2022 Workshop on Benchmarking Cyber-Physical Systems and Internet of Things (CPS-IoTBench) | 978-1-6654-7038-4/22/$31.00 ©2022 IEEE | DOI: 10.1109/CPS-IoTBench56135.2022.00008

Benchmarking Audio-based Deep Learning Models for Detection and Identification of Unmanned Aerial Vehicles

Sai Srinadhu Katta, Sivaprasad Nandyala, Eduardo Kugler Viegas, Abdelrahman AlMahmoud
Secure Systems Research Centre (SSRC)
Technology Innovation Institute (TII)
Abu Dhabi, U.A.E.
{sai, siva, eduardo, abdelrahman}@ssrc.tii.ae

***Abstract*—Over the last few years, Unmanned Aerial Vehicles (UAVs) have become increasingly popular for both commercial and personal applications. As a result, security concerns in both physical and cyber domains have been raised, as a malicious UAV can be used for the jamming of nearby targets or even for carrying explosive assets. UAV detection and identification is a very important task for safety and security. In this regard, several techniques have been proposed for the detection and identification of UAVs, in general, through image, audio, radar, and RF based approaches. In this paper, we benchmark the detection and identification of UAVs via audio data from [1]. We benchmarked with widely used deep learning algorithms such as Deep Neural Networks (DNN), Convolutional Neural Network (CNN), Long Short Term Memory (LSTM), Convolutional Long Short Term Memory (CLSTM) and Transformer Encoders (TE). In addition to the dataset of [1], we collected our own diverse identification audio dataset and experimented with Deep Neural Networks (DNN). In a UAV detection task, our best model (LSTM) outperformed the best model of [1] (CRNN) by over 4% in accuracy, 2% in precision, 4% in recall and 4% in F1-score. In UAV identification task, our best model (LSTM) outperformed the best model of [1] (CNN) by over 5% in accuracy, 2% in precision, 4% in recall and 3% in F1-score.**

***Index Terms*—Audio, Deep learning, Detection, Identification, UAVs**

I. Introduction

The popularity of Unmanned Aerial Vehicles (UAV), commonly referred as drones, have significantly increased over the last few years, followed by the technological advancements of their on-board components. In practice, modern UAVs have enabled the deployment of user customized solutions, which are able to analyze the data from a variety of UAV hardware components including the camera, microphones, LiDAR, accelerometer, GPS among others [2]. This ease of customized UAV solutions has paved the way for several autonomous UAV applications, such as object delivery, field surveillance, and even border control etc. [3].

Unfortunately, modern UAVs can also be used for malicious purposes. Unsurprisingly, drone-based attacks can have a significant negative impact on the economy, safety and security.

As a result, in recent years several works have been proposed for the detection and identification of nearby UAVs [4]–[8]. Detection and identification of nearby UAVs are generally achieved using vision-based analysis [9], radio fingerprint detection [10], radar-based identification [11], and microphone-based approaches [12], [13]. In this work we focus on UAV detection and identification by audio-based approaches. Machine Learning (ML) based techniques are predominantly used for analyzing audio acoustic UAV signatures for UAV detection and identification [12]–[17].

We benchmarked UAV drone detection and identification from publicly available audio dataset [1], via commonly used deep learning algorithms namely DNN, CNN, LSTM, CLSTM and TE. In addition to [1], we collected our diverse identification audio dataset and experimented with DNN models on it.

The main contributions of this paper are as follows:

- We benchmarked various deep learning models namely DNN, CNN, LSTM, CLSTM and TE on publicly available UAV detection and identification audio dataset [1]. Our models perform significantly better than the [1].
- In addition to the dataset of [1], we collected our own diverse identification audio dataset consisting of 7 different categories of UAVs namely no-UAV, drone, helicopter, drone-membo, drone-bebop, airplane, and drone-hovering. We experimented with DNN models on this and results are promising.

The remainder of the paper is organized as follows. Section II describes the literature review. Section III presents the audio-based scheme for UAV detection and identification through deep learning architectures. Section IV talks about experimental results and discussion. Finally, Section V concludes our work.

II. Literature review

Over the last years, several techniques have been proposed for UAV detection and identification, ranging from video [5], radio frequency [6], [18], thermal imaging [7], radar [11], and more recently audio-based [12], [14], [19]–[22]. In such a context, several approaches have been proposed for audio-based UAV detection and identification techniques, as they can be easily deployed due to their negligible equipment costs, and the promising accuracy results reported in the literature.

