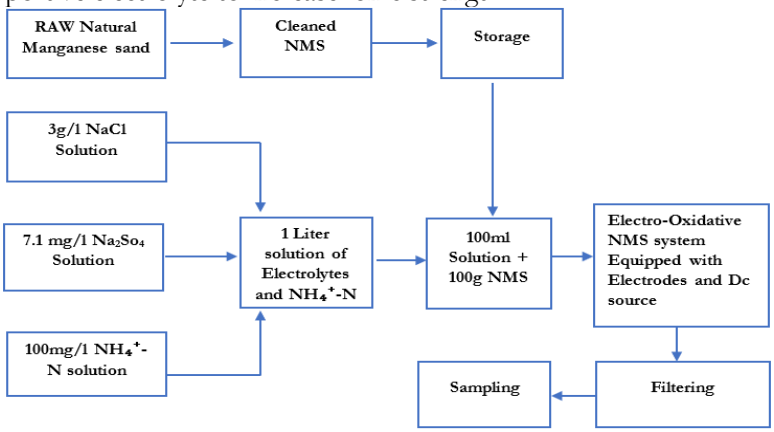


Figure 1. Flow Diagram of Methodology

A catalyst dose of 100 g of manganese sand was added to facilitate the oxidation of ammonium nitrogen. The electrochemical oxidation process was conducted for 120 minutes, with current densities ranging from 5.5 to 15.5 mA/cm². To evaluate the impact of various parameters on ammonium nitrogen removal efficiency, samples were collected at 20-minute intervals throughout the experiment. The pH of the system was maintained between 3 and 11 by adjusting with dilute HCl. After each sampling, the collected solutions were filtered to remove any solid manganese particles before analysis. Ammonium concentrations in the filtered samples were measured using a spectrophotometer coupled with an ion-selective electrode.

Additionally, the effect of varying sodium chloride (NaCl) concentrations was examined, with NaCl concentrations ranging from 1 to 9 g/L. This approach allowed for the evaluation of key factors influencing the efficiency of ammonium nitrogen removal.

Determination of Ammonia Nitrogen:

Ammonium nitrogen was determined using Nessler's spectrophotometry method [14], as shown in Figure 1. A standard curve was created using a known concentration of ammonium nitrogen, in the form of free ammonia or ammonium ions, which reacts with Nessler's reagent to form a brown complex. The absorbance of this complex was directly proportional to the ammonium nitrogen concentration. For analysis, the samples were diluted with 50 mL of deionized water, followed by the addition of 1.0 mL of potassium sodium tartrate solution and 1.0 mL of Nessler's reagent. After the reaction, the absorbance of the samples was measured using the standard curve to determine the ammonium concentration.

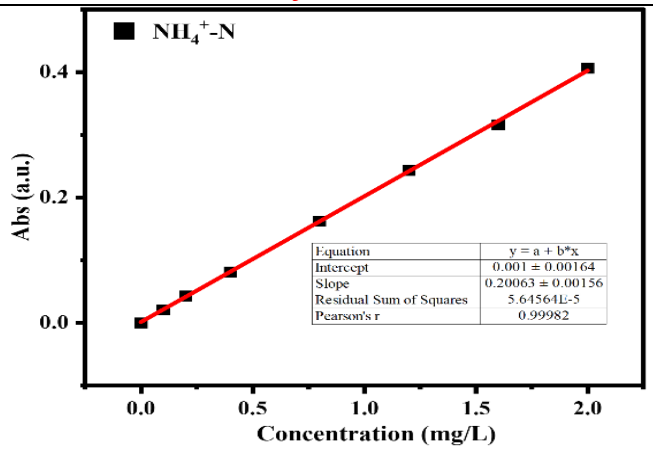


Figure 2. Standard Curve of Ammonium Nitrogen

Result and Discussion:

The catalytic performance of the electro-oxidative natural manganese sand (NMS) system for ammonium nitrogen oxidation was thoroughly investigated by comparing three systems: electro-oxidation alone, electro-oxidation with quartz sand, and advanced electro-oxidation with manganese sand. The results showed that electro-oxidation alone achieved a removal efficiency of 42.7%, and electro-oxidation with quartz sand reached 37.6%. However, with the inclusion of manganese sand, the removal efficiency significantly improved to 86.4% after 120 minutes of electrolysis, demonstrating the superior performance of the electro-oxidative manganese sand system in treating inorganic pollutants.

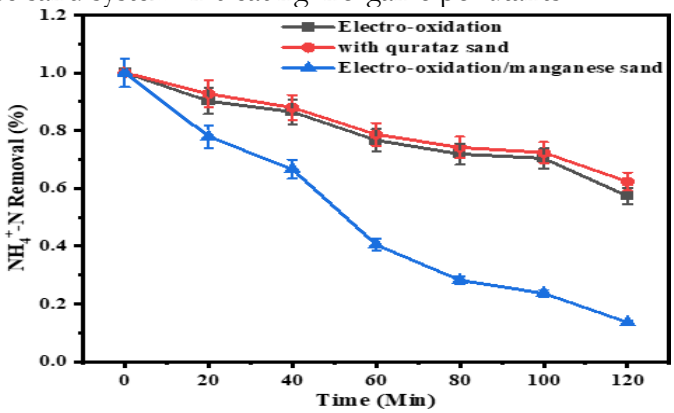


Figure 3. Performance of electro-0xidation manganese sand on removal of Ammonium

Influence of Operating Conditions on Electro-Oxidative NMS Performance:

Effect of Ph on Removal Rate:

The influence of pH on ammonium nitrogen removal was evaluated, and the system showed adaptability to a wide pH range. At higher pH values, the system facilitated the formation of easily removable ammonia (NH_3), while pH levels below 8 made ammonium nitrogen removal more challenging. The system achieved ammonium nitrogen removal efficiencies ranging from 90.45% to 94.4% across a pH range of 3 to 11 after 120 minutes. This enhanced removal efficiency was attributed to the system's ability to effectively oxidize ammonium nitrogen even at lower pH levels, ensuring the stability of the electrochemical manganese sand system.

Effect of Current Density:

The effect of current density on ammonium nitrogen removal was also examined. As the current density increased from 5.5 to 15.5 mA/cm², the removal efficiency increased accordingly. After 60 minutes, removal rates ranged from 55% to 77.3%, while after 120 minutes, the removal efficiency reached up to 94.9% at the highest current density of 15.5

mA/cm². The enhancement in performance was attributed to the increased production of hypochlorite (ClO^-), which facilitated the oxidation of ammonium nitrogen.

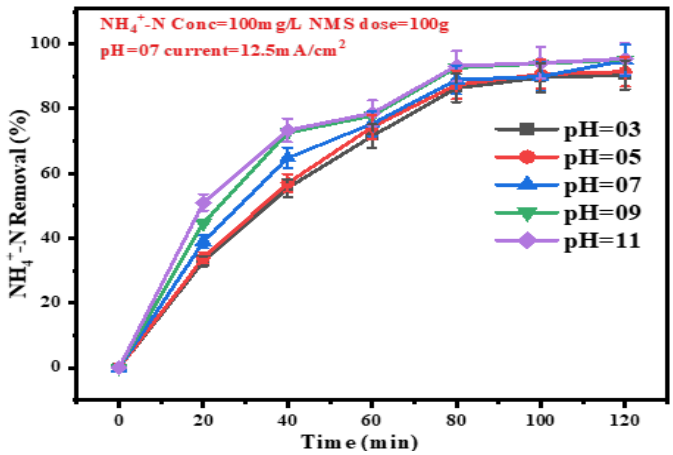


Figure 4. pH vs. NH_4^+-N Removal Efficiency (%)

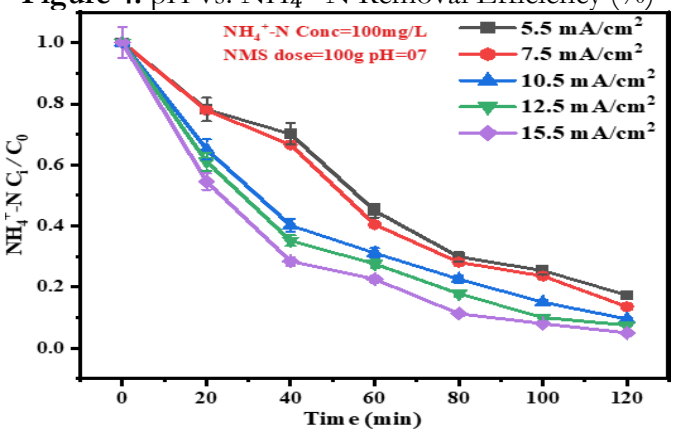


Figure 5. Current density vs. NH_4^+-N Removal Efficiency (%)

Effect of NaCl Concentration:

The role of NaCl concentration in ammonium nitrogen removal was evaluated by increasing NaCl concentration from 1 to 9 g/L. Removal efficiency progressively improved, with the highest efficiency of 96.2% achieved after 120 minutes. NaCl played a crucial role in generating hypochlorous acid (HClO) and hypochlorite (ClO^-) at the anode, which enhanced the oxidation of ammonium nitrogen. This demonstrated the potential of the NMS system as an effective technology for ammonium nitrogen removal, offering high removal rates, adaptability to varying pH conditions, and minimal secondary pollution.

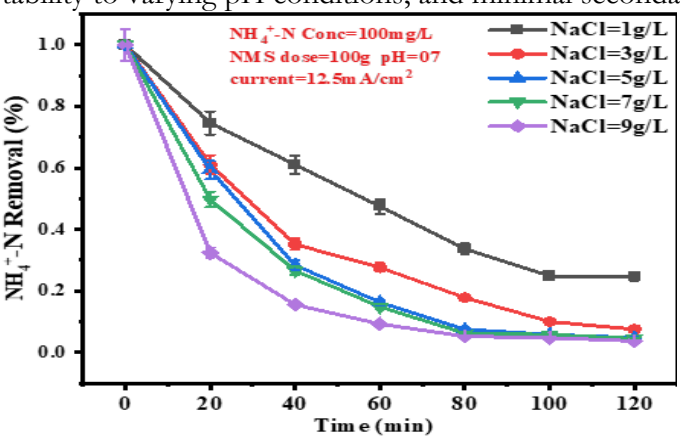


Figure 6. NaCl amount vs. NH_4^+-N Removal Efficiency (%)

Proposed Mechanism:

The most probable reaction mechanism for eliminating ammonium nitrogen from wastewater is given below.

$$2Cl^- \rightarrow Cl_2 + 2e^- \quad (1)$$

$$Cl_2 + H_2O \rightarrow HClO + Cl^- + H^+ \quad (2)$$

$$HClO \rightarrow ClO^- + H^+ \quad (3)$$

$$2H_2O + 2e^- \rightarrow H_2 + 2OH^- \quad (4)$$

$$Cl_2 + H_2O \rightarrow HClO + H^+ + Cl^- \quad (5)$$

$$NH_4^+ + OH^- \rightarrow NH_3 + H_2O \quad (6)$$

$$2NH_3 + 3HClO \rightarrow N_2 + 3HCl + 3H_2 \quad (7)$$

$$2NH_3 + 2ClO^- \rightarrow N_2 + 2HCl + 2H_2O \quad (8)$$

$$NH_4^+ + HClO \rightarrow NH_2Cl + H_2 + H^+ \quad (9)$$

$$NH_2Cl + HClO \rightarrow NHCl_2 + H_2 \quad (10)$$

$$NHCl_2 + HClO \rightarrow NCl_3 + H_2O \quad (11)$$

$$NCl_3 + NHCl_2 + 5OH^- \rightarrow N_2 + 3Cl^- + 3H_2O + 2ClO^- \quad (12)$$

$$2NH_4^+ + 3HClO \rightarrow N_2 + 3H_2O + 5H^+ + 3Cl^- \quad (13)$$

$$2NH_4^+ + 3ClO^- \rightarrow N_2 + 3H_2O + 2H^+ + 3Cl^- \quad (14)$$

This mechanism is proposed to eliminate ammonium nitrogen from wastewater. Electrochemically generated chlorine reacts with water to produce hypochlorous acid (HClO) and hypochlorite (ClO^-) [15]. Ammonium ions (NH_4^+) convert to ammonia (NH_3) in the presence of hydroxyl ions [16]. Intermediate chloramines, such as monochloramine, dichloramine, and trichloramine, are then formed when ammonia combines with HClO and ClO^-. These chloramines further decompose into byproducts like nitrogen gas, hydrogen ions, and chloride ions [9].

Manganese sand acts as a catalyst, providing active sites [17] for ammonia (NH_3), ammonium ions (NH_4^+), hypochlorite (ClO^-), and hypochlorous acid (HClO). It also facilitates electron transfer between ammonia/chloramines, hypochlorous acid, and hypochlorite.

Discussion:

The results of this study demonstrate the significant enhancement in ammonium nitrogen removal achieved through the use of electro-oxidative manganese sand (NMS) systems. In comparison to electro-oxidation alone and electro-oxidation with quartz sand, the inclusion of manganese sand greatly improved the removal efficiency, achieving a remarkable 86.4% after 120 minutes of electrolysis (Figure 3). This indicates that manganese sand plays a crucial role in facilitating the oxidation process, likely by acting as a catalyst that enhances electron transfer and supports the formation of key reactive species such as hypochlorous acid (HClO) and hypochlorite (ClO^-), which are involved in ammonium nitrogen oxidation [15][17].

The impact of operating conditions on the performance of the electro-oxidative NMS system was thoroughly investigated. A key finding was that the system exhibited high adaptability to varying pH levels, with effective ammonium nitrogen removal across a broad pH range (3 to 11), reaching removal efficiencies of over 90% in most conditions (Figure 4). The formation of ammonia (NH_3) at higher pH levels and ammonium ions (NH_4^+) at lower pH was consistent with previous findings [18], indicating that both acidic and alkaline environments are favorable for the oxidation of ammonium nitrogen. The higher removal efficiencies observed at pH values around 8 and above suggest that ClO^- ions, which are more abundant under alkaline conditions, accelerate the oxidation process, further supporting the role of manganese sand in promoting these reactions.

The effect of current density on ammonium nitrogen removal was also explored, revealing a positive correlation between increased current density and enhanced removal efficiency (Figure 5). Higher current densities generate a greater driving force for electrochemical reactions, leading to higher production of ClO^-, which facilitates the oxidation of ammonium ions. These findings align with previous studies that have shown an increased rate of pollutant removal with higher current densities in electrochemical oxidation systems [19].

Additionally, the study examined the role of NaCl concentration in the electro-oxidative NMS system. Increasing NaCl concentration from 1 to 9 g/L led to a progressive increase in ammonium nitrogen removal efficiency, reaching 96.2% at 9 g/L (Figure 6). This can be attributed to the increased production of hypochlorous acid and hypochlorite at the anode, which are key oxidizing agents that promote the oxidation of ammonium ions [20]. The synergy between NaCl concentration and manganese sand highlights the importance of optimizing these parameters to enhance the performance of electrochemical treatment systems for wastewater treatment.

Overall, the electro-oxidative NMS system demonstrates promising potential for efficient ammonium nitrogen removal from wastewater. The system's high removal efficiency, adaptability to varying pH conditions, and the synergistic effect of NaCl and manganese sand suggest that it could be an effective technology for treating ammonium nitrogen contamination in industrial effluents. Future studies should further explore the long-term stability and scalability of this system, as well as its potential application in larger-scale wastewater treatment facilities.

Conclusion:

The removal efficiency of ammonium nitrogen (NH_4^+-N) was improved by adjusting various parameters such as pH, current density, and NaCl concentration. Manganese sand serves as an effective catalyst due to its high redox potential and large surface area, significantly enhancing NH_4^+-N removal. Varying NaCl concentrations (1–9 g/L) notably improved the removal efficiency, reaching 96.2%, due to the formation of reactive chlorine species (ClO^- and HClO). Increasing the current density (from 5.5 to 15.5 mA/cm^2) further enhanced performance, achieving 94.9% removal at the highest current density by generating more active chlorine species (ClO^-). The system is adaptable to a wide pH range (3–11), with the highest removal efficiency (95.2%) observed at pH 9. At this pH, NH_3 and ClO^- are more prevalent, speeding up the oxidation process and shifting the equilibrium toward NH_3 formation. However, in strongly acidic conditions, activated ions from manganese sand may leach, reducing removal efficiency and possibly deactivating the catalyst due to the reaction of byproducts [21]. This electro-oxidative NMS method is characterized by its controllability, scalability, and high removal efficiency, making it ideal for treating toxic waste in metropolitan and industrial settings. Its adaptability to various operational conditions ensures its effectiveness across multiple applications.

References:

[1] T. Zhou, M. Wang, H. Zeng, R. Min, J. Wang, and G. Zhang, "Application of physicochemical techniques to the removal of ammonia nitrogen from water: a systematic review," *Environ. Geochemistry Heal. 2024 469*, vol. 46, no. 9, pp. 1–24, Jul. 2024, doi: 10.1007/S10653-024-02129-6.

[2] N. Sonadia, Z. Iqbal, W. Miran, A. Ul-Hamid, K. S. Ayub, and F. Azad, "Enhanced Electrocatalytic Performance of Erbium-Incorporated Nickel-Based Metal-Organic Frameworks for Water Splitting," *Energy and Fuels*, vol. 38, no. 6, pp. 5397–5406, Mar. 2024, doi: 10.1021/ACS.ENERGYFUELS.3C04609/SUPPL_FILE/EF3C04609_SI_001.PDF.

[3] V. R. Viktor Yushchenko, Elena Velyugo, "Influence of ammonium nitrogen on the treatment efficiency of underground water at iron removal stations," *Groundw. Sustain. Dev.*, vol. 22, p. 100943, 2023, doi: https://doi.org/10.1016/j.gsd.2023.100943.

[4] S. F. S. Lennevey Kinidi, Ivy Ai Wei Tan, Noraziah Binti Abdul Wahab, Khairul Fikri Bin Tamrin, Cirilo Nolasco Hipolito, "Recent Development in Ammonia Stripping Process for Industrial Wastewater Treatment," *Int. J. Chem. Eng.*, 2018, doi: https://doi.org/10.1155/2018/3181087.

[5] R. S. Vinod K Gupta, H. Sadegh, Mehdi Yari, "Removal of ammonium ions from wastewater: A short review in development of efficient methods," Global Journal of Environmental Science and Management. Accessed: Mar. 13, 2025. [Online]. Available: https://www.researchgate.net/publication/270341154_Removal_of_ammonium_ions_from_wastewater_A_short_review_in_development_of_efficient_methods

[6] S. Xiang *et al.*, "New progress of ammonia recovery during ammonia nitrogen removal from various wastewaters," *World J. Microbiol. Biotechnol. 2020 3610*, vol. 36, no. 10, pp. 1–20, Aug. 2020, doi: 10.1007/S11274-020-02921-3.

[7] S. W. Xiaolong Yang, Lihua Liu, "A strategy of high-efficient nitrogen removal by an ammonia-oxidizing bacterium consortium," *Bioresour. Technol.*, vol. 275, pp. 216–224, 2019, doi: https://doi.org/10.1016/j.biortech.2018.12.057.

[8] Y. Cheng, T. Huang, Y. Sun, and X. Shi, "Catalytic oxidation removal of ammonium from groundwater by manganese oxides filter: Performance and mechanisms," *Chem. Eng. J.*, vol. 322, pp. 82–89, 2017, doi: https://doi.org/10.1016/j.cej.2017.04.010.

[9] J. W. Xinyu Liu, "Selective oxidation of ammonium to nitrogen gas by advanced oxidation processes: Reactive species and oxidation mechanisms," *J. Environ. Chem. Eng.*, vol. 11, no. 3, p. 110263, 2023, doi: https://doi.org/10.1016/j.jece.2023.110263.

[10] Z. Abbas *et al.*, "Catalytic nonthermal plasma using efficient cobalt oxide catalyst for complete mineralization of toluene," *Res. Chem. Intermed.*, vol. 47, no. 6, pp. 2407–2420, Jun. 2021, doi: 10.1007/S11164-021-04406-W/METRICS.

[11] L. Z. Fengjiao Quan, Guangming Zhan, Bing Zhou, Cancan Ling, Xiaobing Wang, Wenjuan Shen, Jianfen Li, Falong Jia, "Electrochemical removal of ammonium nitrogen in high efficiency and N2 selectivity using non-noble single-atomic iron catalyst," *J. Environ. Sci.*, vol. 125, pp. 544–552, 2023, doi: https://doi.org/10.1016/j.jes.2022.03.004.

[12] W. W. Lixia Jia, Qi Zhou, Yuanwei Li, "Application of manganese oxides in wastewater treatment: Biogeochemical Mn cycling driven by bacteria," *Chemosphere*, vol. 336, p. 139219, 2023, doi: https://doi.org/10.1016/j.chemosphere.2023.139219.

[13] O. Scialdone, S. Randazzo, A. Galia, and G. Silvestri, "Electrochemical oxidation of organics in water: Role of operative parameters in the absence and in the presence of NaCl," *Water Res.*, vol. 43, no. 8, pp. 2260–2272, 2009, doi: https://doi.org/10.1016/j.watres.2009.02.014.

[14] H. L. Kunning Lin, Yong Zhu, Yuanbiao Zhang, "Determination of ammonia nitrogen in natural waters: Recent advances and applications," *Trends Environ. Anal. Chem.*, vol. 24, p. e00073, 2019, doi: https://doi.org/10.1016/j.teac.2019.e00073.

[15] C. Kim, T. T. Thao, J.-H. Kim, and I. Hwang, "Effects of the formation of reactive chlorine species on oxidation process using persulfate and nano zero-valent iron," *Chemosphere*, vol. 150, p. 126266, 2020, doi: https://doi.org/10.1016/j.chemosphere.2020.126266.

[16] Kwang-Wook Kim, Y.-J. Kim, I.-T. Kim, G.-I. Park, and E.-H. Lee, "The electrolytic decomposition mechanism of ammonia to nitrogen at an IrO2 anode," *Electrochim. Acta*, vol. 50, no. 22, pp. 4356–4364, 2005, doi:

https://doi.org/10.1016/j.electacta.2005.01.046.
[17] X. T. Qian Peng, Yingjie Zhang, Wanling Zhong, Kun Liu, Jiajie Xing, "Facile preparation of manganese sand-based monolithic catalysts with excellent catalytic performance and reusability for activation of peroxymonosulfate: The key role of pre-calcination," *J. Water Process Eng.*, vol. 56, p. 104398, 2023, doi: https://doi.org/10.1016/j.jwpe.2023.104398.
[18] D. . Mehendale, F. V, Clayton, G., Homyer, K., Reynolds, "HOCl vs OCl−: clarification on chlorine-based disinfectants used within clinical settings," *J. Glob. Heal. Reports*, vol. 7, 2023, doi: https://doi.org/10.29392/001c.84488.
[19] S.-E. O. Umesh Ghimire, Min Jang, Sokhee P. Jung, Daeryong Park, Se Jin Park, Hanchao Yu, "Electrochemical Removal of Ammonium Nitrogen and COD of Domestic Wastewater using Platinum Coated Titanium as an Anode Electrode," *Energies*, vol. 12, no. 5, p. 883, 2019, doi: https://doi.org/10.3390/en12050883.
[20] T. D. W. Changyong Zhang, Di He, Jinxing Ma, "Active chlorine mediated ammonia oxidation revisited: Reaction mechanism, kinetic modelling and implications," *Water Res.*, vol. 145, pp. 220–230, 2018, doi: https://doi.org/10.1016/j.watres.2018.08.025.
[21] J. Shu, R. Liu, Z. Liu, H. Chen, and C. Tao, "Leaching of manganese from electrolytic manganese residue by electro-reduction," *Environ. Technol.*, vol. 38, no. 16, pp. 2077–2084, Aug. 2017, doi: 10.1080/09593330.2016.1245789.

Detection of Application-Layer Dos Attacks in IoT Devices Using Feature Selection and Machine Learning Models

Mustabeen Aziz[1], Muhammad Usman Sana[2]*, Tayybah Kiren[3]*, Tahmina Ehsan[1], Alvena Ehsan[1], Fateha Minahil[1]

[1]Department of Information Technology, University of Gujrat, Pakistan

[2]Department of Software Engineering, University of Gujrat, Pakistan

[3]Department of Computer Science (RCET), University of Engineering and Technology Lahore, Pakistan

*** Correspondence:** m.usman@uog.edu.pk, tayybah@uet.edu.pk

Citation | Aziz. M, Sana. M. U, Kiren. T, Ehsan. T, Ehsan. A, Minahil. F, "Detection of Application-Layer Dos Attacks in IoT Devices Using Feature Selection and Machine Learning Models", IJIST, Special Issue pp 197-207, March 2025

Received | Feb 21, 2025 **Revised |** March 05, 2025 **Accepted |** March 11, 2025 **Published |** March 14, 2025.

With technological advancements, innovations like the Internet of Things (IoT) have become widespread, connecting more devices to the Internet. However, as the number of connected devices increases, cyber-attacks—especially Distributed Denial of Service (DDoS) attacks—are also becoming more frequent. This research explores these cyber threats, focusing on DDoS attacks, and proposes strategies to protect IoT devices. It specifically aims to detect DDoS attacks in IoT devices using feature selection methods and machine learning algorithms. The study targets attack detection at the application layer of IoT devices by analyzing a relevant dataset. By applying feature selection techniques and machine learning models, we strive to enhance the accuracy and efficiency of DDoS detection, ultimately improving IoT security.

Keywords: Distributed Denial of Service (DDoS), Cybersecurity, Internet of Things, Feature Selection.

Introduction:

In recent years, millions of IoT devices have been connected for communication, relying on the Internet to transmit data between devices. These include sensors, smart devices, actuators, and RFID devices. However, IoT devices face cybersecurity challenges, particularly cyber-attacks. As the number of connected devices grows, different types of attacks emerge, including Distributed Denial of Service (DDoS) attacks.

ADDoS attack does not steal information but floods the server with excessive traffic, slowing it down. This study focuses on detecting exploitation- and reflection-based DDoS attacks using reduced features [1]. The main goal of a DDoS attack is to overwhelm a targeted website or server with high traffic from multiple sources, rendering it inaccessible. These attacks are particularly harmful because they originate from numerous locations, making them difficult to mitigate [2].

Application-layer DDoS attacks, such as HTTP request-based attacks, are especially dangerous because they require relatively few attacking connections to disrupt a website. Their traffic closely resembles normal traffic, making detection challenging [3]. In APDDoS attacks, attackers send numerous packets to the target server, causing congestion and slowing it down. To launch this attack, each participating system first establishes a TCP connection with the victim's server, requiring a valid IP address [4].

Research Contributions:

- This study focuses on detecting DDoS attacks at the application layer of IoT devices, as shown in Figure 1.
- It employs the Extra Tree feature selection method to identify relevant features for detecting DDoS attacks.
- Hyperparameter tuning and k-fold cross-validation are applied to enhance model robustness and prevent overfitting.
- Machine learning models and feature selection techniques are used to improve IoT security and enhance DDoS attack detection.

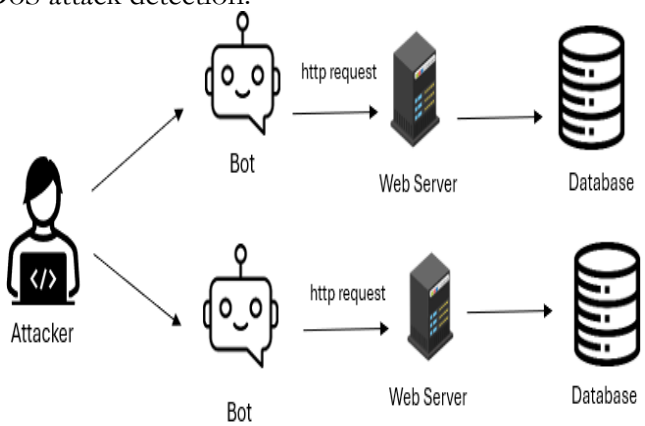


Figure. 1. Application-layer DOS attack

DDoS attacks are categorized into two types:

- **Network Layer DDoS Attacks** – These aim to generate high volumes of traffic to overwhelm the target.
- **Application Layer DDoS Attacks** – These mimic legitimate behavior while using low bandwidth, making them harder to detect and mitigate [5].

DDoS attacks can harm networks in various ways. Most application-layer DDoS (APDDoS) attacks exploit protocols such as HTTPS, HTTP page flood, DNS query flood, and HTTP bandwidth utilization [6].

- **Exploitation attacks** include SYN flood, UDP lag, and UDP flood.

- **Reflection-based attacks** involve SNMP, MSSQL, LDAP, SSDP, DNS, and NETBIOS.

The objective of this paper is to detect DDoS attacks at the application layer of IoT devices. The research contributions include:

- Detecting DDoS attacks using a feature selection method.
- Applying machine learning algorithms for detection.
- Comparing different ML models to improve DDoS attack detection at the application layer.

Denial of Service Attack:

A DDoS attack targets IoT devices by overwhelming them with excessive traffic. Hackers exploit these devices to access resources and disrupt normal operations [7]. In this attack, multiple devices send numerous requests to a server, which treats them as legitimate and responds to each one. As the server receives more requests than it can handle, it slows down. The goal of a DDoS attack is not to steal information but to degrade server performance.

There are three main types of DDoS attacks [8]:

- Protocol-Based Attacks – Exploit network protocols to overwhelm resources.
- Application Layer Attacks – Target specific applications, mimicking legitimate traffic.
- Volume-Based Attacks – Flood the network with massive amounts of traffic.

When multiple computers send requests to a single server beyond its capacity, the server cannot differentiate IP addresses, leading to incorrect responses and further congestion.

Research Focus:

- This study focuses on detecting DoS attacks at the application layer in the IoT environment.
- The Extra Tree feature selection method is used to identify key features for detecting DDoS attacks.
- The dataset includes relevant features of application-layer attacks for accurate detection.
- Hyperparameter tuning and k-fold cross-validation are applied to enhance model robustness and prevent overfitting.
- Machine learning models and feature selection techniques are used to strengthen IoT security and improve DDoS attack detection.

Novelty of the Study:

This research emphasizes application-layer DDoS attacks, which closely resemble normal traffic and are harder to detect than traditional network-layer attacks. By using an application-layer DDoS dataset, this approach provides a practical solution for the early detection of APDDoS attacks in IoT networks.

Objectives of the Study:

Key Objectives of This Study:

- Detect **application-layer DDoS attacks** in IoT devices using machine learning techniques.
- Improve detection accuracy and efficiency by applying Extra Tree feature selection, which reduces dataset dimensionality while preserving essential features.
- Evaluate and compare the performance of Decision Tree, Naïve Bayes, and Logistic Regression classifiers using key metrics such as accuracy, precision, recall, and F1-score.
- Enhance IoT security by identifying the most effective machine learning model for real-time detection and practical deployment in IoT environments.

Literature Review:

In [9], the attacker targets an IP address and sends requests to a server via the Internet. This type of attack is called a reflection attack, where the response size is larger than the request size. DDoS detection methods fall into three categories:

- **Supervised learning** – Uses labeled data for classification.
- **Unsupervised learning** – Works with unlabeled data to identify patterns.
- **Hybrid learning** – Combines both approaches to distinguish DDoS attacks from normal traffic.

A hybrid machine learning method is used for DDoS detection, working in three phases:

- DBSCAN algorithm clusters benign and DDoS network flows.
- Clusters are partitioned and analyzed using statistical measures.
- The CICIDS dataset is used for training, while the CICDDoS2019 dataset is used for testing.
- DBSCAN is used for unsupervised learning, while classification algorithms are applied for supervised learning.

In [10], DDoS attacks are recognized as a major threat to cloud computing, IoT, and 5G networks. While many researchers have studied DoS attacks, they often use outdated datasets that lack modern threats. This study employs an SDN-based (Software-Defined Network) architecture to detect DDoS attacks at the transport and application layers using deep learning (DL) and machine learning (ML) algorithms.

- **Transport layer attacks**: UDP flood, TCP-SYN flood.
- **Application layer attacks:** High- and low-volume HTTP-based attacks.

The study uses machine learning models such as Random Forest (RF), Support Vector Machine (SVM), and K-Nearest Neighbors (KNN), along with deep learning techniques like Multilayer Perceptron (MLP), Convolutional Neural Network (CNN), and Gated Recurrent Unit (GRU).

- The CICDDoS2017 and CICDDoS2019 datasets are used.
- Achieved 98% accuracy for application-layer attacks and 95% accuracy for transport-layer attacks.

In [11], the study highlights DDoS attacks as a major challenge to internet security. These attacks target the transport, application, and network layers using various protocols such as ICMP, HTTP, TCP, and UDP.

- A DDoS taxonomy is proposed to categorize different attack types.
- The CICDDoS2019 dataset is used to provide a feature set for detecting various DDoS attacks.
- Machine learning techniques used: Random Forest, Logistic Regression, Naïve Bayes, and ID3 algorithms.
- The study focuses on detecting SYN, DNS, MSSQL, UDP-Lag, and LDAP attacks.
- In [12], the study addresses IoT security by applying machine learning (ML) and artificial intelligence (AI) techniques.
- IoT devices rely on sensors and communicate via wired and wireless networks.
- The study applies AI-based intrusion detection and ML-based classification to detect anomalies in IoT systems.
- Neural networks are used to train the system to identify invalid traffic.
- Focuses on differences between IoT systems and traditional systems.
- Uses the KDD Cup 1999 dataset, which contains IoT and cybersecurity data.
- In [13], a review of ensemble learning techniques is conducted, comparing different feature selection methods for DDoS detection.
- Evaluates True Positive Rate (TPR), False Positive Rate (FPR), False Negative Rate (FNR), and accuracy.
- Highlights challenges in existing models, such as high false rates and low detection rates.
- Compares traditional ML models with ensemble learning techniques:
- Stacking-based, Bagging-based, and Boosting-based approaches.

- Traditional ML methods include Naïve Bayes (NB), K-Nearest Neighbors (KNN), and Decision Tree (DT).
- Focuses on intrusion detection in smart grids to distinguish malicious activities from normal activities.

In [14], various feature selection methods are analyzed for detecting DDoS attacks, including:

- Chi-Square, ANOVA, Extra Tree, and Mutual Information methods.
- Machine learning algorithms Random Forest (RF) and Decision Tree (DT) are applied.
- The Extra Tree feature selection method is used to extract the most relevant features for detecting different DDoS attack types.

Table 1. Comparison of Published Results

Ref	Problem	Methodology	Accuracy%	Features	Dataset
[9]	Detecting unprecedented DDoS attacks	Hybrid ML-based method	99%	All Features	CICIDS2017 CICDDOS2019
[10]	SDN-based architecture to Detect DDOS attacks at the Application and Transport Layer	ML and DL algorithms	99%	All features	CICDDOS2019
[11]	Real-time detection of different taxonomies of DDOS	To propose a taxonomy of DDOS for application later		Generate dataset	CICDDOS2019
[12]	To secure the IOT system	AI-based intrusion detection and classification in IOT networks using machine learning.	97.77%	All features	KDD Cup 1999
[13]	Intrusion detection in smart grids	Use ensemble learning and ML techniques for intrusion detection	93.4% 97.4%	All features	CICDDOS2019
[14]	Feature selection for detection of DDOS	Feature selection techniques and ML classifiers	82% 61%	Top Ten Features	CICDDOS2019
[15]	Detect DDOS Attack	Extra Tree-Random Forest model is used for detection of DDOS Attack	99%	99%	CICDDOS2019

Feature Selection and DDoS Detection:

In this study, the top 10 features are selected to detect DDoS attacks. The accuracy of detection is analyzed using three feature selection methods:

- Chi-Square

- Extra Tree
- Mutual Information

The performance of these feature selection techniques is evaluated using two machine learning classifiers:

- Decision Tree (DT)
- Random Forest (RF)

Study in [15]

In [15], the author applies the ET-RF model to the CICDDoS2019 dataset for DDoS attack detection. The study is divided into two scenarios:

1. Performance Evaluation of ML Algorithms
   - Different machine learning models, including K-Nearest Neighbors (KNN), Decision Tree (DT), and Random Forest (RF), are compared.
   - The Random Forest (RF) classifier, combined with the ET-RF feature selection method, achieves the highest accuracy.
2. Detection of Different DDoS Attack Types
   - Various DDoS attack types are analyzed independently to improve detection precision.

Methodology:

Research Methodology:

First, we identified the research question, focusing on cybersecurity—specifically, the detection of DDoS attacks in IoT devices**.**

Literature Selection:

To gather relevant studies, we searched platforms like Google Scholar and ResearchGate using keywords such as:

- "Cybersecurity"
- "IoT"
- "Machine Learning"

During this process, we encountered several recurring papers. After filtering, we selected 50 studies most relevant to our topic.

Data Collection & Analysis:

As we reviewed these papers, we identified common challenges and gaps in the field. We then analyzed data from these studies to understand the current research landscape on DDoS attacks in IoT environments. This helped us pinpoint a research gap and refine our focus.

Experimentation:

For experimentation, we applied an Application-Layer DDoS attack dataset. After pre-processing, including handling missing or null values, we used the Extra Trees feature selection technique to identify the most relevant features, as shown in Figure 2**.**

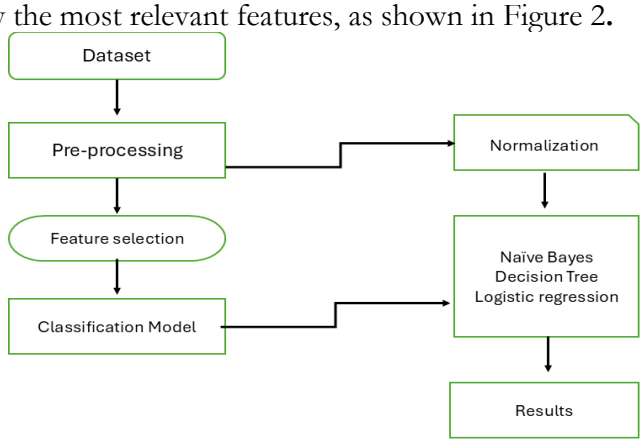


Figure 2. Methodology

Next, we utilized three machine learning classifiers—Naïve Bayes (NB), Decision Tree (DT), and Logistic Regression (LR)—to evaluate the performance of the selected features.

To ensure a fair comparison, we:

- Tuned hyper-parameters for each classifier to optimize accuracy.
- Applied k-fold cross-validation to enhance model robustness and prevent overfitting.

Finally, we compared the performance metrics of NB, DT, and LR. The results, presented in Figure 3, provide a clear assessment of their effectiveness in detecting DDoS attacks.

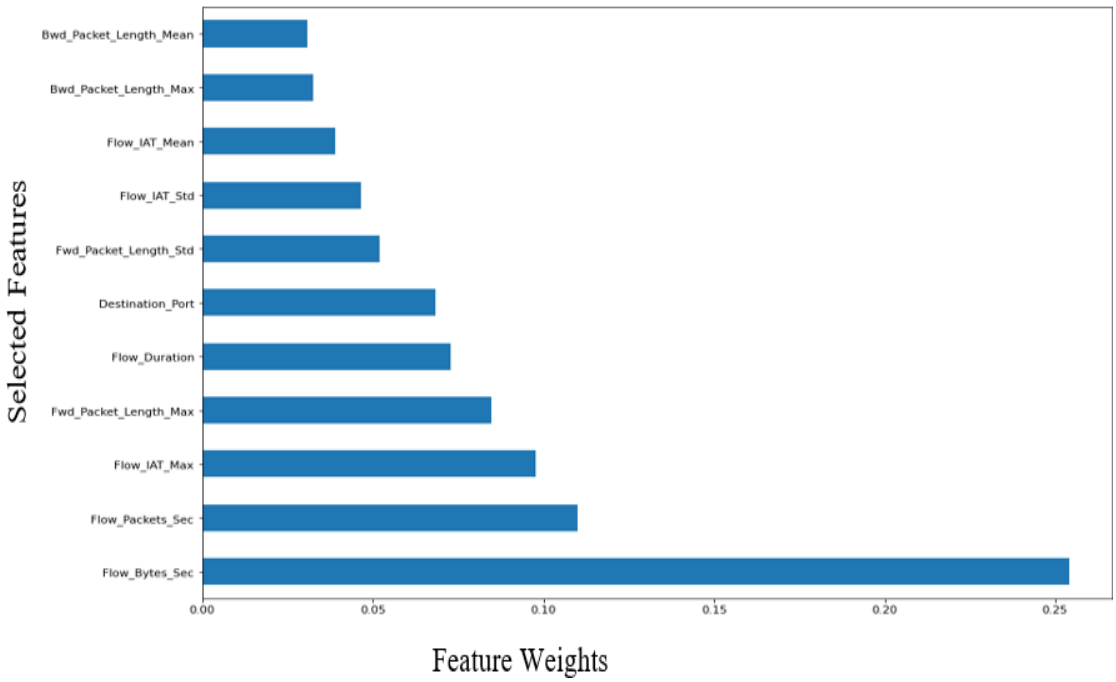


Figure 3. Extra Tree Feature selection

Decision Tree:

The Decision Tree (DT) algorithm is a supervised learning method used for classification. It follows a hierarchical structure to make decisions by partitioning data into subsets based on input values.

DT operates in a tree-like structure, where:

- **Branches** represent feature labels.
- **Leaves** represent class labels.
- **Decisions** are made at each branch, leading to target values at the leaves.

This structured approach helps in effectively classifying data [16].

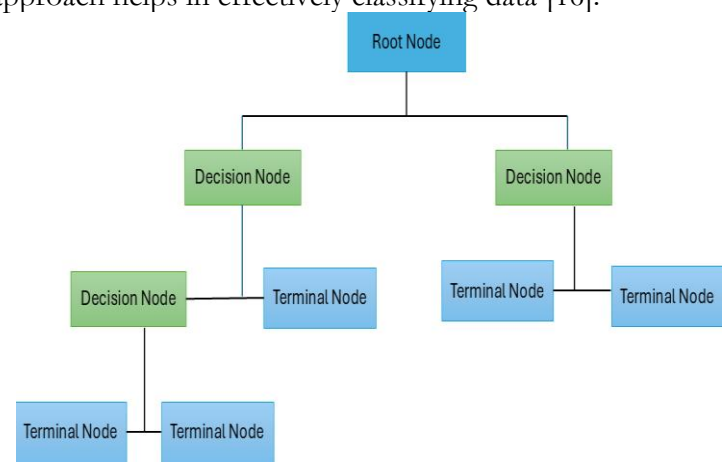


Figure. 4 Decision Tree Learning Architecture

Naive Bayes:

Naïve Bayes is a set of probabilistic algorithms based on Bayes' Theorem, commonly used for classification tasks. The term "Naïve" refers to the assumption that all features in the dataset are independent, though this is rarely true in real-world scenarios.

Despite this simplification, Naïve Bayes classifiers often deliver strong performance, especially in text classification tasks such as:

- **Sentiment analysis**
- **Spam detection**
- **Document categorization**

These characteristics make Naïve Bayes a widely used and effective classification technique [17].

$$P(A/B) = \frac{P(B/A).P(A)}{P(B)} \quad (1)$$

Logistic Regression:

Logistic Regression is a widely used machine learning approach for binary classification tasks. Despite its name, it is a classification model, not a regression one.

The primary goal of Logistic Regression is to estimate the probability that a given input belongs to a specific category. It does this by applying the sigmoid function, which maps predictions to values between 0 and 1, making it ideal for yes/no or true/false classification problems.

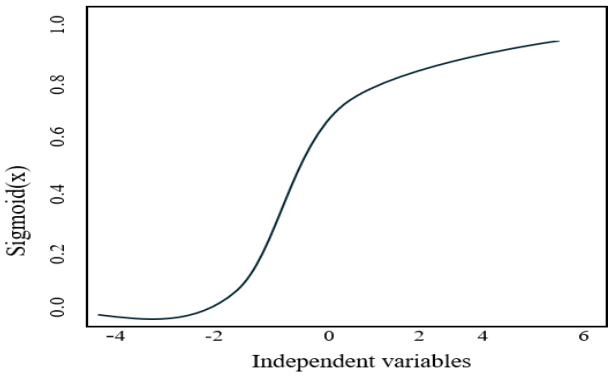


Figure 5 Logistic Regression

Result and Discussion:

The results presented in Table 2 highlight the performance of three machine learning classifiers—Decision Tree (DT), Naïve Bayes (NB), and Logistic Regression (LR)—based on key evaluation metrics. These results demonstrate the effectiveness of different algorithms in detecting application-layer DDoS attacks targeting IoT devices.

The models were evaluated using multiple metrics, including accuracy, precision, recall, F1-score, and cross-validation accuracy. Among these, the Decision Tree (DT) outperformed the other classifiers, achieving an accuracy of 99%. It also demonstrated high precision (98%), recall (99%), and F1-score (98%), making it a highly effective model for distinguishing between legitimate and malicious traffic.

Furthermore, the high recall value indicates that the Decision Tree is capable of detecting the majority of attack instances, making it a reliable choice for real-time security applications in IoT environments.

Table 2: Performance Results of Proposed Machine Learning Models for Detection of DDoS

Algorithm	Accuracy	Precision	Recall	F1-Score	Cv-Accuracy
Decision tree	99%	98	999	98	98.9
Naïve Bayes	94%	99	92	92	94.1
Logistic regression	89%	88	87	87	88

The Naïve Bayes (NB) classifier achieved a notable accuracy of 94%, with an exceptionally high precision of 99%. However, its recall was slightly lower at 92%, indicating that while the model is highly confident in its positive classifications, it may misclassify some attack instances as benign, leading to false negatives. Despite this, its cross-validation accuracy of 94.1% demonstrates consistent performance across different data splits, making it a reliable and generalizable model.

Among the three classifiers, Logistic Regression (LR) was the least effective. While it remains a viable option, its lower recall suggests that it may fail to detect all attack instances, which is critical in cybersecurity applications. However, its high precision indicates that when it does classify an instance as an attack, it is highly likely to be correct.

The bar chart in Figure 6 visually compares the accuracy of the three classification models: Decision Tree, Naïve Bayes, and Logistic Regression. The Decision Tree exhibits the highest accuracy, followed by Naïve Bayes, while Logistic Regression ranks the lowest. This visualization effectively highlights the performance differences among the models, aiding in the selection of the most suitable classifier for detecting application-layer DDoS attacks in IoT environments.

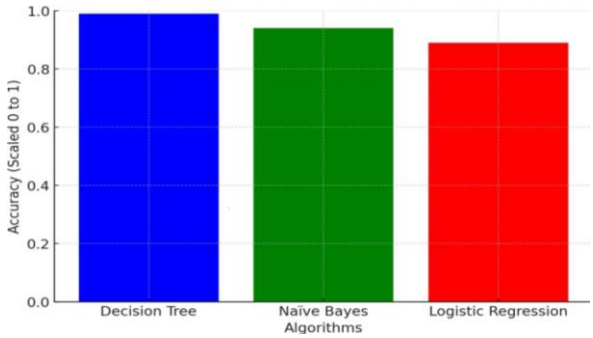


Figure 6. ML Classifier Comparison

Overall, the results indicate that Decision Trees are the most effective models for detecting application-layer DoS attacks in IoT environments, owing to their high accuracy and recall. Naïve Bayes also performs well, particularly in terms of precision, making it useful in scenarios where minimizing false positives is a priority. Logistic Regression, while still useful, is not the best choice when high recall is essential for detecting all attack instances.

The bar chart in Figure 6 visually compares the performance of the proposed machine learning models, clearly highlighting the Decision Tree as the most effective classifier.

Discussion:

This study demonstrates the effectiveness of machine learning-based approaches in detecting application-layer Distributed Denial of Service (DDoS) attacks in Internet of Things (IoT) environments. By leveraging the Extra Tree feature selection method, the proposed framework successfully reduces dataset dimensionality while preserving the most relevant features, enhancing both model accuracy and computational efficiency.

A comparative evaluation of three machine learning classifiers—Decision Tree, Naïve Bayes, and Logistic Regression—highlights the superior performance of the Decision Tree model, which achieved the highest accuracy (99%), precision (98%), recall (99%), and F1-score (98%). This confirms its effectiveness in distinguishing normal and attack traffic, making it a reliable choice for real-time DDoS detection in IoT networks.

The Naïve Bayes classifier also performed well, attaining 94% accuracy and a precision score of 99%. However, its lower recall (92%) suggests it may misclassify some attack instances as benign, leading to false negatives. This trade-off makes Naïve Bayes suitable for applications where minimizing false positives is critical, but less ideal for comprehensive attack detection. Logistic Regression, while achieving 89% accuracy, performed less effectively due to its lower recall and F1 score. This makes it a weaker option for highly imbalanced datasets or scenarios

requiring high recall, though its simplicity and interpretability may still be valuable in specific IoT applications.

Future Directions:

The results emphasize the importance of selecting appropriate machine learning models and feature selection techniques to improve the detection of sophisticated DDoS attacks that mimic legitimate traffic. The Extra Tree feature selection method effectively enhances classification accuracy by identifying key distinguishing features.

However, certain limitations remain, including the need for further optimization of classifiers for real-time deployment in large-scale IoT networks. Future research could explore:

- Advanced deep learning models**,** hybrid algorithms, or ensemble learning techniques to improve detection rates and reduce false positives**.**
- Expanding the dataset scope to include dynamic, real-time data for improved practical applicability in diverse IoT environments.

Conclusion:

Securing IoT Devices Against DDoS Attacks Using Machine Learning:

IoT devices play a crucial role in data collection and communication but remain vulnerable to cyber threats, particularly Denial of Service (DoS) and Distributed Denial of Service (DDoS) attacks. As the number of IoT devices grows, these threats become more severe.

To address this challenge, we explored DDoS attack detection at the application layer using Extra Tree feature selection and machine learning (ML) models. The results showed that the Decision Tree classifier achieved the highest accuracy (99%), outperforming Naïve Bayes (94%) and Logistic Regression (89%). This demonstrates the effectiveness of ML-based feature selection in strengthening IoT security against DDoS attacks.

Future research could focus on:

- **Advanced deep learning techniques** for improved accuracy and robustness.
- Real-time detection systems to enhance practical applications in IoT security.
- Hybrid models combine multiple algorithms for **stronger and more adaptive defense mechanisms** against evolving cyber threats.

Future Work:

Although, this approach works well. There are ways to make it better. One improvement could be making the system faster so it can handle large data easily. Another idea is to use advanced machine and deep learning algorithms to make results more accurate. Future research can also focus on testing this method in real-world situations to see how well it works.

Acknowledgment: Sincere gratitude is expressed to Dr. Muhammad Usman Sana for invaluable guidance and support throughout this research, which greatly contributed to the completion of this paper.

Author's Contribution: Mustabeen Aziz led the research and experiments. Muhammad Usman Sana and Tayybah Kiren supervised and reviewed the work. Alvena Ehsan, Tahmina Ehsan, and Fateha Minahil contributed to data analysis, literature review, and manuscript preparation.

Conflict of Interest: The authors declare no conflict of interest regarding this publication.

References:

[1] D. Kshirsagar and S. Kumar, "A feature reduction based reflected and exploited DDoS attacks detection system," *J. Ambient Intell. Humaniz. Comput.*, vol. 13, no. 1, pp. 393–405, Jan. 2022, doi: 10.1007/S12652-021-02907-5/METRICS.

[2] H. B. and M. F. D. Mohammed Sharif, "Detection of Application-Layer DDoS Attacks Produced by Various Freely Accessible Toolkits Using Machine Learning," *IEEE Access*, vol. 11, pp. 51810–51819, 2023, doi: 10.1109/ACCESS.2023.3280122.

[3] A. Praseed and P. S. Thilagam, "Modelling Behavioural Dynamics for Asymmetric Application Layer DDoS Detection," *IEEE Trans. Inf. Forensics Secur.*, vol. 16, pp. 617–626, 2021, doi: 10.1109/TIFS.2020.3017928.

[4] D. M. S. and M. F. H. Beitollahi, "Application Layer DDoS Attack Detection Using Cuckoo Search Algorithm-Trained Radial Basis Function," *IEEE Access*, vol. 10, pp. 63844–63854, 2022, doi: 10.1109/ACCESS.2022.3182818.

[5] C. Benzaid, M. Boukhalfa, and T. Taleb, "Robust Self-Protection Against Application-Layer (D)DoS Attacks in SDN Environment," *IEEE Wirel. Commun. Netw. Conf. WCNC*, vol. 2020-May, May 2020, doi: 10.1109/WCNC45663.2020.9120472.

[6] A. Munshi, N. A. Alqarni, and N. Abdullah Almalki, "DDOS Attack on IOT Devices," *ICCAIS 2020 - 3rd Int. Conf. Comput. Appl. Inf. Secur.*, Mar. 2020, doi: 10.1109/ICCAIS48893.2020.9096818.

[7] M. Odusami, S. Misra, O. Abayomi-Alli, A. Abayomi-Alli, and L. Fernandez-Sanz, "A survey and meta-analysis of application-layer distributed denial-of-service attack," *Int. J. Commun. Syst.*, vol. 33, no. 18, p. e4603, Dec. 2020, doi: 10.1002/DAC.4603.

[8] V. Gaur and R. Kumar, "Analysis of Machine Learning Classifiers for Early Detection of DDoS Attacks on IoT Devices," *Arab. J. Sci. Eng.*, vol. 47, no. 2, pp. 1353–1374, Feb. 2022, doi: 10.1007/S13369-021-05947-3/METRICS.

[9] S. Z. & S. M. Mohammad Najafimehr, "A hybrid machine learning approach for detecting unprecedented DDoS attacks," *J. Supercomput.*, vol. 78, pp. 8106–8136, 2022, doi: https://doi.org/10.1007/s11227-021-04253-x.

[10] C. V.-R. and J. A. P.-D. N. M. Yungaicela-Naula, "SDN-Based Architecture for Transport and Application Layer DDoS Attack Detection by Using Machine and Deep Learning," *IEEE Access*, vol. 9, pp. 108495–108512, 2021, doi: 10.1109/ACCESS.2021.3101650.

[11] I. Sharafaldin, A. H. Lashkari, S. Hakak, and A. A. Ghorbani, "Developing realistic distributed denial of service (DDoS) attack dataset and taxonomy," *Proc. - Int. Carnahan Conf. Secur. Technol.*, vol. 2019-October, Oct. 2019, doi: 10.1109/CCST.2019.8888419.

[12] M. T. Mahmood, S. R. A. Ahmed, and M. R. A. Ahmed, "Using Machine Learning to Secure IOT Systems," *4th Int. Symp. Multidiscip. Stud. Innov. Technol. ISMSIT 2020 - Proc.*, Oct. 2020, doi: 10.1109/ISMSIT50672.2020.9254304.

[13] T. T. Khoei, G. Aissou, W. C. Hu, and N. Kaabouch, "Ensemble Learning Methods for Anomaly Intrusion Detection System in Smart Grid," *IEEE Int. Conf. Electro Inf. Technol.*, vol. 2021-May, pp. 129–135, May 2021, doi: 10.1109/EIT51626.2021.9491891.

[14] V. Gaur and R. Kumar, "FSMDAD: Feature Selection Method for DDoS Attack Detection," *Proc. Int. Conf. Electron. Renew. Syst. ICEARS 2022*, pp. 939–944, 2022, doi: 10.1109/ICEARS53579.2022.9752308.

