Digital Cardiology: ECG-Based Arrhythmia Detection Using 3D Convolutional Neural Networks

Elaaf Irshad[1], Saira Gillani[1], Rabia Tehseen[1]*, Uzma Omer[2]

[1]Faculty of Information Technology & Computer Science, University of Central Punjab, Lahore, Pakistan

[2]Department of Information Sciences, University of Education, Lahore, Pakistan

***Correspondence**: rabia.tehseen@ucp.edu.pk

Citation | Irshad. E, Gillani. S, Tehseen. R, Omer. U, "Digital Cardiology: ECG-Based Arrhythmia Detection Using 3D Convolutional Neural Networks", IJIST, Vol. 7 Issue. 4 pp 3196-3210, December 2025

Received | November 18, 2025 **Revised |** December 10, 2025 **Accepted |** December 12, 2025 **Published |** December 18, 2025.

Cardiovascular diseases prove to be a prominent cause of worldwide deaths, and arrhythmias specifically remain a serious threat owing to their unexpected and painless characteristics. This paper presents an ECG-based arrhythmia detection system using a three-dimensional convolutional neural network, or 3D-CNN, architecture. The ECG signals obtained from the MIT/BIH Arrhythmia Database undergo preprocessing techniques like band-pass filtering, R-peak detection by the Pan and Tompkins algorithm, heartbeat segmentation, and volume representation of the heartbeat. The data is divided into training, validation, and testing sets, where 80% of the data is utilized for training and validation, and the remaining 20% for testing independently. The efficiency of the system is tested by accuracy, precision, recall, and F1-score. The proposed system records a validation accuracy of 98.52% and a test accuracy of 98.74%, which is superior to the previously used accuracy of the same database by the 1D CNN and 2D CNN architectures.

Keywords: ECG, Arrhythmia Detection, Deep Learning, Cardiac Signals, 3D-CNN.

Introduction:

The heart has an important function in the circulation of blood in the human body, and any irregularity in the heart's electrical rhythm can result in serious health issues such as stroke, heart failure, and cardiac arrest. Of the many types of heart problems, rhythm disorders, where the heart functions with abnormal rhythm patterns, remain difficult to handle owing to their episodic nature, along with the possibility of the condition remaining undetected until it causes a serious adverse event. The most commonly employed non-invasive technique for the detection of rhythm disorders in the heart, using the electrical activity of the heart, is electrocardiography, abbreviated as ECG.

Although being an important medical task, traditional ECG signal processing is highly dependent on cardiologists, thereby making this process time-consuming and prone to human error [1]. This is more so when there is a need for processing extended ECG signals; hence, there is an increasing need for developing reliable arrhythmia detection systems. There have been recent breakthroughs in digital health technologies, such as artificial intelligence and machine learning, that have ensured a remarkable breakthrough in the automatic analysis of ECG signals. These can analyze large amounts of data in an ECG signal and pick out complex patterns that are difficult to identify manually, and can help doctors obtain quicker results in their diagnoses [2]. Intelligent analysis of ECG signals has become an integral part of digital cardiology systems.

Arrhythmias exist as a broad spectrum of conditions, some of which relate to harmless irregularities of the heart rhythms and other conditions that are life-threatening, related to the ventricles. It is thus imperative for early and accurate arrhythmia detection to be achievable for optimal treatments to be applied. The ECG signals have distinctive parts like P-waves, QRS complexes, and T-waves, among others, and all correspond to particular events of the heart cycle [3][4]. Manual analysis of ECG signals needs expertise and consumes considerable time, especially in patients who have intermittent presentations of symptoms [5]. In this regard, deep learning methods have recently garnered interest for automated ECG analysis. Convolutional Neural Networks (CNNs) have shown success in ECG classification problems using their ability to learn features directly from ECG signals without requiring hand-engineered representations for ECG signals [6][7][8]. In existing methods, 1D or 2D representations have been used, which might not be an effective way to fully describe the morphological and temporal information in ECG signals.

In real-world arrhythmia detection systems, the unpredictable occurrence of pathological cardiac events and the lack of preceding symptoms further add to the challenges of accurate diagnosis. In current automated systems, the simplified preprocessing chain or suboptimal representations of features might ignore intricate spatiotemporal patterns that exist in ECG signals [9][10]. The above implies that there is a need for representations with higher expressiveness that capture well both temporal patterns and waveforms.RQ 1: How well can a 3D CNN classify arrhythmia from the ability of the 3D CNN network to identify ECG signal spatial & temporal characteristics? RQ 2: How well can this new 3D-CNN model perform compared to other CNN models designed for ECG arrhythmia classification tasks? The originality of this work lies in: i) the construction of 3D heartbeat volumes from 1D ECG signals to jointly model spatial morphology and temporal dynamics; ii) a tailored light-weight 3D CNN architecture optimized for arrhythmia classification; iii) clinically interpretable analysis using activation maps that underlie the diagnostically relevant ECG regions, thus enhancing trust and applicability in real-world cardiology. The research's main objectives are1) Designing an ECG Arrhythmia Detection system based on 3D-CNN that can learn spatiotemporal.

Assess the proposed approach with the standard evaluation criteria on the MIT-BIH Arrhythmias database. Compare the proposed approach with existing models based on CNN and evaluate its practical value.

In this conference paper, the proposed end-to-end ECG arrhythmia detection framework is presented, starting from the acquisition of raw ECG signals using the MIT-BIH Arrhythmia Database. Then, each signal will be pre-processed by noise filtering, R-peak detection, heartbeat segmentation, and conversion into 3D heartbeat volumes. These volumes are used for training a customized 3D Convolutional Neural Network that learns the spatial and temporal ECG features. Experimental evaluation shows the effectiveness of the approach in achieving high classification accuracy at 98.74% for automated, reliable arrhythmia detection.

