Int. J. Nonlinear Anal. Appl. 13 (2022) 2, 2709–2722
ISSN: 2008-6822 (electronic)
http://dx.doi.org/10.22075/ijnaa.2022.27556.3648

Detection of drones with YOLOv4 deep learning algorithm

Ahmed S. Naseri*, Nada Hussein M. Ali

Computer Science Department, College of Science, University of Baghdad, Baghdad, Iraq

(Communicated by Javad Vahidi)

Abstract

Drones or unmanned aerial vehicles (UAVs) have rapidly spread all over the world and are becoming widely popular in major cities for personal and commercial use. It has also been widely used for military purposes in the last decade. Thus, it has become difficult to maintain control over them and the risks they pose to privacy and security. In this paper, we present a solution to detect drones before they can reach a sensitive area or residence using the latest YOLOv4 deep learning algorithm while using Darknet as a backbone. We trained our model on different images at different distances and climatic conditions and trained our model to detect birds and aircraft that are very similar to drones at higher distances that may cause confusion, and also train the system at close distances and at very low and high image quality. For all available cases, our dataset was collected from three global and certified datasets in aircraft detection systems and the result was a dataset containing all cases. However, the collection of drones, birds and aircraft datasets is not easy to obtain. The proposed method achieved an accuracy of 98.3% with the main challenge of detecting similar small objects near and far in all conditions.

Keywords: Drones, Unmanned Aerial Vehicles, YOLOv4, Deep learning, Backbone, Dataset
2020 MSC: 68T07

1 Introduction

Drones, also known as unmanned aerial vehicles (UAVs), have captivated the curiosity of hobbyists and investors in recent years [12]. Due to their small size and ability to fly autonomously, such drones have a wide number of commercial applications, including agriculture, photography, and a variety of public services [4]. Simultaneously, they can be used to carry out chemical, and biological attacks, or to assist in the smuggling of drugs or illegal immigrants across the border, posing security risks to public safety due to their small size and ability to fly low enough to evade conventional radar detection [2], Commercial drone companies such as DJI, Parrot, and 3DRobotics are rapidly growing in popularity as a result of the continuous advancement of associated technology. Drones have been widely deployed in both military and civilian environments, and their use has surged in recent years as a result of their low cost and ease of operation. [1], Drone purchases totalled 1.9 million in 2016, according to the "Federal Aviation Administration (FAA)", and were expected to reach 4.3 million by 2020 and their spread and use has increased in 2021 [6]. Object identification methods and categorization are based on the items' shared characteristics. This means that traits are defined by their appearance as well as their behavior and movement patterns. The problem of recognizing unmanned aerial vehicles (UAVs) in the sky focuses on both the items that may appear in the sensing region as well as the UAVs themselves [12].and The problem of recognizing "unmanned aerial vehicles (UAVs)" in the sky focuses on both

*Corresponding author
Email addresses: ahmad.samir101@gmail.com (Ahmed S. Naseri), nada.husn@sc.uobaghdad.edu.iq (Nada Hussein M. Ali)

Received: November 2021 *Accepted:* January 2022

the objects that may appear in the sensing region as well as the UAVs themselves. A UAV's form is one of its most distinguishing visual qualities. Each type of "UAV (from the tricopter to the octocopter)" appears strikingly similar within their own categories [3]. A tricopter is shaped like an equilateral triangle, a quadcopter like a square, and so on. Additionally, each UAV is constructed from a rigid construction that has individual visual characteristics. Typically, the control board is located in the structure's center, with three to eight arms depending on the number of propellers [10]. Each arm is completed by the engine from which the propeller is attached. This is the most fundamental UAV design, and it can be used to practically build any UAV [13]. Propellers and motors can be fitted directly on the center panel of the aforementioned UAV, but there may be an exception in the nano and small categories [15].