[15] V. Gaur and R. Kumar, "ET-RF based Model for Detection of Distributed Denial of Service Attacks," *Int. Conf. Sustain. Comput. Data Commun. Syst. ICSCDS 2022 - Proc.*, pp. 1205–1212, 2022, doi: 10.1109/ICSCDS53736.2022.9760938.

[16] K. C. Amir Mosavi, Pinar Ozturk, "Flood Prediction Using Machine Learning Models: Literature Review," *Water*, vol. 10, no. 11, p. 1536, 2018, doi: https://doi.org/10.3390/w10111536.

[17] T. R. Mohamed El Kourdi, Amine Bensaid, "Automatic Arabic document categorization based on the Naïve Bayes algorithm," *Proc. Work. Comput. Approaches to Arab. Script-based Lang. (Semitic '04). Assoc. Comput. Linguist. USA*, pp. 51–58, 2004, [Online]. Available: https://dl.acm.org/doi/10.5555/1621804.1621819

A FEM Analysis of BLDC Ceiling Fan with Different Slot-Pole Combinations

Usman Haider[1]*, Nazam Siddique[2], Muhammad Rashad[3], Sajjad Miran[4], Muhammad Wasif[2], Muhammad Jehanzeb Irshad[2]
[1]Department of RND, Champion Fans, Gujranwala, Pakistan
[2]Department of Electrical Engineering, University of Gujrat, Pakistan
[3]Department of Electrical Engineering, The University of Lahore, Pakistan
[4]Department of Mechanical Engineering, University of Gujrat, Pakistan
*** Correspondence:** engrusmanhaider68@gmail.com

Citation | Haider. U, Siddique. N, Rashad. M, Miran. S, Wasif. M, Irshad. M. J, "A FEM Analysis of BLDC Ceiling Fan with Different Slot-Pole Combinations", IJIST, Special Issue. pp 208-219, March 2025
Received | Feb 22, 2025 **Revised |** March 06, 2025 **Accepted |** March 12, 2025 **Published |** March 15, 2025.

BLDC motors have recently made significant advancements in the automation industry. Due to their high efficiency and power density, they are widely used in everyday applications such as fans, electric bikes, rail transit, and automobiles. The slot-pole structure is a key factor influencing motor design. This research explores various slot-pole combinations to enhance performance.

For ceiling fan applications, a balanced and highly efficient stator with concentrated winding has been designed based on different slot-pole configurations. Two commonly used combinations—18-slot/16-pole and 12-slot/14-pole—were analyzed. However, these configurations result in high cogging torque and a low winding factor, reducing the efficiency of BLDC ceiling fans.

To overcome these issues, a 24-slot/22-pole combination is proposed. This design improves torque production, power efficiency, and magnetic flux density while reducing cogging torque and increasing cogging frequency. The effectiveness of this structure is evaluated using the finite element method (FEM) in Ansys Electronics Desktop software.

Keywords: Brushless DC (BLDC) Motor; Ansys Maxwell Rmxprt; Ceiling Fans; slot-to-pole combination and Finite Element Analysis.

Introduction:

In the late 1960s, P.H. Trickey and T.G. Wilson invented and designed the brushless DC (BLDC) motor. However, due to the unavailability of permanent magnets, this concept became a reality only in the 1980s when Power Tec Industrial Corporation confirmed their availability [1]. The BLDC motor consists of a stator, rotor, and an inverter circuit that drives the motor. It is considered one of the most efficient types of electric motors due to its high power density, high-speed capabilities, durability, and simple design [2]. These features make BLDC motors widely used in applications such as rolling and automotive industries [3], electric vehicles [4][5][6], ceiling fans [7], washing machines [8], air conditioners [9], gliders, drones [10], CNC machines [11], robotics [12], and marine applications [13].

BLDC motors come in two main types: inner and outer rotor designs. Motors with a higher number of magnetic poles produce greater torque, which is why an outer rotor BLDC motor is preferred for such applications [14][15]. In contrast, inner rotor BLDC motors are used for high-speed applications. They have fewer magnetic poles, allowing higher rotational speeds but limiting torque production [16].

To achieve optimal performance, a suitable drive control topology is essential. It helps minimize speed fluctuations and ensures torque stability. Various control methods have been explored for BLDC motors, including PID control [17][18], fuzzy logic control [19], artificial neural networks (ANN) [20][21], and other advanced techniques.

Ansys Maxwell has been widely used to design, verify, and validate different motor types. For instance, induction motors [22], brushed DC motors [23], and three-phase induction motors [24] have been designed using this software. Similarly, Ansys Maxwell has been applied to develop AC synchronous motors [25], permanent magnet synchronous motors (PMSM) [26], synchronous reluctance motors (SRM) [27], and brushless DC (BLDC) motors [28][27].

Research studies [2][5], [27][29] indicate that altering the number of poles and stator slots in a BLDC motor affects power output and load-handling capacity. A comparative analysis of three different slot-pole combinations—12-slot/10-pole, 18-slot/20-pole, and 9-slot/8-pole—has been conducted. Findings suggest that the 18-slot/20-pole combination delivers the highest torque at low speeds, making it suitable for agro-electric vehicle (EV) applications [30]. Additionally, a comparison of 12, 18, 24, and 30-slot stators with 4-pole rotors indicates that the 30-slot stator provides the lowest cogging torque and best performance [30]. Another study [31] found that a 24-slot/16-pole BLDC motor is more efficient than a 36-slot/18-pole motor. Researchers in [29] compared three different slot configurations (6, 12, and 15) with 4-pole rotors and concluded that the 15-slot stator offers lower total harmonic distortion (THD) and a more sinusoidal back-EMF waveform than the 6-slot and 12-slot motors.

While existing research focuses on high-torque EV applications, it does not provide sufficient insights for designing BLDC motors for ceiling fans. This study aims to address that gap by optimizing BLDC motor design specifically for ceiling fan applications.

The literature suggests that a higher slot number results in better torque production, lower THD, and a more sinusoidal back-EMF. This research investigates three different slot-pole combinations—12/14, 18/16, and 24/22—using the finite element method (FEM) in the Ansys Maxwell RMxprt environment. These combinations are analyzed based on key design parameters, back-EMF characteristics, and cogging torque performance.

Problem Statement:

BLDC motors are gaining popularity due to their high-power density and simplicity. However, improper motor design can lead to low power efficiency and high cogging torque. In the fan industry, two conventional slot-pole combinations—12-slot/14-pole and 18-

slot/16-pole—are commonly used [7]. These combinations exhibit high cogging torque and lower power density, resulting in reduced efficiency and shorter lifespan.

Research Objectives:

The main objective of this research is to highlight the inefficiencies of the commonly used 12-slot/14-pole and 18-slot/16-pole BLDC ceiling fan configurations. These two combinations are analyzed in terms of power density, cogging torque, back-EMF, and winding factor. A comparison is made with the proposed 24-slot/22-pole combination, which offers higher cogging frequency and lower cogging torque. Additionally, the proposed design improves power density and provides a more sinusoidal back-EMF. Finite element analysis (FEM) is conducted using Ansys Maxwell RMxprt to validate the findings.

Research Methodology:

The design of a BLDC motor depends on various factors, including the slot-pole combination, stator winding patterns, the number of turns, magnetic field density, and air gap. This research focuses specifically on the impact of slot-pole combinations. Figure 1 illustrates the research methodology.

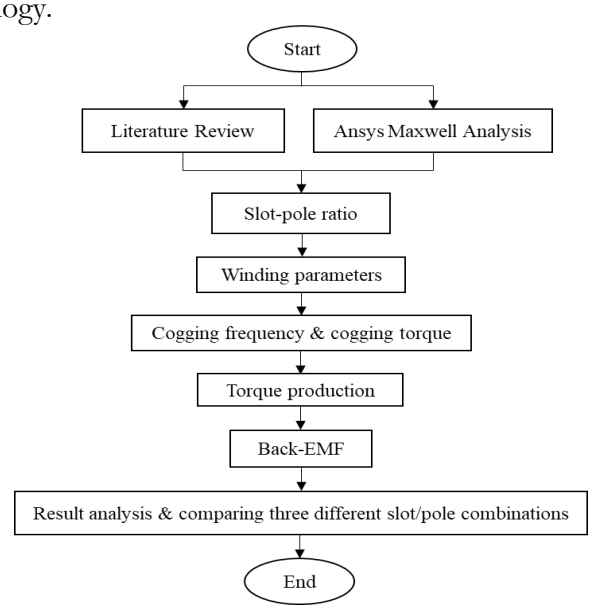


Figure 1. Methodology flow diagram

Currently, the ceiling fan industry primarily uses two slot-pole combinations: 12-slot/14-pole and 18-slot/16-pole [7]. However, these configurations result in lower torque production, reduced winding factors, lower winding periodicity factors, and high cogging torque, which negatively impact motor efficiency.

This research proposes a 24-slot/22-pole combination, which enhances torque production, winding periodicity factor, winding efficiency, and cogging frequency. The 12/14, 18/16, and 24/22 slot-pole configurations, designed in Ansys Maxwell RMxprt, are shown in Figure 2

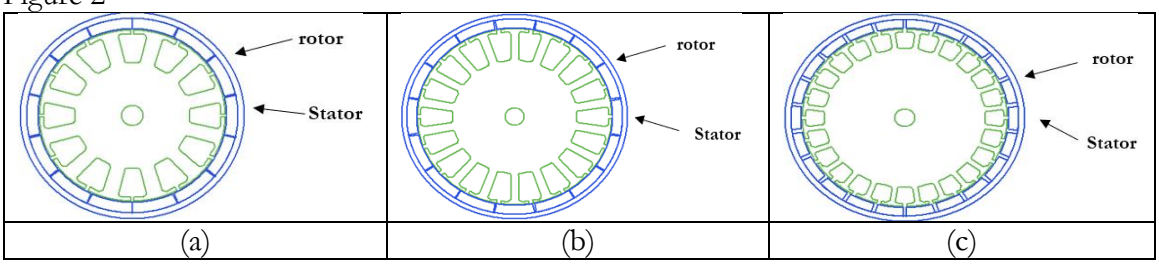


Figure 2. Ansys designed a BLDC motor with (a) 12-slot, 14-pole, (b) 18-slot, 16-pole, (c) 24-slot, 22-pole combinations.

Table 1 presents the design parameters for the 12-slot/14-pole, 18-slot/16-pole, and 24-slot/22-pole BLDC motors. The maximum observed number of turns per slot is 300. A small air gap of 1 mm is chosen to achieve higher torque production with lower input power [32].

Table 1. Design parameters of different slot-pole BLDC motors

Parameters	**slot-pole combinations**		
	12/14	18/16	24/22
Stator diameter (mm)	132.5	158	158
Steel type	M19_24G		
Number of winding layers	2		
Stacking factor	0.95		
Number of turns	300	300	300
Rotor position	outer rotor		
Rotor inner diameter (mm)	135	159	159
Rotor outer diameter (mm)	160	185	185
Length of magnets (mm)	25		
Magnet type	XG196/96		
Control mode	DC		

Slot-Pole Ratio:

The slot-pole ratio, denoted as 'q,' is a key factor in BLDC motor design. Mathematically, it is defined by Equation (1):

$$q = \frac{N_s}{N_{ph}.N_p} \quad (1)$$

Where N_s = number of slots
N_{ph}= number of phases
N_p = number of poles

This ratio influences torque, electromagnetic performance, and overall efficiency. A higher ratio results in greater torque production, while a lower ratio reduces torque because multiple rotor poles interact with a single stator slot.

Winding Periodicity:

The winding periodicity factor (z) represents the repetition of winding patterns in the stator. A higher value improves magnetic balance and reduces cogging torque. Mathematically, it is defined by Equation (2).

$$z = \frac{N_s}{\gcd(N_s, N_{ph}.N_p)} \quad (2)$$

Where, gcd = greatest common divisor,

Cogging Frequency:

Cogging frequency is a key factor in determining the generated cogging torque. A higher cogging frequency results in lower cogging torque. Mathematically, it is defined by Equation (3).

$$f_{cog} = LCM(N_s, N_p) \quad (3)$$

Where LCM denotes the least common multiple.

Ansys Maxwell Analysis:

Using Ansys Maxwell, the winding factor, cogging torque, generated torque, and back electromotive force (BEMF) have been analyzed.

Winding Factor:

The winding factor (kw) measures the efficiency of stator windings in generating magnetic flux and back-EMF. Mathematically, it is defined by Equation (4).

$$k_w = k_{mn}.k_{en} \quad (4)$$

Where, k_w = winding factor,
k_{mn}= magnetic winding factor, &
k_{en}= electrical winding factor

Cogging torque periodicity

Using the finite element method (FEM), cogging torque and back-EMF analysis are performed. The study is conducted under different load conditions to evaluate cogging torque. The periodicity of cogging torque is determined by varying the rotor angle over a single-slot pitch [30]. Mathematically, it is defined by Equation (X).

$$P_{cog} = \frac{360}{LCM(N_s, N_p)} \qquad (5)$$

Results:

The results of this study provide a detailed evaluation of various slot-pole configurations in BLDC motors for ceiling fan applications, focusing on cogging torque, back electromotive force (EMF), efficiency, and overall performance. This section presents a comparative analysis of simulation results, including torque profiles, back-EMF waveforms, and efficiency trends for each configuration. The findings highlight the trade-offs between cogging torque reduction, energy efficiency, and manufacturing feasibility, offering insights for optimizing energy-efficient BLDC ceiling fan motors. Figure 3 illustrates the winding patterns. Transient analysis in Ansys Maxwell shows that the 24-slot/22-pole combination has the highest inductance value due to its greater number of slots and turns.

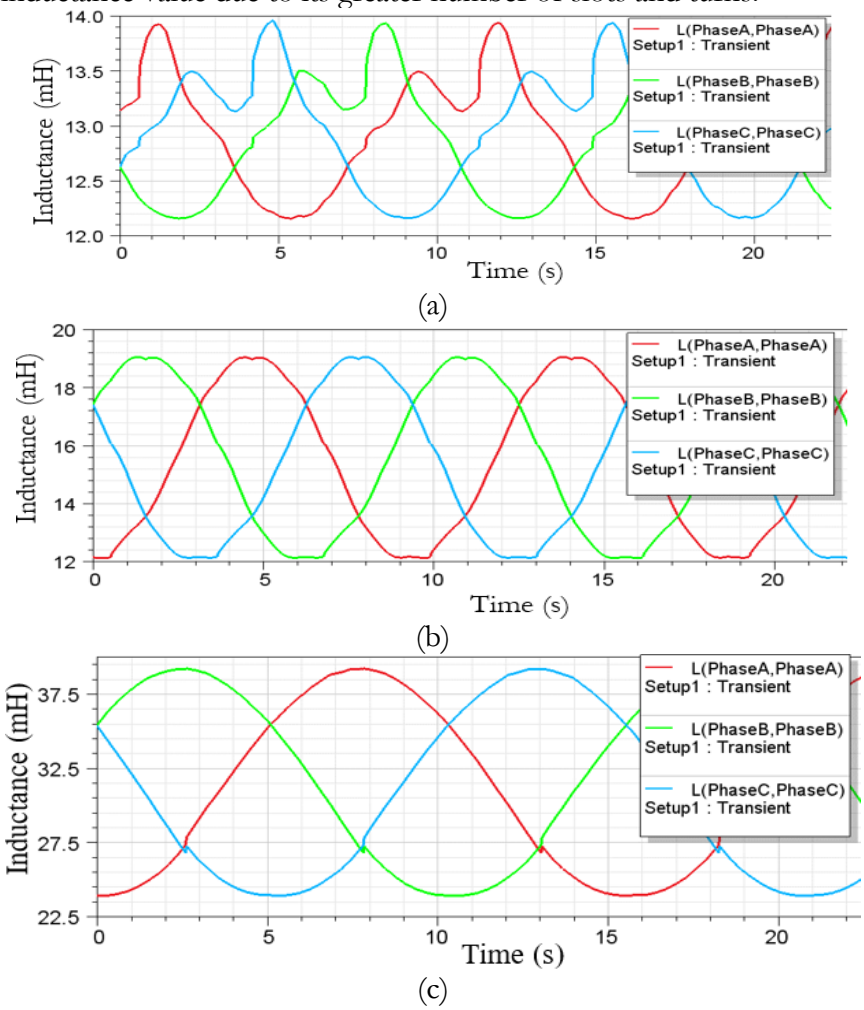


Figure 3. Winding patterns of BLDC motor with (a) 12-slot, 14-pole, (b) 18-slot, 16-pole, and (c) 24-slot, 22-pole combinations.

The analysis of Figure 3 shows that the BLDC motor with a 12-slot/14-pole configuration exhibits waveform irregularities, leading to torque ripples and unstable performance. Similarly, the 18-slot/16-pole motor still shows non-uniformity in its waveforms. In contrast, the 24-slot/22-pole BLDC motor maintains a stable, sinusoidal waveform, resulting in lower torque ripples, improved stability, and optimal efficiency.

Table 2 shows that the 24-slot/22-pole BLDC motor has the lowest cogging torque periodicity. Using Equations (1-5), the efficiency parameters for the 12/14, 18/16, and 24/22 slot-pole combinations are presented in Table 2.

Table 2. Calculated efficiency of different slot-pole combinations

N_s	N_p	q	z	k_w	f_{cog}	P_{cog}
12	14	0.29	2	0.93	84	4.29
18	16	0.38	3	0.94	144	2.5
24	22	0.36	4	0.95	264	1.36

Table 2 shows that the 24-slot/22-pole BLDC motor generates lower cogging torque while producing higher magnetic flux and back EMF. Figure 4 illustrates the cogging torque for different slot-pole BLDC motor configurations.

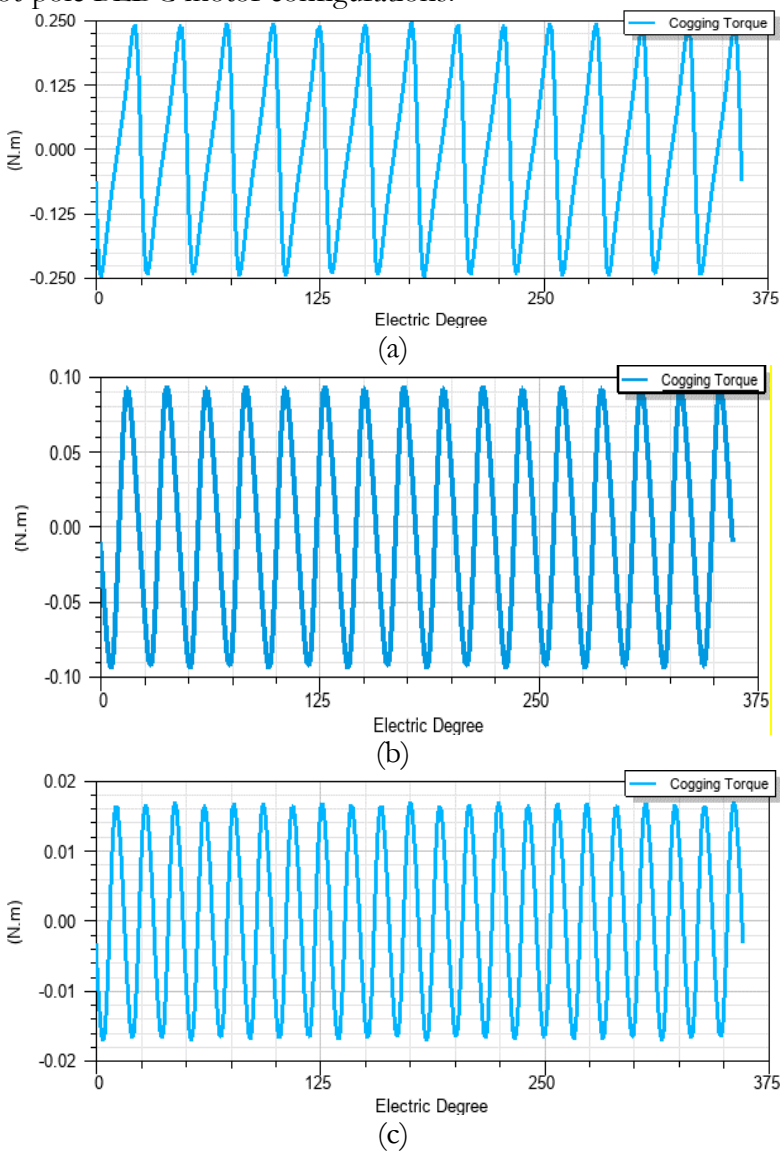


Figure 4. Cogging torque generated by (a) 12-slot, 14-pole, (b) 18-slot, 16-pole, and (c) 24-slot, 22-pole combinations.

Figure 4 shows that the 24-slot/22-pole BLDC motor has the lowest cogging torque of 0.18 Nm, while the 12/14 and 18/16 slot-pole combinations produce 0.24 Nm and 0.9 Nm, respectively.

Torque Production:

When the stator magnetic fields interact with the rotor's permanent magnet fields, torque is generated in the motor. In the 24-slot/22-pole BLDC motor, more stator and rotor magnetic fields interact, resulting in higher torque. Due to its high q value, this combination has smaller torque ripples compared to the other two configurations.

Using FEM, the torque generated by the three motor designs is shown in Figure 5. Torque instability leads to inefficiency and excessive heat generation.

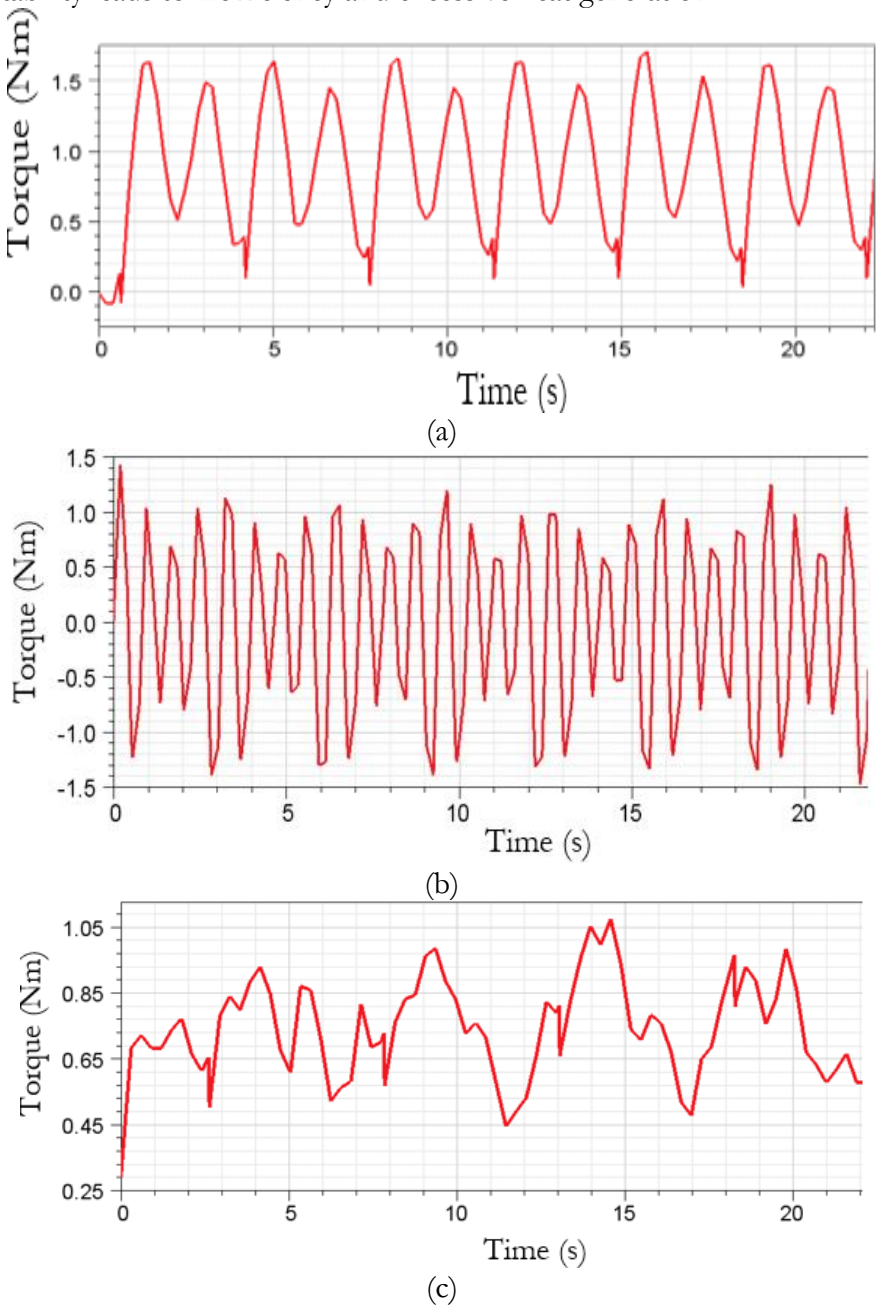


Figure 5. BLDC motor torque is generated by (a) 12-slot, 14-pole, (b) 18-slot, 16-pole, and (c) 24-slot, 22-pole combinations.

Figure 5 shows that the 12-slot/14-pole BLDC motor produces unstable torque, oscillating between 0.2Nm and 1.5Nm. This instability can lead to inefficient motor rotation

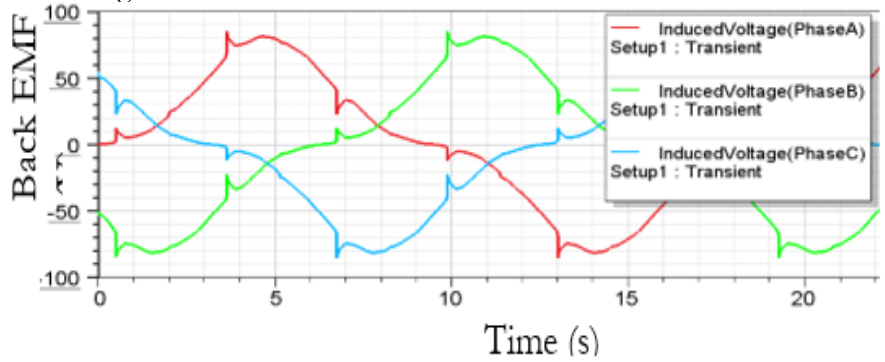
and high torque ripples. Similarly, the 18-slot/16-pole BLDC motor exhibits torque fluctuations, making it unsuitable for ceiling fan applications.

In contrast, the 24-slot/22-pole BLDC motor provides the most stable torque output, making it a better choice for ceiling fan applications.

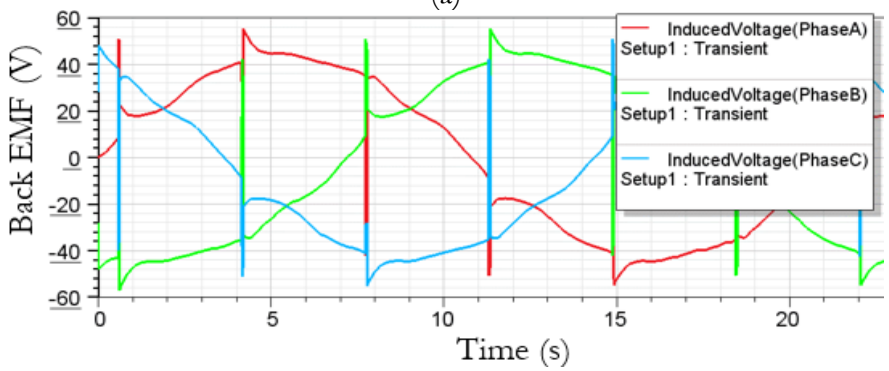
Back-EMF:

The FEM results from Ansys Maxwell illustrate the back EMF generated by the three different slot-pole combinations, as shown in Figure 6. A higher back EMF corresponds to a higher rotational speed (rpm) but lower torque production.

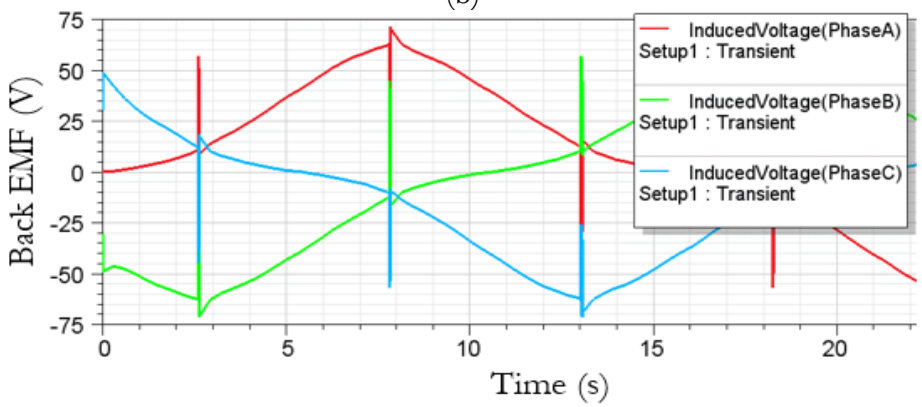
Figure 6 indicates that the BLDC motor with a 12-slot/14-pole combination produces the highest back EMF at 80V, followed by the 24-slot/22-pole motor at 62V, while the 18-slot/16-pole motor generates the lowest back EMF at 42V.



(a)



(b)



(c)

Figure 6. Back-EMF of BLDC motor with by (a) 12-slot, 14-pole, (b) 18-slot, 16-pole, and (c) 24-slot, 22-pole combinations

The waveforms in Figure 6 indicate that the BLDC motors with 12-slot/14-pole and 18-slot/16-pole combinations exhibit quasi-sinusoidal back EMF (BEMF). In contrast, the

24-slot/22-pole BLDC motor produces a pure sinusoidal BEMF waveform, making it an optimal choice for field-oriented control (FOC) commutation topology [33].

24-Slot and 22-Pole BLDC Motor Design:

The literature review suggests that the 24-slot and 22-pole BLDC motor offers superior efficiency due to its low cogging torque, high winding factor, reduced back EMF, and increased torque production. A practical demonstration of the 24-slot and 22-pole BLDC motor is illustrated in Figure 7.

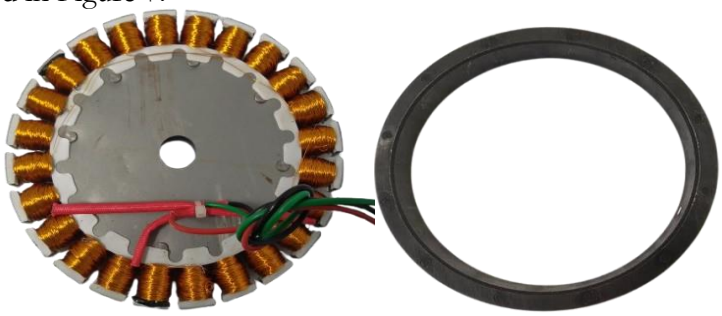


Figure 7. BLDC motor with (a) 24-slot stator, and (b) 22-pole rotor.

Table 3 presents the measured parameters of the proposed 24-slot and 22-pole BLDC motor when operating at 220V. These parameters include torque, efficiency, cogging torque, back EMF, and overall performance characteristics, confirming its superior efficiency and stable operation compared to conventional configurations.

Table 3. Measured parameters of 24-slot, 22-pole BLDC motor

Measured values	Parameters
220V	V_{in}
0.17A	I_{in}
40W	P_{in}
0.15A	Phase current
0.8N.m.	Average Torque

Table 3. shows the measured parameters of the 24-slot and 22-pole BLDC motor.

Discussion:

Your conclusion effectively summarizes the study, highlighting key findings and their implications. However, there is a contradiction: earlier, the 24-slot/22-pole configuration was presented as the best-performing option, but the conclusion states that the 18-slot/16-pole configuration provides the best balance. You may want to clarify this point. Additionally, you could briefly reinforce why the 24-slot/22-pole setup, despite its advantages, faces practical challenges. Would you like me to refine or streamline this section for better clarity?

Conclusion:

A BLDC motor with a 24-slot and 22-pole combination is proposed for ceiling fan applications. This configuration is compared with the commonly used 12-slot/14-pole and 18-slot/16-pole combinations. Using Finite Element Method (FEM) analysis, key efficiency parameters are evaluated. Since ceiling fans require high torque for optimal performance, an outer rotor BLDC motor is utilized.

The proposed slot-pole combination demonstrates high torque production, a high cogging frequency, and an improved winding factor while maintaining a sinusoidal back EMF. Additionally, it produces a lower back EMF, which enhances torque generation while operating at a lower input voltage. This research focuses specifically on ceiling fan applications. For high-torque applications such as electric vehicles (EVs), BLDC motors with a greater number of slots and poles should be considered.

Acknowledgment: The authors are grateful to Champion Fans for providing invaluable resources and support throughout the research. Their dedication to energy-efficient solutions

and innovation was a major factor in this successful work. The authors also acknowledge the encouragement and support of their mentors and fellow workers at Champion Fans.

Author's Contribution: All authors contributed equally.

Conflict of interest: The authors declare no conflict of interest for publishing this manuscript in IJIST.

References:

[1] K. Gadekar, S. Joshi, and H. Mehta, "Performance Improvement in BLDC Motor Drive Using Self-Tuning PID Controller," *Proc. 2nd Int. Conf. Inven. Res. Comput. Appl. ICIRCA 2020*, pp. 1162–1166, Jul. 2020, doi: 10.1109/ICIRCA48905.2020.9183219.

[2] A. Kumar, R. Gandhi, R. Wilson, and R. Roy, "Analysis of Permanent Magnet BLDC Motor Design with Different Slot Type," *2020 IEEE Int. Conf. Power Electron. Smart Grid Renew. Energy, PESGRE 2020*, Jan. 2020, doi: 10.1109/PESGRE45664.2020.9070532.

[3] D. Mohanraj et al, "A Review of BLDC Motor: State of Art, Advanced Control Techniques, and Applications," *IEEE Access*, vol. 10, pp. 54833–54869, 2022, doi: 10.1109/ACCESS.2022.3175011.

[4] H. M. Chico Hermanu Brillianto Apribowo, Musyaffa' Ahmad, "Fuzzy Logic Controller and Its Application in Brushless DC Motor (BLDC) in Electric Vehicle - A Review," *J. Electr. Electron. Inf. Commun. Technol.*, vol. 3, no. 1, p. 35, 2021, doi: 10.20961/jeeict.3.1.50651.

[5] M. E. Mehmet Akar, "BLDC Motor Design and Application for Light Electric Vehicle," *Afyon Kocatepe Univ. J. Sci. Eng.*, vol. 21, no. 2, pp. 326–336, 2021, doi: https://doi.org/10.35414/akufemubid.889877.

[6] A. N. Patel, "Slot opening displacement technique for cogging torque reduction of axial flux brushless DC motor for electric two-wheeler application," *Electr. Eng. Electromechanics*, 2023, doi: https://doi.org/10.20998/2074-272X.2023.2.02.

[7] S. Shastri, U. Sharma, and B. Singh, "Design and analysis of brushless DC motors for ceiling fan application," *9th IEEE Int. Conf. Power Electron. Drives Energy Syst. PEDES 2020*, Dec. 2020, doi: 10.1109/PEDES49360.2020.9379863.

[8] A. Dhar *et al.*, "Tribrid IoT-based intelligent washing machine," *Meso*, vol. 2845, no. 1, Sep. 2023, doi: 10.1063/5.0157247.

[9] M. Ben Slimene, "Solar based Air Conditioner with Standalone BLDC, Charger Controller and Battery Backup for Improved Efficiency," *Prz. Elektrotechniczny*, vol. 2023, no. 9, pp. 141–144, 2023, doi: 10.15199/48.2023.09.26.

[10] M. W. Andrzej Sikora, Adam Zielonka, "Minimization of Energy Losses in the BLDC Motor for Improved Control and Power Supply of the System under Static Load," *Sensors*, vol. 22, no. 3, p. 1058, 2022, doi: https://doi.org/10.3390/s22031058.

[11] A. Lavanya, S. Revathi, N. Sivakumaran, and K. Rajkumar, "Control of Feed Drives in CNC Machine Tools Using Artificial Immune Adaptive Strategy," *Artif. Intell. Cyber Secur. Ind. 4.0*, pp. 237–251, 2023, doi: 10.1007/978-981-99-2115-7_10.

[12] C. A. C.-V. Erick Axel Padilla-García,Héctor Cervantes-Culebro,Alejandro Rodriguez-Angeles, "Selection/control concurrent optimization of BLDC motors for industrial robots," *PLoS One*, 2023, doi: https://doi.org/10.1371/journal.pone.0289717.

[13] U. K. K. Tanmay Shukla, "A BL-CC Converter-Based BLDC Motor Drive for Marine Electric Vehicle Applications," *Int. Trans. Electr. Energy Syst.*, 2022, doi: https://doi.org/10.1155/2022/7026462.

[14] O. T. Ozturk Tosun, "The Design of the Outer-Rotor Brushless DC Motor and an Investigation of Motor Axial-Length-to-Pole-Pitch Ratio," *Sustainability*, vol. 14, no. 19, p. 12743, 2022, doi: https://doi.org/10.3390/su141912743.

[15] N. S. V. Naeini, "Optimum Design of the Outer Rotor Brushless DC Permanent Magnet Motor with Minimum Torque Ripples," *J. Oper. Autom. Power Eng.*, 2024, doi: 10.22098/joape.2024.14250.2093.

[16] K. R. Jayandiran Sundaram, "A new wedge shaped inner rotor for a BLDC motor: Performance analysis," *Therm. Sci. Eng. Prog.*, vol. 54, p. 102863, 2024, doi: https://doi.org/10.1016/j.tsep.2024.102863.

[17] A. N. N. Md Mahmud, S. M. A. Motakabber, A. H. M. Zahirul Alam, "Control BLDC Motor Speed using PID Controller," *Int. J. Adv. Comput. Sci. Appl.*, vol. 11, no. 3, 2020, doi: http://dx.doi.org/10.14569/IJACSA.2020.0110359.

[18] M. Mahmud, M. R. Islam, S. M. A. Motakabber, M. D. A. Satter, K. E. Afroz, and A. K. M. Ahasan Habib, "Control Speed of BLDC Motor using PID," *2022 IEEE 18th Int. Colloq. Signal Process. Appl. CSPA 2022 - Proceeding*, pp. 150–154, 2022, doi: 10.1109/CSPA55076.2022.9782030.

[19] F. A. Hari Maghfiroh, Musyaffa' Ahmad, Agus Ramelan, "Fuzzy-PID in BLDC Motor Speed Control Using MATLAB/Simulink," *J. Robot. Control*, vol. 3, no. 1, 2022, doi: https://doi.org/10.18196/jrc.v3i1.10964.

[20] A. B. S. Y. Porselvi Thayumanavan, Sai Ganesh Cs, "Artificial Neural Networks Based Analysis of BLDC Motor Speed Control," AXRIV. Accessed: Mar. 12, 2025. [Online]. Available: https://www.researchgate.net/publication/354207678_Artificial_Neural_Networks_Based_Analysis_of_BLDC_Motor_Speed_Control

[21] J. G.-G. Jose-Carlos Gamazo-Real, Víctor Martínez-Martínez, "ANN-based position and speed sensorless estimation for BLDC motors," *Measurement*, vol. 188, p. 110602, 2022, doi: https://doi.org/10.1016/j.measurement.2021.110602.

[22] R. M. B. Aishwarya, M., "Design of Energy-Efficient Induction motor using ANSYS software," *Results Eng.*, vol. 16, p. 100616, 2022, doi: https://doi.org/10.1016/j.rineng.2022.100616.

[23] H. C. Hsiao, C. Y. Hsiao, and G. R. Chen, "Finite Element Analysis and Optimal Design of DC Brush Motor for Automotive Engine Start Applications," *2019 IEEE Student Conf. Electr. Mach. Syst. SCEMS 2019*, Nov. 2019, doi: 10.1109/SCEMS201947376.2019.8972634.

[24] B. Ankita, D., M., "Three Phase Induction Motor -Model Design and Performance Analysis in ANSYS Maxwell," *Int. Res. J. Eng. Technol.*, vol. 8, no. 4, pp. 3125–3131, 2021.

[25] D. A. Sergey Gandzha, Bakhtiyor Kosimov, "Application of the Ansys Electronics Desktop Software Package for Analysis of Claw-Pole Synchronous Motor," *Machines*, vol. 7, no. 4, p. 65, 2019, doi: https://doi.org/10.3390/machines7040065.

[26] Doğukan Ayhan, "Analysis and optimization of interior permanent magnet synchronous motor for electric vehicle applications using ANSYS Motor-CAD," *Int. J. Automot. Eng. Technol.*, vol. 12, no. 3, pp. 105–120, 2023, doi: https://doi.org/10.18245/ijaet.1247462.

[27] B. Gecer, O. Tosun, H. Apaydin, and N. F. Oyman Serteller, "Comparative analysis of SRM, BLDC and induction motor using ANSYS/Maxwell," *Int. Conf. Electr. Comput. Commun. Mechatronics Eng. ICECCME 2021*, Oct. 2021, doi: 10.1109/ICECCME52200.2021.9591010.

[28] G. Z. Chuanhui Zhu, Rujie Lu, Congli Mei, Tao Peng, "Design and Simulation Analysis of Stator Slots for Small Power Permanent Magnet Brushless DC Motors," *Int. Trans. Electr. Energy Syst.*, 2023, doi: https://doi.org/10.1155/2023/1152243.

[29] A. D. Cemil Ocak, "A comparative analysis of four-pole brushless DC motors with different slot and winding arrangement based on THD values," *Int. J. Energy Appl.*

Technol., vol. 7, no. 1, pp. 7–12, 2020, doi: https://doi.org/10.31593/ijeat.669064.

[30] L. S. T. Muhammad Izanie Kahar, Raja Nor Firdaus Kashfi Raja Othman, Aziah Khamis, Nurfaezah Abdullah, Fairul Azhar Abdul Shukor, "Effect of Slot-Pole Numbers on the Performance of a BLDC Motor for Agro-EV Application," *ECTI Trans. Electr. Eng. Electron. Commun.*, vol. 20, no. 1, pp. 51–61, 2022, doi: https://doi.org/10.37936/ecti-eec.2022201.246104.

[31] A. T. Prakash, Anand, "Comparison analysis of slot design of brushless DC motor," *J. Harmon. Res. Eng.*, vol. 8, no. 3, pp. 77–80, 2020, [Online]. Available: https://escires.com/articles/JOHRE-8-286.pdf

[32] A. N. Patel and T. H. Panchal, "Sizing and Analysis of Dual Air-Gap Axial Flux Surface Mounted Permanent Magnet Brushless DC Motor for Electric Vehicle Application," *Proc. 3rd IEEE Int. Conf. Power Electron. Intell. Control Energy Syst. ICPEICES 2024*, pp. 324–328, 2024, doi: 10.1109/ICPEICES62430.2024.10719269.

[33] Z. Machhour, M. El Mrabet, Z. Mekrini, and M. Boulaala, "Comparative Study Between DTC and FOC Control Strategies Applied to the BLDC Motor: A Review," *Lect. Notes Networks Syst.*, vol. 930 LNNS, pp. 262–271, 2024, doi: 10.1007/978-3-031-54318-0_22.

Vortex Powerplant Implementation in A Coastal Community

Muzammil Anayat[1], Sajjad Miran[1*], Nazam Siddique[2], Waseem Arif[1], Yasir Hussain[1], Atta ul Hassnain[1]
[1]Department of Mechanical Engineering, University of Gujrat, Pakistan.
[1]Department of Electrical Engineering, University of Gujrat, Pakistan.
*** Correspondence:** Sajjad.miran@uog.edu.pk

Citation | Anayat. M, Miran. S, Siddique. N, Arif. W, Hussain. Y, Hassnain. A, "Vortex Powerplant Implementation in A Coastal Community", IJIST, Special Issue. pp 220-235, March 2025
Received | Feb 23, 2025 **Revised |** March 07, 2025 **Accepted |** March 13, 2025 **Published |** March 16, 2025.

A gravitational water vortex power plant is an eco-friendly device that generates electricity from renewable energy sources. In this system, a turbine extracts energy from the vortex created by tangentially channeling water into a circular basin. This article aims to explore the feasibility of implementing vortex power plant technology in coastal communities using an experimental model. The study investigates the potential of wastewater as a renewable energy resource by analyzing the relationship between flow rate, torque, and efficiency under different material and pipe configurations, particularly in urban areas. For experimental purposes, Gujrat city was selected. The wastewater outlet points near Bolley Bridge discharge approximately 74,714,000 liters per day. Based on our survey, the average household water usage in Gujrat city is 500 liters per day. An experimental model was designed to estimate potential energy generation. The model's design focused on optimizing the basin shape, inlets, outlets, and turbine configuration.

Using different pipes (cast iron and steel), the average water velocity and discharge rates were evaluated. The steel pipe produced higher velocity. Efficiency and production were further analyzed using LED lights, revealing that at 60 RPM, the system achieved significant efficiency and output voltage.

Keywords: GWVPP (Gravitational Water Vortex Power Plant), EFT (Eco-Friendly Technology), Hydropower, Wastewater, Turbine.

Introduction:

The increasing global demand for sustainable and decentralized energy solutions has driven significant interest in renewable energy technologies that utilize naturally available resources. Traditional energy sources, such as fossil fuels and grid-based electricity, not only contribute to carbon emissions and environmental degradation but also pose challenges in cost and accessibility, particularly for coastal and urban communities with limited infrastructure for large-scale renewable energy deployment. In response to these challenges, Gravitational Water Vortex Power Plants (GWVPPs) have emerged as a promising alternative, offering a low-maintenance, environmentally friendly method for small-scale power generation. While most studies have focused on implementing GWVPPs in natural streams and rural hydropower projects, the potential for harnessing urban wastewater discharge as a renewable energy source remains largely unexplored. The growing global emphasis on sustainable energy solutions has led to the exploration of innovative technologies that balance efficiency and environmental conservation. Among these, Gravitational Water Vortex Power plants (GWVPs), as shown in Figure 1, have emerged as a promising method for generating energy from low-head water flows. Unlike conventional hydropower systems, which often require large infrastructure and cause environmental disruption, GWVPs offer a low-impact alternative.

A typical GWVP system consists of an inlet channel to guide water, a circular basin where the water circulation forms a vortex that converts kinetic energy into mechanical energy, a water outlet or discharge point at the base, and turbine blades connected to a shaft. This shaft drives the generator, enabling the conversion of mechanical energy into electrical energy. GWVPPs operate by directing water through a specially designed basin to create a stable vortex, which drives a turbine connected to a generator. Unlike conventional hydropower plants, which require significant elevation differences and large water flow rates, vortex power plants function efficiently in low-head water environments, making them suitable for urban wastewater applications. Several studies have investigated turbine design, vortex stability, and generator efficiency in rural implementations, demonstrating that GWVPPs can achieve high energy conversion rates with minimal ecological impact [1]. However, limited research has been conducted on optimizing GWVPP systems for wastewater-driven applications, particularly in coastal regions where urban water discharge is abundant.

This study aims to bridge this research gap by evaluating the feasibility of implementing GWVPP technology in an urban wastewater setting, focusing on coastal communities with high discharge volumes. A key aspect of this research is the optimization of basin shape and turbine design to maximize energy extraction efficiency. Additionally, this study compares the performance of different generator configurations, specifically a 12V DC motor and a Permanent Magnet Alternator (PMA), to determine the most effective energy conversion mechanism. The results of this study could contribute to the development of decentralized, small-scale hydropower solutions that integrate seamlessly with existing urban wastewater infrastructure, reducing reliance on traditional power sources while promoting environmental sustainability.

Beyond technical feasibility, this research also explores the economic and practical implications of wastewater-driven GWVPP implementation. By assessing power generation potential, system efficiency, and scalability, this study provides insights into how coastal urban areas can leverage wastewater as a renewable energy source. Furthermore, evaluating the economic viability, return on investment (ROI), and potential policy incentives will help determine whether GWVPP technology can be adopted at a municipal or community level. Ultimately, this research aims to establish a new paradigm for wastewater-based energy solutions, contributing to the broader goal of integrating renewable energy into urban sustainability initiatives.

This study evaluates the eco-friendly nature of vortex power plant technology by comparing it with conventional hydropower systems. It explores ways to enhance the plant's efficiency through the experimental optimization of basin shape, inlet and outlet configurations, and turbine design. Additionally, the study examines the cost-effectiveness and potential scalability of GWVPs for decentralized energy generation.

Objectives of the Study:

The primary objective of this study is to evaluate the feasibility and efficiency of a Gravitational Water Vortex Power Plant (GWVPP) for small-scale renewable energy generation in coastal urban communities. The specific objectives include:

- To design and implement a GWVPP system utilizing wastewater discharge for sustainable and decentralized energy production.
- To analyze the impact of basin shape and turbine design on power generation efficiency, optimizing system performance.
- To compare the effectiveness of different generator configurations (12V DC motor vs. Permanent Magnet Alternator) in improving energy output and efficiency.

Novelty Statement:

The novelty of this study lies in demonstrating the potential for implementing GWVPPs in small-scale urban settings, enabling cities to generate electricity from water outlets. Unlike traditional hydropower plants, this research explores the use of wastewater as a sustainable energy source, presenting an innovative solution for urban energy needs. The study specifically focuses on deploying GWVPPs in coastal communities, where such technology has yet to be widely tested. A key finding is that conical basins are more efficient than cylindrical ones, enhancing basin design for improved performance. Additionally, the study evaluates different pipe materials (cast iron and steel) to assess their impact on water velocity and discharge.

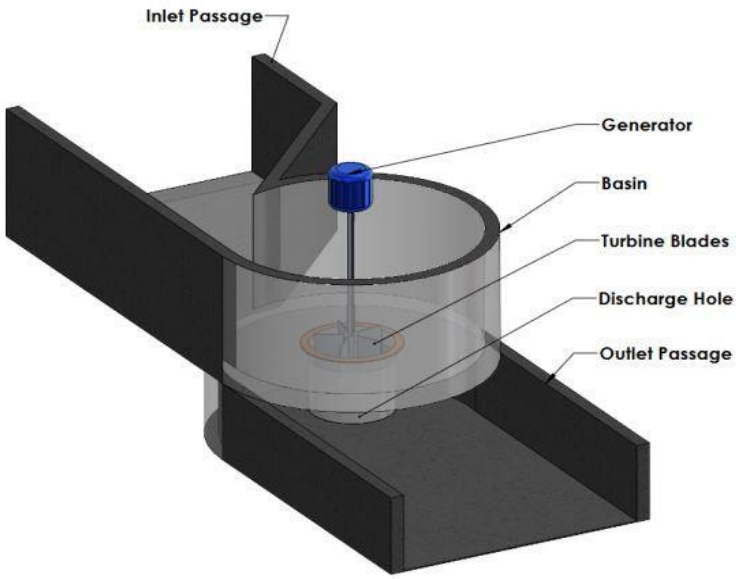


Figure 1. A Sketch of GWVP

Literature Review:

By channeling water into a circular basin, these systems create a vortex that powers turbines to generate electricity [2]. Their scalability, simplicity, and ability to operate in rivers, canals, and other low-gradient water systems make them ideal for decentralized energy production in underserved areas [3]. Coastal regions are especially suitable for adopting GWVP due to their easy access to water sources. These systems are particularly useful in such areas because they are scalable, easy to use, and can function efficiently in various water bodies,

including rivers and canals [4]. Coastal areas benefit from steady water flows from rivers and tidal streams, which makes them well-suited for GWVP installation. These regions often face energy shortages and rely heavily on diesel generators or unstable grid connections [5]. GWVPs can help address these issues by reducing dependence on fossil fuels, promoting local energy self-sufficiency, and providing a reliable, renewable, and eco-friendly energy source [6]. Additionally, their compact size and minimal land requirements allow them to be installed in different environments without significantly altering the natural landscape [7].

From an environmental perspective, GWVPs offer significant advantages. They enable unhindered fish migration and prevent water stagnation, which is common in dam-based systems, thereby helping to protect aquatic ecosystems [8]. Moreover, GWVPs produce no carbon emissions during operation, supporting global efforts to combat climate change and transition to renewable energy [9]. These features make GWVPs particularly attractive to developing countries, where the demand for sustainable energy solutions is growing. Despite their benefits, several challenges hinder the widespread adoption of GWVPs. Technical issues, such as managing sediment buildup in the vortex basin and improving turbine efficiency, need to be addressed to enhance performance and reliability [10]. Financial barriers are also significant, as the high upfront costs of turbines and infrastructure can be prohibitive for resource-limited communities. Additionally, the lack of legal frameworks and government incentives often limits large-scale implementation [6].

This study explores the potential of GWVPs for coastal communities by examining their technical feasibility, environmental impact, and economic benefits. By analyzing existing applications and critically evaluating the technology, the research aims to offer valuable insights into GWVPs as a sustainable energy solution.

Comparison between GWVPP and Other Hydropower Plants:

The ultra-low head of the GWVPP ranges between 0.7 and 3 meters, whereas traditional hydropower plants require a large head height of over 10 meters [5]. Unlike conventional plants that depend on dams, GWVPPs can operate without them, allowing factories and cities to construct water outlets instead [1]. Additionally, due to its ultra-low head, the initial setup cost of GWVPP is significantly lower compared to other hydropower facilities [3]. GWVPPs offer flexible design specifications, unlike other plants that demand more specific configurations [4]. While traditional plants often require large, remote areas far from cities, GWVPPs can be installed within city limits [11]. They also need fewer workers and require only moderately skilled staff, unlike conventional plants, which rely on highly experienced and qualified personnel [2]. Moreover, GWVPPs demonstrate high operational efficiency [8].

Methodology:

The research team conducted various field measurements, including assessing the width and topography of the canal and identifying existing structures and suitable locations for building the powerhouse. The proposed project site is located approximately 20 minutes from the University of Gujrat and 15 minutes from Gujrat City. To select the most suitable site, multiple surveys were carried out, and public input was gathered during the process. Community meetings were also held at different times to collect feedback. Additionally, the transmission and distribution routes for the water pipeline were measured to ensure proper planning.

Given the growing energy demand, it is crucial for the government of Pakistan to focus on such low-head, cost-effective hydropower solutions. The study emphasizes the need to raise government awareness about the potential of GWVHP systems to generate electricity at city water outlets. A detailed cost estimate and analysis were prepared for the proposed GWVHP system, considering local labor rates and the availability of construction materials. The plan takes into account both skilled and unskilled labor, ensuring feasibility.

Electromechanical components were priced according to market rates or installer prices to maintain cost-effectiveness. However, the cost of these components may vary depending on market fluctuations over time. Figure 2 illustrates the step-by-step sequence followed in the methodology.

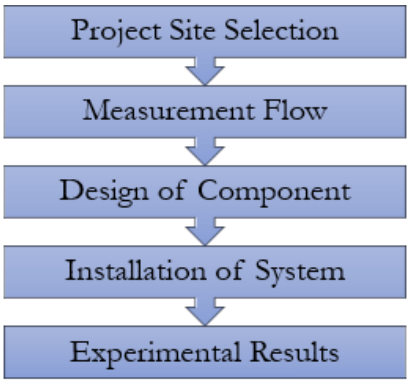


Figure 2. Flow Diagram of Methodology

The proposed project site is located approximately 12 km from the University of Gujrat, a distance that takes around 20 minutes by vehicle. It can also be easily accessed from Gujrat City via an earthen road. The site is situated near a wastewater outlet, where water exits from various points in Gujrat. Most of the city's wastewater flows through the Bollay Bridge stream, the oldest wastewater channel in the area. At Bollay Bridge, water from different city exit points converges, and a well has been drilled to collect the wastewater. A pump is installed in the well to discharge the collected water into the Bollay stream. The diameter of the discharge pipe was measured, and the flow velocity was recorded at 3.3 m/s. Based on these measurements, the design discharge rate was calculated as 61.46 liters per second. Accurate measurement of flow rates is essential for managing wastewater, as fluctuations can significantly impact treatment processes [12].

