EasyChair Preprint

№ 9335

Deep Learning Based Malicious Drone Detection Using Acoustic and Image Data

Juann Kim, Dongwhan Lee, Youngseo Kim, Heeyeon Shin, Yeeun Heo, Yaqin Wang and Eric T. Matson

November 18, 2022

Deep Learning Based Malicious Drone Detection Using Acoustic and Image Data

Juann Kim
Dept. Software
Sangmyung University
Cheonan, Republic of Korea
201920951@sangmyung.kr

Dongwhan Lee
Software Convergence
Kyung Hee University
Yongin, Republic of Korea
derick_lee@khu.ac.kr

Youngseo Kim
Dept. Human Centered AI
Sangmyung University
Seoul, Republic of Korea
201910787@sangmyung.kr

Heeyeon Shin
Computer Engineering
Kyung Hee University
Yongin, Republic of Korea
567didi@khu.ac.kr

Yeeun Heo
Software Engineering
Soongsil University
Seoul, Republic of Korea
gjdpdms2005@soongsil.ac.kr

Yaqin Wang*
Computer and Information Technology
Purdue University
West Lafayette, United States
wang4070@purdue.edu

Eric T. Matson*
Computer and Information Technology
Purdue University
West Lafayette, United States
ematson@purdue.edu

***Abstract*—Autonomous drones have been studied in a variety of industries including delivery services and disaster protection. As the supply of low-cost drones has been increasing, a CUAS (Counter unmanned aerial systems) is critical to manage autonomous drone traffic control and prevent drone flights in secured areas. For these systems, drone detection is one of the most important steps in the overall process. The goal of this paper is to detect a drone using the microphone and the camera by training deep learning models based on image and acoustic features. For evaluations, three methods are used: visual-based, audio-based, and the decision fusion of both features. The decision fusion of audio and vision-based features is used to obtain higher performance on drone-to-drone detection. Image and audio data were collected from the detecting drone, by flying two drones in the sky at a fixed Euclidean distance of 20m. In addition, deep learning methods are applied to investigate an optimal performance. CNN (Convolutional Neural Network) was used for acoustic data, and YOLOv5 was used for computer vision. From the result, the decision fusion of audio and vision-based features showed the highest accuracy among the three evaluation methods.**

***Index Terms*—drone detection, audio classification, computer vision, counter unmanned aerial systems, deep learning**

I. Introduction

Recently, the demand for drones has been increasing significantly. With the growing number of drones, the importance of small and low-cost drones has considerably been expanded. The benefits of drones are enormous: operating without a pilot, applying diverse fields, no high-cost infrastructure, etc. As reported by global market research publishing and management consulting firm *Grand View Research*, the size of the global commercial drone market is expected to flourish to a CAGR (compound annual growth rate) of 57.5% from 2021 to 2028 [1].

As many drones are commercialized, these are also used for malicious reasons. Drones have been used to attack and invade, such as an assassination attempt on the president of Venezuela in 2018 [2]. To prevent similar events from occurring, CUAS has been further developed. In CUAS, it is critical to detect, track, and eventually destroy malicious drones [3]. Hence, it is needed to detect drones in order to solve problems of detecting, tracking, and removing malicious drones, etc.

Researchers have shown that cameras or microphones, placed on the ground, were handled for malicious drones [4]. However, little is researched based on both vision and acoustic. Hence, an unprecedented drone detection system using the two features is introduced. The proposed system detects a flying drone in the air via image and acoustic data. It is conducted in the moving area, expanding the CUAS invasion system between moving drones with a fusion of two methodologies.

Overall, the main contributions of this work can be summarized as follows:

- The high quality of drone-to-drone audio and image data were collected by distance of 20 to 100 meters manually.
- This paper proposes a novel drone detection scheme that reduces the error rate, which using decision fusion.

II. Related Work

A. CUAS (Counter Unmanned Aerial Systems)

In recent years, drone flights in the AEZ (Air Exclusion Zone) have repeatedly occurred. In 2015, a man was detained, since he flew his drone 100 feet above Lafayette Park near the White House [5]. Hence, CUAS has been conducted to prevent these occurrences. To illustrate, the largest air show in Europe, Airpower 22, successfully provided two eight-hour performances to 275,000 spectators with AARTOS, which is the world's best anti-drone system [6]. CUAS is a system to detect and track drones that approach protected or secure areas. Following [7], researchers present a study on a shooting system using Class 1 drones, defined as small and transportable, with a human-in-loop, an autonomous and vision-based system. If a target drone is continuously captured in several

*Corresponding authors.

