Sensors and Materials, Vol. 37, No. 11 (2025) 5043–5053
MYU Tokyo

S & M 4231

Cyber-physical System-based Wide-area IoT for Illegal Forest Logging Monitoring and Alert System

Shu-Han Liao,[1*] Yan-Jia Huang,[2] Shiuan-Jen Yang,[3] Yu-Wei Li,[1]
Pei-Chi Chung,[1] Hsin-Wen Wei,[1] and Cheng-Fu Yang[4,5**]

[1]Department of Electrical and Computer Engineering, Tamkang University, Tamsui 25137, Taiwan
[2]Department of Geography, National Taiwan University, Taipei, Taiwan
[3]Department of Electrical Engineering, National Taiwan University, Taipei, Taiwan
[4]Department of Chemical and Materials Engineering, National University of Kaohsiung, Kaohsiung 811, Taiwan
[5]Department of Aeronautical Engineering, Chaoyang University of Technology, Taichung 413, Taiwan

(Received April 19, 2025; accepted November 18, 2025)

Keywords: environmental sound classification, LoRa, forest monitoring, illegal logging detection, poaching detection, machine learning, IoT, Mel spectrogram, sustainable forest management

Illegal logging and poaching in forests threaten the forest ecosystem and cause biodiversity loss and environmental degradation. However, traditional monitoring methods lack scalability and real-time responsiveness. To overcome these limitations, we developed a long-range (LoRa)-based environmental sound classification alert system for real-time forest surveillance and alerting. With low-power hardware components such as Raspberry Pi, Arduino, and Dragino LoRa modules, the developed alert system captured and transmitted audio data. It also converted and compressed sound signals into Mel spectrograms using dimensionality reduction for the low-bandwidth LoRa network. A machine learning model then classified the sounds into "dangerous" (e.g., chainsaws and gunshots) and "safe" categories. The system showed an accuracy higher than 85% in detecting threats and presented reliable performance under noisy field conditions in the experiment. The alert system can operate efficiently in large forested areas with minimal maintenance. While challenges remain such as limited transmission capacity and occasional misclassifications, it is necessary to explore convolutional neural network-based models, edge AI integration, and deployment at scale. The developed forest alert system enables scalable and cost-effective IoT-AI solutions for forest protection and supports data-driven sustainable ecosystem management.

1. Introduction

Illegal poaching and logging in the forest endanger the global forest ecosystem and damage biodiversity, carbon sequestration, and the livelihoods of local communities irreversibly. Forests play a crucial role in maintaining ecological balance, yet they are increasingly vulnerable to exploitation. Traditional monitoring methods of illegal poaching and logging require manual patrolling and stationary surveillance cameras, which are inefficient in addressing large-scale

*Corresponding author: e-mail: shliao@gms.tku.edu.tw
*Corresponding author: e-mail: cfyang@nuk.edu.tw
https://doi.org/10.18494/SAM5705

ISSN 0914-4935 © MYU K.K.
https://myukk.org/

and geographically dispersed threats. These conventional methods also demand high operational costs with scalability issues and the delayed detection of incidents in real time.

Researchers have actively developed sound-based detection systems to monitor illegal logging. Mporas *et al.* proposed an acoustic surveillance system to detect chainsaw sounds in the natural environment using support vector machines (SVMs).(1) Girbau-Llistuella *et al.* developed a real-time detection model based on environmental sound classification (ESC), utilizing low-power IoT devices for long-term deployment in forests.(2,3) The previously developed systems represent the growing interest and effectiveness of sound recognition technology in forest surveillance. Figure 1 provides an overview of the proposed cyber-physical system from the microphone node and feature extraction, through the LoRa uplink, to server-side inference and the monitoring graphical user interface (GUI). Advancements in IoT and AI have strengthened the potential of such systems.(4) Among the related technologies, ESC has been widely used as its cost-effective and scalable approach for identifying illegal activities. In using ESC, deep learning models such as convolutional neural networks (CNNs) and transformer-based models are adopted for sound classification under complex conditions.(5,6) Such models analyze audio data collected from remote forests and distinguish sounds from chainsaw operations, gunshots, and background noises including animal calls or wind.(7)