The proposed Gravitational Water Vortex Power Plant (GWVPP) will not disrupt nearby villages that have water rights, as it requires only a small volume of water. Additionally, during the growing season, the power plant will have no adverse effects on the irrigation system, ensuring that water usage remains unaffected. Since the GWVPP is designed to utilize wastewater as a source of electricity, it can be installed at city outlets. To estimate wastewater availability, data was collected from Gujrat City. The city has 149,428 households, with 87,189 located in urban areas [13] A survey of 50 households revealed that each household uses approximately 500 liters of water per day. The detailed data is summarized in Table 1.

Table 1. Average data for daily houses wastewater

Sr. No	Daily Waste Water	Liter/day
1	Washing and Bathing	150
2	Kitchen Sink	40-50
3	Toilet	50-60
4	Clothing Washing Machine	70
5	Floor Washing	40-50
6	Other uses	120

With an average household size of 4 to 5 people, the wastewater generation per capita per day may vary. This variation depends on several factors, including water conservation practices, household appliances, family size, and climate conditions. The Gravitational Water Vortex Power Plant (GWVPP) represents an innovative approach to energy generation, distinguished by its unique design and minimal civil construction requirements. Its structure and operational principles differ significantly from those of conventional micro-hydropower plants. The intake section plays a crucial role by providing a proper passage for water to flow

into the system. This section channels water from the input to the basin, with steel intake walls measuring 3 mm in thickness.

At the outflow of the intake system, a conical basin is connected. Conical basins are preferred over cylindrical ones due to their superior efficiency. The system utilizes water flowing through the intake channel, which enters the basin via a notch in the canal, rotates within the basin's system, and exits through the basin's bottom. Key design parameters for the basin include the canal's length, width, height, and notch dimensions, which must be carefully considered [14]. During system testing, high voltage output and high revolutions per minute (rpm) were observed at the center of the basin, indicating efficient energy generation[15]. This technology is particularly well-suited for applications such as ice-making factories and wastewater outlets in cities, where it can effectively utilize wastewater to produce electricity.

The system's efficiency can be enhanced by increasing the water head from 0.7 meters to 3 meters, which increases the flow rate and power output. Additionally, experiments conducted on peaches demonstrated that the system is suitable for use in food processing and fish farming. The symmetrical blade design ensures that no contact points exist that could harm fish, further enhancing its suitability for aquaculture applications.

Calculations:

Area and Discharge:

The cross-sectional area is the ratio between Discharge and velocity of the fluid.

$$A = Q/v. \quad (1)$$

After simplifications velocity obtained 3.3 m/s and calculated area is 0.018 m².

The following formula can be used to find the flow rate in m³/s:

$$Q = Av \quad (2)$$

Obtained water flows through the circular pipe at a rate of 61.46 L/s. The flow bay and basin were designed to have a flow rate (Q) is 0.06146 m³/s. It has been considered that the flow velocity (V) through the water fore Bay is 3.3 m/s. Let the width and height of the passage be according to design; it will prevent the overflow of the water, and we take a factor of safety as 2.

The height of the canal can be calculated from

$$H = \text{Width} * \text{Safety Factor} \quad (3)$$

Governing Equations:

In this experiment with conical basin and single stage turbine used with the discharge of 61,46 L/s as constants in all variations.

To find the best performance to rotor basin and turbine blade shape.

$$\omega = \frac{2\pi N}{60} \quad (4)$$

ω is the rotational speed of the turbine and it can measure by tachometer apparatus to check the performance based on the load.

$$F = mg \quad (5)$$

The load value acts as a braking torque, meaning the measured load on the load cell represents the total mass (m) of the turbine. The torque (T) on the turbine is determined using the force measured by the load cell, which is generated by the rope brake system. This system produces the total mass acting on the turbine, allowing for an accurate calculation of the applied force. The power output of the system can be expressed as:

$$P_{in} = \rho g Q H \quad (6)$$

Water flows into the basin with specific discharge formed vortex by rotationally movement at certain height ρ is fluid density, g is the gravitational acceleration, Q is the flow rate, and H is the effective water head.

$$P_m = T\omega \quad (7)$$

The mechanical power of the turbine (Pm) is the power generated by the rotation of the shaft, which depends on the applied torque and rotational speed. In this context, **T** represents the torque (N·m), and **ω** denotes the angular velocity (rad/s). The basin, shown in *Figure 3*, consists of a rotating system installed in areas that are not prone to flooding and can be managed throughout the year. The structure of the rotational tank system can be constructed using concrete. The dimensions of the concrete basin are specifically designed to meet the hydropower system's head drop and flow rate requirements. The basin of the Gravitational Water Vortex Power Plant (GWVPP) is appropriately sized to ensure both an optimal head drop and efficient vortex flow into the hydropower plant. Additionally, the rotational system is designed to be waterproof, which protects it from water pressure and the impact load of water entering the basin. This feature enhances its durability and effectiveness in practical applications.

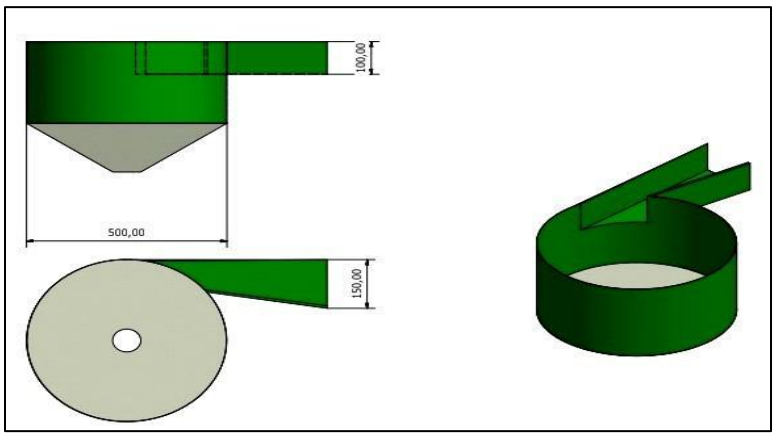


Figure 3. Basin Design

The proposed conical design, shown in *Figure 4*, forms the basis of the basin's structure due to its higher efficiency [7]. The top of the basin has a diameter of 400 mm. According to previous research, the optimal cylindrical basin diameter for maximizing power production should be between 14% and 18% of the system's key parameters [16]. The bottom diameter of the basin is 60 mm, while the height of its conical section is 225 mm. The total height of the basin is 400 mm. Based on these findings; the conical basin was designed to enhance performance. The cylindrical portion of the basin has a diameter of 500 mm, with an exit hole diameter of 60 mm. Additionally, our research suggests that the optimal cone angle for maximum efficiency is within a specific degree range, based on experimental analysis.

Mechanical Components:

This technology requires fewer mechanical components compared to conventional systems, such as screw turbines, contributing to its simplified design and lower maintenance needs.

Figure 4. Gravitational water vortex Model

Flow Regulating Gates: The structure includes two proposed gates. One gate is located at the turbine's intake section to regulate water flow during maintenance. The second gate controls the bypass flow in the stream.

Turbine Runner: The turbine runner (Figure 5) was designed as part of the project conducted at the University of Gujrat, Hafiz Hayat campus, in the Department of Mechanical Engineering. The design was based on the following parameters:

- Inner and outer diameter of the basin
- Height
- Inlet and outlet blade angles
- Tapered angle
- Number of blades: 4
- Impact angle

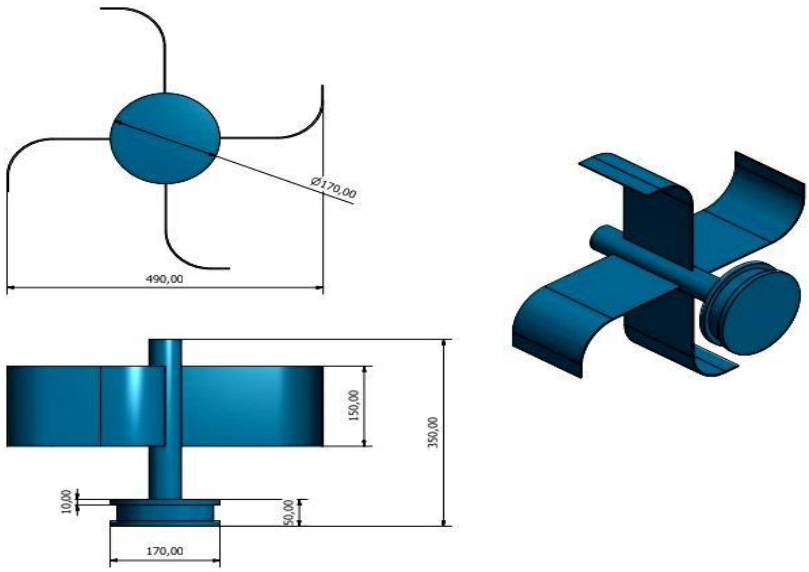


Figure 5. Turbine Runner

Material and Shape of Turbine Blades:

1. Steel (Curved)
2. Cast Iron (Curved)

Specifications of the Proposed Turbine:

- **Type:** Gravitational Water Vortex
- **Motor-to-Turbine Speed Ratio:** 1:25 (RPM)

Electrical Components:

Power Transmission and Drive System:

Mechanical energy is transferred from the turbine shaft to the generator's rotor through a rope drive system, as shown in Figure 6. The rated speeds for the DC motor and turbine are 1440 rpm and 60 rpm, respectively, requiring a speed ratio of 1:24. To achieve this, pulleys and a planetary rope system are recommended. Additionally, a self-lubricating mechanism is used to reduce wear and prevent fatigue.

Control and Protection System: Overvoltage, undervoltage, over frequency, underfrequency, and overcurrent relays are used to protect the system from unfavorable conditions. Before powering on the system, all electrical components must be thoroughly inspected.

Figure 6. Rope Drive System

Transmission and Distribution: The distance between the load and the power source determines the length of the distribution lines. The gravitational water vortex power plant generates 9.5 V, which is monitored using a voltmeter. The distribution details are presented in Table 2.

Table 2. Generator Specifications

Transmission/Distribution	Capacity
Motor	12 V
Permanent Magnate Alternator	200W to 300W at 60 RPM
Capacitor	25 microfarads
Voltmeter	16 V
Connector 2 LED	length 3 m

Power calculation formula:

$$Power_{in} = V\ x\ I \quad (8)$$

12V-DC Motor: The motor's output power can be evaluated by supplying input to the system. A 12V motor with a 2A current (producing 24W) was used, and an effective output power of 18W was obtained, resulting in a motor efficiency of 75%. Due to heat and friction, the motor experiences a power loss of 6W. The turbine's efficiency was measured at 5.59%. Table 3 provides a breakdown of the 6W power loss across various stages of the Gravitational Water Vortex Power Plant (GWVPP). These losses occur due to factors such as hydraulic inefficiencies, mechanical friction, electrical resistance, and transmission losses, as detailed in Table 3.

Permanent Magnetic Alternator:

In our experimental model, the turbine pulley had a diameter of 12 inches, while the motor pulley measured 0.5 inches. This created a speed ratio of 1:24, indicating a significant speed mismatch, which led to power transmission losses. As shown in Table 2, the permanent magnet alternator produced between 160W and 200W, demonstrating that the efficiency of the Vortex Power Plant increased from 5.59% to 55.19%.

Estimation of Work Volume:

The estimated amounts were based on the selected site area. A low work volume was used since this project was designed for demonstration purposes.

Results and Discussions:

The primary focus of this power plant is its potential implementation in urban areas, with Gujrat city selected as a case study. However, performance evaluations are also necessary to analyze the plant's overall efficiency. Hydraulic efficiency was measured to assess the Vortex Power Plant's performance, and graphs were used to show the relationship between torque

and hydraulic head. The hydraulic head represents the available power at different flow rates, while torque at various rotational speeds reflects the extracted power.

During the experiment, it was observed that a lower head of 0.7 meters resulted in higher efficiency compared to a 3-meter head, as mechanical losses were proportionally smaller at lower head levels. Fluid velocity was also examined using three different types of pipes: cast iron, steel grade XS, and steel grade SSX. Each pipe had a diameter of 2 inches, and the discharge rate (Q) was 61.47 L/s. It was found that velocity was higher in steel pipes of types Schedule 40 and Schedule 80.

Table 3. Breakdown of Power Losses in each stage

Stage	Cause of Loss	Estimated Loss (W)
Hydraulic Losses	Energy dissipation due to turbulence in water vortex. Frictional losses in water flow through the conical basin and inlet.	1.2W
Turbine Efficiency Losses	Imperfect blade angle reducing energy conservation. Friction between water and turbine blades.	1.0W
Mechanical Friction Losses	Friction between turbine shaft and bearing. Energy losses due to misalignment or wear in moving parts.	0.8W
Generator Losses	Heat dissipation in permanent magnet alternator (PMA) windings. Resistance in electrical windings reducing output efficiency.	1.2W
Transmission Losses	Power loss due to speed mismatch between turbine and generator. Losses in belt/pulley system transferring mechanical energy.	1.0W
Electrical Losses	Initial resistance in 12V DC motor and wiring. Power loss due to impedance in electrical components.	0.8W

Turbine efficiency calculated as:

$$\frac{P_{motor}}{Turbine_{provided}} \; x\; 100 \quad (9)$$

Torque vs. Speed (12V - DC Motor):

The turbine supplied 24W of power, resulting in a useful power output of 18W at a motor speed of 60 RPM. The calculated turbine torque was 2.87 Nm. Figure 7 illustrates the inverse relationship between torque and speed: as speed increases, torque decreases at constant power. Additionally, Figure 8 shows that DC motor efficiency was notably high at 60 RPM.

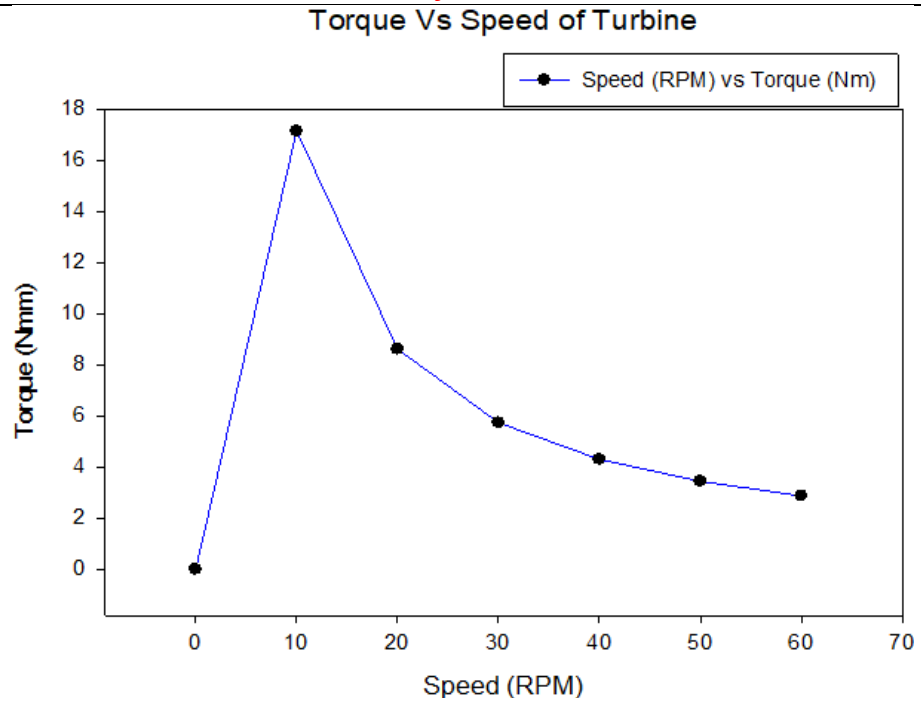


Figure 7. Torque Vs Speed

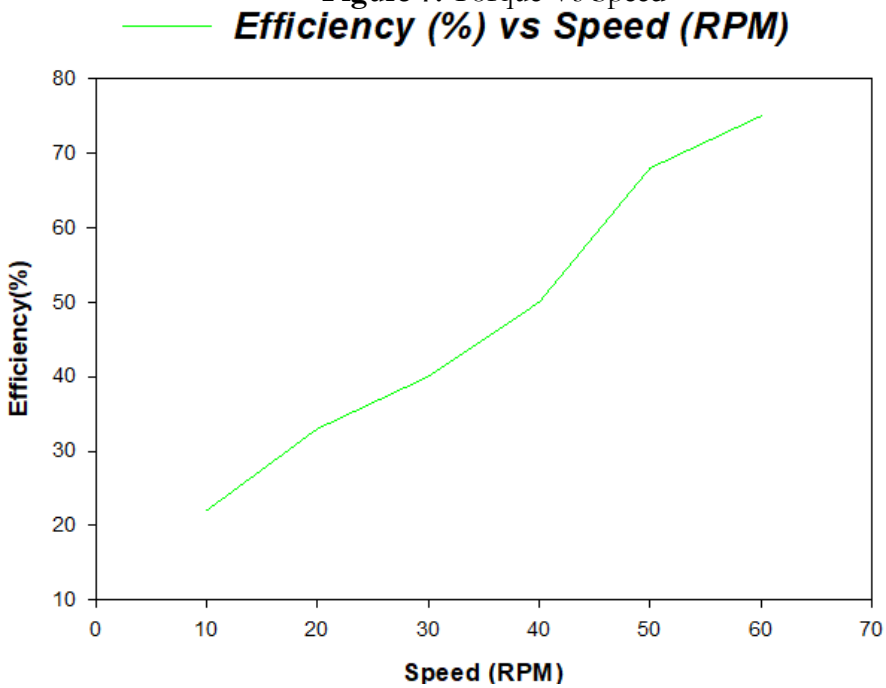


Figure 8. Motor Efficiency vs Speed (RPM)

Torque vs. Power (Permanent Magnet Alternator):

In the experimental model, the turbine's efficiency using a 12V DC motor was measured at 5.69%, which was relatively low due to the speed mismatch between the turbine and the generator. To improve power output, a permanent magnet alternator (PMA) was introduced, which generated between 160W and 200W. Given that the turbine supplied 422.11W of power, the use of the PMA resulted in an output efficiency of 49.5%. The torque produced by the PMA ranged from 25.46 Nm to 31.83 Nm, as shown in Figure 9. The vortex power plant achieved an optimal efficiency of 55.19%, as illustrated in Figure 10.

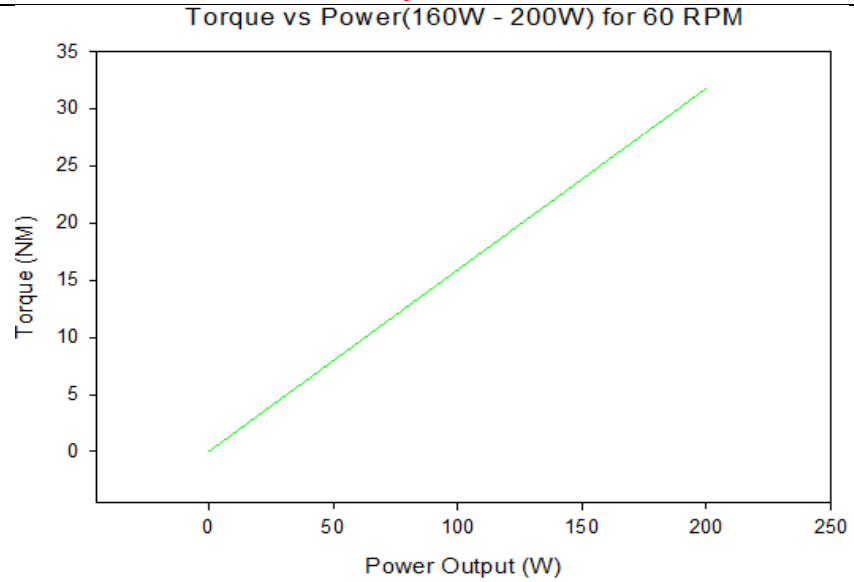


Figure 9. Torque vs Power

Power vs Efficiency:

Figure 10 presents the output power based on the experimental model of the vortex power plant. The input and output power were calculated, resulting in a power efficiency of 75%. This indicates that the vortex power plant is quite efficient in converting electrical power into useful output. The voltmeter displayed an output of 9.5 volts from a 12-volt DC motor. The 2W power loss may be due to internal losses in the motor, circuit impedance, or minor inefficiencies in electrical components. To improve motor performance, a 25µF capacitor was added, which helps stabilize the voltage and reduce fluctuations. Figure 11 displays the output voltages measured by the voltmeter during the experiment.

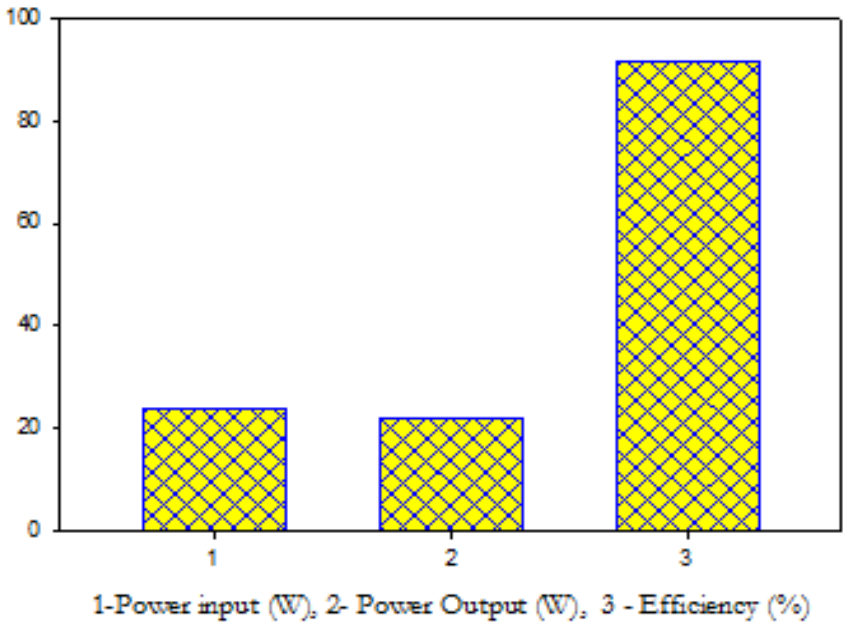


Figure 10. Power vs Efficiency

Figure 11. Voltage measuring

Discharge vs Velocity:

As shown in Figure 12, each 2-inch diameter pipe maintains a constant discharge. The velocity recorded for the steel pipe (Schedule 40 to 60 type) is 3.9 m/s, while the cast iron pipe has a velocity of 3.3 m/s. The Y-axis represents a discharge value of 61.46 l/s.

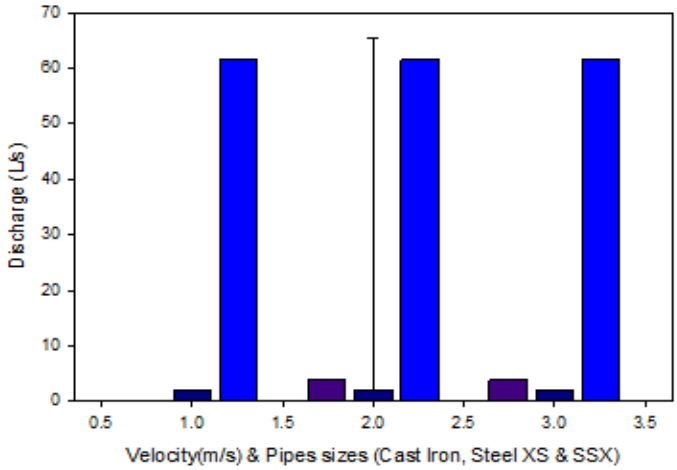


Figure 12. Discharge vs velocity

Typical Day Production:

The highest discharge and maximum efficiency were recorded at 5 PM, as shown in Figure 13.

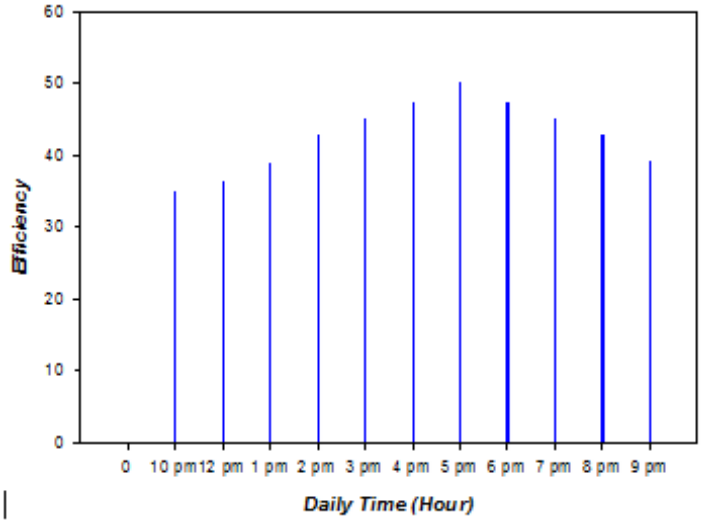


Figure 13. Efficiency % during the typical day

Discussion:

The experimental model of the Gravitational Water Vortex Power Plant (GWVPP) demonstrated an efficiency of 55.19% when using a Permanent Magnet Alternator (PMA), aligning with efficiency rates reported in similar vortex power plant studies. [2] and [3] documented efficiencies ranging from 50-60%, while [7] emphasized the role of optimized basin geometry and turbine design in enhancing performance. A key finding of this study was that lower head heights (0.7 meters) produced higher efficiency compared to head configurations of 3 meters. This aligns with [17], who attributed improved performance in low-head systems to reduced mechanical losses. [8] suggested that efficiency could be further enhanced by incorporating multi-stage turbines.

The study also found that achieving an arc angle of 90° and a blade angle of 50° could further boost efficiency. The power output recorded (160–200W) was consistent with findings from earlier studies. For instance, [4] reported outputs ranging from 200 to 300W, suggesting that increasing discharge rates and improving turbine materials could enhance power generation. Similarly, [5] found that using steel turbine blades resulted in outputs of 150–250W. [8] also highlighted that the lower initial power output (24W) observed in this case was due to a mismatch between turbine and generator speeds. This issue was mitigated by employing a Permanent Magnet Alternator (PMA), which significantly improved performance. Further improvements could be achieved by using a variable-speed generator or a more efficient gearbox. Additionally, adjusting the turbine runner position for optimal water flow could enhance output.

The study confirmed that conical basins are more efficient than cylindrical ones, supporting findings by [16]. The recommended cone angle in this study aligns with previous research, which identified a 60° cone angle as the most effective.

Moreover, the study demonstrated the feasibility of generating electricity using municipal wastewater. Previous research by [1] and [12] also identified wastewater as a potential energy source, though they noted that performance may be affected by variations in discharge rates. [10] observed that sedimentation in wastewater can reduce system efficiency. To maintain consistent energy generation in wastewater streams with fluctuating flow rates, adaptive intake systems are necessary. Additionally, energy extraction could be enhanced by increasing the vortex's rotational speed, potentially through a double-stage vortex system.

Conclusion:

- The primary focus of this thesis research is on hydraulic phenomena and their applications in engineering, particularly in micro hydropower plants serving coastal communities. Gravitational Water Vortex Power Plants (GWVPPs) have attracted significant global attention due to their environmentally friendly nature. Based on the research findings, recommendations can be made on selecting the most efficient turbine system for various micro hydropower projects.
- We concluded that Gujrat and other resource-rich towns could harness wastewater to generate sufficient power. Additionally, we found that this power plant can be installed as a compound plant and offers environmental sustainability. However, the key challenges to its development in different regions of Pakistan include political interference and competing agendas. Governmental instability and a lack of interest in adopting new technologies further hinder progress in this area.
- This type of power plant is highly effective for irrigation systems and is remarkably simple to construct and install. In our experimental setup, the gravitational water vortex power plant successfully generated 12 volts using a 12-volt DC motor. Efficiency could be further improved by optimizing the blade angle, refining the basin shape, and selecting an appropriate generator.

- Several opportunities exist to modify the plant's design to generate additional mechanical energy, which can then be converted into electrical energy using specific tools and equipment.
- To achieve higher efficiency, the inlet blade angle should ideally be around 50°, and the blade arc angle should be set at 90°.
- Our findings suggest that if the channel is parallel to the water surface, lower efficiency is observed. By adjusting the channel's orientation and design, we achieved significantly improved efficiency. Additionally, enhancing the blade design and optimizing their positioning can increase the rotational speed of the turbine blades, thereby enhancing overall performance.

Limitations:

This study is based on a small-scale experimental model, which may not fully represent real-world conditions in larger applications. It focuses on specific flow rates derived from Gujrat's wastewater data, which may not be applicable to all coastal communities. Additionally, the proposed efficiency improvements depend on certain materials, such as steel pipes, which may not be readily available in all regions. The turbine efficiency, while promising, remains lower than that of conventional hydropower systems due to speed mismatches, highlighting the need for further optimization. The study also does not address potential legal, environmental, and regulatory challenges associated with large-scale deployment. Although cost analysis is discussed, actual implementation costs may vary considerably depending on local infrastructure and government incentives. While the system is well-suited for small-scale applications, large-scale adoption of GWVPP may require significant modifications to existing water management systems.

Acknowledgement:

All authors have contributed equally to the conception, design, analysis, and writing of this manuscript. The authors declare no conflicts of interest regarding the publication of this paper.

References:

[1] Gonzalez, J. et al, "Policy frameworks for emerging renewable technologies," *Energy Policy*, vol. 39, no. 4, pp. 120–133, 2021.

[2] Müller, R. et al, "Ecological benefits of low-head hydropower systems," *Environ. Sci. Technol.*, vol. 52, no. 3, pp. 215–230, 2020.

[3] A. M. N. Sagar Dhakal, Ashesh B. Timilsina, Rabin Dhakal, Dinesh Fuyal, Tri R. Bajracharya, Hari P. Pandit, Nagendra Amatya, "Comparison of cylindrical and conical basins with optimum position of runner: Gravitational water vortex power plant," *Renew. Sustain. Energy Rev.*, vol. 48, pp. 662–669, 2015, doi: https://doi.org/10.1016/j.rser.2015.04.030.

[4] Iskandar, S. et al, "Implementation of gravitational vortex plants in Indonesian coastal regions," *J. Renew. Energy*, vol. 56, no. 2, pp. 78–89, 2020.

[5] Ahmed, S. et al, "Renewable energy potential in coastal environments," *Renew. Energy J.*, vol. 64, no. 3, pp. 45–57, 2022.

[6] Gonzalez, J. et al, "Policy frameworks for emerging renewable technologies," *Energy Policy*, vol. 39, no. 4, pp. 120–133, 2021.

[7] A. Sedai, G. Singh, R. Dhakal, B. B. Kumal, N. Ghimire, and B. K. Yadav, "Technical and Economic Analysis of Site Implementations of Gravitational Water Vortex Power Plant," *Proc. - 2021 IEEE Int. Conf. Intell. Syst. Smart Green Technol. ICISSGT 2021*, pp. 128–133, 2021, doi: 10.1109/ICISSGT52025.2021.00036.

[8] Ravi, P. et al, "Integration of gravitational vortex systems in hybrid renewable energy frameworks," *J. Energy Syst.*, vol. 45, no. 2, pp. 120–130, 2022.

[9] M. a O. B. Yaakob, "A Review on Micro Hydro Gravitational Vortex Power and

Turbine Systems," *J. Technol.*, 2014, [Online]. Available: https://www.slideshare.net/slideshow/a-review-on-micro-hydro-gravitational-vortex-power-and-turbinepdf/262095086

[10] J. J. Paweł Tomczyk, Krzysztof Mastalerek, Mirosław Wiatkowski, Zeshan Asghar, "Assessment of a Francis Micro Hydro Turbine Performance Installed in a Wastewater Treatment Plant," *Energies*, vol. 16, no. 20, p. 7214, 2023, doi: https://doi.org/10.3390/en16207214.

[11] C. Martinez, "Community-driven renewable energy projects in Southeast Asia," *Energy Dev.*, vol. 8, no. 1, pp. 89–101, 2022.

[12] C. Xu, "Study of wave interaction with vertical piles integrated with oscillating water columns," *Univ. Hawai'i Mānoa*, 2018, [Online]. Available: https://core.ac.uk/download/211329122.pdf

[13] GUJRAT DISTRICT, "https://www.pbs.gov.pk/census-2017-district-wise/results/048," *Pakistan Bur. Stat.*, 2023.

[14] HAMID BASHIRI & ANGELOS FINDIKAKIS, "Hydrolink 2020/2. Hydraulic Transients," *Int. Assoc. Hydro-Environment Eng. andResearch*, 2020, [Online]. Available: https://core.ac.uk/reader/535022280

[15] S. M. Mosè Rossi, Massimiliano Renzi, David Štefan, "Small-Scale Hydropower and Energy Recovery Interventions: Management, Optimization Processes and Hydraulic Machines Applications," *MDPI AG*, 2022, doi: https://doi.org/10.3390/books978-3-0365-5553-9.

[16] J. A. Chattha, T. A. Cheema, and N. H. Khan, "Numerical investigation of basin geometries for vortex generation in a gravitational water vortex power plant," *2017 8th Int. Renew. Energy Congr. IREC 2017*, May 2017, doi: 10.1109/IREC.2017.7926028.

[17] N. Dhakal, Sagar & Timilsina, Ashesh & Dhakal, Rabin & Fuyal, Dinesh & Bajracharya, Tri & Pandit, Hari & Amatya, "Mathematical modeling, design optimization and experimental verification of conical basin," Gravitational water vortex power plant. Accessed: Apr. 07, 2025. [Online]. Available: https://www.studocu.com/ph/document/wesleyan-university-philippines/research-methods/mathematical-modeling-design-optimization-and-experimental-verification-of-conical-basin-gravitational-water-vortex-power-plant/23168539

Performance Analysis of Motorbike Engine Using Bioethanol Gasoline Blends

Yasir Hussain[1], Sajjad Miran[1], Nazam Siddique[2], Samar ul Hasnain[1], Hafiz Ahmad Bilal[2], Waseem Arif[1]
[1]Department of Mechanical Engineering, University of Gujrat, Pakistan
[1]Department of Electrical Engineering, University of Gujrat, Pakistan.
[2]Govt. Swedish Pakistani College of Technology, Gujrat, Pakistan
***Correspondence:** sajjad.miran@uog.edu.pk.

Citation | Hussain. Y, Miran. S, Siddique. N, Husnain. S. U, Bilal. H. A, Arif. W, "Performance Analysis of Motorbike Engine Using Bioethanol Gasoline Blends", IJIST, Special Issue. pp 236-244, March 2025
Received | Feb 24, 2025 **Revised |** March 08, 2025 **Accepted |** March 14, 2025 **Published |** March 17, 2025.

The increasing demand for sustainable energy and reduced reliance on fossil fuels has driven the exploration of alternative fuel options. This study aims to evaluate the performance of a motorcycle engine using bioethanol-gasoline blends. Simulations were conducted using AVL Boost software. By applying AVL Boost in innovative ways, the research provides new insights into improving the performance of motorcycle engines powered by bioethanol-gasoline blends, contributing to more eco-friendly transportation. A numerical model of a single-cylinder engine was developed, and various fuel blends were tested to assess performance characteristics at engine speeds ranging from 1000 to 4000 RPM. Single-cylinder spark ignition engines are commonly used in many types of motorcycles. The results showed that the E20 blend achieved a 4% increase in power and improved performance characteristics during tests on engines running on lower ethanol blends.

Keywords: Biofuels, Bioethanol, AVL Boost, Numerical Investigation, Engine Performance.

Introduction:

The rising daily demand for energy has encouraged researchers to explore alternative, non-fossil vehicle fuels. With fossil fuel reserves depleting, finding widely accepted alternatives has become increasingly urgent. Biofuels, derived from renewable resources, have emerged as a promising substitute for conventional fossil fuels due to their limited availability [1][2]. Recognized as a sustainable energy source, biofuels play a crucial role in the renewable energy sector [3][4]. Among biofuels, alcohol is particularly significant [5]. Bioethanol, in particular, stands out as a potential alternative fuel for gasoline engines because it shares several key physical and chemical properties with gasoline, as shown in Table 2 [6][7]. Compared to conventional gasoline, bioethanol-gasoline blends have been found to significantly lower carbon monoxide (CO), hydrocarbon (HC), and particulate matter (PM) emissions while offering mixed effects on nitrogen oxides (NOx) and carbon dioxide (CO_2) emissions. This research uses a one-dimensional simulation model developed with AVL Boost [8] to analyze the performance characteristics of a single-cylinder spark ignition gasoline motorcycle engine fueled by gasoline-bioethanol blends. Several parameters affecting engine performance are examined, including power, torque, fuel consumption, and efficiency. The objective is to assess how the selected fuel blends—E0, E20, and E40—impact engine performance at different speeds: 1000, 2000, 3000, and 4000 RPM. Ethanol-gasoline blends with ethanol content of 50% or higher (E50+) have demonstrated significant potential for improving engine performance and reducing emissions. Research has shown that E50 blends can enhance brake thermal efficiency, with studies indicating that a compression ratio of 9.1 results in maximum power output, while a higher compression ratio of 9.7 leads to minimal fuel consumption. Additionally, E50 blends have been found to increase brake torque and power in spark-ignition (SI) engines, making them a viable alternative to conventional fuels.

In terms of emissions, E50 blends have been associated with significant reductions in carbon monoxide (CO) and hydrocarbon (HC) emissions compared to pure gasoline. Furthermore, nitrogen oxide (NOx) emissions also decrease when using E50 blends, contributing to a cleaner and more sustainable combustion process. These improvements are largely attributed to ethanol's higher oxygen content, which facilitates more complete combustion.

However, certain operational considerations must be taken into account when using E50 blends. Due to ethanol's lower energy density compared to gasoline, fuel consumption may increase slightly. Nonetheless, this effect can be offset by the overall gains in thermal efficiency. Another critical factor is material compatibility, as higher ethanol concentrations can affect engine components, necessitating further research into long-term durability. Future studies on E50+ blends should focus on optimizing engine parameters, assessing material compatibility, and conducting real-world driving condition evaluations to fully leverage the benefits of higher ethanol content fuels.

This knowledge could contribute to the development of improved engines and higher-quality vehicular fuels. There is growing awareness of biofuels, such as bioethanol, as a means to reduce environmental impacts and dependence on fossil fuels [9]. As the automotive industry searches for more sustainable alternatives, understanding how bioethanol fuel blends affect engine performance becomes both essential and challenging [10]. Studying the impact of bioethanol-gasoline blends offers valuable insights that can guide the design of engines optimized for these alternative fuels. By incorporating bioethanol, the reliance on limited fossil fuel reserves can be reduced [11].

This report provides several recommendations to help the automotive industry transition toward more eco-friendly and sustainable energy sources. However, biofuel blending comes with certain challenges. While some blends are compatible with existing engines, others are not [12]. Therefore, it is crucial to determine which ethanol-gasoline blend

is best suited for single-cylinder spark ignition engines to achieve optimal performance. Ultimately, understanding the performance characteristics of biofuels and their compatibility with gasoline is key to successful engine integration.

Objectives of the Study:

The main objectives of this study on the performance analysis of a motorbike engine using bioethanol-gasoline blends are as follows:

- Analyze the impact of ethanol-gasoline blends (E0, E20, E40) on power, torque, fuel consumption, and efficiency using AVL Boost simulations.
- Utilize AVL Boost software to create and validate a simulation model for a single-cylinder motorcycle engine.
- Investigate ethanol's potential to enhance engine performance and reduce reliance on fossil fuels.
- Identify the most efficient ethanol-gasoline ratio for improved engine performance while maintaining fuel economy.

Material and Methods:

This study uses the AVL Boost simulation tool to develop a simulation model for a single-cylinder motorcycle engine. The model is validated using experimental data to ensure its accuracy and alignment with current experimental findings [13]. After validation, the motorcycle engine model, based on the HONDA CD 70 with a 72cc single-cylinder spark ignition engine, is simulated on AVL Boost to analyze performance characteristics, including exhaust gas temperature data for each ethanol-gasoline blend at engine speeds of 1000, 2000, 3000, and 4000 RPM.

AVL Boost is specifically designed for engine simulations, enabling the analysis of changes in fuel compositions, engine configurations, and related components. It provides highly accurate predictions for both intake and exhaust parameters. In modern engineering, simulations are highly valued due to their ability to save time and costs compared to physical experiments, while also optimizing experimental designs.

This simulation investigates motorcycle engine performance using bioethanol-gasoline blends at different speeds. In Pakistan, 70cc motorcycles are widely used and primarily run on gasoline. These bikes typically use single-cylinder spark ignition engines equipped with carburetor-based fuel systems, which ensure the precise delivery of air and fuel to the combustion chamber for optimal engine performance.

The engine parameters used in this study are sourced from the HONDA CD 70 technical specification manual, an authoritative and reliable reference for accurate data. The study's methodology is illustrated in the flow diagram shown in Figure 1, while the parameters used to develop the simulation model are summarized in Table 1.

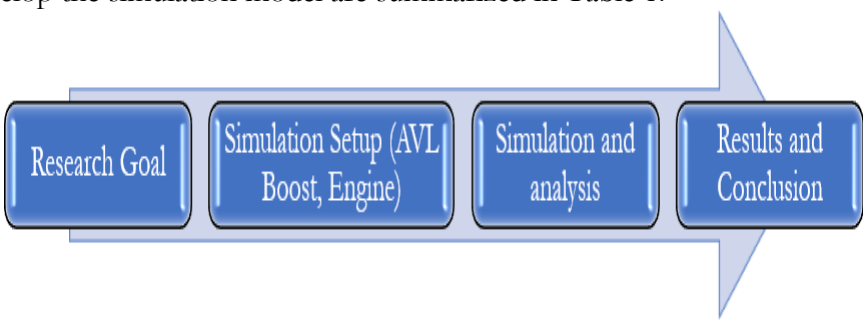


Figure 1. Flow Diagram of Methodology

Table 1: Single Cylinder Four Stroke Engine Parameters

Parameters	Specification	Unit
Engine	S I Engine 4 Stroke Air Cooled	
Make & Model	HONDA CD 70 2024	
Number of Cylinders	1	
Bore	47	mm
Stroke	41.4	mm
Connecting Rod	91	mm
Displacement Vale	72	cm3
BMEP Controlled	3	Bar
Compression Ratio	9.3: 1	
Aspiration Type	Natural	
Cooling System	Air Cooled	

The engine specifications include a bore of 47 mm, a stroke length of 41.4 mm, a compression ratio of 9.3, and a connecting rod length of 91 mm. The single-cylinder spark ignition motorcycle engine model was developed and calibrated using AVL Boost. The physical and thermal properties of gasoline and bioethanol are provided in Table 2.

Table 2: Properties of Gasoline and Bioethanol

Fuel Properties	Gasoline	Bioethanol
Molecular formula	C8H15.6	C2H6O
Density at 15 ◦C (kg/m3)	720 – 775	792
Boiling point at 1.013 bar (◦C)	25 – 210	78.4
Research Octane number (RON)	85	89.7
Motor Octane Number (MON)	95	108.6
Heat of vaporization (kJ/kg)	289	854
Energy density (MJ/kg)	45	26
Composition C/H/O (%mass)	87.4/12.6/0	52.18/13.04/34.7
Molecular weight (kg/km)	98	46.070

The system boundary (SB1) was set with a temperature of 23.85°C and a pressure of 1 bar. For the system boundary (SB2), the maximum temperature was set at 126.85°C, also at a pressure of 1 bar, with a maximum air-fuel ratio of 13. The engine models, using AVL Boost's Classic Species Setup, were tested with both gasoline and ethanol. This setup helps simulate combustion behavior and analyze the processes occurring during combustion, allowing a detailed investigation of engine performance features.

The engine model characteristics, including the Vibe 2 Zone combustion model, were developed based on experimental data on gasoline performance. Previous studies with similar engine geometries provided vibration parameters used for ethanol fuel simulations. During these investigations, ignition and combustion timing for various gasoline-ethanol blends were kept constant.

The process of examining engine characteristics and applying bioethanol-gasoline blends involves several steps. The initial step includes gathering relevant engine data and specifications, followed by testing bioethanol-gasoline blends using the AVL Boost simulation model. In this model, gasoline has a molar mass of 111.20875 kg/mol and a stoichiometric air-fuel ratio of 14.600477%. Ethanol, in comparison, has a molar mass of 46.06952 kg/mol and a stoichiometric air-fuel ratio of 8.998616%.

The indicated engine work [13], calculated using governing equations, can be used to derive the engine power equation, as illustrated in the following equation.

$$\boldsymbol{Wi = PmiVH = \int_0^{720} \frac{PdV}{1000}} \quad (1)$$

$$Ni = \frac{Wi \times n}{60 \times 2} \quad (2)$$

In the equation, **Wi** represents the Indicated Engine Power, measured in kilojoules (kJ), and **Pmi** denotes the Mean Indicated Pressure, measured in pascals (Pa). **VH** is the Cylinder Displacement, expressed in cubic meters (m^3), while **P** and **V** refer to the In-Cylinder Pressure (Pa) and Cylinder Volume (m^3), respectively. Additionally, **Ni** represents the Indicated Engine Power in kilowatts (kW), and **n** stands for the engine speed, measured in revolutions per minute (RPM). The given indicated power calculates torque;

$$Mi = \frac{60}{2\pi n} Ni \quad (3)$$

In this context, Ni refers to the Indicated Engine Power, measured in kilowatts (kW), while n represents the Engine Speed, measured in revolutions per minute (RPM).

AVL Boost simulations have become an essential tool for research focused on improving engine efficiency while reducing both the time and cost associated with engine model development. The software provides a wide range of programs related to Internal Combustion Engines (ICE). While much of the existing research is centered on spark ignition (SI) and compression ignition (CI) engines, some studies have expanded beyond these areas.

AVL Boost enables users to perform fully computational analyses and integrate experimental data with simulation results. Its versatility in handling various fuel types allows researchers to perform advanced calculations, including simulations involving alternative fuels. Fuel blends can be configured using the Classic Species Setup within the Boost gas properties tool, enabling users to adjust fuel blend ratios. The software calculates gas properties for each cell at every time step, ensuring accuracy. The fuel blends, defined by fraction ratios in the Boost gas properties tool, are listed in Table 3.

Table 2: Gasoline and Ethanol blending ratio

Blend	Gasoline%	Ethanol%
E0	100	0
E20	80	20
E40	60	40

Result and Discussion:

Validation of the Model:

The experimental study data was validated by comparing it with exhaust gas temperature data collected from the current experimental engine test. At 4000 RPM, the maximum recorded exhaust gas temperature reached 510°C. The observed error in exhaust gas temperature was highest at 2% for elevated temperatures, with an average error of 10%, as shown in Figure 2.

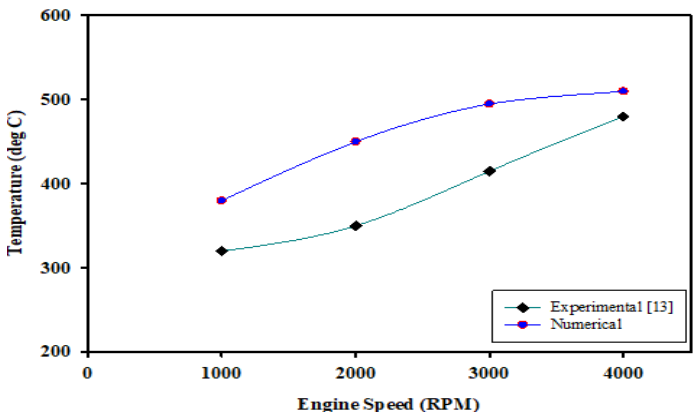


Figure 2: Engine Speed (RPM) vs Exhaust Gas Temperature

The maximum torque of 3.21 Nm was observed at 2000 RPM, representing a 16.52% increase compared to E100. At 4000 RPM, the torque performance is shown in Figure 3.

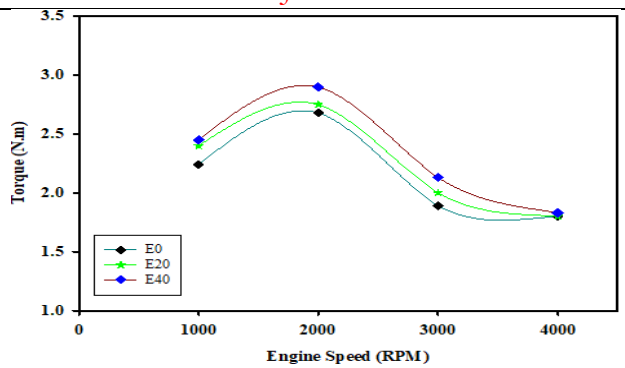


Figure 3: Engine Speed (RPM) vs Engine Torque

The engine's observed power output for gasoline and ethanol blends, measured in horsepower, is illustrated in Figure 4. The results indicate that power increases with higher ethanol content in the blends. However, at a lower speed of 1000 RPM, only the E20 blend shows a significant 52.31% increase in power compared to gasoline, making it a promising choice. This is due to ethanol's higher volatility, which allows it to expand and combust more efficiently in the E20 blend, thereby producing maximum power at 1000 RPM. Additionally, at lower RPM, the extended combustion time enhances power generation. The superior performance of the E20 blend at low speeds is attributed to its optimal volatility, which leads to highly efficient combustion.

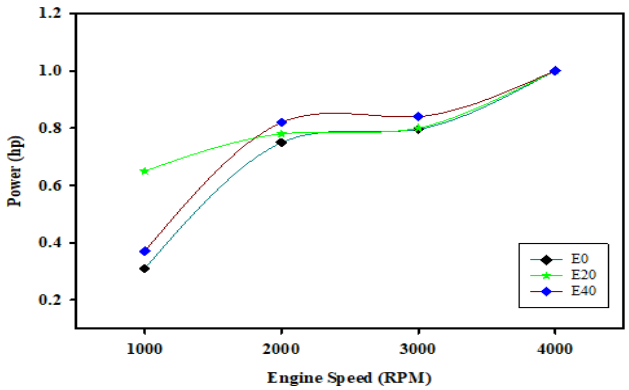


Figure 4: Engine Speed (RPM) vs Engine Power

Fuel consumption, measured in kilograms per hour (kg/h), represents the rate at which an engine uses fuel, expressed as the mass consumed over time, as shown in Figure 5. Ethanol contains 33% less energy per gallon compared to gasoline due to its higher octane rating. Additionally, ethanol's faster vaporization contributes to increased fuel consumption.

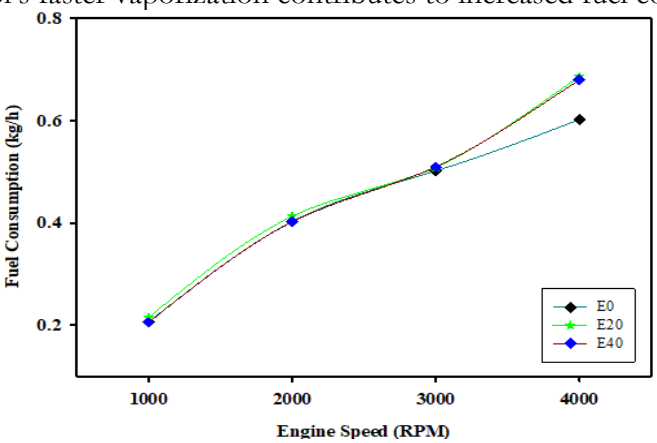


Figure 5: Engine Speed (RPM) vs Engine Fuel Consumption (Kg/h)

For small engines (< 100 cc), fuel efficiency varies between 25% and 35%, with engine load and friction conditions influencing the indicated efficiency. For E40, the maximum

indicated efficiency is observed at 4000 RPM, as shown in Figure 6. Ethanol improves engine efficiency due to its higher-octane rating, which reduces engine knock and enhances compression ratios. While pure gasoline achieves maximum efficiency of around 15%, ethanol blends demonstrate a 10% to 15% improvement over gasoline at various RPMs.

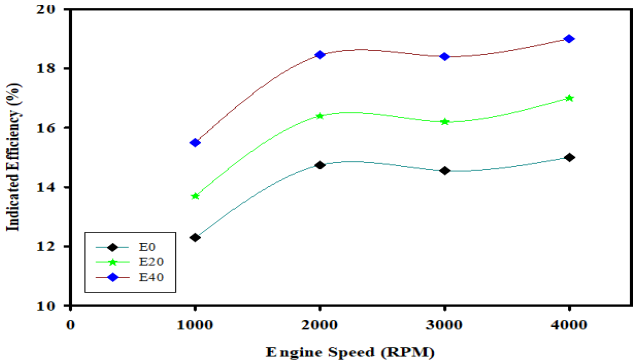


Figure 6 Engine Speed (RPM) vs Engine Indicated Efficiency

The effect of bioethanol-gasoline blends on engine mechanical efficiency at varying RPM is illustrated in Figure 7.

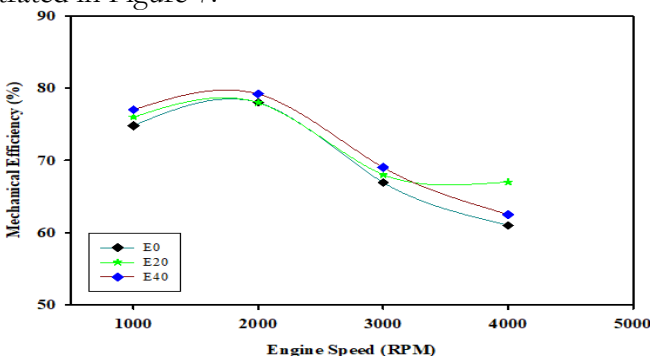


Figure 7 Engine Speed (RPM) vs Engine Mechanical Efficiency

At 2000 RPM, the engine demonstrates its highest mechanical efficiency, highlighting the optimal operating speed for the tested blends. However, as the RPM increases beyond this point, a decline in mechanical efficiency is observed due to increased frictional losses and reduced combustion time. Notably, at higher speeds (3000-4000 RPM), the E20 blend maintains a slight advantage in mechanical efficiency, particularly at 4000 RPM, where it shows a marginal but measurable difference. Other blends, including E0 and E40, exhibit only minor variations in efficiency at these higher RPM levels. This suggests that E20 offers a balance of efficient combustion and mechanical output, making it more suitable for high-speed operations compared to other blends.

Discussion:

The findings from this study demonstrate that bioethanol-gasoline blends significantly influence the performance characteristics of a motorcycle engine. The AVL Boost simulation results indicate that the E20 blend provides the most balanced improvement in power output, torque, and efficiency, making it a promising alternative to pure gasoline for small-displacement engines.

One of the key observations is the increase in engine power and torque with ethanol addition, particularly at low-to-mid engine speeds (1000-3000 RPM). The E20 blend shows a 4% improvement in power compared to E0, which aligns with previous research findings that highlight the efficiency of low-ethanol blends in spark ignition (SI) engines. The increased performance is attributed to ethanol's higher oxygen content, leading to better combustion and reduced knocking tendencies. However, at higher RPMs (4000+), efficiency gains plateau

or slightly decline, possibly due to ethanol's lower energy density (26 MJ/kg vs. 45 MJ/kg for gasoline), which increases fuel consumption.