2 Methods

Many approaches have been used to eliminate this menace, and many of them have shown promising results, but some of them are limited by distance, weather conditions, or object size. Detection with still photos or videos provides several advantages over other methods, including tool and equipment accessibility, a near to medium range, and a high detection rate. The detection range for various approaches is shown in table 1 below [23, 14]:

Table 1: The import technology in detection approaches [19]

Technology	Advantages	Disadvantages	Range
Image	Mature technology Equipment accessible.	A network of high-resolution cameras is required. Subject to issues with visibility.	100–1000 m
Audio	Sensors are cheap. Processing is limited. Equipment-accessible.	Sensitized to ambient noise Range is restricted, an array of distributed microphones is required. Large datasets are required for training.	40–300 m
Radar	Simple to install. long range.	Either monostatic with a limited resolution or multi-static with a scattered pattern. Detecting small drone cross- sections can be difficult. Expensive.	$\leq$ 3000 m
RF	Simple to install. Sensing devices are cheap.	To accomplish detection, a high SNR is required. Irregularities are possible. The range is constrained by the quality of the underlying links.	$\leq$ 1000 m

Computer vision technologies have been evolving in the past years, and machine learning has become more accurate in distinguishing objects by their visual features, although it has a noticeable difficulty with the recognition of smaller objects (e.g., birds and drones), with the development of detection algorithms and captured media quality, the time has come for computer vision to be developed to capture the small or sometimes tiny (35×35) objects in media such as videos or still pictures. Numerous technologies such as rapid R-CNN, Retina-Net, single-shot multi-box detector (SSD), and You only look once (YOLO) can be used to recognize objects using neural networks [5, 20]. We worked on one of the newest, most advanced, and most accurate algorithms in this study to detect small objects in the sky (drones) using a deep learning approach (YOLOv4).

3 YOLO Algorithm

YOLO [5] stands for "you only look once", this is a cutting-edge object detection system capable of real-time object recognition. In YOLO, object detection is a regression issue. Additionally, YOLO recognizes objects in real-time by utilizing a convolutional neural network (CNN) with a single forward propagation. This means that a single algorithm is used to forecast an image, and YOLO is available in a variety of variants, including YOLOv3 and tiny YOLO. In their article "You Only Look Once: Unified, Real-Time Object Detection" [20], Joseph Redman introduced the first version of YOLO in 2015. RCCN models were accurate, but they were sluggish since they were a multi-step procedure to determine the recommended region for objects and then classify them; they also had to undertake post-processing to improve the result. They remained the most popular methods for detecting objects, and YOLO's major goal is to eliminate multistage detection and complete object detection in a single run [3]. It raised inference time but decreased training time. YOLOv1 was a big step forward at the time, with 63.4 mAP and an inference time of 22 milliseconds per image, compared to competing algorithms that took 20 to 143 milliseconds per image. YOLO improves speed, accuracy, and learning capabilities. It improves speed by predicting objects in real-time, and accuracy by providing minimal background errors [24]. YOLO architecture for its different models are similar, defined below [21]:

- Backbone: A convolutional neural network that produces visual features with many shapes and sizes.
- Neck: A set of layers that connect and mix characteristics before delivering them to the prediction layer.
- Head: performs classification and regression on the features and the bounding boxes, after receiving the features and the bounding boxes from the neck to finish the detection process.

YOLOv4 is considered the fastest, most accurate algorithm for object detection, it has a Darknet53 backbone, the same as YOLOv3 [16]. YOlOv4 coined two critical concepts: the bag of freebies and the bag of specials. The most well-known object that embodies or falls inside the purview of this modern algorithm is [29, 28]:

- Bag of freebies (BOF): refers to the different techniques that improve the model without increasing the inference, different concepts like Cut mix (cutting and mixing different images that contain objects we want to detect), Mix-up (randomize images in a dataset), Cutout and Mosaic data augmentation. Bounding box regression box (different types of bounding boxes), regularization, and Normalization (cross mini-batch normalization which increases accuracy and GPU normalization).
- Bag of specials (BOS): techniques that increase accuracy as well as the computation cost such as spatial attention modules (SAM) which generate feature maps using the inter-spatial relationship for different features. Non-max suppression (NMS) is used for multiple objects that are grouped to minimize false prediction boxes. Non-linear activation functions, "weighted residual connections (WRC)" or "cross-stage partial connection (CSP)".