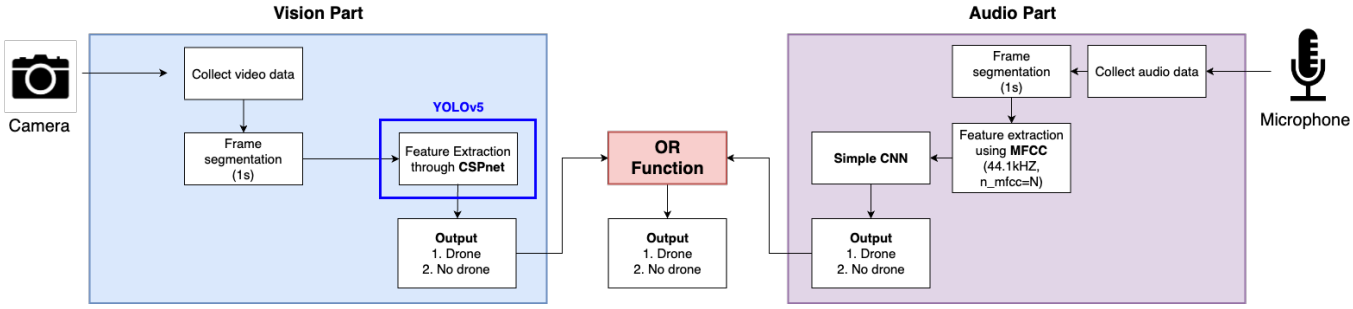

Fig. 1. Overview of the drone detection system

frames, a drone pilot changes the mode to autonomous, and the detecting drone approaches the target drone. A complete procedure is to move the detecting drone towards the target drone and put the target drone into an inoperable status. This research demonstrates that the proposed system depends on a pilot to be practiced or skilled.

B. Drone Detection using Sensors

1) Radar: Various methods have been used for drone detection, including Radar, LiDAR, Computer vision, and Acoustic sensor. Each method has its own strengths and limitations when detecting drones. Most commercial products that are widely utilized for drone detection are based on radar. More specifically, radar is used for binary drone and multi-drone detection. However, radar is not optimized for detecting drone made of plastic material or small drones at widely varying ranges [8]. Furthermore, a lot of false positives are recorded as it is difficult to distinguish the difference between drones and other flying objects such as birds [9]. On the other hand, in this paper, acoustic and vision-based features, which are not significantly affected by the size or the materials of drones, are used for drone detection.

2) Camera: In the past few decades, studies related to drone detection using computer vision have already been conducted [10][11]. Vision-based object detection method is accurate enough to classify the binary classes (drone, no drone) and can localize the actual location. Moreover, a single camera, a small, self-contained, and portable device, is even more accessible to perform detection tasks rather than LiDAR or Radar. *Craye* [10] and *Ulzhalgas* [11] used a single camera to detect drones on the ground, using a computer vision method. Each author generated an accuracy of 73.5%, and 74.2% respectively on detecting objects from particular frames of image data based on CNN.

3) Microphone: Using a iPhone as a microphone is easy to utilize and more affordable for drone detection than other sensors, such as radar, LiDAR, etc. The author in [4] presented a drone detection system using multiple acoustic nodes along with machine learning models. This system evidences that the models are able to recognize the acoustic signals in a wider range under 3-dimensional spaces. By using the low-priced multiple acoustic nodes, this technique can detect drones up to 75m. Additionally, the paper above deduces that based on audio features, a method using deep learning has higher performance than that of machine learning.

4) Fusion: As previously stated, diverse sensors can be employed to detect drones. As proposed by several papers, a fusion technique of more than two domains critically influenced academic dialogues on improving the performances of drone detection. The combination of radar and audio sensors, the suggested method in [12], can detect and track rotor types of drones. An electro-optical and acoustic-based fusion system were deployed to detect, localize, and track drones [13]. In [14], vision information and IMU (inertial measurement unit) were collected using a monocular camera. Collected data processed by end-to-end deep neural network architecture with feature fusion resulted in less than error of 3%. Also, [15] achieved an accuracy of 75% that shows the feasibility of a sensor fusion (RF data and image data) based technique for drone detection.

