Performance Analysis of YOLOv7 and YOLOv8 Models for Drone Detection

[1] Khushi Agarwal
Department Computer Science
CHRIST UNIVERSTIY
Bangalore, Karnataka, India
khushi.agarwal@science.christuniversity.in

[2] Ashik Sanyo M S
Department Computer Science
CHRIST UNIVERSTIY
Bangalore, Karnataka, India
ashik.ms@science.christuniversity.in

[3] Srishti Bakshi
Department Computer Science
CHRIST UNIVERSTIY
Bangalore, Karnataka, India
srishti.bakshi@science.christuniversity.in

[4] Vinay M
Department Computer Science
CHRIST UNIVERSTIY
Bangalore, Karnataka, India
vinay.m@christuniversity.in

[5] Jayapriya J
Department Computer Science
CHRIST UNIVERSTIY
Bangalore, Karnataka, India
jayapriya.j@christuniversity.in

[6] Deepa S
Department Computer Science
CHRIST UNIVERSTIY
Bangalore, Karnataka, India
deepa.s@christuniversity.in

Abstract- **Drone detection techniques are used to detect unmanned aerial systems (UAS) also commonly known as drones. A rapid increase in these drones has limited the airspace safety and so the research for drone detection has emerged. This study compares between the two widely used deep-learning models, previously used YOLOv7 and the latest YOLOv8. The overall finding of this study suggests that the YOLOv8 deep-learning model appears to be more promising and may make valuable contributions on their own. We got the result that for 10 epochs YOLOv8 gave 50.16% accuracy while YOLOv7 gave 48.16% accuracy making YOLOv8 more promising for the task. As a practical application for future work, we intend to deploy YOLOv8 on edge devices to achieve real-time drone detection in critical security applications.**

Keywords- Yolo (You Only Look Once), UAS (Unmanned Aerial System), Drone Detection, unmanned **aerial vehicles (UAVs),** ***Computer Vision, Yolov7, Yolov8, CNN (convoluted neural network)***

I. INTRODUCTION

Over the years, drone technology has accelerated quickly, allowing the widespread use of unmanned aerial vehicles for a number of factors including commercial, military, and civil. Drones have abundant benefits such as low cost and great mobility, however their unchecked usage can pose a significant threat to public as well as private safety. The need for reliable drone detection systems has grown above all other requirements. The process of finding and following UAVs in a specific area is known as drone detection. Drone detection's main objective is to stop unauthorized drone activities and safeguard infrastructure and people from harm. Acoustic sensors, radio frequency (RF) scanners, and visual systems based on computer vision techniques are just a few of the drone detection approaches that have been developed. One such computer vision method is the cutting-edge object detection algorithm YOLO (You Only Look Once), which is based on convolutional neural networks (CNNs). YOLO is a popular option for drone detection because of its accuracy, speed, and real-time performance. The most recent iterations of the YOLO algorithm, YOLOv7 and YOLOv8, have demonstrated encouraging results in object detection tasks.

Other drone detection approaches are also given significant attention in the context of this research article. For instance, in supplemented datasets used in 3D LADAR systems, the V-RBNN (Variable Radially Bounded Nearest Neighbour) technique has demonstrated promise for tiny drone identification. When compared to the traditional RBNN approach, V-RBNN shows a significant improvement of about two times, with an accuracy of between 0.6 and 0.7. Drone Detection and Pose Estimation is a noteworthy additional method. Using Parrot Bebop-2 quadrotor keypoints as the dataset, Relational Graph Networks uses a combination of a deep learning model (Xception-like model) and RGNs. Although accuracy was achieved, there is still room for improvement, especially with regard to 6D pose estimate accuracy for small targets. Future iterations are intended to improve the findings by utilizing adversarial training techniques in order to overcome this constraint.Additionally, automated drone recognition utilizing YOLOv4, which is based on the MS COCO dataset of Microsoft Common Objects in Context, exhibits astounding accuracy with a stunning 89.32%. This accomplishment outperforms the results of its forerunner, YOLOv3, by 5.18%.

The study also offers a high-resolution background difference-based drone detection technique that uses SAG-YOLOv5s, a YOLOv5 model version. This method achieves a remarkable mean Average Precision (mAP) of 97.6% using a dataset of 12,188 drone targets and an average of 1.08 targets per image.In this article, we analyze the performance of YOLOv7 with YOLOv8 for drone detection using a dataset of UAV images obtained in various different environments. We also examine the two deep-learning model's accuracy, responsiveness as well as their pertinence to actual drone detection. Our research seeks to explain the advantages and the drawbacks of YOLO-based drone

979-8-3503-0082-6/23/$31.00 ©2023 IEEE

2023 International Conference on Network, Multimedia and Information Technology (NMITCON) | 979-8-3503-0082-6/23/$31.00 ©2023 IEEE | DOI: 10.1109/NMITCON58196.2023.10276343

detection systems and to provide a route for the development of the future.

II. LITERATURE REVIEW

The article "Machine Learning-Based Drone Detection and Classification: State-of-the-Art in Research" discusses about various techniques for drone detection and classification, such as the use of 3-dimensional Histograms of Gradients (HoG3D) and a CNN model for image analysis, and Gaussian Mixture Model (GMM), CNN, and RNN classification for sound data analysis. The study's accuracy rate for identifying drone sounds was 96.4%[1]. In order to detect objects on drone photos, the paper "VisDrone-DET2018: The Vision Meets Drone Object Detection in Image Challenge Results" compares two-stage and one-stage CNN-based algorithms. The top detector scored 31.88% AP, which is impressive[2].A study titled "Using Deep Networks for Drone Detection" assesses the effectiveness of convolutional neural networks (CNN) as a stand-alone method for drone detection, obtaining an accuracy of roughly 0.9 [3]. A pre-trained YOLOv5s model by Ultralytics is used in "Platooning control of drones with real-time deep learning object detection" to enable successful drone platooning in an indoor setting [4]. "Object Detection-Based System for Traffic Signs on Drone-Captured Images" detects traffic signs on drone-captured photos using CNN-based object detectors, leading to a 55% increase in safety[5].“Special Vehicle Detection from UAV Perspective via YOLO-GNS Based Deep Learning Network" improves special vehicle detection accuracy by 4.4% using the SEVE dataset and the YOLOv7 algorithm[6].Different neural network architectures are used in "Deep learning approaches to building rooftop thermal bridge detection from aerial images" to detect thermal bridges on building rooftops with an accuracy of 87% [7]. "CoDerainNet: Collaborative Deraining Network for Drone-View Object Detection in Rainy Weather Conditions" investigates a semi-supervised learning framework for deraining images gathered in actual rainy scenarios, achieving 1.73% and 1.12% detection errors using YOLOv5 and YOLOv7 on synthetic drone-captured datasets[8].”DroMOD: A Drone-Based Multi-Scope Object Detection System" examines the system's accuracy, real-time processing, and resource efficiency, attaining an accuracy of 93%. The system leverages YOLOv3 for object detection on a multiple eye-sky dataset[9]. In "Drone Detection Using YOLOv5", a collection of drone and bird photos is created, and the accuracy of three different object detectors is examined. The CNN model achieves the highest accuracy of 93% in this study[10].Phat Thai, Alam, Nimrod Lilith, Binh Nguyen, and Ho Chi Min's paper "Small Flying Object Detection and Tracking in Digital Airport Tower through Spatial-Temporal ConvNets" employs the R-CNN model for detecting small flying objects in airport towers. A notable 99.69% Average Precision (AP) score on the Leesburg airport dataset highlights its advancement in aerial item detection [11]. "Automated Drone Detection using YOLOv4," by Subhobroto Singha and Buchanan Aydin, reveals a comprehensive drone detection methodology. They merge drone and bird image datasets with the YOLOv4 model, yielding exceptional metrics: a mean Average Precision (mAP) of 74.36%, precision at 0.95, recall of 0.68, and F1-score of 0.79. Notably, their technique achieves noteworthy frame rates, hitting 20.5 frames per second for DJI Phantom III and 19.0 frames per second for DJI Mavic Pro. The study strongly underscores YOLOv4's drone detection competence, promising precision and real-time processing [12].