4 Related work

There are many approaches to UAVs detection, each with its own advantages and disadvantages. as drones become more widely used, researchers developed more sophisticated, more advanced and more accurate approaches for detecting them. In [27] suggested system employs a deep learning approach to create a real-time drone detector. To be more specific, enhanced the performance of a well-performing deep learning model, YOLOv2, by altering its structure and tweaking its parameters to better accommodate drone detection. Given that a robust detector requires a large number of training images, also proposed a semi-automatic dataset labelling method based on the Kernel zed Correlation Filters tracker to accelerate the pre-processing of the training images and train the network using the USC Drone Dataset and another type of dataset. The combined datasets yield a recall of 85.44% and a precision of 88.35% for the system. as well as the combined datasets have a precision of 60.59% and a recall of 62.44%, indicating the suggested system's limitation. The proposed system does not work on backgrounds, but only on drones within a close distance. And In [25] To protect important places, a proposed system that utilizes computer vision and machine learning to power an autonomous drone monitoring system is presented. A wide-angle, high-resolution daylight camera and a somewhat narrow-angle thermal camera mounted on a rotating turret comprised the system. The wider viewing daylight camera detects intruders as small as 20 pixels with an extremely low false alarm rate. The primary detection is based on YOLO convolutional neural network (CNN) rather than conventional background subtraction algorithms due to its low false alarm rate performance. At the same time, the tracked flying objects are tracked by the rotating turret and classified by the narrow-angle, zoomed thermal camera, where the classification algorithm is also based on CNNs. The training of the algorithms is performed by artificial and augmented datasets due to the scarcity of infrared videos of drones and the result of the system is 0.91 accuracy and zero false alarm. And In [21] The suggested system recognizes low-altitude drone identification methods using the YOLOv4 model and also introduces the YOLOv4 algorithm for the first time in low-altitude UAV object detection. To address the lack of a standard data set, a sample collection of drones flying low attitude photographs is contracted through shooting, downloaded from the internet, to increase the current data. and the experimental results indicate that YOLOv4, YOLOv3, and SSD are all one-stage algorithms, with YOLOv4 achieving an accuracy of 89.32% and a recall of 92.48%, while YOLOv3 achieves an accuracy of 84.14 percent and a recall of 89.27 percent. And 79.52% accuracy and 85.31% recall in SDD. The limitation of the proposed system is that it only works on low-distance drones, and does not work at long distances, which are the most difficult to detect. In [22] the proposed system used YOLOv4 to create an automated drone detection system. The model was trained using datasets from drones and birds. On the testing dataset, evaluation of the trained YOLOv4 model was done by using mean average precision (mAP), frames per second (FPS), precision, recall, and F1-score as evaluation metrics. Following that, two types of drone footage were collected, conducted drone detections, and calculated the

FPS to determine the detection speed at three altitudes. the proposed method has outperformed previous comparable studies, attaining an mAP of 74.36%, a precision of 0.95, a recall of 0.68, and an F1 score of 0.79. For video detection, it achieved a frame rate of 20.5 frames per second on the DJI Phantom III and 19.0 frames per second on the DJI Mavic Pro, and the suggested system's drawback is that it does not function on a wider variety of drones. The image dataset was utilized to improve the results further, and the suggested system is not focused on recognizing drones at various altitudes, which can be challenging due to their small size, high altitude and speed, and the presence of drone-like items.

5 Objective of the study

The main objective of the research can be summarized in two points:

1. Training the model to distinguish between drones, birds and planes.
2. Detection of drones that are at long distances and close distances as well in complex environments.

