Combatting Illegal Logging with AI-powered IoT Devices for Forest Monitoring

Abdullah Khan, Hamza Ali, Maham Jadoon, Zain Ul Abideen, Nasru Minuallah
Computer Systems Engineering University of Engineering and Technology, Peshawar, Pakistan
***Correspondence**: 20pwcse1916@uetpeshawar.edu.pk, hamzaali.dcse@gmail.com, 20pwcse1875@uetpeshawar.edu.pk, zainikhan3434@gmail.com, n.minallah@uetpeshawar.edu.pk.

Citation | Khan. A, Ali. H, Jadoon. M, Abideen. Z, Minuallah. N, "Combatting Illegal Logging with AI Powered IoT Devices for Forest Monitoring", IJIST, Special Issue pp 134-142, May 2024

Received | May 05, 2024, **Revised |** May 12, 2024, **Accepted |** May 18, 2024, **Published |** May 22, 2024.

This research presents a comprehensive strategy for tackling illegal logging by leveraging Artificial Intelligence (AI) and Internet of Things (IoT) technologies. In high-risk forestry areas, sensors-equipped Internet of Things devices are used to continuously monitor and detect the sound of the surroundings. The AI component uses machine learning methods to identify potential unlawful logging activities by accurately detecting and distinguishing sound patterns associated with chainsaw and logging operations such as tree cutting and also detecting natural disasters like wildfires. When such activities are detected by these smart AI-powered IoT devices installed in the forest, real-time notifications are generated after such activity which allows surrounding enforcement agencies, such as the forest department, to intervene promptly. By providing a targeted and prompt solution to the issue of illicit logging, this strategy supports biodiversity preservation and sustainable forest management.

Keywords: Forest Monitoring; AI Against Illegal Logging; Real-time Alerts; Environmental Conservation; IoT for Anti-Logging.

Introduction:

These Forests are woven into our history. It is the source of fresh air, strong materials for houses, fresh water, and fertile soil for growing crops. But forests are being challenged. Illegal logging driven by uncontrolled greed for money challenges them. It throws nature's balance off give and take. The results are damaging to our ecosystems causing disasters like habitat losses as well as social entity indifferences. Deforestation initiates climate imbalance messing up the weather regime as well as disturbing the process of equilibrium in the ecosystem. All this careless act performed risks the same resources for its gain. Big-time cutting of the trees combined with illegal chopping of the wood provisions to the problems in saving worldwide vine variance, keeping equal ecology, and protecting the homes of countless wild beings. This destroys the system of world waters, reduces nature's distinction by the destruction of homes, and creates a lot of conflicts and other effects on social matters. Alarmingly, every year an expanse of country-sized land is afflicted by deforestation either from illegal logging or fires, especially in growing countries like Brazil and Indonesia. The records of vast forests have always been a challenge in the history books. Walking through dense woods is risky, requires a number of people, and usually leaves out far places. Regular surveillance like cameras can help, but they don't capture everything and can miss things as an action is taking place. The quiet noises in the forest, often covered by leaves and distance, can slip on by. This allows lasting damage to happen before we can step in. What if tech could boost these soft sounds, turning them into warnings of unfairness and signals of hope for our vanishing woods? Imagine hearing - really hearing – the hidden messages in the moving leaves. This is the big dream driving our project. We try to appeal to the cutting edge of the Internet of Things (IoT) and voice classification power [1]. We want to stand guard for peaceful green defenders. Now let's imagine intelligent sensors scattered in a forest, as shown in Figure 1. They're run by the chip ESP32 [2], a powerful, widely-used engine. Their job is to beckon and learn, with advanced tech that recognizes voices. They're always on the lookout for injurious things like noisy chainsaws, falling logs, or people talking up to no good. Before, these damaging noises would go unnoticed. Now they are translated into useful information. Straightaway, this information fires through the unseen paths of the IoT to a main cloud base. The soft sounds would rise into a loud noise in the woods, slowly but surely. Alerts go off immediately, leading forest rangers to exactly where the action was improper. Now, lawbreakers who prowl and come unseen are seen; their quiet deeds demand responsibility loudly and effect action, reverberating in the digital world.