III. METHODOLOGY

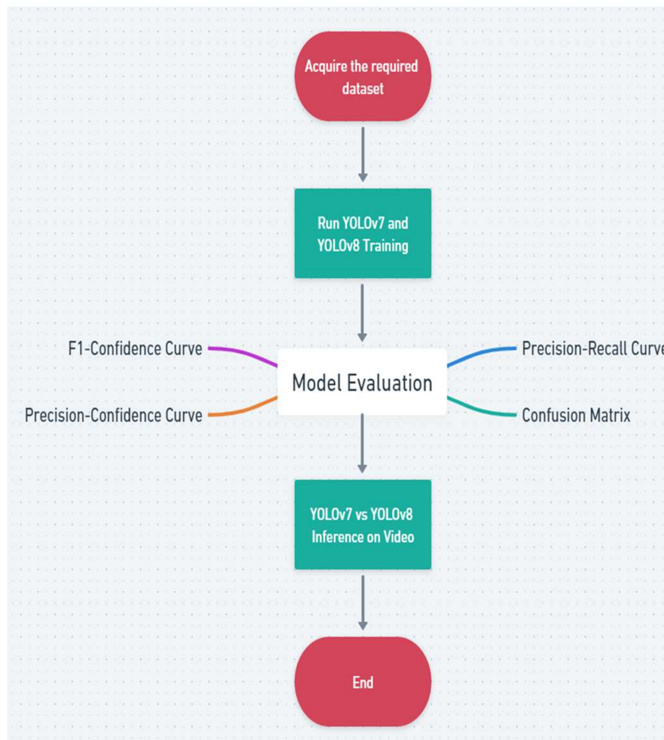

Fig. 1. Process Steps

Here are the process steps given in Fig.1. for the study on drone image detection using YOLOv7 and YOLOv8 models:

A) Dataset Description

The YOLO Drone Detection Dataset, which is a collection of pictures and annotations of drones shot in diverse outdoor settings and taken from RoboFlow, is the dataset that was used. This dataset includes approximately 3.5k training images, 1k validation images and 510 testing images, each of which includes a bounding box indicating where the drone is located in the image.The dataset was made in order to recognise and track drones using the real-time object detection technique YOLO (You Only Look Once), which utilizes convolutional neural networks (CNNs). Due to its quick detection time and great accuracy, YOLO is a well-liked object identification technique.

B) Run YOLOv7 and YOLOv8 Training

In this study, we used the recent versions of the YOLO algorithm: YOLOv7 and YOLOv8. A powerful and quick object identification method, YOLOv7, finds things in real time. Another model is the YOLOv8, an appealing option for a variety of visual AI because it is built with a heavy emphasis on speed, size, and accuracy.

1) YOLOv7

The number of computations, the computational density, and a few other characteristics are frequently optimized while creating an effective network. Based on ELAN (efficient layer aggregation network), the YOLOv7 architecture design. In order for deeper networks to converge and train efficiently, ELAN takes into account constructing an efficient network by managing the shortest and longest gradient paths. Below is a diagram of ELAN's modules. In the Fig. 2, the bottom block's input is fed straight into a 1x1 convolution, whereas the input for the other two connections to the block was convolved with 2 and 4 3x3 convolution blocks using the same channel multiplier. In the bottom block, 1x1 convolution is used to learn rich information after concatenating all the features.

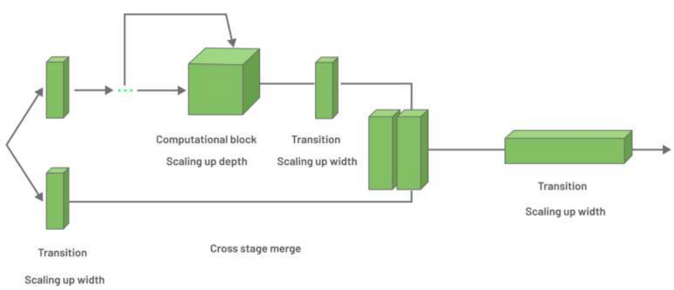

Fig. 2. YOLOv7 Architecture [30]

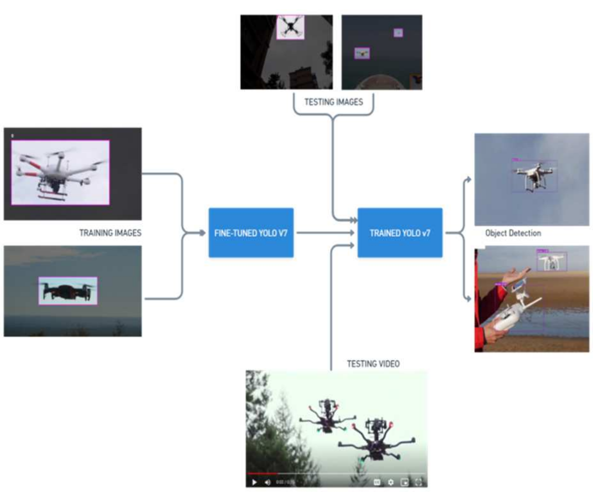

Fig. 3. YOLOv7 Methodology

Fig.3 gives the basic YOLOv7 methodology which is we first collected Training images from Roboflow and fine-tuned our YOLOv7 then we Trained our model using the fine-tuned yolov7 and testing images and video and finally evaluated our model on the validation images and obtained the bounding boxes which will be used for evaluation and our comparative study.