978-1-6654-7038-4/22/$31.00 ©2022 IEEE
DOI 10.1109/CPS-IoTBench56135.2022.00008

To achieve such a goal, UAV audio-based detection and identification is often implemented through four sequential modules. First, the *Data Acquisition* module collects sound samples from a given microphone. In general, the collected data is evaluated according to a given predefined time interval, e.g. every 1 second. Then, the *Feature Extraction* module extracts a set of behavioral features from the analyzed audio sample, compounding a feature set. Several techniques can be used to fulfill such a task in audio-based detection, including the extraction of the audio spectrogram [17], [23], and the building of the audio coefficients [8]. Finally, the *Detection/Identification* module classifies the built feature set in one of the selected classes.

In recent years, a plethora of highly accurate audio-based detection and identification approaches have been proposed for the detection and identification tasks in UAV detection and identification [1], [14], [18], [20], [24]. In general, authors resort to ML approaches, typically implemented through pattern recognition techniques. To fulfill such a task, the operator typically relies in a two-phase process, namely *training*, and *testing*.

The *training* step aims at training of the ML model with the training dataset and choosing the best model on the validation dataset. The *testing* phase evaluates the final model detection and identification performance metrics. In practice, the performance measurements obtained at *testing* phase are expected to be evidenced when the designed system is deployed in production environments.

III. Proposed approach

The tasks here are detection and identification of UAV via UAV audio signatures. The overall procedure of our method is shown in Figure 1.

Almost all the related existing literature works [12], [13], [15]–[17] exclusively followed the representation learning techniques, we have also picked that as a design choice for the implementation. Previous results on various settings of this task gave state-of-the-art results. The audio data is sent through feature extraction, the extracted features are passed through the deep learning models giving the identification and detection results. The end-to-end model is as follows: first an audio is sent through mel frequency cepstral coefficients (MFCC) feature extraction,followed by a deep learning model of choice (DNN, CNN, LSTM, CLSTM, TE) giving us the result of detection and identification. Based on the result safety/alert actions can be performed accordingly. The detailed deep learning architectures are as follows:

1) Deep Neural Network (DNN): The DNN is made of fully-connected layers and non-linear activations. The input to the DNN is the flattened MFCC features, which feeds into a stack of hidden fully-connected layers. At the output is a linear layer followed by a softmax layer generating the output probabilities of the classes.

2) Convolutional Neural Network (CNN): CNNs exploit the local temporal and spectral correlation in the features via 2D convolution. The input to the CNN is the MFCC features, which feeds into a stack of convolution layers. At the output is a linear layer followed by a softmax layer generating the output probabilities of the classes.

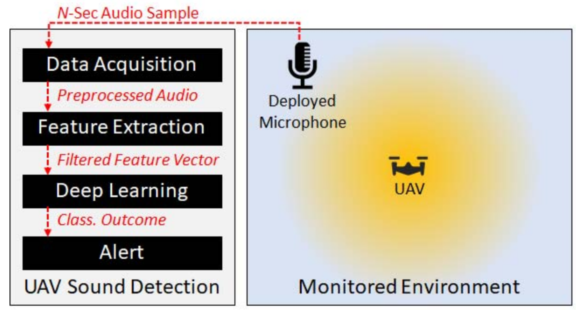

Fig. 1. Overview of the proposed deep learning model for audio-based detection and identification of UAVs.

3) Long Short-Term Memory (LSTM): LSTMs are known to model long term dependencies and are shown to work very well on various sequence modelling tasks. The input to the LSTM is the MFCC features, and the whole flattened output sequence is fed to a linear layer followed by softmax for output probabilities of the classes.

4) Convolutional Long Short-Term Memory (CLSTM): CLSTMs are combination of convoultion followed by LSTMs. This has the benefits of both CNNs and LSTMs. CLSTMs exploit the local temporal and spectral correlation, and model the long term dependencies well. The input to the CLSTM is the MFCC features, which has convolution followed by LSTM, and the whole flattened output sequence is fed to a linear layer followed by softmax for output probabilities of the classes

5) Transformer Encoder (TE): Transformers are shown to be the fundamental block for SOTA on various sequence modelling tasks across various domains. In this work we use only the Transformer Encoder part. The input is the MFCC features, and the whole output sequence is fed to a linear layer followed by softmax for output probabilities of the classes.