III. Methodology

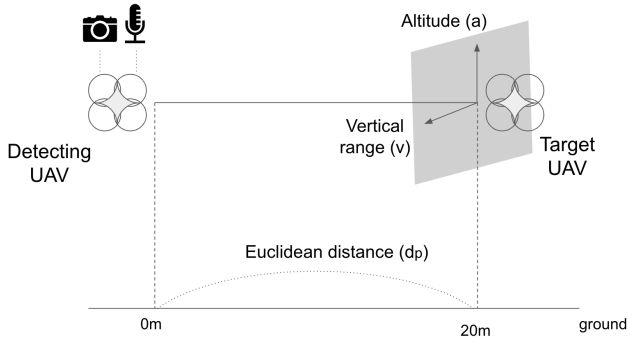

Fig. 2. A view of two drones while flying at the same time

A. System Overview

Drone data was sectioned into two classes: drone and no drone. Drone classification was performed to identify the existence of drones. A required processing flow can be segmented into five parts displayed in Fig 1.

The proposed system has a single camera and microphone. The data was collected using a camera of a drone. iPhone 6 was also used as an microphone to record acoustic data.

The conducted experiments were assigned to two drones to encounter each other in the air. Acoustic and image data were collected by the detecting drone while hovering. The target drone moved with a fixed euclidean distance as Fig 2.

Fig. 3. iPhone 6 attached to DJI Mavic 2 Pro

B. Data Collection

1) Environment setting: DJI Matrice 200 V2 and *DJI Mavic 2 Pro* were used to collect acoustic and image data. Acoustic data were collected using *iPhone 6* attached to *Mavic 2 Pro* as shown in Fig 3, and image files were collected using the built-in camera, while both drones were flying in the air. *DJI Mavic 2 Pro*, which has a built-in camera and a microphone, is defined as a *Detecting Drone*, and *DJI Matrice 200 V2* is defined as a *Target Drone*. This paper analyzes not only the hovering data but also the data with the target drone moving within the screen of the detecting drone.

The acoustic signal was recorded in a "wav" format, and the sampling rate is 44.1kHz. The noises contain the sounds of insects, airplanes, human voices, animals including birds, and ground vehicles such as tractors and trucks. For the vision part, the image data were recorded in the same environment with different weather conditions.

C. Feature extraction

1) Audio: In [16], splitting acoustic data into one second showed the highest performance than other time intervals when training the deep learning model. Therefore, the acoustic data are split into one second for audio segmentation [17]. Pitch shifting is used for data augmentation. Pitch shifting is a method to raise or lower the pitch of the audio sample without affecting the speed of the sound. In [18], pitch shifting augmentation showed the greatest positive impact on performance and is the only one that does not have any negative impact on any classes of environmental sound classification. In order to classify the drone sound, MFCC (Mel Frequency Cepstral Coefficients) is used, which is a non-linear mapping of the original frequency according to the auditory mechanism of the human ear [8].

Furthermore, MFCC is widely used for audio classification and successfully used with machine learning [19][20] and deep learning approaches [21][22]. Also, it offers useful features for capturing periodicity from the fundamental frequencies brought on by the rotor blades of a drone [23]. In the acoustic analysis, the MFCC features are extracted using Librosa[24] from the Python package. In [21][23], an experiment was conducted with various numbers of MFCC such as 20 or 40, by extracting various features. In this paper, results are extracted with four different sizes of MFCC features. These values include 20, 40, 80, and 120 sizes to obtain diverse results. Each data for one second is composed of 44xN size.

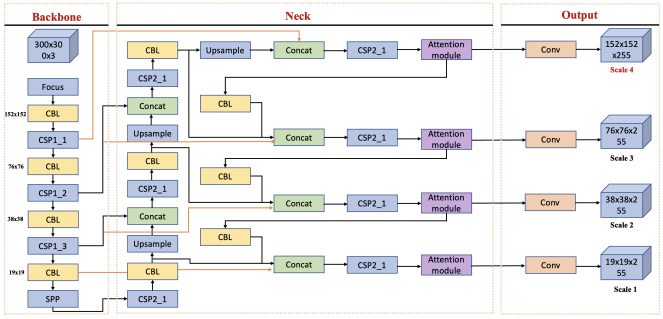

Fig. 4. YOLOv5 Architecture

2) Vision: In order to train the model for drone detection, all the ground truth objects in the images need to be labeled first. This dataset is labeled using the "LabelImg" [25], which is an open source tool.