2) YOLOv8

With the introduction of YOLOv8, a model that establishes a new state-of-the-art for object recognition and instance segmentation, the field of computer vision progresses.

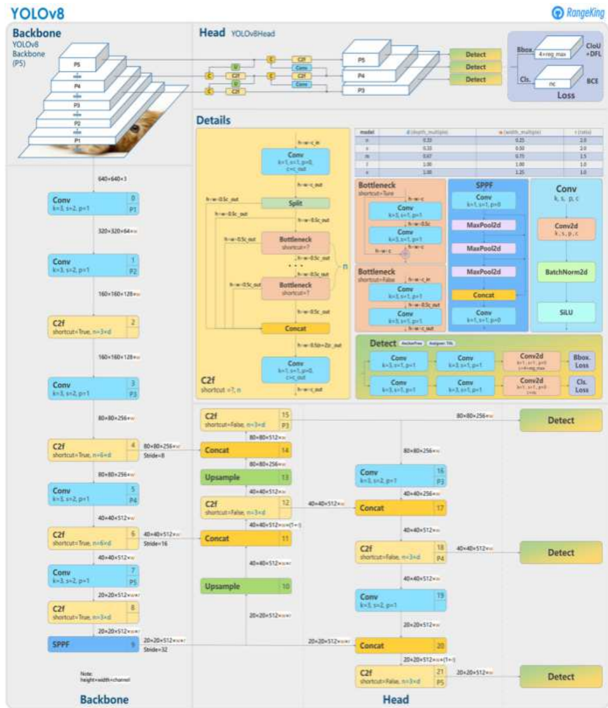

Fig. 4. YOLOv8 Architecture [29]

Fig. 4 , Fig. 5 and Fig. 6 depicts YOLOV8 architecture that is an anchor-free model. In other words, rather than predicting an object's offset from a known anchor box, it predicts the object's center directly.

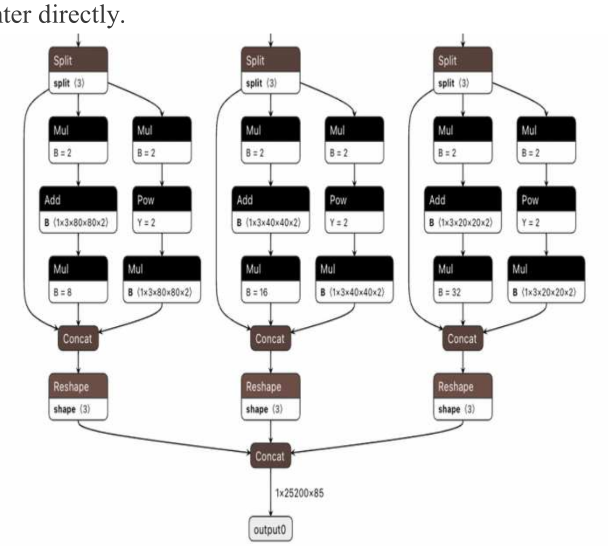

Fig. 5. YOLOv8 Architecture [29]

In order to speed up Non-Maximum Suppression (NMS), a challenging post-processing step that sorts through candidate detections following inference, anchor free detection lowers the number of box predictions.

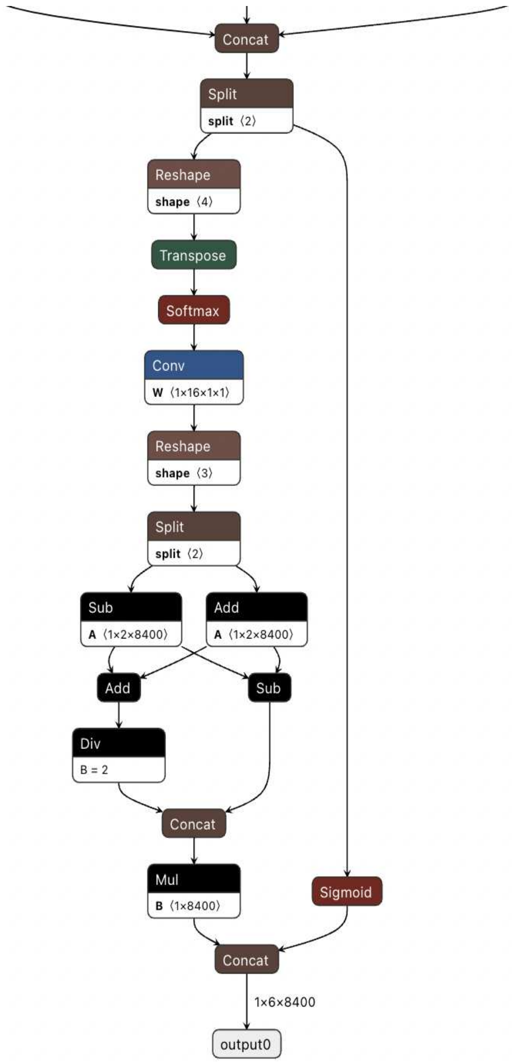

Fig. 6. YOLOv8 Architecture [29]

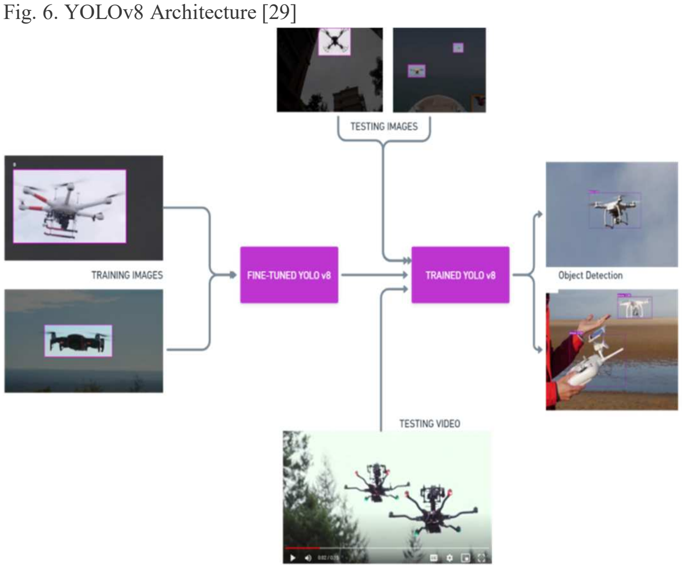

Fig. 7. YOLOv8 Methodology

In Fig.7 similar to our YOLOv7 methodology we give the basic YOLOv8 methodology which is we first collected Training images from Roboflow and fine-tuned our YOLOv8 then we Trained our model using the fine-tuned yolov8 and testing images and video and finally evaluated our model on the validation images and obtained the bounding boxes which will be used for evaluation and our comparative study.