IV. Experimental results and discussion

Our proposed scheme was evaluated and compared considering both literature and our own built dataset. The evaluation aims at answering the following research questions, as follows: (***RQ1***) *What is the detection accuracy of audio-based techniques in a publicly available dataset?* (***RQ2***) *What is the identification accuracy when different UAV types are considered?* (***RQ3***) *What is the identification performance impact when more diverse flying devices are considered?*.

The proposed scheme for an audio-based detection and identification of nearby UAVs is implemented through a pattern recognition pipeline. Therefore, we consider a given microphone deployed in a monitored environment (Fig. 1, *Deployed Microphone*), which will be used for the collection and periodic sending of the environment audio samples. For instance, collection of 1-sec batches of audio in a predefined format. The main assumption is that nearby flying UAVs will produce audio noises that will be captured by the microphone for further analysis. The collected audio sample is analyzed by the feature extraction module, which goal is to extract a set of UAV related features. To achieve such a goal the module applies an audio filtering technique (e.g., filtering UAV-related audio signals through the Mel-Frequency Cepstral Coefficients (MFCC)), before using it for the detection and identification tasks. As a result, a portion of non UAV-related audio can be removed from the analyzed sample, improving the system generalization even in highly noise environments. Finally, the extracted filtered feature vector (Fig. 1, *Filtered Feature Vector*) is used as input by a deep learning classifier, which outputs a corresponding event class. In such a case, the used deep learning model can output the analyzed event label in a two-class setting, e.g. *normal* or *UAV*, or even in a multiclass setting, e.g. outputting the type of UAV that generated the analyzed audio.

The next subsections further describes the built datasets and the performed experiments.

A. UAV Audio Datasets

In general, to properly evaluate ML-based techniques for audio detection and identification, it is necessary the provision of huge amounts of labeled data. Unfortunately, due to privacy issues, only few datasets are publicly available.

In light of this, our work makes use of Abdulla Al-Ali *et. al.* [1] dataset which provides more than 1300 audio clips of drone sounds. We used the dataset mainly for two tasks, one is detection and other is identification. In detection, the classes are Drone and Not a Drone. Number of drone samples are 1332 and not a drone samples are 10372. In classification task the classes are membo, bebop, and not a Drone, the number of samples are 666, 666, and 10372 respectively. Every file is of 1sec duration. The data set is split on file level basis. The drone data is a good representation of the real-world drone audio.

To increase the model performance, the audio dataset is also augmented through the introduction of noise data to ensure that the system will be able to detect and identify the drone's sound from similar noises in an environment. From [1], the SNR levels are not available. For the dataset publicly available, it was collected in a quite indoor environment with drone flying and hovering.

Apart from the publicly available dataset, we have also collected a new dataset with additional diverse flying devices audio sounds from multiple open sources to evaluate the impact of model identification performance metrics. The dataset was built, with 7 UAV types, including *no-UAV*, *drone*, *helicopter*, *drone-membo*, *drone-bebop*, *airplane*, and *drone-hovering*. The majority of audio files for each UAV type was collected for 5 minutes.

For both selected datasets the audio data is made with a sampling rate of 48 kHz and a linear encoding with 16 bits for sample. Each input sound window is further segmented into sub-frames of 20ms using a moving Hamming window with overlap of 10ms. The sub-frames are processed by a bank of filters to compute the short term feature in both temporal and frequency domains [20]. For the deep learning algorithms, If the majority of the frames is labelled with the tag "drone" into the given audio segment, it is assumed to recognize a flying drone in the surrounding environment.

Both datasets are split into 80% for training, 10% for validation and 10% for testing. The audio files are broken into chunks of 1 second.

B. Model Building

For the *Feature Extraction* module, for each audio frame, we compute the Mel-Frequency Cepstral Coefficients (MFCC) that are commonly used in audio analysis. We make use of Mel-scale in audio analysis to capture the comparatively higher energy in lower frequencies compared to higher frequencies in the range for the compounding of 40 MFCC features, which are used as input by the selected deep learning model (Fig. 1, *Filtered Feature Vector*).

We evaluate 5 commonly used deep learning algorithms for the audio-based classification task (Fig. 1, *Deep Learning*), namely *Deep Neural Network* (DNN), *Convolutional Neural Network* (CNN), *Long Short-Term Memory* (LSTM), *Convolutional Long Short-Term Memory* (CLSTM), and *Transformer Encoder* (TE).