To extract image features, CBL (Convolution with Batch normalization and Leaky ReLU), SPP (Spatial Pyramid Pooling), and CSP (Cross Stage Partial) were used in the backbone layer of YOLOv5 [26]. The backbone network extracts feature maps of various sizes from input images through the convolution layer and pooling layer. The overall structure is shown in Fig 4. First, CBL is a block that is essentially used to extract features consisting of the convolution layer, batch normalization, and leaky ReLU. SPP improves performance by pooling various sizes of feature maps with filters and then merging them again. The CSP divides the feature map of the base layer into two parts to reduce the heavy inference computations caused by duplicate gradient information. Then, it was combined again in the cross-stage hierarchy method proposed in [27]. This way, the extended gradient information can have a large correlation difference by switching the concatenation and transformation steps. Furthermore, CSP can significantly reduce computational effort and improve inference speed and accuracy.

5 Backbone networks - YOLOv5-n,s,m,l,x are used. Each model is distinguished by *depth_multiple* and *width_multiple*. The larger the *depth_multiple* value, the more BottleneckCSP() is repeated to become a deeper model. The larger the *width_multiple*, the higher Convolution filter number of the corresponding layer.

D. Deep learning models

1) Audio: Among various classifiers, CNN is used as shown high performance in audio signal classification with spectral features such as MFCC [28]. The architecture of our CNN model is shown in Table I.

The learning rate set 0.0001 with TensorFlow Keras optimizer, Adam. In addition, Early Stopping is used for preventing the model from being overfitted. The two activation

functions, sigmoid and softmax, are used for evaluating final performance. The activation function of Sigmoid shows the best performance.

TABLE I
THE CNN MODEL SUMMARY WITH 44X80 INPUT SIZE

Layer Type	Output Shape	Parameters
Conv2D	(None, 44, 80, 32)	832
Conv2D	(None, 44, 80, 32)	25632
MaxPooling2D	(None, 22, 40, 32)	0
Dropout	(None, 22, 40, 32)	0
Conv2D	(None, 22, 40, 64)	18496
Conv2D	(None, 22, 40, 64)	36928
MaxPooling2D	(None, 11, 20, 64)	0
Dropout	(None, 11, 20, 64)	0
Flatten	(None, 14080)	0
Dense	(None, 256)	3604736
Dropout	(None, 256)	0
Dense	(None, 2)	514

2) Vision: Image classification generally refers to images in which only one object is visible and analyzed. In contrast, object detection includes classification and localization tasks for analyzing more realistic situations where multiple objects may be present in an image [29].

The object detection model can be divided into two main types: one-stage model and two-stage model. Compared to other two-stage object detection models including R-CNN[30] and Faster R-CNN [31], one-stage models such as YOLO [32] can calculate fast enough to conduct real-time object detection tasks. Therefore, YOLOv5 is selected as the appropriate model for drone detection in this research. Since object detection for CUAS have to be implemented in real-time.

The YOLOv5 model can be represented by YOLOv5-n,s,m,l,x depending on the capacity of the model and the number of parameters. Models with large capacities such as YOLOv5x can increase accuracy but have a slow operation. Conversely, a lightweight model such as YOLOv5n is fast, but it can not get outstanding performance in accuracy.

IV. EXPERIMENT

TABLE II
THE NUMBER OF ENTIRE DATASET

Type of data	Class	Audio	Image	Augmented	Total times (s)
Train	drone	1055	1055	1055	4220
	no drone	1055	1055	1055	
Validation	drone	300	300	300	1200
	no drone	300	300	300	
Test	drone	154	154	-	308
	no drone	154	154	-	

A. Audio

1) setup: In total, there are 4220 samples for training, 1200 samples for validation, and 308 samples for testing. For CNN model, the input data size is 44xN which is the same as each size of MFCC features. The CNN model is trained with a different number of MFCC features and activation functions. Fig 5. shows the result of confusion matrix.