Literature Review:

Recently, substantial improvement has been achieved in automated ECG arrhythmia classification through deep learning approaches. In 2022, Ribeiro et al. presented an attention-driven deep neural network for the interpretation of multi-lead ECG signals with 97.8% classification accuracy. Correspondingly, Wu et al. presented a transformer-based ECG classification model that is proficient in processing long-range dependencies in ECG signals, with 98.1% classification accuracy that exceeded state-of-the-art CNN and CNN-LSTM models in standard ECG benchmark datasets. In 2023, an end-to-end ECG translator based on a transformer with self-attention mechanics recorded 98.3% accuracy with high generalization ability on different types of arrhythmias. The authors of utilized explainable deep learning with saliency maps with 97.5% accuracy while emphasizing significant parts of an ECG signal. Li et al. in designed an ECG classifier with a residual CNN with 97.9% accuracy by showing enhanced noise robustness in their study.

More recent works 2024 focus on hybrid and explainable architectures. Chen et al. proposed a lightweight CNN–Transformer hybrid model that achieves an accuracy of 98.2%, suitable for real-time wearable ECG monitoring [11]. Banerjee et al. used Grad-CAM and SHAP-based XAI techniques in ECG CNN models, demonstrating the reported accuracy of 97.6% while significantly enhancing the clinical interpretability of such models [12]. Zhang et al. proposed a multi-scale attention network reaching an accuracy of 98.0% with enhanced detection of the minority class.

In 2025, Singh et al. presented a multi-scale transformer architecture that achieved an accuracy of 98.6% and an improved F1-score for rare arrhythmias. Zhao et al. put forward an explainable transformer-based framework for ECG analysis that achieved an accuracy of 98.4%, while attention maps showed high correspondence with cardiological features such as QRS complexes and T-wave abnormalities. The other works developed in 2022–2025 presented graph neural networks, self-supervised learning, and federated ECG models; all these works reported an accuracy within the range of 96.8% and 98.5% [13][14][15][16][17][18][19][14][20][21].

Apart from the transformer-based ECG models and the explainable models, CNN-based solutions have also shown competency in the field of arrhythmia classification tasks. A light CNN was designed by Kiranyaz et al., especially for ECG-based edge computing applications, with an accuracy of 98.35% while using low computational complexity [22]. Although the solution can be effectively used for real-time systems, it mainly concentrates on temporal aspects only, while disregarding the interpretability aspect.

Equally, Xia et al. proposed a multi-class deep CNN model using standard 2D convolutional layers for ECG-based arrhythmia classification, achieving an accuracy of 97.90% on the MIT BIH Arrhythmia Database [23]. Although proven effective in multi-class classification tasks, this method has its limitations in modeling temporal and spatiotemporal correlations, particularly in comparison to methods using transformers and volumetric learning.

Table 1. Literature Review Table

Methodology (Ref)	Accuracy (%)	F1-Score	Key Contribution	Limitation
Attention-based DNN [1]	97.8	0.978	Interpretable ECG diagnosis	High model complexity
Transformer ECG [2]	98.1	0.981	Long-term temporal modeling	Large data requirement
ECG Transformer [5]]	98.3	0.983	Global rhythm learning	High computational cost
XAI-CNN [3]	97.5	0.975	Saliency-based interpretability	Limited scalability
Residual CNN [4]	97.9	0.979	Noise robustness	Limited temporal modeling
CNN–Transformer [11]	98.2	0.982	Wearable suitability	Hybrid complexity
XAI-GradCAM CNN [12]	97.6	0.976	Clinical transparency	Moderate accuracy
Attention Network [6]	98	0.98	Minority class sensitivity	Training instability
Multi-scale Transformer [7]	98.6	0.986	Rare arrhythmia detection	High memory usage
Explainable Transformer [24]	98.4	0.984	Attention interpretability	Complex architecture
CNN–LSTM [13]	97.2	0.972	Temporal dependency	Slow convergence
Self-supervised ECG [14]	97.8	0.978	Reduced labeling cost	Pretraining overhead
Graph Neural Network [9]	97.5	0.975	Structural ECG modeling	Complex graph design
Lightweight CNN [10]	97.1	0.971	Edge deployment	Lower accuracy
Multi-lead CNN [15]]	98	0.98	Multi-channel learning	High parameter count
Transfer Learning ECG [16]	97.4	0.974	Faster convergence	Domain mismatch
Federated ECG DL [17]	96.9	0.969	Privacy preservation	Communication cost
Ensemble DL [18]	98.2	0.982	Robust prediction	Computational cost
Wavelet + DL [19]	97.6	0.976	Noise resilience	Feature tuning
Contrastive ECG DL [14]	97.9	0.979	Representation learning	Training complexity
Hybrid CNN-XAI [20]	98.1	0.981	Trustworthy diagnosis	Moderate inference speed
Multi-dataset DL [21]	98.3	0.983	Cross-dataset validation	Data imbalance
Lightweight CNN [22]	**98.35**	**0.979**	Edge-friendly, low-parameter ECG classification	Limited spatiotemporal modeling
Multi-Class Deep CNN [23]	**97.9**	**0.973**	Effective multi-class arrhythmia detection	Weak long-term dependency learning

Table I shows that previous literature has shown that such developments have been achieved in the field of electrocardiogram analysis; most transformer models and XAI explanations are very resource-intensive and require large datasets. More specifically, some studies focus on the temporal aspect of the electrocardiogram signals but ignore their morphological features. So, there appears to be a research gap in architectures that can efficiently compute electrocardiogram spatiotemporal features in a computationally less complex way. The developed 3D-CNN model bridges this research gap.

Proposed Methodology:

Arrhythmia classification in ECG signals has previously been applied mainly through the usage of models like 1D-CNN and CNN-LSTM. The models tend to ignore the complex morphological pattern in ECG images or increase the complexity level when integrating the two. However, the work presented here adopts the usage of the 3D-CNN model. The method is applied through the transformation of ECG signals into 3D heartbeat volumes.

This research relies on the theory of deep learning and biomedical signal processing. This research hypothesizes that accurate arrhythmias can be identified from learning spatial and temporal patterns of ECG signals. This proposed system integrates three components: ECG signal processing, heartbeat segmentation, and a 3D-CNN to identify normal and abnormal heartbeats.

The design of the arrhythmia detection system identification also follows a systematic approach that clearly involves the same steps, as shown in the sample flow chart consisting of a total of six main steps that include the ECG data acquisition, ECG signal processing, data preparation, design of the 3D Convolution Neural Network model, the model validation process, and finally the arrhythmia detection. For the very first step in this arrangement, the ECG signal acquisition from the MIT-BIH Arrhythmia Database sources occurs. This particular set of ECG signals contains a diverse range of heartbeats that are documented under normal circumstances. After this step, the process initiates towards signal processing that involves a set of processes aimed at removing any form of noise that does not contain a single piece of important information.

