

Audio-Based Drone Detection Via CRNN: An Investigation into Threshold Sensitivity and Stability

Zainab Ali Mezher*¹  Kadim Karim Mohsen¹ 

Received: 26th November 2025/Accepted: 26th July 2025/ Published Online: 1st June 2026

© The Author(s), under exclusive license to the University of Thi-Qar

Abstract

Unmanned Aerial Vehicles (UAVs), or commonly known as drones, have been utilized in numerous applications, including civilian and military environments, due to their low cost and high level of usability. These devices have triggered significant security and protection issues, demanding the development of reliable detection and classification systems. However, UAVs have been used to achieve illegal operations, resulting in emerging threats to personal and public security. The traditional detection method suffers from small targets, uncertain situations, and stealthy cases. Consequently, due to their accuracy and efficacy in detecting drones in a diverse range of scenarios, Artificial Intelligence (AI) has become a significant solution for drone detection and classification. In this paper, an audio-based mechanism has been developed using a convolutional recurrent neural network (CRNN) algorithm to exploit the exclusive audio fingerprint of drones for detecting and identifying them. The model achieved 94.85% accuracy with a 94.41% F1-score from different types of drone sound files collected from the DADS site. It operates with a three-seed and thresholds to guarantee the stability and dependability of the model. This paper aims to validate the use of this algorithm for drone detection in an actual environment and train the model with different seeds to measure the model's stability against random variation.

Keywords — Machine Learning, Drone Security, Deep Learning, Acoustic Feature

1 Introduction

Over the last few years, there has been a significant increase in drone research due to their increasing feasibility in various applications (Digulescu et al., 2020). Drones have become more attractive in both civil and military environments, owing to their low cost and variety of sizes (Coluccia et al., 2020). However, they pose a genuine threat when obtained by an unauthorized entity (Chang et al., 2018), so there is an urgent need to enhance the detection and classification of drone functions in the aerial domain (Frid et al., 2024). Traditional detection methods, including Radar, Radio Frequency (RF), Optical, and Acoustic sensors, are widely utilized to detect and identify drones (the pros and cons are mentioned in Table 1). Consequently, these created a demand for other efficient ways of detection and classification (Almasri, 2021). Machine Learning (ML) and Deep Learning (DL) based object detection techniques have achieved notable success due to their high accuracy and readily available processing capabilities (Singha & Aydin, 2021). Since most traditional methods for identifying or detecting undesirable UAVs haven't achieved a high enough prevention rate during detection, ML has shown considerable advantages in object

Zainab Ali Mezher

zainab.a.mezher@utq.edu.iq

Kadim Karim Mohsen

dkadim2020@utq.edu.iq

¹ Department of Mechanical Engineering, College of Engineering, University of Thi-Qar, Thi-Qar, Iraq

recognition and categorization across various disciplines, where it can recognize patterns without human assistance. For several reasons, including human limits in recognizing small objects and the possibility of focus problems due to boredom or tiredness, decreased reliance on human assistance is desirable. ML can identify forms through paradigms that are completely invisible to human sight, and these include traditional methods that are undetectable by human sensory systems (Rahman et al., 2024). Acoustic detection eliminates the need for line of sight, making it productive in complex environments or low-altitude drone operations (Tejera-Berengue et al., 2023). This study suggests using deep learning methods to detect drones using audio inputs. It aims to improve drone detection by merging Artificial Intelligence (AI) with a conventional identifying approach, especially the audio one. This will enhance the defense system for both military and civilian applications.

This paper is organized as follows: Section 2 mentions the related works. The proposed framework is in Section 3. Section 4 presents the experimental results. Some evaluation metrics are defined in Section 5. Finally, the paper concluded in Section 6 with a summary of the results and recommendations for further research.

Table 1. Pros and Cons of Traditional Methods

| Technology | Pros | Cons |
|-------------------------|---|--|
| Radar | Radar systems are advancing due to their intrinsic advantages (De Quevedo et al., 2018), such as operating day and night, long range, being less affected by weather, and operating even when the drone navigates autonomously (absent RF emissions) (Famili et al., 2024; Park et al., 2021). | However, these systems often encounter difficulties in detecting drones with a small radar cross-section (RCS), a measurement of a target's ability to scatter radar signals, broadcasting permission, and spectrum verification are necessary to avoid complications (Nemer et al., 2021; Samuell Aiad Saleip et al., 2021). |
| RF Signals | RF-based systems provide enhanced range and coverage, rendering them ideal for detecting UAVs over large areas (Vasant Ahirrao et al., 2024). It is relatively inexpensive (Medaiyese et al., 2022), shows a minimal rate of false alarms, and is easier to find UAVs and their operators (Sayed et al., 2024). | On the other hand, radio frequency demands an extensive database of RF drone/RC signals (Aouladhadj et al., 2023) and can be affected by dangerous or legally altered RF that will surpass receiver capabilities (Chiper et al., 2022). |
| Acoustic Sensors | Acoustic sensors are useful for differentiating drones from birds based on acoustic signals (Coluccia et al., 2020). Also show superior performance in low-visibility conditions (Zamri et al., 2024). | Acoustic sensors are sensitive to surrounding sounds, especially in noisy environments, are unsuccessful in windy conditions, and require a database of acoustic signatures for various drones for training and testing (Narayanan et al., 2023; Taha & Shoufan, 2019). |
| Optical Sensors | In optical detection, the developed cameras can already be used, with a low cost for essential optical sensors. Moreover, software applications identify and categorize objects (Guvenc et al., 2018; Matić et al., 2020). | Visual information presents several challenges, including the operation of UAVs in convoluted environments, which can result in target obstruction and unstable flight paths (Opromolla et al., 2018). Furthermore, the precision performance declines in adverse environments such as poor light, fog, dust, and other conditions, perhaps resulting in full failure at night, and also requires a line of sight (LOS) (Khan et al., 2022; Seidaliyeva et al., 2024). |