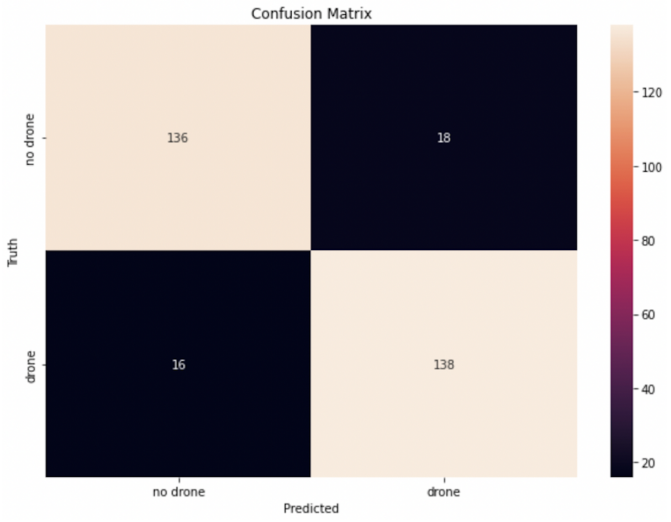

Fig. 5. Result of Confusion Matrix of Acoustic data

2) training & testing: The overall results are described in Table III with the different activation functions and the number of MFCC features. From the results, the model with the highest performance showed an accuracy of 88.9%. Also, the model was well trained without being overfitted by using early stopping technique.

TABLE III
RESULTS FOR CNN MODELS IN ACCURACY

Activation	n_mfcc 20	n_mfcc 40	n_mfcc 80	n_mfcc 120
Softmax	0.866	0.886	0.886	0.866
Sigmoid	0.870	0.873	0.879	0.889

B. Vision

1) setup: The training and validation dataset for the vision task includes 5420 images, and the dataset for testing includes 308 images, totally equaled with the number of acoustic data. In addition, data augmentation was introduced to prevent overfitting. Three types of methods, horizontal flip, noising, and blur, were mixed. Input images are fixed with the size of 640 x 640 demanded by YOLOv5. For the hyperparameters, the batch size is 16, and epochs are 50. SGD (Stochastic Gradient Descent) is used for YOLOv5 as an optimizer. In this paper, the comparison of experiments with 5 different YOLOv5 models figures out which model is appropriate for the drone-to-drone detection tasks.

2) training & testing: In order to compare five different YOLOv5 models, all experiment environments were totally set to be same, shown to Table IV and Table V. Comparing among five models with Model inference time, the YOLOv5n model showed the best performance.

TABLE IV
TRAINING RESULTS FOR YOLOV5 MODELS IN MAP, PRECISION, RECALL, F1-SCORE

Models	mAP_0.5	mAP_0.5:0.95	Precision	Recall	F1-score
YOLOv5n	0.840	0.390	0.780	0.861	0.82
YOLOv5s	0.870	0.377	0.763	0.870	0.81
YOLOv5m	0.860	0.378	0.806	0.944	0.81
YOLOv5l	0.821	0.358	0.784	0.991	0.79
YOLOv5x	0.851	0.372	0.835	0.991	0.80

TABLE V
TESTING RESULTS FOR YOLOV5 MODELS IN MAP, PRECISION, RECALL, F1-SCORE

Models	mAP_0.5	mAP_0.5:0.95	Precision	Recall	F1-score
YOLOv5n	0.904	0.574	0.940	0.696	0.82
YOLOv5s	0.922	0.583	0.855	0.824	0.81
YOLOv5m	0.904	0.570	0.793	0.809	0.81
YOLOv5l	0.822	0.602	0.694	0.971	0.79
YOLOv5x	0.902	0.636	0.669	0.926	0.80

C. Result

Therefore, the decision fusion of vision and acoustic features is considered the stronger drone detection method that can compensate for each other's shortcomings. Also, this method shows higher accuracy compared to the results of each previous experiment.

For the decision fusion, "OR" operation is used to fuse the final result of CNN and YOLOv5, which is referred to [33]. Before processing the fusion, the accuracy for vision is 90.26% and the accuracy for acoustic is 88.96%. However, after applying the OR function to the result of vision and acoustic, the accuracy is improved to 92.53%. The result can be also described as a graph shown in Fig 6.