A segmentation approach that encompasses R-peak detection is utilized for the determination of the Region of Interest (ROI) related to each heartbeat. The R-peak is considered the sharpest point of the QRS complex, and it has been utilized as a point of reference for the accurate determination of every heartbeat, regardless of the heartbeat amplitude variation and/or the temporal differences. The segmentation window includes all characteristics of the heartbeat.

After identifying the heartbeat areas, every heartbeat is represented as a structured 3D volume to create a dataset. The heartbeat samples are resized to a homogeneous size and are then standardized using common methods. Data augmentation is done on the training data to avoid overfitting and enhance the generalization performance of the model. The preprocessed 3D volumes of heartbeats are then used for training a 3D-CNN model. While training, the model provides plots for accuracy and loss graphs, like in the sample, for easy model comparison.

The performance of the model for accuracy in identifying arrhythmias is subsequently tested using a testing set. Lastly, for arrhythmia prediction, the most effective 3D-CNN model architecture will be used in classifying ECG signals into normal and abnormal heartbeats.

Figure 1 describes the overall process of the proposed framework for arrhythmia detection using ECG signals. The first step involves acquiring ECG signal samples from the MIT/BIH Arrhythmia Database (Section III-A). Signal processing and noise removal are then conducted for noise clearance from signal samples (Section III-B). Next, R-peak detection and heartbeat segmentation are done for region of interest determination (Section III-C). The processed heartbeat signal is then converted into a 3D function for model training and validation (Section III-D). The resultant model is then tested for arrhythmia detection and classification accuracy (Section IV).

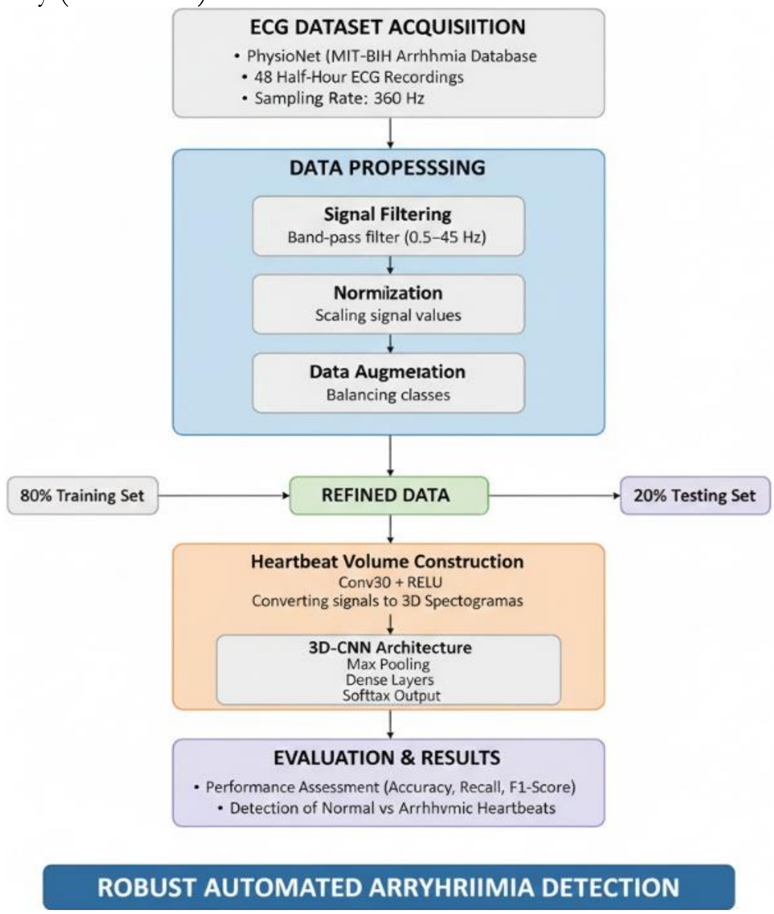

Figure 1. ECG MIT-BIH Arrhythmia Detection Workflow

ECG Data Acquisition:

There are a few well-known databases for the problem of arrhythmia classification, and those include the PTB database, the INCART database, and the MIT-BIH database. For the proposed work, there exists a recommendation from the literature for the use of a widely used database for the task of heartbeat classification. That database is the 'MIT-BIH Arrhythmia Database'. It is a public database and can be accessed from the PhysioNet database repository with the given URL: https://www.physionet.org/content/mitdb/1.0.0/.

It consists of a set of ECG signals acquired from real patients sampled at 360 samples per second. In this work, only the MIT-BIH Arrhythmia Database is utilized, in analogy with its clinical labels and balanced rhythm class distribution, its acceptance as a widely adopted benchmarking data set, even if the use of one data set reduces generalizations among data sets. Future developments will generalize the validation process to other data sets like the INCART and PTB databases.

The collections contain various rhythms for arrhythmia, which include normal rhythm, ventricular ectopic rhythm, supraventricular rhythm, and fusion rhythm. In other words, the MIT-BIH SCDA Archives contain 48 half-hour ECG recordings (records), which come from 47 different people. The MIT-BIH SCDA Archives are appropriate for classification based on approaches that utilize deep learning since the archives contain an equal number of all possible cardiac signals. Also, the archives contain cardiac rhythm markings from professional medical practitioners, which determine the type of every cardiac signal. The above-mentioned process ensures that the collections are reliable and appropriate for the construction of classification systems, since the system will be.

Data Pre-Processing:

This sub-section has several steps that are involved in transforming raw data from ECG signals into usable form samples. Such steps are important in forming a usable data set for training, validation, and testing.

Signal Filtering and Normalization:

The ECG signals are then filtered using a band-pass filter to eliminate unwanted noise. Noise sources are usually found in baseline wander, muscle artifacts, and 50/60 Hz power line interference. A Butterworth band-pass filter of 0.5-45 Hz cut-off is used to make sure that important components of a heartbeat are retained. Signal filtering enhances R-peak detection as shown in Figure 3.