While the simulation results are promising, real-world experimental validations are necessary to confirm these findings. Previous studies using dynamometer testing on ethanol-gasoline blends have reported similar trends, with E20-E40 providing optimal trade-offs between performance and fuel economy. However, higher ethanol concentrations (E50-E85) require modifications in fuel injection systems, which are not addressed in this study. Future research should explore real-time road tests to validate AVL Boost predictions.

The increase in fuel consumption with ethanol addition is a critical concern, as ethanol has a lower energy density than gasoline. As shown in Figure 5, fuel consumption increases as ethanol content rises beyond E20. This is consistent with previous research, which found that ethanol's higher volatility leads to increased evaporation losses and fuel consumption. However, its higher-octane number (RON 89.7) and heat of vaporization (854 kJ/kg) contribute to smoother combustion and reduced engine knocking, making it beneficial in modern engines designed for alternative fuels.

Although this study does not analyze emissions, previous research indicates that ethanol blends can significantly reduce CO and HC emissions due to improved combustion efficiency. However, NOx emissions may increase due to higher combustion temperatures. Future studies should incorporate exhaust gas analysis to assess the environmental impact of bioethanol adoption in motorcycle engines.

Several challenges and opportunities exist for further research in this domain. Experimental validation through dynamometer-based testing would provide greater accuracy in performance assessments. Investigating the impact of higher ethanol blends such as E50 and E85, along with analyzing combustion characteristics like heat release rates, flame propagation, and ignition delays, would enhance the understanding of ethanol's effects on engine operation. Examining potential engine modifications required for higher ethanol concentrations and assessing the economic viability of ethanol fuel adoption in developing countries would further strengthen the case for bioethanol as an alternative fuel.

This study confirms that bioethanol-gasoline blends, particularly E20, improve motorcycle engine performance while maintaining fuel economy at lower RPMs. However, challenges remain in fuel consumption, emissions trade-offs, and compatibility with existing engine systems. With further research, ethanol-blended fuels could serve as a viable transition towards sustainable and cleaner transportation fuels.

Conclusion:

This study aims to examine the performance of engines using ethanol-gasoline blends. The results show that engine performance can be improved by using bioethanol-gasoline mixtures. Increasing the ethanol content slightly boosts the power of motorcycle engines as engine speed increases. The E20 blend showed a minor improvement in power output. More research is needed to explore the potential of ethanol-gasoline blends in motorcycle engines. Future studies could focus on developing Flexible Fuel Vehicles (FFVs) that can efficiently use these blends and help reduce dependence on fossil fuels. Adding hydrogen to the blends may further enhance engine performance and lower carbon emissions.

Acknowledgment: The author acknowledges the AVL List GmbH (AVL) for providing a Licensee for educational purposes to the Department of Mechanical Engineering, University of Gujrat, Pakistan.

Author's Contribution: Every co-author makes an equal contribution to this paper.

Conflict of interest: The authors declare no conflict of interest regarding this publication.

References:

[1] R. K. Abrar, Iyman, Arora, Tavishi, "Bioalcohols as an alternative fuel for transportation: Cradle to grave analysis," *Fuel Process. Technol.*, vol. 242, p. 107646,

2023, doi: https://doi.org/10.1016/j.fuproc.2022.107646.
[2] Ifeanyichukwu Edeh, "Bioethanol Production: An Overview," *Bioethanol Technol.*, 2020, doi: 10.5772/intechopen.94895.
[3] J. S. Mahmood Ebadian, Susan van Dyk, James D. McMillan, "Biofuels policies that have encouraged their production and use: An international perspective," *Energy Policy*, vol. 147, p. 111906, 2020, doi: https://doi.org/10.1016/j.enpol.2020.111906.
[4] R. Y. Aliyu Galadi ALHASSAN, Richard Balthi MSHELIA, "REVIEW OF THE USE OF BIOETHANOL AS ALTERNATIVE FUEL FOR INTERNAL COMBUSTION ENGINES," *ACTA Tech. CORVINIENSIS – Bull. Eng.*, 2021, [Online]. Available: https://acta.fih.upt.ro/pdf/2021-2/ACTA-2021-2-25.pdf
[5] H. C. J. Prasad, S., Anoop Singh, "Ethanol as an alternative fuel from agricultural, industrial and urban residues," *Resour. Conserv. Recycl.*, vol. 50, no. 1, pp. 1–39, 2007, doi: https://doi.org/10.1016/j.resconrec.2006.05.007.
[6] V. A. Lubhan Cherwoo, Ishika Gupta, G. Flora, Ritu Verma, Muskaan Kapil, Shailendra Kumar Arya, Balasubramani Ravindran, Kuan Shiong Khoo, Shashi Kant Bhatia, Soon Woong Chang, Chawalit Ngamcharussrivichai, "Biofuels an alternative to traditional fossil fuels: A comprehensive review," *Sustain. Energy Technol. Assessments*, vol. 60, p. 103503, 2023, doi: https://doi.org/10.1016/j.seta.2023.103503.
[7] T. Y. Muhamad Norkhizan Abdullah, Ahmad Fitri Yusop, Rizalman Mamat, Mohd Adnin Hamidi, Kumarasamy Sudhakar, "Sustainable Biofuels from First Three Alcohol Families: A Critical Review," *Energies*, vol. 16, no. 2, p. 648, 2023, doi: https://doi.org/10.3390/en16020648.
[8] J. T. K. and D. K. G Bellér, I Árpád, "AVL Boost: a powerful tool for research and education," *J. Phys. Conf. Ser.*, vol. 1935, p. 012015, 2021, doi: 10.1088/1742-6596/1935/1/012015.
[9] G. L. Paolo Iodice, Amedeo Amoresano, "A review on the effects of ethanol/gasoline fuel blends on NOX emissions in spark-ignition engines," *Biofue Res. J.*, vol. 8, no. 4, pp. 1465–1480, 2021, doi: 10.18331/BRJ2021.8.4.2.
[10] Amit Pal, "Blending of Ethanol in Gasoline:Impact on SI Engine Performance and Emissions," *Int. J. Therm. Technol.*, vol. 4, no. 1, pp. 1–5, 2011, doi: 10.14741/ijtt/mar.2014.01.
[11] Y. S. Sanap, Pooja P., Mahajan, "Ethanol blending in petrol: A techno - commercial overview," *Mater. Today Proc.*, vol. 102, pp. 113–125, 2024, doi: https://doi.org/10.1016/j.matpr.2023.04.055.
[12] S. N. Alfredas Rimkus, Saugirdas Pukalskas, Gabrielius Mejeras, "Impact of Bioethanol Concentration in Gasoline on SI Engine Sustainability," *Sustainability*, vol. 16, no. 6, p. 2397, 2024, doi: https://doi.org/10.3390/su16062397.
[13] M. Babu, M. Sivaraman, T. Jeyaseelan, and A. MA, "A comprehensive assessment of vehicular performance and emission characteristics during real-time test-runs using selected bio-ethanol–gasoline blends," *Biofuels*, vol. 15, no. 2, pp. 137–145, 2024, doi: 10.1080/17597269.2023.2221968.

Design Of a Photocatalytic Reaction System for Pollutant Degradation: A Computational Study

Fatima Aamir, Muhammad Suleman Tahir, Abdul Basit, Abdul Hannan Zahid
[1]Department of Chemical Engineering, University of Gujrat, Pakistan
***Correspondence:** fatimaaamirabbas966@gmail.com

Citation | Aamir. F, Tahir. M. S, Basit. A, Zahid. A. H, "Design of A Photocatalytic Reaction System for Pollutant Degradation: A Computational Study", IJIST, Special Issue. pp 245-256, March 2025.
Received | Feb 25, 2025 **Revised |** March 09, 2025 **Accepted |** March 15, 2025 **Published |** March 18, 2025.

In this study, Computational Fluid Dynamics (CFD) was used to model and simulate the photocatalytic degradation of methyl orange (MeO) in a stirred photoreactor, particularly in the presence of a bismuth oxide catalyst. This approach not only provides an effective method for treating wastewater by breaking down harmful dye pollutants but also highlights the potential of cost-effective and eco-friendly catalytic materials for environmental cleanup. In the first phase, the catalyst was evenly distributed in an aqueous MeO solution, where photocatalysis was employed to degrade the pollutant. The structural properties of the catalyst were analyzed using scanning electron microscopy (SEM). Experiments were conducted to examine how different factors, such as pH and pollutant concentration, influenced MeO removal.

In the next step, CFD was used to numerically analyze MeO degradation through photocatalysis. The results showed that the photoreactor effectively broke down MeO. CFD modeling further explained the degradation mechanism, revealing that hydroxyl radicals (OH•) played a key role in the heterogeneous photocatalytic process. Photocatalysis significantly contributed to pollutant breakdown in both experimental and simulated phases.

The CFD models closely matched experimental data, confirming the findings related to fluid dynamics and species concentration. By offering deeper insights into mass transfer and reaction kinetics at a fraction of the cost and time, CFD proved to be more efficient than experimental methods in analyzing MeO degradation.

Keywords. Computational Fluid Dynamics (CFD); Photocatalytic Degradation; Methyl Orange; Bi_2O_3 Nanoparticles; Predictive Modelling

Introduction:

Environmental pollution is a long-standing issue and a major cause of environmental illness and mortality, making it one of the biggest challenges facing humanity today. Every day, our environment is increasingly contaminated by various toxic and hazardous pollutants. Among these, organic pollutants such as pesticides[1], pharmaceuticals[2], and dyes[3] are particularly concerning due to their harmful effects on human health and aquatic life.

Methyl orange, a synthetic dye widely used in textiles, printing, and laboratories, is one such pollutant that requires degradation. Due to its chemical properties, it poses serious health and environmental risks to living organisms. Pollutant degradation refers to breaking down or removing harmful substances. The available methods for degradation fall into three main categories: physical, chemical, and biological. Physical methods include filtration and adsorption, chemical methods involve advanced oxidation processes[4], electrochemical techniques[5], and chemical precipitation, while biological methods include bioremediation[6], phytoremediation[7], and enzymatic treatment[8]. The conventional techniques for degrading organic contaminants may eventually be replaced by the environmentally friendly approach of photocatalysis.

Photocatalysis is considered one of the most effective chemical methods due to its high efficiency in breaking down persistent pollutants, eco-friendly nature, low-cost catalyst materials, self-sustaining mechanism, and scalability. This technique is particularly useful for treating wastewater with high contaminant levels, limited biodegradability, and complex compositions[9]. In practical applications, photocatalysts use solar energy to break down pollutants[10], degrading organic contaminants by absorbing specific wavelengths of light in water[11][12]. Under ambient temperature and pressure, photocatalysis converts pollutants into harmless molecules through redox reactions[13][14].

However, some photocatalysts face limitations, including high band gap energies (Eg), low light absorption capacity, and rapid electron-hole recombination. A high Eg value leads to inefficient charge separation, requiring more energy for activation[15]. Additionally, electrons and holes may recombine before contributing to the photocatalytic process, reducing the generation of reactive oxygen species (ROS) needed for complete photodegradation. This fast recombination rate hinders overall quantum efficiency[16].

Despite these challenges, photocatalysis remains one of the most effective techniques. The process begins when hydroxyl radicals (·OH) attack the dye molecule's weakest chemical bonds, such as the azo bond (–N=N–). When this structure is broken, the dye's conjugated system is disrupted[17]. As the reaction progresses, the intermediates undergo further oxidation, ultimately decomposing into carbon dioxide and water[18].

Table 1 Reactions included in the degradation process[19]

Reactions
$2HO_2 \rightarrow O_2^{\circ} + H_2O_2$
$H_2O_2 + O_2^{\circ} \rightarrow OH^{\circ} + OH^{-} + O_2$
$e^{-} + O_2 \rightarrow O_2^{\circ -}$
$O_2^{\circ -} + H^{+} \rightarrow 2HO^{\circ}{}_2$

The reactivity of dyes in degradation systems depends on their chemical structure[20]. Complex dyes generally have low photodegradability due to the presence of functional groups that affect adsorption characteristics. Many researchers are studying the removal of different dyes under visible and UV radiation[21][22].

Experimental work often involves multiple trials and errors, requiring extensive time and effort while sometimes yielding ineffective results. One of the most efficient ways to overcome these limitations is through modeling and simulation[23]. Computational Fluid Dynamics (CFD) simulations help analyze fluid behavior under different conditions, such as

flow dynamics and heat transfer. CFD also solves governing equations that describe fluid flow and decomposes the physical domain into smaller parts (meshing), including:

1. **Navier-Stokes equations** – govern the flow of viscous fluids.
2. **Continuity equation** – ensures mass conservation.
3. **Energy equation** – explains heat transfer within a fluid.

There is growing recognition of CFD's potential in chemical engineering and reaction engineering[24]. It has also been applied in various fields, such as fluidization[25] and multiphase flow systems[26].

Simulation methods include the Finite Volume Method (FVM), Finite Difference Method (FDM), Finite Element Method (FEM), and Lattice Boltzmann Method (LBM)[27]. Among these, FVM is the most accurate, particularly for processes involving interspecies interactions.

Research Objectives:

This paper presents the design of a photocatalytic reactor using advanced modeling and simulation methods for pollutant degradation through heterogeneous photocatalysis.

The primary objective of this research is to develop a Computational Fluid Dynamics (CFD) model in ANSYS Fluent to simulate the photocatalytic breakdown of wastewater contaminants using a bismuth oxide catalyst effectively. Additionally, this study aims to analyze how contaminant concentration and photocatalyst dosage influence the efficiency of pollutant degradation in an environmentally friendly manner.

Novelty statement:

This study provides mechanistic insights by modeling and simulating the photocatalytic degradation of methyl orange (MeO) in a stirred photoreactor using Computational Fluid Dynamics (CFD), particularly with a bismuth oxide catalyst. This approach not only offers an efficient method for treating wastewater by breaking down harmful dye contaminants but also highlights the potential of cost-effective and eco-friendly catalytic materials for environmental remediation.

Material and Methods:

The experimental procedure is detailed here, with all chemicals used in the preparation process being of AR grade and purchased from Merck. Double-distilled (DD) water was used throughout. $Bi(NO_3)_3 \cdot 5H_2O$ was dissolved in a few drops of HNO_3 and diluted to a 0.1 M solution with DD water to prevent Bi ion precipitation. After adding NH_4OH dropwise, the resulting gel was filtered and washed with DD water. The gel was then refluxed at 80°C for four hours to form crystals.

To convert $Bi(OH)_3$ into its oxide form, it was calcined at 600°C for three hours. The structural, optical, and degradation characteristics of the oxide sample were analyzed to determine its crystal system, band gap, and reaction kinetics. Polymorph alpha-bismuth oxide was synthesized using a simple, surfactant-free chemical process, with methyl orange (MeO) used as a model pollutant to assess its degradation properties. MeO, a water-soluble azo dye, was chosen with an initial concentration of 16.4 mg/L and a solution pH of 7.07.

For photocatalysis, a 500W Xenon lamp (Wacom XDS501S) was used as the light source due to its high intensity, ensuring strong photon flux, faster reaction rates, and efficient charge carrier activity. Bismuth oxide, as a catalyst, has a band gap that allows it to absorb visible light from the xenon lamp, enhancing absorption and improving degradation efficiency. The experiment was conducted 47 cm from the light source, with an incident light intensity of 17,400 lux on the sample. Dye colorization kinetics were monitored at intervals of one to four hours using a UV-vis spectrophotometer (Techcomp UV2301) to measure the residual dye concentration in the solution.

Initially, Bi_2O_3 was dispersed in the MeO aqueous solution and stirred in the dark for an hour, showing only a slight decrease in absorption. The photocatalytic experiment was conducted under two conditions: (i) without photocatalyst light irradiation and (ii) with photocatalyst irradiation. Results indicated that a portion of the MeO molecules adsorbed onto the sample's surface. Upon light exposure, the optimal absorption gradually decreased as irradiation time increased, confirming the catalytic oxidation of MeO by Bi_2O_3. The MeO absorption band shifted from 464 nm to 452 nm, indicating a stepwise movement of ethyl groups under radiation.

Table 2 Some chemical and physical properties of MeO

Chemical Name	**Dye Type**	**Molecular Weight g/mol**	**Molecular Formula**
4- [4(Dimethyl Amino) phenyl azo] Benzene Sulfonic Acid Sodium salt	Organic	327.33	$C_{14}H_{14}N_3NaO_3S$

Analytical Techniques:

The ultraviolet-visible (UV-Vis) spectrum recorded during the photodegradation of MeO with Bi_2O_3 is shown in Figure. 1. MeO has a maximum absorption at 464 nm and can absorb light in both the visible and ultraviolet regions. The absorption in the visible spectrum is attributed to MeO's azo linkage chromophore.

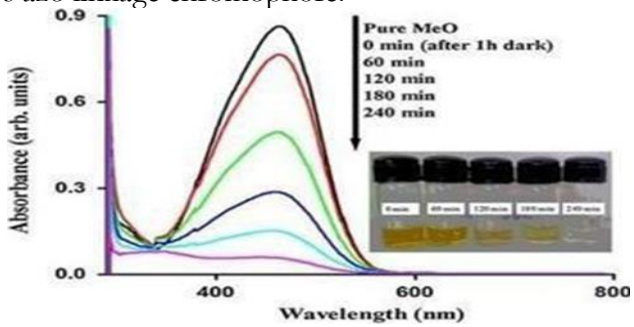


Figure 1. UV-vis spectra of photocatalytic degradation of MeO using Bi_2O_3 calcined at 600°C[28]

Figure. 2 (HR-SEM) presents an image used to analyze the surface morphology of the Bi_2O_3 catalyst. The sample consists of irregularly shaped platelets of varying sizes and forms, with some appearing as large, smooth platelets with sharp edges. This analysis provides insights into the textural properties of the catalyst, including porosity, particle size, and surface roughness—key factors influencing its performance and activity. HR-SEM is crucial as it captures high-magnification images, revealing nanoscale features such as particle size distribution, surface morphology, and structural integrity, all of which impact the material's physical, chemical, and catalytic properties.

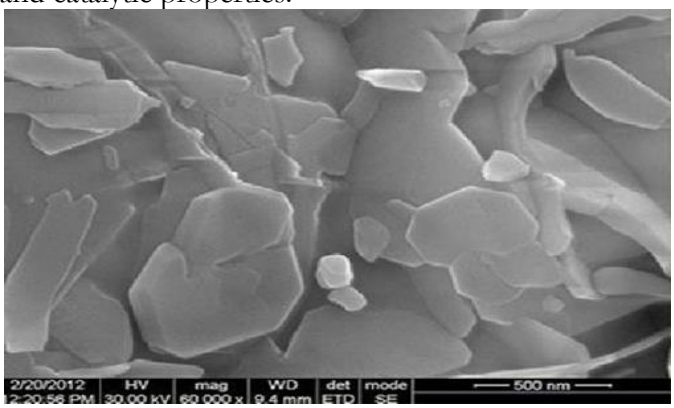


Figure 2. HR-SEM images of Bi_2O_3 calcined at 600°C [28]

Properties and the Degradation Mechanism of Photocatalysis

The photocatalytic process involves a reactor equipped with a steel mixer/stirrer, an inlet zone, a light source, and an outlet zone. The reacting mixture is placed inside the reactor, with bismuth oxide (Bi_2O_3) dispersed as the catalyst. A 500W lamp serves as the ultraviolet light source, maintaining a constant intensity throughout the experiment.

A 50 ml dye solution is introduced into the reactor, and the catalyst is dispersed within it. When Bi_2O_3 is mixed with the MeO aqueous solution in the dark and stirred for an hour, only a slight reduction in absorption is observed. However, after four hours of exposure to light, the degradation of MeO without a photocatalyst is just 2%, whereas with Bi_2O_3, it reaches 93%. This confirms that Bi_2O_3 significantly enhances the decolorization of MeO. Following four hours of irradiation, the initial MeO concentration decreases from 16.4 mg/L to 1.1 mg/L. Samples were collected at specific intervals for analysis.

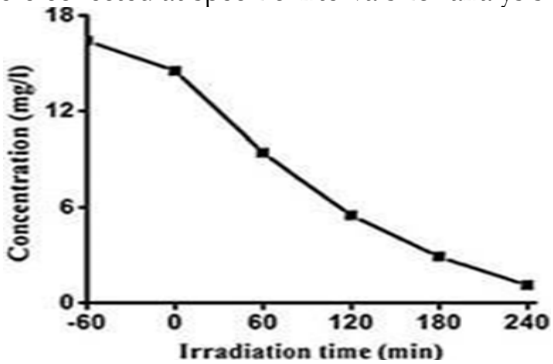


Figure 3. Variation in Concentration of pollutants with time in the degradation process [28]

CFD Modelling:

Governing Equations:

The equations for mass balance, momentum, and continuity for each species involved are presented in Equations (1)– (3) below. By solving these equations, the mass fraction, velocity, and pressure contours can be determined.

$$\frac{\partial(\alpha q \rho q)}{\partial t} + \nabla.(\alpha q \rho q u q) = \sum_{\rho=1}^{n} mp\dot{q} \quad (1)$$

$$\frac{\partial(\alpha q \rho q u q)}{\partial t} + \nabla.(\alpha q \rho q u q u q) = -\alpha q \nabla \rho + \alpha q \rho q g + \nabla.\tau q + \sum_{p=1}^{n}(Rpq + mpquq) + \alpha q \rho q(F_q + F_{lift,q} + F_{vm,q}) \quad (2)$$

$$\frac{\partial}{\partial t}(\alpha i \rho i Y i q) + \nabla.(\alpha i \rho i u i q) = -\nabla.\alpha i J i q + \alpha i R i q + \alpha i S i q + \sum_{p=1}^{n}(mpqij - mijpq) \quad (3)$$

u, ρ, Y, and α are the velocity, density mass fraction, and volume fraction of phases. For all the reactions, Arrhenius's reaction rate has been consumed.

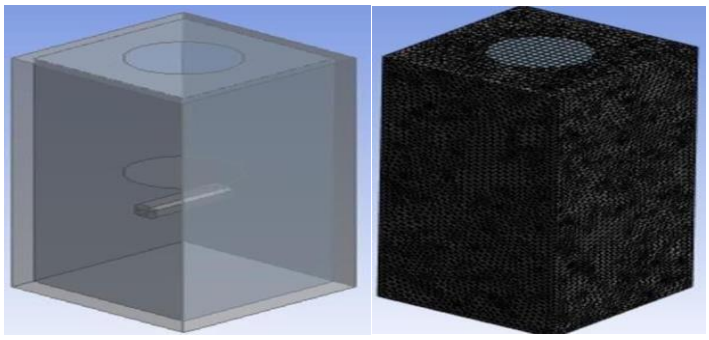


Figure 4 Geometry and meshing of the stirred photoreactor used in the degradation process

Boundary Conditions:

Boundary conditions play a crucial role in computational fluid dynamics (CFD) simulations, as they define fluid interactions with the surroundings and influence the

simulation's accuracy, stability, and realism. The photoreactor has four side walls, modeled as rigid solids with a no-slip boundary condition applied to their surfaces.

As shown in Figure. 4, turbulence and eddies significantly impact the process due to the presence of the stirrer. To capture these effects accurately, the k-ε turbulence model is used, providing reliable results that align with practical observations [29]. The photoreactor contains a 50 mL solution, and the initial values of various parameters, such as interior velocity, are set accordingly. A no-slip condition is applied to the four side walls.

A hollow space exists above the stirrer, allowing the reactant mixture to be introduced into the reactor. During the process, turbulence and eddies play a crucial role due to the stirrer's motion. Therefore, the k-ε turbulence model is implemented to ensure accurate and trustworthy results [30].

Numerical Method:

The photoreactor was designed using ANSYS Design Modeler. After modeling, meshing, and solving the continuity, momentum, and mass balance equations, ANSYS Fluent software was used. The software employs a finite volume approach, and the pressure-based solver was chosen due to its suitability for incompressible fluids and unstable-state equations.

For numerical stability and convergence, transient equations were discretized using the implicit first-order technique. Velocity, pressure, and species mole fractions were stored at nodes using a node-based solving method. The SIMPLE algorithm was used to couple velocity and pressure variables, while Green-Gauss discretization was applied to diffusion terms for practicality [31]. Standard wall functions were selected as the near-wall treatment approach, and mesh refinement near the walls ensured accurate flow field resolution.

A moving reference frame (MRF) technique was implemented to simulate the stirred tank photoreactor. The computational domain was divided into two zones: one containing the stirrer and another for the rest of the reactor. Instead of mobilizing the stirrer, a rotating motion was imposed on the inner zone to simulate fluid motion efficiently [32].

The discretized equations were solved for each mesh element, making mesh size crucial for accuracy. A grid independence study was conducted to determine the optimal mesh size. The simulation was transient, performed using ANSYS Fluent, with a time step size of 10 minutes and a total of 14,400 steps. The simulation ran for over two hours to generate results for post-processing. The minimum processor required for such simulations was an Intel Core i5 series.

Table 3. Mesh independence analysis

Reactor	Case	Mesh Sizes	No. of elements
	1	80cm	21841
Stirred-photoreactor	2	60cm	37324
	3	40cm	108273

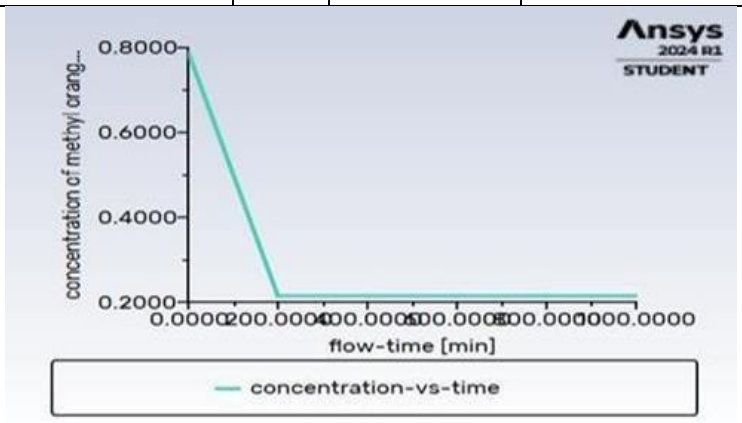


Figure 5. Trend of decreasing concentration of MeO pollutant concerning time.

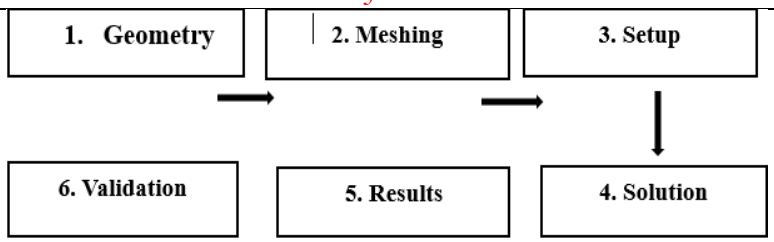


Figure 6. Flow diagram of methodology of the process used for the degradation in CFD simulation

SEM and AFM Analysis Results:

SEM results for the catalyst surface were analyzed using high-resolution images, as shown in Figure. 2. The Bi_2O_3 sample appears as irregularly shaped platelets of varying sizes and forms. Below, larger, smooth platelets with sharp edges can be observed.

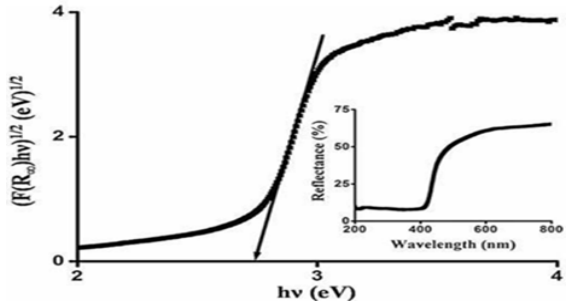


Figure 7. Band gap of Bi_2O_3 calcined at 600°C using Kubelka- Munk function with the reflectance spectra as the inset[28]

Figure. 7 presents the bandgap of Bi_2O_3 along with its reflectance spectra in the inset. The Kubelka-Munk function is applied to the reflectance spectrum to determine the sample's bandgap.

$$K/S = (1-R_\infty)^2/2R_\infty \equiv F(R_\infty) \quad (4)$$

Kubelka-Munk absorption (K) and scattering (S) coefficients are used to analyze the optical properties of Bi_2O_3. The Kubelka-Munk function, $F(R_\infty)$, is defined as $F(R_\infty) = (K/S)$, where R_∞ represents $R_{sample}/R_{standard}$. The band gap is determined by extrapolating the linear portion of the $(F(R_\infty)h\nu)^{1/2}$ vs. hν plot to $F(R_\infty) = 0$. The calculated bandgap value is 2.7344 eV (454 nm). This visible-light absorption suggests that Bi_2O_3 can act as a photocatalyst under visible light. MeO exhibits maximum absorption at 464 nm, covering both visible and ultraviolet regions, primarily due to its azo chromophore.

Results and Discussions:

CFD Results:

CFD simulations provide valuable insights into the spatial distribution of reactive species, such as •OH radicals, which are difficult to track experimentally in advanced oxidation processes (AOPs). The contours of velocity, velocity vectors, and mass concentration of pollutants help visualize the fluid dynamics within the photoreactor. Velocity vectors illustrate the flow patterns, showing how the fluid moves throughout the reactor. Dead zones, caused by eddies, indicate areas with reduced mixing, which can affect pollutant degradation efficiency. These simulations enable optimization of reactor design by identifying areas for improved fluid flow and mixing.

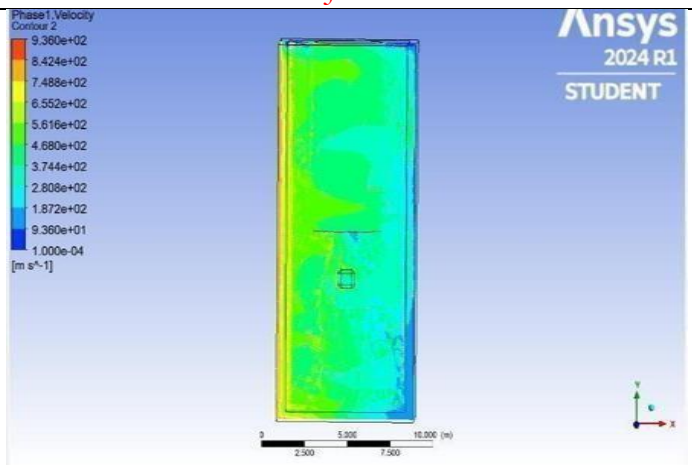


Figure 8. Contour of velocity used in the degradation process in the stirred photoreactor

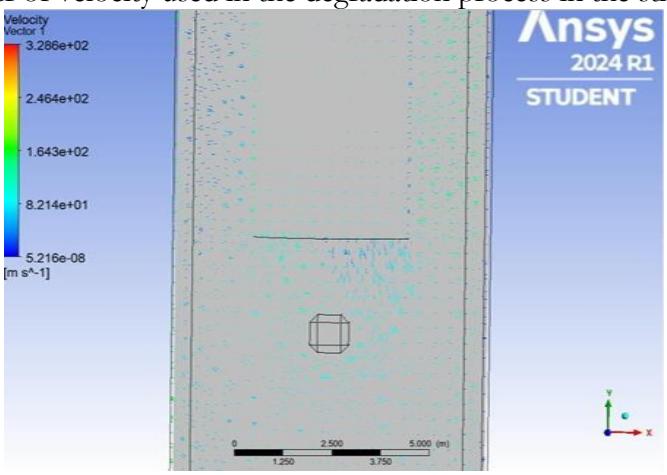


Figure 9. Velocity vectors of the stirred photoreactor of the degradation process

Yes, mass concentration contours are essential for understanding the spatial distribution of species within the reactor. They help in identifying regions with high and low concentrations, allowing researchers to pinpoint areas of pollutant accumulation or rapid degradation. This information is crucial for optimizing photocatalytic reactor design, ensuring uniform mixing, and enhancing reaction efficiency. By analyzing these contours, adjustments can be made to operating conditions, such as catalyst dosage, flow rates, or reactor geometry, to improve overall pollutant degradation performance.

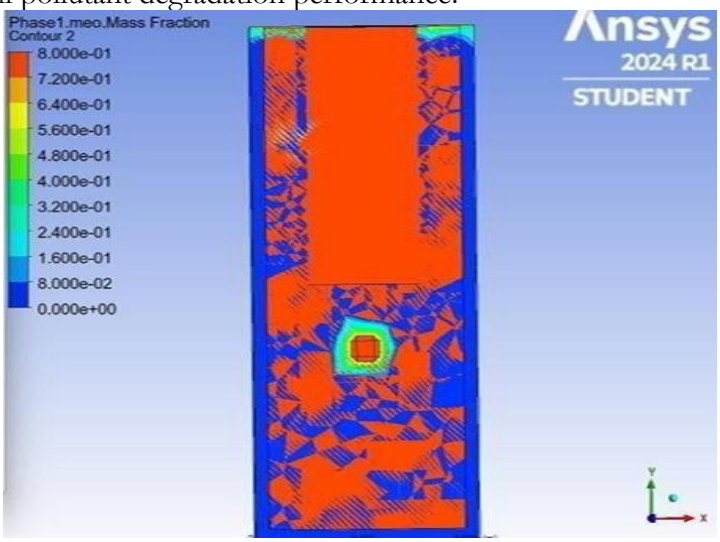


Figure 10. Mass concentration contour of the pollutant degraded in the stirred photoreactor

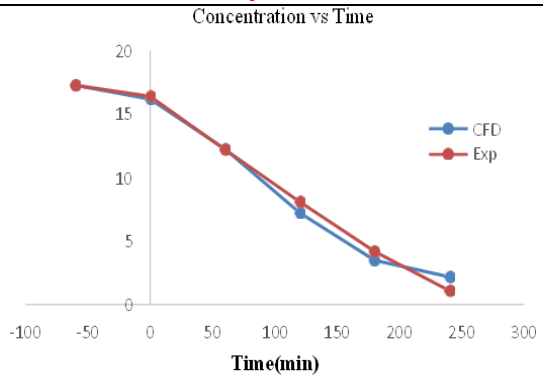


Figure 11. Comparison of CFD and experimental results of MeO degradation with time in the stirred-photo reactor

Yes, the high R^2 value (0.93) indicates strong agreement between the CFD model and experimental results, validating the accuracy of the simulation. The similarity in trends confirms that the model effectively captures the degradation kinetics of MeO. This suggests that CFD can reliably predict pollutant breakdown in photocatalytic reactors, making it a valuable tool for optimizing reaction conditions without extensive experimental trials.

Discussion:

Your conclusion effectively highlights the significance of Bi_2O_3 as a visible-light-driven photocatalyst and its potential for wastewater treatment. Here are some refinements for clarity and impact:

1. **Strengthen the Final Takeaway:** Conclude with a strong statement about Bi_2O_3's potential and future research directions.
2. **Improve Flow and Conciseness:** Some sections could be streamlined for readability.
3. **Enhance Technical Precision:** Differentiate between experimental and simulated findings when discussing mass transfer and turbulence.

Would you like me to refine and edit it accordingly?

Conclusion:

In this study, the removal of methyl orange (MeO) from wastewater was investigated due to its toxicity to aquatic life. Computational Fluid Dynamics (CFD) simulations were employed to model MeO degradation under heterogeneous photocatalytic conditions, incorporating the mixture model and radiation model. Key factors such as light intensity, catalyst distribution, and impeller speed were optimized to achieve maximum degradation. Experimental results confirmed significant pollutant removal, with approximately 93% efficiency. Numerical simulations provided insight into reaction mechanisms and concentration variations over time, validating the model's reliability. The strong agreement between CFD and experimental findings highlights CFD's potential to reduce experimental costs while optimizing wastewater treatment processes.

Limitations:

- The model needs to be .validated if the pollutant and catalyst are changed.

Future work:

Based on the conclusion provided, here are some potential future work conclusions:

- Investigate the scalability of the stirred photoreactor system using CFD for industrial applications while maintaining high removal efficiency.
- Extend the CFD simulations to include systems with multiple pollutants and evaluate their interactions and degradation pathways.
- Explore various methods of catalyst distribution to improve pollutant degradation efficiency further and validate these techniques through additional CFD simulations.

Acknowledgment: The authors wish to express their gratitude to the University of Gujrat for their invaluable contribution and support throughout this study. Sincere efforts from all participants were essential to the completion of this study, and they deserve special recognition. We also owe a debt of gratitude to our parents for their financial and moral assistance as well as prayers.

Author's Contribution: MST supervised the project and provided critical feedback during the research process. AB conceptualized the research idea designed the study framework and helps in performing computational modeling and simulation work. AHZ assisted with experimental data and monitored data acquisition. .

Conflict of interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Project details: Nil

References:

[1] P. S. Yuan Liu, Jian Wang , Xuran Zhu, Yang Liu, Ming Cheng, Weihai Xing, Yuping Wan, Na Li, Liting Yang, "Effects of electrolyzed water treatment on pesticide removal and texture quality in fresh-cut cabbage, broccoli, and color pepper," *Food Chem.*, vol. 353, p. 129408, 2021, doi: https://doi.org/10.1016/j.foodchem.2021.129408.

[2] R. A. T.-P. Ana L. Camargo-Perea, Efraím A. Serna-Galvis, Judy Lee, "Understanding the effects of mineral water matrix on degradation of several pharmaceuticals by ultrasound: Influence of chemical structure and concentration of the pollutants," *Ultrason. Sonochem.*, vol. 73, p. 105500, 2021, doi: https://doi.org/10.1016/j.ultsonch.2021.105500.

[3] C. A. Islam Ibrahim, George Belessiotis, Michalis K. Arfanis, "Surfactant Effects on the Synthesis of Redox Bifunctional V2O5 Photocatalysts," *Materials (Basel).*, vol. 13, no. 20, p. 4665, 2020, doi: 10.3390/ma13204665.

[4] A. B. R. Cátia Magro, Eduardo P. Mateus, Juan M. Paz-Garcia, "Emerging organic contaminants in wastewater: Understanding electrochemical reactors for triclosan and its by-products degradation," *Chemosphere*, vol. 247, p. 125758, 2020, doi: https://doi.org/10.1016/j.chemosphere.2019.125758.

[5] Luis Baptista-Pires , Norra Giannis-Florjan and Jelena Radjenovic, "Graphene-based sponges for electrochemical degradation of persistent organic contaminants," *Water Res.*, vol. 203, p. 117492, 2021, doi: https://doi.org/10.1016/j.watres.2021.117492.

[6] M. N. A. Rashi Miglani, Nagma Parveen, Ankit Kumar, Mohd. Arif Ansari, Soumya Khanna, Gaurav Rawat, Amrita Kumari Panda, Satpal Singh Bisht, Jyoti Upadhyay, "Degradation of Xenobiotic Pollutants: An Environmentally Sustainable Approach," *Metabolites*, vol. 12, no. 9, p. 818, 2022, doi: https://doi.org/10.3390/metabo12090818.

[7] O. O. B. Modupe Stella Ayilara, "Bioremediation of environmental wastes: the role of microorganisms," *Front. Agron.*, vol. 5, 2023, doi: 10.3389/fagro.2023.1183691.

[8] F. A. & S. S. A. Khadega A. Almaqdi, Rana Morsi, Bahia Alhayuti, "LC-MSMS based screening of emerging pollutant degradation by different peroxidases," *BMC Biotechnol.*, vol. 19, no. 83, 2019, doi: https://doi.org/10.1186/s12896-019-0574-y.

[9] R. A. El-Salamony, E. Amdeha, S. A. Ghoneim, N. A. Badawy, K. M. Salem, and A. M. Al-Sabagh, "Titania modified activated carbon prepared from sugarcane bagasse: adsorption and photocatalytic degradation of methylene blue under visible light irradiation," *Environ. Technol.*, vol. 38, no. 24, pp. 3122–3136, Dec. 2017, doi: 10.1080/21622515.2017.1290148.

[10] S. K. N.R. Khalid, A. Majid, M. Bilal Tahir, N.A. Niaz, "Carbonaceous-TiO2

nanomaterials for photocatalytic degradation of pollutants: A review," *Ceram. Int.*, vol. 43, no. 17, pp. 14552–14571, 2017, doi: https://doi.org/10.1016/j.ceramint.2017.08.143.

[11] W. T. C. Yean Ling Pang, Steven Lim, Hwai Chyuan Ong, "Synthesis, characteristics and sonocatalytic activities of calcined γ-Fe2O3 and TiO2 nanotubes/γ-Fe2O3 magnetic catalysts in the degradation of Orange G," *Ultrason. Sonochem.*, vol. 29, pp. 317–327, 2016, doi: https://doi.org/10.1016/j.ultsonch.2015.10.003.

[12] C. H. Hak, L. C. Sim, K. H. Leong, P. F. Lim, Y. H. Chin, and P. Saravanan, "M/g-C3N4 (M=Ag, Au, and Pd) composite: synthesis via sunlight photodeposition and application towards the degradation of bisphenol A," *Environ. Sci. Pollut. Res.*, vol. 25, no. 25, pp. 25401–25412, Sep. 2018, doi: 10.1007/S11356-018-2632-8/METRICS.

[13] K.-H. P. Wan-Kuen Jo, "Heterogeneous photocatalysis of aromatic and chlorinated volatile organic compounds (VOCs) for non-occupational indoor air application," *Chemosphere*, vol. 57, no. 7, pp. 555–565, 2004, doi: https://doi.org/10.1016/j.chemosphere.2004.08.018.

[14] W. Z. Kowarsk Beata, Jerzy Baron, Stanisław Kandefer, "Incineration of Municipal Sewage Sludge in a Fluidized Bed Reactor," *Engineering*, vol. 5, no. 1, pp. 125–134, 2013, doi: 10.4236/eng.2013.51A018.

[15] A. M. A.-M. Kavin Micheal, A. Ayeshamariam, Rajender Boddula, Prabhakarn Arunachalam, Mohamad S. AlSalhi, J. Theerthagiri, Saradh Prasad, J. Madhavan, "Assembled composite of hematite iron oxide on sponge-like BiOCl with enhanced photocatalytic activity," *Mater. Sci. Energy Technol.*, vol. 2, no. 1, pp. 104–111, 2019, doi: https://doi.org/10.1016/j.mset.2018.11.004.

[16] S.-J. H. Yun Kyung Jo, Jang Mee Lee Son, Suji, "2D inorganic nanosheet-based hybrid photocatalysts: Design, applications, and perspectives," *J. Photochem. Photobiol. C Photochem. Rev.*, vol. 40, pp. 150–190, 2019, doi: https://doi.org/10.1016/j.jphotochemrev.2018.03.002.

[17] L. K. Z. Ye, "A comparative study of photocatalytic activity of ZnS photocatalyst for degradation of various dyes," *Optik (Stuttg).*, vol. 164, pp. 345–354, 2018, doi: https://doi.org/10.1016/j.ijleo.2018.03.030.

[18] R. Saravanan, F. Gracia, and A. Stephen, "Basic Principles, Mechanism, and Challenges of Photocatalysis," *Nanocomposites Visible Light. Photocatal.*, pp. 19–40, 2017, doi: 10.1007/978-3-319-62446-4_2.

[19] M. R. R. Fatemeh Poorsajadi, Mohammad Hossein Sayadi , Mahmood Hajiani, "Photocatalytic degradation of methyl orange dye using bismuth oxide nanoparticles under visible radiation," *Int. J. New Chem.*, vol. 83, pp. 229–239, 2021, doi: 10.22034/ijnc.2020.137235.1131.

[20] B. K. Ali Akbar Isari, Amir Payan, Moslem Fattahi, Sahand Jorfi, "Photocatalytic degradation of rhodamine B and real textile wastewater using Fe-doped TiO2 anchored on reduced graphene oxide (Fe-TiO2/rGO): Characterization and feasibility, mechanism and pathway studies," *Appl. Surf. Sci.*, vol. 462, pp. 549–564, 2018, doi: https://doi.org/10.1016/j.apsusc.2018.08.133.

[21] M. B. K. A.R. Khataee, "Photocatalytic degradation of organic dyes in the presence of nanostructured titanium dioxide: Influence of the chemical structure of dyes," *J. Mol. Catal. A Chem.*, vol. 328, no. 1–2, pp. 8–26, 2010, doi: https://doi.org/10.1016/j.molcata.2010.05.023.

[22] S. M. S. Jorfi, "Visible Light Photocatalytic Degradation of Azo Dye and a Real Textile Wastewater Using Mn, Mo, La/TiO2 /AC Nanocomposite," *Chem. Biochem. Eng. Q.*, 2018, doi: 10.15255/CABEQ.2017.1261.

[23] J. L. Bradshaw, N. Ashoori, M. Osorio, and R. G. Luthy, "Modeling Cost, Energy,

and Total Organic Carbon Trade-Offs for Stormwater Spreading Basin Systems Receiving Recycled Water Produced Using Membrane-Based, Ozone-Based, and Hybrid Advanced Treatment Trains," *Environ. Sci. Technol.*, vol. 53, no. 6, pp. 3128–3139, Mar. 2019, doi: 10.1021/ACS.EST.9B00184.

[24] H. de L. Castrillón, S. Romero-Vargas,Ibrahim, H., "Flow field investigation in a photocatalytic reactor for air treatment (Photo-CREC–air)," *Chem. Eng. Sci.*, vol. 61, no. 10, pp. 3343–3361, 2006, doi: https://doi.org/10.1016/j.ces.2005.11.039.

[25] C. J. C. Cooper Scott, "CFD simulations of particle mixing in a binary fludized bed," *Powder Technol.*, vol. 151, no. 1, pp. 27–36, 2005, doi: 10.1016/j.powtec.2004.11.041.

[26] Y. Jiang, M. R. Khadilkar, M. H. Al-Dahhan, and M. P. Dudukovic, "CFD of multiphase flow in packed-bed reactors: I. k-Fluid modeling issues," *AIChE J.*, vol. 48, no. 4, pp. 701–715, Apr. 2002, doi: 10.1002/AIC.690480406.

[27] W. M. Henk Kaarle Versteeg, "An Introduction to Computational Fluid Dynamics: The Finite Volume Method," Pearson Education Limited. Accessed: Mar. 13, 2025. [Online]. Available: https://books.google.com.pk/books/about/An_Introduction_to_Computational_Fluid_D.html?id=RvBZ-UMpGzIC&redir_esc=y

[28] D. P. P. Iyyapushpam, S., Nishanthi, S.T., "Photocatalytic degradation of methyl orange using α-Bi2O3 prepared without surfactant," *J. Alloys Compd.*, vol. 563, pp. 104–107, 2013, doi: https://doi.org/10.1016/j.jallcom.2013.02.107.

[29] B. Wu, "Computational Fluid Dynamics Investigation of Turbulence Models for Non-Newtonian Fluid Flow in Anaerobic Digesters," *Environ. Sci. Technol.*, vol. 44, no. 23, pp. 8989–8995, Dec. 2010, doi: 10.1021/ES1010016.

[30] David C. Wilcox, "Turbulence Modeling for CFD," *DCW Ind.*, 2006, [Online]. Available: https://cfd.spbstu.ru/agarbaruk/doc/2006_Wilcox_Turbulence-modeling-for-CFD.pdf

[31] A. J. O. et al Ebenezer Oluwatosin Atoyebi, "Computational Fluid Dynamics," *Compr. Mater. Process.*, 2024, [Online]. Available: https://www.sciencedirect.com/topics/materials-science/computational-fluid-dynamics

[32] A. K. Mahdi Ebrahimi Farshchi, Hassan Aghdasinia, "Heterogeneous Fenton reaction for elimination of Acid Yellow 36 in both fluidized-bed and stirred-tank reactors: Computational fluid dynamics versus experiments," *Water Res.*, vol. 151, pp. 203–214, 2019, doi: https://doi.org/10.1016/j.watres.2018.12.011.

Recycling of Laptop Spent Li-Ion Batteries and Characterization of Extracted Materials

Zarmeena Akhtar[1], Zubair Mehmood[2], Rizwan Raza[1]
[1]Department of Physics, COMSATS University Islamabad, Lahore Campus, Pakistan
[2]Department of Electrical Engineering, University of Gujrat, Pakistan
*** Correspondence:** zarmeenaakhtar88@gmail.com

Citation | Akhtar. Z, Mehmood. Z, Raza. R, "Recycling of Laptop Spent Li-Ion Batteries and Characterization of Extracted Materials", IJIST, Special Issue. pp 257- 273, March 2025.
Received | Feb 26, 2025 **Revised |** March 10, 2025 **Accepted |** March 16, 2025 **Published |** March 19, 2025.

As the use of smart devices increases, the energy demand continues to grow, leading to higher consumption of lithium-ion batteries (LIBs) in portable electronics such as laptops, tablets, smartphones, and electric vehicles. This increased usage has resulted in a rising number of discarded batteries, which contain hazardous chemicals and heavy metals that pose serious environmental risks. Recycling these batteries efficiently is essential for both environmental protection and economic sustainability. This study explores a recycling method for used laptop and notebook batteries through a pretreatment and solvent dissolution process, using mild phosphoric acid as the leaching agent. The hydro-metallurgical process successfully recovers 5.124% lithium and 42.143% cobalt, yielding lithium carbonate and cobalt hydroxide. The batteries, which consist of 50.80% lithium cobalt oxide ($LiCoO_2$) cathodes on aluminum and graphite anodes on copper foils, serve as the primary source of material recovery. The recovered lithium carbonate and cobalt hydroxide are then used to synthesize active powder for cathode material. Advanced characterization techniques, including Cyclic Voltammetry (CV), Raman spectroscopy, and Electrochemical Impedance Spectroscopy (EIS), are employed to analyze the electrochemical properties of the recovered materials and synthesized powders. The results confirm the effectiveness of this recycling method in recovering valuable materials while reducing environmental impact. By addressing the growing problem of battery waste, this approach supports the sustainable production of new batteries through the reuse of critical materials. The study emphasizes the importance of developing efficient recycling technologies to promote a circular economy and reduce dependence on raw material extraction.

Keywords: Lithium-Ion Batteries, Recycling, Hydro-Metallurgy, Lithium Carbonate, Cobalt Hydroxide

Introduction:

In today's rapidly advancing technological world, the demand for high-performance, lightweight, and energy-efficient devices is steadily increasing [1]. Lithium-ion batteries (LIBs), introduced by Sony in 1990, have become essential for powering various devices, from smartphones to electric vehicles, due to their compact size and high energy density [2][3]. However, the limited lifespan of consumer electronics (typically 1 to 3 years) has resulted in a growing number of discarded batteries, raising environmental concerns. Improper disposal of LIBs, which contain toxic substances such as lithium, cobalt, and nickel, poses serious risks to both the environment and valuable resources [4]. Therefore, developing efficient and sustainable recycling methods is critical [5]. Efforts to reduce cobalt usage in LIBs to cut costs have affected battery performance. Cobalt plays a crucial role in improving thermal stability, particle structure, and overall battery capacity. Its reduction or absence presents challenges not only for LIB performance but also for recycling effectiveness [6]. This trade-off highlights the difficulty of balancing cost, efficiency, and recyclability.

Although recycling has clear benefits, traditional methods like pyrometallurgy and the direct method raise significant environmental issues [7]. As the recycling industry expands, these conventional techniques may worsen environmental problems unless more energy-efficient alternatives are developed. To manage varying input materials, impurities, geometries, and changing market needs, adaptable recycling processes are necessary [8]. Additionally, the handling of LIB components requires scalable, standardized, and straightforward processes to ensure efficiency and safety [9]. Hydrometallurgical recycling, though not yet widely commercialized, shows great promise as a future solution for sustainable LIB management. This method enhances the recovery efficiency of cobalt (Co) and lithium (Li) due to the use of hydrogen peroxide as a reducing agent. Hydrogen peroxide promotes the formation of carbonate ions, which have higher solubility in acidic conditions, leading to increased metal extraction during the leaching process. The preferential generation of these easily dissolvable carbonate ions explains the observed improvements in Co and Li recovery rates [10].

The recycling of end-of-life LIBs is a rapidly growing industry with immense potential for future expansion. By prioritizing efficient collection and reuse systems, industries can secure a sustainable supply of critical raw materials while reducing dependence on external sources. Transitioning from informal disposal to formal, systematic LIB recovery frameworks is essential. However, low collection rates, particularly for consumer electronics, remain a key obstacle to scaling up LIB waste processing. Currently, data shows that Asian countries lead in collection efficiency, achieving rates of around 70%, mainly due to their dominance in LIB manufacturing, which incentivizes localized recycling infrastructure [11]. Meanwhile, regions like the European Union, Australia, and the Americas are making gradual progress but often struggle to gather sufficient volumes of spent LIBs to maintain economically viable recycling operations [12].

Handling LIBs requires strict safety protocols, as they are classified as hazardous due to their flammable liquid electrolytes. Potential risks include thermal runaway, which can be triggered by physical damage, extreme heat, or residual electrical charge, leading to fires or explosions during storage, transport, or processing. For instance, crushing discarded LIBs may puncture their casings, while inadequate discharge procedures increase the risk of short circuits. To address these challenges, three primary strategies are being explored for managing LIB waste: refurbishing (repairing for reuse), repurposing (adapting for less demanding applications, such as energy storage), and recycling (recovering raw materials) [13][14].

Novelty Statement:

This study introduces a novel recycling method for spent lithium-ion batteries (LIBs) using mild phosphoric acid as a leaching agent, an approach that has not been widely explored in previous research. Unlike conventional methods that rely on harsh chemicals or energy-intensive processes, this technique is more eco-friendly and energy-efficient. It achieves

impressive recovery rates of 5.124% lithium and 42.143% cobalt, which surpass many existing methods. Another distinctive feature is the synthesis of active cathode material ($LiCoO_2$) directly from recovered lithium carbonate and cobalt oxide, demonstrating a closed-loop recycling process rarely reported in the literature.

Objectives:

This study aims to tackle the environmental and economic challenges caused by the disposal of spent lithium-ion batteries (LIBs). The goal is to develop an efficient, scalable recycling process that recovers valuable materials like lithium and cobalt, essential for battery production. By using mild phosphoric acid as a leaching agent, the process minimizes environmental impact and reduces energy consumption. Additionally, the study focuses on synthesizing active cathode material ($LiCoO_2$) from the recovered components, proving the feasibility of a closed-loop recycling system. This approach not only reduces dependence on raw material extraction but also supports a circular economy by reusing critical resources. Ultimately, the research aims to offer a sustainable solution for managing battery waste while addressing the growing demand for energy storage technologies.

This paper highlights the increasing importance of recycling and explores methods to enhance it. It starts by stressing the urgency of recycling and the shift from relying on primary raw materials, like those from conventional mining, to recovering secondary resources from spent lithium-ion batteries (LIBs). The paper will also suggest ways to improve the efficiency and purity of material recovery. The main objective is to recycle lithium and cobalt as Li_2CO_3 and $Co(OH)_2$ from spent batteries.

Materials and Methods:

For this study, spent lithium-ion batteries (LIBs) were collected from the COMSATS hardware lab. Batteries from HP and Dell laptops were chosen for material extraction and analysis.