IV. RESULTS AND DISCUSSION

From the below images of Fig. 8, Fig. 9, Fig. 10, Fig. 11 used for the study. The results of the bounding boxes of YOLOv8 are much better at predicting the drone images correctly than YOLOv7 whereas YOLOv7 was misclassifying not drone images as drone images which were YOLOv8. The YOLOv8 accuracy was 50% whereas YOLOv7 gave 48% accuracy, both were trained for 10 epochs.

We trained our models for 10 epochs. We can achieve even better accuracy if we train for more epochs, since the purpose of this paper was to evaluate the performance of YOLOv7 and YOLOv8 for drone detection tasks

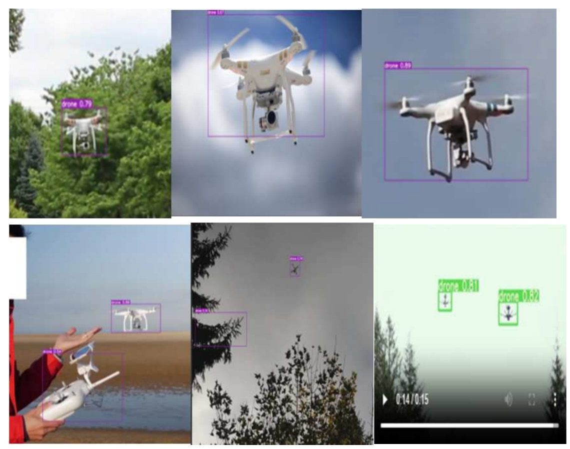

Fig. 8. YOLOv7 bounding boxes on Images

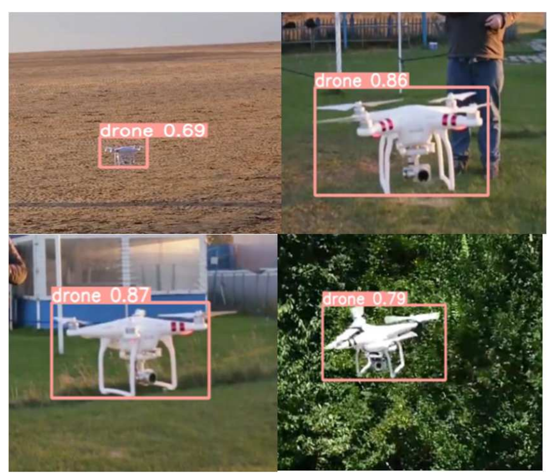

Fig. 9. YOLOv8 bounding boxes on Images

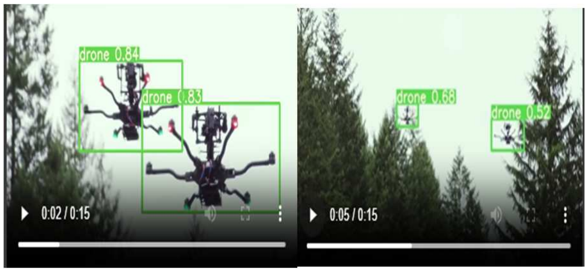

Fig. 10. YOLOv7 bounding boxes on Video

Fig. 11. YOLOv8 bounding boxes on Video

A. Evaluation Parameters

1). F1-Confidence curve

The relationship between a model's prediction confidence and F1 score is shown graphically by the F1 confidence curve. The F1 score, which considers both precision and recall, is a statistic frequently used to assess the effectiveness of a classification algorithm. Recall is the percentage of genuine positives among all actual positives, whereas precision measures the percentage of true positives among all projected positives. The harmonic mean of recall and precision is the F1 score.

The F1 confidence curve graphs a model's F1 score versus the confidence threshold used to categorize predictions as positive or negative. The threshold denotes the minimal degree of certainty necessary for the model to correctly categorize a

specific event as positive. The algorithm gets more cautious in its predictions as the threshold rises, identifying fewer examples as positives but with higher confidence. Table 1 refers the relationship between F1 and Confidence for YOLOV7 model.

TABLE 1 F1-CONFIDENCE CURVE YOLOV7

Parameter	***F1***		
Confidence	***0 (No drone)***	***Drone***	***All Class***
0.0	0.1	0.09	0.1
0.2	0.62	0.51	0.58
0.4	0.64	0.53	0.59
0.6	0.6	0.52	0.58
0.8	0.29	0.4	0.29
1.0	1.0	1.0	1.0

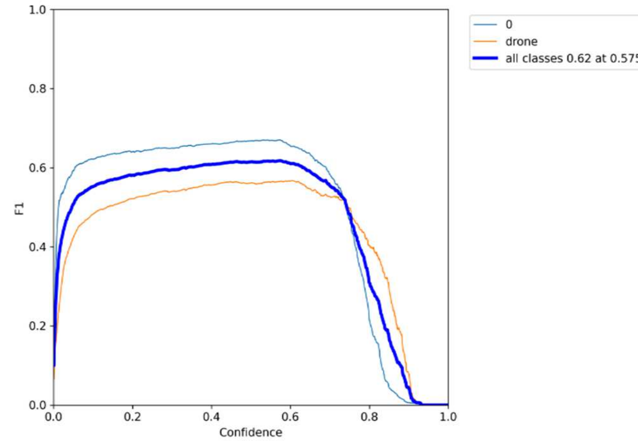

Fig. 12. F1-Confidence Curve for YOLOv7

In Fig.9., it can be observed that as the confidence is increasing beyond a certain point, in our case, 0.7, the F1 measure drone label drops drastically and reaches almost 0. It can be concluded that the model performs its best around a confidence interval of 0.6 to 0.8. The model can predict with high confidence if the object is a drone or not, but it cannot predict if an image contains a drone or not with more than 90% confidence. Fig. 12, represents the relationship of F1 score with respect to confidence for the three classes. Table 2 refers the relationship between F1 and Confidence for YOLOV8 model.