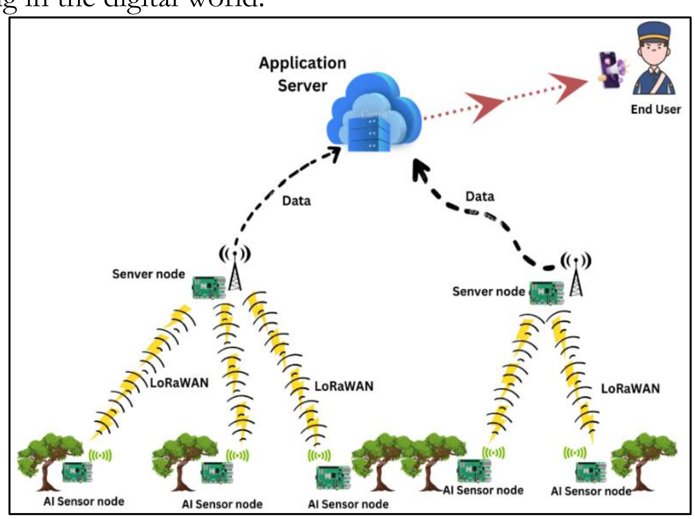

Figure 1: Use Case Diagram

Illegal logging is a pressing threat to global forests, imperiling their sustainability. A solution leveraging artificial intelligence (AI) and Internet of Things (IoT) devices [3] is proposed to combat this issue. The project aims to reduce unauthorized logging through real-time monitoring, and deploying IoT devices strategically. AI algorithms [4] will analyze data, identifying patterns and potential logging incidents. Preserving biodiversity is a key focus, preventing deforestation from illegal logging. The project enables data-driven decision-making for forest management, emphasizing collaboration among stakeholders. Capacity-building initiatives prioritize enhancing local communities' capabilities in monitoring and combating illegal logging. Awareness initiatives highlight the impacts of illegal logging, emphasizing forest conservation. Specific objectives include deploying IoT devices [3], implementing AI-powered analysis, and fostering collaboration. Advocacy for supportive policies is integral to combating illegal logging. Community engagement involves training local communities in monitoring and protecting forests. Ongoing evaluation ensures the effectiveness of AI-powered IoT devices in combating illegal logging.

Literature Review:

In the world of audio classification, which has made great strides since the last decades of the 20th century, applications have been found for distance learning and digitized libraries [5]. The new generation methodologies that have been pioneered include the Hidden Markov Model by Lu Jian with its intricate sorting of audio files and also label-free approach by Zhao Xueyan. Li, S.Z provides a refined sound identification system using Support Vector Machines (SVMs) and Mel Frequency Cepstral Coefficients (MFCC) in the domain of sound understanding [6]. Features from the time and frequency domain form the focus area of the study for sorting the sounds. Using Short-Time Average ZCR and Short-Time Energy in the time domain, the important features for sound or silence determination are extracted on the paper. In the frequency domain, there exist repeated features like the Centroid of the Audio Frequency Spectrum and Sub-Band Energy Ratio that give useful pointers. The idea of the suggested method of sound sorting is grounded in SVM, and thus it requires the careful preprocessing of audio samples and many features to be extracted, including MFCC. The system proved effectual as it was tested for a mixed set of 2500 sound pieces and obtained an average commendable accuracy of 92.14%. In conclusion, this article reveals some insight into the historical context of audio sorting, a deep sound sorting system mixing time and frequency traits, augmented with SVM processing. Tiny Machine Learning (Tiny ML) challenges the problem of pitching machine learning into microcontroller integration in domains such as sound recognition [5], image classification, and motion classification/anomaly detection [1]. In sound recognition, the most powerful keyword spotting (KWS) takes center stage whereby individuals gain hands-free access control for better user accessibility [6]. In projects that depict exemplary Tiny ML applications, such as the project with XIAO nRF52840 Sense, implementation of Tiny ML applications as feasible is demonstrated even in cases of limited resources. Regarding other projects, Tiny ML has an application in image classification where XIAO ESP32S3 Sense additionally supports camera access and therefore can be used in projects related to object recognition too. Tiny ML enhanced by an Ensenso Pico for motion classification/anomaly detection enabled the real-time analysis of sensor data and has been exemplified in an innovative cargo damage prevention tool utilizing XIAO nRF52840 Sense. These, as well as those purpose-developed in this area, illustrate the transforming potential of Tiny ML hence empowering low-power embedded systems across various sectors.