2 Related Work

Several research projects have focused on detecting drones using their auditory features, as acoustic-based detection is still a relatively new field of study. Seo et al. analyzed convolutional neural networks (CNNs) when using the Short-Time Fourier Transform (STFT) properties of the drone's acoustic signal for drone detection. The DJI Phantom 3 or Phantom 4 hovering drone is used for experiments conducted in an open area. The percentage of false alarms and detections for the 100-epoch model is 98.97% and 1.28, respectively, whereas the 10-epoch model shows a detection rate of 98.77% and a false alarm rate of 1.62% (Seo et al., 2018). They didn't utilize the CRNN architecture and didn't mention evaluation metrics such as F1-score. Ohlenbusch et al. classified UAVs according to their acoustic signature. They used the Support Vector Machine (SVM) decision algorithm on various datasets collected from different environments. The model achieved excellent results with an accuracy of 92.63% and a 93.19% F1-score. The model's results were not effective in a far distance and quiet situation. The authors excluded comprehensive evaluation metrics, like precision and recall (Ohlenbusch et al., 2021).

Al-Emadi et al. demonstrated a drone detection approach utilizing various algorithms, including a Convolutional Neural Network (CNN), a Recurrent Neural Network (RNN), and a Conventional Recurrent Neural Network (CRNN). The paper analyzes their performance on the author's recording drone dataset. The results outlined that CNN surpassed the RNN model with 96.38% accuracy and 95.90% F1-score. In all evaluation metrics, the CRNN model exceeded the RNN model. They considered that the CRNN model is more practical due to its processing speed and coherent results, even with the CNN's high general performance (Al-Emadi et al., 2019). Jasim and Hreshee proposed a similar strategy by introducing a three-hybrid model for drone detection and classification, namely CNN-Conformer, CRNN-Conformer, and RNN-Conformer. A dataset of 1,500 drone audio clips gathered from different drone types, including DJI Phantom 4 Pro and DJI Mavic 2 Pro. The results show a superiority of CNN-Conformer with 98% accuracy and 97% for CRNN-Conformer. The research exhibited the efficacy of drone detection using acoustic features with a hybrid DL model and presents a reliable solution for scenarios where other methods are limited (Jasim & Hreshee, 2025). Some approaches extract features using the Fast Fourier Transform (FFT), while others add more microphones to address potential issues (Busset et al., 2015; Svanstrom et al., 2020).

3 Proposed Framework

3.1 Dataset

This paper used a large number of drone audio files to train the model. The dataset was obtained by downloading files of the Drone Audio Detection Dataset (DADS) (GeronimoBasso, n.d.), a publicly available audio dataset that is specifically aimed at researching the classification of sounds of a drone to allow the model to learn sound features and avoid overfitting. The dataset consists of various audio clips recorded in different environments and situations with various drone types and background sounds. The data were split into two classes: 1) Drone with 5000 clips and 2) No-Drone with 5000 clips. Each audio clip was checked to ensure its quality and integrity. The data were split into 70% training, 15% validation, and 15% testing to develop a coherent and readily assessable model.

3.2 Data Preprocessing

The audio files were formatted into the WAV format, mono channel, and sampling rate of 16 kHz. The length utilized was a 1-second clip, where the defective and entirely quiet file was deleted. Additionally, segment the file whose length exceeds 1 second and add a loop for the shorter one. The audio recordings were segmented and transformed into Mel-spectrograms with constant parameters using the STFT. Feature values were then normalized to improve convergence and stable model training. The 1-second segmentation behaved better than the 2- and 5-second ones (Al-Emadi et al., 2019). Even though the original audio features may be lost, the only commercially available technique for training machine learning or deep learning algorithms on audio data is converting audio samples into spectrograms.

3.3 Data Labeling

Each audio segment was given a ground truth label following data collection and preprocessing, to facilitate supervised training and analysis.

- Segments with drone noises were identified using the keyword "drone".
- Segments that had additional background or ambient sounds—such as typing, traffic, human activity, or ambient noise—were marked as "no-drone".

For steady model training and balanced distribution, the 10,000 one-second audio segments in the dataset are split equally between 5,000 drone and 5,000 no-drone samples. Evaluation is possible by contrasting model predictions with the actual ground truth, thanks to the maintenance of class labels in a ground_truth.csv file. In Table 2, the data per label are shown.

Table 2. Data per Label

| Type of data | Records | | | Total |
|--------------|----------|-----------|-------|-------|
| | Original | Augmented | Total | |
| Drone | 5000 | 0 | 5000 | 10000 |
| No-drone | 5000 | 0 | 5000 | |

3.4 Neural Networks

This study computed one-second Mel-spectrograms with CRNN to identify drone sounds. Convolutional Neural Networks (CNNs) are employed to extract discriminative spectral features from time–frequency representations of audio signals, while Recurrent Neural Networks (RNNs) are used to model the temporal dependencies across consecutive frames. Using a Convolutional Recurrent Neural Network (CRNN) design by concatenating both, the system can utilize both spectral and temporal attributes of sound events, and thus achieve better detection (Çakır et al., 2017; Choi et al., 2016). The last fully connected layer provides a single detection probability as a sigmoid function, which is the probability of the presence of drones. To perform a more generalization, dropout regularization and the Adam optimizer are used with binary cross-entropy loss. Fig.1 represents the CRNN Architecture.