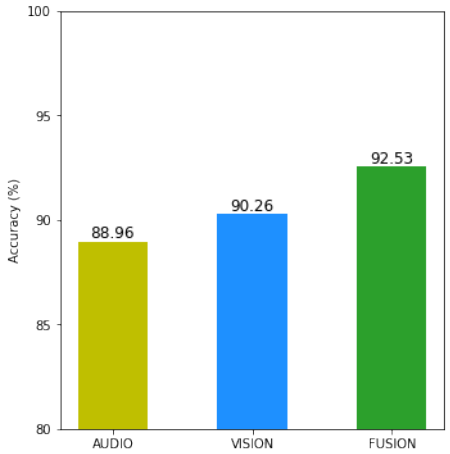

Fig. 6. The graph showing the result of OR function

V. CONCLUSION AND FUTURE WORK

A. Conclusion

Utilizing both the acoustic and vision-based features enabled the drone detection system that is indispensable for CUAS. Two drones were flying at the same time for drone detection, rather than recording from the ground. The result shows that it is possible to classify the target drone sound regardless of the noise of the detecting drone.

B. Limitation

The proposed research has a limitation of hovering a detecting drone at the same location rather than moving in the air. In addition, the microphone of the iPhone 6 for audio recording had a lower performance than that of the latest iPhones, or other professional microphones. Nevertheless, this is also a disproof that the drone-to-drone detection task could still show sufficient results only by using the acoustic device with low performance.

C. Future work

It is necessary to develop a complete real-time method connected to one end-to-end process detection system through further research. Moreover, drone detection, using different numbers of drones by changing the movement of drones, needs to be researched to make future research applicable in a real environment or general situations.

Indeed, the range of availability of the drone-to-drone task needs to be specified. Expanding the experiments based on the paper, a wider range of Euclidean distance, rather than 20m, would be conducted to collect target drone data. Eventually, finding the maximum Euclidean distance range available for drone detection can be researched.

REFERENCES

[1] Grand View Research, "Commercial Drone Market Size, Share Trends Analysis Report by Product, by Application, by End-use, by Region, and Segment Forecasts, 2021-2028", Apr. 26, 2021. [Online]. Available: https://www.grandviewresearch.com/Filters-search=Commercial-Drone Marketsearch

[2] Science and Technology. "Counter-Unmanned Aircraft Systems (C-UAS)", Jul.22, 2022. [Online]. Avaliable: https://www.dhs.gov/science-and-technology/counter-unmanned-aircraft-systems-c-uas.

[3] X. GUAN et al., "A survey of safety separation management and collision avoidance approaches of civil UAS operating in integration national airspace system." *Chinese Journal of Aeronautics*, Apr.27, 2020

[4] Yang, Bowon, et al. "UAV detection system with multiple acoustic nodes using machine learning models." *2019 Third IEEE International Conference on Robotic Computing (IRC)*, 2019.

[5] H. Abdullah, "Man Detained for Flying Drone Near White House". NEWS, May. 15, 2015. [Online]. Available: https://www.nbcnews.com/news/us-20news/20man-20detained-20trying-20fly-20drone-20near-20white-20house-20n359011

[6] Press, "AARTOS protected the largest air show in Europe from illegal drones" sUAS News, Oct.11, 2022. [Online]. Available: https://www.suasnews.com/2022/10/aartos-protected-the-largest-air-show-in-europe-from-illegal-drones/

[7] A. R. Wagoner, D. K. Schrader and E. T. Matson, "Towards a vision-based targeting system for counter unmanned aerial systems (CUAS)," *2017 IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA)*, 2017.

[8] Liu, Hao, et al. "Drone detection based on an audio-assisted camera array." *2017 IEEE Third International Conference on Multimedia Big Data (BigMM)*, 2017.

[9] B. Taha and A. Shoufan, "Machine Learning-Based Drone Detection and Classification: State-of-the-Art in Research," *IEEE Access*, vol. 7, pp. 138669-138682, 2019.

[10] Craye, Celine, and Salem Ardjoune. "Spatio-temporal semantic segmentation for drone detection." *2019 16th IEEE International conference on advanced video and signal based surveillance (AVSS)*, 2019.

[11] U. Seidaliyeva, D. Akhmetov, L. Ilipbayeva, and E. T. Matson, "Real-Time and Accurate Drone Detection in a Video with a Static Background," *Sensors*, vol. 20, no. 14, p. 3856, Jul. 2020.