The DNN was implemented with 3 hidden layers, each with 256 units, coped with a *relu* activation function, while the output layer relies in a *softmax* activation function. The CNN was implemented with 3 convolutional layers, each with a *relu* activation function, a kernel size of $3X3$, and a $1X1$ stride, followed by a hidden layer with a *softmax* activation function. There are no pooling layers used in the CNN we implemented, the input dimension size to the CNN is [1X20X33], given the small size there was no need to use the pool layers. The CNN model architecture is not provided in [1], due to this, we are not able to do a proper comparison in terms of architecture, we made our design choice to keep the model as smaller as possible.

The LSTM was implemented with one LSTM layer with 128 units, followed by a hidden output layer with a *softmax* activation function. Similarly, the CLSTM makes use of a convolutional layer with a *relu* activation function, a kernel size of $3X3$, and a $1X1$ stride, followed by a LSTM layer with 128 units, followed by a hidden output layer with a *softmax* activation function. Finally, the TE was implemented with a positional encoder, number of attention heads as 2, followed by a encoder layer with 128 units, and a output layer with a *softmax* activation function,

For the model building procedure, for all selected deep learning algorithms 100 epochs are executed with a batch size of 128. As for number of epochs, 100 is similar to the related works. It is important to note that the used set of parameters, were set similarly to related works, and no significant differences were found while varying them. There is 80% train data, 10% val data, 10% test data. In each epoch train and val data are used, model is trained on train data and

TABLE I
DETECTION RESULTS OF THE DEEP LEARNING ALGORITHMS AS MEASURED AT THE PUBLICLY AVAILABLE DATASET [1]

Deep Learning Algorithm	Accuracy (%)	Precision	Recall	F1-Score
Recurrent Neural Network (RNN [1]	75.00	0.7592	0.6801	0.6838
Convolutional Neural Network (CNN) [1]	96.38	0.9624	0.9560	0.9590
Convolutional Recurrent Neural Network(CRNN) [1]	94.72	0.9502	0.9308	0.9393
Deep Neural Network (DNN)	98.35	0.9661	0.9549	0.9604
Convolutional Neural Network (CNN)	98.85	0.9753	0.9696	0.9724
Long Short-Term Memory (LSTM)	**98.93**	**0.9759**	**0.9731**	**0.9745**
Convolutional Long Short-Term Memory (CLSTM)	97.78	0.9460	0.9486	0.9473
Transformer Encoder (TE)	98.35	0.9634	0.9489	0.9606

TABLE II
IDENTIFICATION RESULTS OF THE DEEP LEARNING ALGORITHMS AS MEASURED AT THE PUBLICLY AVAILABLE DATASET [1]

Deep Learning Algorithm	Accuracy (%)	Precision	Recall	F1-Score
Recurrent Neural Network (RNN [1]	57.16	0.5964	0.5716	0.5562
Neural Network (CNN) [1]	92.94	0.9275	0.9263	0.9263
Convolutional Recurrent Neural Network (CRNN) [1]	92.22	0.9254	0.9223	0.9225
Deep Neural Network (DNN)	98.52	**0.9589**	0.9439	0.9508
Convolutional Neural Network (CNN)	98.60	0.9553	0.9447	0.9510
Long Short-Term Memory (LSTM)	**98.60**	0.9480	**0.9603**	**0.9540**
Convolutional Long Short-Term Memory (CLSTM)	98.11	0.9457	0.9188	0.9314
Transformer Encoder (TE)	98.19	0.9405	0.9369	0.9386

best model on val data is updated. Finally the best val model is tested on the unseen test data.

C. Evaluation

The selected deep learning algorithms were evaluated with respect to their accuracy, precision, recall and F1 scores. To achieve such a goal, the following classification performance metrics were used:

- *True-Positive* (TP): number of UAV-related audio samples correctly classified as UAV-related.
- *True-Negative* (TN): number of normal samples correctly classified as normal.
- *False-Positive* (FP): number of normal samples incorrectly classified as UAV-related.
- *False-Negative* (FN): number of UAV-related audio samples samples incorrectly classified as normal.

The F1 score was computed as the harmonic mean of precision and recall values while considering UAV-related as positive samples and normal as negative samples, as shown in Eq. 3.

$$Precision = \frac{TP}{TP + FP} \tag{1}$$

$$Recall = \frac{TP}{TP + FN} \tag{2}$$

$$F1 = 2 * \frac{Precision * Recall}{Precision + Recall} \tag{3}$$

From the literature [1], there were no experiments conducted on the publicly available dataset with DNN, LSTM and Transformer techniques. We performed them for a proper comparison to asses the overall improvement for the detection and classification tasks.