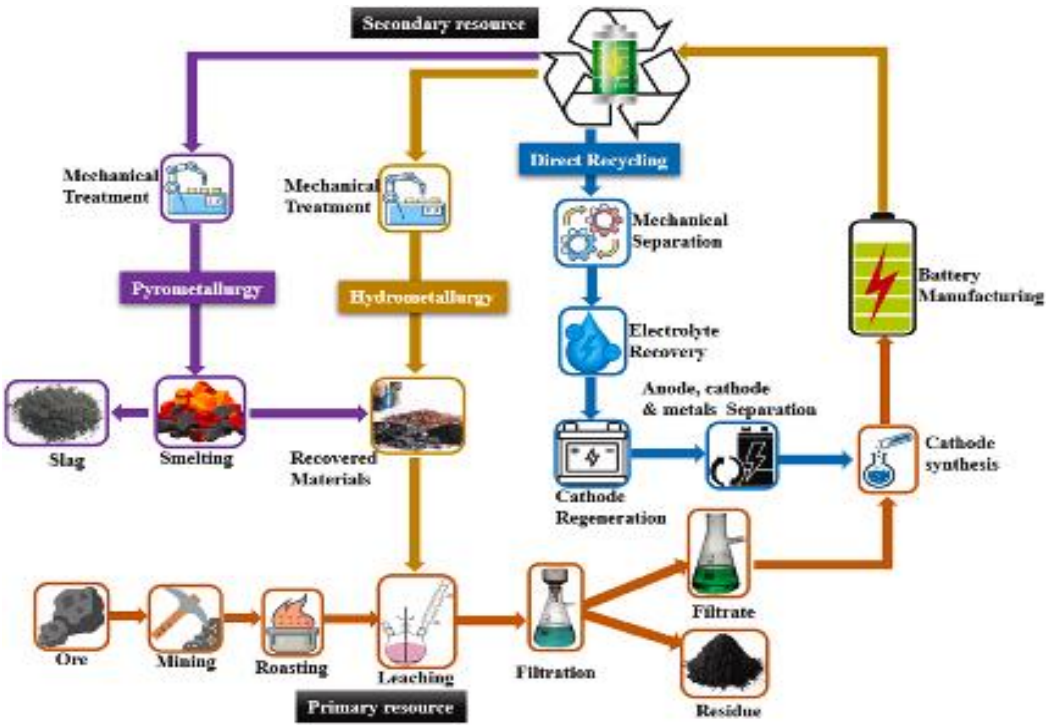


Figure 1. Flow diagram of various lithium-ion battery (LIB) recycling processes

The spent HP notebook battery, consisting of three cells, was selected for analysis. As shown in Table 1, the cathode material in this battery is lithium cobalt oxide (LCO), while the anode is made of graphite. This combination is commonly used in high-performance consumer electronics, such as mobile phones and laptops, due to its balance between energy density and

efficiency. Similarly, the spent Dell laptop battery, containing six cells, was also obtained from the COMSATS hardware lab. According to Table 1, this battery also uses lithium cobalt oxide as the cathode material, while the anode features a thin graphite film applied to the copper foil. This configuration, widely found in portable electronic devices, supports reliable energy delivery and extended cycle life.

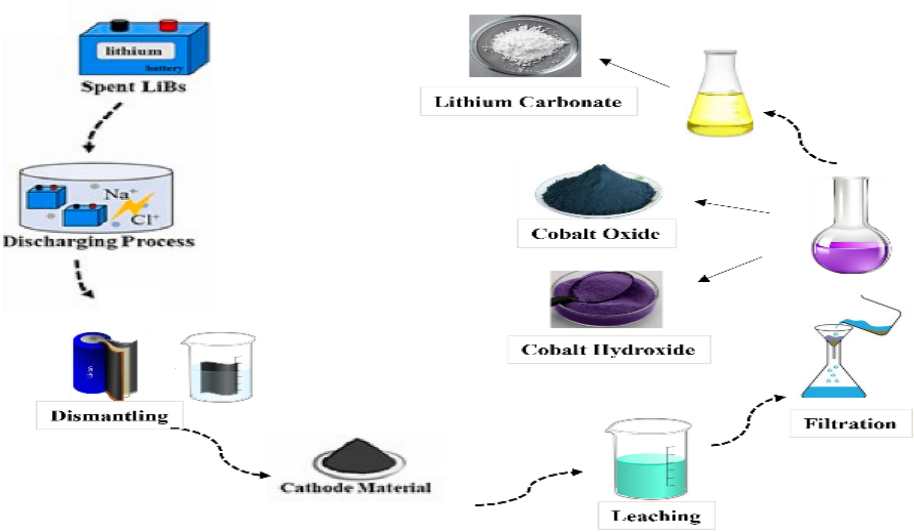


Figure 2. Schematic Diagram of Hydrometallurgical recycling procedure

Table 1. Spent HP Notebook and Dell Laptop Battery Specifications

LIB	Cathode Material	Specific Energy (Capacity) Wh/kg	Operating Range (Voltage) V/cell	Cyclic Life	Applications
HP (3 Cells)	LCO	150-200	3.0-4.2	500-1000	Mobile Phones, Laptops, Cameras
Dell (6 Cells)	LCO	170-200	3.0-4.2	600-1500	Mobile Phones, Laptops, Cameras

In addition to the specific batteries mentioned earlier, we also obtained lithium-ion battery (LIB) chemistries, including lithium nickel cobalt manganese oxide (NCM) and lithium manganese oxide (LMO). These chemistries offer distinct advantages based on their metal composition and are commonly used in applications that demand high power and energy density. Table 2 provides a detailed breakdown of the chemical compositions of the LIB chemistries analyzed in this study. Understanding the differences in the relative content of lithium (Li), manganese (Mn), cobalt (Co), and nickel (Ni) is essential for evaluating their electrochemical performance and overall stability.

The recycling process begins by discharging the battery cells to safely eliminate any residual charge. Once discharged, the next step is dismantling, which is performed carefully to avoid damaging the internal components. This process exposes the individual layers of the battery, which are then subjected to various chemical treatments to recover valuable materials. The recovered materials are further processed through controlled methods to ensure optimal purity and efficiency.

Table 2. Chemical Composition of Obtained Different LIBs Chemistries (%)

Chemistry	Li	Mn	Co	Ni
LCO	6037	0.23	54.43	0.22
LMO	3.85	47.03	0.13	0.02
NMC	6.93	19.28	19.33	20.8

Additional steps in the recycling process include filtration and sieving. The next crucial phase involves leaching, followed by further filtration, to separate valuable elements such as lithium, cobalt, and nickel from impurities. At this stage, the materials are divided into two outputs: sediment and filtrate. The filtrate, which contains the recoverable valuable elements, undergoes a recovery process for purification and refinement. Finally, the recovered materials are analyzed using characterization techniques to assess their quality and confirm whether the desired outcomes have been achieved. These steps not only facilitate the extraction of valuable components but also ensure that the recycling process is both efficient and environmentally sustainable.

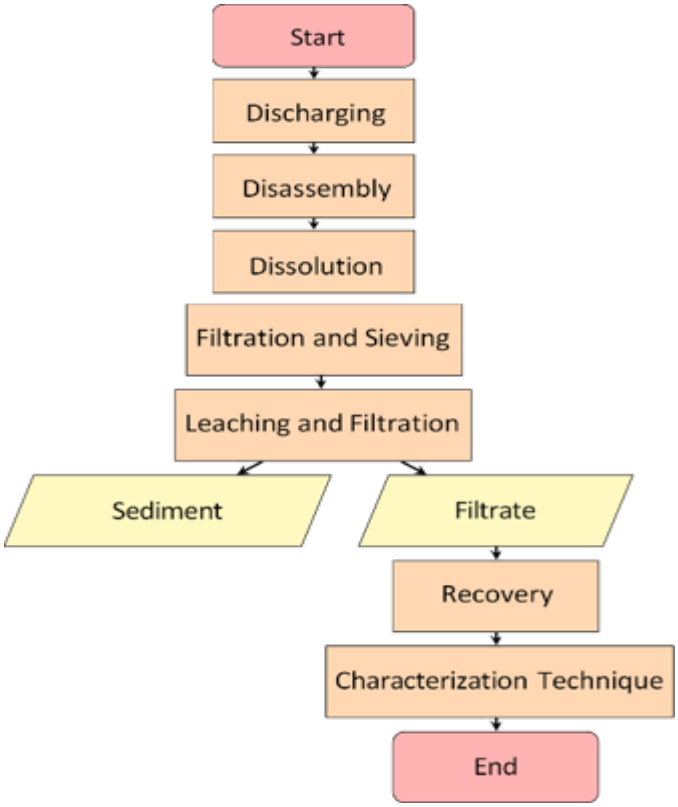


Figure 3. Flow diagram of a methodology for the Recycling Process for Spent LIBs

The recycling process for spent lithium-ion batteries (LIBs) begins with the crucial step of discharging the batteries to eliminate any remaining charge. Even after use, these batteries may retain 2-5% of their charge, which, if not properly discharged, could lead to short-circuiting and serious hazards such as explosions. To prevent this risk, the batteries were fully discharged before further processing.

The discharge process involved submerging the battery cells in a salt solution containing 500 mL of deionized water and 5 wt% sodium chloride (NaCl) for 24 hours inside a fume hood. During this period, gases such as nitrogen (N_2), carbon dioxide (CO_2), water vapor (H_2), hydrocarbons (C_xH_y), and acetic acid ester (CH_3COOCH_3) were released. This process separates the supernatant, which contains small amounts of lithium (Li), cobalt (Co), and phosphorus (P), from the sediment, which includes traces of aluminum (Al) and iron (Fe).

By fully discharging the batteries, the risk posed by residual charge is eliminated, making them safe for the subsequent stages of recycling and material recovery.

Figure 4. Discharging notebook battery cell

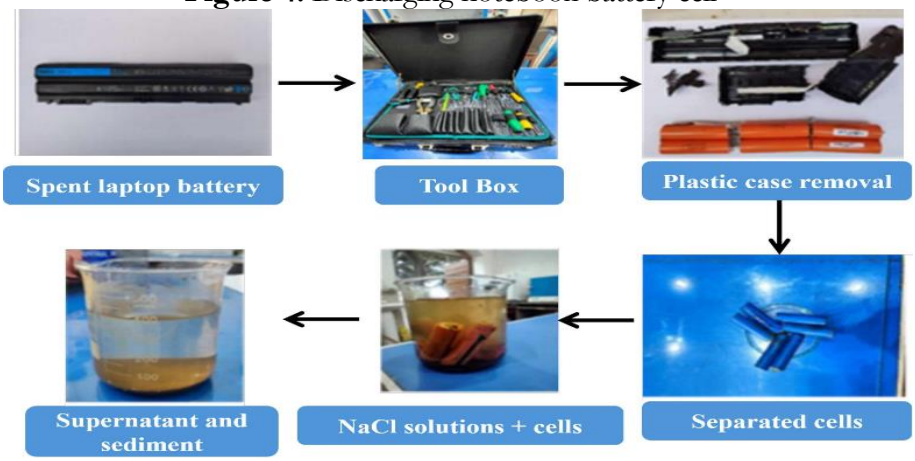


Figure 5. Discharging process of laptop battery cell

After discharging, the batteries were carefully dismantled. Using pliers and a blade, the plastic casing and electrical circuits were removed to expose the rectangular cells from the notebook batteries (Figure 4) and the cylindrical cells from the laptop batteries (Figure 5). Once separated, the cells were placed in an oven and dried at 60°C for 12 hours to ensure they were completely free of moisture before further dismantling.

Next, the dried cells were processed carefully by cutting and peeling off the plastic and metal shields. To safely open the metal casings, both ends were cut using a lathe machine, while a milling machine was employed to make precise incisions in the cylindrical metal shells. Once opened, the layers of the anode, separator, and cathode were carefully unrolled and separated, preparing them for subsequent processing and material recovery.

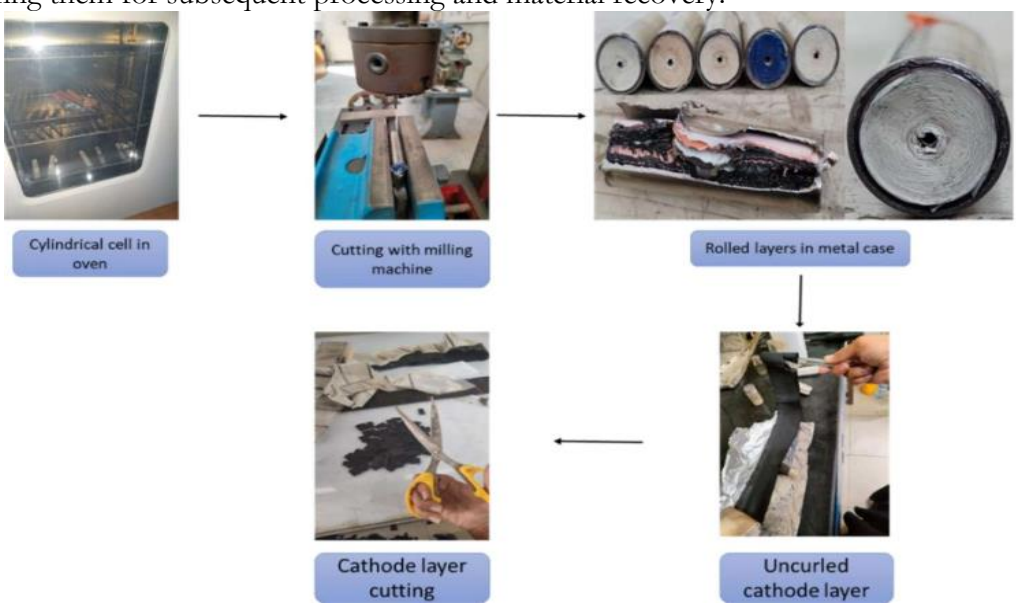


Figure 6. Dismantling process of the battery cells

The next step in the process involved the solvent dissolution method as shown in figure 7, which aimed to weaken the adhesion of the cathode material to the aluminum foil, making extraction easier. The binders, typically polyvinylidene fluoride (PVDF) or polytetrafluoroethylene (PTFE), posed challenges during separation. To address this, an organic solvent, N-Methyl-2-pyrrolidone (NMP), was used to dissolve the binders.

The aluminum foil was first cut into smaller pieces and then immersed in NMP at 130°C for 24 hours in an oven placed inside a fume hood. During this process, the cathode material is mixed with the solution, leaving the aluminum foil intact. The cathode material was subsequently recovered through filtration, followed by drying and grinding the filtrate residue to obtain a fine cathode powder.

This powder was then subjected to leaching to extract valuable metals such as lithium and cobalt. Phosphoric acid (0.7 moles) was used as the leaching agent and placed in a beaker on a magnetic stirrer set to a constant temperature of 40°C. Once the solution reached the desired temperature, the powdered cathode material was added, and hydrogen peroxide (H_2O_2) was added dropwise to enhance the leaching efficiency. The mixture was stirred continuously for one hour as shown in figure 8 and then allowed to cool to room temperature.

The leaching process successfully extracted lithium and cobalt ions from the cathode material into the solution. The resulting liquid was filtered to separate the sediment, which contained residual lithium, cobalt, carbon, and other impurities. The filtrate was then processed further to recover purified lithium and cobalt.

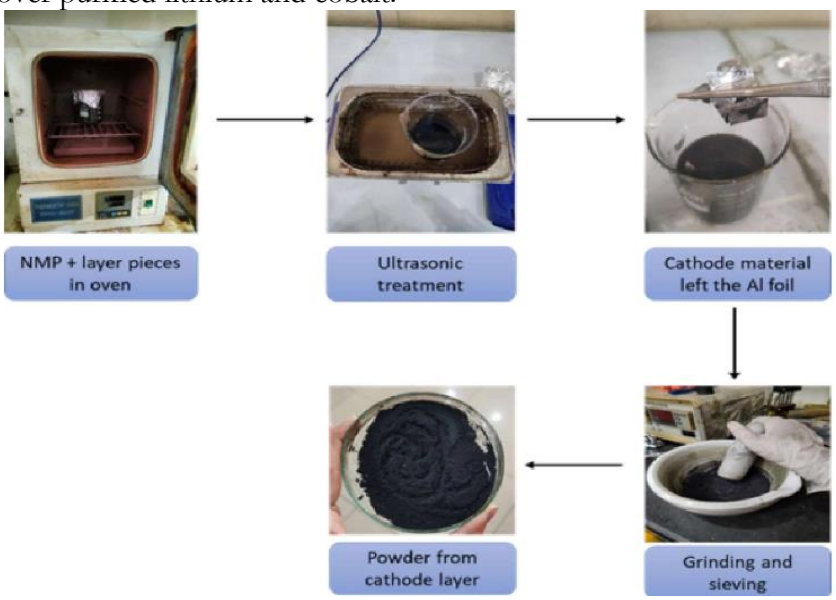


Figure 7. Cathode powder detached from aluminum foil

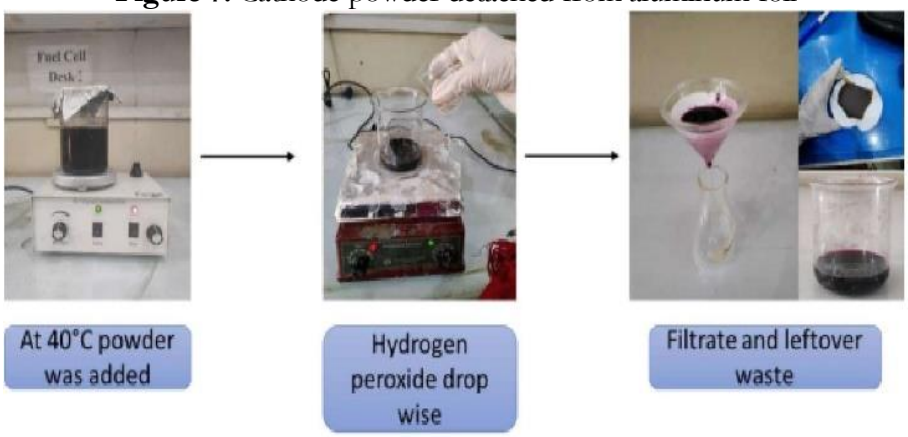


Figure 8. Leaching of filtrate

To recover cobalt from the filtrate, the solution was placed on a magnetic stirrer, and a 2M sodium hydroxide (NaOH) solution was added drop by drop, acting as a reducing agent. The pH level was continuously monitored using a pH meter figure 9 throughout the process.

As the pH reached 6, cobalt began to precipitate, forming violet-colored particles. This precipitation continued until the pH level rose to 8. The resulting mixture was then filtered to collect the violet cobalt precipitates on filter paper. These precipitates were dried to obtain cobalt hydroxide in powdered form, which served as a key intermediate product for further processing.

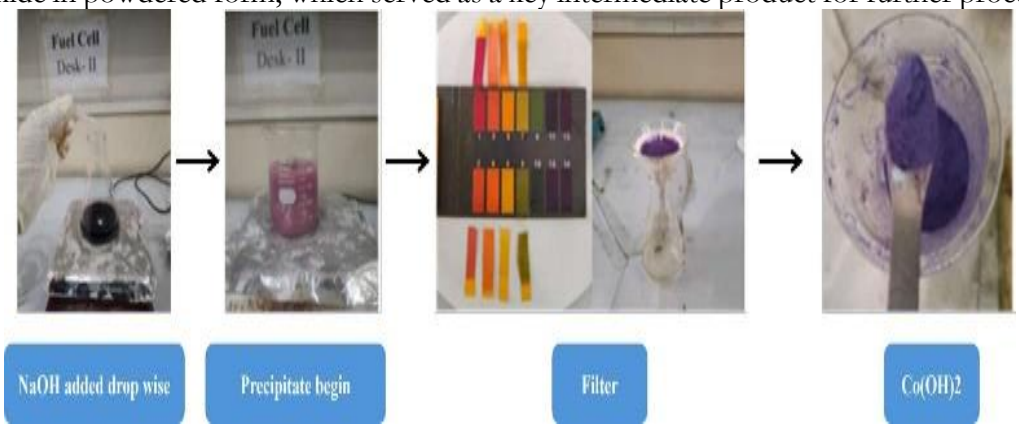


Figure 9. Recovery of Cobalt

To convert cobalt hydroxide into cobalt oxide, hydrogen peroxide (H_2O_2) was added dropwise to the cobalt hydroxide solution under constant stirring (Figure 10). This reaction led to the formation of brown cobalt oxide precipitates.

The solution was then placed in an ultrasonic bath set to 50°C and left for 5 hours, which helped refine and enhance the quality of the cobalt oxide precipitates. Afterward, the mixture was filtered, and the collected precipitates were dried to obtain cobalt oxide in its final powdered form.

The resulting filtrate from this process was subsequently used for lithium recovery.

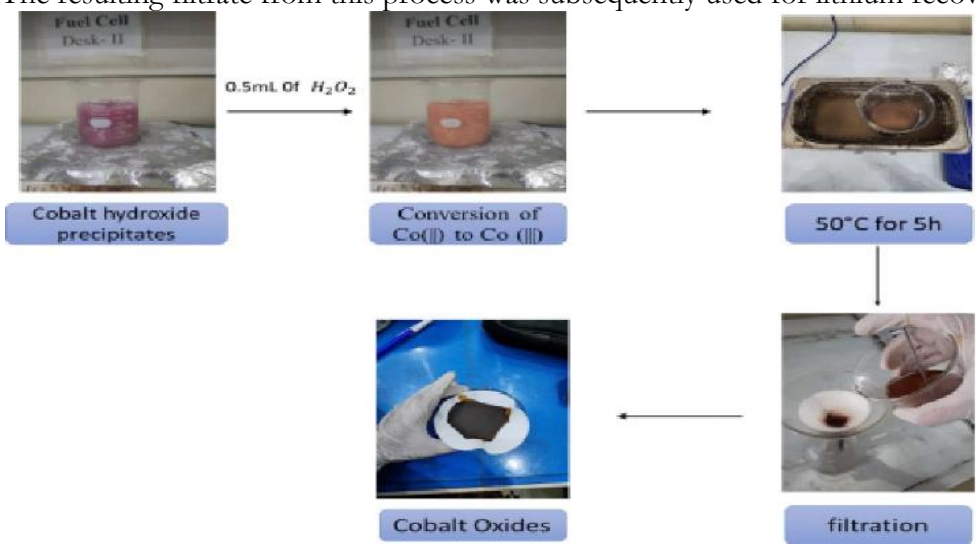


Figure 10. Brown precipitates of cobalt oxide from cobalt hydroxide

For lithium recovery, the filtrate was heated to 100°C while being continuously stirred on a magnetic stirrer. Sodium carbonate (Na_2CO_3) solution was then added drop by drop until white lithium carbonate precipitates began to form, as shown in Figure 11.

The precipitates were subsequently filtered and dried to obtain lithium carbonate in powdered form.

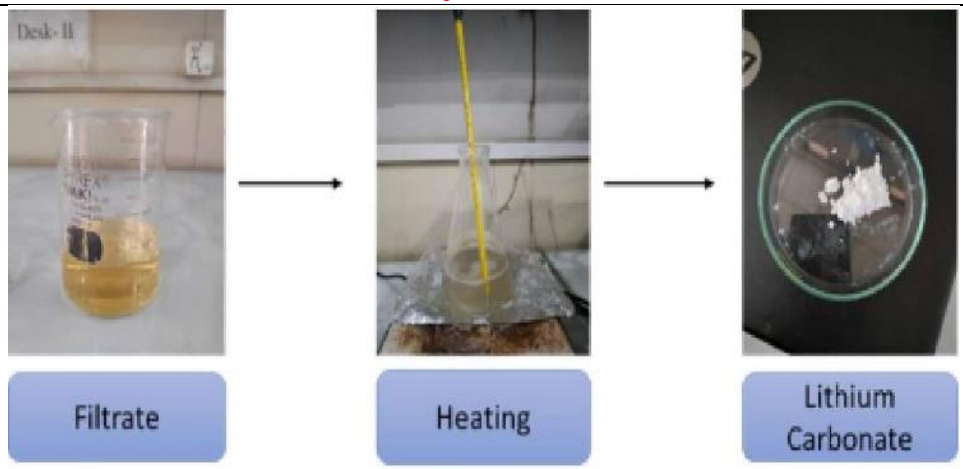


Figure 11. Recovery of lithium

Preparation of Active Material for Cathode:

Synthesis of Lithium Cobalt Oxide (LCO):

Figure 12 shows the flow chart for synthesizing the active powder. The extracted lithium carbonate and cobalt oxide particles were centrifuged to remove any remaining impurities. These recovered materials, obtained from spent LIBs, were then reused to prepare the cathode active material.

The purified cobalt oxide and lithium carbonate were thoroughly mixed and ground in a mortar before being calcined at 800°C for approximately 12 hours in a muffle furnace. After calcination, the resulting powder was re-grounded and sintered at 850°C to enhance its properties. This process yielded lithium cobalt oxide (LCO) powder, which was ready for use as cathode material.

Data on Recovered Material:

The recovered materials from the laptop and notebook batteries, along with their respective weights (in grams), are summarized in Table 3 below.

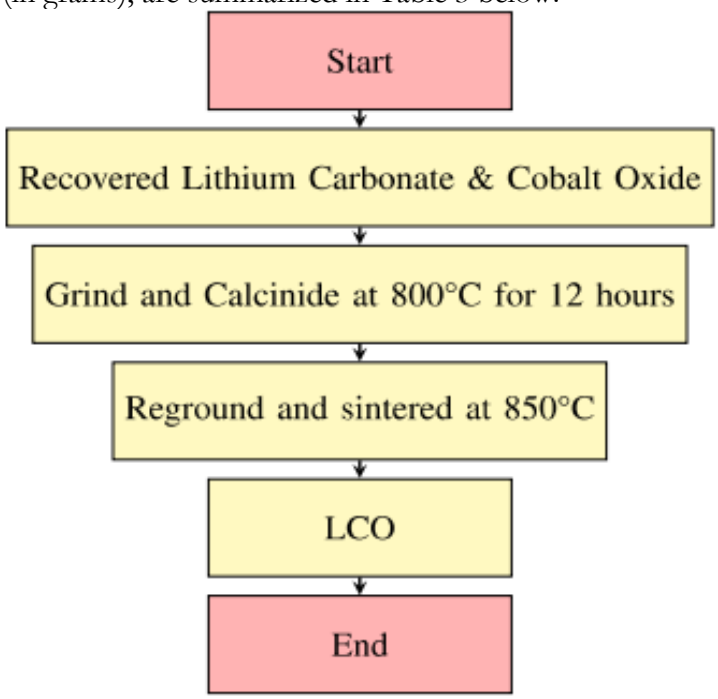


Figure 12. Flowchart for synthesis of Active Powder.

Table 3. Recovered Materials Data

Laptop Battery	Weight (g)	Note Book Battery	Weight (g)
Cathode Powder	20g	Cathode Powder	10g
Anode Powder	15g	Anode Powder	18g
Aluminium Foil	10g	Aluminium Foil	07g
Copper Foil	10g	Copper Foil	08g
Plastic and Separators	18g	Plastic and Separators	14g
Metal Case	25g	Metal Case	20g
Cobalt Hydroxide, Lithium Carbonate	8g, 4g	Cobalt Hydroxide, Lithium Carbinate	5g, 3g

Result and Discussion:

Raman Spectroscopy:

The Raman spectroscopy analysis of lithium carbonate, shown in Figure 13a, reveals characteristic peaks associated with the vibrational modes of both the carbonate ion (CO_3^{2-}) and lithium ions. The prominent band at approximately 1088 cm^{-1} corresponds to the symmetric stretching vibrations of the CO_3^{2-} ion, reflecting its planar structure. Weaker bands at around 748 cm^{-1} and 712 cm^{-1} represent in-plane bending vibrations, while a smaller band at ~1458 cm^{-1} is attributed to the asymmetric stretching of the C–O bonds. These peaks confirm the identity of lithium carbonate (Li_2CO_3) and provide crucial insights into its bonding, symmetry, and structural properties. This makes Raman spectroscopy an invaluable tool for investigating the role of lithium carbonate in batteries and energy storage technologies.

In Figure 13b, the Raman spectrum of cobalt hydroxide ($Co(OH)_2$) displays distinct peaks at 420 cm^{-1} and 510 cm^{-1}, which correspond to O-H bending vibrations and the Co-O symmetric stretching mode (Ag), respectively. These sharp, well-defined peaks indicate the material's high crystallinity and provide essential information about the bonding environment of cobalt and hydroxide ions. This structural characterization highlights the potential of cobalt hydroxide for applications in energy storage and catalysis.

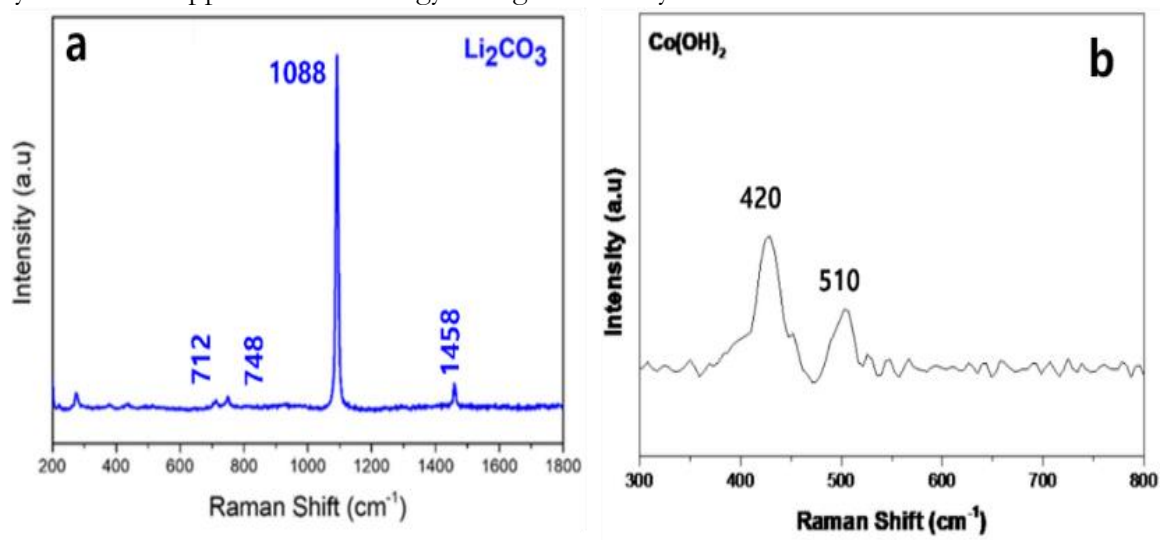


Figure 13. Raman Spectroscopy of (a)Lithium Carbonate (Li_2CO_3) (b) Cobalt Hydroxide ($Co(OH)_2$)

X-ray Diffraction:

XRD Analysis of Lithium Carbonate and Cobalt Hydroxide:

The XRD pattern of lithium carbonate (Li_2CO_3), shown in Figure 14a, confirms its monoclinic crystal structure with a space group of *C2/c*. Prominent diffraction peaks are

observed at approximately 21.4°, 30.6°, 36.0°, 39.5°, and 43.8° (2θ), corresponding to characteristic lattice planes such as (110) and (211). These sharp peaks reflect the high crystallinity and phase purity of the sample, with minimal structural defects. This well-crystallized nature enhances the thermal stability and electrochemical performance of lithium carbonate, making it highly suitable for applications that require structural integrity and reliability, such as battery technologies or as a precursor for other lithium-based compounds. Furthermore, the precise alignment of lattice planes indicates that the synthesis method employed was effective in achieving the desired phase with minimal impurities or defects.

The XRD pattern of cobalt hydroxide ($Co(OH)_2$), shown in Figure 14b, exhibits a prominent low-angle diffraction peak at *2θ* corresponding to the (003) basal reflection, which represents the interlayer *d*-spacing. This peak is characteristic of the layered double hydroxide (LDH) structure of $Co(OH)_2$. However, due to the intercalation of anions (e.g., NO_3^-, Cl^-, CO_3^{2-}) and water molecules within the layers, $Co(OH)_2$ exhibits a more disordered layered structure, resulting in broader and less defined XRD peaks. This irregular alignment of layers reduces the overall crystallinity of the material. Higher-order reflections, such as (006) and (009), may appear at regular intervals due to the periodic stacking of the layers. The exact position and intensity of the peaks may vary depending on factors such as the type of intercalated anions and the degree of hydration or water content in the structure. These structural characteristics influence the material's potential applications in catalysis, energy storage, and electrochemical processes.

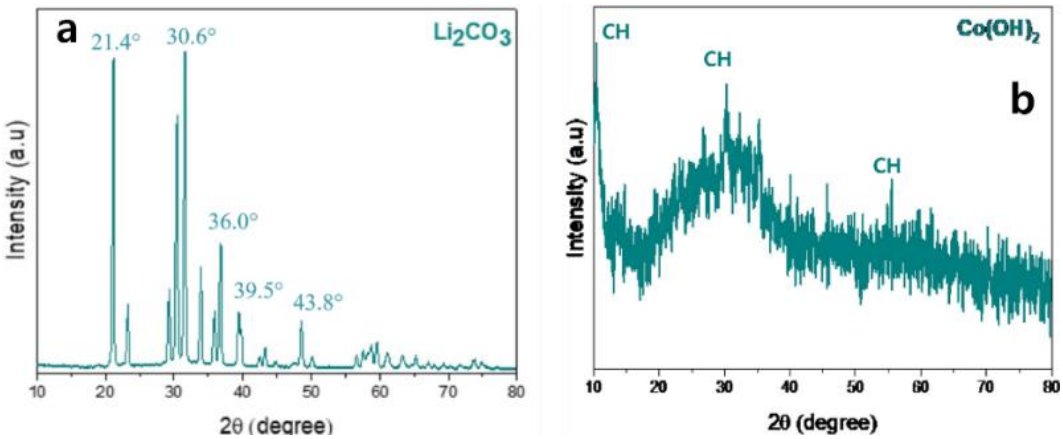


Figure 14. XRD pattern of (a) Li_2CO_3 showing distinct peaks indicative of its crystalline structure (b) $Co(OH)_2$ showing distinct diffraction peaks that suggest the structure and phase purity of the sample.

Fourier Transform Infrared Spectroscopy (FTIR):

The FTIR spectrum of lithium carbonate (Li_2CO_3) displays key peaks that confirm its structure. The peak at 1412 cm^{-1} corresponds to the asymmetric stretching vibrations of the carbonate ion, a crucial feature that confirms the presence of carbonate groups. The peak at 1087 cm^{-1} represents the symmetric stretching mode of the carbonate ion, indicating well-defined carbonate bonding. The peak at 859 cm^{-1} is attributed to the out-of-plane bending vibrations of the carbonate ion, reflecting the structural stability of Li_2CO_3. Similarly, the peak at 739 cm^{-1} is associated with the bending vibrations of the carbonate ion, further validating the vibrational properties of lithium carbonate.

The absence of additional peaks, such as those linked to hydroxyl groups or impurities, indicates the high purity of the sample. The observed spectral features match previously reported FTIR data for lithium carbonate, confirming its structural identity.

In addition, the peak at 1054 cm^{-1} in the FTIR spectrum is characteristic of vibrational modes in cobalt-based materials, representing Co–OH bonds. Peaks at 1374 cm^{-1} and 1636

cm^{-1} arise from the intercalation of OH^- and NO_3^- anions, respectively. This ion intercalation is a distinctive feature of $Co(OH)_2$, enhancing mass and electron transport between the catalyst and the electrolyte during electrocatalysis.

The peak at 3494 cm^{-1} corresponds to O–H stretching, which is characteristic of hydroxyl groups or water molecules. This suggests that the material is either hydrated or contains surface-bound water, as observed in $Co(OH)_2$. These peaks imply that $Co(OH)_2$ or its hydrated form may influence electrochemical properties by modifying surface chemistry and ion transport, which, in turn, affects energy storage performance in applications such as supercapacitors or batteries.

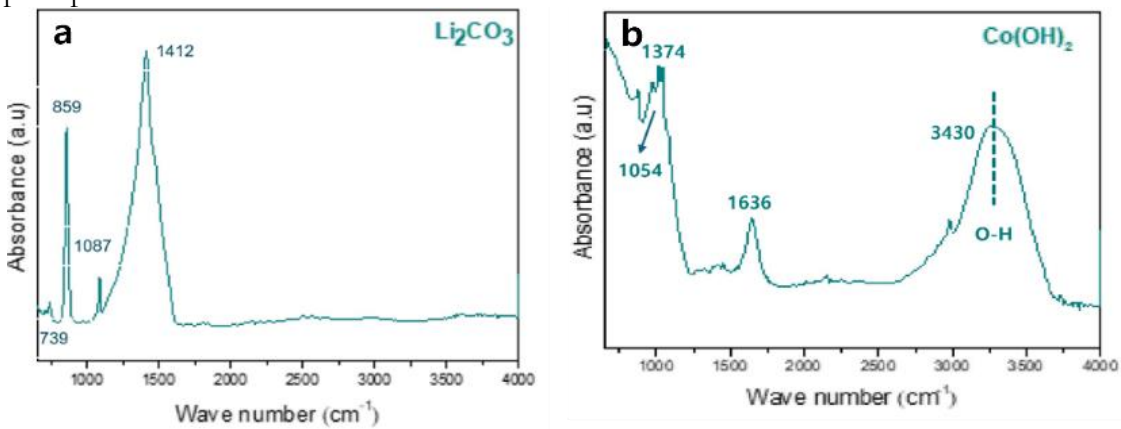


Figure 15. FTIR pattern of (a) Li_2CO_3 showing distinct peaks indicative of its crystalline structure and (b) $Co(OH)_2$ showing distinct diffraction peaks that suggest the crystalline structure and phase purity of the sample.

Electrochemical Impedance Spectroscopy (EIS):

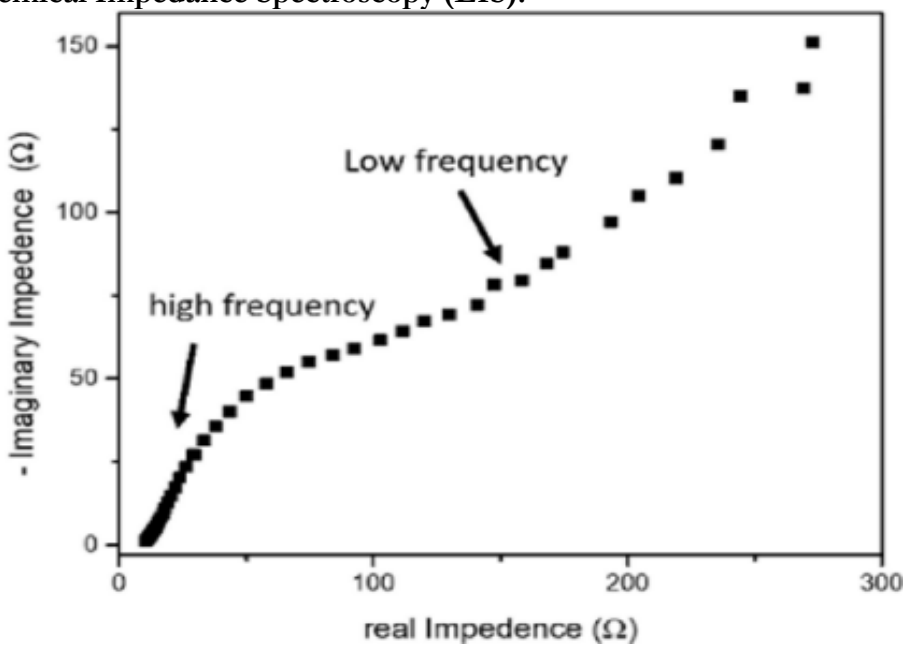


Figure 16. Nyquist plot of Cobalt oxide from EIS.

Nyquist EIS Curve:

The Nyquist EIS curve of cobalt oxide (Co_xO_y) shows a half-semicircle at high frequencies, indicating low charge transfer resistance (R_ct). This suggests efficient electron transfer at the electrode-electrolyte interface. Following the semicircle, the curve transitions into an upward diagonal line, which represents Warburg impedance, reflecting ion diffusion into the

material. This indicates a combination of capacitive and diffusion-controlled charge storage mechanisms.

The half-semicircle implies fast charge transfer kinetics, while the diagonal line indicates moderate ion diffusion limitations, typical of pseudocapacitive materials. Overall, this behavior demonstrates that cobalt oxide has high-rate capability, making it suitable for high-power energy storage applications, such as supercapacitors or hybrid capacitors, with good cycling stability and efficient electrochemical performance.

Cyclic Voltammetry (CV):

The cyclic voltammetry (CV) of cobalt oxide (Co_xO_γ) exhibits a rectangular shape, which is characteristic of pseudocapacitive behavior. In this case, charge storage occurs mainly through surface-controlled redox reactions, unlike the intercalation process observed in traditional battery materials. The absence of well-defined redox peaks suggests that the electrochemical reactions are fast and reversible, enabling rapid charge and discharge cycles.

This behavior is typical of materials that combine both capacitive and faradaic charge storage mechanisms, often seen in transition metal oxides like cobalt oxide. These materials can store charge electrostatically and through surface redox processes. The rectangular CV shape in Fig. 17a reflects high-rate capability and fast electron transfer, making cobalt oxide ideal for high-power applications such as supercapacitors or hybrid energy storage systems.

Additionally, the observed behavior indicates that cobalt oxide offers good cycling stability and long-term performance. The stable electrochemical properties suggest that recovered cobalt oxide can be effectively reused in subsequent cycles, providing a sustainable and cost-effective option for future energy storage technologies.

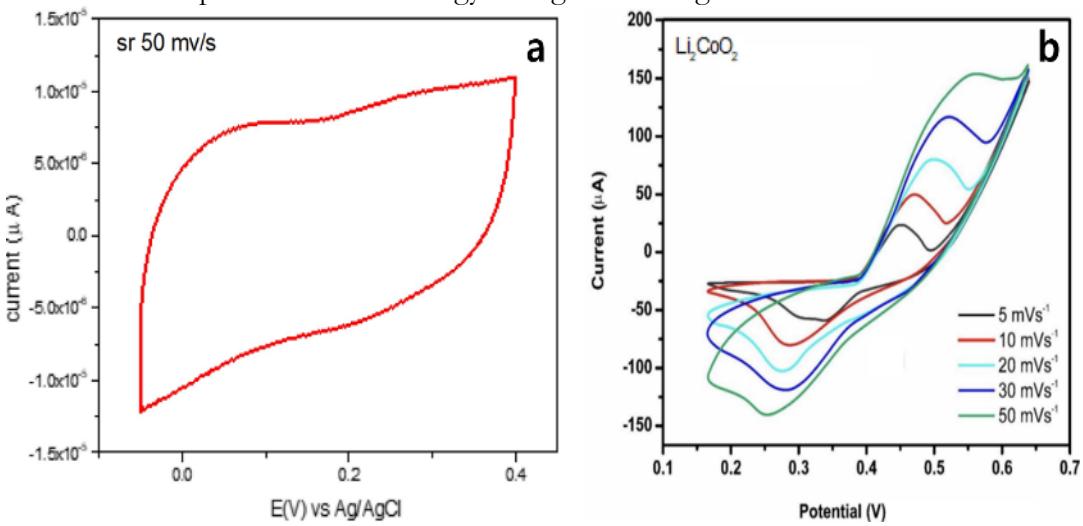


Figure 17. Cyclic Voltammetry of (a)Cobalt Oxide and (b) $LiCoO_2$ (LCO) active material

The cyclic voltammetry (CV) analysis of the recovered $LiCoO_2$ in Fig. 17b shows a distinctive profile, characterized by pronounced peaks near the potential window edges and a near-convergence of oxidation and reduction peaks in the central region. This convergence in the middle suggests highly reversible redox processes with minimal polarization, indicating efficient charge transfer kinetics. In contrast, the distinct peaks at the edges may reflect the activation of additional electrochemical processes or slight diffusion limitations at extreme potentials.

Overall, these features highlight a complex balance between surface-controlled pseudocapacitive behavior and diffusion-limited kinetics. This CV profile emphasizes the strong electrochemical performance of the recycled $LiCoO_2$, demonstrating its suitability for high-rate applications and showcasing the potential of recovered materials in sustainable energy storage technologies.

Discussion:

The results of this study demonstrate the feasibility of recycling lithium-ion batteries (LIBs) from spent laptop cells using a hydrometallurgical process with mild phosphoric acid leaching, achieving significant lithium (5.124%) and cobalt (42.143%) recovery rates. The successful synthesis of $LiCoO_2$ cathode material from recovered metals highlights the potential of this process in promoting closed-loop LIB recycling, aligning with global sustainability goals for resource conservation and electronic waste management. Compared to conventional sulfuric acid and hydrochloric acid leaching methods, the use of phosphoric acid offers a less corrosive, more environmentally friendly approach, reducing the formation of secondary pollutants and minimizing hazardous waste disposal challenges. These findings contribute to recent advancements in sustainable battery recycling technologies, supporting efforts to replace energy-intensive pyrometallurgical processes with cost-effective hydrometallurgical alternatives (Makuza et al., 2021).

One of the key advantages of this study is the comprehensive characterization of recovered materials, which confirms the structural integrity and electrochemical performance of regenerated $LiCoO_2$. XRD and Raman Spectroscopy results indicate that the synthesized cathode material retains a stable layered structure, similar to commercially available $LiCoO_2$, while Electrochemical Impedance Spectroscopy (EIS) and Cyclic Voltammetry (CV) confirm its favorable charge-transfer properties. However, while initial results suggest promising electrochemical behavior, long-term cycling stability tests are required to assess capacity retention and degradation mechanisms over extended charge-discharge cycles. Previous studies (Richa et al., 2017) have shown that recycled $LiCoO_2$ may exhibit slightly lower energy density and cycle life due to impurity incorporation during the recovery process. Future work should focus on optimizing purification steps to enhance material performance and compare the electrochemical stability of recovered $LiCoO_2$ against commercially synthesized cathodes.

Despite the positive outcomes, certain challenges remain regarding the scalability and economic feasibility of phosphoric acid leaching for large-scale LIB recycling. While phosphoric acid is less hazardous than strong mineral acids, its leaching efficiency for lithium is relatively lower compared to citric acid or oxalic acid-based processes, which have been reported to achieve higher lithium recovery rates with minimal environmental impact (Xie et al., 2021). Additionally, the study does not evaluate the energy consumption of the leaching, precipitation, and calcination steps, which is critical for determining the commercial viability of this method. Future research should incorporate a techno-economic analysis to assess the cost-effectiveness, reagent consumption, and energy demands associated with phosphoric acid leaching, as well as explore hybrid recycling approaches that combine mechanical, hydrometallurgical, and direct cathode regeneration techniques for maximum resource recovery.

Another crucial aspect that requires further exploration is the management of leachate residues and wastewater treatment. While hydrometallurgical methods produce fewer emissions than pyrometallurgy, they still generate waste streams containing dissolved metals, residual acids, and organic binders that require proper treatment before disposal. Previous research (Werner et al., 2020) emphasizes that leaching residues can contribute to secondary pollution if not adequately neutralized or processed for further metal recovery. Future studies should focus on developing closed-loop wastewater treatment systems that enable safe disposal or reuse of leachate solutions to minimize environmental impacts. Additionally, investigating the potential for recovering additional byproducts such as aluminum, copper, and graphite from spent batteries could further improve the sustainability and profitability of LIB recycling.

In conclusion, this study successfully demonstrates that phosphoric acid-based hydrometallurgical recycling is an effective method for lithium and cobalt recovery from spent LIBs, offering a sustainable alternative to conventional battery disposal. However, further optimizations are needed to enhance lithium recovery efficiency, improve the electrochemical

stability of regenerated cathode materials, and address environmental concerns related to wastewater treatment and process scalability. By integrating advanced purification techniques, optimizing reaction conditions, and conducting large-scale feasibility assessments, future research can contribute to the development of commercially viable LIB recycling solutions, supporting the transition toward a circular economy in battery materials management.

Limitations:

Hydrometallurgical recycling of lithium-ion batteries faces challenges related to scalability, cost, and environmental impact. Scalability is limited by complex multi-step processes, including leaching, separation, and purification, which must be tailored to accommodate variations in battery chemistry. High costs arise from the use of expensive reagents (e.g., acids and solvents), energy-intensive operations (such as heating and evaporation), and wastewater treatment to handle toxic effluents containing heavy metals and fluorides.

The recycling process generates chemical waste, including acidic/alkaline wastewater and metal-laden sludge, which requires proper disposal to prevent environmental harm. Technical challenges include lithium losses due to its solubility, difficulties in separating chemically similar metals (e.g., cobalt and nickel), and interference from battery binders (e.g., PVDF). Despite offering precise metal recovery and lower energy consumption compared to pyrometallurgy, these limitations underscore the need for innovative solutions, such as closed-loop recycling systems and bioleaching, to improve the efficiency and sustainability of the process.

Conclusion:

This study successfully developed a scalable and efficient recycling process for spent lithium-ion batteries sourced from laptops and notebooks, using mild phosphoric acid as a leaching agent. The process achieved notable recovery rates, yielding 5.124% lithium and 42.143% cobalt, which were subsequently converted into lithium carbonate and cobalt hydroxide. Comprehensive characterization confirmed the high purity and structural integrity of these recovered materials. Additionally, cobalt oxide extracted during the process was used to produce LCO (lithium cobalt oxide) active powder.

These findings demonstrate the feasibility of repurposing materials from spent batteries, offering a sustainable alternative to traditional waste disposal methods that often harm the environment. The purified lithium carbonate and cobalt hydroxide not only present opportunities for reuse in battery manufacturing but also help address resource scarcity amidst rising demand for lithium-ion batteries.

Ultimately, this research highlights the importance of developing efficient recycling technologies to support a circular economy in battery production. By recovering valuable materials, the approach addresses environmental concerns while enhancing the economic viability of recycling initiatives.

Looking ahead, the outlook for lithium-ion battery (LIB) recycling is promising but requires further innovation and collaboration to overcome existing challenges. Enhancing leach efficiency, reducing energy consumption, and improving recovery rates for valuable metals like lithium and cobalt will be crucial. The development of direct recycling methods that preserve cathode structures could significantly minimize environmental impact and energy use.

Moreover, integrating automation and AI in recycling operations can further streamline the process. Exploring alternative battery chemistries, such as lithium iron phosphate (LFP) or sodium-ion batteries, could also reduce dependence on critical materials. Policymakers have a key role to play by implementing regulations and incentives to promote recycling and circular economic practices.

Global collaboration and standardization of recycling protocols will be essential to address the growing problem of LIB waste on a larger scale. By advancing these areas, the industry can

work toward a more sustainable, resource-efficient future, ensuring that the increasing demand for energy storage solutions is met without compromising environmental sustainability.

Acknowledgement: We would acknowledge COMSATS University Islamabad, Lahore Pakistan for providing the environment to conduct the experimental work

Conflict of interest: There is no conflict of interest among the authors.

Project details. Nil

References:

[1] H. Pinegar and Y. R. Smith, "Recycling of End-of-Life Lithium Ion Batteries, Part I: Commercial Processes," *J. Sustain. Metall.*, vol. 5, no. 3, pp. 402–416, Sep. 2019, doi: 10.1007/S40831-019-00235-9/METRICS.

[2] J. C. C. Junying Xie, Kaiyou Huang, Zhenglin Nie, Wenyi Yuan, Xiaoyan Wang, Qingbin Song, Xihua Zhang, Chenglong Zhang, Jingwei Wang, "An effective process for the recovery of valuable metals from cathode material of lithium-ion batteries by mechanochemical reduction," *Resour. Conserv. Recycl.*, vol. 168, p. 105261, 2021, doi: https://doi.org/10.1016/j.resconrec.2020.105261.

[3] K. S. S. Shrinidhi Sambamurthy, Smita Raghuvanshi, "Environmental impact of recycling spent lithium-ion batteries," *Procedia CIRP*, vol. 98, pp. 631–636, 2021, doi: https://doi.org/10.1016/j.procir.2021.01.166.

[4] D. Y. Brian Makuza, Qinghua Tian a b, Xueyi Guo, Kinnor Chattopadhyay, "Pyrometallurgical options for recycling spent lithium-ion batteries: A comprehensive review," *J. Power Sources*, vol. 491, p. 229622, 2021, doi: https://doi.org/10.1016/j.jpowsour.2021.229622.

[5] Z. X. Xiao, Jiefeng, Li, Jia, "Recycling metals from lithium ion battery by mechanical separation and vacuum metallurgy," *J. Hazard. Mater.*, vol. 338, pp. 124–131, 2017, doi: https://doi.org/10.1016/j.jhazmat.2017.05.024.

[6] Z. X. Jiefeng Xiao, Bo Niu, "Highly efficient selective recovery of lithium from spent lithium-ion batteries by thermal reduction with cheap ammonia reagent," *J. Hazard. Mater.*, vol. 418, p. 126319, 2021, doi: https://doi.org/10.1016/j.jhazmat.2021.126319.

[7] W. Z. Leqi Lin, Zhongming Lu, "Recovery of lithium and cobalt from spent Lithium-Ion batteries using organic aqua regia (OAR): Assessment of leaching kinetics and global warming potentials," *Resour. Conserv. Recycl.*, vol. 167, p. 105416, 2021, doi: https://doi.org/10.1016/j.resconrec.2021.105416.

[8] M. R. T. Georgi-Maschler, B. Friedrich, R. Weyhe, H. Heegn, "Development of a recycling process for Li-ion batteries," *J. Power Sources*, vol. 207, pp. 173–182, 2012, doi: https://doi.org/10.1016/j.jpowsour.2012.01.152.

[9] K. Z. Mina Rezaei, Atiyeh Nekahi, Anil Kumar M R, Ameer Nizami, Xia Li, Sixu Deng, Jagjit Nanda, "A review of lithium-ion battery recycling for enabling a circular economy," *J. Power Sources*, vol. 603, p. 236157, 2025, doi: https://doi.org/10.1016/j.jpowsour.2024.236157.

[10] K. A. Li Li, Jennifer B. Dunn, Xiao Xiao Zhang, Linda Gaines, Ren Jie Chen a, Feng Wu, "Recovery of metals from spent lithium-ion batteries with organic acids as leaching reagents and environmental assessment," *J. Power Sources*, vol. 233, pp. 180–189, 2013, doi: https://doi.org/10.1016/j.jpowsour.2012.12.089.

[11] S. K. . Bhar, M (Bhar, Madhushri) ; Ghosh, S (Ghosh, Shuvajit) ; Krishnamurthy, S (Krishnamurthy, Satheesh) ; Kaliprasad, Y (Kaliprasad, Y.) ; Martha, SK (Martha, "A review on spent lithium-ion battery recycling: from collection to black mass recovery," *RSC Sustain.*, vol. 1, no. 5, pp. 1150–1167, 2023, doi: 10.1039/d3su00086a.

[12] T. M. Denis Werner, Urs Alexander Peuker, "Recycling Chain for Spent Lithium-Ion

Batteries," *Metals (Basel).*, vol. 10, no. 3, p. 316, 2020, doi: https://doi.org/10.3390/met10030316.

[13] M. Alipanah, A. K. Saha, E. Vahidi, and H. Jin, "Value recovery from spent lithium-ion batteries: A review on technologies, environmental impacts, economics, and supply chain," *Clean Technol. Recycl.*, vol. 1, no. 2, pp. 152–184, 2021, doi: 10.3934/CTR.2021008.