We provide audio classification in this literature review [1], from understanding basic processes to the application of advanced machine learning models. It focuses on voice conversion into quality digital format, stress on analysis through spectrogram and other tool selections like Mel spectrogram and Short-Time Fourier Transform (STFT). Within the study, the loudness differences are normalized using techniques, and ideas of Data Augmentation and also the use of mechanisms such as Depth Separable Convolutional Neural Networks (DS-CNN) and Fast GRNN are presented to improve the overall capabilities of listening to audio tasks.

This investigation into Enhanced Sound Recognition (ESR) with the application of Deep Learning (DL) techniques underscores the emergent interest in leveraging DL in enhancing Machine Learning (ML) models [5]. This study will review the different approaches toward sound recognition from spectrograms/MFCCs as inputs to ANN/CNN classifiers and raw waveform directly. In one of the most illustrative experiments, sound classifying in a Wireless Sensor Network (WSN) with Raspberry Pi (RPi) nodes is performed using a Convolutional Neural Network (CNN). The research has proven that even feature extraction and sound classification can be executed on embedded high-level devices, so it highly accentuates searching computational capabilities [6]. Addressing the dearth in discourse of the benefits and the feasibility of the introduction to machine learning techniques of sound classification [1]. Other proposed solutions for the Internet of Things (IoT) are specially designed models like CNN [3], with a separate focus on model inference time and resource efficiency on low-power microcontrollers. Hardware accelerators such as Tensor Processing Units (TPUs) and Field Programmable Gate Arrays (FPGAs) [7] are exploited to enhance the performance of audio applications of Deep Learning. Lastly, the study concludes with a better approach for looking at damage identification and location in multiple classes by considering the various levels of severity of damages as well as scenarios. These techniques along with the windowing data augmentation and a unique majority voting technique accompanied by the global CNN-1D model prove to be adequate for dealing with the problems of limited data.

Methodology:

The methodology consists of a series of related steps from the data collection to the real-world deployment designed to develop an effective system to fight against the issue of illegal logging with the help of AI-enabled IoT devices [8]. In our proposed methodology cutting-edge technology is used by utilizing the low-power microcontroller and making it a powerful tool with the use of Machine learning. Like in the first step through the data collection and preprocessing before training the supposed model.

Data Collection and Preprocessing:

First, a rich dataset comprising 75 samples of each of the following audio classes - Silence, Axe, Chainsaw, and Fire has been collected from different websites. Each audio recorded voice has 5 seconds and the total data collected is shown in Table [1]. The process involved in data collection uses Mel-filter bank energy features that are very important in audio classification. These features provide vital information about environmental sounds and extract speech signals for use in recognition tasks like those associated with illegal logging.

Table 1: Data Collection

Data Collected	Train / Test Split
25m 0s	79% / 21%

Training Deep Learning Model:

In this step, the convolutional neural network model (CNN) is trained and selected filter bank energy features (MEF) as a feature to detect the surrounding voices and make meaningful insight from recorded voices. In Step 1, obtain a balanced dataset where each class would contain at least 75 samples: Silence, Axe, Chainsaw. The audio samples consisted of files with durations of exactly 5 seconds each. This dataset is important in training a strong classifier for the accurate classification of environmental sounds associated with the activities of illegal logging.