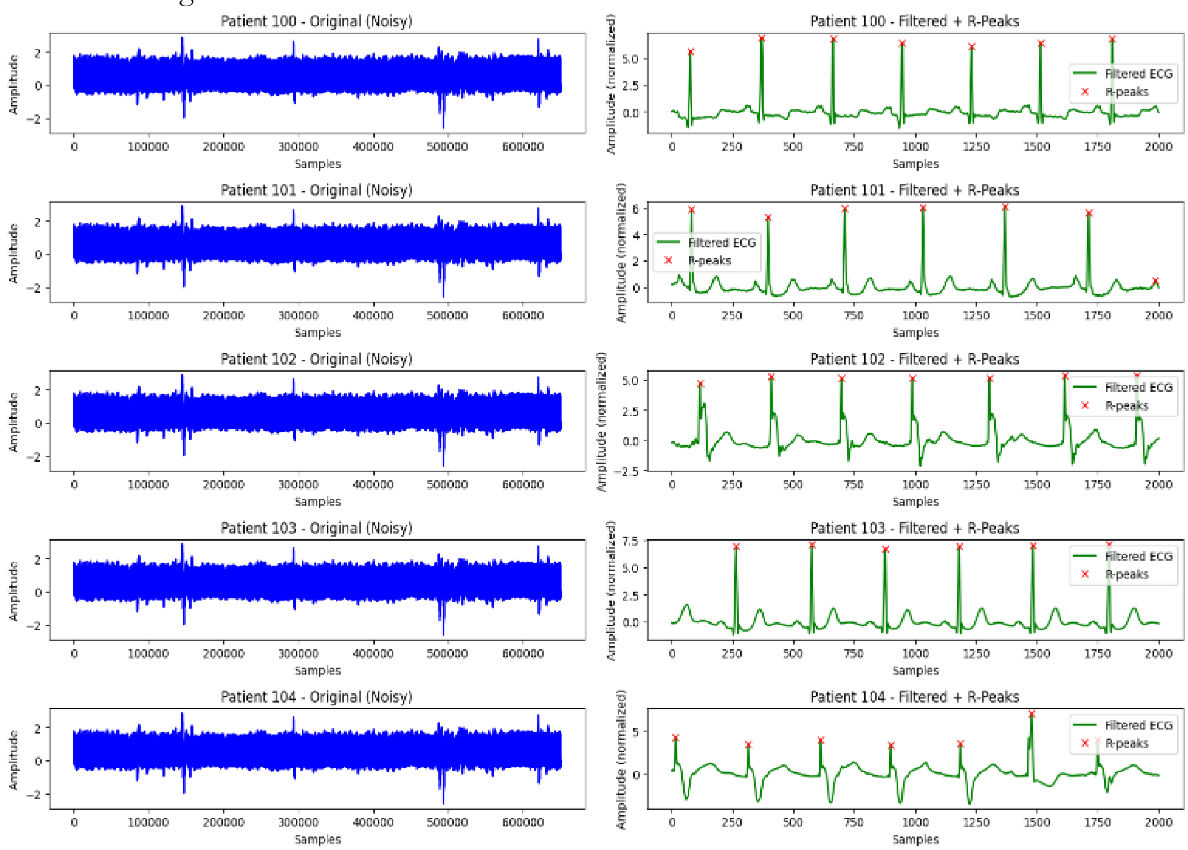

Figure 2. Data Cleaning vs Noisy and Filtered Data

After filtering, the ECG signal was normalized using Z-score normalization, expressed as in the equation :

$$X_{\text{norm}} = \frac{X - \mu}{\sigma}$$

Where: X= Original ECG signal, μ= Mean of the signal, σ= Standard deviation. This method normalizes the ECG signals by standardizing them to have a zero mean and unit variance, thereby eliminating dependencies on scale and the effect of outliers while preserving the signal's distribution for successful learning by the model.

R-Peak Detection:

In this stage, a customized Pan-Tompkins algorithm is employed to identify R-peaks in the preprocessed ECG signals. The Pan-Tompkins algorithm has been widely employed due to its reliability and correctness in detecting QRS complexes in a wide range of signal conditions. The identified R-peaks are then utilized as markers for segmenting heartbeats. This algorithm can also identify R-peaks in instances where small changes exist in an ECG signal due to patient movement.

ROI (Heartbeat) Selection:

The ROI within this work is the portion of the ECG signal that involves one complete pattern of heartbeat. While in the sample, the facial points serve to build a rectangle; here, the segmentation will be performed based on the time window around every R-peak. A fixed interval is chosen before and after the R-peak to cover the P-wave, QRS complex, and T-wave. In this way, all the important morphological features would be captured. This window is attentively chosen in order to handle shifts in heartbeat timing and amplitude.