Despite its huge potential, employing the sound-recognition-based monitoring system in the forest presents technical challenges. First, a robust and accurate sound recognition model requires access to well-labeled datasets specific to forest soundscapes. Bandara *et al.* introduced the forest sound classification (FSC22) dataset,(8) a benchmark dataset containing more than 2000 annotated forest sound clips, which has been widely adopted in ESC research. Second, the real-time transmission of sound features from remote areas is difficult owing to the sound's low bandwidth constraints and intermittent connectivity. Third, for long-term deployment, energy-efficient design and low-power hardware are required.(9) To tackle these issues, researchers have employed long-range (LoRa) communication technology for audio data transmission,(9,10) which offers long-range, low-power wireless transmission. Therefore, LoRa is regarded as ideal for applications in forests and rural areas with limited infrastructure. Ferrari *et al.* explored the feasibility of LoRa in a smart waste system,(11) whereas Wang *et al.* implemented LoRa for forest

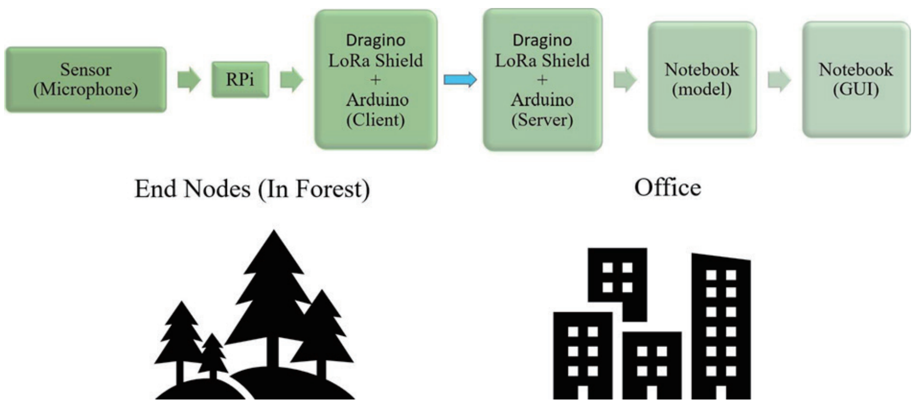

Fig. 1. (Color online) System architecture.

weather monitoring. These studies validated LoRa's energy efficiency and robustness in the natural environment.(9)

In this study, we developed a LoRa-based ESC alert system for real-time forest monitoring and illegal activity detection. We integrated machine learning (ML) algorithms, low-power communication protocols, and edge computing capabilities in the alert system to provide a scalable and cost-effective solution for monitoring illegal poaching and logging. The alert system provides real-time alerts, enhances detection accuracy, and supports large-scale deployment across diverse forest ecosystems. The developed alert system reliably distinguishes dangerous sounds from natural sounds using advanced sound recognition models.(5,6) To ensure robust signal transmission in the forest environment with limited connectivity,(2,11) we optimized the hardware and software of the alert system to minimize power consumption and enable long-term deployment.(2) The alert system functions effectively in noisy and unpredictable outdoor environments and offers a scalable, cost-effective, and real-time forest monitoring solution.(7) It also shows a high illegal activity detection accuracy and presents feasibility for real-world deployment. By refining the alert system, communication protocols can be further optimized to enhance its performance in diverse forest environments.

2. Literature Review

In previous research, acoustic sensing technologies were mainly used for environmental monitoring, particularly in detecting anthropogenic disturbances such as illegal logging and poaching. Early systems relied on classical ML models such as SVMs for sound detection.(1) However, the models were only effective under controlled conditions and often struggled to maintain accuracy in highly variable and noisy outdoor environments. To overcome these limitations, deep learning models have been adopted as they are capable of learning complex temporal and spectral patterns from raw or preprocessed audio signals. CNNs and transformer-based models accurately classify environmental sounds even with prominent performance on large-scale audio datasets such as AudioSet and UrbanSound8K.(5,6) Forest-specific training data to train the models in an actual forest environment are lacking. To solve this problem, Bandara *et al.* introduced the FSC22 dataset, which contains a curated and labeled collection of forest-related sound clips for the development and evaluation of the models tailored to forest soundscapes.(8) With model development, the deployment infrastructure for in situ audio monitoring has been implemented widely.