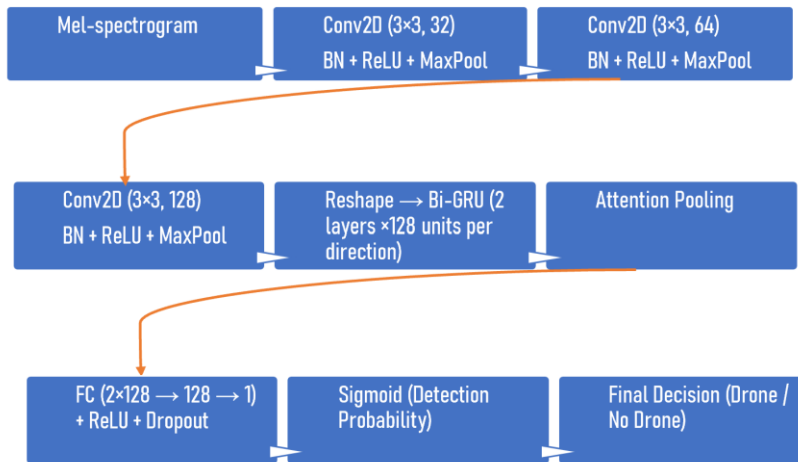


Figure 1. CRNN Architecture

3.1 Methodology and Setup

The suggested pipeline starts by dividing all audio recordings into one-second segments. Every clip is normalized, transformed into a Mel-spectrogram using STFT, and then fed into a CRNN. Fully connected layers with sigmoid activation generate a binary classification (drone/no-drone), recurrent layers record temporal dependencies, and convolutional layers extract time-frequency information. To ensure that the findings only represent the model's capacity to learn from actual data, no data augmentation was used. The neural network was trained with a batch size of 32 and an initial learning rate of 0.001 for 150 epochs in the experimental configuration using the Adam optimizer. Early stopping with a 10-epoch patience was utilized to avoid overfitting. For possible real-time deployment, it will also evaluate the model's stability across a range of thresholds and examine its computing efficiency in terms of training and inference time. The time signal and spectral analysis (FFT) indicate that the sound of the drone has a unique time-frequency pattern as compared to noise, which necessitates the application of the time-frequency analysis and deep-learning methods in the detection process (see Fig. 2). To display the process's output, see Fig. 3, which shows a one-second sample of drone sound. Sample of a random sound, like a human laugh, in Figure 4.

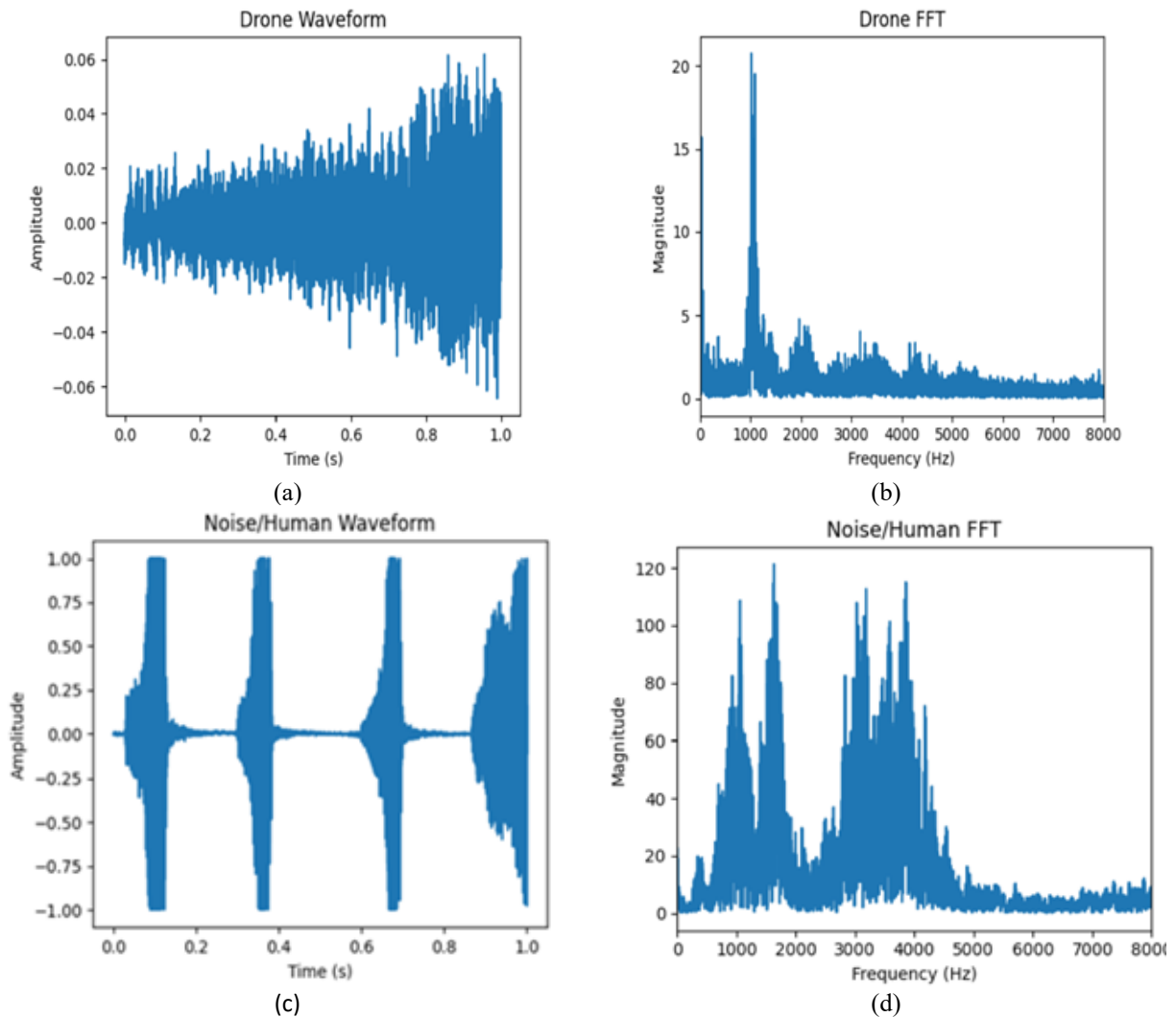


Figure 2. Time-domain waveforms and frequency-domain spectra (FFT) of (a), (b) drone audio and (c), (d) non-drone (human) audio