TABLE III
INDIVIDUAL ACCURACIES IN OUR OWN BUILT DATASET.

UAV type	Accuracy (%)	Prec.	Rec.	F1-Score
no-UAV	85.71	1.00	0.86	0.92
drone	100	1.00	1.00	1.00
helicopter	66.66	0.67	0.67	0.67
drone-membo	100	1.00	1.00	1.00
drone-bebop	100	1.00	1.00	1.00
airplane	100	1.00	1.00	1.00
drone-hovering	100	0.88	1.00	0.93
OVERALL Dataset	95	0.96	0.95	0.95

The first experiment, aims at answering *RQ1* and evaluates the identification accuracy of the selected deep learning algorithms over the publicly available dataset [1]. The evaluation goal is to measure how the selected techniques perform when using publicly available dataset for audio-based detection of UAVs. In Table I shows, modeling results of [1] and our own benchmark results for drone detection task. All our models outperformed the models in [1] by 4% to more than 20% in accuracy. Our best model (LSTM) outperformed best model of [1] (CRNN) by over 4% in accuracy, 2% in precision, 4% in recall and 4% in F1-score. Our best detection model, LSTM classifier achieved the highest accuracy of 98.93%, precision of 0.9759, recall of 0.9731 with an F1-Score of 0.9745.

The second experiment, aims at answering *RQ2*, and evaluates the UAV identification accuracy of the selected deep learning algorithms when different UAV types are considered. To achieve such a goal, the publicly available dataset [1] is evaluated with different types of UAVs. In Table II shows, modeling results of [1] and our own benchmark results for drone identification task. All our models outperformed the models in [1] by 6% to more than 30% in accuracy. Our best model (LSTM) outperformed best model of [1] (CNN) by over 5% in accuracy, 2% in precision, 4% in recall and 3% in F1-score. Our best identification model, LSTM classifier achieved the highest accuracy of 98.60%, precision of 0.9480, recall of 0.9603 with an F1-Score of 0.9540. For instance, our identification LSTM model was able to achieve 98.60% of accuracy, a decrease of only 0.33% when compared to the detection scenario (Table I $vs.$ Table II).

The DNN model in table 3 is trained and tested on the newly collected 7 class UAV dataset. The third experiment aims at answering *RQ3*, and evaluates the identification accuracy in our own built dataset. Table III shows that the classification accuracy of the DNN classifier in our own built dataset. The DNN classifier achieved a high accuracy of 95.20%. Compared to its counterpart, trained in the publicly available dataset, the DNN decreased the accuracy by 3.32%, a marginal decrease considering that the classifiers is being used in a different and difficult setting.

V. CONCLUSIONS

In this work, we benchmarked the detection and identification of UAVs via audio, through multiple deep learning algorithms namely DNN, CNN, LSTM, CLSTM, and TE. In addition to dataset of [1], we also collected our own identification dataset and built DNN model over it. We have demonstrated that all of the DNN, CNN, LSTM, CLSTM and TE algorithms are able to provide significantly higher performance metrics in comparison to [1]. In detection task, our best model (LSTM) outperformed the best model of [1] (CRNN) by over 4% in accuracy, 2% in precision, 4% in recall and 4% in F1-score. In identification task, our best model (LSTM) outperformed the best model of [1] (CNN) by over 5% in accuracy, 2% in precision, 4% in recall and 3% in F1-score.

REFERENCES

[1] Abdulla Al-Ali et. al. Sara Al-Emadi, "Drone audio dataset," *https://github.com/saraalemadi /DroneAudioDataset*, 2018.

[2] Eleni Diamantidou Samaras, Stamatios and Antonios Lalas et al., "Deep learning on multi sensor data for counter uav applications—a systematic review.," *Sensors 19*, vol. 22, pp. 4387, 2019.

[3] Austin Reg, "Unmanned aircraft systems: Uavs design, development and deployment," *John Wiley & Sons*, pp. 1–365, 2011.

[4] Beomhui Jang Jangwon Jung Seo Yoojeong and Sung bin Im, "Uav detection using the cepstral feature with logistic regression," *Sensors*, pp. 219–2222, 2018.

[5] Changwen Zheng Peng Junkai and Si Lingyu et.al., "Using images rendered by pbrt to train faster r-cnn for uav detection," *Computer Science Research Notes CSRN 2802*, pp. 770–778, 2018.