However, previous sound classification systems have severe energy constraints and limited connectivity in the forest environment. Therefore, LoRa technology has been introduced owing to its low power consumption and long-range data transmission capacity. Ferrari *et al.*(11) and Wang *et al.*(9) proved LoRa technology's potential for remote forest weather monitoring and highlighted its applicability in off-grid deployments. Despite such research results, challenges remain. First, the previous systems cannot present low-latency signal processing and reliable data transmission over unstable wireless channels, which are essential in real-time detection. Second, the model's complexity and low computational efficiency with limited processing power do not ensure accurate sound classification outdoors. Third, environmental noises, such as wind, rain, and overlapping animal calls, introduce significant ambiguity in audio classification, demanding robust preprocessing and noise-resilient models. In summary, previous research underscored the potential of combining ESC with IoT technologies to enhance forest protection. However, systems must integrate optimized sound recognition models with energy-efficient transmission protocols such as LoRa in a modular, field-ready architecture. Therefore, this study aims to design a low-power, real-time ESC system customized for forest monitoring to address such requirements and offer a feasible and scalable implementation in the forest.

3. Materials and Methods

3.1 Sound recognition mode

3.1.1 FSC22 dataset

The FSC22 dataset is a publicly available database specifically developed for ESC in the forest.[8] It was established to address the scarcity of open datasets focusing on forest-related acoustic events. The FSC22 dataset contains 2,025 labeled audio clips, each with a fixed duration of 5 s. The dataset is available on www.FreeSound.org. To capture the acoustic diversity in the natural forest, the FSC22 dataset is categorized into mechanical, animal, environmental, vehicle, forest threat, and human sounds. Each category is further subdivided into 34 subcategories, encompassing typical forest sounds such as chainsaws, bird calls, wind, gunshots, and human voices. In its initial version, the dataset included approximately 75 samples for each subcategory. The structure and granularity of the FSC22 dataset are appropriate for training ML models in complex and noisy forests.[7]

3.1.2 YAMNet

YAMNet is a lightweight, pretrained deep-learning model for general-purpose sound event classification. It is built on the MobileNet V1 architecture and trained on the extensive AudioSet-YouTube corpus, which includes more than 2 million annotated audio segments covering more than 500 event classes.[12] The model accepts log-mel spectrogram inputs and outputs a probability distribution across 521 predefined sound classes. As a pretrained network, YAMNet provides a strong feature representation for various acoustic events, making it highly transferable to domain-specific tasks such as forest sound classification.[4] The model is implemented in Keras and TensorFlow, enabling easy integration with custom training pipelines. YAMNet performs well as a backbone for ESC in low-resource environments owing to its compact architecture and efficient inference speed.[5] To use YAMNet for the specific needs of ESC in the forest, we fine-tuned the model using the FSC22 dataset. The process began by randomly extracting 1 s segments from the original 5 s audio files.

These clips were transformed into Mel spectrograms, a widely used time-frequency representation for sound event analysis. To reduce data dimensionality and support low-bandwidth transmission, frequency averaging was applied to compress the spectrograms into manageable feature vectors. The dataset was partitioned into the training, validating, and testing datasets in a ratio of 80:10:10. The Adam optimizer was used to train the model at a learning rate of 0.001 and a batch size of 64 for 15 epochs. The loss function was set to categorical cross-entropy. This configuration offered a balance between training stability and generalization. Compared with training a CNN model, the transfer learning approach used in this study significantly reduced training time while achieving superior classification performance.(5,7) The final layer of the model was modified to output probabilities across 27 softmax-activated nodes, corresponding to the 27 sound subclasses defined in the FSC22 dataset.