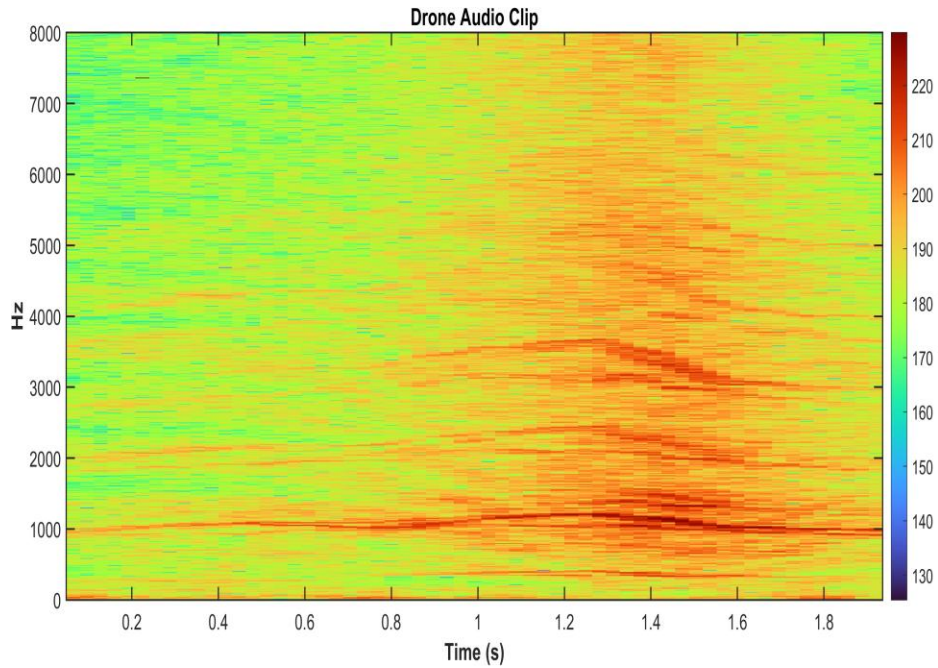


Figure 3. Example of drone noise in spectrogram representation

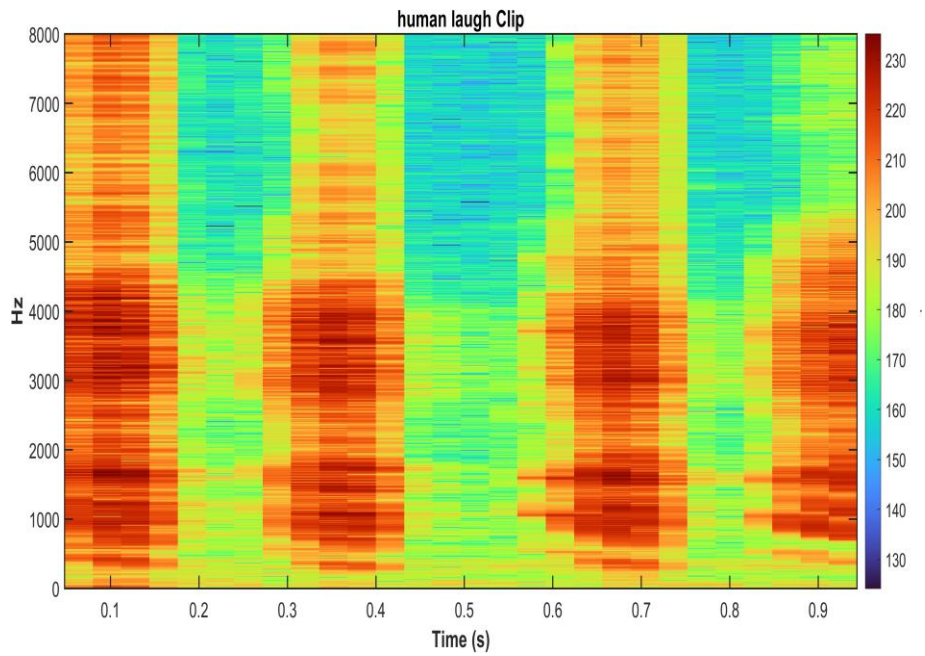


Figure 4. Example of a human laugh in spectrogram representation.

In ML research, seed 42 is typically utilized as a default seed. Therefore, to demonstrate that the model is real, steady, and reproducible, three adjacent seed types (42, 43, and 44) were used, as shown in Table 3. Evaluation metrics were used to exhibit the consistency of the results, where the AUC values exceeded 99.6% for all seeds with ± 0.02 F1-score. The results demonstrate the model's robustness, as it remained unaffected by random weight initialization, thereby ensuring accurate and stable learning. The model was checked using common classification

measures, including accuracy, F1 score, precision, and recall, to meet these goals. The training process took just over an hour on a CUDA-enabled GPU. Details on the setup of the environment are presented in Table 4.

Table 3. Performance Stability Across Different Random Seeds

| Random Seed | F1 | Precision | Recall | AUC |
|-------------|-------|-----------|--------|-------|
| 42 | 0.942 | 0.900 | 0.988 | 0.998 |
| 43 | 0.982 | 0.979 | 0.985 | 0.997 |
| 44 | 0.985 | 0.985 | 0.985 | 0.997 |

Table 4. Details of the Environment Setup

| | |
|--------------------------------|---|
| Operating System | Windows 11 Enterprise, Version 24H2 (OS Build 26100.6584). (ASUS laptop) |
| Programming language | Python 3.12 |
| CPU | Intel(R) Core (TM) Ultra 7 155H (22 CPUs), ~3.8 GHz |
| Number of CPU Cores | 16 |
| Framework and libraries | PyTorch 2.0.0, TorchAudio 2.0.0, Librosa 0.10.0, Scikit-learn 1.3.0 |

Al-Emadi et al. tested various dataset split ratios for testing, validation, and training, and found that the split ratio didn't affect performance variance. So, a 70:30 mix is utilized. The data is with 70% for training and 15% for both validation and testing (Al-Emadi et al., 2019). The model flowchart is illustrated in Fig.5.