TABLE 2. F1-CONFIDENCE CURVE YOLOV8

Parameter	***F1***		
Confidence	***0 (No drone)***	***Drone***	***All Class***
0.0	0.021	0.09	0.022
0.2	0.61	0.49	0.58
0.4	0.58	0.5	0.56
0.6	0.2	0.51	0.39
0.8	0.11	0.5	0.29
1.0	0.01	1.0	1.0

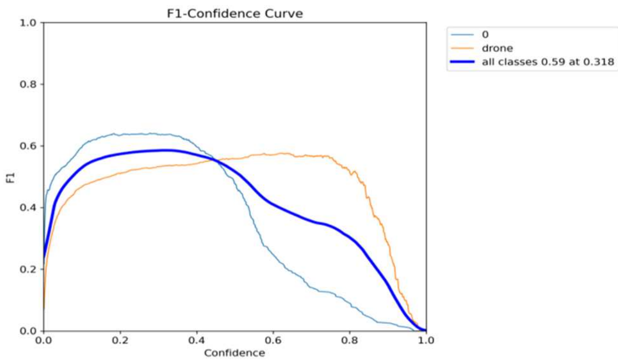

Fig. 13. F1-Confidence Curve for YOLOv8

Similarly, in Fig.13, it can be observed that as the confidence is increasing beyond a certain point, in yolov8's case, 0.8, the F1 measure drone label drops drastically and reaches almost 0. It can be concluded that the model performs its best around a confidence interval of 0.7 to 0.9. The model can predict with high confidence if the object is a drone or not, but it cannot predict if an image contains a drone or not with more than 100% confidence.

In both the above cases we can conclude YOLOv8 performs better than YOLOv7 and is able to predict if the image is drone or not with 90% of confidence but YOLOv7 cannot.

2) Precision-Recall curve

A binary classifier's performance is graphically depicted by a precision-recall (PR) curve. It is a depiction of the recall (sensitivity) and precision (positive predictive value) for various classification thresholds. The classification threshold of the classifier, which establishes how confident the classifier must be before issuing a positive label, is changed to produce the PR curve. At each threshold, precision and recall are calculated, and a curve with recall on the x-axis and precision on the y-axis is then shown. Table 3 refers the relationship between precision and recall for YOLOV7 model.

TABLE 3 PRECISION-RECALL CURVE YOLOV7

Parameter	**Precision**		
Recall	***0 (No drone)***	***Drone***	***All Class***
0.0	1.0	0.042	0.07
0.2	0.96	0.42	0.67
0.4	0.9	0.42	0.63
0.6	0.62	0.42	0.52
0.8	0.39	0.42	0.49
1.0	1.0	0.15	1.0

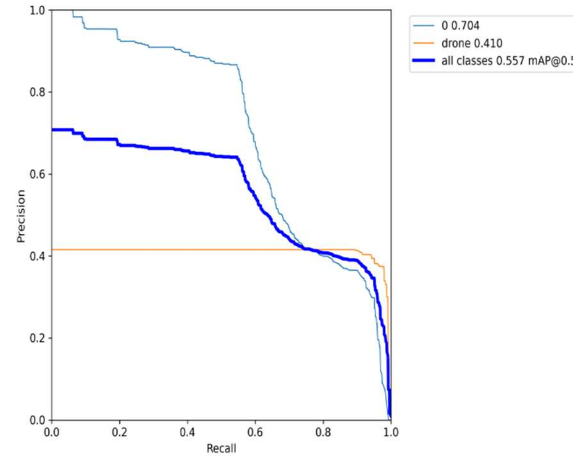

Fig. 14. Precision-Recall Curve for YOLOv7

In Fig.14, it can be observed that the precision is at max 0.4 and as the confidence increases the precision value stays constant so it would be better to set the threshold as 0.8 or 0.9 to predict whether the image is drone or not for better accuracy. Table 4 refers the relationship between precision and recall for YOLOV8 model.

TABLE 4 PRECISION-RECALL CURVE YOLOV8

Parameter	**Precision**		
Recall	**0 (No drone)**	**Drone**	**All Class**
0.0	1.0	1.0	1.0
0.2	0.78	0.52	0.78
0.4	0.74	.0.5	0.75
0.6	0.46	0.49	0.49
0.8	0.38	0.47	0.39
1.0	1.0	1.0	1.0

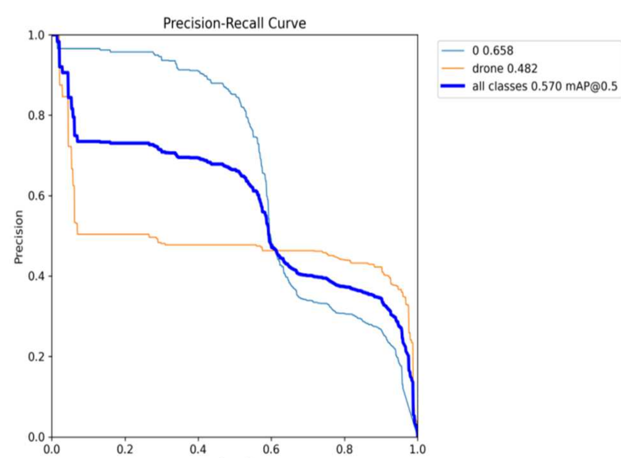

Fig. 15. Precision-Recall Curve for YOLOv8

Similarly, In Fig. 15., it can be observed that the precision is high in the beginning which is the ideal case but the recall is low which will give poor F1 score and after that we can observe that precision stays at max 0.4 and as the confidence increases the precision value stays constant so it would be better to set the threshold as 0.8 or 0.9 same as in yolov7 and then predict whether the image is drone or not for better accuracy. Both perform kind of similar in this evaluation metric.

2) Precision-Confidence curve

The link between a binary classifier's precision and the confidence in its predictions is depicted graphically by a Precision-Confidence curve, or PCC. It is used to assess a classification model's quality and to aid in deciding on a suitable threshold for the classifier. The projected probabilities of positive class membership by the classifier are sorted in decreasing order to get a PCC. The classifier's precision is then calculated for each of the bins that were created from the sorted probabilities. The average probability of the positive class for each bin is how the confidence is commonly defined. Next, the PCC is plotted with confidence on the x-axis and accuracy on the y-axis. Table 5 refers the relationship between precision and confidence for YOLOV7 model.