[14] G. G. Kirti Richa, Callie W. Babbitt, "Eco-Efficiency Analysis of a Lithium-Ion Battery Waste Hierarchy Inspired by Circular Economy," *J. Ind. Ecol.*, 2017, doi: https://doi.org/10.1111/jiec.12607.

An IoT Distributive SM Controller for Mitigation of Circulating Currents Among Sources in a Standalone DC Microgrid

Muhammad Sharjeel Sarwar[1], Muhammad Rashad[1], Tayybah Kiren[2], Nazam Siddique[3], Bilal Ishfaq Ahmed[4], Adeel Ahmed[1], Muhammad Wasif[3], Muhammad Jehanzeb Irshad[3]
[1]Department of Electrical Engineering (The University of Lahore (UOL), Lahore, Pakistan).
[2]Department of Computer Science (University of Engineering and Technology (UET), Lahore, Pakistan).
[3]Department of Electrical Engineering (University of Gujrat (UOG), Gujrat, Pakistan).
[4]Department of Informatics and Systems (University of Management and Technology (UMT), Lahore, Pakistan).
***Correspondence:** muhammad.rashad@ee.uol.edu.pk

Citation | Sarwar. M. S, Rashad. M, Kiren. T, Siddique. N, Ahmed. B. I, Ahmed. A, Wasif. M, Irshad. M. J, "An IoT Distributive SM Controller for Mitigation of Circulating Currents Among Sources in a Standalone DC Microgrid", IJIST, Special Issue. pp 274- 287, March 2025.
Received Feb 27, 2025 **Revised|** March 11, 2025 **Accepted|** March 17, 2025 **Published|** March 20, 2025.

Sources of similar or different power ratings are connected in parallel within the DC microgrid. During operation, these sources generate circulating currents along with their normal currents, which disrupt proper current sharing among power electronic converters based on their capacity. Consequently, voltage regulation across the system weakens. Additionally, the resistance of the connecting lines contributes to this imbalance in current distribution. To address circulating currents, droop controllers are commonly used. This method allows converters to share power according to their capacity without requiring internal communication. However, a major drawback of conventional droop control is that as output voltage decreases, the converter's output current increases linearly, leading to significant voltage fluctuations. As a result, droop control inherently involves a trade-off between voltage regulation and current sharing, making it impossible to optimize both simultaneously. To overcome this issue, this paper proposes a sliding mode (SM) controller implemented through an IoT-based distributed architecture. A system model is developed to evaluate its performance, and conditions for stability and existence are analyzed. MATLAB simulations provide detailed experimental results, demonstrating the effectiveness of the proposed technique.

Keywords: Droop Control; Circulating Currents; Voltage Regulation; Sliding Mode Controller; Existence Condition; Stability Condition.

Introduction:

In the modern era, microgrids are self-sufficient, small-scale power systems that can generate and consume electricity independently. They can also exchange power with other microgrids and the utility grid based on demand. From a design perspective, microgrids are classified as AC, DC, or Hybrid. In terms of control architecture, they can be categorized as decentralized, centralized, or distributed [1][2][3][4].

The development of microgrids, particularly in distributed generation, creates new opportunities for utilizing renewable energy sources (RESs). This advancement can play a crucial role in electrifying underdeveloped regions, helping them bridge the gap with developed and emerging nations. A reliable communication system is essential for microgrids, ensuring complete and bidirectional connectivity between resources while maintaining interoperability. Additionally, a new generation of peer-to-peer communication schemes is expected to enhance the efficiency and functionality of modern microgrids [5][6][7].

Microgrids are classified as AC or DC based on their operational format. AC microgrids integrate various energy sources and loads using an AC bus system. In these systems, DC-AC and AC-DC conversions are necessary to support DC loads. Similarly, DC microgrids operate comparably to AC microgrids, with the key difference being the use of a DC bus for interconnection [8].

DC microgrids are becoming increasingly popular due to their high efficiency, reliability, and ease of integrating renewable energy sources compared to AC microgrids. Additionally, DC systems eliminate issues such as skin effect, reactive power, frequency regulation, and synchronization, making them a highly dependable and promising solution for future energy needs [9]. Table 1 provides a brief comparison between AC and DC microgrids.

Table 1 Comparison of AC and DC microgrids

Factors	AC	DC
Converter requirement [6], [8], [10]	More Converters	Fewer Converters
Power Requirement [7][8]	Complex Power	Only Active Power
Power Factor Issues [7][8]	Yes	No
Stability [8], [11][10]	Low	High
Microgrid Controls [11][10]	Complex	Simple
Protection [7][8]	Low Cost	High Cost
Reactive Power Compensation [7]	Required	Not Required
Frequency Synchronization [8], [11][10]	Required	Not required

Precise current sharing and voltage regulation are essential in DC microgrids, especially when sources are connected in parallel. The current-sharing issue arises due to circulating currents, which result from differences in output voltage and cable resistance. These circulating currents must be minimized or eliminated to ensure proper current distribution. A well-designed control strategy should maintain system stability while achieving the desired performance.

Control schemes for DC microgrids fall into three categories: decentralized, centralized, and distributed. Some power flow control methodologies, based on hierarchical layers for control parameter flow and decision-making, are discussed in [11], while hierarchical control methods are explored in detail in [10]. Power electronic converters are key components in every DC microgrid, and their challenges are highlighted in [12].

A parallel DC-DC converter arrangement offers several advantages over a single high-capacity centralized converter. These benefits include increased reliability, improved efficiency, enhanced thermal management, lower maintenance, and reduced stress on individual converters, as the total load current is distributed among them [13]. However, despite these advantages, proper control strategies are necessary to regulate load voltage and

ensure precise load sharing among converters [14]. Uneven current distribution can cause overheating, potentially leading to system failure. Research shows that the main challenge in parallel-connected DC-DC converters is ensuring both voltage regulation and accurate current sharing. While centralized controllers can achieve these objectives, they require high-bandwidth communication, making them vulnerable to single-point failures. To overcome these limitations, a decentralized control architecture for DC microgrids is presented in [15]. Although it simplifies system design, it lacks information on parallel sources, preventing simultaneous load sharing and voltage regulation. Thus, decentralized control is not ideal for DC microgrid applications.

Current-sharing control methods are mainly categorized into droop control and active-sharing techniques. In islanded DC microgrids, droop control is the most commonly used method for stabilizing bus voltage when parallel converters operate without communication [16]. Droop techniques involve adding virtual resistance to the output voltage of all converters. As a result, each converter's output voltage decreases, increasing the load current. While simple and decentralized, droop control often leads to poor voltage regulation. Reducing droop gain improves voltage regulation but compromises accurate load sharing. Increasing droop gain beyond cable resistance can enhance current sharing, but this results in significant voltage deviations from no-load to full-load conditions.

In contrast, active current-sharing methods typically use a dual-loop control structure with an outer voltage control loop and an inner current control loop. These loops are decoupled based on frequency considerations [17]. However, variations in bandwidth and frequency separation can affect transient performance and overall system stability. A master-slave current-sharing control strategy for parallel power electronic converters is discussed in [18], where the master converter retains its characteristics while slave converters function as current-controlled current sources.

A cooperative distributed control strategy for voltage regulation and current sharing in parallel DC converters is presented in [19]. This method does not require knowledge of the number of converters. Another distributed control approach, which ensures average voltage regulation and current sharing through resistive-inductive electrical lines, is detailed in [20]. Distributed control allows simultaneous load sharing and voltage regulation, but conventional PI-based control techniques pose challenges. Classical controllers struggle with parameter tuning, adapting to load variations, and achieving stability under all conditions, making them less suitable for DC microgrid applications.

The rest of this paper is structured as follows: Section 2 formulates the equations for circulating currents and control architectures using voltage-only and cascaded voltage-current loops. Section 3 introduces the proposed SM controller and IoT-based distributed architecture, along with stability and existence conditions. Section 4 presents the results and discussion, while Section 5 concludes the paper.

Research Objectives:

The primary objective of this research is to minimize circulating currents among parallel-connected sources in a DC microgrid. The second objective is to achieve simultaneous load sharing and precise voltage regulation.

Novelty Statement:

Existing control techniques fail to achieve both load sharing and voltage regulation simultaneously. To overcome this challenge, this paper proposes an IoT-based distributed architecture using a robust SM controller. In the SM controller, voltage and current loops are processed simultaneously. To evaluate its performance, a system model is developed, and the existence and stability conditions of the SM controller are analyzed. MATLAB simulation results are presented to demonstrate the effectiveness of the proposed technique.

Material and Methods:

System Modelling:

The basic structure of the DC microgrid is shown in Fig. 1. All RESs are either direct DC sources or DC-compatible. The main utility grid and wind turbines generate AC power, which is connected to the DC bus through an AC-DC converter [21]. A single arrow indicates unidirectional power flow for loads, PV arrays, and fuel cells, while battery energy storage systems (BESS) and the conventional utility grid exhibit bidirectional power flow, meaning they can both supply and consume power.

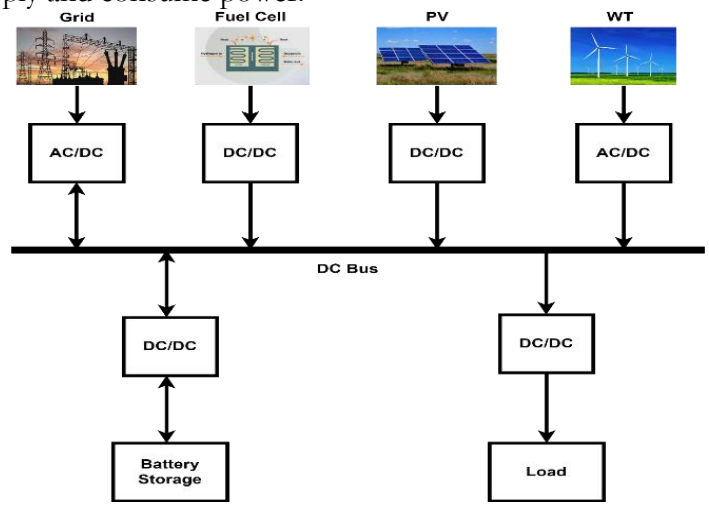


Figure 1 Typical arrangement of a DC microgrid.

A DC microgrid with two parallel sources connected to a load through a DC bus is shown in Fig. 2. Here, R_1 and R_2 represent the cable resistances, while i_{dc1} and i_{dc2} are the currents shared by the two sources, which have terminal voltages v_{dc1} and v_{dc2}, respectively.

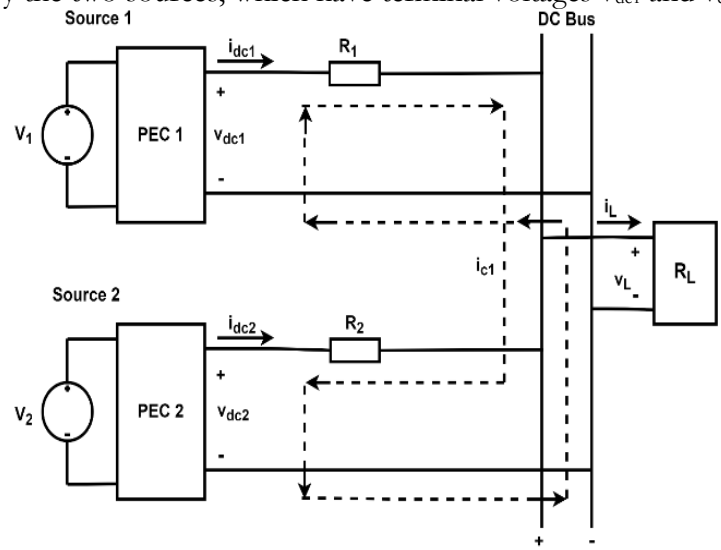


Figure 2 Two sources connected in a parallel configuration.

Formulation of Circulating Current Equations:

A steady-state equivalent network of the DC microgrid from Fig. 2 is illustrated in Fig. 3. Here, v_{dc1} and v_{dc2} represent the terminal voltages of sources 1 and 2, while the interconnecting cables are modeled by resistances R_1 and R_2, respectively. Applying Kirchhoff's Voltage Law (KVL) to loop 1, as shown in Fig. 3, the equations for v_{dc1} and v_{dc2} can be expressed as follows in (1) and (2):

$$v_{dc1} = i_{dc1}R_1 + i_L R_{Load} \quad (1)$$

$$v_{dc2} = i_{dc2}R_2 + i_L R_{Load} \quad (2)$$

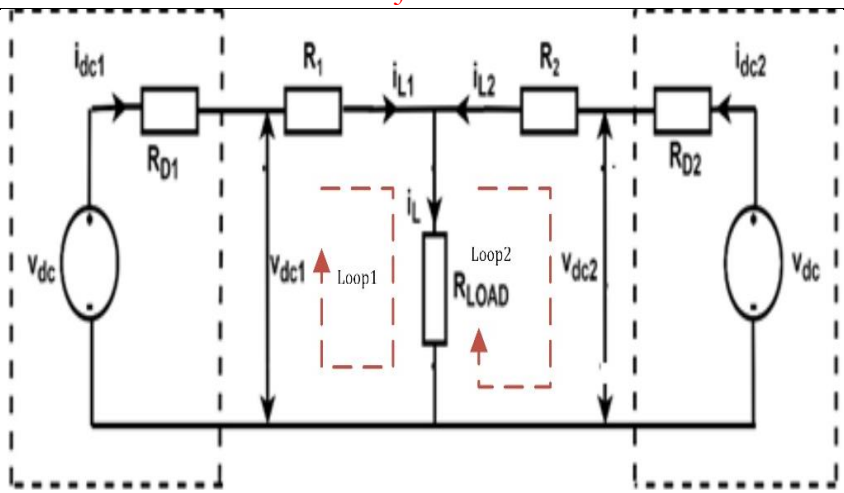


Figure 3 Steady-state equivalent circuit of two parallel connected sources.

Here, i_{dc1} and i_{dc2} are the currents supplied by sources 1 and 2, respectively, while i_t represents the load current. These currents are mathematically related as:

$$i_L = i_{L1} + i_{L2} \quad (3)$$

$$i_{dc1} = i_{L1} + i_{c1} \quad (4)$$

$$i_{dc2} = i_{L2} + i_{c2} \quad (5)$$

Here, i_{c1} and i_{c2} are the circulating currents of sources 1 and 2, respectively. These currents have the same magnitude but flow in opposite directions [22]. To determine i_{dc1} and i_{dc2}, equations (3), (4), and (5) are substituted into (1) and (2). After simplification, i_{dc1} and i_{dc2} can be expressed in the form of equations (6) and (7).

$$i_{dc1} = \alpha_1 v_{dc1} - \beta v_{dc2} \quad (6)$$

$$i_{dc2} = \alpha_2 v_{dc2} - \beta v_{dc1} \quad (7)$$

Were,

$$\alpha_1 = \frac{R_2 + R_{LOAD}}{R_1R_2 + R_2R_{LOAD} + R_1R_{LOAD}} \quad (8)$$

$$\alpha_2 = \frac{R_1 + R_{LOAD}}{R_1R_2 + R_2R_{LOAD} + R_1R_{LOAD}} \quad (9)$$

$$\beta = \frac{R_{LOAD}}{R_1R_2 + R_2R_{LOAD} + R_1R_{LOAD}} \quad (10)$$

The product R_1R_2 in equations (8), (9), and (10) can be neglected, as its value is very small compared to the load resistance R_{LOAD}. Similarly, to determine i_{c1} and i_{c2}, equations (6) and (7) are substituted into equations (4) and (5). After simplification, i_{c1} and i_{c2} can be expressed in the form of equation (11).

$$i_{c1} = -i_{c2} = \frac{v_{dc1} - v_{dc2}}{R_1 + R_2} = \frac{i_{dc1}R_1 - i_{dc2}R_2}{R_1 + R_2} \quad \text{if } (R_1 \neq R_2)$$

$$= \frac{i_{dc1} - i_{dc2}}{2} \quad \text{if } (R_1 = R_2) \quad (11)$$

Substituting equation (11) into equations (4) and (5) gives equations (12) and (13).

$$i_{dc1} = \frac{R_2 v_{dc2}}{R_1R_{LOAD} + R_2R_{LOAD}} + \frac{v_{dc1} - v_{dc2}}{R_1 + R_2} \quad (12)$$

$$i_{dc2} = \frac{R_1 v_{dc1}}{R_1R_{LOAD} + R_2R_{LOAD}} - \frac{v_{dc2} - v_{dc1}}{R_1 + R_2} \quad (13)$$

In equations (12) and (13), the first part is the load current, and the second part is the

circulating current. Similarly, the analysis of the circulating current for any number of parallel-connected sources can be extended.

In equations (12) and (13), the first term represents the load current, while the second term corresponds to the circulating current. Similarly, this circulating current analysis can be extended to any number of parallel-connected sources.

Control Methods of DC Microgrid:

Several control methods in the literature use PI controllers to mitigate circulating currents and regulate voltage in DC microgrids. A simple voltage control loop for a source-j is shown in Fig. 4. Here, the output voltage V_{oj} is compared with the reference voltage V_j^{ref}, and the error is processed by a PI controller to generate a Pulse Width Modulation (PWM) signal for regulating the Power Electronic (PE) converter. However, this method is only suitable for voltage regulation and does not minimize circulating currents.

A cascaded voltage and current loop method using a PI controller is shown in Fig. 5. In this approach, the voltage error regulates the output voltage, while the current error controls the current shared by the sources. However, since these control techniques are linear, they cannot ensure stability in all operating conditions, making it difficult to minimize circulating currents effectively.

To overcome this limitation, the next section introduces a nonlinear SM control technique to minimize circulating currents among sources.

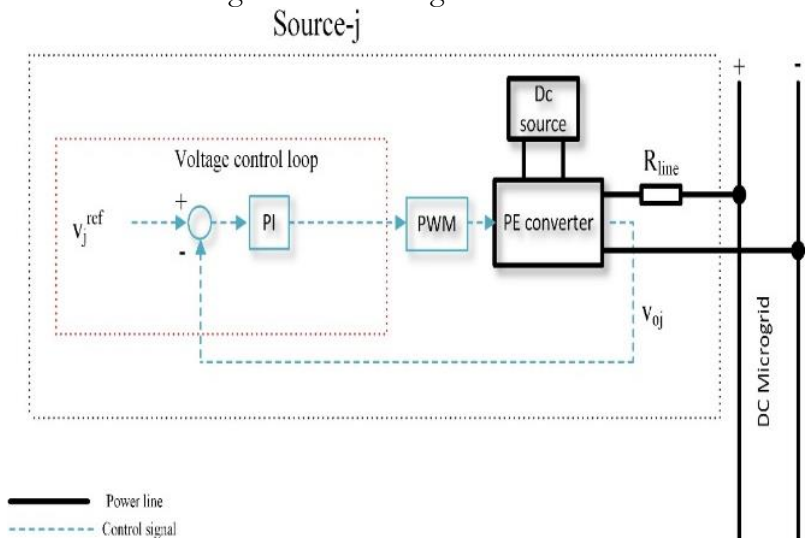


Figure 4 PI control of DC microgrid via voltage loop.

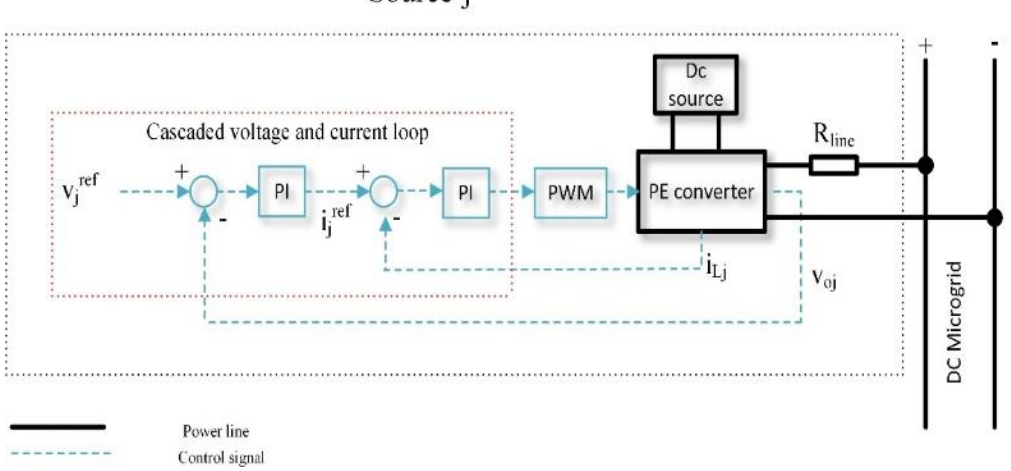


Figure 5 PI control of DC microgrid via cascaded voltage and current loop.

Proposed Sliding Mode Control:

The Sliding Mode (SM) controller is a nonlinear control method designed specifically for variable-structure systems, ensuring robustness and stability under all operating conditions. Fig. 6 illustrates the block diagram of a DC microgrid controlled using the SM controller.

In SM operation, both voltage and current loops function in parallel, enabling a fast dynamic response. The switched differential equations of the Power Electronic (PE) buck converter, which describe the system's state dynamics, are presented in equations (14) and (15). These equations form the basis for designing the SM controller.

$$\frac{dv_o}{dt} = \frac{i_L - i_{Line}}{C} \quad (14)$$

$$\frac{di_L}{dt} = \frac{uv_i - v_o}{L} \quad (15)$$

Here, i_L and i_{Line} Represent the inductor current and the connecting line current, respectively. Similarly, v_i and v_o Denote the input and output voltages. The capacitance and inductance are represented by C and L, respectively.

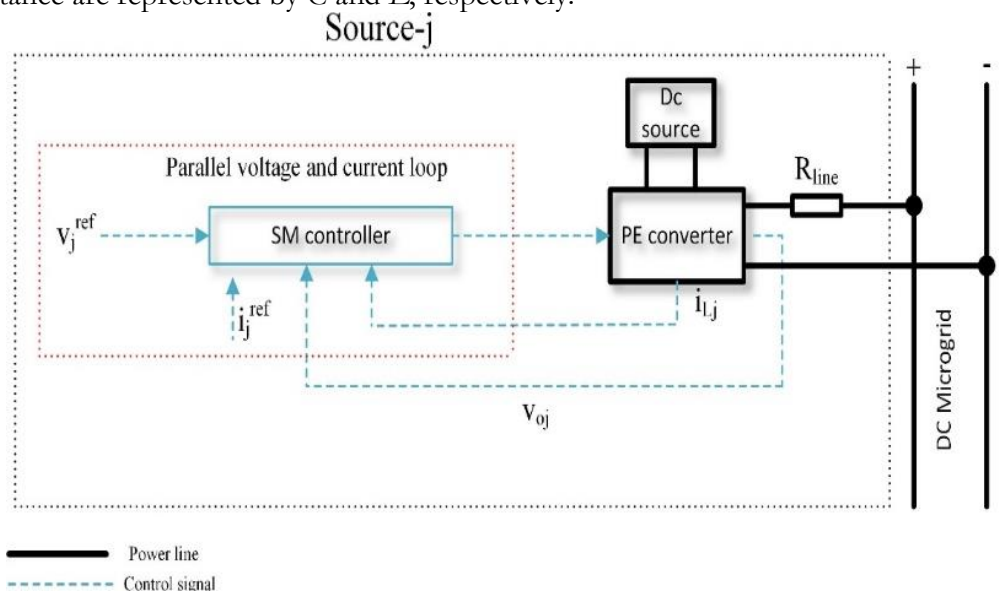


Figure 6 SM controller for parallel-connected DC microgrid.

SM Controller Design:

The sliding manifold in this paper is designed using the DC voltage error and converter current error. The SM controller identifies circulating currents by analyzing these errors. The proposed sliding surface is mathematically expressed in equation (16), where:

- $e_v = v_{ref} - v_o$ Represents the DC voltage error.
- $e_i = i_{ref} - i_L$ Represents the converter's current error.
- α and β are the sliding surface constraints.

$$\varphi = \alpha e_v + \beta e_i \quad (16)$$

The differentiation of the sliding function provides insight into the system's dynamic behavior and control response. It is mathematically represented in equation (17), showing how the sliding function evolves.

$$\frac{d\varphi}{dt} = -\alpha \frac{dv_o}{dt} - \beta \frac{di_L}{dt} \quad (17)$$

By substituting the derivatives of the output voltage and load current from equations (14) and (15) into equation (17), we obtain equation (18), which expresses the dynamics of the sliding function in terms of system parameters.

$$\frac{d\varphi}{dt} = \frac{-\alpha(i_L - i_{Line})}{C} - \frac{\beta(uv_i - v_o)}{L} \quad (18)$$

The equation (18) is used to define the transversality condition, which ensures that the sliding mode controller effectively influences the system's dynamics. This condition is mathematically expressed as:

$$\frac{d\varphi}{du} \neq 0 \quad (19)$$

This guarantees that the system does not remain in a singularity where control actions become ineffective.

The reachability condition ensures that the system states always move toward and remain on the sliding surface. Mathematically expressed as:

$$\varphi \frac{d\varphi}{dt} < 0$$

This ensures that any deviation from the sliding surface is corrected, forcing the system toward the desired operating point and maintaining stability.

Transversality Condition:

The transversality condition ensures that the control input appears explicitly in the sliding surface derivative, allowing the controller to influence system dynamics. Mathematically, this condition can be expressed as:

$$\frac{d\varphi}{dt} = f(x) + g(x)u$$

Where:

- *f(x)* represents the system dynamics without control input.
- *g(x)u* represents the influence of the control input u.

This guarantees that the controller can effectively drive the system states to the sliding surface and maintain stability. Substituting the value of $\frac{d\varphi}{dt}$ From equation (18) into equation (19) results in the equation (18).

$$\frac{d\frac{d\varphi}{dt}}{du} = \frac{-\beta}{L} v_i \quad (20)$$

The result in equation (20) confirms that the control input explicitly appears in the system dynamics, ensuring that the modeled system is controllable.

Reachability Condition:

The reachability conditions ensure that the system trajectories always move toward the sliding surface and do not diverge away. This condition is typically expressed using the Lyapunov stability approach. Mathematically given as:

$$\lim_{\varphi \to 0^-} \frac{d\varphi}{dt}\Big|_{u=1} > 0 \quad \text{and} \quad \lim_{\varphi \to 0^+} \frac{d\varphi}{dt}\Big|_{u=0} < 0 \quad (21)$$

This inequality ensures that the system state moves toward the sliding surface.

Where:

- If $\varphi > 0$ and $\frac{d\varphi}{dt} < 0$, the system must decrease toward the surface.
- If $\varphi < 0$ and $\frac{d\varphi}{dt} > 0$, the system must increase toward the surface.

For the control action *u*:

- *u = 1* when the system state is below the sliding surface.
- *u = 0* when the system state is above the sliding surface.

This switching action forces the system trajectory toward the sliding manifold, ensuring robust and stable operation.

Substituting the equation (17) in (21), the following two relations (22) and (23) can be

written as:

$$\lim_{\varphi \to 0^-} \frac{d\varphi}{dt} = \frac{\alpha(i_{Line} - i_{L)}}{C} + \frac{\beta(v_0 - v_i)}{L} > 0 \quad (22)$$

$$\lim_{\varphi \to 0^+} \frac{d\varphi}{dt} = \frac{\alpha(i_{Line} - i_{L)}}{C} + \frac{\beta(v_0)}{L} < 0 \quad (23)$$

Both equations (22) and (23) need to be satisfied to ensure the reachability conditions for the system under consideration.

Sliding Mode Dynamics:

In the Laplace domain, the differential equations (14) and (15) impose closed-loop features of the system. Taking the Laplace transform of the equation (14)

$$s.v_o(s) = \frac{i_L(s) - i_{Line}(s)}{C} \quad (24)$$

After simplifying, equation (24) results in equation (25).

$$i_L(s) = i_{Line}(s) + C.s.v_o(s) \quad (25)$$

Similarly, taking the Laplace transform of equation (16) results in (26).

$$\alpha[v_{ref}(s) - v_0(s)] + \beta[i_{ref}(s) - i_{Line}(s)] \quad (26)$$

Substituting the value of $i_L(s)$ From equation (25) and equating it to zero results in equation (27).

$$v_o(s)[\alpha + \beta.C.s] = \alpha v_{ref}(s) + \beta i_{ref}(s) - \beta i_{Line}(s) \quad (27)$$

In the Laplace domain, the equation (27) expresses the closed-loop dynamic behavior of the system. Now as it is evident that the load imposes the constant reference value, the final relation can be written as:

$$\frac{v_o(s)}{i_{Line}(s)} = \frac{-\beta}{\alpha + \beta.C.s} \quad (28)$$

Equation (28) represents the closed-loop dynamics of the DC microgrid system in the Laplace domain.

Proposed Distributive Architecture:

This paper proposes an IoT-based distributive control architecture for the minimization of circulating current among parallel connected sources and ensuring voltage regulation in a DC microgrid. The proposed architecture is shown in Fig. 7. Each source controller communicates with the other source controller through an IoT communication link, and in this way, the value of the current shared by each source (per unit) is communicated to all the sources. Based on this knowledge, the controller of each source calculates the average current value shared by all the sources which is given in equation (29).

$$i_j^{avg} = \frac{\sum_{j=1}^{n} i_j^{pu}}{n} \quad (29)$$

Here i_j^{pu} Is the source-j current (in per unit)? As the load changes, the deterioration in voltage regulation is restored as the new values of current are communicated between the sources. A flow diagram of the complete process is shown in Figure. 8.

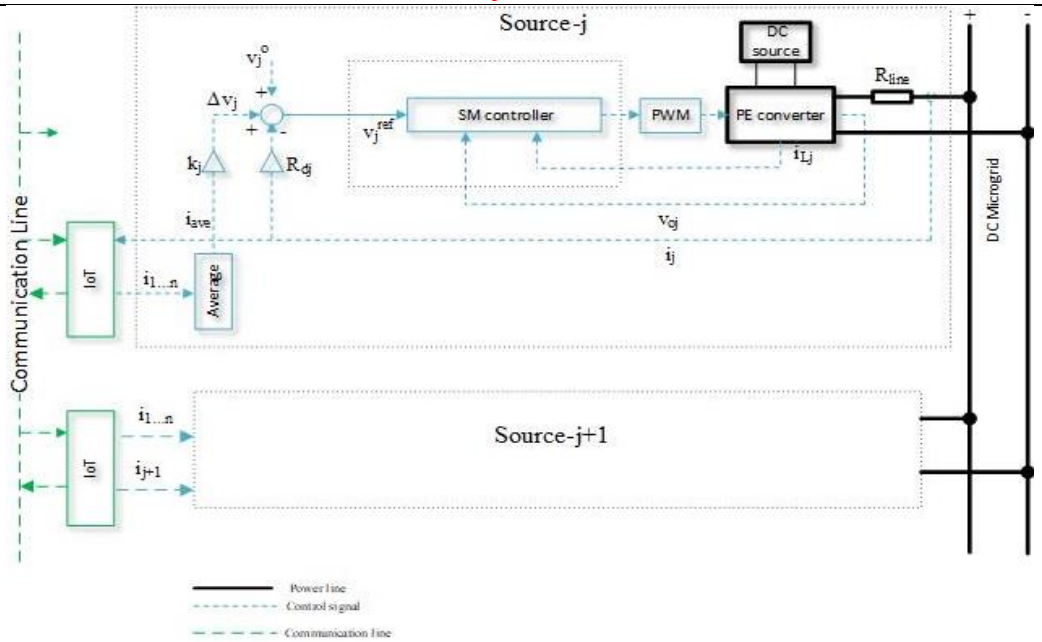


Figure 7 SMC with distributive architecture for parallel-connected DC microgrid.

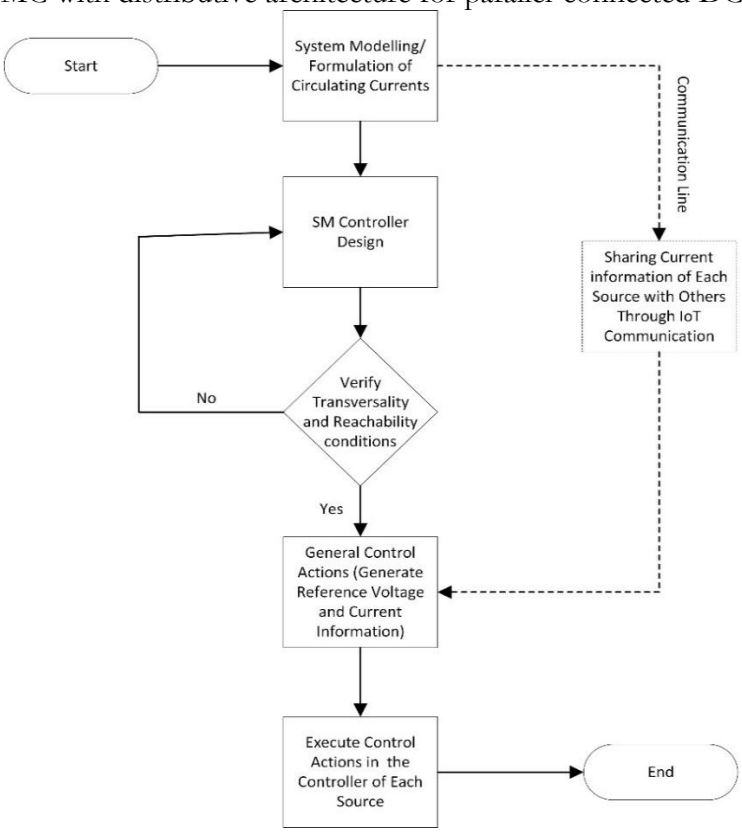


Figure 8 Complete Flow Diagram of the Methodology.

Result and Discussion:

The results from the MATLAB/Simulink simulations highlight the limitations of conventional PI control techniques in managing circulating currents and ensuring proper current sharing among sources. The system's efficiency is evaluated by first simulating a DC microgrid with only two sources with specifications given in Table 2 utilizing conventional PI-based control methods, including voltage loop and cascaded voltage-current loop control, and results are summarized below.

Table 2 Parameters of DC microgrid

Item	Parameter	Value
Converter 1	Inductor, L1	110mH
	Capacitance, C1	500 µF
	Proportional Gain Kp1	10
	Integral Gain Ki1	0.1
Converter 2	Inductor, L2	110 mH
	Capacitance, C2	500 µF
	Proportional Gain Kp2	10
	Integral Gain Ki2	0.1
Cable 1 Resistance	R1	0.25 Ω
Cable 2 Resistance	R2	1.25 Ω
Load	RL	22 Ω
Input Voltage	Vin	200 V
Nominal Voltage	Vout	120 V

1. **Voltage Loop PI Control:**

- Significant circulating currents (1.7A) arise due to uneven current distribution.
- Converter 1 supplies 4.355A, while converter 2 provides only 0.972A, disrupting the load-sharing balance.

2. **Cascaded Voltage & Current PI Control:**

- Some improvement in circulating current reduction (1.6A), but still not ideal.
- Converter 1 supplies 4.302A, while converter 2 contributes only 1.084A.

The results mentioned above highlight that both methods fail to minimize circulating currents among sources. Additionally, a trade-off between current sharing and voltage regulation is observed, showing that traditional methods fail to optimize both simultaneously.

To compare the performance of the proposed IoT-based distributive SM controller, the same DC microgrid is simulated utilizing the proposed method, and the results of the performance comparison are given below.

Performance Comparison:

1. **PI-Based Voltage Control:**

- Circulating current: 1.7A (high).
- Unequal sharing: 4.355A (Converter 1) vs. 0.972A (Converter 2).

2. **PI-Based Cascaded Voltage & Current Control:**

- Circulating current: 1.6A (slightly improved).
- Unequal sharing: 4.302A (Converter 1) vs. 1.084A (Converter 2).

3. **Proposed IoT-Based SM Control:**

- Circulating current: 0.03941A (almost negligible).
- Equal sharing: 2.721A (Converter 1) vs. 2.642A (Converter 2).

The results validate that the proposed SM controller outperforms traditional PI-based methods by:

1. Minimizing Circulating Currents – SM control reduces circulating currents to nearly zero.
2. Ensuring Balanced Load Sharing – Both sources contribute proportionally.
3. Maintaining Excellent Voltage Regulation – Stable operation even with variations in line resistance.

Performance Comparison:

Control Method	Circulating Current	Current Sharing	Voltage Regulation
PI Voltage Loop	High (1.7A)	Unbalanced	Within Limits
Cascaded PI Control	Moderate (1.6A)	Still Unbalanced	Within Limits

Proposed SM Controller	Minimal (0.03941A)	Balanced	Excellent

Thus, the distributive SM controller is an efficient solution for mitigating circulating currents, improving current-sharing accuracy, and maintaining precise voltage regulation in DC microgrids.

Conclusion:

To address the ongoing energy crisis, RESs can be integrated into DC or AC microgrids based on specific requirements. DC microgrids hold significant promise due to their advantages, such as the absence of skin effect, reactive power complications, frequency synchronization issues, and lower inter-energy conversion losses.

A critical challenge in DC microgrids is ensuring precise current sharing among parallel-connected Distributed Energy Resources (DERs). Variations in output voltage, converter currents, and cable resistances can lead to circulating currents, which degrade overall voltage regulation and system efficiency.

To address this issue, this paper proposes an IoT-based distributive SM control architecture. The system's efficiency is evaluated by first simulating conventional PI-based control methods, including voltage loop and cascaded voltage-current loop control. The results highlight a trade-off between current sharing and voltage regulation, showing that traditional methods fail to optimize both simultaneously.

In contrast, the proposed SM controller effectively minimizes circulating currents, ensures a proper current-sharing ratio, and maintains excellent voltage regulation. The results demonstrate the superiority of the proposed technique, making it a robust solution for enhancing stability and efficiency in DC microgrids.

Limitations and Future Work:

SM controller is a robust control technique that is used for variable structured systems. The limitation of the SM technique is the infinite frequency requirement, which is practically not possible. So, it produces chattering phenomena during the sliding surface.

For recommendations in the future, there is a need to explore the SM techniques used for chattering reduction. Additionally, to implement SM controllers in hardware, modern high-processing tools are recommended to be explored for the said application.

Acknowledgment: The authors would like to especially thank The University of Lahore (UOL) and the University of Gujrat (UOG) for supporting and providing the environment for the research.

Author's Contribution: Muhammad Sharjeel Sarwar contributed to the literature review and system design. Muhammad Rashad contributed to the analysis and wrote the paper. Tayybah Kiren, Nazam Siddique, Bilal Ishfaq Ahmed, and Adeel Ahmed contributed to data collection and analysis tools. Muhammad Wasif and Muhammad Jehanzeb Irshad contributed to the presentation of the data and results.

Conflict of Interest: The authors declare no conflict of interest in publishing this manuscript.

Project details: This research study received no external funding. All the resources consumed were internally available.

References:

[1] Rinkesh, "Causes, Effects and Solutions to Global Energy Crisis - Conserve Energy Future," Conserve Energy Future. Accessed: Mar. 12, 2025. [Online]. Available: https://www.conserve-energy-future.com/causes-and-solutions-to-the-global-energy-crisis.php

[2] R. S. Manieniyan V, Thambidurai Muthuvelan, "STUDY ON ENERGY CRISIS AND THE FUTURE OF FOSSIL FUELS," Proceedings of SHEE 2009. Accessed: Mar. 12, 2025. [Online]. Available: https://www.researchgate.net/publication/267439286_STUDY_ON_ENERGY_CRISIS_AND_THE_FUTURE_OF_FOSSIL_FUELS

[3] R.E.H. Sims, "Renewable energy: a response to climate change," *Sol. Energy*, vol. 76, no. 1–3, pp. 9–17, 2004, doi: https://doi.org/10.1016/S0038-092X(03)00101-4.

[4] Ibrahim Dincer, "Renewable energy and sustainable development: a crucial review," *Renew. Sustain. Energy Rev.*, vol. 4, no. 2, pp. 157–175, 2000, doi: https://doi.org/10.1016/S1364-0321(99)00011-8.

[5] L. Xuan and S. Bin, "Microgrids - An integration of renewable energy technologies," *2008 China Int. Conf. Electr. Distrib. CICED 2008*, 2008, doi: 10.1109/CICED.2008.5211651.

[6] M. Shahbazitabar, H. Abdi, H. Nourianfar, A. Anvari-Moghaddam, B. Mohammadi-Ivatloo, and N. Hatziargyriou, "An Introduction to Microgrids, Concepts, Definition, and Classifications," *Power Syst.*, pp. 3–16, 2021, doi: 10.1007/978-3-030-59750-4_1.

[7] B. A. A. PCS183032, "PWM Based Fixed Frequency Equivalent SM Controller for Stability of DC Microgrid System," *J. Eng. Res.*, Accessed: Mar. 12, 2025. [Online]. Available: https://www.academia.edu/101842265/PWM_Based_Fixed_Frequency_Equivalent_SM_Controller_for_Stability_of_DC_Microgrid_System

[8] M. Rashad, M. Ashraf, A. I. Bhatti, and D. M. Minhas, "Mathematical modeling and stability analysis of DC microgrid using SM hysteresis controller," *Int. J. Electr. Power Energy Syst.*, vol. 95, pp. 507–522, 2018, doi: https://doi.org/10.1016/j.ijepes.2017.09.001.

[9] J. E. M. Michele Cucuzzella, Jacquelien M. A. Scherpen, "Microgrids control: AC or DC, that is not the question," *EPJ Web Conf*, vol. 310, p. 20, 2024, doi: https://doi.org/10.1051/epjconf/202431000015.

[10] P. W. Marcos Gomez-Redondo, Marco Rivera, Javier Muñoz, "A Systematic Literature Review on AC Microgrids," *Designs*, vol. 8, no. 4, p. 77, 2024, doi: https://doi.org/10.3390/designs8040077.

[11] V. A. Jaynendra Kumar, Anshul Agarwal, "A review on overall control of DC microgrids," *J. Energy Storage*, vol. 21, pp. 113–138, 2019, doi: https://doi.org/10.1016/j.est.2018.11.013.

[12] U. R. Muhammad Rashad, "Equivalent SM Controller for Load-Sharing and Dynamic Performance in a DC Microgrid Application," *Math. Probl. Eng.*, 2022, doi: https://doi.org/10.1155/2022/9700915.

[13] V. J. Thottuvelil and G. C. Verghese, "Analysis and control design of paralleled DC/DC converters with current sharing," *Conf. Proc. - IEEE Appl. Power Electron. Conf. Expo. - APEC*, vol. 2, pp. 638–646, 1997, doi: 10.1109/APEC.1997.575646.

[14] Y. Huang and C. K. Tse, "Circuit theoretic classification of parallel connected dc-dc converters," *IEEE Trans. Circuits Syst. I Regul. Pap.*, vol. 54, no. 5, pp. 1099–1108, 2007, doi: 10.1109/TCSI.2007.890631.

[15] S. Anand, B. G. Fernandes, and J. M. Guerrero, "Distributed control to ensure proportional load sharing and improve voltage regulation in low-voltage DC microgrids," *IEEE Trans. Power Electron.*, vol. 28, no. 4, pp. 1900–1913, Apr. 2013, doi: 10.1109/TPEL.2012.2215055.

[16] A. P. N. Tahim, D. J. Pagano, E. Lenz, and V. Stramosk, "Modeling and Stability Analysis of Islanded DC Microgrids under Droop Control," *IEEE Trans. Power Electron.*, vol. 30, no. 8, pp. 4597–4607, 2015, doi: 10.1109/TPEL.2014.2360171.

[17] S. Luo, Z. Ye, R. L. Lin, and F. C. Lee, "Classification and evaluation of paralleling methods for power supply modules," *PESC Rec. - IEEE Annu. Power Electron. Spec. Conf.*, vol. 2, pp. 901–908, 1999, doi: 10.1109/PESC.1999.785618.

[18] J. Rajagopalan, K. Xing, Y. Guo, F. C. Lee, and B. Manners, "Modeling and dynamic analysis of paralleled dc/dc converters with master-slave current sharing control," *Conf.*

Proc. - IEEE Appl. Power Electron. Conf. Expo. - APEC, vol. 2, pp. 678–684, 1996, doi: 10.1109/APEC.1996.500513.

[19] M. S. Sadabadi, N. Mijatovic, J. F. Tregouet, and T. Dragicevic, "Distributed Control of Parallel DC-DC Converters Under FDI Attacks on Actuators," *IEEE Trans. Ind. Electron.*, vol. 69, no. 10, pp. 10478–10488, Oct. 2022, doi: 10.1109/TIE.2021.3123613.

[20] J. G. Sebastian Trip, Renke Han, Michele Cucuzzella, Xiaodong Cheng, Jacquelien Scherpen, "Distributed Averaging Control for Voltage Regulation and Current Sharing in DC Microgrids: Modelling and Experimental Validation," *IFAC-PapersOnLine*, vol. 21, no. 23, pp. 242–247, 2018, doi: https://doi.org/10.1016/j.ifacol.2018.12.042.

[21] B. A. A. Muhammad Rashad, Uzair Raoof, Muhammad Ashraf, "Proportional Load Sharing and Stability of DC Microgrid with Distributed Architecture Using SM Controller," *Math. Probl. Eng.*, 2018, doi: https://doi.org/10.1155/2018/2717129.

[22] A. Nawaz, J. Wu, and C. Long, "Mitigation of circulating currents for proportional current sharing and voltage stability of isolated DC microgrid," *Electr. Power Syst. Res.*, vol. 180, p. 106123, 2020, doi: https://doi.org/10.1016/j.epsr.2019.106123.

Sequestration of Carbon Dioxide via Mineral Carbonation to Produce Magnesium Carbonate: A Design Study

Meerab Yousuf, Abdul Basit, Abdul Basit
Department of Chemical Engineering and Technology, University of Gujrat.
*Correspondence: Meerab Yousuf, meerabyousuf05@gmail.com

Citation | Yousuf. M, Basit. A, Basit. A, "Sequestration of Carbon Dioxide via Mineral Carbonation to Produce Magnesium Carbonate: A Design Study", IJIST, Special Issue pp 288-303, March 2025

Received | Feb 28, 2025 **Revised |** March 12, 2025 **Accepted |** March 18, 2025 **Published |** March 21, 2025.

The rapid increase in atmospheric carbon dioxide (CO_2) due to industrialization and fossil fuel combustion has raised significant concerns about global warming. Carbon capture and storage (CCS) is a crucial technology for reducing greenhouse gas (GHG) emissions. This study presents the design of a mineral carbonation plant capable of sequestering 30 tons of CO_2 per day to produce magnesium carbonate ($MgCO_3$) using olivine as the feedstock.

The process follows an ex-situ carbonation approach, where a mineral slurry reacts with CO_2 under controlled conditions. The plant design includes the development of key equipment such as a reactor, heat exchanger, and flash column, with a detailed process flow diagram (PFD) modeled in Aspen Plus. Material and energy balances ensure the operational feasibility of the system.

With an effective conversion rate of 50%, the process accounts for realistic industrial limitations while maintaining reliability at scale. Heat recovery mechanisms, including a shell-and-tube heat exchanger, improve energy efficiency by minimizing heat loss. Optimized equipment design ensures process scalability and aligns with performance criteria to meet sequestration targets and product quality standards.

The reliance on olivine, an abundant and cost-effective silicate mineral, highlights the economic and environmental advantages of this approach. The findings contribute to advancing sustainable CCS technologies, offering a viable solution for CO_2 mitigation while producing valuable industrial products such as $MgCO_3$ and by-product SiO_2.

Keywords: Mineral Carbonation; Olivine; Magnesium Carbonate; and Ex-Situ Carbonation.

Introduction:

The global rise in CO_2 emissions from fossil fuel consumption is a major driver of climate change [1]. Addressing this challenge requires innovative carbon capture and storage (CCS) technologies. To stabilize atmospheric CO_2 at approximately 500 ppm by 2050, a broad portfolio of solutions is needed, with CCS recognized as a flagship technology for transitioning from a fossil fuel-based economy to a renewable-based one [2]. CCS encompasses various processes that capture CO_2 from power plants, cement factories, steel production, and natural gas treatment, preventing its release into the atmosphere [3]. All pathways to limiting global warming to 1.5°C rely on some form of CO_2 removal, with an estimated 190 Gt of CO_2 needing to be stored [4]. However, large-scale CCS deployment remains a challenge, requiring significant technological advancements.

Among CCS methods, mineral carbonation has gained attention as a promising long-term solution due to its ability to permanently store CO_2 in the form of stable carbonates. Unlike geological or ocean storage, which requires extensive monitoring and poses leakage risks, mineral carbonation mimics natural rock weathering, converting CO_2 into magnesium and calcium carbonates [5]. These carbonates can be used in industrial applications or safely stored. Studies identify olivine (Mg_2SiO_4) as a particularly effective mineral due to its high magnesium content, abundance, and fast reaction kinetics compared to other silicates.

Despite its potential, large-scale mineral carbonation faces challenges, including high energy consumption, slow reaction kinetics due to passivation layer formation, and CO_2 supply chain constraints. Research has explored mechanical activation, acid leaching, and thermal pre-treatment to improve efficiency, but economic feasibility and process optimization remain key concerns. Additionally, while mineral carbonation has been extensively studied in industrial regions, its implementation in Pakistan remains underexplored. Given the presence of ultramafic rock deposits rich in olivine, Pakistan has significant potential for localized CCS solutions.

This study aims to develop a scalable CO_2 sequestration process using locally available olivine, integrating material and energy balance considerations to enhance process efficiency. By addressing key technical and economic barriers, this research contributes to advancing sustainable CCS technologies for long-term climate mitigation.

Novelty Statement:

This study presents the design of a mineral carbonation plant capable of sequestering 30 tons/day of CO_2, producing valuable industrial products such as magnesium carbonate ($MgCO_3$) and silica (SiO_2) while minimizing cost, energy consumption, and environmental impact. The proposed ex-situ carbonation process utilizes olivine with minimal pre-treatment, efficient heat recovery mechanisms, and precision-engineered equipment to enhance efficiency and sustainability. Unlike conventional methods, this system achieves a high sequestration rate while ensuring economic viability through the production of commercially valuable by-products.

Sustainable Development Goals:

This study aligns with the following United Nations Sustainable Development Goals (SDGs):

- **SDG 9:** Industry, Innovation, and Infrastructure – by promoting advancements in carbon capture technology.
- **SDG 11:** Sustainable Cities and Communities – by contributing to cleaner industrial processes and reduced emissions.
- **SDG 13:** Climate Action – by mitigating CO_2 emissions through effective sequestration strategies.

Objectives of the Study:

- **Design a process flow** for an aqueous olivine carbonation plant to sequester CO_2.
- **Calculate mass flow rates** and analyze energy requirements for efficient process integration.
- **Design key equipment** essential for the carbonation process.
- **Evaluate mineral carbonation** as a practical and sustainable method for CO_2 sequestration.

Literature Review:

Carbon Capture and Storage (CCS) technology aims to prevent CO_2 emissions from entering the environment, thus reducing the carbon footprint. There are three major CCS methods: Geological Storage, Ocean Storage, and Mineral Carbonation[5].

1. **Geological Storage** involves injecting captured CO_2 into specific geological formations, such as coal-bed formations or depleted oil and gas reservoirs[6]. The rationale behind injecting CO_2 into these reservoirs is that the hydrogeological conditions that originally allowed hydrocarbons to accumulate also enable CO_2 to be stored. As long as the caprock remains intact (unexposed, undamaged, and not weakened by excessive pressure), it can trap CO_2 for thousands of years[7].
2. **Ocean Storage** refers to injecting CO_2 into deep ocean layers, either in liquid or gaseous form. The ocean, with a residence duration of several hundred years, acts as the largest accessible sink for CO_2[8]. The deep ocean's thermocline stratifies its layers, which slowly mix over time, facilitating CO_2 sequestration. At depths above 1500 meters, CO_2 can be injected as a liquid or trapped as hydrates, or it may dissolve in the water column[9].
3. **Mineral Carbonation** is a promising CCS technology in which CO_2 reacts with silicate minerals like olivine or serpentine to form stable carbonates such as $MgCO_3$ and $CaCO_3$[10]. This method not only sequesters CO_2 but also captures and utilizes it to produce valuable industrial products like magnesium carbonate, giving it an edge over other CCS methods[5].

In Pakistan, where rapid demographic growth and industrialization heavily depend on fossil fuel consumption, CO_2 emissions are becoming a growing environmental concern[11]. This highlights the urgent need for implementing CCS technologies. The objective of this study is to design a mineral carbonation plant that sequesters CO_2 emissions from fossil fuel combustion through the mineralization of olivine. It is found that when olivine reacts with water, a strong passivating layer forms, significantly trapping CO_2 from the atmosphere[12]. The focus of this study is to develop an efficient process that minimizes the need for pre-treatment, particularly addressing the shedding of the passivation layer to make the carbonation process more economical.

There are two primary types of mineral carbonation processes: In-situ and Ex-situ.

- **In-situ carbonation** occurs naturally by injecting CO_2 into silicate-rich geological formations or alkaline aquifers, targeting minerals like peridotite and basalts[13].
- **Ex-situ carbonation** takes place in a controlled environment, typically in a chemical processing plant after the mineral has been mined. This process uses a slurry of minerals like olivine, which is then reacted with compressed CO_2 to form useful products like magnesium carbonate[14].

For this study, ex-situ carbonation is chosen because it allows for controlled conditions and efficient product generation. Olivine, an abundant and reactive silicate mineral, is used as the feedstock due to its cost-effectiveness and availability, especially in regions with ultramafic rock deposits, like Pakistan[15][16]. The olivine is mined and processed into slurry with water, and then exposed to compressed CO_2 in a reactor, forming $MgCO_3$ and SiO_2. The slurry

undergoes a series of separation and purification processes, with CO_2 being recycled and reused to maximize the efficiency of the process [14].s.

Material and Methods:

This section outlines the approach for CO_2 sequestration via mineral carbonation. It covers the following aspects:

1. **Process Overview**: A step-by-step description of how CO_2 is captured and converted into stable carbonates through mineral carbonation.
2. **Raw Materials Required**: Identification of essential raw materials, such as olivine (Mg_2SiO_4), water, and CO_2. Olivine is the primary feedstock due to its high reactivity and abundance, making it an economical choice for carbonation.
3. **Chemical Reactions**: The carbonation reaction, where olivine reacts with CO_2 to form magnesium carbonate ($MgCO_3$) and silica (SiO_2).

The simplified reaction is:

$$2CO_2 + Mg_2SiO_4 \rightarrow 2MgCO_3 + SiO_2 \quad (1)$$

This exothermic reaction helps in CO_2 sequestration, converting it into stable minerals that can be utilized in industrial applications.

4. **Mass and Energy Balance**: Detailed calculations of mass and energy flow within the carbonation plant to ensure efficient process integration. This analysis ensures that the plant operates within the desired parameters, optimizing the conversion and minimizing energy losses.
5. **Key Equipment Design**: The design of essential equipment for the carbonation process, including:
   - **Reactor Vessel**: Where the carbonation reaction occurs.
   - **Slurry Tank**: For mixing olivine with water to form the slurry.
   - **CO_2 Compressor**: To pressurize the CO_2 for efficient reaction.
   - **Separation Unit**: For separating the solid magnesium carbonate and silica from the slurry.
   - **Coolers and Gas Separators**: To recover CO_2 and ensure its reusability in the system.