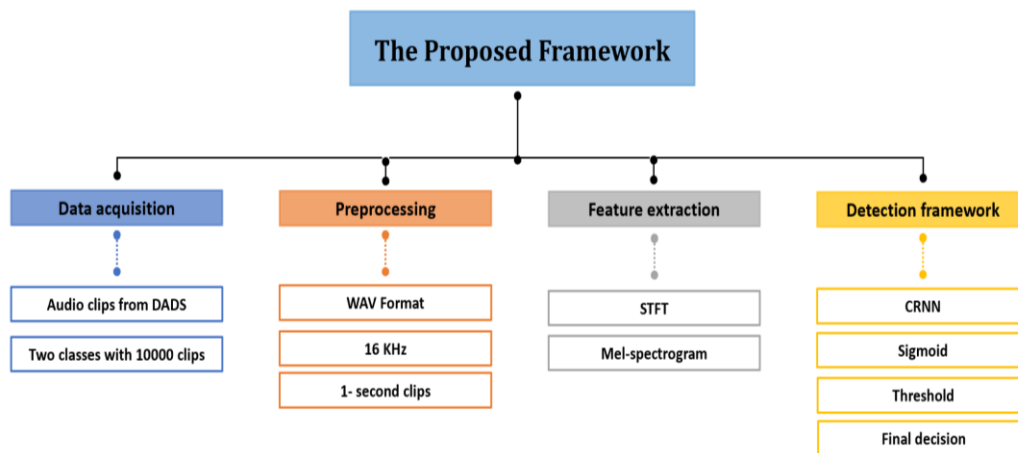


Figure 5. The system flowchart

4 Simulation Results

The experimental outcomes of training and assessing the CRNN model on the drone audio dataset are discussed in this section. Three separate random seeds (42, 43, and 44) were used in the training, and three decision thresholds (0.40, 0.50, and 0.65) were applied for the best seed only to measure the study's resilience. The (0.50) threshold is the standard boundary for detection, balancing precision and recall. A lower threshold of (0.40) raised sensitivity for detecting potential drones with reduced precision. In contrast, a higher threshold of 0.65 prioritized precision but minimized recall. Experiments using multiple random seeds were conducted to assess the stability of results. The 42 seed achieved the best results; consequently, it was utilized in the model. Table 5 illustrates the three seed results.

Table 5. Drone Detection Results

| Evaluation Metrics (%) | CRNN (0.40) | CRNN (0.50) | CRNN (0.65) |
|------------------------|-------------|-------------|-------------|
| Accuracy | 91.59 | 94.86 | 83.17 |
| Precision | 89.42 | 95.87 | 97 |
| Recall | 93 | 93 | 66 |
| F1-Score | 91.17 | 94.41 | 78.57 |

The three evaluation thresholds showed clear differences in the results. At the lowest one (0.40), the model was able to detect most real cases, where the recall reached 93% while the precision reduced to 89% due to the increase in FP. With an F1-score of 0.94 and an overall accuracy of 0.95, the model performed most consistently and balancedly at threshold 0.50, demonstrating the ideal trade-off between recall and precision. Recall dropped precipitously to 0.66 at threshold 0.65, lowering the F1-score to 0.78, but precision increased to 0.97. These findings are consistent with the expected trade-off behavior: whereas raising the threshold reduces false positives but increases false negatives, decreasing the threshold boosts recall but also increases the chance of false positives. Numerical results from Table 5, which looked at thresholds of 0.4, 0.5, and 0.65, confirm this conclusion. The 0.5 threshold achieved the best balance for the F1 score, while the 0.65 threshold prioritizes accuracy and renders the model more conservative. The improvement in recall due to a 0.40 threshold. Fig. 6 shows both Accuracy and F1-score for three seeds, which were the best at a 0.5 threshold.

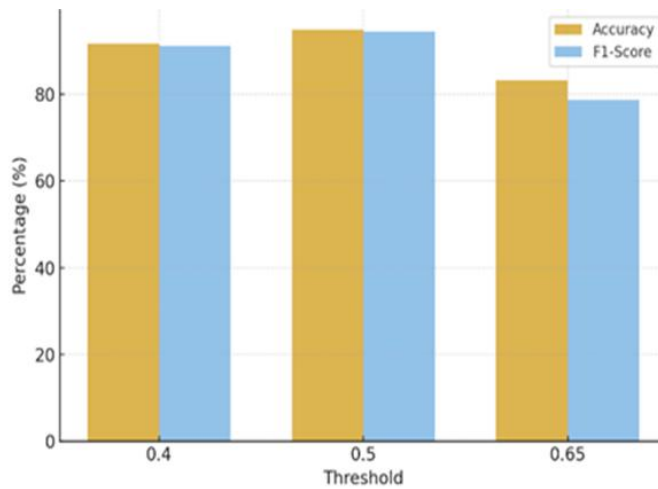


Figure 6. Accuracy and F1-Score Across Thresholds

The plot in Fig. 7 illustrates the increase in precision to 97% with a large decrease in recall (66%). Therefore, the selection of thresholds depends on the primary objective of the system, whether it is to detect all targets, a security system, or an early warning system.