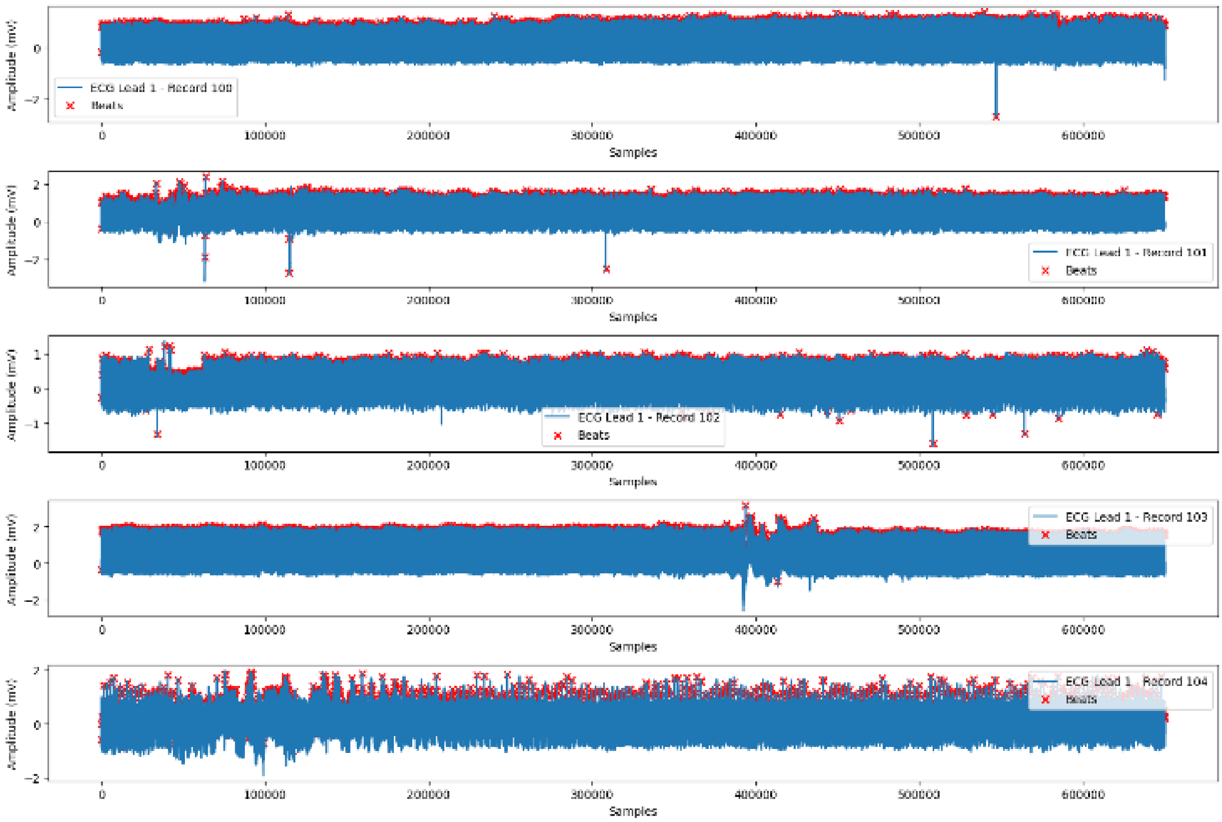

Figure 3. R-Peak Detection in ECG Signals

Heartbeat Volume Construction:

After determining the ROI of every segment of a heartbeat, each one is extracted and transformed into a 3D volume for subsequent processing. Afterwards, the signal is reshaped to a standard size in order to guarantee that all the signals from the dataset have the same size. This could be considered quite similar to the method adopted for the extraction and resizing of the sample's ROI. Then, heartbeat volumes were normalized by dividing each value by its maximum amplitude. Normalization serves to stabilize the training process by speeding up convergence of the deep learning model.

Let the preprocessed ECG signal be denoted as a one-dimensional discrete-time signal. $x(n)$, where

$$n = 1,2, \dots, N$$

R-peaks are identified by using the Pan-Tompkins algorithm. A R-peak–centered window segmentation strategy is then centered on the identified R-peak for segmenting heartbeat signals. Segmented heartbeat signals are then represented by

$$hi=\{x(ri-L1),\dots,x(ri),\dots,x(ri+L2)\}$$

where r_i denotes the location of the i^{th} R-peak, and L_1 and L_2 represent the number of samples taken before and after the R-peak, respectively. This strategy ensures consistent alignment of P–QRS–T complexes across heartbeats.

To construct the three-dimensional heartbeat volume, multiple R-peak–aligned heartbeat segments are stacked along the third dimension, forming a volumetric ECG representation:

$$V=\{h1,h2,\ldots,hK\}\in R(L1+L2+1)\times W\times K$$

Where W represents the width generated by reshaping operations, and K represents the number of heartbeat stacks. This volumetric modeling makes it possible to learn jointly about ECG characteristics in both space and time.

Dataset Creation:

After the extraction of the heartbeat ROIs, the procedure involves developing a dataset for training, validation, and testing in 3DCNN. Each heartbeat is processed to have a similar dimension corresponding to the required size for modeling in deep learning. A numerical description of heartbeat samples is resized to have a fixed dimension in accordance with 3DCNN models. Each sample is normalized by dividing it by its utmost amplitude to ensure similarity in dimensioning.

The training set is further expanded using the concept of data augmentation to avoid overfitting. This includes the addition of a small amount of Gaussian noise, signal amplitude scaling, time shifting, and slight stretching/contraction of the signal. These steps enable the network to become robust to different ECG patterns that are likely to be absent in the training set.

The information is presented in two categories:

Training/ validation dataset 80%

Testing set 20%

This division enables a systematic method for training and testing the model. A well-organized dataset enables greater accuracy and reliability of the outcomes acquired for the prediction of arrhythmia.

3D-CNN Training Experiments:

Dataset Processing:

After developing the data set, it will then be used for training, validation, and testing of a 3D-CNN model. The heartbeat volume data will then be trained using a supervised learning technique where every heartbeat will be labeled in accordance with the type of arrhythmia it possesses. This step aims to help the model learn all complicated patterns in heartbeats for proper classification.

Model Architecture and Training:

The training data, consisting of the augmented and normalized volumes of heartbeat sounds, is then input into the proposed 3D-CNN model. The model has many convolution, pooling, and dense layers. The convolution layers are responsible for feature extraction from a heartbeat sound, and then the pooling layer reduces the spatial size of these features. A softmax layer follows as the output layer for classifying a heartbeat.

The model uses Adam optimizer and categorical cross-entropy as a loss function. The model also creates graphs of accuracy and loss as it trains. The training result shows how well a model has learned. The model is then run on a test set for evaluation of its accuracy in practical situations.

Training Configuration and Implementation Details:

The proposed 3D-CNN model, implemented in a supervised learning framework, was trained in batches containing 32 samples. During training, the database was divided into three different sets: training, validation, and testing, with 80% utilized for training and validation, while the remaining portion was used for independent testing. The convergence of the model was guaranteed by setting the number of epochs to 50. The network was optimized using the Adam optimizer with an initial learning rate of 0.001, while the loss function of categorical cross-entropy was used for multi-class arrhythmia classification. Early stopping was followed based on the validation loss in order to prevent overfitting. The 3D convolutional layers had kernel sizes of (3 × 3 × 3) and a stride of (1 × 1 × 1), which greatly enabled the effective extraction of spatiotemporal features across the 3D heartbeat volumes. Dimensionality reduction was performed with the help of max-pooling layers having a pool size of (2 × 2 × 2). A softmax activation function was implemented in the final classification layer.