TABLE 5 PRECISION-CONFIDENCE CURVE YOLOV7

Parameter	**Precision**		
Confidence	***0 (No drone)***	***Drone***	***All Class***
0.0	0.017	0.011	0.012
0.2	0.67	0.35	0.51
0.4	0.76	0.37	0.56
0.6	0.83	0.38	0.57
0.8	0.91	0.34	0.6
1.0	1.0	1.0	1.0

In Fig. 16., it can be observed that the precision drops at around 0.9 confidence but then sharply shoots up at 0.9 confidence. Table 6 refers the relationship between precision and confidence for YOLOV8 model.

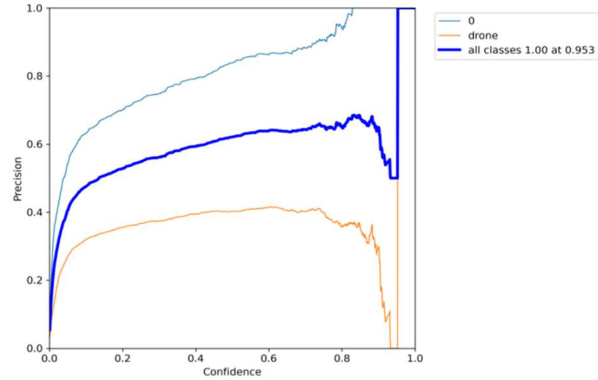

Fig. 16. Precision-Confidence Curve for YOLOv7

TABLE 6: PRECISION-CONFIDENCE CURVE YOLOV8

Parameter	**Precision**		
Confidence	***0 (No drone)***	***Drone***	***All Class***
0.0	0.019	0.012	0.019
0.2	0.76	0.3	0.5
0.4	0.83	0.33	0.51
0.6	0.89	0.36	0.62
0.8	0.86	0.38	0.68
1.0	1.0	1.0	1.0

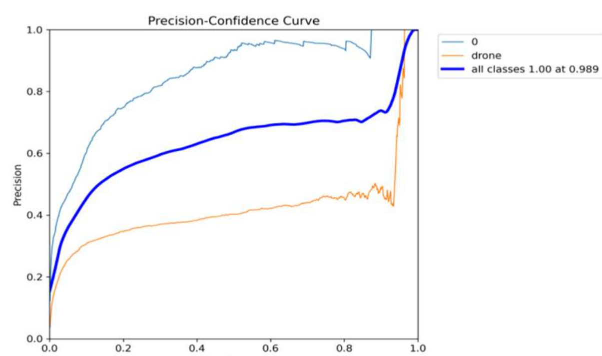

Fig. 17. Precision-Confidence Curve for YOLOv8

Similarly, In Fig. 17., it can be observed that the precision rises as the confidence interval increases which is a good thing as it can predict with around 0.9 confidence and give high precision if the image is drone or not.

3) Confusion Matrix

Confusion matrix is the representation of a tabular layout of the different outcomes of the prediction and to evaluate the binary classifier performance. There are four values in a confusion matrix: True Positive (TP), False Positive (FP), True Negative

(TN) and False Negative (FN). The confusion matrix provides helpful information about the binary classifier's effectiveness. It can also be used to calculate the F1 Score, recall accuracy and precision.

a) Confusion Matrix YOLOv7

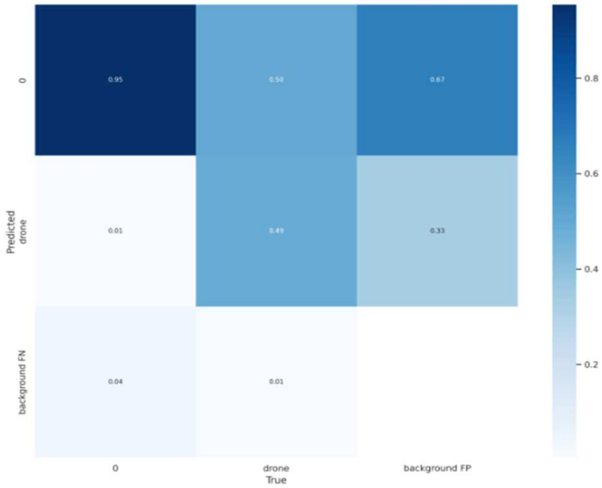

Fig. 18. Confusion Matrix for YOLOv7

Confusion Matrix Interpretation:

True Positives (TP) for each class:

Class 1: TP1 = 0.95 (value at the intersection of Class 1 true and predicted)

Class 2: TP2 = 0.49 (value at the intersection of Class 2 true and predicted)

Class 3: TP3 = 0 (value at the intersection of Class 3 true and predicted)

False Positives (FP) for each class:

Class 1: FP1 = 0.50 + 0.67 = 1.17 (sum of values in the predicted Class 1 column except for TP1)

Class 2: FP2 = 0.01 + 0.04 = 0.05 (sum of values in the predicted Class 2 column except for TP2)

Class 3: FP3 = 0.33 + 0.67 + 0 = 1.00 (sum of values in the predicted Class 3 column except for TP3)

False Negatives (FN) for each class:

Class 1: FN1 = 0.01 + 0.04 = 0.05 (sum of values in the true Class 1 row except for TP1)

Class 2: FN2 = 0.50 + 0.01 + 0.01 = 0.52 (sum of values in the true Class 2 row except for TP2)

Class 3: FN3 = 0.95 + 0.50 + 0.49 = 1.94 (sum of values in the true Class 3 row except for TP3)

True Negatives (TN) for each class:

Class 1: TN1 = 0.49 + 0.33 + 0.01 + 0.01 = 0.84 (sum of values not in the true Class 1 row and not in the predicted Class 1 column)

Class 2: TN2 = 0.95 + 0.67 + 0.04 = 1.66 (sum of values not in the true Class 2 row and not in the predicted Class 2 column)

Class 3: TN3 = 0.95 + 0.50 + 0.67 + 0.01 + 0.04 + 0.01 = 3.18 (sum of values not in the true Class 3 row and not in the predicted Class 3 column)

Fig.18, is the obtained Confusion matrix of the YOLOv7's performance in classifying if the image is drone or not. Here the model predicts correctly that the image is not a drone with True Negative value 0.95 but does not do that well in predicting correctly if the image is a drone as it gives the True positive value of 0.49.