[12] S. Park et al., "Combination of radar and audio sensors for identification of rotor-type Unmanned Aerial Vehicles (UAVs)," *2015 IEEE SENSORS*, pp. 1-4, 2015.

[13] Christnacher, Frank, et al., "Optical and acoustical UAV detection." *Electro-Optical Remote Sensing X*. vol. 9988, SPIE, 2016.

[14] Zhang, Xupei, et al., "VIAE-Net: An End-to-End Altitude Estimation through Monocular Vision and Inertial Feature Fusion Neural Networks for UAV Autonomous Landing." *Sensors*, 2021

[15] Aledhari, Mohammed, et al. "Sensor Fusion for drone detection." 2021 IEEE 93rd Vehicular Technology Conference (VTC2021-Spring). IEEE, 2021.

[16] S. Al-Emadi, A. Al-Ali, A. Mohammad and A. Al-Ali, "Audio Based Drone Detection and Identification using Deep Learning," *2019 15th International Wireless Communications Mobile Computing Conference (IWCMC)*, pp. 459-464, 2019.

[17] P. Casabianca and Y. Zhang, "Acoustic-Based UAV Detection Using Late Fusion of Deep Neural Networks," Drones, vol. 5, no. 3, p. 54, Jun. 2021, doi: 10.3390/drones5030054.

[18] J. Salamon and J. P. Bello, "Deep Convolutional Neural Networks and Data Augmentation for Environmental Sound Classification," in IEEE Signal Processing Letters, vol. 24, no. 3, pp. 279-283, March 2017

[19] R. Serizel, V. Bisot, S. Essid, and G. Richard, "Acoustic features for environmental sound analysis," *Computational Analysis of Sound Scenes and Events"*, pp. 71–101, Springer, Cham.

[20] Y. Wang, F. E. Fagian, K. E. Ho and E. T. Matson, "A Feature Engineering Focused System for Acoustic UAV Detection," 2021 Fifth IEEE International Conference on Robotic Computing (IRC), 2021, pp. 125-130

[21] S. Jeon, J. W. Shin, Y. J. Lee, W. H. Kim, Y. Kwon and H. Y. Yang, "Empirical study of drone sound detection in real-life environment with deep neural networks," 2017 25th European Signal Processing Conference (EUSIPCO), 2017.

[22] Q. Dong, Y. Liu, X. Liu. "Drone sound detection system based on feature result-level fusion using deep learning" "Multimedia Tools and Applications", 2022, doi:10.1107/s11042-022-12964-3

[23] S. Seo, S. Yeo, H. Han, Y. Ko, K. E. Ho and E. T. Matson, "Single Node Detection on Direction of Approach," *2020 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*, pp. 1-6, 2020.

[24] McFee, Brian, Colin Raffel, Dawen Liang, Daniel PW Ellis, Matt McVicar, Eric Battenberg, and Oriol Nieto. "librosa: Audio and music signal analysis in python." In Proceedings of the 14th python in science conference, pp. 18-25. 2015.

[25] heartexlabs, "labelImg", github.com https://github.com/heartexlabs/labelImg

[26] Ultralytics, "yolov5", github.com https://github.com/ultralytics/yolov5 (accessed Aug. 1, 2022)

[27] Wang, Chien-Yao, et al., "CSPNet: A new backbone that can enhance learning capability of CNN." *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*, 2020.

[28] Li, Shulin, et al., "Convolutional Neural Networks for Analyzing Unmanned Aerial Vehicles Sound." *2018 18th International Conference on Control, Automation and Systems (ICCAS)*, 2018, pp. 862-866.

[29] T. Wu, T. Wang and Y. Liu, "Real-Time Vehicle and Distance Detection Based on Improved Yolo v5 Network," *2021 3rd World Symposium on Artificial Intelligence (WSAI)*, 2021, pp. 24-28, doi: 10.1109/WSAI51899.2021.9486316.

[30] Girshick, Ross, et al. "Rich feature hierarchies for accurate object detection and semantic segmentation." *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014.

[31] Ren, Shaoqing, et al. "Faster r-cnn: Towards real-time object detection with region proposal networks." *Advances in neural information processing systems 28*, 2015.

[32] Redmon, Joseph, et al. "You only look once: Unified, real-time object detection." *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016.

[33] X. Shi et al., "Anti-drone system with multiple surveillance technologies," IEEE Communications Magazine, vol. 56, no. 4, 2018.