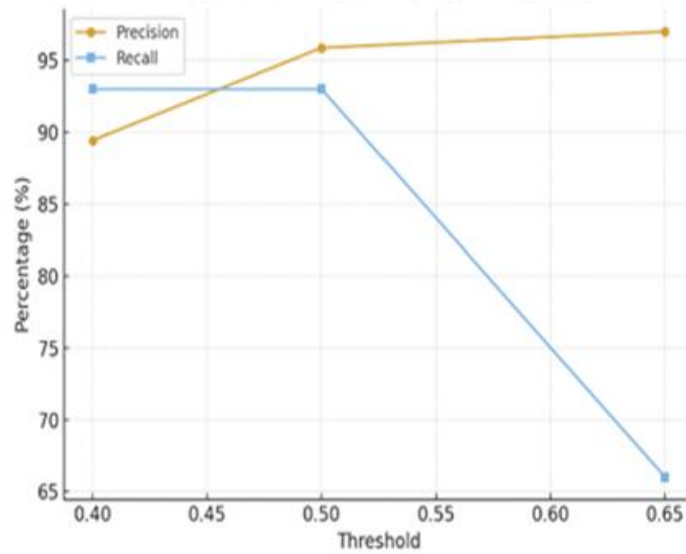
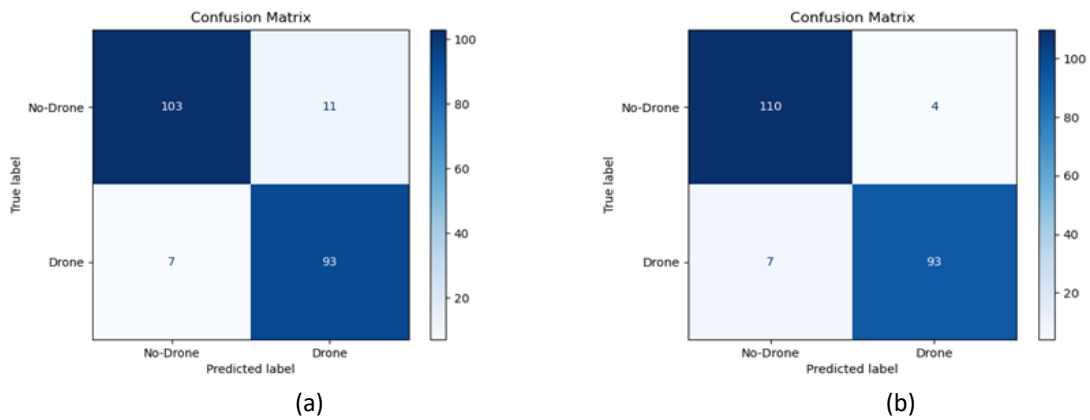
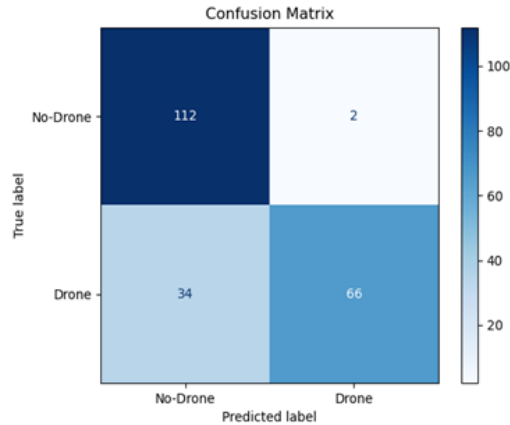


Figure 7. Precision and Recall Across Thresholds

The confusion matrices of all seeds are shown in Figures 8.





(c)

Figure 8. Confusion Matrix for Threshold values of (a) 0.4, (b) 0.5, (c) 0.6

Some studies have shown a greater level of accuracy, but they are tested on small or proprietary data, with fewer types of drones. Conversely, the suggested approach is tested on the more varied DADS dataset, and it shows good generalization, as shown in Table 6.

Table 6: The evaluation results of the proposed CRNN are compared with those of previous research

| Method | Dataset | Drone Types | Model | Accuracy |
|----------------------------|---------------|-----------------------------|-------------------------------|----------|
| (Seo et al., 2018) | Self-recorded | DJI Phantom 3 and Phantom 4 | CNN | 98.87% |
| (Al-Emadi et al., 2019) | Self-recorded | Parrot Mambo, Parrot Bebop | CNN, RNN, CRNN | 94.72% |
| (Casabianca & Zhang, 2021) | Self-recorded | Multicopter UAV | CNN + late fusion | 94.7% |
| (Frid et al., 2024) | Self-recorded | Multi types | LSTM (RF + Acoustic) | 91% |
| (Mapara et al., 2025) | Self-recorded | (c) ified | Random Forest (RF) | 89% |
| (Ren et al., 2025) | Self-recorded | ... types | CNN + adaptive feature fusion | 94.50% |
| Proposed Method | DADS | Multi types | CRNN | 94.86% |

5 Evaluation Metrics

There is an urgent need to evaluate the model's performance, especially in DL and ML (Dalianis, 2018).

Some evaluations are mentioned below:

True positive: Expected the drone class correctly.

$$TP = \frac{TP}{TP+FN} \tag{1}$$

False Positives: When a no-drone class is identified as a drone.

$$FP = \frac{FP}{TN+FP} \tag{2}$$

False negative: The model incorrectly classified the drone as a no-drone.

Recall: Present how many positive samples are predicted as positive.

$$\text{Recall} = \frac{\text{TP}}{\text{TP}+\text{FN}} \times 100\% \quad (3)$$

Precision: Indicates how many of the positives predicted are actually positive.

$$\text{Precision} = \frac{\text{TP}}{\text{TP}+\text{FP}} \times 100\% \quad (4)$$

F1-score: Is a merged value between precision and recall. Gives a balanced evaluation of the real model situation.

$$\text{F1-score} = \frac{\text{PRECISION} \times \text{RECALL}}{\text{PRECISION} + \text{RECALL}} \times 100\% \quad (5)$$

6 Conclusion and Future Work

The paper introduces a drone detection model based on acoustics and implemented as a CRNN model. It is trained on the publicly available DADS dataset. The experimental findings proved that the suggested method has a good level of detection ability in a variety of acoustic situations, with the most significant outcomes being recorded with the decision threshold of 0.5, the optimal level of trade-off between accuracy and recall. In addition, the use of multiple random seeds confirmed the stability of the training process and showed that the obtained results are not dependent on random initialization. The CRNN model has a high performance with a 94% F1-score and 94.8% accuracy, proving the robustness of the model. The threshold sensitivity analysis revealed that the choice of a suitable value has a direct impact on the precision/detection rate relationship, and that it is reasonable to change this value based on the practical application needs. The proposed framework has shown the capability of generalizing, despite the challenges of noise and the variety of recording conditions, when applied to unseen data. The proposed method achieves a competitive performance, despite being based solely on acoustic signals, as compared to recent methods that often rely on self-recorded datasets or less diverse datasets. The proposed method can be utilized in real-world surveillance and security systems because it is low-cost, therefore, allowing it to be used in the security of critical infrastructure and restricted zones. For future work, merging several algorithms will be utilized to enhance the detection, such as computer vision or machine learning approaches. Using augmentation approaches will enhance the model's resilience and improve the ability to detect distant or weakly signaled drones.