Mel-filter bank Energy Features (MEF):

The Mel-filter bank energy features (MEF) have been used as the choice feature for this audio classification. The method of extraction of the features used in this work is to capture the frequency content of the audio signals using the Edge Impulse Platform [9]. This process can be achieved by using a filter bank to divide this audio signal into several narrow frequency bands. The resulting energy values across these bands serve as essential input features for the machine learning model. Parameters for MEF are listed below

- **Frame Length:** 0.02 seconds
- **Frame Stride:** 0.01 seconds
- **Filter Number:** 40
- **FFT Length:** 256
- **Low Frequency:** 0

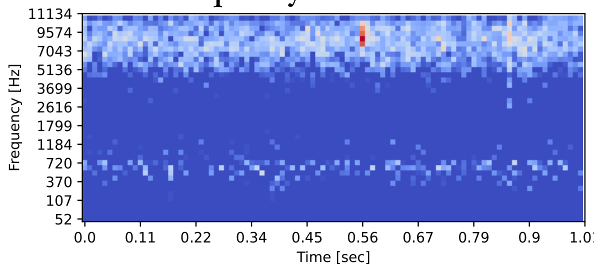

Figure 2: MEL Energies (DSP Output)

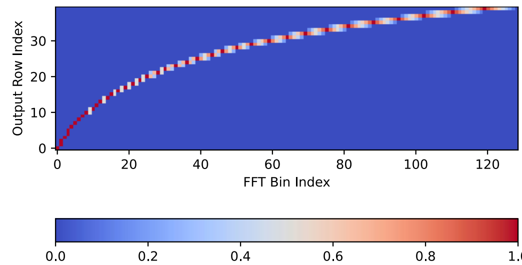

Figure 3: FFT Bin Weightings

Figures 2 and 3 illustrate the Mel Eng Mel Energies and FFT Bin Weighting of some of the voices from a data set. These parameters should be chosen very carefully in such a way as to optimize the extraction of Mel filter bank [10] energy features from the audio data, such that ultimately a set of very representative and informative features is obtained for the purpose of training the deep learning models.

Neural Network Architecture:

The neural network architecture was designed to effectively process the Mel- filter bank energy features classifying the audio into their corresponding classes. The architecture includes:

- **Input Layer (3,960 features):** The model reshapes the features so that it can be processed by the model.
- **Reshape Layer (40 columns):** 40 filters that are part of the MEF.
- **1D Conv/Pool Layer (8 neurons, 3 kernel size, 1 layer):** First incidence of convoluting.
- **Dropout (Rate 0.25):** Introduces regularization to prevent overfitting.
- **Conv/Pool Layer (16 neurons, 3 kernel size, 1 layer):** Extra convolutions for better feature extraction.
- **Dropout (Rate 0.25):** Further regularization.
- **Flatten Layer:** It converts the output into a one-dimensional array/tensor for the next series of layers.
- **Additional Layer**: Introduces complexity and abstraction for improved model performance.
- **Output Layer (4 classes):** Represents the four classes (Silence, Axe, Chainsaw, Fire).