[6] Faith Erden Martins Ezuma and Ismail Guvenc et.al., "Micro-uav detection and classification from rf fingerprints using machine learning techniques," *2019 IEEE Aerospace Conference*, pp. 1–13, 2019.

[7] Emmanuel Zenou Eren Unlu and Nicolas Riviere, "Using shape descriptors for uav detection," *Electronic Imaging*, vol. 1-5, 2018.

[8] Amin Ullah Jamil Sonain, MuhibUr Rahman and Seyed Sajad Mirjavadi et.al., "Malicious uav detection using integrated audio and visual features for public safety applications," *Sensors 20*, p. 3923, 2020.

[9] Liu Hao and Yunfang Ren et al., "Drone detection based on an audio-assisted camera array," *2017 IEEE Third International Conference on Multimedia Big Data(BigMM)*, p. 402–406, 2017.

[10] Erden F. Ezuma M. and I. et.al. Guvenc, "Micro-uav detection and classification from rf fingerprints using machine learning techniques," *2019 IEEE Aerospace Conference*, pp. 1–13, 2019.

[11] Justin Matheson Philip Church, Christopher Grebe and Brett Owens, "Aerial and surface security applications using lidar. in laser radar technology and applications," *International Society for Optics and Photonics*, vol. 10636, pp. 27–38, 2018.

[12] Chaoqun Yang Chang Xianyu and Zhiguo Shi. et. al., "A surveillance system for drone localization and tracking using acoustic arrays," *IEEE 10th Sensor Array and Multichannel Signal Processing Workshop (SAM)*, pp. 573–577, 2018.

[13] Ald rich Martinez-Carranza Cabrera-Ponce and Caleb Rascon, "Detection of nearby uavs using a multi-microphone array on board a uav," *International Journal of Micro Air Vehicles*, vol. 12, pp. 1–10, 2020.

[14] Abdulla Al-Ali Al-Emadi, Sara and Abdulaziz Al-Ali, "Audio-based drone detection and identification using deep learning techniques with dataset enhancement through generative adversarial networks," *Sensors 21*, p. 4953, 2021.

[15] Wansong Zhu Chen Liu and Minghui Zheng, "Audio-based fault diagnosis method for quadrotors using convolutional neural network and transfer learning," *2020 American Control Conference*, pp. 1367–1372, 2020.

[16] Sara Abdulla Al-Ali Amr Mohammad Al-Emadi and Abdul aziz Al-Ali, "Audio based drone detection and identification using deep learning," *15th International Wireless Communications and Mobile Computing Conference (IWCMC)*, pp. 459–464, 2019.

[17] Vemula Hari Charan., "Multiple drone detection and acoustic scene classification with deep learning," *PhD diss., Wright State University*, pp. 1–48, 2018.

[18] H. Truong P. Nguyen and T. Vu.Matthan et.al., "Drone presence detection by identifying physical signatures in the drone's rf communication," *in Proceedings of the 15th Annual International Conference on Mobile Systems, Applications, and Services, ACM*, vol. 5, pp. 211–224, 2017.

[19] J Shin S Jeon and H. Yang et.al., "Empirical study of drone sound detection in real-life environment with deep neural networks," *2017 25th European Signal Processing Conference (EUSIPCO)*, 2017.

[20] Emiliano Pallotti Bernardini Andrea, Federica Mangiatordi and Licia Capodiferro, "Drone detection by acoustic signature identification," *Electronic Imaging*, pp. 60–64, 2017.

[21] Kolamunna Harini and thilini et al., "Droneprint: Acoustic signatures for open-set drone detection and identification with online data," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, pp. 1–31, 2021.

[22] Darren Haddad Alexander Sedunov and Alexander Yakubovskiy et.al., "Stevens drone detection acoustic system and experiments in acoustics uav tracking," *2019 IEEE International Symposium on Technologies for Homeland Security (HST)*, pp. 1–7, 2019.

[23] C Park J Kim and JC Gallagher et.al., "Real-time uav sound detection and analysis system," *2017 IEEE Sensors Applications Symposium (SAS)*, pp. 1–5, 2017.

[24] Amr Mohammad Al-Emadi, Sara Abdulla Al-Ali and Abdulaziz Al-Ali, "Audio based drone detection and identification using deep learning," *2019 15th International Wireless Communications and Mobile Computing Conference (IWCMC)*, pp. 459–464, 2019.