3.2 System development

3.2.1 LoRa

The LoRa technology is used in a proprietary low-power wide-area network.(10) Designed for long-range and low-data-rate applications, LoRa is well suited for IoT and machine-to-machine (M2M) communication in remote areas. It operates in unlicensed sub-gigahertz frequency bands (e.g., 868 or 915 MHz) and utilizes chirp spread spectrum modulation, enabling robust communication over distances up to 15 km with minimal energy usage. Compared with Wi-Fi, Zigbee, or Bluetooth, LoRa ensures a wider range and superior energy efficiency with lower data throughput. These properties are ideal for forest-based deployments where stable infrastructure and power sources are limited.(11) Implementations in smart waste monitoring and forest-based weather stations have validated LoRa's effectiveness in outdoor and low-maintenance environments.(2,11) Commercial solutions such as Tektelic sensors further demonstrate LoRa's appropriateness for multiyear, low-power operation.(2) In this study, LoRa is used for the long-distance transmission of compressed audio features from devices in remote areas to a central monitoring unit. Its low power consumption and wide coverage solve previous problems in forest monitoring, making LoRa an integral part of the developed real-time ESC alert system in this study.

3.2.2 System architecture

The system architecture demonstrates a low-power, long-range audio monitoring and alert solution tailored for the forest environment. Environmental audio is captured using a high-sensitivity microphone and locally processed by Raspberry Pi (RPi), which performs feature extraction and signal compression. To ensure continuous operation in remote areas, the alert system is powered by a small solar panel and rechargeable lithium battery module, enabling autonomous energy supply without frequent maintenance. The average power consumption of the sensing node, including the RPi, LoRa transceiver, and microphone, is approximately 1.5–2.0 W during active operation and less than 0.3 W in standby mode. This energy-efficient design allows the system to maintain long-term monitoring and communication while minimizing power demand in off-grid forest environments. To address environmental challenges such as storms and heavy rainfall, the sensing units are enclosed in weather-resistant housings with protective mesh openings designed to reduce direct wind impact on the microphone. Nevertheless, the risk raised regarding reduced sound collection performance under extreme weather is valid.

The current sound database used for model training includes a wide range of background noise conditions but does not systematically model storm-level wind, heavy rain, or enclosure-induced attenuation. Consequently, actual field deployments may experience reduced signal-to-noise ratio and minor shifts in frequency response. Future work will expand the environmental sound database to include severe weather conditions and apply adaptive noise filtering techniques to further improve robustness in adverse environments. The processed data is then transmitted via an Arduino board equipped with a Dragino LoRa shield. This client module, deployed in remote forest areas, communicates with a server module located in a base station or an office environment using LoRa wireless technology. The server receives the transmitted audio features and classifies them using a pretrained ML model. The results are then visualized on GUI, allowing users to monitor potential threats such as illegal logging. This modular and distributed design enables scalable, real-time forest surveillance with minimal energy consumption and infrastructure requirements.

The developed system comprises the following hardware components to achieve low-power, long-range audio monitoring in the forest. Arduino UNOs are deployed on the field (client side) and at the base station (server side) for serial communication and interfacing with LoRa modules. Dragino LoRa shields enable long-distance wireless communication by transmitting and receiving feature vectors over the LoRa network, being mounted on the Arduino boards [Fig. 2(a)]. RPi 4B acts as the local processing unit on the client side and captures raw audio, performs signal preprocessing (e.g., Mel spectrogram generation), and communicates with the Arduino via USB. A high-sensitivity external microphone (Yeti USB) is used to capture environmental

(a) (b)

Fig. 2. (Color online) (a) Dragino LoRa shield with antenna and (b) microphone (Logitech G Yeti).

sounds with clarity, particularly important for recognizing faint or distant sounds such as chainsaws and gunshots [Fig. 2(b)]. USB cables are employed for data transfer between RPi and Arduino and also serve as power supply connections. This hardware configuration has been designed to ensure modularity, cost-efficiency, and compatibility with field deployment in remote areas lacking conventional infrastructure.

3.2.3. End node transmission

RPi records approximately 6 s of audio every 30 s using an external microphone. The audio is trimmed to 5 s by removing the excess at the end, and a 1 s segment is randomly extracted from the middle. RPi processes the audio by converting it into a Mel spectrogram through a logarithmic transformation to reduce dimensionality by averaging values across different frequency bands. The resulting values are negative floating-point numbers (e.g., –xx.xx). To meet the 255-character limit of LoRa transmissions, these values are scaled by −10 and converted to integers, resulting in three-digit positive integers. At each field node, RPi captures approximately 6 s of audio every 30 s using the external USB microphone. By adopting this intermittent recording, power usage is reduced while maintaining the temporal coverage of acoustic events. After trimming excess time from the ends, a random 1 s segment is extracted from the middle to serve as the analysis window. This preprocessing is conducted to preserve relevant spectral features while minimizing bandwidth and energy consumption.