Evaluation Metrics:

The accuracy, precision, recall rate, and F1 score are some examples of metrics used for the assessment of the efficiency of the model. The metrics can be calculated using the value of true positives, false positives, true negatives, and false negatives that can be derived by using the confusion matrix. While accuracy helps to determine the overall correctness of the model, the precision and recall rate help determine the exact efficiency of the model regarding classes associated with certain types of arrhythmias.

Arrhythmia Detection:

After identifying the most accurate model, it is then tested using real-time data from heartbeats. The model will estimate the likelihood that each heartbeat is associated with a given sort of arrhythmia pattern. A heartbeat will be identified as arrhythmic if it exceeds a certain threshold for the possibility of an abnormal heartbeat pattern identification. This approach enables rapid arrhythmic pattern identification, which is potentially dangerous, and it is possible to implement it using real-time systems.

Experiment and Results:

Experimental Setup:

The experimental verification of the design 3D-CNN architecture would be performed within a high-performance computing setup. The hardware components of the setup would feature a high performance of GPU along with 16 GB System RAM, which would be sufficient for 3D heartbeat volumes computation. The software tools would be developed utilizing a Python coding framework that would enable support for TensorFlow and Keras deep learning platforms. The design architecture would be trained through an Adam optimization technique, having a learning rate of 0.001, while a Categorical Cross-entropy function could serve as an optimal loss function, suitable for a pattern recognition task that involves classification of different classes, like heart rhythm classes. Moreover, within this environment, graphs would also be produced to demonstrate the convergence of accuracy/loss of the design architecture.

CNN Validation:

The results of validation show that the new 3D CNN model has robust learning functionality. The new model has an accuracy of 98.52% on validation. Precision, recall, and F1 measure values are all over 98%, showing satisfactory arrhythmia classification with less error. The proximity of accuracy and recall values supports robust validation functionality. Consistency among validation results for different iterations of training further supports the good-generalization functionality of the new model.

CNN Testing Evaluation:

Based on that, performance evaluation of the proposed 3D CNN was performed on totally unseen ECG data to ensure an unbiased assessment. Testing metrics have been summarized in Table 1. Accuracy, precision, recall, and F1-score for the major arrhythmia categories have been listed. The proposed model has a high level of generalization capability with a testing accuracy of 98.74% in classifying ECG beats into proper classes outside the training dataset.

The confusion matrix constructed during testing showed that the model accurately identified most arrhythmia classes, especially having great performances in Normal (N) and Premature Ventricular Contraction (V) classes. The precision and recall rates for these classes were well over 98% because there was little chance of misclassification. The other classes, like

Fusion Beats (F) and Atrial Premature Beats (A), reported a slight decrease in precision because there were similarities among the waveforms, which again is consistent with the results reported by previous studies.

Table 2. Testing Metrics

Metric	Traning	Validation	Testing
Accuracy	99.02%	98.52%	**98.74%**
Precision	0.981	0.982	**0.980**
Recall	0.987	0.983	**0.982**
F1-Score	0.984	0.983	**0.981**

Table II. shows distributions of the performances across the test experiments: the all-over accuracy is constant within all the tests performed, with a median accuracy of 98.7% and recall values concerning the arrhythmia classes, indicating strong class-level sensitivity during testing. In this way, these results show that the performance of the testing by using 3D CNN is stable and reliable.

Class-wise Performance Analysis:

This subsection provides an analysis of class-wise performance of the proposed 3D-CNN approach on the MIT-BIH Arrhythmia Dataset. Precision, recall, and F1 measures are provided to verify the efficiency of the proposed method to handle class imbalance problems.

Table 3. Class-wise Performance Evaluation

Arrhythmia Class	Precision (%)	Recall (%)	F1-Score (%)
Normal (N)	98.9	99.1	99.0
Supraventricular (S)	97.8	97.4	97.6
Ventricular (V)	98.2	98.0	98.1
Fusion (F)	96.5	96.1	96.3
Unknown (Q)	95.9	95.4	95.6

Table III shows that the proposed model consistently high recall value along with an F1 score in the case of majority as well as minority classes of arrhythmia. To be precise, in the case of minority classes, such as supraventricular & fusion beats, it is performing well, proving the effectiveness of the proposed 3D CNN architecture in class imbalance problems of ECG arrhythmia classification.

CNN Visual Result:

In this work, the decision-making behaviour of the proposed 3D CNN model is analysed by using activation maps produced with a modified Grad-CAM technique. These maps highlight regions in an ECG waveform that are most influential for the classification task. For instance, for Normal beats, the model concentrates mostly on QRS complexes around the R-peaks. At an early stage, increased arrhythmic activity was observed with heightened attention to irregular P-wave and T-wave regions. In contrast, for other arrhythmic beats like PVCs, the activation maps emphasized broadened QRS morphology. The model clearly identified abnormal ventricular depolarization patterns. The same activation behaviour is repeated for several test samples. Consistent activation implies stable learning of ECG features. Constructive morphological patterns relevant clinically are effectively captured by the model. The overall visualization result improves the interpretability and diagnostic reliability of the proposed system.

Preprocessing Results:

The key driving metric was the computational performance, which can be captured through three top-level metrics: training time, model size, and inference speed. Table 2 summarizes the results of processing obtained during the experiments.

Table 4. CNN Processing Results

Metric	3D CNN
Training Time per Epoch	2.3 min
Total Parameters	3.4 million
Model Size (MB)	42.7 MB
Inference Time (ms)	18.4 ms

Table IV. shows the model's efficient processing capabilities in all trials. The average time for training of 2.3 minutes per epoch shows the feasibility of computation. The model size of 42.7 MB shows a compact model despite the capabilities and depth of the model. The inference time of less than 20 ms per ECG segment shows the efficiency of the model for real-time ECG analysis systems.

ECG Detection Results:

The suggested 3D CNN is emphasised as a promising method for recognizing and classifying arrhythmia signals from ECG data based on experimental results. Nine consecutive ECG signals are analyzed to demonstrate how emerging flaws cause normal P-QRS signals to change into arrhythmic signals. When the abnormal morphology continues for a series of heartbeats, the computer program recognizes ventricular abnormalities and declares the signal arrhythmic, triggering an alert to go out. In addition to this, the proposed model can recognize the return of the heart to the normal rhythm, indicating that the technique effectively works.