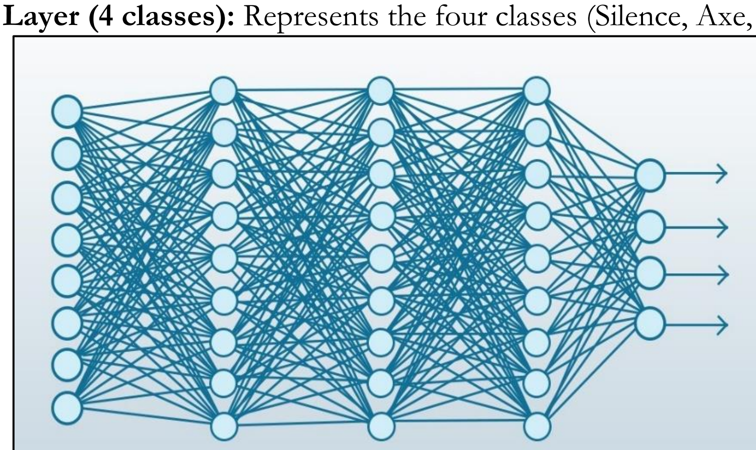

Figure 4: Model with Hidden Layer

Table 2: Evaluation Metrics

Accuracy	Loss
96.2%	0.10

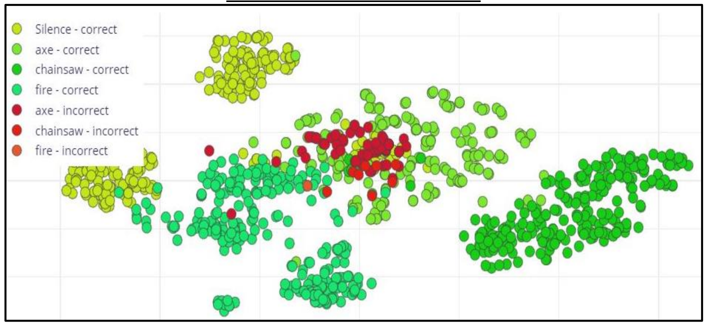

Figure 5: Data Explorer

Training Settings:

- **Number of Training Cycles:** 100
- **Learning Rate:** 0.005

In processing that, it is possible to realize the consensus and accuracy during the training process to see that the network learns effectively the Mel-filterbank energy features. Figure 5 below shows the model training accuracy which also shows the incorrect classification with red points.

Integration Model with ESP32 Microcontroller as a Sensor Node:

Using the trained model that we have developed in the previous section, the model is exported as an Arduino and deployed to the ESP32 [2] (ESP32 is a series of low-cost, low-power systems on a chip microcontroller with integrated Wi-Fi and dual-mode Bluetooth) microcontroller functioning as a sensor node. For the recording of audio data of surroundings, an enabled INMP441 microphone has been used The INMP441 is a high-performance, low-power, digital-output, omnidirectional MEMS microphone. The recorded audio data, therefore undergoes processing by use of a device machine learning model. This integration ensures a classification of sounds that falls into categorization related to the activities of illegal logging by ESP32 autonomously.

Table 3: On Device Performance

Interfacing Time	Peak RAM Usage	Flash Usage
21ms	10.6Kb	32.6Kb

In Table 3, on-device performance is illustrated. On ESP32 [2] SoC. The latter uses LoRa communication for a secure transmission of the voice identified, and node location (longitude, latitude). In this scenario, the ESP32 being a node sensor, sends its data, using LoRa technology to the central server node which is another ESP32, and works as a server as it is handling the database of that project. This way of communication is needed to establish networking abilities over wide forested areas so that monitoring can occur in real-time.

Server Node (ESP32) and Database Integration

In this step, ESP32 works as the server node and is to receive signals from the ESP32 sensor nodes. The server node connects Firebase which is a cloud platform by Google. Firebase is used for real-time data storage, updating node locations, and other relevant information that exists in the field of operations. This integration ensures a centralized and accessible repository of data. The mobile application is built to interface with the Firebase database and give real-time notifications on detected activities and node locations. It is the user interface for the app, where different users like members of the local community, law enforcement agencies, as well as environmental groups interact. Figure 6 illustrates the project top overview and clear flow of our project. This is the real-world-based scenario that is used to test how good the model is and has trained on several localization and classification tasks in a real-world case. The testing phase also tests whether LoRa communication is robust as well as ensuring data moves from LoRa to Firebase successfully.