3.2.4 Receiver side

Once transmitted via LoRa, the data are received by the server-side Arduino and forwarded to a central computer (e.g., a laptop computer) through serial communication. The system then rescales the received integer values by multiplying them with 0.1 to restore their original floating-point precision (Fig. 3). The reconstructed feature vector is reformatted to match the input dimensions required by the classification model. In processing, the system outputs the

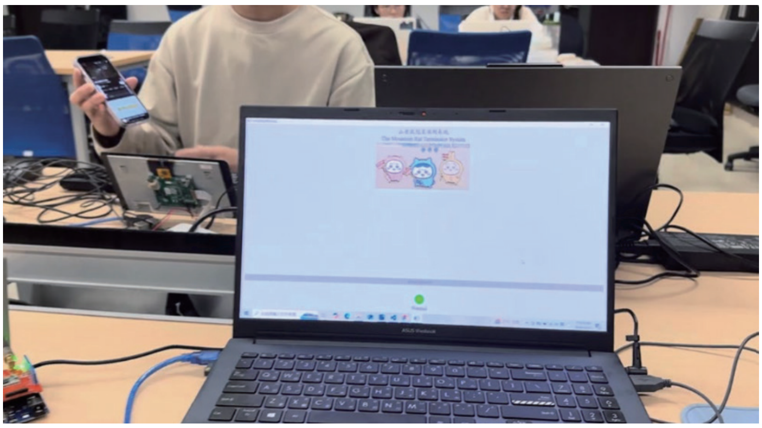

Fig. 3. (Color online) Demonstration of the developed forest sound monitoring system. The laptop displays the user interface showing classification results, while background sound forest audio clips (chainsaw sounds) are played on a smartphone to simulate real-world acoustic events.

predicted class label (e.g., “chainsaw” or “birdsong”), which is stored in a continuously updated comma-separated values (CSV) log. These results are displayed on GUI in real time, allowing users to monitor sounds continuously (Fig. 4). This entire cycle from audio capture to classification is in a loop, enabling uninterrupted forest surveillance until the program is manually terminated.

The recorded 5 s audio is trimmed, and a 1 s segment is randomly selected, as shown in Fig. 5. The segment is converted into a Mel spectrogram on RPi, and dimensionality is reduced by averaging across frequency bands. The resulting feature vector consists of negative decimal values representing spectral intensities, typically in the format of two-digit negative floating-point numbers (e.g., –xx.xx).

Since each LoRa message is limited to 255 characters, all floating-point values are multiplied by –10 and converted to positive integers by taking only the integer part. The resulting sequence consists of three-digit positive integers in the transmitted data sent to the Arduino module, as shown in Fig. 6.

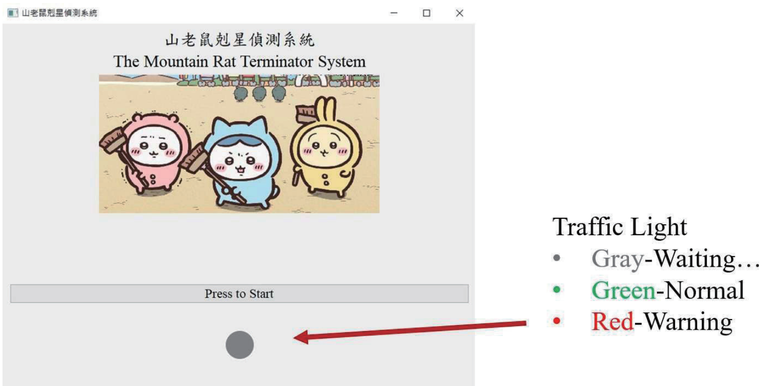

Fig. 4. (Color online) GUI of developed system.

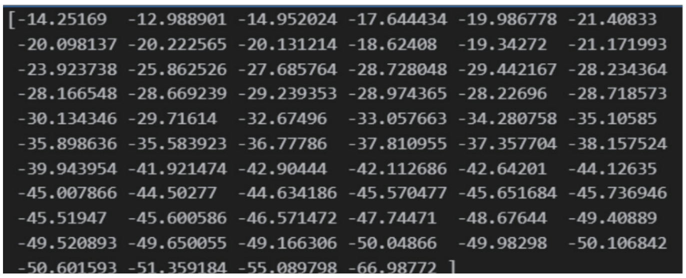

Fig. 5. Trimmed 5 s audio and randomly selected 1 s segments.