From the above confusion matrix, we calculated the accuracy by formula:

$$(TP+TN)/(TP+TN+FP+FN) \qquad (1)$$

Accuracy: (TP1 + TP2 + TP3) / Total = (0.95 + 0.49 + 0) / (0.95 + 0.50 + 0.67 + 0.01 + 0.49 + 0.33 + 0.04 + 0.01 + 0) = 1.44 / 2.99 ≈ 0.4816 (approx. 48.16%)

Therefore, we got 48.16% accuracy for our YOLOv7 model trained with 10 epochs.

b) Confusion Matrix YOLOv8

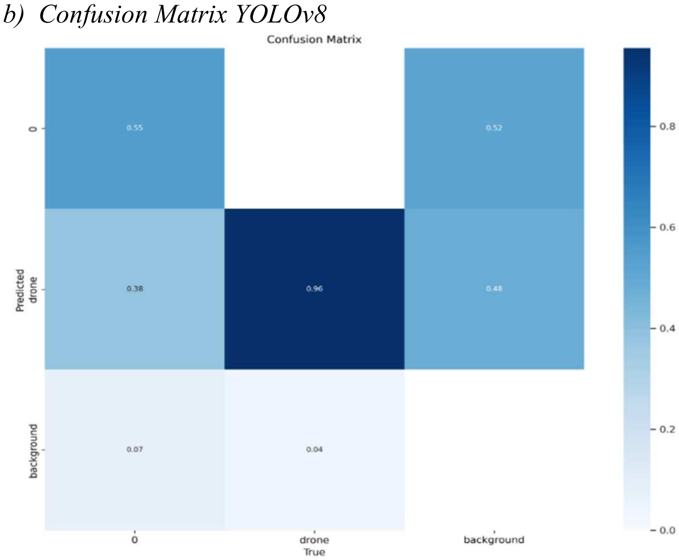

Fig. 19. Confusion Matrix for YOLOv8

Confusion Matrix Interpretation:

True Positives (TP) for each class:

Class 1: TP1 = 0.55 (value at the intersection of Class 1 true and predicted)

Class 2: TP2 = 0.96 (value at the intersection of Class 2 true and predicted)

Class 3: TP3 = 0 (value at the intersection of Class 3 true and predicted)

False Positives (FP) for each class:

Class 1: FP1 = 0 + 0.49 = 0.49 (sum of values in the predicted Class 1 column except for TP1)

Class 2: FP2 = 0.38 + 0.07 = 0.45 (sum of values in the predicted Class 2 column except for TP2)

Class 3: FP3 = 0.52 + 0.49 + 0 = 1.01 (sum of values in the predicted Class 3 column except for TP3)

False Negatives (FN) for each class:

Class 1: FN1 = 0.38 + 0.07 = 0.45 (sum of values in the true Class 1 row except for TP1)

Class 2: FN2 = 0.52 + 0.07 + 0.04 = 0.63 (sum of values in the true Class 2 row except for TP2)

Class 3: FN3 = 0.55 + 0.38 + 0.96 = 1.89 (sum of values in the true Class 3 row except for TP3)

True Negatives (TN) for each class:

Class 1: TN1 = 0.96 + 0.49 + 0.04 = 1.49 (sum of values not in the true Class 1 row and not in the predicted Class 1 column)

Class 2: TN2 = 0.55 + 0.52 + 0 = 1.07 (sum of values not in the true Class 2 row and not in the predicted Class 2 column)

Class 3: TN3 = 0.55 + 0.38 + 0.96 + 0.49 + 0.07 + 0.04 = 2.49 (sum of values not in the true Class 3 row and not in the predicted Class 3 column) Fig. 19., is the obtained Confusion matrix of the YOLOv8's performance in classifying if the image is drone or not. Here the model predicts correctly that the image is a drone with True Positive 0.96 but does not do that well in predicting correctly if the image is not a drone as it gives the True Negative value of 0.55.

From the above confusion matrix in Fig. 19, we calculated the accuracy by formula as seen in [eq. (1)]
Accuracy: (TP1 + TP2 + TP3) / Total = (0.55 + 0.96 + 0) / (0.55 + 0 + 0.52 + 0.38 + 0.96 + 0.48 + 0.07 + 0.04 + 0) ≈ 0.5016
Therefore, we got 50.16% accuracy for our YOLOv8 model trained with 10 epochs.

TABLE 7. ACCURACY RECORDED FOR RESPECTIVE MODELS

YOLOv7 Accuracy	YOLOv8 Accuracy
48.16%	50.16%

Table 5 gives the accuracy comparison between YOLOV7 and YOLOV8 models. We can conclude that YOLOv8 outperforms YOLOv7 with a slightly higher accuracy of 50.16% compared to 48.16%.

Like many other object detection algorithms, YOLO v7 has some issues with small objects. For example, it may not be able to detect small objects in busy scenes or when the object is far from the camera. YOLO v7 can be computationally intensive, that is, running it in real-time can be challenging on devices with limited resources, such as smartphones or edge devices.

Similarly, YOLO v8 may not be able to detect objects when the scene is cluttered or when the objects are partially obscured. It may also be unable to detect small objects or low contrast objects.
YOLOv8 does not support models trained in 1280 (in pixels) resolution, so if you are planning to run the inference at high px resolution, then it is not advisable to run the inference using YOLOv8.

V. CONCLUSION AND FUTURE WORK

In our research, we focused on performing image detection on drone images using YOLOv7 and YOLOv8 models. Our main objective was to classify whether a given image contains a drone or not. To achieve this, we first fine-tuned the pre-trained YOLOv7 and YOLOv8 models on our specific drone dataset. This involved adapting the models to perform well on the new task of drone detection. After fine-tuning, we proceeded to train both models on the drone dataset to further improve their accuracy in detecting drones. We then evaluated the models using various metrics to determine their performance. Based on the evaluation results, we found that YOLOv8 performed slightly better than YOLOv7 in accurately classifying drone images. We were particularly impressed with the robustness of YOLOv8, as it demonstrated the ability to handle variations in the data and generalize to unseen examples better than any previous version of YOLO. Our research highlights the superior performance of YOLOv8 compared to YOLOv7 in drone image detection. We also look forward to the practical implementation of our findings by deploying the YOLOv8 model on edge devices for real-time drone detection in important security applications. In future enhancement to implement data augmentation techniques to generate additional training samples, thereby enriching the dataset and enhancing the models' ability to handle variations in real-world drone images. Extend the research to perform multi-class detection, where the models can detect and differentiate various types of UAVs, enabling a more comprehensive approach to drone surveillance.

REFERENCES

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

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

[3]. Pengfei Zhu et al. "VisDroneDET2018: The Vision Meets Drone Object Detection in Image Challenge Results". 23 January 2019.

[4]. C. Aker and S. Kalkan, "Using deep networks for drone detection," 2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)", Lecce, Italy, 2017, pp. 1-6.

[5]. Xin Dai & Masaaki Nagahara (2023), "Platooning control of drones with real-time deep learning object detection", Advanced Robotics, 37:3, 220-225.