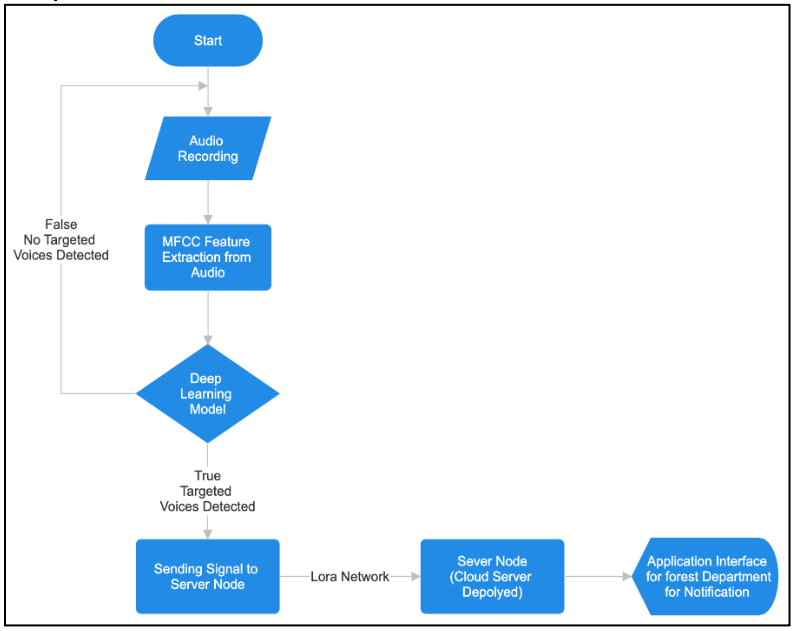

Figure 6: Workflow diagram

Results:

The model performed exceptionally well in various aspects. It achieved an accuracy of 96.2% on the validation set along with low loss, that is, 0.10. The confusion matrix reflected the accurate different class abilities of the model, with high percentages for SILENCE, AXE, CHAINSAW, and FIRE. The corresponding efficiency in terms of F1 scores was further validated through the table below. During the on-device performance, the model showed efficient inferencing that took very short of 21 MS. Minimal resource usage was identified by peak RAM at 10.6K and flash usage at 32.6K as already shown in Table 3. This makes for deployment in real-time constrained devices.

Table 4: 82% Accuracy

	Silence	Axe	Chainsaw	Fire	Uncertainty
Silence	73.5%	9.8%	0%	0%	16.7%
Axe	1.8%	70.2%	0%	4.0%	24.1%
Chainsaw	0%	2.9%	95.2%	0%	1.9%
Fire	0%	1.2%	0%	98.4%	0.4%
F1 score	0.84	0.77	0.98	0.96	-

Table 5: Model Efficiency in Utilization

	MFE	CLASSIFIER	TOTAL
LATENCY	450MS	21MS	471MS
RAM	20.4K	10.6K	20.4K
FLASH	-	32.6K	-
ACCURACY	-	-	-

In each of the testing scenarios, the model maintained a high level of accuracy at 82.57%. The performance behavior of the classification algorithm in classifying into groups the audio samples under the different classes including instances of uncertainty was depicted by the confusion matrix. The F1 scores for each class also confirmed the reliability of the model under different setup conditions during testing. The latencies distribution during the inference process of the model displays as shown in Table 5 a well-distributed figure, with an overall of 471 MS made up of 450 MS that are contributed by the Mel-filter bank energy features (MFE), another 21 MS for the classifier. Resource usage in terms of RAM and FLASH stood at 20.4K and 32.6K respectively unveiling further efficacy of the model in memory allocation as well as storage. Overall, the results underscore the strong generalization ability of the model showing it as being promising in a real-world setting combating illegal logging with its effective audio classification on IoT devices.

Conclusion:

The application of the Internet of Things (IoT) and Artificial Intelligence (AI) technology [4] offers an achievable approach to combating illegal logging, which is a major danger to world forests and biodiversity. We have demonstrated an innovative strategy for real-time monitoring and detecting actions related to illegal logging, such as chainsaw running and tree cutting down, through the deployment of smart sensors installed with AI algorithms [4]. By harnessing advanced machine learning techniques which include convolutional neural networks (CNNs) trained on Mel-filter bank energy features (MEF), the developed model attained a spectacular accuracy of 96.2% on the validation set. The model efficient inferencing process with minimal resource usage on low-powered microcontrollers like the ESP32 which spectacle its suitability for real-world deployment in forest environments. Integration with IoT devices like ESP32 [2] microcontroller serving as a sensor node enables autonomous classification of environmental noises related to illegal logging activities. Employing LoRa communication technology, data transmission to a centralized server node deployed with the Firebase database ensures real-time data storage and accessibility. The model's overall accuracy of 82.57% as well as successful memory allocation and memory resource usage have been confirmed by the testing scenarios' results, which also highlight the model's robustness and generalization capacity. The model is a useful tool for monitoring and preventing illicit logging in real-world forest situations because of its low-latency inference process and ability to reliably classify audio samples into distinct classifications. In conclusion, our research demonstrates the effectiveness of AI-powered IoT devices in preserving biodiversity promoting sustainable forest management, and also safeguarding against the dangerous effects of illegal logging. By providing timely detection and intervention capabilities this strategy contributes to the conservation of forests, ecosystems, and a lot of benefits they provide to humanity and the planet as a whole.

References:

[1] C. Anjanappa, S. Parameshwara, M. K. Vishwanath, M. Shrimali, and C. Ashwini, "AI and IoT based Garbage classification for the smart city using ESP32 cam," Int. J. Health Sci. (Qassim)., vol. 6, no. S3, pp. 4575–4585, May 2022, doi: 10.53730/IJHS.V6NS3.6905.

[2] "ESP32 Wi-Fi & Bluetooth SoC | Espressif Systems." Accessed: May 06, 2024. [Online]. Available: https://www.espressif.com/en/products/socs/esp32

[3] S. B. Calo, M. Touna, D. C. Verma, and A. Cullen, "Edge computing architecture for applying AI to IoT," Proc. - 2017 IEEE Int. Conf. Big Data, Big Data 2017, vol. 2018-January, pp. 3012–3016, Jul. 2017, doi: 10.1109/BIGDATA.2017.8258272.

[4] M. Ali, Y. S. Kwon, C.-H. Lee, J. Kim, and Y. Kim, Eds., "Current Approaches in Applied Artificial Intelligence," vol. 9101, 2015, doi: 10.1007/978-3-319-19066-2.

[5] M. Merenda, C. Porcaro, and D. Iero, "Edge Machine Learning for AI-Enabled IoT Devices: A Review," Sensors 2020, Vol. 20, Page 2533, vol. 20, no. 9, p. 2533, Apr. 2020, doi: 10.3390/S20092533.

[6] X. Yang, H. Xing, and X. Su, "AI-based sound source localization system with higher accuracy," Futur. Gener. Comput. Syst., vol. 141, pp. 1–15, Apr. 2023, doi: 10.1016/J.FUTURE.2022.10.023.

[7] D. Roggen, R. Cobden, A. Pouryazdan, and M. Zeeshan, "Wearable FPGA Platform for Accelerated DSP and AI Applications," 2022 IEEE Int. Conf. Pervasive Comput.

Commun. Work. other Affil. Events, PerCom Work. 2022, pp. 66–69, 2022, doi: 10.1109/PERCOMWORKSHOPS53856.2022.9767398.

[8] I. Ahmad, S. E. Shariffudin, A. F. Ramli, S. M. M. Maharum, Z. Mansor, and K. A. Kadir, "Intelligent Plant Monitoring System Via IoT and Fuzzy System," 2021 IEEE 7th Int. Conf. Smart Instrumentation, Meas. Appl. ICSIMA 2021, pp. 123–127, Aug. 2021, doi: 10.1109/ICSIMA50015.2021.9526312.

[9] "Edge Impulse - The Leading edge AI platform." Accessed: May 06, 2024. [Online]. Available: https://edgeimpulse.com/

[10] S. R. Madikeri and H. A. Murthy, "Mel filter bank energy-based slope feature and its application to speaker recognition," 2011 Natl. Conf. Commun. NCC 2011, 2011, doi: 10.1109/NCC.2011.5734713.