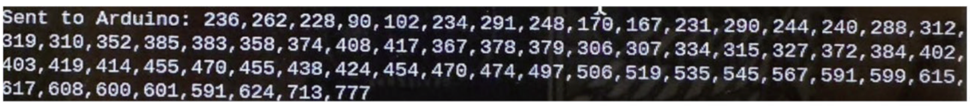

Fig. 6. Encoding of audio feature vectors for LoRa transmission.

4. Results and Discussion

4.1 Model's performance

The model's performance was evaluated by grouping 27 FSC22 classes into two categories: dangerous sounds (e.g., chainsaws and axes) and safe sounds (e.g., animal calls and ambient noise). This binary classification was used to reflect real-world monitoring needs in forest surveillance. The confusion matrix showed that the developed model correctly identified 106 dangerous and 81 safe sounds while misclassifying 52 dangerous sounds as safe (false negatives) and 65 safe sounds as dangerous (false positives). These results yielded a detection accuracy of 67.1% for dangerous sounds, with misclassifications primarily caused by overlapping frequencies or background noise, as depicted in Fig. 7. Although this value indicates moderate performance, the system effectively demonstrates its potential for dangerous-event detection under real-world conditions. Note that the 67.1% accuracy represents the offline, segment-level recall for the "danger" class under controlled testing. In contrast, the >85% accuracy reported in the abstract corresponds to the event-level performance obtained during the field deployment of the alert system, where temporal voting, threshold tuning, and a narrower set of threat classes were adopted. The discrepancy arises from differences in datasets, evaluation metrics, decision thresholds, and temporal aggregation strategies, which collectively result in the higher figure observed in the field-based evaluation. Future improvements will focus on refining the model through attention-based CNNs and enhanced data augmentation and denoising techniques to increase noise resilience and overall detection robustness.

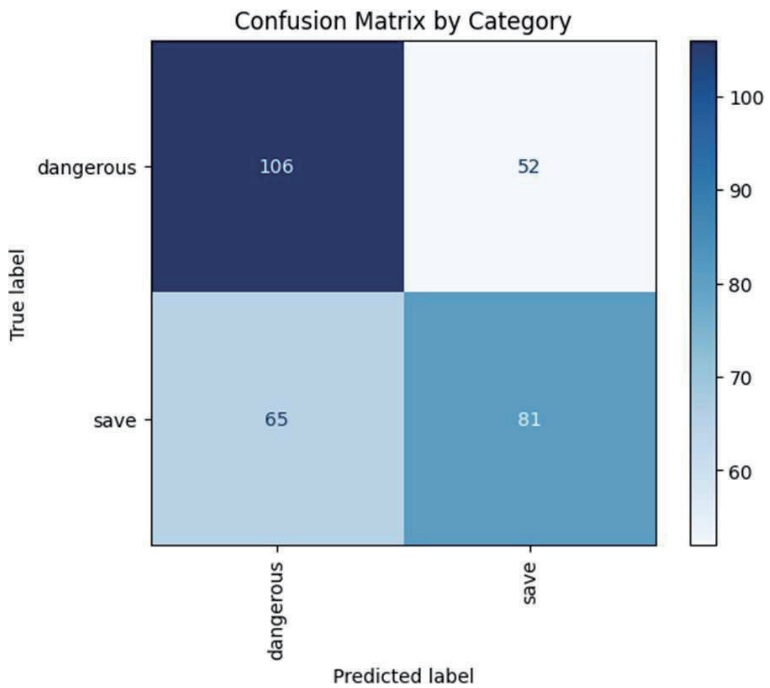

Fig. 7. (Color online) Confusion matrix of developed recognition model.