References

- Al-Emadi, S., Al-Ali, Abdulla, Mohammad, A., & Al-Ali, Abdulaziz. (2019). Audio based drone detection and identification using deep learning. *2019 15th International Wireless Communications and Mobile Computing Conference, IWCMC 2019*, 459–464. <https://doi.org/10.1109/IWCMC.2019.8766732>
- Almasri, M. (2021). Deep learning for RF-based drone detection and identification using welch's method. *Proceedings of the 10th International Conference on Data Science, Technology and Applications, DATA 2021*, 208–214. <https://doi.org/10.5220/0010530302080214>
- Aouladhadj, D., Kpre, E., Deniau, V., Kharchouf, A., Gransart, C., & Gaquière, C. (2023). Drone Detection and Tracking Using RF Identification Signals. *Sensors*, 23(17). <https://doi.org/10.3390/s23177650>
- Busset, J., Perrodin, F., Wellig, P., Ott, B., Heutschi, K., Rühl, T., & Nussbaumer, T. (2015). Detection and tracking of drones using advanced acoustic cameras. *Unmanned/Unattended Sensors and Sensor Networks XI; and*

- Advanced Free-Space Optical Communication Techniques and Applications*, 9647, 96470F. <https://doi.org/10.1117/12.2194309>
- Çakır, E., Parascandolo, G., Heittola, T., Huttunen, H., & Virtanen, T. (2017). *Convolutional Recurrent Neural Networks for Polyphonic Sound Event Detection*. <https://doi.org/10.1109/TASLP.2017.2690575>
- Casabianca, P., & Zhang, Y. (2021). Acoustic-based UAV detection using late fusion of deep neural networks. *Drones*, 5(3). <https://doi.org/10.3390/drones5030054>
- Chang, X., Yang, C., Wu, J., Shi, X., & Shi, Z. (2018). A Surveillance System for Drone Localization and Tracking Using Acoustic Arrays. *Proceedings of the IEEE Sensor Array and Multichannel Signal Processing Workshop, 2018-July*, 573–577. <https://doi.org/10.1109/SAM.2018.8448409>
- Chiper, F. L., Martian, A., Vladeanu, C., Marghescu, I., Craciunescu, R., & Fratu, O. (2022). Drone Detection and Defense Systems: Survey and a Software-Defined Radio-Based Solution. In *Sensors* (Vol. 22, Number 4). MDPI. <https://doi.org/10.3390/s22041453>
- Choi, K., Fazekas, G., Sandler, M., & Cho, K. (2016). *Convolutional Recurrent Neural Networks for Music Classification*. <http://arxiv.org/abs/1609.04243>
- Coluccia, A., Parisi, G., & Fascista, A. (2020). Detection and classification of multirotor drones in radar sensor networks: A review. In *Sensors (Switzerland)* (Vol. 20, Number 15, pp. 1–22). MDPI AG. <https://doi.org/10.3390/s20154172>
- Dalianis, H. (2018). Evaluation Metrics and Evaluation. In *Clinical Text Mining* (pp. 45–53). Springer International Publishing. https://doi.org/10.1007/978-3-319-78503-5_6
- De Quevedo, Á. D., Urzaiz, F. I., Menoyo, J. G., & López, A. A. (2018). Drone Detection and RCS Measurements with Ubiquitous Radar. *2018 International Conference on Radar, RADAR 2018*. <https://doi.org/10.1109/RADAR.2018.8557320>
- Digulescu, A., Despina-Stoian, C., Stănescu, D., Popescu, F., Enache, F., Ioana, C., Rădoi, E., Rîncu, I., & Șerbănescu, A. (2020). New approach of uav movement detection and characterization using advanced signal processing methods based on uwb sensing. *Sensors (Switzerland)*, 20(20), 1–18. <https://doi.org/10.3390/s20205904>
- Famili, A., Stavrou, A., Wang, H., Park, J. M., & Gerdes, R. (2024). Securing Your Airspace: Detection of Drones Trespassing Protected Areas. In *Sensors* (Vol. 24, Number 7). Multidisciplinary Digital Publishing Institute (MDPI). <https://doi.org/10.3390/s24072028>
- Frid, A., Ben-Shimol, Y., Manor, E., & Greenberg, S. (2024). Drones Detection Using a Fusion of RF and Acoustic Features and Deep Neural Networks. *Sensors*, 24(8). <https://doi.org/10.3390/s24082427>
- GeronimoBasso. (n.d.). *Drone audio detection samples (DADS)*. Hugging Face. Retrieved <https://huggingface.co/datasets/geronimobasso/drone-audio-detection-samples>
- Guvenc, I., Koohifar, F., Singh, S., Sichitiu, M. L., & Matolak, D. (2018). Detection, Tracking, and Interdiction for Amateur Drones. *IEEE Communications Magazine*, 56(4), 75–81. <https://doi.org/10.1109/MCOM.2018.1700455>
- Jasim, S. W., & Hreshee, S. S. (2025). Real Time Drone Detection Based on Acoustics Using Hybrid Deep Learning Models. *Journal of Internet Services and Information Security*, 15(2), 673–693. <https://doi.org/10.58346/JISIS.2025.12.046>