Recommendations and Practical Implications:

The dataset is divided into training, validation, and test datasets, where 80% of the dataset is used for model training and validation, and the remaining 20% for independent testing. The proposed 3D-CNN-based ECG arrhythmia detection framework has shown great potential in real-world practice in both clinical and everyday healthcare applications. By virtue of its high classification accuracy and strong recall performance, the model could be employed in clinical practice by offering support in making decisions while considering cardiologists' interpretations for early detection of arrhythmias by manual workload reduction.

This can also be practically implemented on continuous cardiac monitoring platforms, including systems for hospital-based ECG and wearable health gadgets. In this paper, the automated preprocessing and volumetric representation of heartbeats allow the proposed framework to process ECG recordings of long duration efficiently, which is quite essential for real-time and remote patient monitoring.

It is suggested that future work should continue the validation with multi-center and multi-dataset ECG repositories for generalizability across diverse patient populations. Further, optimization of computational complexity and inference latency can easily render the deployment on resource-constrained devices such as wearable sensors. Future enhancements may also involve noise-adaptive preprocessing in real-time and personalized model tuning for increased robustness in real-world applications. These suggestions provide testimony on the feasibility of translating the proposed research framework into practical cardiac health monitoring solutions.

Comparison and Discussion:

A comparison in Table V reveals that the presented 3D-CNN framework ensures more accurate results in arrhythmia classification compared to other existing approaches for lightweight and multi-class CNN architectures on the MIT-BIH Arrhythmia Database. In practical clinical applications, besides the measure of accuracy, other factors like recall and precision are also crucial because the issue of false negative results may cause difficulties in effective diagnosis and treatment of the problem in the form of arrhythmias.

Table 5. Comparison of Arrhythmia Detection Methods

Methodology	Targeted Features	Accuracy	Recall	F1-score	Precision
Lightweight_CNN Architecture [22]	Focused on minimized parameter set suitable for deployment on mobile devices with a focus on the edge	98.35%	98.10%	97.95%	97.80%
Multi-Class Deep CNN [23]	Employ standard 2D convolution layers for multi-level classification of arrhythmia patterns	97.90%	97.20%	97.30%	97.40%
Proposed 3D CNN Model	Passages the ability of Spatial and Temporal by translating 1D ECG signals into 3D heartbeat volumes	98.74%	98.65%	98.57%	98.50%

The baseline models were chosen to reflect the popular CNN-based, hybrid CNN & LSTM, and recent models incorporating the Attention mechanism in ECG classification methods reported in the literature. These models were chosen for their relevance in the field of arrhythmia detection, the existence of results on the MIT-BIH Arrhythmia Database, and their use in previous works as baseline models for comparison.

The performance improvements can be explained by the fact that the volume expressed by the heartbeat in 3D represents learning both morphological and ECG temporal information. By contrast, traditional methods involving the use of only 1D CNN learn mostly from the time information, and this spatiotemporal information is efficiently represented by the 3D-CNN architecture.

The Grad-CAM maps enable the provision of clinical insights about the decision made by the proposed model. The dense areas in the Grad-CAM maps relate to the most relevant aspects of ECG signals for diagnoses and include the QRS and, in some cases, the ST segment and T-wave portions in particular. Cardio logically speaking, an aberrant QRS morphology is known to be highly correlated with abnormalities in the ventricle's arrhythmic patterns, whereas an aberrant morphology in the segments related to repolarization is related to pathologies in such a segment. The correspondence between dense model areas and meaningful cardiac waveform portions indicates that the proposed model is based on physiologically meaningful aspects and is not sensitive to noise or unnecessary portions of a cardiac signal.

Despite these benefits, the model has been tested on a single dataset, making generalization over a broader population of patients difficult. Also, the use of multi-stage processing and possible class imbalance difficulties in identifying rare arrhythmic may pose robustness issues under practical scenarios. Future improvements would incorporate overcoming these issues through validation on more than one dataset. Even though the proposed method shows better results in terms of performance, the formal test for significance and estimation of confidence intervals was not performed. This part would be addressed in future work to validate the performance improvement. With regards to its usability in realistic applications, aspects like latency in inference, noise robustness, memory usage, and compatibility with wearable devices are to be accounted for. Although it has been shown that the proposed model has lower latency (<20 ms), in further research, attention will be focused on noise-adaptive preprocessing steps, energy-optimised model reduction, and testing on wearable ECG devices.

Conclusion and Future Direction:

This research was a success, having developed a 3D Convolutional Neural Network (3D CNN), aiding in improving automatic detection/classification of Cardiac Arrhythmias. The process of analyzing ECG signals as a volumetric object helps to effectively identify morphologic differences as well as the relationships between the heartbeats. The algorithm was found to be very accurate with an accuracy of 98.74%, which indicates better performance with no chances of overfitting. Its efficiency, which has an inference time of < 0.02 seconds, also helps it to be very useful in real-time applications. This study has managed to offer a useful decision-making aid system in the field of cardiology.

Future research will involve testing the validity of the proposed model with different publicly available ECG datasets, including INCART and PTB, to establish its generality over different population groups. Besides, there will also be research involving combining 3D CNNs with other recurrent models like LSTMs or Transformers to improve the model's understanding of longer-term intra-cycle relationships. Another area of research will involve optimizing the model into a lightweight model suitable for IoT devices to improve its applicability to health IoT devices. Finally, research will also involve incorporating advanced EXPLAINABLE AI (XAI) tools like SHAP Values and Relevance Propagation into the model to improve its interpretability by healthcare practitioners. There will also be research involving multi-modal learning involving ECG signals and other modal physiological signals like blood pressure and respiration rates.

References:

[1] T. B. S. & A. L. P. R. Antônio H. Ribeiro, Manoel Horta Ribeiro, Gabriela M. M. Paixão, Derick M. Oliveira, Paulo R. Gomes, Jéssica A. Canazart, Milton P. S. Ferreira, Carl R. Andersson, Peter W. Macfarlane, Wagner Meira Jr., "Automatic diagnosis of the 12-lead ECG using a deep neural network," *Nat. Commun.*, 2020, doi: https://doi.org/10.1038/s41467-020-15432-4.