4.2 System demonstration

To validate the end-to-end functionality of the developed alert system—which integrates the RPi-based end node for feature extraction, the LoRa modules for transmission, and the server-side model for classification —a real-time demonstration was conducted. In this demonstration, the system successfully identified dangerous sounds, including the sound of a chainsaw, in real time. Once a chainsaw sound was captured by the microphone and preprocessed by RPi, the model successfully classified it into the "dangerous" category. The server-side component of the alert system then triggered an alert within seconds, updating the monitoring GUI and highlighting the detected sound with a corresponding label and timestamp. This immediate feedback mechanism allows forest rangers or stakeholders to respond promptly to potential illegal logging or other hazardous activities. The alert system operates continuously and autonomously for deployment in remote areas with limited human access. A demonstration of this real-time detection and alerting process of the developed alert system is available in the video linked to https://youtu.be/3iSb10vEZIE?si=1nAOwTmZadfgtm7j.

5. Conclusions

We developed a LoRa-based (ESC) system by integrating a LoRa-based transmission system with an ML model for real-time forest monitoring. The developed system accurately detects dangerous sounds such as the sound of chainsaws and triggers immediate alerts in a lightweight, energy-efficient hardware architecture. This enables rapid responses to illegal logging or other environmental threats, even in remote forests with limited infrastructure. The modular design, based on RPi and Arduino platforms, enables cost-effective deployment and scalability across different terrains. Additionally, LoRa ensures long-range communication with minimal power consumption for long-term operation. Although the developed system demonstrated real-time forest monitoring by leveraging LoRa technology and ML for sound classification, model optimization is required to improve detection accuracy, particularly in distinguishing subtle environmental and dangerous sounds. By incorporating data augmentation, dynamic thresholding, or noise-resilient architectures such as attention-based CNNs, false positives can be minimized and robustness can be enhanced for diverse soundscapes. Moreover, triangulation algorithms or other sensing modalities such as vibration sensors, optical cameras, or temperature detectors can be added to improve source localization and provide a richer environmental context, enabling enhanced situational awareness and proactive response.

System-level safeguards such as watchdog timers, emergency stop conditions, or redundancy modules need to be integrated into the developed system to ensure safety and resilience in the case of power loss, network failure, or hardware malfunction. It is also necessary to assess the durability, communication stability, and real-time inference of the developed system in the forest and under uncontrolled conditions. Through field validation, sensor placement strategies can be optimized depending on terrain, vegetation density, or weather variability. Edge AI or satellite-based data relay (e.g., via Starlink) needs to be considered to expand the deployment range and autonomous forest monitoring of the developed system. These refinements enhance scalability and long-term sustainability for environmental conservation. At the same time, classification accuracy can be enhanced by integrating attention-based CNNs or recursive neural networks and expanding the dataset to include a broader range of dangerous and safe sounds. Combining IoT and AI supports data-driven decision-making in sustainable environmental management and conservation efforts.

References

1 I. Mporas, I. Perikos, V. Kelefouras, and M. Paraskevas: Appl. Sci. **10** (2020) 7379.
2 M. Girbau-Llistuella, A. Caballé, and J. Silva-Campillo: Sensors **21** (2021) 7593.
3 A. Andreadis, G. Giambene, and R. Zambon: Sensors **21** (2021) 7593.
4 M. Bandara, R. Jayasundara, I. Ariyarathne, D. Meedeniya, and C. Perera: Heliyon **8** (2022) e09319.
5 Y. Wang, J. Wang, and Z. Wang: Sci. Rep. **11** (2021) 21045.
6 Y. Wang, J. Wang, and Z. Wang: Comput. Mater. Continua **73** (2022) 1001.
7 D. Meedeniya, I. Ariyarathne, M. Bandara, R. Jayasundara, and C. Perera: ACM Comput. Surv. **56** (2023) 66.
8 M. Bandara, R. Jayasundara, I. Ariyarathne, D. Meedeniya, and C. Perera: Sensors **23** (2023) 2032.
9 H. Wang, Y. Wang, and Z. Zhang: Sensors **23** (2023) 4678.
10 M. Bor, S. Yousefi, and H. Visser: IEEE Trans. Wireless Commun. **20** (2021) 6994.
11 F. Ferrari, M. Zimmermann, and L. Thiele: Sensors **21** (2021) 2600.
12 https://vtiya.medium.com/yamnet-for-audio-classification-867ff0ca9197 (accessed Oct. 2024).