6 Proposed System

In the proposed system, work was done on the principle of deep learning through the use of an algorithm (you only look once or (YOLO)). And the primary objective of the proposed system is to identify drones in more complicated environments and at greater distances, as well as to teach the model to distinguish between drones, birds, and planes. However, it is very difficult to extinguish drones, birds and planes from a far distance due to their similar shape when it comes to planes and birds, or the similar size drones and birds. This actually becomes more difficult with bad weather or bad lighting.

The proposed system in general proposes collecting many different photos with different distances, lighting and weather conditions to help our model learn better about our objects. We also used YCbCr colour space, with the help of OpenCV, our model can understand the lighting patterns better because it is the native format, which represents colours as combinations of a brightness signal and two chroma signals. While RGB represents colours as combinations of red, green and blue signals. YCbCr helps with using specific filters on the dataset in preprocessing. the dataset has been collected from different sources which include images from the web and images from other open-source datasets. Our dataset contains about Eight thousand images with approximately twelve thousand annotated objects. Keeping in mind that the machine we are using for training is a personal laptop with 6GB VRAM, which makes it more difficult to have a large training dataset since we were using GPU for training, this was one of the challenges we had to deal with. Since our dataset was collected online, it had different image sizes and aspect ratios. We used one resolution for all images (512×512) to reduce our total dataset size, however, this change in resolution cropped images and distorted them which in return cropped some objects out of the image and distorted their features, simple solution for that issue was using a letter box on all images to keep everything inside while freely changing the resolution. Some images must also include birds and drones, planes and drones. . . etc. or include all of the three objects together, the reason behind this is to train our model to have these objects in one image, otherwise, it will not be able to detect them together in one environment. We have made our dataset open source and it can be downloaded from the web for future studies and more improvements on the detection of these objects. Furthermore, we have divided our dataset into 5 K-folds for more accurate results and the training was done using the K-folds method. The pre-processing stage included dividing our dataset into 5 k-folds randomly, to make sure that our results will be as accurate as possible. Figure 1 shows the block diagram of the proposed system:

6.1 Standard Dataset Sources

The data set presents a major challenge to the proposed system due to the lack of a data set that contains within it drones, birds, planes, and far-away objects or complex environmental factors to identify between them. So a dataset containing 50 thousand diverse drone images [18] was used to collect different random drone images and the dataset expands adding multiclass image classification and object detection datasets (ImageNet, MS-COCO, PASCAL VOC, anti-UAV), with other online sources to collect the desired dataset, and the result was 4000 drones images distributed proportionally. and using another standard dataset (PASCAL Visual Object Classes Challenge 2011 (VOC2011) Complete Dataset [17], Caltech-UCSD Birds-200-2011 [7], Stanford STL-10 Image Dataset [8]) to collect birds and planes images. the result of the new dataset collected dataset is (50% drones only, 25% plans, 25% birds). All of these three data sets were freely available to researchers, and based on them, a new data set of 8000 images is created. a mixture of birds, planes and drones. to evaluate the work of the algorithm adopted in the proposed system. This type