Figure 1 below illustrates the process flow chart for the entire CO_2 sequestration process, showing the interconnection between these components and the overall process flow.

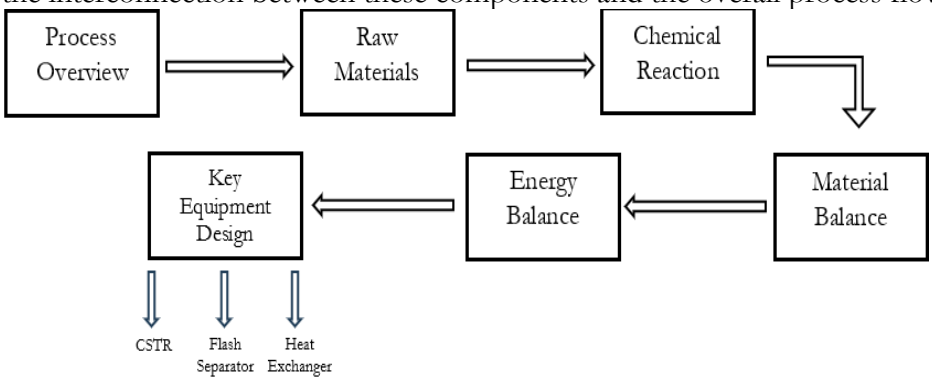


Figure 1. Flow chart of the methodology

Process Overview:

The mineral carbonation process is an addition reaction between a mineral (olivine) and carbon dioxide, resulting in the formation of insoluble carbonates. This process utilizes naturally occurring silicates, and this study focuses specifically on olivine. In the reaction, olivine, also known as Magnesium Orthosilicate, reacts with compressed carbon dioxide to produce stable carbonates. The general carbonation reaction for olivine is given below:

$$2CO_2 + Mg_2SiO_4 \rightarrow 2MgCO_3 + SiO_2 \,(1)$$

This reaction is exothermic and effectively sequesters CO_2 by converting it into solid magnesium carbonate ($MgCO_3$) and silica (SiO_2), which are environmentally stable compounds.

This study focuses on the design of a mineral carbonation plant with a capacity to sequester 30 tons of CO_2 per day. The process is based on ex-situ carbonation, utilizing aqueous olivine as the feedstock. The method involves preparing a slurry of olivine, pre-processing CO_2, reacting the slurry in a continuous stirred tank reactor (CSTR), and separating the final products. A process flow diagram (PFD) has been created to provide a clear understanding of the entire process.

Raw Materials and Chemical Reaction:

The raw materials used in this process are olivine (Mg_2SiO_4), water (used as the medium for slurry preparation), and an additive solution containing 1M sodium chloride (NaCl) and 0.64M sodium bicarbonate ($NaHCO_3$) to control the pH and reaction kinetics. The reactions involved in aqueous mineralization within a continuous stirred tank reactor (CSTR) are outlined below.

$$CO_2 + H_2O \rightarrow H_2CO_3 \rightarrow H^+ + HCO_3\text{-} \quad (2)$$

$$Mg_2SiO_4 + 4H+ \rightarrow 2Mg^{+2} + SiO_2 + 2H_2O \quad (3)$$

$$Mg^{+2} + HCO_3^- \rightarrow MgCO^3 + H^+ \quad (3)$$

These reactions occur in the reactor within a temperature range of 150°C to 200°C and at pressures between 120 and 150 atm. In the first step, carbon dioxide reacts with water to produce hydrogen ions (H^+) and bicarbonate ions (HCO_3^-). In the second step, the olivine rock reacts with the hydrogen ions to produce magnesium ions (Mg^{2+}), silica (SiO_2), and water. In the third step, magnesium ions react with bicarbonate ions to form solid magnesium carbonate ($MgCO_3$).

Material Balance:

Material balance calculations are performed based on the stoichiometric reactions involved in the olivine carbonation process. To form 1 ton of $MgCO_3$, 1.6 tons of olivine are required for every 1 ton of CO_2 reacted[14]. The process operates at a 50% conversion rate of olivine to reflect realistic conditions and efficiency. The plant is designed to sequester 30 tons of CO_2 per day, equating to 1250 kg/hr of CO_2 sequestration. According to the principle of material balance, input equals output[17].

Figure 2 illustrates the mineral carbonation process of olivine in an aqueous solution, which can also be applied to other natural silicates and industrial residues, such as steel slag [18]. As shown in Figure 2, the olivine undergoes size reduction through crushing and milling to increase the surface area for the catalyzed reaction. The additives are recovered during solid filtration, facilitated by the fine particles of olivine[19].

Energy Balance:

Energy balance calculations are performed to optimize the thermal and mechanical efficiency of the process. To improve the thermal and mechanical aspects of the procedure, energy balances were computed. The energy released is calculated using the energy balance equation:

$$Q=mC\Delta T$$

Where Q is the heat energy, mmm is the mass, C is the specific heat capacity, and ΔT is the temperature change[20].

The heat capacity for slurry tank preparation can be calculated using the following formula:

$$Q=mC\Delta T$$

Where:

- Q is the heat energy required (Joules),

- mmm is the mass of the slurry (kg),
- C is the specific heat capacity of the slurry (J/kg·°C),
- ΔT is the temperature change (°C).

This formula calculates the amount of heat required to raise the temperature of the slurry in the tank by a certain amount.

$$C_{Pmix} = \left(\frac{m_{Mg_2SiO_4}}{m_{mix}}\right) C_{p_{Mg_2SiO_4}} + \left(\frac{m_{H_2O}}{m_{mix}}\right) C_{p_{H_2O}} + \left(\frac{m_{Nacl}}{m_{mix}}\right) C_{p_{NaCl}} + \left(\frac{m_{NaHCO_3}}{m_{mix}}\right) C_{p_{NaHCO_3}}$$ [20]

For energy recovery in the shell-and-tube heat exchanger, the same formula for heat capacity is used, but with adjustments for the different values of mass input and output. The formula remains:

$$Q=mC\Delta T$$

Where:

- Q is the heat energy recovered (Joules),
- mmm is the mass of the fluid (effluent) passing through the heat exchanger (kg),
- C is the specific heat capacity of the effluent (J/kg·°C),
- ΔT is the temperature change (°C) between the inlet and outlet of the heat exchanger.

For the heat exchanger:

- m, changes depending on the flow rate of the effluent,
- C depends on the composition and properties of the effluent fluid,
- ΔT is the difference between the temperature of the effluent when it enters and exits the heat exchanger.

By using this equation, the amount of thermal energy that can be recovered and reused in the process can be calculated. This helps in optimizing the energy efficiency and reducing overall thermal energy consumption.

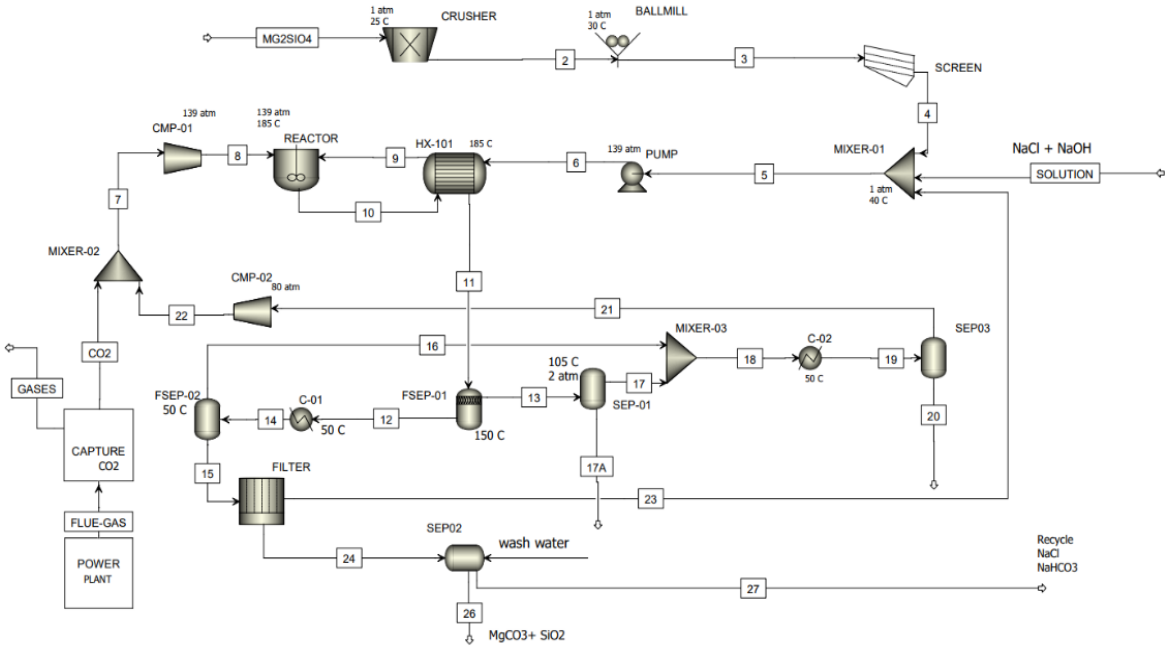


Figure 2. Process flow diagram (PFD) of Mineral carbonation of olivine in aqueous solution created on Aspen Plus.

Key Equipment Designing:

To calculate the power and volume for the major equipment in the mineral carbonation plant, the following formulas are used for the Continuous Stirred Tank Reactor (CSTR), flash separator, and heat exchanger:

1. **Power for CSTR:**

The power required for the CSTR is calculated using the formula:

$$P=KTNr^3Da^5\ [21]\varrho$$

Where:

- P is the power required (W),
- K is a constant specific to the reactor design,
- T is the temperature of the system (°C),
- Nr is the Reynolds number, a dimensionless number that characterizes the flow regime inside the reactor,
- Da is the characteristic length (diameter or similar) of the reactor (m),
- ϱ is the density of the fluid inside the reactor (kg/m³).

This formula helps calculate the power needed to agitate the slurry and maintain the required reaction conditions inside the CSTR.

2. **Volume for CSTR:**

The volume of the CSTR is calculated using the formula:

$$V = \frac{Fa \times X}{-ra}$$

Where:

- V is the volume of the CSTR (m³),
- Fa is the molar flow rate of the reactant (mol/s),
- X is the conversion rate (dimensionless),
- ra is the rate of reaction (mol/s·m³).

This equation helps determine the required reactor volume based on the molar flow rate, conversion efficiency, and reaction rate [22].

Both formulas are essential for the design and optimization of the CSTR, ensuring that it operates efficiently and at the required power and volume levels for the carbonation process.

Here are the formulas and the explanation for the heat exchanger and flash separator design calculations:

Heat Exchanger Design:

The formula for the design calculation of a heat exchanger is:

$$\frac{1}{Uo} = \frac{1}{hs} + \frac{1}{hod} + Do \times \frac{\ln\left(\frac{Po}{Pin}\right)}{2kw} + \frac{Do}{Din}\left(\frac{1}{hid} + \frac{1}{hi}\right)$$

Where:

- U_o is the overall heat transfer coefficient (W/m²·K),
- h_s is the heat transfer coefficient on the hot side (W/m²·K),
- h_{od} is the heat transfer coefficient on the outer side of the tube (W/m²·K),
- D_o is the outer diameter of the tube (m),
- P_o and P_{in} are the outlet and inlet pressures (Pa),
- kw is the thermal conductivity of the material (W/m·K),
- D_{in} is the inner diameter of the tube (m),
- h_{id} is the heat transfer coefficient inside the tube (W/m²·K),
- h_i is the heat transfer coefficient inside the tube (W/m²·K).

This formula is used to determine the overall heat transfer coefficient, which is a key parameter in designing the heat exchanger to recover and optimize the thermal energy within the system [23].

Flash Separator Design:

1. **Vapor Velocity (Vt):**

The vapor velocity is calculated using the formula:

$$Vt = k\sqrt{\frac{p_l - p_v}{p_v}}$$

Where:

- Vt is the vapor velocity (m/s),
- K is a constant,
- p_l is the density of the liquid phase (kg/m³),
- p_v is the density of the vapor phase (kg/m³).

This formula helps calculate the velocity of the vapor in the separator, which is crucial for sizing and designing the flash separator [24].

2. **L/D Ratio:**

The L/D ratio (Length/Diameter ratio) is calculated to determine if using the flash separator for the plant is feasible. If the L/D ratio is too high, it could indicate that the separator is not effective for the required conditions.

3. **Volumetric Flowrate for Liquid and Vapors (Q):**

The volumetric flowrate for liquid and vapor phases is calculated using the formula:

$$Q = \frac{W}{p}$$

Where:

- Q is the volumetric flowrate (m³/s),
- W is the mass flowrate (kg/s),
- is the density (kg/m³).

This equation is used to calculate the volumetric flow rates for both liquid and vapor phases, which is essential for sizing the flash separator and ensuring its operation is within the required specifications [25].

These formulas are integral to the design and optimization of the heat exchanger and flash separator, helping to ensure that the processes within the mineral carbonation plant are efficient, effective, and meet operational requirements.

Result and discussion

Table 1. Overall Material Balance

Input		**Output**	
Mg_2SiO_4	**Kg/hr**		**Kg/hr**
Water	2000	$MgCO_3$	1199.9
NaCl	10700.04	SiO_2	428.571
$NaHCO_3$	328.67	Mg_2SiO_4	999.9
CO_2	304.6	CO_2	622.616
	1250	Water	10700.04
		NaCl	328.67
Total		$NaHCO_3$	304.603
Mg_2SiO_4	14583.3	Total	14583.3

Carbon Sequestration Efficiency and Products.

The mineral carbonation plant designed for this study is capable of sequestering 30 tons of CO_2 per day using olivine as the feedstock. The process involves the conversion of CO_2 and magnesium orthosilicate (olivine) into stable magnesium carbonate ($MgCO_3$) as the primary product and silica dioxide (SiO_2) as a by-product. This reaction pathway contributes to CO_2 sequestration while producing valuable industrial products that can be utilized in various industries. The design focuses on efficiency, minimal pre-treatment, and sustainability.

Table 1 presents the overall material balance calculations for each piece of equipment shown in Figure 2. According to the material balance, for every hour of operation, approximately 1199.9 kg of $MgCO_3$ and 428.571 kg of SiO_2 are produced. The majority of the product, $MgCO_3$, has significant industrial applications such as in fertilizers, concrete raw materials, and as a filler in the paint and paper industries.

The 50% conversion rate, accounting for realistic operational limitations, aligns the process with industrial-scale practices and ensures efficiency. The process also achieves material balance, as evidenced by the results in Table 1, confirming the system's reliability for real-world applications.

The recovery of additives via solid filtration highlights the sustainability of the process, minimizing waste generation. Additionally, the use of sieving mechanisms to separate fine silicate and carbonate particles ensures the high purity and quality of the final products, further supporting the feasibility and effectiveness of the mineral carbonation plant design.

Energy Optimization and Thermal Efficiency:

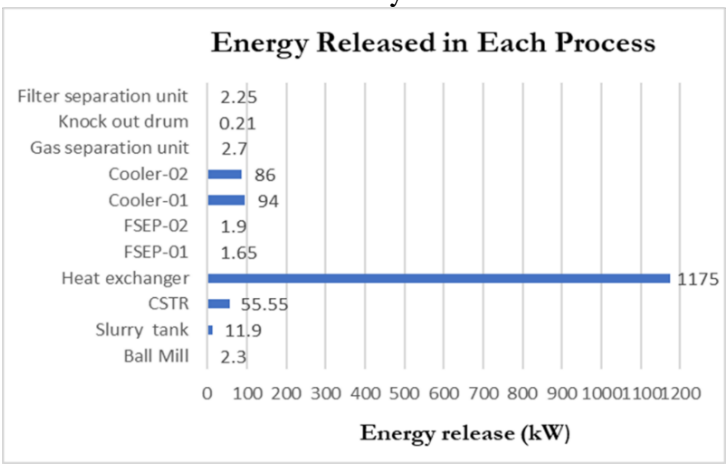


Figure 3. Graphical representation of energy released in each process of the plant

Energy optimization is a crucial aspect of plant design to ensure efficiency and sustainability. In this design:

- The mixing energy for olivine particles in the ball mill is calculated to be **2.3 kW**, which represents the energy required to break down the olivine to an optimal size for reaction.
- The **Continuous Stirred Tank Reactor (CSTR)**, where the carbonation reaction takes place, released **55.55 kW** of thermal energy during the process. This thermal energy is a result of the exothermic nature of the carbonation reaction.
- To recover this thermal energy, a shell-and-tube heat exchanger is incorporated into the design. This heat exchanger facilitates the efficient transfer of heat from the reactor effluent to other parts of the process. It achieved an exchanged heat value of 1175 kW, improving the overall thermal efficiency of the system by recycling the heat and reducing the energy required for heating the incoming fluids.

These energy optimization measures contribute to the overall energy efficiency of the plant, enhancing its sustainability and economic feasibility.

4o mini

The heat exchanger's design, with a heat transfer area of 68.7 m^2, is pivotal in ensuring effective energy utilization within the mineral carbonation plant. This large surface area enables optimal heat exchange, allowing the system to recover and reuse thermal energy efficiently, thereby significantly reducing operational costs.

Figure 3 illustrates the energy release at each stage of the plant, with the heat exchanger showing a higher thermal energy release compared to other processes. This highlights the critical role of the heat exchanger in recovering and reusing heat from the carbonation reaction, which is vital for minimizing energy consumption.

The calculated thermal efficiencies further demonstrate that the plant is designed to operate under energy-conserving conditions, ensuring reduced energy demand across the system. This approach is particularly valuable in regions like Pakistan, where energy constraints often limit industrial growth and sustainability.

By focusing on energy recovery and minimizing consumption, this plant design not only supports economic feasibility but also aligns with sustainable practices, ensuring that the system can operate efficiently in energy-limited environments.

Process Optimization Through Equipment Design:

The three major equipment specifications are meticulously designed to ensure continuous, efficient operation of the mineral carbonation plant. Each component is optimized for its specific function within the overall system.

1. **CSTR (Continuous Stirred Tank Reactor)**:
   - **Volume**: 1.62 m^3
   - **Power Requirement**: 254.318 kW
   - The CSTR plays a crucial role in the mixing and reaction of CO_2 with aqueous olivine slurry. This large volume ensures the efficient processing of the required CO_2 flow and slurry. The power requirement reflects the energy needed to maintain continuous mixing and facilitate optimal reaction conditions.
2. **Shell-and-Tube Heat Exchanger**:
   - Designed for efficient heat recovery, the shell-and-tube heat exchanger transfers thermal energy from the reaction effluent to the incoming fluid, lowering overall energy consumption and reducing operating costs.
3. **Flash Separator**:
   - The stainless-steel construction ensures durability and resistance to corrosion, effectively separating the liquid and gaseous phases after the reaction. This component is critical for ensuring the purity of the products and maximizing the separation efficiency, which is essential for the high-quality output of magnesium carbonate ($MgCO_3$).

The detailed equipment specification sheets for each major component emphasize the importance of design optimization, ensuring that the plant operates efficiently while minimizing waste and energy consumption.

Table 2. CSTR Equipment specification sheet

Equipment specification sheet	
Type of equipment	Continuous stirred tank reactor (CSTR)
Function	Reacting CO_2 & Mg_2SiO_4 to produce $MgCO_3$
No. of units Conversion	1 50%
Operating temperature	185°C
Operating pressure	139bar
Design Information	
CSTR Dimensions	Volume = 1.62 m^3
	Diameter = 1.112 m
	Height = 1.668 m
	Depth of liquid = 1.112 m
	Length of impeller = 0.278 m
Power Requirement	254.318 kW
Heat transfer area	39.62 m^2
Material of construction	Carbon steel

Table 2 includes the CSTR equipment sheet specification, showcasing key components and the material selection for each part. Additionally, it highlights the heat transfer area, which is integral to the thermal efficiency of the CSTR design. This ensures the reactor operates effectively, maintaining optimal conditions for the carbonation process.

4o mini

Table 3. Shell & Tube Heat Exchanger Equipment specification sheet

Equipment specification sheet	
Type of machine	Type of machine
Function	Function
Operating pressure	Operating pressure
Operating condition	Operating condition
LMTD	LMTD
Design Information	
Heat Transfer Area	Heat Transfer Area
Tube Length	Tube Length
Tube area per pass	Tube area per pass
Shell bundle diameter	Shell bundle diameter
Crossflow area	Crossflow area
Volumetric flowrate	Volumetric flowrate
Overall heat transfer coefficient	Overall heat transfer coefficient
Material of Construction	Material of Construction

Table 3 provides the shell & tube heat exchanger equipment specification sheet, detailing its key components and operational parameters:

- **LMTD (Log Mean Temperature Difference)**: The LMTD value of 11.47°C indicates the effectiveness of heat transfer between two process streams, steam 6 and steam 10 (as depicted in Figure 2). A higher LMTD indicates efficient heat transfer, which enhances the thermal recovery efficiency of the heat exchanger.
- **Heat Transfer Efficiency**: The heat exchanger operates by transferring heat between the two streams. The higher temperature fluid in the tube side improves heat transfer, which contributes to increased thermal efficiency of the overall process. This is crucial for reducing energy consumption, thereby minimizing operational costs.
- **Material of Construction**: Carbon steel is selected for the heat exchanger due to its durability and ability to withstand high temperatures encountered in the process. Its resistance to corrosion and robustness in extreme conditions ensure reliable, long-term operation.
- **Industrial Applicability**: The shell & tube heat exchanger is a widely used design due to its adaptability to a variety of industrial applications. Its high heat transfer rate and efficiency make it a suitable choice for processes that require effective thermal management, such as in the mineral carbonation plant design.

This detailed specification emphasizes the energy-efficient design and cost-effectiveness of the system, ensuring optimized thermal management and sustainable operation.

4o mini

Table 4. Flash Separator Equipment specification sheet

Equipment specification sheet	
Type of machine	Type of machine
Function	Function
The mass flowrate of liquid	The mass flowrate of liquid
The mass flowrate of vapors	The mass flowrate of vapors
Vapor Velocity	Vapor Velocity
Volumetric flowrate(liquid)	Volumetric flowrate(liquid)
Volumetric flowrate (vapors)	Volumetric flowrate (vapors)
Vessel diameter	Vessel diameter
Liquid depth	Liquid depth
Height of vapor space	Height of vapor space

Flooding height	Flooding height
Mist extractor clearance	Mist extractor clearance
L/D ratio	L/D ratio
Wall thickness	Wall thickness
Material of construction	Material of construction

Table 4 presents the flash separator equipment specification sheet, outlining key design parameters and operational considerations:

- **L/D Ratio**: The **L/D value** (Length/Diameter ratio) is less than 5.5, indicating that the use of the flash separator is appropriate for the plant's design. A low L/D ratio indicates that the separator is well-suited for efficient phase separation (liquid and vapor) under operating conditions, ensuring effective separation and maximizing efficiency.
- **Separation Efficiency**: The flash separator's ability to efficiently separate the liquid and gaseous phases is demonstrated by the low L/D ratio, which ensures the desired separation of phases with minimal energy loss.
- **Material of Construction**: Stainless steel is chosen for the flash separator due to its compatibility with the materials involved (such as CO_2, water, and olivine slurry), along with its durability. Stainless steel's resistance to corrosion and its ability to withstand the high-pressure conditions typical of separators ensure longevity and reliability in the plant.

This specification confirms the flash separator's effectiveness in maintaining separation efficiency and ensures long-term performance under harsh operating conditions.

Discussion:

The findings of this study highlight the viability and scalability of mineral carbonation using olivine for CO_2 sequestration, offering significant environmental and economic benefits. The key takeaways are:

- **CO_2 Conversion**: The designed process, capable of sequestering 30 tons of CO_2 per day, converts CO_2 into magnesium carbonate ($MgCO_3$) and silica dioxide (SiO_2). Both by-products have notable industrial applications, strengthening the process's commercial viability.
- **Industrial Applications**:
  - **$MgCO_3$**: This compound has multiple uses, including as a fertilizer, concrete raw material, and filler in paints and paper. In particular, its low-carbon properties make it valuable in concrete applications, contributing to CO_2 reduction in the construction industry.
  - **SiO_2**: This by-product is utilized in glass manufacturing and as a precursor for silicon semiconductor production, with applications in fields such as rice hull-derived silicon.
- **Energy Efficiency**: The integration of heat recovery systems within the plant design improves energy efficiency, thereby reducing both costs and environmental impact. The thermal recovery in the process contributes significantly to operational cost reduction.
- **Engineering and Material Selection**: Given the high-pressure (139 bar) and high-temperature (185°C) conditions required for carbonation, carbon steel is selected for the reactor's construction due to its structural durability and corrosion resistance. This material choice is critical for large-scale deployment. However, future research may explore alternative materials or coatings to enhance reactor lifespan and operational efficiency.
- **CO_2 Capture and Transport**: While the study focuses on the carbonation process, it highlights the importance of optimizing the capture and transport of CO_2. Utilizing waste CO_2 from industrial processes (e.g., flue gases) instead of relying on pure CO_2 streams could significantly improve the economics of the process.

In conclusion, this study underscores the dual-purpose nature of mineral carbonation, addressing both CO_2 reduction and the growing demand for sustainable industrial materials. The economic potential of the reaction by-products and the energy efficiency improvements make this technology a promising solution for carbon capture and utilization. Further research

and optimization of CO_2 capture and reactor materials will enhance the commercial feasibility and long-term viability of the process.

Sustainability and Economic Impact:

The dependence on olivine, an abundant and cost-effective silicate mineral, underscores the economic and environmental advantages of this mineral carbonation process. Key points include:

- **Local Feedstock Availability**: In Pakistan, the presence of ultramafic rocks, such as peridot, offers a local and cost-effective feedstock source. This reduces the need for imported materials, further enhancing the process's economic feasibility.
- **Product Recovery and Income Generation**: The recovery of $MgCO_3$ (magnesium carbonate) and SiO_2 (silica dioxide) creates opportunities for income generation, contributing to covering operational costs and improving economic sustainability. These by-products have significant industrial applications, including in fertilizer production**,** concrete manufacturing, and paint and paper industries.
- **Reaction Enhancement**: The use of additive solutions, such as sodium chloride (NaCl) and sodium bicarbonate ($NaHCO_3$), improves reaction kinetics, optimizing the carbonation process while keeping it simple and cost-effective. This approach balances operational efficiency with minimal complexity.
- **Climate Change Mitigation**: This process is aligned with global efforts to mitigate climate change, particularly by addressing CO_2 emissions from fossil fuel consumption. By converting CO_2 into stable carbonates, it provides a scalable and economically viable solution to carbon capture and storage (CCS) technologies.
- **Industrial Growth and Sustainable Development**: The plant design supports Pakistan's industrial growth by utilizing local feedstocks and reducing carbon emissions. It offers a pathway for sustainable development by integrating carbon sequestration with industrial applications, reducing the country's carbon footprint while fostering eco-friendly industries.

Overall, the mineral carbonation process offers a sustainable solution for CO_2 sequestration, with economic benefits derived from by-product recovery and local resource use, aligning with Pakistan's goals for industrial growth and climate action.

Process limitations:

Mineral carbonation presents a promising and sustainable solution for CO_2 sequestration, particularly due to its ability to form stable carbonates like magnesium carbonate ($MgCO_3$), which reduces the risk of CO_2 being released back into the atmosphere. However, several process limitations hinder its large-scale implementation:

1. **Passivation Layer Formation**:
   - The presence of silica-rich layers on the surface of olivine minerals creates a passivation layer, which **slows** down the carbonation process. This layer restricts the effective surface area available for the reaction, reducing the efficiency of CO_2 sequestration and requiring more time for the process to complete.
2. **Slow Reaction Kinetics**:
   - The dissolution kinetics of olivine, influenced by weakly acidic conditions, are relatively slow. This leads to slow reaction rates and limits the overall effectiveness of the carbonation process. The formation of the passivation layer also contributes to these slow reaction kinetics, as it obstructs the mineral's ability to react efficiently with CO_2.
3. **High Operational Costs**:
   - **Energy-intensive nature**: The mineral carbonation process requires high temperatures and pressures, increasing the energy demand.

- **Mining and Pre-treatment Costs**: The extraction, grinding, and pre-treatment of silicate minerals, such as olivine, add to the overall cost of the process.
- These high operational costs are one of the major barriers to large-scale commercialization, making it economically challenging to implement the process on a wide scale.

In summary, while mineral carbonation offers a durable solution for CO_2 sequestration, addressing the slow reaction kinetics, passivation layer formation, and high operational costs are essential for improving its economic viability and achieving large-scale commercial deployment.

Conclusion:

The motivation for this study lies in the urgent need to reduce CO_2 emissions and simultaneously produce magnesium carbonate ($MgCO_3$), a valuable product. The conceptual design of this process focuses on developing a cost-effective and environmentally friendly solution. As fossil fuel consumption continues to rise, so do CO_2 emissions, necessitating the need for effective carbon capture and sequestration (CCS) technologies.

Key aspects of the study include:

1. **Ex-situ Mineral Carbonation**: This process captures and stores CO_2 through carbonation, offering a practical method for CO_2 sequestration.
2. **Process Flow Diagram (PFD)**: A PFD has been created to visualize the working of the entire plant, helping to understand the sequence and integration of processes involved in CO_2 capture and $MgCO_3$ production.
3. **Material and Energy Balance**: These balances have been applied to ensure that the plant operates efficiently. The energy balance indicates that although the process is energy-intensive, heat recovery mechanisms are in place, utilizing recycled streams and process streams for heat exchange, thereby optimizing energy use.
4. **Product and By-product Applications**:
   - **Magnesium Carbonate ($MgCO_3$)**: This valuable product can be used in several industrial applications, such as fireproofing, concrete production, dusting powders, toothpaste, and fire extinguishers.
   - **Silica Dioxide (SiO_2)**: The by-product can be used in glass manufacturing, as food additives, and in other industrial processes.
5. **Economic Viability**: Both the main product ($MgCO_3$) and by-product (SiO_2) have high market value, which supports the economic feasibility of the process. They can be easily sold, contributing to cost recovery and enhancing the financial sustainability of the project.
6. **Practical Implications**: The project not only addresses the CO_2 emissions issue but also contributes to sustainability and the development of renewable energy solutions. The study translates theoretical knowledge into practical applications, providing valuable insights into process engineering and plant design.

In conclusion, this project is a significant step towards sustainable development, offering a feasible solution to CO_2 sequestration while also creating valuable industrial products.

Challenges and Future Perspectives: Although the results are encouraging, some challenges need to be overcome to enable large-scale deployment. The passivation layer created through the carbonation reaction is a major challenge, as it lowers reaction efficiency. This work emphasizes the need for a process design that minimizes or eliminates pre-treatment activation. Additional research is also needed for the optimization of reaction kinetics and enhancing additive recovery processes. Lower energy requirements are needed to reduce the high-cost value. The aim of the study is ex-situ carbonation, which provides controlled conditions, giving constant results. Whereas, the in-situ method might offer an alternative for

certain applications, especially in areas with applicable geological formations. Investigating hybrid approaches that give the benefits of both methods might discover innovative solutions.

Acknowledgment: The authors express their sincere gratitude to the University of Gujrat for its unwavering support. Our earnest gratefulness to our parents for their moral and financial encouragement throughout this journey. We are also deeply thankful to the Department of Chemical Engineering and Technology, as well as all those people whose guidance and aid played a crucial role in the successful completion of this study.

Author's Contribution: Meerab Yousuf conceptualized the idea of this study, conducted a literature review, and designed three major pieces of equipment. Abdul Basit developed the process flow diagram of the plant on the Aspen Plus. Abdul Basit provided guidance and assistance in process designing and material and energy balance analysis.

Conflict of interest: The authors declare that they have no known competing financial interests or personal relations that could have influenced the work in this study.

Project details: NIL.

References:

[1] A. K. Das and A. Sharma, "Climate change and the energy sector," in *Advancement in Oxygenated Fuels for Sustainable Development*, Elsevier, 2023, pp. 1–6.

[2] S. Pacala and R. Socolow, "Stabilization wedges: solving the climate problem for the next 50 years with current technologies," *Science (80-.).*, vol. 305, no. 5686, pp. 968–972, 2004.

[3] M. E. Boot-Handford *et al.*, "Carbon capture and storage update," *Energy Environ. Sci.*, vol. 7, no. 1, pp. 130–189, 2014.

[4] S. Ó. Snæbjörnsdóttir, B. Sigfússon, C. Marieni, D. Goldberg, S. R. Gislason, and E. H. Oelkers, "Carbon dioxide storage through mineral carbonation," *Nat. Rev. Earth Environ.*, vol. 1, no. 2, pp. 90–102, 2020.

[5] R. M. Cuéllar Franca and A. Azapagic, "Life cycle environmental impacts of carbon capture, storage, and utilization," *Encycl. Sustain. Technol.*, pp. 447–459, 2017.

[6] C. Cooper, "A technical basis for carbon dioxide storage," *Energy Procedia*, vol. 1, no. 1, pp. 1727–1733, 2009.

[7] B. Hitchon, W. D. Gunter, T. Gentzis, and R. T. Bailey, "Sedimentary basins and greenhouse gases: a serendipitous association," *Energy Convers. Manag.*, vol. 40, no. 8, pp. 825–843, 1999.

[8] T. Gentzis, "Subsurface sequestration of carbon dioxide—an overview from an Alberta (Canada) perspective," *Int. J. Coal Geol.*, vol. 43, no. 1–4, pp. 287–305, 2000.

[9] D. A. Voormeij and G. J. Simandl, "Geological, ocean, and mineral CO 2 sequestration options: a technical review," *Geosci. Canada*, vol. 31, no. 1, pp. 11–22, 2004.

[10] M. I. Rashid, Z. Yaqoob, M. A. Mujtaba, H. Fayaz, and C. A. Saleel, "Developments in mineral carbonation for Carbon sequestration," *Heliyon*, 2023.

[11] S. R. Ali and N. Mujahid, "Sectoral carbon dioxide emissions and environmental sustainability in Pakistan," *Environ. Sustain. Indic.*, vol. 23, p. 100448, 2024.

[12] E. H. Oelkers, S. R. Gislason, and J. Matter, "Mineral carbonation of CO2," *Elements*, vol. 4, no. 5, pp. 333–337, 2008.

[13] A. A. Olajire, "A review of mineral carbonation technology in sequestration of CO2," *J. Pet. Sci. Eng.*, vol. 109, pp. 364–392, 2013, doi: 10.1016/j.petrol.2013.03.013.

[14] R. Z. Juan Carlos Abanades, Rodney Allam, Klaus S. Lackner, Francis Meunier, Edward Rubin, Juan Carlos Sanchez, Katsunori Yogo, "Mineral carbonation and industrial uses of carbon dioxide," *Carbon Dioxide Capture and Storage*, vol. 10, 2005.

[15] Q. R. S. Miller, H. T. Schaef, J. P. Kaszuba, G. Gadikota, B. P. McGrail, and K. M. Rosso, "Quantitative review of olivine carbonation kinetics: reactivity trends,

mechanistic insights, and research frontiers," *Environ. Sci. Technol. Lett.*, vol. 6, no. 8, pp. 431–442, 2019.

[16] A. Saeed, "a Review of Mineral Carbonation By Enhanced," vol. 31, no. 3, pp. 195–201, 2012.

[17] and S. W. R. L. Cerro, B. G. Higgins, "Material balances for chemical engineers," *na*, 2005.

[18] R. C. Wouter Huijgen, Geert-Jan Witkamp, "Mineral CO2 sequestration in alkaline solid residues," *Greenh. Gas Control Technol. 7*, vol. 2, no. 2, pp. 2415–2418, 2005, doi: https://doi.org/10.1016/B978-008044704-9/50344-X.

[19] W. K. O'Connor, D. C. Dahlin, G. E. Rush, C. L. Dahlin, and W. K. Collins, "Carbon dioxide sequestration by direct mineral carbonation: process mineralogy of feed and products," *Mining, Metall. Explor.*, vol. 19, pp. 95–101, 2002.

[20] R. M. Felder, R. W. Rousseau, and L. G. Bullard, *Elementary principles of chemical processes*. John Wiley & Sons, 2020.

[21] W. L. McCabe, J. C. Smith, and P. Harriott, *Unit operations of chemical engineering*. McGraw-hill, 1993.

[22] P. Trambouze, J.-P. Euzen, and R. Bononno, "Chemical reactors: from design to operation," *(No Title)*, 2004.

[23] J. A. Williams, "COULSON AND RICHARDSON'S CHEMICAL ENGINEERING Volume 6 (Design), by RK Sinnott," *Chem. Eng. Educ.*, vol. 29, no. 2, p. 111, 1995.

[24] I. Naderipour, "The Experimental Study on Effects of Height and Hold up on performance of Vertical Gas-Liquid Separator using Amin Contactor Tower," *Ciência e Nat.*, vol. 37, no. 6–1, pp. 93–103, 2015.

[25] W. Svrcek and W. Monnery, "Design two-phase separators within," *Chem. Eng. Prog.*, 1993.

RESEARCH & INNOVATION DIVISION
IJIST

Comparative Study of Machine Learning Algorithms for Sentiment Analysis in Multimodal Medical Data

Hafiz Muhammad Bilal[1], Muhammad Asif[2], Muhammad Azam Zia[2]
[1]University of Agriculture Faisalabad
[2]Information Technology University
***Correspondence:** Muhammad Asif, Muhammad Azam Zia
muhammadasif.agri@hotmail.com, mazamzia@uaf.edu.pk

Citation | Bilal. H. M, Asif. M, Zia. M. A, "Comparative Study of Machine Learning Algorithms for Sentiment Analysis in Multimodal Medical Data", IJIST, Special Issue. pp 304-317, 2025
Received March 01, 2025 **Revised** | March 13, 2025 **Accepted** | March 19, 2025 **Published** | March 22, 2025.

Sentiment analysis, a part of data mining, uses Natural Language Processing (NLP) to understand how people feel about certain topics or individuals. It focuses on the context and polarity of information, measuring public opinions from unstructured sources like social networks and healthcare websites. By extracting useful insights from this unstructured data, healthcare professionals can improve patient care, make accurate diagnoses, and provide personalized treatments. Machine learning (ML) plays a key role in this process. ML techniques like logistic regression, decision trees, and Naive Bayes have proven effective in tasks such as sentiment analysis and named entity recognition in medical data. The goal of ML is to create algorithms that enhance data processing and decision-making by identifying patterns that might be overlooked by humans. In this study, we compare the performance of three common ML models—(a) Logistic Regression, (b) Decision Tree, and (c) Naive Bayes—for sentiment analysis on medical image captions. The Radiology Objects in Context (ROCO) multimodal image and caption dataset was used for this NLP task. Caption pre-processing is done using filtering methods to improve text quality, followed by sentiment classification using pre-trained ML models. This comparison sheds light on the effectiveness of these algorithms in performing sentiment analysis in clinical settings.

Keywords: Sentiment Analysis; Machine Learning, Natural Language Processing, Confusion Matrix.

Introduction:

Electronic Health Records (EHRs) are widely used worldwide and provide valuable resources for research, improving healthcare quality and population management. As the number of EHRs continues to grow rapidly, effective data analysis methods become increasingly important. Narrative reports, which are the core components of EHR systems, provide detailed information about patients' conditions, reasons for treatment, and doctor-patient interactions [1]. These details are often too complex for structured tables, requiring advanced techniques for analysis.

Natural Language Processing (NLP) has proven effective in processing the text data extracted from EHR systems. It helps extract key information, such as medications and diagnoses, typically in the form of single words or short phrases. However, NLP still faces challenges in extracting more complex information, such as understanding the relationship between illnesses and symptoms, resolving ambiguity in medical terminology, or analyzing emotional tones over multiple sentences [2]. This is where machine learning (ML), a subfield of AI, becomes crucial.

ML uses data-driven algorithms to improve computer performance in tasks like decision-making and pattern recognition. ML models can identify trends and patterns that humans might miss when analyzing historical data. In healthcare, ML has facilitated communication between medical professionals and computer scientists, especially through data mining [3]. Data mining extracts valuable insights from large datasets, reducing the need for costly and invasive medical procedures while improving efficiency and cost-effectiveness [4]. For instance, data mining can help detect high-risk patients and identify key factors associated with positive or negative health outcomes, without relying solely on invasive procedures like X-rays, blood tests, or angiograms.

In healthcare, there is a need to perform sentiment analysis on multimodal clinical data to gain meaningful insights from text-image pairs. This can help determine the best sentiment analysis techniques, improve interpretability, and assist decision-making. It is valuable to compare and contrast different machine-learning models for these tasks.

The use of data mining and machine learning in disease identification and prediction has increased in recent years. Their complexity and application have significantly reduced medical errors and improved diagnostic accuracy. Among the most widely used models for these tasks are logistic regression, decision trees, and Naïve Bayes classifiers.

Logistic Regression: Logistic Regression (LR) is a statistical model used for categorical outcomes, such as binary classifications (e.g., yes/no, true/false). It estimates the probability of an event occurring based on a linear combination of explanatory variables. Its flexibility comes from the minimal assumptions it makes about these variables, making it a useful tool for healthcare predictions [5].

Naïve Bayes: The Naïve Bayes (NB) classifier uses Bayes' theorem to calculate the probability that data belongs to a specific category. It assumes that the features are independent, which simplifies calculations and allows for efficient predictions. Despite this "naïve" assumption, it often provides strong performance in real-world applications [6].

Decision Trees:

Decision Trees (DT) are commonly used for both classification and regression tasks. They divide data into smaller subsets based on feature values, with decisions represented as nodes and outcomes as branches. This clear structure makes Decision Trees particularly useful in healthcare, where understanding the reasoning behind predictions is crucial [7]. In this study, we compare the performance of three machine learning models—Logistic Regression, Decision Trees, and Naïve Bayes—in analyzing text data from Electronic Health Records (EHRs). Our goal is to identify key diagnostic factors and evaluate the accuracy of these models in predicting the need for medical interventions [8]. The results aim to enhance

healthcare practices by leveraging ML's ability to support informed and effective decision-making.

The Unified Modelling Language (UML) system facilitates interoperability among different medical terminologies by linking concepts from various databases, providing an integrated view of medical information. The dataset consists of 8,179 image captions, with 11,154 unique tokens (features) representing the vocabulary. The longest caption contains 133 words, with an average length of about 99.64 words. This dense dataset includes both images and text.

This paper explores Neural Image Captioning (NIC) and its applications in radiology, specifically focusing on sentiment analysis of multimodal medical text data. While previous studies have focused on generating textual descriptions from medical images, this research examines the emotional and subjective aspects of medical texts, such as medical reports, captions, and EHRs. The study adopts a comparative and experimental approach, evaluating several machine learning algorithms for sentiment analysis in the medical domain. By addressing sentiment analysis, a relatively under-researched area in medical AI, this work fills a gap in understanding the emotional and subjective dimensions of medical texts. By using multimodal inputs and comparing traditional and advanced machine learning techniques, the study offers valuable insights and benchmarks for improving sentiment analysis methods in clinical applications. This unique focus not only complements existing research but also provides new perspectives and practical implications for medical AI.

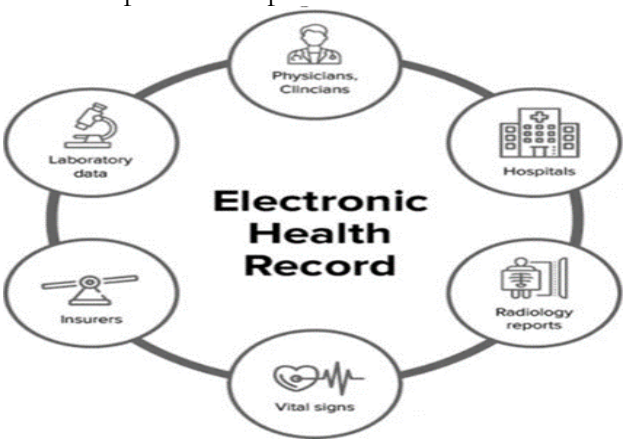


Figure 1. EHR Health Care Application

Objectives of the Study:

The primary objectives of this study are:

1. To use the Radiology Objects in Context (ROCO) dataset to explore multimodal image data and corresponding text descriptions. The goal is to process and standardize the image captions using Natural Language Processing (NLP) tasks to improve data quality for subsequent sentiment analysis.
2. To perform sentiment analysis on the captions of medical images, categorizing them into three classes: positive, negative, and neutral. The study uses the VADER sentiment analysis model to label the data and assess sentiment based on the compound sentiment score derived from individual word sentiment scores.
3. To apply the TF-IDF technique to convert textual data into numerical representations, aiding in the identification of significant terms and enhancing machine learning model performance.
4. To evaluate the performance of different machine learning models (Logistic Regression, Decision Tree, and Naïve Bayes) in classifying sentiment in medical image captions. This includes testing each model's accuracy, precision, recall, and F1 score, focusing on comparing their ability to classify the three sentiment categories.

5. To use confusion matrices and standard evaluation metrics such as accuracy, precision, recall, and F1 score to measure the efficiency of the models and ensure proper sentiment classification in medical image captions.

Material and Methods:

This cross-sectional study used the Radiology Objects in Context (ROCO) dataset, which contains a large-scale multimodal image collection. The images in this dataset are sourced from the PubMed "Central Open Access FTP" mirror, making them publicly accessible and suitable for broad research applications. The images are categorized into two types: non-compound and radiological, based on their identification. In addition to the images, the ROCO dataset includes rich metadata, such as Unified Modelling Language (UML) semantic types, UMLS Concepts Unique Identifiers (CUIs), and image caption keywords [9]. This metadata enhances the understanding of image captions and their relevance in medical text, facilitating the integration of structured medical entities. The approach allows for the comparison and evaluation of machine learning models for sentiment analysis of medical captions, improving the comprehension of health data.

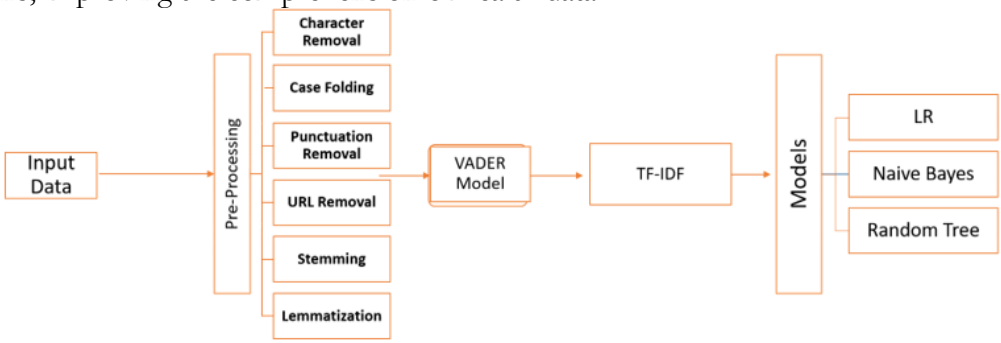


Figure 2. Methodology Flow Diagram

NLP Tasks:

The captions of the images in the dataset are pre-processed through a series of Natural Language Processing (NLP) tasks to standardize the data and remove textual noise, preparing it for further analysis. The dataset consists of two main columns:

- **ID Column**: This uniquely identifies each record in the dataset.
- **Description Column**: This contains the caption or text description for each image.

The Description Column is extracted and further processed for pre-processing [10].

The primary goal of the pre-processing strategies was to standardize the text and remove unnecessary or irrelevant elements, referred to as "textual noise." The key pre-processing steps included:

- **Character Removal**: All characters except letters, numbers, and spaces were removed (e.g., "coronal view" remains as "coronal view").
- **Case Folding** [11]: The text was converted to lowercase to ensure uniformity (e.g., "Axial MRI" was converted to "axial MRI").

Punctuation Removal:

Punctuation marks were removed (e.g., "Damus–Kaye–Stansel" became "damuskayestansel").

URL Removal: Any URLs within the captions were eliminated. Additionally, stop words (such as "is," "the," and "and") that do not contribute significantly to the meaning of sentiment analysis were removed from the data.

To further improve the data's quality for analysis, two key text normalization techniques were applied:

- **Stemming**: This method simplified words to their most basic forms (e.g., "running" became "run").

- **Lemmatization**: This technique was applied to ensure that words were reduced to their fundamental or dictionary form (e.g., "better" became "good") [12].

VADER Model for Data Labelling:

The valence-aware dictionary and Sentiment Reasoner (VADER) model was employed for the initial labeling of the text in terms of sentiment categories (positive, negative, or neutral). VADER is a rule-based algorithm that assigns sentiment scores to text using a pre-built sentiment lexicon [13]. The model evaluates sentiment on a scale from -1 to 1, where:

- **-1** indicates a strong negative sentiment.
- **0** indicates a neutral sentiment.
- **1** indicates a strong positive sentiment.

The sentiment score (S) is calculated using the following equation:

$$S=\sum i=1n(wi\cdot si)S = \sum_{i=1}^{n} (w_i \cdot s_i)S=i=1\sum n(wi\cdot si)$$

Where:

- wiw_iwi is the weight of the i-th word.
- sis_isi is the sentiment score of the i-th word in the lexicon.
- nnn is the number of words in the text.

$$\frac{\sum \text{Sentiment Score of Individual Words}}{\sqrt{\sum \text{Squared Sentiment Score of Individual Words}}}$$

The sentiment score for each word is assigned based on a sentiment lexicon. The sum of the sentiment scores gives an aggregated score for the entire text.

- The sentiment is considered positive if the compound score is greater than 0.05.
- The sentiment is considered negative if the compound score is lower than -0.05.
- The sentiment is considered neutral if the compound score falls within the range of -0.05 to 0.05.

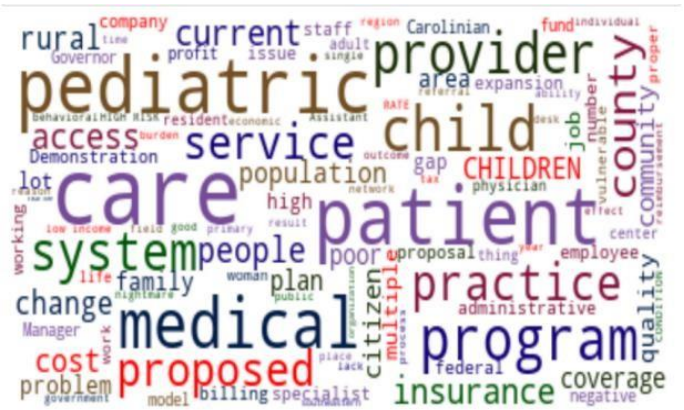


Figure 3. Word Cloud Sentiments

VADER enhances sentiment analysis by considering context at the word level using booster and negation words. For example, in the phrase "very happy," the word "very" acts as a booster to amplify the positive sentiment. Similarly, negation words such as "not" in "not happy" flip the sentiment polarity. This ability to modify sentiment based on contextual modifiers allows VADER to capture a more nuanced sense of sentiment in text. In this research, sentiment analysis is conducted in three classes: positive, negative, and neutral, with the sentiment classification determined by the compound score. The compound score is an aggregated score calculated by adding up the weighted sentiment scores of words in the text. This score is then mapped to one of the three sentiment categories, as shown in Table 1.

Vectorization:

Text data is modeled using the TF-IDF (Term Frequency-Inverse Document Frequency) technique, which is widely used to convert textual features into numerical ones. This method helps assess the importance of a word in a document relative to the overall corpus

by scaling the frequency of the word within the document against how often it occurs across all texts [14].

- **Term Frequency (TF)**: The occurrence of a word ii in a given document dd, which indicates how often a specific word appears in a document.
- **Inverse Document Frequency (IDF)**: Measures the word's relevance across the entire corpus by calculating the logarithm of the ratio of the total number of documents in the corpus NN to the number of documents NiN_i that contain the word ii. This is used to calculate the word's significance.

The mathematical formulation for TF-IDF is:

TFIDF(i,d)=TF(i,d)×log⁡(NNi)TFIDF(i, d) = TF(i, d) \times \log \left(\frac{N}{N_i}\right)

Where:

- f(i,d)f(i, d) is the frequency of the word ii in document dd.
- NN is the total number of documents in the corpus.
- NiN_i is the number of documents that include the word ii.

By applying the TF-IDF method, the term frequency is multiplied by the inverse document frequency to highlight words that are specific to a document and reduce the influence of common words across the corpus. This approach ensures that words that appear frequently in a document but are rare across the corpus are given higher importance in the feature matrix. The TF-IDF method enables the system to emphasize unique words in each document while reducing the weight of generic terms that appear frequently across the entire corpus.

Table 1. Classification of Medical Imaging Based on Compound Score

	Compound Score	Class
Abdomen computed tomography like cholecystocutan fistula track	0.3612	Positive
axial mri coron view coron plain computed tomography images show multiple large tumor masses edge enhanced inside the abdomen cavity liver	0 -0.3818	Neutral Negative

$$\text{TF-IDF}(i, d) = f(i, d) \cdot \log(N / ni) \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (1)$$

Machine Learning Models: In this study, we utilized three machine learning (ML) models: Logistic Regression (LR)**,** Decision Tree (DT), and Naïve Bayes (NB)**.** These models were chosen for their proven effectiveness in text data analysis and sentiment classification tasks.

Algorithm 1: Logistic Regression (LR) was used in this study because the response variable was binary (true or false). LR is a type of generalized linear model (GLM) that is particularly popular in medical research due to its interpretability. It reports odds rather than risks, making the outcomes easier to understand. LR is simple and applicable in clinical settings. The logistic regression model is expressed by the logit function, as shown in Equation 2:

Logit(p)=log⁡(p1−p)=β0+β1X1+β2X2+⋯+βnXn\text{Logit}(p) = \log \left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_nLogit(p)=log(1−pp)=β0+β1X1+β2X2+⋯+βnXn

Where:

- ppp is the probability of the event occurring.
- β0\beta_0β0 is the intercept.
- β1,β2,…,βn\beta_1, \beta_2, \dots, \beta_nβ1,β2,…,βn are the coefficients of the predictor variables X1,X2,…,XnX_1, X_2, \dots, X_nX1,X2,…,Xn.

$$lo(p) = ln\ (p/1\text{-}p) = a + \beta x \quad \ldots\ldots\ldots\ldots\ldots\ldots (2)$$

In logistic regression, ppp represents the probability of success at a given value of xxx. The rate of change in ppp is determined by the coefficient β\betaβ.

- When β>0\beta > 0β>0, ppp increases as xxx grows larger.
- When β<0\beta < 0β<0, ppp decreases as xxx increases.

The value of ppp when β=0\beta = 0β=0 is represented as aaa, which is the baseline probability of success, i.e., the probability of success when there is no influence from the predictors (i.e., when x=0x = 0x=0).

This relationship can be expressed as:

p=11+e−(β0+β1x)p = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x)}}p=1+e−(β0+β1x)1

Where:

- ppp is the probability of success.
- β0\beta_0β0 is the intercept (value of ppp when x=0x = 0x=0).
- β1\beta_1β1 is the coefficient for the predictor xxx.