[6]. Naranjo, M.; Fuentes, D.; Muelas, E.; Díez, E.; Ciruelo, L.; Alonso, C.; Abenza, E.; Gómez-Espinosa, R.; Luengo, I. "Object Detection-Based System for Traffic Signs on Drone-Captured Images" .2023, 7, 112.

[7]. Z. Qiu, H. Bai, and T. Chen, "Special Vehicle Detection from UAV Perspective via YOLO-GNS Based Deep Learning Network," *Drones*, vol. 7, no. 2, p. 117, Feb. 2023.

[8]. Zoe Mayer, James Kahn, Yu Hou, Markus Götz, Rebekka Volk, Frank Schultmann, "Deep learning approaches to building rooftop thermal bridge detection from aerial images", Automation in Construction, Volume 146, 2023, 104690.

[9]. Xi, Y.; Jia, W.; Miao, Q.; Feng, J.; Liu, X.; Li, F. "CoDerainNet: Collaborative Deraining Network for Drone-View Object Detection in Rainy Weather Conditions". Remote Sens. **2023**, 15, 1487.

[10]. T. Abdellatif, M. A. Sedrine and Y. Gacha, "DroMOD: A Drone-Based Multi-Scope Object Detection System," in *IEEE Access*, vol. 11, pp. 26652-26666, 2023.
[11]. [10] Aydin, B.; Singha, S. “Drone Detection Using YOLOv5”. *Eng* **2023**, *4*, 416-433.
[12]. Thai P, Alam S, Lilith N, Nguyen B. “Small Flying Object Detection and Tracking in Digital Airport Tower through Spatial-Temporal ConvNets”. Research Square; 2023.
[13]. Singha, S.; Aydin, B. “Automated Drone Detection Using YOLOv4. Drones” **2021**, *5*, 95.
[14]. Wu, Manjia & Jiayang, Xie & Shi, Xiufang & Shao, Panyu. (2018). “Real-Time Drone Detection Using Deep Learning Approach”: Third International Conference, MLICOM 2018, Hangzhou, China, July 6-8, 2018.
[15]. D. Lee, W. Gyu La and H. Kim, "Drone Detection and Identification System using Artificial Intelligence," 2018 International Conference on Information and Communication Technology Convergence (ICTC), Jeju, Korea (South), 2018, pp. 1131-1133.
[16]. Liu, Hao & Qu, Fangchao & Liu, Yingjian & Zhao, Wei & Chen, Yitong. (2018). “A drone detection with aircraft classification based on a camera array”. IOP Conference Series: Materials Science and Engineering. 322. 052005.
[17]. B. Kim, D. Khan, C. Bohak, W. Choi, H. Lee, and M. Kim, “V-“RBNN Based Small Drone Detection in Augmented Datasets for 3D LADAR System”, Sensors, vol. 18, no. 11, p. 3825, Nov. 2018.
[18]. Jin R, Jiang J, Qi Y, Lin D, Song T. Drone Detection and Pose Estimation Using Relational Graph Networks. *Sensors*. 2019; 19(6):1479. https://doi.org/10.3390/s19061479
[19]. H. Chen, Z. Wang and L. Zhang, "Collaborative spectrum sensing for illegal drone detection: A deep learning-based image classification perspective," in China Communications, vol. 17, no. 2, pp. 81-92, Feb. 2020, doi: 10.23919/JCC.2020.02.007.
[20]. Mahdavi, Fatemeh & Rajabi, Roozbeh. (2020). Drone Detection Using Convolutional Neural Networks. 1-5. 10.1109/ICSPIS51611.2020.9349620.
[21]. D. K. Behera and A. Bazil Raj, "Drone Detection and Classification using Deep Learning," 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, 2020, pp. 1012-1016, doi: 10.1109/ICICCS48265.2020.9121150.
[22]. M. S. Allahham, T. Khattab and A. Mohamed, "Deep Learning for RF-Based Drone Detection and Identification: A Multi-Channel 1-D Convolutional Neural Networks Approach," 2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT), Doha, Qatar, 2020, pp. 112-117, doi: 10.1109/ICIoT48696.2020.9089657.
[23]. D. T. Wei Xun, Y. L. Lim and S. Srigrarom, "Drone detection using YOLOv3 with transfer learning on NVIDIA Jetson TX2," 2021 Second International Symposium on Instrumentation, Control, Artificial Intelligence, and Robotics (ICA-SYMP), Bangkok, Thailand, 2021, pp. 1-6, doi: 10.1109/ICA-SYMP50206.2021.9358449.
[24]. Çetin, Ender & Barrado, C. & Pastor, E.. (2021). Improving real-time drone detection for counter-drone systems. The Aeronautical Journal. 125. 1-26. 10.1017/aer.2021.43.
[25]. M. Aledhari, R. Razzak, R. M. Parizi and G. Srivastava, "Sensor Fusion for Drone Detection," 2021 IEEE 93rd Vehicular Technology Conference (VTC2021-Spring), Helsinki, Finland, 2021, pp. 1-7, doi: 10.1109/VTC2021-Spring51267.2021.9448699.
[26]. Lee, Dong-Hyun. (2021). CNN-based single object detection and tracking in videos and its application to drone detection. Multimedia Tools and Applications. 80. 1-12. 10.1007/s11042-020-09924-0.
[27]. Lv, Yaowen et al. “High-Resolution Drone Detection Based on Background Difference and SAG-YOLOv5s.” *Sensors (Basel, Switzerland)* vol. 22,15 5825. 4 Aug. 2022, doi:10.3390/s22155825
[28]. K. Agarwal, S. Deepa, R. V. SivaBalan and C. Balakrishnan, "Performance Analysis of Various Machine Learning Classification Models Using Twitter Data: National Education Policy," 2023 International Conference on Intelligent and Innovative Technologies in Computing, Electrical and Electronics (IITCEE), Bengaluru, India, 2023, pp. 862-870, doi: 10.1109/IITCEE57236.2023.10091034..
[29]. Putro, Muhamad Dwisnanto & Priadana, Adri & Jo, Kang-Hyun. (2023). “Efficient Multi-Receptive Pooling for Object Detection on Drone.” 10.1007/978-981-99-4914-4_2.
[30]. Solawetz, J. (2023, January 11). What is YOLOv8? “The Ultimate Guide. Roboflow” Blog. https://blog.roboflow.com/whats-new-in-yolov8.
[31]. Rizwan, M. (n.d.).” YOLOv7 Architecture Explanation. Cameralyze - Connect to AI with API”. https://www.cameralyze.co/blog/yolov7-architecture explanation.