- Khan, M. A., Menouar, H., Eldeeb, A., Abu-Dayya, A., & Salim, F. D. (2022). On the Detection of Unauthorized Drones - Techniques and Future Perspectives: A Review. *IEEE Sensors Journal*, 22(12), 11439–11455. <https://doi.org/10.1109/JSEN.2022.3171293>
- Mapara, T. M., Sesham, S., & Sesham, P. K. (2025). *Lightweight Machine Learning Models for Drone Detection Using Acoustic and Optical Features*. <https://doi.org/10.21203/rs.3.rs-6688717/v1>
- Matić, V., Kosjer, V., Lebl, A., Pavić, B., & Radivojević, J. (2020). *Methods for Drone Detection and Jamming*.
- Medaiyese, O. O., Ezuma, M., Lauf, A. P., & Guvenc, I. (2022). *Wavelet Transform Analytics for RF-Based UAV Detection and Identification System Using Machine Learning* *. <https://www.elsevier.com/open-access/userlicense/1.0/>
- Narayanan, R. M., Tsang, B., & Bharadwaj, R. (2023). Classification and Discrimination of Birds and Small Drones Using Radar Micro-Doppler Spectrogram Images †. *Signals*, 4(2), 337–358. <https://doi.org/10.3390/signals4020018>
- Nemer, I., Sheltami, T., Ahmad, I., Yasar, A. U. H., & Abdeen, M. A. R. (2021). Rf-based UAV detection and identification using hierarchical learning approach. *Sensors*, 21(6), 1–23. <https://doi.org/10.3390/s21061947>
- Ohlenbusch, M., Ahrens, A., Rollwage, C., & Bitzer, J. (2021). Robust drone detection for acoustic monitoring applications. *European Signal Processing Conference, 2021-January*, 6–10. <https://doi.org/10.23919/EUSIPCO47968.2020.9287433>
- Opromolla, R., Fasano, G., & Accardo, D. (2018). A vision-based approach to uav detection and tracking in cooperative applications. *Sensors (Switzerland)*, 18(10). <https://doi.org/10.3390/s18103391>
- Park, S., Kim, H. T., Lee, S., Joo, H., & Kim, H. (2021). Survey on Anti-Drone Systems: Components, Designs, and Challenges. *IEEE Access*, 9, 42635–42659. <https://doi.org/10.1109/ACCESS.2021.3065926>
- Rahman, M. H., Sejan, M. A. S., Aziz, M. A., Tabassum, R., Baik, J. I., & Song, H. K. (2024). A Comprehensive Survey of Unmanned Aerial Vehicles Detection and Classification Using Machine Learning Approach: Challenges, Solutions, and Future Directions. In *Remote Sensing* (Vol. 16, Number 5). Multidisciplinary Digital Publishing Institute (MDPI). <https://doi.org/10.3390/rs16050879>
- Ren, J., Wang, Z., Zhao, J., & Liu, X. (2025). A Drone Sound Recognition Approach Using Adaptive Feature Fusion and Cross-Attention Feature Enhancement. *Electronics (Switzerland)*, 14(8). <https://doi.org/10.3390/electronics14081491>
- Samuell Aiad Saleip, G. N. A., Maurer, P., Hassan, A., Frangenberg, M., & Granig, W. (2021). A Deep learning approach for multi-copter detection using mm-wave radar sensors. *ACM International Conference Proceeding Series*, 106–111. <https://doi.org/10.1145/3505688.3505706>
- Sayed, A. N., Ramahi, O. M., & Shaker, G. (2024). Frequency-Modulated Continuous-Wave Radar Perspectives on Unmanned Aerial Vehicle Detection and Classification: A Primer for Researchers with Comprehensive Machine Learning Review and Emphasis on Full-Wave Electromagnetic Computer-Aided Design Tools. In *Drones* (Vol. 8, Number 8). Multidisciplinary Digital Publishing Institute (MDPI). <https://doi.org/10.3390/drones8080370>
- Seidaliyeva, U., Ilibayeva, L., Taissariyeva, K., Smailov, N., & Matson, E. T. (2024). Advances and Challenges in Drone Detection and Classification Techniques: A State-of-the-Art Review. In *Sensors* (Vol. 24, Number 1). Multidisciplinary Digital Publishing Institute (MDPI). <https://doi.org/10.3390/s24010125>

- Seo, Y., Jang, B., & Im, S. (2018). Drone Detection Using Convolutional Neural Networks with Acoustic STFT Features. *Proceedings of AVSS 2018 - 2018 15th IEEE International Conference on Advanced Video and Signal-Based Surveillance*. <https://doi.org/10.1109/AVSS.2018.8639425>
- Singha, S., & Aydin, B. (2021). Automated drone detection using YOLOv4. *Drones*, 5(3). <https://doi.org/10.3390/drones5030095>
- Svanstrom, F., Englund, C., & Alonso-Fernandez, F. (2020). *Real-Time Drone Detection and Tracking With Visible, Thermal and Acoustic Sensors*. <http://arxiv.org/abs/2007.07396>
- Taha, B., & Shoufan, A. (2019). Machine Learning-Based Drone Detection and Classification: State-of-the-Art in Research. *IEEE Access*, 7, 138669–138682. <https://doi.org/10.1109/ACCESS.2019.2942944>
- Tejera-Berengue, D., Zhu-Zhou, F., Utrilla-Manso, M., Gil-Pita, R., & Rosa-Zurera, M. (2023). Acoustic-Based Detection of UAVs Using Machine Learning: Analysis of Distance and Environmental Effects. *2023 IEEE Sensors Applications Symposium, SAS 2023 - Proceedings*. <https://doi.org/10.1109/SAS58821.2023.10254127>
- Vasant Ahirrao, Y., Yadav, R. P., & Kumar, S. (2024). RF-Based UAV Detection and Identification Enhanced by Machine Learning Approach. *IEEE Access*, 12, 177735–177745. <https://doi.org/10.1109/ACCESS.2024.3502754>
- Zamri, F. N. M., Gunawan, T. S., Kartiwi, M., Pratondo, A., Yusoff, S. H., & Mustafah, Y. M. (2024). Deep Learning Techniques for Advanced Drone Detection Systems: A Comprehensive Review of Techniques, Challenges and Future Directions. *Indonesian Journal of Electrical Engineering and Informatics*, 12(4), 818–857. <https://doi.org/10.52549/ijeei.v12i4.6028>