[2] J. S. L. Donghyeon Kim, Kyoung Ryul Lee, Dong Seok Lim, Kwang Hyun Lee, "A novel hybrid CNN-transformer model for arrhythmia detection without R-peak identification using stockwell transform," *Sci. Reports Vol.*, vol. 15, 2025, [Online]. Available: https://www.nature.com/articles/s41598-025-92582-9

[3] T. M. Patrick Wagner, "Explaining deep learning for ECG analysis: Building blocks for auditing and knowledge discovery," *Comput. Biol. Med.*, 2024, [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/38749322/

[4] X. Y. Fahad Khan, "ECG classification using 1-D convolutional deep residual neural network," *PLoS One*, 2023, [Online]. Available: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0284791

[5] Aymane Edder, Fatima Ezzahraa Ben-Bouazza, "Self Attention-Driven ECG Denoising: A Transformer-Based Approach for Robust Cardiac Signal Enhancement," *Signals*, vol. 6, no. 2, p. 26, 2025, [Online]. Available: https://www.mdpi.com/2624-6120/6/2/26

[6] Yaqoob Ansari, Omar Mourad, "Deep learning for ECG Arrhythmia detection and classification: an overview of progress for period 2017–2023," *Front. Physiol.*, 2023, [Online]. Available: https://www.frontiersin.org/journals/physiology/articles/10.3389/fphys.2023.1246746/full

[7] M. L. Xin Zhang, "MSFT: A multi-scale feature-based transformer model for arrhythmia classification," *Biomed. Signal Process. Control*, vol. 100, p. 106968, 2025, [Online]. Available: https://www.sciencedirect.com/science/article/abs/pii/S1746809424010267

[8] Y. W. Zhuozheng Wang, "Multi-branch GAT-GRU-transformer for explainable EEG-based finger motor imagery classification," *Front. Hum. Neurosci.*, vol. 19, 2025, [Online]. Available: https://www.frontiersin.org/journals/human-neuroscience/articles/10.3389/fnhum.2025.1599960/full

[9] J. L. Myeonghun Lee, "ECG-GraphNet: Advanced arrhythmia classification based on graph convolutional networks," *Hear. Rhythm O2*, vol. 6, no. 8, pp. 1199–1211, 2025, [Online]. Available: https://www.sciencedirect.com/science/article/pii/S266650182500162X
[10] K. Tzevelekakis, Z. Stefanidi, and G. Margetis, "Real-Time Stress Level Feedback from Raw Ecg Signals for Personalised, Context-Aware Applications Using Lightweight Convolutional Neural Network Architectures," *Sensors 2021, Vol. 21, Page 7802*, vol. 21, no. 23, p. 7802, Nov. 2021, doi: 10.3390/S21237802.
[11] H. Z. T. Chen, Y. Sun, "Hybrid CNN–Transformer architecture for real-time ECG arrhythmia detection," *IEEE Sens. J.*, vol. 24, no. 3, pp. 4121–4132, 2024.
[12] S. P. R. Banerjee, A. Mitra, "Explainable AI-enabled ECG arrhythmia detection using Grad-CAM," *IEEE Access*, vol. 12, 2024.
[13] S. R. Yuwei Wang, "Hybrid CNN-LSTM model for ECG-based arrhythmia detection in internet of medical things," *Discov. Internet Things*, vol. 6, no. 7, 2026, [Online]. Available: https://link.springer.com/article/10.1007/s43926-025-00131-7
[14] C. L. Jianqiang Hu, "A novel multimodal self-supervised framework for ECG arrhythmia classification," *Comput Biol Med*, 2025, [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/41056617/
[15] B. C. Challagundla, "Advanced Neural Network Architecture for Enhanced Multi-Lead ECG Arrhythmia Detection through Optimized Feature Extraction," *arXiv:2404.15347*, 2024, [Online]. Available: https://arxiv.org/abs/2404.15347
[16] T. O. F. Conrad, Kuba Weimann, "Transfer learning for ECG classification," *Sci. Rep.*, 2021, [Online]. Available: https://www.nature.com/articles/s41598-021-84374-8
[17] H. M. H. Daniel Mauricio Jimenez Gutierrez, "Application of federated learning techniques for arrhythmia classification using 12-lead ECG signals," *arXiv:2208.10993v3*, 2024, [Online]. Available: https://arxiv.org/html/2208.10993v3
[18] B. S. Vaishnavi, P. Nayak, R. Janapati, V. Trivedi, J. S. Naga Jyothi, and P. Jugge, "Ensemble Deep Learning Framework for Robust Arrhythmia Detection," *2025 Third Int. Conf. Ind. 4.0 Technol.*, pp. 1–7, Sep. 2025, doi:10.1109/I4TECH64670.2025.11277843.
[19] M. A. M. Shadhon Chandra Mohonta, "Electrocardiogram based arrhythmia classification using wavelet transform with deep learning model," *Sens. Bio-Sensing Res.*, vol. 37, p. 100502, 2022, [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2214180422000319
[20] S. H. A. Hassan, M. Amin, "Hybrid CNN with explainable AI for ECG arrhythmia diagnosis," *IEEE Access*, vol. 12, pp. 21456–21468, 2024.
[21] A. O. M. Mohammed Guhdar, "Advanced deep learning framework for ECG arrhythmia classification using 1D-CNN with attention mechanism," *Knowledge-Based Syst.*, vol. 315, p. 113301, 2025, [Online]. Available: https://www.sciencedirect.com/science/article/abs/pii/S095070512500348X
[22] S. Kiranyaz, T. Ince, and M. Gabbouj, "Real-Time Patient-Specific ECG Classification by 1-D Convolutional Neural Networks," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 3, pp. 664–675, Mar. 2016, doi: 10.1109/TBME.2015.2468589.
[23] M. Dhilsath Fathima, S. Sithsabesan, K. Jayanthi, and S. U. Dhanyaa, "ACAC: Automatic Cardiac Arrhythmia Classification Based on 1D-Deep Resnet," *Proc. InC4 2024 - 2024 IEEE Int. Conf. Contemp. Comput. Commun.*, 2024, doi:10.1109/INC460750.2024.10649362.
[24] A. I. Sunnia Ikram, "Transformer-based ECG classification for early detection of cardiac arrhythmias," *Front. Med.*, vol. 12, p. 1600855, 2025, [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC12411431/