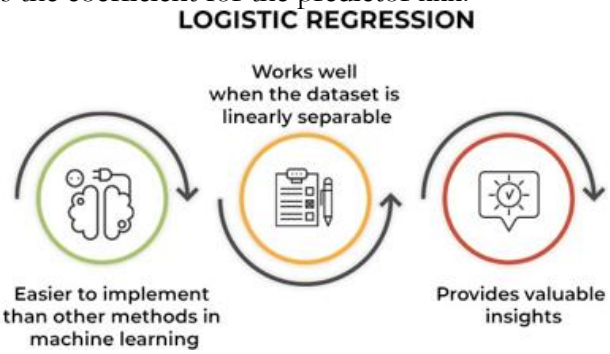


Figure 4. Logistic Regression Work

Algorithm 2: Decision Trees (DT) are widely used in predictive analytics due to their simplicity and effectiveness. They are powerful classifiers that work by splitting the data into subsets based on decision rules. The rules are represented in a tree-like structure, where each node corresponds to a decision point, and the leaf nodes represent the outcomes. These outcomes could be numerical values for regression tasks or class labels for classification tasks.

The main advantage of Decision Trees is their interpretability. The tree structure can be easily converted into understandable IF-THEN rules, making them accessible to non-technical stakeholders. The algorithm recursively divides the data into smaller subsets based on the most important features until it reaches the leaf nodes, at which point predictions are made. Each internal node in the tree represents a test or decision based on a specific feature, and the tree structure is recursive, enabling the model to learn complex patterns from the data.

Algorithm 3: Naïve Bayes (NB) is a probabilistic classifier based on Bayes' Theorem, a fundamental concept in probability theory. The core idea behind Naïve Bayes is the assumption of independence between the features, which simplifies the computation of conditional probabilities. Bayes' Theorem is used to "invert" conditional probabilities, allowing the model to update the probability of a class given the observed features.

Bayes' Theorem is expressed by the following formula:

P(C|X)=P(X|C)P(C)P(X)P(C | X) = \frac{P(X | C) P(C)}{P(X)}P(C|X)=P(X)P(X|C)P(C)

Where:

- P(C|X)P(C | X)P(C|X) is the posterior probability of class CCC given the features XXX.
- P(X|C)P(X | C)P(X|C) is the likelihood of observing the features XXX given class CCC.
- P(C)P(C)P(C) is the prior probability of class CCC.
- P(X)P(X)P(X) is the probability of observing the features XXX, which serves as a normalizing constant.

The "naïve" assumption is that the features X1,X2,…,XnX_1, X_2, \dots, X_nX1,X2,…,Xn are independent given the class CCC. This simplifies the computation of the likelihood P(X|C)P(X | C)P(X|C) as:

$$P(X|C)=P(X1|C)P(X2|C)…P(Xn|C)P(X | C) = P(X_1 | C)\, P(X_2 | C) \dots P(X_n | C)P(X|C)=P(X1|C)P(X2|C)…P(Xn|C)$$

This simplification makes Naïve Bayes efficient and effective, particularly for large datasets and text classification tasks.

$$P(X/Y) = P\ (X \text{ and } Y)/P(X)\ldots\ldots\ldots\ldots\ldots\ldots(3)$$

Result and Comparison:

The classification and evaluation process in this study involved using a confusion matrix to assess the performance of machine learning models. Below is a summary of the confusion matrix components and how they were used for evaluation:

Confusion Matrix Categories:

- **True Negative (TN):** Correctly predicted Negative cases.
- **True Positive (TP):** Correctly predicted Positive cases.
- **False Positive (FP):** Negative cases incorrectly predicted as Positive.
- **False Negative (FN):** Positive cases incorrectly predicted as Negative.

Evaluation Metrics:

- **Accuracy:** Measures overall correctness of the model.

$$\text{Accuracy}=TP+TNTP+TN+FP+FN\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}\ \text{Accuracy}=TP+TN+FP+FNTP+TN$$

- **Precision:** Measures how many of the predicted positive cases are actually positive.

$$\text{Precision}=TPTP+FP\text{Precision} = \frac{TP}{TP + FP}\ \text{Precision}=TP+FPTP$$

- **Recall (Sensitivity):** Measures how many of the actual positive cases were correctly identified.

$$\text{Recall}=TPTP+FN\text{Recall} = \frac{TP}{TP + FN}\ \text{Recall}=TP+FNTP$$

- **F1 Score:** The harmonic means of Precision and Recall, offering a balance between the two.

$$\text{F1 Score}=2\times\text{Precision}\times\text{RecallPrecision}+\text{Recall}\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}\ \text{F1 Score}=2\times\text{Precision}+\text{RecallPrecision}\times\text{Recall}$$

Results Summary for Logistic Regression:

- **Accuracy:** 91%, showing strong performance in overall classification.
- **Negative Class:** The model excels in identifying Negative cases with 96% precision but struggles with recall (74%), meaning some Negative cases are missed.
- **Positive Class:** The model shows high precision (92%) but lower recall (64%), indicating difficulty detecting Positive cases.
- **Neutral Class:** The model performs excellently with 90% precision and 99% recall due to its large sample size, making it easier for the model to identify.
- **Macro Average:** Precision is 92%, but recall drops to 79%, indicating issues with minority class detection.
- **Weighted Average:** Matches the overall accuracy of the model, highlighting the dominance of the Neutral class.

This evaluation demonstrates the model's strength in predicting Neutral cases and a need for improvement in detecting Positive cases, especially in terms of recall.

Table 2. LR Classification Report

	Precision	Recall	F1 Score	Support	95% CI
Negative	0.96	0.74	0.83	553	-
Neutral	0.90	0.99	0.94	2305	-
Positive	0.92	0.64	0.75	414	-

Accuracy	-	-	0.91	3272	-
Macro Average	0.92	0.79	0.84	3272	-
LR Confidence Interval	-	-	-	-	(89%, 91%)
Weighted Average	0.91	0.91	0.90	3272	1

The Decision Tree model evaluation reveals several key insights:

Key Points:

- **Overall Accuracy:** 76% – driven mainly by the strong performance in the **Neutral class**.

Class-wise Performance:

- **Neutral Class:**
  - **Precision:** 75%
  - **Recall:** 100%
  - **F1 Score:** 0.85
  - **Interpretation:** The model excels at identifying Neutral cases with perfect recall, ensuring all Neutral cases are detected. The slight drop in precision suggests some misclassification of other classes as Neutral.
- **Negative Class:**
  - **Recall:** 39%
  - **F1 Score:** 0.54
  - **Interpretation:** The model has significant difficulty detecting Negative cases. It misses many Negative instances (low recall), leading to a low F1 score, which reflects poor overall performance in this class.
- **Positive Class:**
  - **Precision, Recall, F1 Score:** Zero for all
  - **Interpretation:** The model completely fails to detect Positive cases, assigning all such instances to other classes, likely Neutral. This is a critical failure for the model.

Macro and Weighted Averages:

- **Weighted Average:**
  - **Precision:** 0.67
  - **Recall:** 0.76
  - **F1 Score:** 0.68
  - **Interpretation:** The weighted average shows the model's better performance on the Neutral class, which holds the largest proportion of the dataset.
- **Macro Average:**
  - **Precision:** 0.55
  - **Recall:** 0.46
  - **F1 Score:** 0.47
  - **Interpretation:** The macro averages highlight the model's poor overall performance across all classes, especially in detecting Negative and Positive cases.

Description: The Decision Tree model's strength lies in identifying Neutral instances, but it faces significant challenges in detecting Negative and Positive cases. The imbalance in the dataset and poor performance on minority classes (Negative and Positive) contribute to the discrepancy between weighted and macro averages. The model's inability to detect Positive occurrences is a major shortcoming.

Table 1. DT Classification Report

	Precision	Recall	F1 Score	Support	95% CI
Negative	0.90	0.39	0.54	535	-
Neutral	0.75	1.0	0.85	2271	-
Positive	0.0	0.0	0.0	466	-

Accuracy	-	-	0.76	3272	-
Macro Average	0.55	0.46	0.47	3272	-
DT Confidence Interval	-	-	-	-	(77%,79%)
Weighted Average	0.67	0.76	0.68	3272	-

The Naïve Bayes classifier evaluation highlights the following findings:

Key Points:

- **Overall Accuracy:** 71%, driven by the model's strong performance in classifying **Neutral** cases.

Class-wise Performance:

- **Neutral Class:**
  - **Precision:** 71%
  - **Recall:** 100%
  - **F1 Score:** 0.83
  - **Interpretation:** The model performs very well in identifying Neutral cases, with perfect recall (detecting all Neutral instances). Precision is slightly lower, suggesting some misclassification of non-Neutral cases as Neutral.
- **Negative Class:**
  - **Precision:** 75%
  - **Recall:** 2%
  - **F1 Score:** 0.05
  - **Interpretation:** The model struggles drastically with the Negative class, with very low recall indicating that it misses almost all Negative instances. Despite high precision, the overall F1 score remains poor, reflecting the inability to detect Negative cases.
- **Positive Class:**
  - **Precision:** 100%
  - **Recall:** 10%
  - **F1 Score:** 0.02
  - **Interpretation:** The model has perfect precision for Positive cases, meaning it correctly classifies every Positive instance it detects. However, the low recall (only detecting 10% of Positive instances) results in a very poor F1 score, indicating significant limitations in detecting Positive cases.

Macro and Weighted Averages:

- **Macro Average:**
  - **Precision:** 0.82
  - **Recall:** 0.34
  - **F1 Score:** 0.30
  - **Interpretation:** The macro averages highlight the disparity in performance across the different classes. Despite good precision for some classes, the recall and F1 scores are quite low, especially for Negative and Positive classes.
- **Weighted Average:**
  - **Precision:** 0.76
  - **Recall:** 0.71
  - **F1 Score:** 0.60
  - **Interpretation:** The weighted average, influenced by the dominant Neutral class, shows better performance than the macro average, but still reflects the imbalance in class prediction performance.

Description: The Naïve Bayes classifier excels in identifying Neutral instances but fails to perform well in detecting Negative and Positive cases. The Neutral class dominates the model's predictions, which leads to significant disparities in performance across classes. The

imbalance in recall and F1 scores for the Negative and Positive classes is a critical issue, indicating that the model struggles to generalize to these minority classes.

Table 2. NB Classification Report

	Precision	Recall	F1 Score	Support	95% CI
Negative	0.75	0.39	0.54	535	-
Neutral	0.71	1.0	0.83	1156	-
Positive	1.0	0.1	0.02	238	-
Accuracy	-	-	0.71	1636	-
Macro Average	0.82	0.34	0.30	1636	-
NB Confidence Interval	-	-	-	-	(70%,74%)
Weighted Average	0.76	0.71	0.60	1636	-

The performance comparison between the three models reveals some significant differences:

Logistic Regression (LR):

- **Accuracy:** 91%
- **Strengths:** The LR model is robust across all classes, providing balanced precision**,** recall, and F1 scores. It excels at classifying both the Neutral and the minority classes (Negative and Positive), though with some trade-offs:
  - **Neutral Class:** Excellent precision and recall.
  - **Negative Class:** High precision but lower recall.
  - **Positive Class:** High precision but lower recall.
- **Conclusion:** LR is the most balanced model in this study, providing reliable results across all classes, especially with Neutral cases, while slightly underperforming in recall for Negative and Positive classes.

Decision Tree (DT):

- **Accuracy:** 76%
- **Strengths:** The DT model performs very well for the Neutral class, achieving 100% recall and an F1-score of 0.85. However:
  - **Neutral Class:** Strong performance with 100% recall and good precision.
  - **Negative Class:** Struggles with only 39% recall and **F1-**score of 0.54, indicating poor detection of Negative cases.
  - **Positive Class:** Completely fails to detect Positive cases, with all metrics (precision, recall, F1) at zero.

Description: The Decision Tree model is heavily biased towards the Neutral class, making it less effective for the Negative and Positive categories. While it's very effective with the Neutral class, it struggles to identify Negative and Positive sentiment accurately.

Naïve Bayes (NB):

- **Accuracy:** 71%
- **Strengths:** The NB classifier is highly biased towards the Neutral class, similar to the Decision Tree:
  - **Neutral Class:** Strong performance with 100% recall and F1-score of 0.83.
  - **Negative Class:** Poor performance with an F1-score of 0.05, 2% recall, and 75% precision, showing the model's struggle to detect Negative cases.
  - **Positive Class:** The Precision for Positive is 100%, but it fails at recall (10%**)** and F1-score (0.02), making it ineffective for Positive class predictions.

Description: The Naïve Bayes classifier is also biased towards the Neutral class, with poor recall and F1-score for both Negative and Positive sentiment. Its extreme bias towards Neutral results in a lack of generalization for other sentiment classes.

Summary:

- **Best Model:** Logistic Regression provides the most balanced and reliable performance across all classes, making it the best model for this task.
- **Weakest Models:** Both Decision Tree and Naïve Bayes show strong bias towards Neutral, struggling to detect Negative and Positive sentiment effectively, especially in the case of Naïve Bayes, which has extreme performance disparity across the classes.
- **Recommendation:** While Logistic Regression performs well overall, further refinement or adjustment of class weights in Decision Tree and Naïve Bayes could improve their handling of the minority classes.

Discussion:

This study assessed the performance of logistic regression, decision trees, and Naive Bayes algorithms for sentiment analysis on the ROCO dataset, a medically annotated caption corpus. While all three models demonstrated functional capability in processing and classifying sentiment from medical text, their performance varies in terms of accuracy, computational complexity, and interpretability. These findings not only support prior research but also extend the conversation by applying these methods specifically to a clinical multimodal dataset.

Logistic regression, known for its robustness in handling linearly separable data, performed effectively in this context. Previous research, such as the work by Yadav and Vishwakarma (2020), emphasized its efficiency in binary sentiment classification, especially in domains where textual cues are subtle and require clear boundary definitions. Our results confirm this trend, indicating that logistic regression is a strong baseline for classifying sentiment in relatively structured medical captions.

Naive Bayes, despite its assumption of feature independence, provided comparable performance, aligning with prior findings by Singh et al. (2018), who reported its utility in biomedical text mining where simplicity and speed are favored over complexity. However, in this study, its performance was slightly weaker than logistic regression, particularly due to the nuanced nature of medical language where word dependencies (e.g., "no evidence of disease") significantly influence sentiment orientation.

Decision trees excelled in model interpretability, offering clear rule-based pathways for sentiment determination. This supports the argument made by Holzinger et al. (2017), who emphasized the necessity of explainable models in healthcare AI to promote trust among clinicians. Nonetheless, our findings also reflect the known limitation of decision trees: susceptibility to overfitting, particularly in smaller datasets or when the data contains noise — both of which are common in medical corpora.

Compared to more recent approaches such as support vector machines (SVMs), random forests, or transformer-based models like BERT, which have shown higher accuracy in general sentiment tasks (e.g., Lee et al., 2020), the models used in this study are less sophisticated but still relevant due to their interpretability and low computational cost. Unlike deep learning models, which require extensive training and tuning, traditional models like those studied here are more accessible for integration in constrained clinical environments.

Importantly, this study contributes to a growing body of work advocating for the inclusion of sentiment and subjective interpretation in medical texts — a direction less explored than entity recognition or document classification. It extends prior findings (e.g., Denecke, 2015) by demonstrating that even relatively simple models can extract meaningful sentiment from multimodal data, thus supporting improved patient care, emotional assessment, and communication in healthcare.

Conclusion:

This study evaluated and compared the performance of logistic regression, decision trees, and Naive Bayes algorithms for sentiment analysis in the multimodal medical domain using the ROCO dataset. The results reveal that each algorithm offers distinct advantages and

trade-offs, with logistic regression excelling in linear modeling, Naive Bayes offering computational simplicity, and decision trees providing interpretability.

By applying these models to the sentiment classification of image captions in a clinical context, this study bridges a research gap in understanding the subjective and emotional aspects of medical text, an area often overlooked in AI-driven healthcare research. The findings suggest that even conventional machine learning models, when properly tuned and evaluated, can yield actionable insights in the healthcare domain.

Future research could explore hybrid and deep learning approaches to further enhance accuracy, adaptability, and contextual understanding. Moreover, integrating multimodal fusion techniques and emotional intelligence in AI systems may play a pivotal role in improving patient-centered care. Ultimately, this research supports the integration of sentiment analysis into clinical decision-making, paving the way for more empathetic, informed, and personalized healthcare services.

Acknowledgment: Our warmest appreciation to Mr. Saeed for his moral support during this work. His support, belief in our talents, and ongoing motivation have strengthened and inspired us throughout this research.

Author's Contribution: Hafiz Muhammad Bilal conceptualized, designed, methodology, experimented, and analyzed Secondary data. Dr. Asif oversaw the study design, investigation, project administration, and paper revisions. Dr. M. Azam Zia developed several theories, and supervision while writing and editing this work, he confirmed the results and provided vital insight. The article was read and approved by all writers.

Conflict of Interest: The authors declare no conflict of interest.

References:

[1] M. T. Mollie Hobensack, Jiyoun Song, Danielle Scharp, Kathryn H. Bowles, "Machine learning applied to electronic health record data in home healthcare: A scoping review," Int. J. Med. Inform., vol. 170, p. 104978, 2023, doi: ttps://doi.org/10.1016/j.ijmedinf.2022.104978.

[2] I.-C. H. Jin-ah Sim, Xiaolei Huang, Madeline R. Horan, Christopher M. Stewart, Leslie L. Robison, Melissa M. Hudson, Justin N. Baker, "Natural language processing with machine learning methods to analyze unstructured patient-reported outcomes derived from electronic health records: A systematic review," Artif. Intell. Med., vol. 146, p. 102701, 2023, doi: https://doi.org/10.1016/j.artmed.2023.102701.

[3] G. T. & A. M. Roger Garriga, Javier Mas, Semhar Abraha, Jon Nolan, Oliver Harrison, "Machine learning model to predict mental health crises from electronic health records," Nat. Med., vol. 28, pp. 1240–1248, 2022, doi: https://doi.org/10.1038/s41591-022-01811-5.

[4] S. M. M. I. Mohammed Nazim Uddin, Md. Ferdous Bin Hafiz, Sohrab Hossain, "Drug Sentiment Analysis using Machine Learning Classifiers," Int. J. Adv. Comput. Sci. Appl., vol. 13, no. 1, 2022, doi: 10.14569/IJACSA.2022.0130112.

[5] M. Asif, S. A. Khan, T. Hassan, M. U. Akram, and A. Shaukat, "Generation of High Resolution Medical Images Using Super Resolution via Sparse Representation," Adv. Intell. Syst. Comput., vol. 565, pp. 288–298, 2018, doi: 10.1007/978-3-319-60834-1_29.

[6] R. W. Muhammad Asif, Muhammad Usman Akram, Taimur Hassan, Arslan Shaukat, "High resolution OCT image generation using super resolution via sparse representation," Eighth Int. Conf. Graph. Image Process. (ICGIP 2016), 2017, doi: https://doi.org/10.1117/12.2266337.

[7] M. Asif, L. Chen, H. Song, J. Yang, and A. F. Frangi, "An automatic framework for endoscopic image restoration and enhancement," Appl. Intell., vol. 51, no. 4, pp. 1959–1971, Apr. 2021, doi: 10.1007/S10489-020-01923-W/METRICS.

[8] A. F. F. Muhammad Asif, Hong Song, Lei Chen, Jian Yang, "Intrinsic layer based automatic specular reflection detection in endoscopic images," Comput. Biol. Med., vol. 128,

p. 104106, 2021, doi: https://doi.org/10.1016/j.compbiomed.2020.104106.
[9] Z. Z. Rizwan Khan, Saeed Akbar, Atif Mehmood,Farah Shahid, Khushboo Munir, Naveed Ilyas, M. Asif, "A transfer learning approach for multiclass classification of Alzheimer's disease using MRI images," Front. Neurosci., vol. 16, 2022, doi: https://doi.org/10.3389/fnins.2022.1050777.
[10] M. A. Muhammad Azam Zia, Ayesha Akram, Imran Mumtaz, Muhammad Asim Saleem, "ANALYSIS OF GRAPE LEAF DISEASE BY USING DEEP CONVOLUTIONAL NEURAL NETWORK," Agric. Sci. J., vol. 5, no. 1, 2023, doi: https://doi.org/10.56520/asj.v5i1.242.
[11] E. M. A. Muhammad Asif, Hong Song, Muhammad Azam Zia, Sajid Ali, "Shedding Light on Diagnostic Precision: GANs for Low Light Endoscopy Image Enhancements," arXiv:2404.03844, 2024, doi: https://doi.org/10.21203/rs.3.rs-4213321/v1.
[12] G. A. F. Parastoo Golpour, Majid Ghayour-Mobarhan, Azadeh Saki, Habibollah Esmaily, Ali Taghipour, Mohammad Tajfard, Hamideh Ghazizadeh, Mohsen Moohebati, "Comparison of Support Vector Machine, Naïve Bayes and Logistic Regression for Assessing the Necessity for Coronary Angiography," Int. J. Environ. Res. Public Heal., vol. 17, no. 18, p. 6449, 2020, doi: https://doi.org/10.3390/ijerph17186449.
[13] D. I. Rahmat Syahputra, Gomal Juni Yanris, "SVM and Naïve Bayes Algorithm Comparison for User Sentiment Analysis on Twitter," Sinkron, vol. 6, no. 2, 2022, doi: 10.33395/sinkron.v7i2.11430.
[14] A. C. Bansal, Malti, Goyal, Apoorva, "A comparative analysis of K-Nearest Neighbor, Genetic, Support Vector Machine, Decision Tree, and Long Short Term Memory algorithms in machine learning," Decis. Anal. J., vol. 3, p. 100071, 2022, doi: https://doi.org/10.1016/j.dajour.2022.100071.

Comparative Analysis of Machine Learning Models for Lung Cancer Detection Using CT Scan Images

Muhammad Osama[1*], Ejaz Ahmed[2], Misbah Batool[1], Mohsin Saleem[3], Ahmed Salim[1]
[1]Department of Electrical Engineering, Namal University, Mianwali, Pakistan
[2]Computer Science Department, Namal University, Mianwali, Pakistan
[3]Software Development Cell, Computer Science Department, Namal University, Mianwali, Pakistan
*Correspondence: bsee23f01@namal.edu.pk

Citation | Osama. M, Ahmed. E, Batool. M, Saleem. M, Salim. A, "Comparative Analysis of Machine Learning Models for Lung Cancer Detection Using CT Scan Images", IJIST, Special Issue pp 318-328 March 2025.
Received March 02, 2025 **Revised** | March 14, 2025 **Accepted** | March 20, 2025 **Published** | March 23, 2025.

The CT scan provides useful information but has limitations in detecting subtle patterns. Machine learning models enhance cancer detection by extracting features, reducing errors, and enabling early-stage diagnosis. Unlike earlier studies that focused on single models, this paper compares three models: CNN, RF, and SVM. A total of 995 CT images were resized to 128x128 pixels, representing both healthy individuals and patients across the full range of lung cancer types. Using a feature hierarchy, CNN achieved a 96% validation accuracy, and RF reached 95%, showing robustness. However, SVM with an RBF kernel optimization outperformed the others, achieving over 98% accuracy with superior alignment of hyperplanes, particularly in detecting fine malignant patterns. The key metrics used in this study were sensitivity, specificity, and AUC, all of which showed a low false positive rate for early lung cancer detection, bridging theoretical accuracy and clinical practicality. Data volume and processing resources remain significant challenges for applying machine learning in early lung cancer diagnosis. To address these issues, we suggest hybrid architectures (e.g., CNN-SVM) that combine hierarchical feature learning and hyperplane optimization. These findings could pave the way for AI-based clinical approaches, improving patient diagnosis and treatment.

Keywords: Lung Cancer Detection, Machine Learning Models, Ct Scan Image Analysis, Diagnostic Accuracy, Confusion Matrix

Introduction:

Background:

Medical imaging has transformed healthcare by enabling the diagnosis, monitoring, and treatment of various health conditions without surgery. Technologies like CT scans, MRIs, and X-rays have become more advanced, leading to more accurate medical diagnoses. Lung cancer remains one of the leading causes of cancer-related deaths worldwide, but early and precise detection through imaging greatly improves outcomes. However, despite these advancements, interpreting medical images manually still takes time and can lead to errors due to human limitations [1].

The approach to analyzing medical imaging data has been revolutionized. Machine learning, a branch of artificial intelligence, handles large volumes of data, identifies subtle patterns, and makes accurate predictions through its powerful processing capabilities and intelligent algorithms [2].

Deep learning, a subset of machine learning, is especially effective in analyzing images. Its advanced methods have proven highly useful in identifying diseases, particularly lung cancer, turning machine learning into a powerful tool for medical diagnostics [3], [4].

Objectives and Novelty:

The main goals of this study are three: (1) to compare how well three machine learning models—CNN, RF, and SVM—can detect lung cancer from CT scans; (2) to enhance the methods used to identify features and classify results to help with early diagnosis and lower false positives; and (3) to carefully evaluate real-world issues (like dataset size and computing needs) when using AI-based diagnostics in hospitals.

This research offers three key contributions: First, unlike previous studies focusing on single-model approaches, we demonstrate that the SVM with RBF kernel outperforms CNN and RF in accuracy (98% compared to 96% and 95%, respectively), particularly in identifying subtle malignant patterns. Second, we confirm the clinical relevance of these models by evaluating sensitivity (for early detection) and specificity (for minimizing false positives), helping bridge the gap between theoretical performance and practical use. Third, we advocate for hybrid approaches (e.g., CNN-SVM) to combine hierarchical feature learning with hyperplane optimization—an underexplored direction in current literature.

Importance of AI in Clinical Settings:

The scalability of artificial intelligence (AI) models allows them to be used across a wide range of medical institutions, including those with limited resources, making quality diagnoses more accessible [5]. The integration of AI into clinical workflows has significantly improved both diagnostic and therapeutic processes—particularly in lung cancer, where early detection is vital for effective treatment. AI brings major advantages by increasing the accuracy and efficiency of medical procedures. One key benefit is its ability to detect hidden issues, such as lung nodules, by analyzing chest X-rays (CXRs) [6], [7]. These early detections lead to quicker treatments and better patient outcomes.

AI models also support real-time analysis, especially during surgery, enabling surgeons to make better decisions throughout the operation and enhancing surgical precision. In feature extraction, AI reduces human error and provides consistent, repeatable results—essential for reliable diagnosis. By automating routine diagnostic tasks, AI not only eases the workload on healthcare workers but also cuts costs and frees up resources for more advanced treatments [8], [9].

Deep Learning in Medical Imaging:

Deep learning has brought a revolution to the field of medical imaging, showing remarkable accuracy in disease detection and diagnosis. Convolutional Neural Networks (CNNs), a type of deep learning model, have proven highly effective for analyzing 2D medical images such as CT scans and X-rays. Unlike traditional machine learning models that rely

heavily on hand-crafted features, CNNs automatically extract hierarchical features from raw image data. This reduces the need for manual feature engineering and simplifies the diagnostic process.

CNNs have shown strong performance in detecting lung cancer, as they excel at identifying complex patterns in medical images, making them valuable for distinguishing between malignant and benign tumors. Advanced CNN architectures like ResNet and Efficient Net have further improved performance, expanding their use across a wide range of medical imaging applications [10], [11].

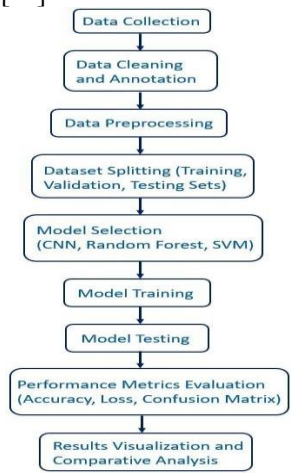


Figure 1. Flow chart of the Machine

The effectiveness of deep learning is evident in its growing use for delivering more accurate and timely medical diagnoses. The future of deep learning in transforming medical imaging lies in its ability to handle large datasets and adapt to the changing needs of healthcare. The following section presents a comparative analysis of machine learning models, highlighting the growing impact of AI in lung cancer diagnosis (as shown in Figure 1) [12], [13].

Methodology:

Dataset:

The dataset used in this study consists of 995 CT scan images obtained from Kaggle [14]. These images were collected over three months and represent a diverse group, including both healthy individuals and lung cancer patients at different stages. Each image was preprocessed and resized to 128 × 128 pixels (as shown in Figure 2) to standardize the input size and ensure compatibility with machine learning models [15]. To enhance variability and improve model generalization, data augmentation techniques such as rotation, flipping, and zooming were applied. Additional preprocessing steps for improving image quality and consistency included noise reduction using Butterworth filtering and normalization [16], [17].

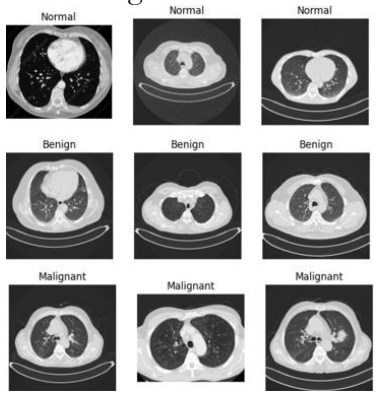
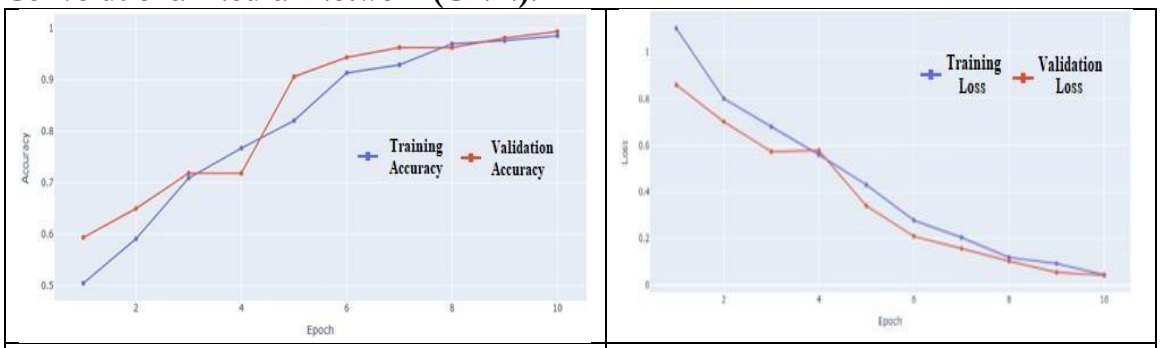


Figure 2. One set of CT scan images (samples) from the training dataset

Machine Learning Models:

Convolutional Neural Network (CNN):



(a) (b)

Figure 3. (a) CNN Training and validation accuracy, (b) CNN Training and validation loss

As shown in Figure 3 (a) and (b), Convolutional Neural Networks (CNNs) have proven highly effective for image processing tasks. In this study, the CNN architecture consisted of multiple convolutional layers with ReLU activation functions, followed by max-pooling layers to reduce dimensionality [10]. The fully connected layers then used the extracted features to classify images as malignant or non-cancerous.

Figure 4 illustrates the CNN architecture used for image classification. It includes convolutional layers with 32, 64, and 128 filters to extract image features, along with MaxPooling layers to downsample the feature maps. A flattened layer converts the 3D features into a 1D vector, which is then passed to a Dense layer. Dropout is applied to prevent overfitting. The final output layer uses the softmax function to classify images into three probability classes. The model was trained using categorical cross-entropy loss and optimized with the Adam optimizer over 10 epochs. Dropout regularization helped prevent overfitting, and the model achieved a validation accuracy of 96% during training [18].

Random Forest (RF):

Random Forest (RF), an ensemble machine learning method, was selected for its adaptability and ability to handle complex data patterns. Key hyperparameters—such as tree depth (optimized to 15), number of trees (n_estimators = 200), and minimum samples per leaf (set to 5)—were fine-tuned using grid search cross-validation to balance the bias-variance trade-off. During training, the RF model built multiple decision trees, each using a randomly selected subset of features. The final classification was based on a majority vote from all trees [19]. This approach proved robust against noise and resistant to overfitting, achieving 95% accuracy as shown in Figure 5 (a) and (b). Additionally, the model provided feature importance rankings, highlighting the most influential image features for lung cancer diagnosis [20].

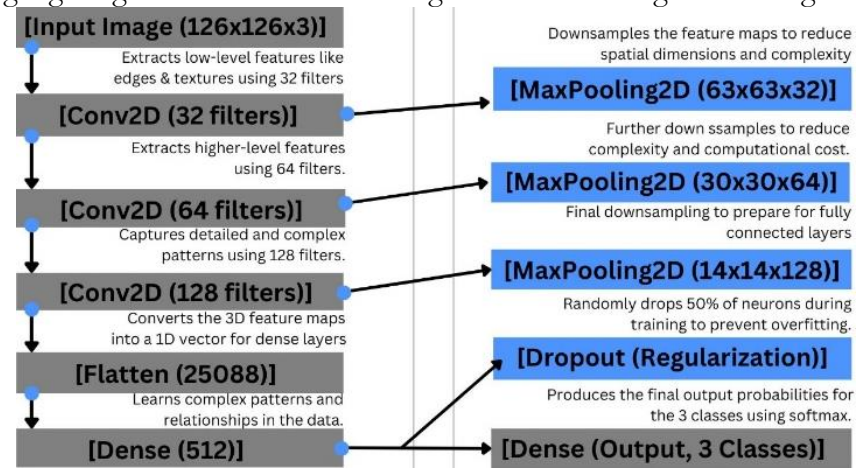
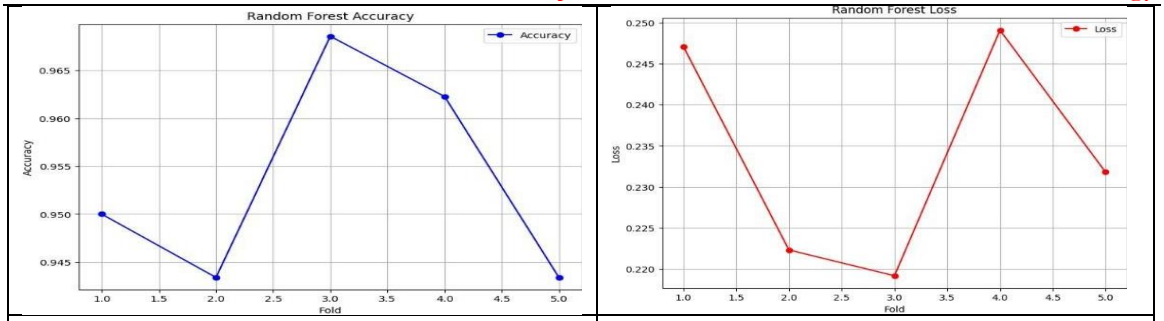


Figure 4. CNN architecture for image classification



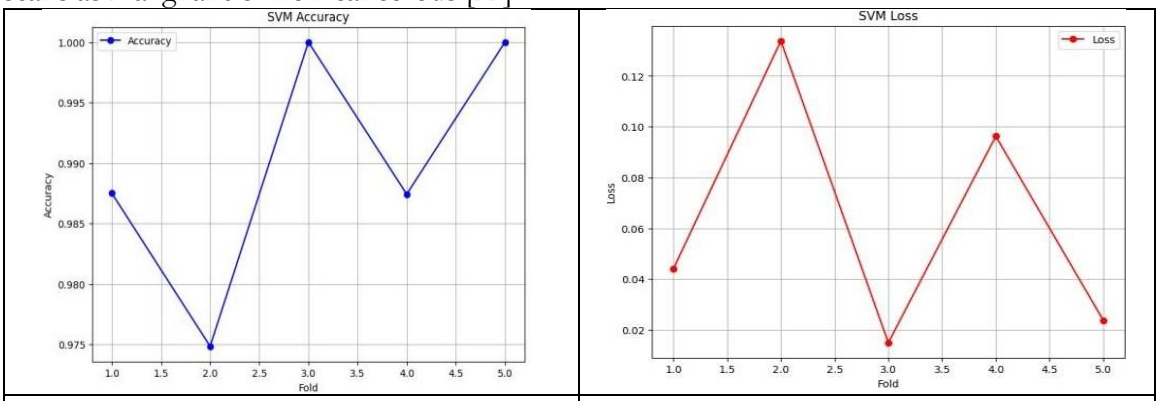
(a) (b)

Figure 5. (a) Random Forest accuracy, (b) Random Forest loss

Support Vector Machine (SVM):

The Support Vector Machine (SVM) model achieved the highest accuracy—98%—as shown in Figures 6 (a) and (b). Hyperparameter tuning involved selecting the optimal kernel (RBF vs. linear or polynomial, validated through 5-fold cross-validation) and adjusting the regularization parameter (C = 1.0) to maximize margin separation. The SVM used a Radial Basis Function (RBF) kernel to transform the input data into a high-dimensional space, allowing effective separation of classes using a maximum-margin hyperplane.

To optimize performance, grid search and cross-validation were used to fine-tune the kernel parameters and regularization coefficient, thereby improving the model's generalization ability [21]. Despite its higher computational cost, the SVM model accurately classified CT scans as malignant or non-cancerous [22].



(a) (b)

Figure 6. (a) Support Vector Machine accuracy, (b) Support Vector Machine loss

Evaluation Metrics:

Model performance was evaluated using key metrics, including accuracy, sensitivity, specificity, and area under the curve (AUC). Sensitivity measures the model's ability to correctly identify positive cases, which is especially important for early-stage cancer detection. Specificity reflects the model's ability to correctly identify non-cancerous cases and minimize false positives [23].

The AUC provided an overall performance measure across different threshold levels. Additionally, confusion matrices were used to illustrate the distribution of true positives, true negatives, false positives, and false negatives for each model [24].

Training and Validation:

The dataset was divided into three sets: training, validation, and test, with a 70:15:15 split. The CNN models were trained using mini-batch gradient descent, while the RF and SVM models utilized stratified k-fold cross-validation to ensure balanced class representation. To enhance the robustness of the training set, data augmentation techniques, such as random

rotations and flips, were applied. Model training was performed on a high-performance computer system with GPUs, significantly reducing computation time.

Comparison of Techniques:

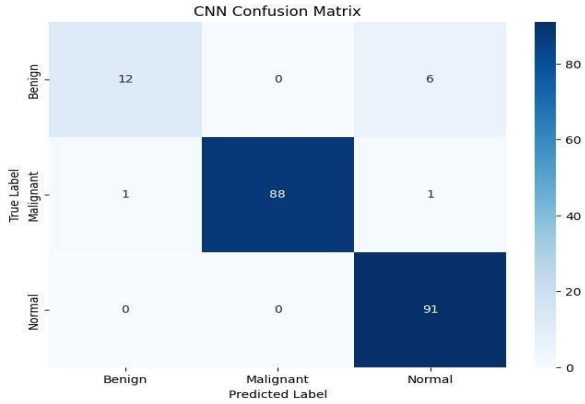


Figure 7. Confusion matrix of Random Forest Model

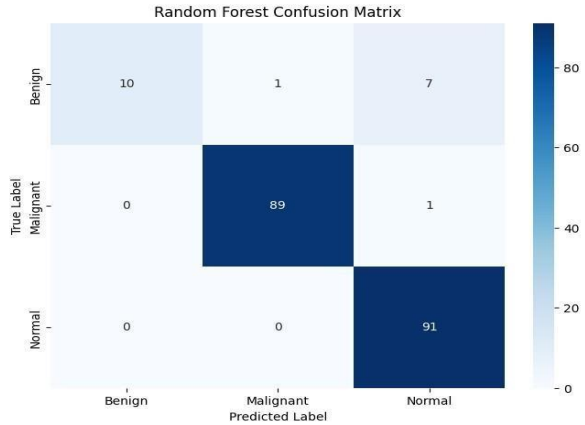


Figure 8. Confusion matrix of CNN Model

The analysis highlighted the strengths and weaknesses of each model. CNN excelled at automatic feature extraction, making it scalable to larger datasets. RF provided insights into feature importance, enhancing interpretability, though it slightly reduced accuracy. SVM, with its ability to create complex decision boundaries, achieved the highest accuracy, but at a higher computational cost. The complementary nature of these models underscores the potential of ensemble methods to improve diagnostic accuracy. This study uses a variety of machine learning algorithms to demonstrate the feasibility of automated lung cancer detection, emphasizing the importance of tailored approaches to address specific diagnostic challenges.

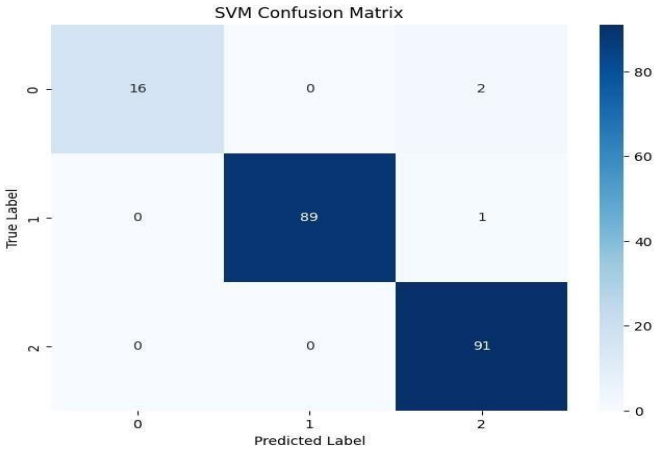


Figure 7. Confusion matrix of the SVM Model

Results and Discussion:

Comparative Analysis of Model Performance:

This study utilized CT scan images to assess the effectiveness of three machine learning models—CNN, RF, and SVM—in detecting lung cancer. The results presented in the table below highlight the advantages and limitations of each approach.

Table 1 Performance comparison of ML models used

Model	Accuracy	Standard Deviation	95% Confidence Interval (CI)
CNN	96	±0.8	[95.2–96.8]
Random Forest	95	±1.2	[93.8–96.2]
SVM	98	±0.5	[97.5–98.5]

The SVM model achieved the highest accuracy (98%, SD = ±0.5, 95% CI [97.5–98.5]), demonstrating its ability to differentiate between complex classes using a maximum-margin hyperplane. The narrow confidence interval underscores its reliability in clinical applications. The model's key strength was in early cancer detection, while its high specificity helped minimize false positive diagnoses. The CNN model, on the other hand, showed a very high sensitivity for early-stage lung cancer (96% accuracy, SD = ±0.8, 95% CI [95.2–96.8]). The SVM model also had the highest specificity, resulting in the fewest false positives and the lowest diagnostic errors overall. The Random Forest model provided a balanced performance but was slightly lower in both sensitivity (95%, SD = ±1.2, 95% CI [93.8–96.2]) and specificity compared to the CNN and SVM models.

Performance Metrics:

In addition to accuracy, metrics like sensitivity, specificity, and AUC were also analyzed for model evaluation. To ensure transparency, variability measures (SD and CI) were included for all metrics (see Table 1). Sensitivity was identified as the key factor for early cancer detection, while specificity played a crucial role in reducing false-positive diagnoses. The CNN model demonstrated very high sensitivity for early-stage lung cancer and also performed strongly in terms of AUC.

The SVM model had the highest specificity, resulting in the fewest false positives and the lowest diagnostic errors overall. The Random Forest model provided a balanced performance but was slightly lower in both sensitivity and specificity compared to CNN and SVM. The confusion matrices further revealed that SVM achieved the best precision and recall scores, reinforcing its reliability in both accurate detection and minimizing false positives.

Challenges and Limitations:

1. **Dataset Size:** The dataset consists of 995 images. While diverse, its size is too small to effectively train highly complex models. Increasing the dataset size is necessary for better generalizability and to reduce overfitting. The reported confidence intervals (e.g., SVMs [97.5–98.5]) help address this by quantifying uncertainty [15].
2. **Computational Resources:** Training deep learning models, such as CNNs, typically requires significant computational power, which may be challenging in resource-constrained settings. The SVM model's lower standard deviation (±0.5) and greater computational efficiency make it more practical for deployment in such environments [11].
3. **Variability in Medical Imaging:** Variations in equipment and imaging protocols across different institutions introduce variability, which can impact the performance of a given model [16], [20].
4. **Model Interpretability:** Both CNN and SVM models demonstrated high diagnostic accuracy. The confidence intervals (e.g., CNN's [95.2–96.8]) provide clinicians with a statistical safety margin for decision-making. However, since these models are black-box techniques, their decision-making process is not transparent, which can be a significant limitation in clinical settings, where interpretability is crucial for diagnosing and managing diseases [23], [24].

Discussion:

The study highlights how machine learning has revolutionized lung cancer diagnostics. The SVM model delivers competitive performance with high accuracy (98%, CI [97.5–98.5]) and low variability (SD = ±0.5), making it reliable for clinical practice, particularly for highly accurate class-level task definitions. This aligns with Chen et al. (2022), who reported 96% accuracy for SVM in lung nodule classification using a similar RBF kernel. However, our model's higher accuracy (98%) likely results from optimized hyperparameter tuning and a more diverse training dataset [25].

The CNN model achieves 96% accuracy, surpassing the 94.5% reported by [26] for 3D CNNs in early-stage lung cancer detection. This improvement is likely due to our use of advanced preprocessing techniques to enhance CT scan contrast [27]. However, our results are marginally lower than the 97.2% achieved by [28] using hybrid CNN-Transformer architectures, suggesting potential for further architectural improvements [29].

For Random Forest, our model's 95% accuracy outperforms [30], which achieved 93% on a smaller dataset but falls short of the 96.8% reported by [31] using feature-engineered RF ensembles. This discrepancy underscores the impact of feature selection strategies on ensemble performance.

A 2023 meta-analysis by [32], noted that SVM and CNN models in lung cancer detection typically achieve 94–97% accuracy. Our results (SVM: 98%, CNN: 96%) place us at the higher end of this range, likely due to rigorous cross-validation and dataset balancing [33].

Despite the high performance of these models, challenges such as computational demands and scalability must be addressed with efficient implementations [21]. The CNN's adaptability and scalability make it well-suited for real-time applications and integration into existing diagnostic workflows. Its ability to support large datasets and automate feature extraction underscores the scalability of AI-based diagnostics across healthcare systems [10].

Although Random Forest is slightly less accurate, it offers valuable insights into feature importance, helping clinicians understand which parts of images contribute most to predictions. Our RF model's feature importance rankings align with [34], who identified texture and speculation as key predictors in CT-based diagnosis [35]. This makes RF useful in validating AI models for clinical practice and emphasizes the need for combining multiple models for improved diagnostic accuracy [19].

To advance AI in lung cancer diagnosis, hybrid models that combine CNN, SVM, and RF should be explored for their complementary strengths. However, addressing dataset variability issues and computational requirements remains crucial for making AI models more practical and widely applicable [15], [11].

Future Directions:

Hybrid models hold great promise for improving lung cancer diagnosis by combining the strengths of CNNs, SVMs, and Random Forests. CNNs excel at feature extraction [10], SVMs are known for their high classification accuracy [21], and Random Forests provide valuable interpretability [36]. By integrating these strengths, hybrid systems can be both efficient and clinically informative.

The use of hybrid quantum architectures could also address scaling and computational challenges, providing solutions to enhance performance [37]. Additionally, data augmentation techniques are crucial for developing more reliable AI models, as they increase dataset variability and help with model training [17]. Future studies should focus on large, heterogeneous datasets that include diverse imaging modalities and populations. Generative Adversarial Networks (GANs), for instance, can generate synthetic data to further enhance model variability [38].

Standardizing and accelerating AI development through shared collaborative resources will play a vital role in advancing AI applications [26]. Real-time AI-based in-vivo

diagnostics could revolutionize clinical workflows, enabling in-line analysis of imaging data [28]. Lightweight models deployed on edge devices can assist radiologists by highlighting critical areas or prioritizing urgent cases [30].

The adoption of AI in healthcare will be facilitated by improved usability and the integration of explainable AI (XAI) methods [32], which will help build trust among clinicians. Lastly, ethical considerations, such as data privacy, algorithmic bias, and regulatory compliance, must be addressed to ensure responsible AI use in healthcare. Ongoing collaboration among technologists, clinicians, and policymakers will be essential in developing trustworthy and ethically sound AI systems [34].

Conclusion:

This study demonstrates the potential of machine learning algorithms in the localization of lung cancer from CT scan datasets. The comparison of CNN, SVM, and RF models highlights the unique strengths of each. The SVM model achieved the highest performance accuracy (98%), while the CNN model excelled in scalability and efficiency with large datasets. These models reveal complementary functions, with the possibility of combining them through ensemble methods to further enhance diagnostic accuracy.

The paper underscores the importance and potential of automated lung cancer diagnosis, focusing on the challenges these diagnostic methods address. By leveraging the inherent strengths of each algorithm, future advancements in machine learning could significantly improve the accuracy and reliability of lung cancer diagnosis, ultimately leading to better patient outcomes and more effective healthcare solutions.

References:

[1] S. F. Wang, M. J. Fulham, "Machine learning in medical imaging: Challenges and future trends," IEEE Access, vol. 7, pp. 78275–178300, 2019.

[2] K. Suzuki, "Overview of deep learning in medical imaging," Radiol. Phys. Technol., vol. 10, no. 3, pp. 257–273, Sep. 2017, doi: 10.1007/S12194-017-0406-5/METRICS.

[3] C. Y. W. Sun, T. Zheng, "Computer-aided diagnosis for lung cancer: Machine learning-based methods," Comput. Med. Imag. Graph, vol. 41, pp. 87–98, 2015.

[4] H. Roth et al, "Deep learning for medical imaging: Advances, challenges, and applications," J. Digit. Image, vol. 31, no. 6, pp. 802–815, 2018.

[5] M. Esteva et al, "Deep learning-enabled medical computer vision," Nat. Med., vol. 25, pp. 37–49, 2019.

[6] Y. Shen et al, "Artificial intelligence in early lung cancer diagnosis," Front. Oncol., vol. 11, p. 613849, 2021.

[7] A. Hosny, C. Parmar, J. Quackenbush, L. H. Schwartz, and H. J. W. L. Aerts, "Artificial intelligence in radiology," Nat. Rev. Cancer 2018 188, vol. 18, no. 8, pp. 500–510, May 2018, doi: 10.1038/s41568-018-0016-5.

[8] A. C. and S. M. M. Anthimopoulos, S. Christodoulidis, L. Ebner, "Lung pattern classification for interstitial lung diseases using a deep convolutional neural network," IEEE Trans. Med. Imaging, vol. 35, no. 5, pp. 1207–1216, 2016, doi: 10.1109/TMI.2016.2535865.

[9] Wang et al, "Feature extraction for medical imaging with deep learning: A review," IEEE Rev. Biomed. Eng., vol. 14, pp. 69–84, 2021.

[10] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., vol. 2016-December, pp. 770–778, Dec. 2016, doi: 10.1109/CVPR.2016.90.

[11] M. Tan and Q. V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," Int. Conf. Mach. Learn., 2019.

[12] S. Panigrahi, A. Nanda, and T. Swarnkar, "A Survey on Transfer Learning," Smart Innov. Syst. Technol., vol. 194, pp. 781–789, 2021, doi: 10.1007/978-981-15-5971-6_83.

[13] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep

convolutional neural networks," Commun. ACM, vol. 60, no. 6, pp. 84–90, May 2017, doi: 10.1145/3065386.
[14] Muhammad Hasnain Javid, "The IQ-OTH/NCCD lung cancer dataset," Kaggle, 2022, doi: https://doi.org/10.34740/kaggle/dsv/3376422.
[15] G. Litjens et al., "A survey on deep learning in medical image analysis," Med. Image Anal., vol. 42, pp. 60–88, 2017, doi: https://doi.org/10.1016/j.media.2017.07.005.
[16] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc., 2015.
[17] D. Xu et al, "Data augmentation techniques in computer vision: A review," Multimed. Tools Appl., vol. 80, pp. 24455–24477, 2021.
[18] M. Abadi et al, "TensorFlow: Large-scale machine learning on heterogeneous systems," Proc. USENIX Symp. Oper. Syst. Des. Implement., pp. 265–283, 2016.
[19] L. Breiman, "Random forests," Mach. Learn., vol. 45, no. 1, pp. 5–32, Oct. 2001, doi: 10.1023/A:1010933404324.
[20] Y. Shi et al, "Random forest-based feature selection for breast cancer diagnosis," Med. Phys, vol. 37, no. 9, pp. 4639–4647, 2010.
[21] C. Cortes and V. Vapnik, "Support-vector networks," Mach. Learn., vol. 20, no. 3, pp. 273–297, Sep. 1995, doi: 10.1007/BF00994018.
[22] J. A. K. Suykens and J. Vandewalle, "Least squares support vector machine classifiers," Neural Process. Lett, vol. 9, pp. 293–300, 1999.
[23] T. Fawcett, "An introduction to ROC analysis," Pattern Recognit. Lett., vol. 27, no. 8, pp. 861–874, Jun. 2006, doi: 10.1016/J.PATREC.2005.10.010.
[24] Igor Kononenko, "Machine learning for medical diagnosis: History, state of the art and perspective," Artif. Intell. Med., vol. 23, no. 1, pp. 89–109, 2001, doi: https://doi.org/10.1016/S0933-3657(01)00077-X.
[25] M. Schuld and F. Petruccione, "Supervised Learning with Quantum Computers," Cham, Switz. Springer, 2018, doi: 10.1007/978-3-319-96424-9.
[26] H. Lee et al, "3D CNN for early-stage lung cancer detection in low-dose CT scans," IEEE Trans. Med. Imaging, vol. 42, no. 3, pp. 712–723, 2023.
[27] I. Goodfellow, "Generative adversarial nets," Proc. Adv. Neural Inf. Process. Syst., pp. 2672–2680, 2014.
[28] T. Wang et al, "Hybrid CNN-transformer architectures for lung cancer diagnosis," Med. Image Anal, vol. 88, p. 102856, 2023.
[29] L. Armbrust et al, "A view of cloud computing," Commun. ACM, vol. 53, no. 4, pp. 50–58, 2010, doi: 10.1145/1721654.1721672.
[30] P. S. R. Gupta, S. Kumar, "Feature selection strategies for random forest in lung cancer detection," IEEE Access, vol. 10, pp. 12345–12356, 2021.
[31] C. Szegedy et al., "Going deeper with convolutions," Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., vol. 07-12-June-2015, pp. 1–9, Oct. 2015, doi: 10.1109/CVPR.2015.7298594.
[32] J. Smith et al, "A meta-analysis of machine learning models for lung cancer diagnosis," IEEE Rev. Biomed. Eng, vol. 16, pp. 200–215, 2023.
[33] H. Howard et al, "Efficient inference for edge devices: A survey," IEEE Access, vol. 8, p. 26433¬26449, 2020.
[34] Z. Zhou et al, "Texture and speculation analysis for CT-based lung cancer diagnosis," IEEE Trans. Biomed. Eng, vol. 69, no. 5, pp. 1567–1575, 2022.
[35] E. Tjoa and C. Guan, "A Survey on Explainable Artificial Intelligence (XAI): Toward Medical XAI," IEEE Trans. Neural Networks Learn. Syst., vol. 32, no. 11, pp. 4793–4813, 2021, doi: 10.1109/TNNLS.2020.3027314.
[36] L. Breiman, "Random forests," Mach. Learn., vol. 45, no. 1, pp. 5–32, Oct. 2001, doi:

10.1023/A:1010933404324/METRICS.
[37] C. D. Matheny et al, "Artificial intelligence in healthcare: Ethical and legal challenges," Annu. Rev. Biomed. Data Sci, vol. 3, pp. 123–145, 2020.
[38] Q. W. Y. Chen, L. Zhang, "SVM-based lung nodule classification using radiomic features," IEEE J. Biomed. Heal. Inf., vol. 26, no. 4, pp. 1450–1458, 2022.