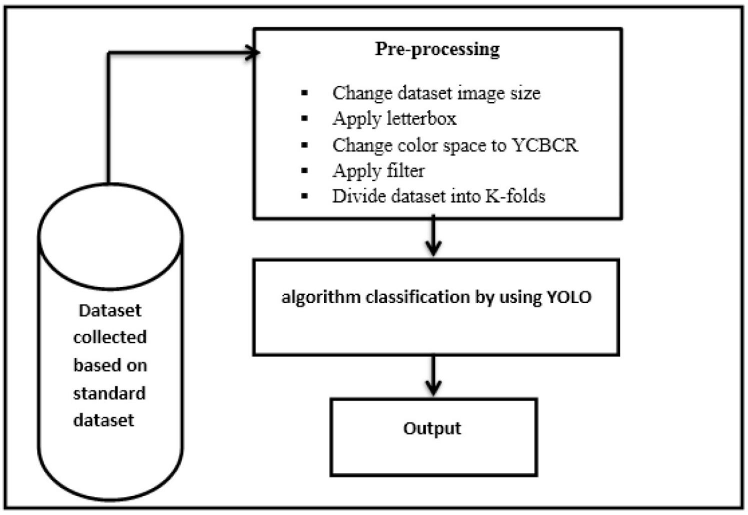


Figure 1: Block Diagram of Proposed System

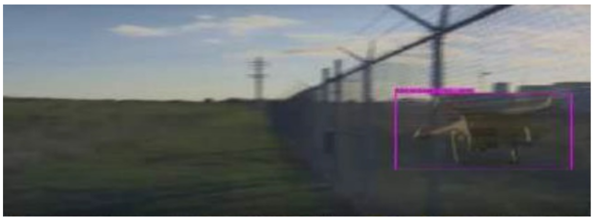
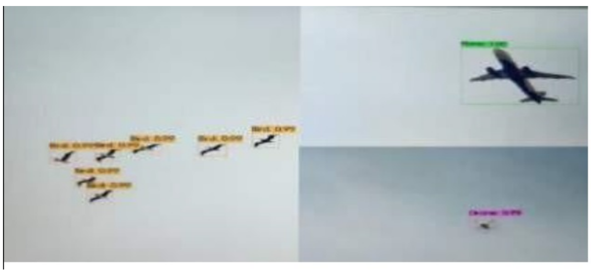
of collection worked well due to the unavailability and absence of a three-factor data set in order to train and evaluate the system. And the next table 2 shows the different types of images in the dataset

Table 2: image samples from the collected dataset

Image	Description
	Fog: Fog plays an important challenge factor, due to the difficulty of distinguishing the drones in a foggy atmosphere, especially if the drones are of light colors.
	Night: The problem of darkness is the problem that most drone detection systems suffer from, due to the difficulty of detecting and identifying drones at night.
	Distance: Detection of drones from very long distances plays an important factor for drone detection systems in order to report danger before it occurs. This factor is considered one of the most important challenges that detection systems face.

6.2 Pre-processing in proposed system

In the proposed system, preprocessing was done in the following order:

	Planes: The idea can be simplified that aircraft have a structure similar to drones at greater distance, also, some drones have shapes similar to planes.
	Birds: The idea can be simplified that birds have a structure similar to drones at close or great distances.
	Noise: Most of the drone photos suffer from noise, contrast, and have different angles
	Camouflage: This condition is very difficult to distinguish in most systems. drones have the ability to move flexibly, hide, camouflage and are highly controllable.
	Triple Distinguish: The challenge is in the event that the three factors meet and the ability of the proposed system to distinguish between them in an efficient manner

- Change dataset image size.
- Apply letterbox.
- Change color space to YCBCR.
- Apply filter.

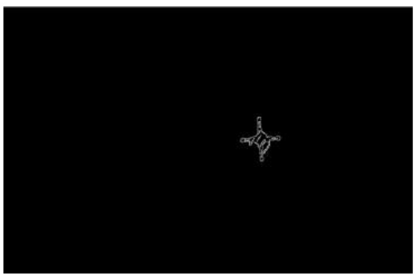
In the proposed system, the image dimension was converted to (512×512) as it's an ideal size, does not cause high processing and training time and it's memory efficient. Letterbox preserves images aspect ratio. It is a very important factor because when it is not implemented, part of the image will be cropped or images will be distorted. That is, its function in the proposed system is to preserve the image's dimension and object shape. that is, when the image has rectangular dimensions in length, it maintains the elongation in length, and if it is rectangular in width, it maintains its elongation in width. The main purpose of the letterbox is to preserve the image from distortion Result of resizing. primary colour space is converted from RGB to YCbCr because it is the native format, which represents colours as combinations of brightness signal and too chroma signal while the RGB repents the colour as combination (Red, Green, Blue) signals and YCbCr helps with using canny Filter. One reason is that images are represented in RGB format is to reduce their size. However, it is not worth storing or transmitting information in this colour space representation, once it has a large bandwidth. Thus, all pixels are converted to YCbCr. The following is a table showing the difference between the original image and the YCbCr range. In the proposed system, (Y) was adopted, due to the effect of the light factor and its cycle, which helps it to improve the quality of detection as the light factor enters. and the final step in Pre-processing proposed system utilized (Canny edge detector), which is an image processing method for detecting edges in an image while suppressing noise. Table 3 shows the steps of pre-processing in the proposed system:

Table 3: The steps of pre-processing in proposed system

Image	**Description**
	Original Size
	512 × 512 note the difference between the two images. is that the original was of rectangular dimensions and the dimensions of the image in the data set were of different sizes and dimensions.
	letterbox: when the image has rectangular dimensions in length, it maintains the elongation in length, and if it is rectangular in width, it maintains its elongation in width. The main purpose of letterbox is to preserve the image from distortion Result of resizing by maintaining its aspect ratio.

6.3 Classification by using YOLO algorithm

YOLOv4 requires annotation for all objects in the dataset, CVAT annotation tool helps a lot to speed up the annotation processes, especially with the automatic annotation tool in CVAT, by Using pre-trained models to have

	RGB changed the primary color space from RGB to YCbCr because it is the native format, which represents colors as combinations of brightness signal and two chroma signals while the RGB is repents the color as combination (Red, Green, Blue).
	Y In the proposed system, (Y) was adopted, due to the effect of the light factor and its cycle, which helps it to improve the quality of detection because of the light factor.
	Cr Ignored.
	Cb Ignored.
	Contouring use Canny filter in order to extract the borders of the shape or object.

	Applying in (Y) image borders from the previous step were used and applied to an image of type Y
	Applying in RGB image detect drones with high accuray in color images.

annotation for our objects. In the end, we made sure all the objects are annotated perfectly since the pre-trained models are not 100% accurate. YOLOv4 in the proposed system uses five layers for training to give the best training results, however, this was not optimal in our case because our dataset is only about 8000 thousand images and we are using a personal machine for training. It was more optimal to use only three layers as the extra two layers added more time cost and little accuracy. We can summarise the process in steps as follows:

- Download and build darknet.
- Download, build and install the tools needed for GPU training and for preprocessing (OpenCV, CUDA, Cudnn)
- Collect datasets from various datasets.
- Install annotation tool used to create labelled images for YOLO (Annotation), those labelled images are used during training and testing.

7 Evaluation

The recommended befuddlement matrix [**?**] evaluation approach will be used. Which will be performed in error rate and will summarize the number of occurrences predicted correctly or incorrectly by a classification model. The following language is routinely used to refer to the counts generated by a confusion matrix:

- **Precision:** It is measured using the standard deviation of a collection of data, while bias is quantified using the difference between the collection's mean and the known value of the object being quantified. Precision is defined as the percentage of records in a group that the classifier has claimed to be a positive class that are really positive, as indicated by the precision equation below [11].

$$Precision(p) = \frac{TP}{TP + FP} \tag{7.1}$$

By applying equation (7.1) to our results we calculate precision:

$$precision = \frac{9591}{9869} = 0.97$$

- **Recall:** Recall quantifies the proportion of positive examples properly predicted by the classification, and its value is identical to the genuine positive rate [26].

$$Recall(r) = \frac{Tp}{TP + FN} \tag{7.2}$$

By applying equation (7.2) to our results we calculate Recall:

$$Recall = \frac{9591}{9652} = 0.99$$

- **F1-Score:** The harmonic mean of accuracy and recall, as well as the following equation, denote F1, and the equation is [9].

$$F1 = \frac{2rp}{r + p} = \frac{2 \times TP}{2 \times TP + FP + FN} \tag{7.3}$$

By applying equation (7.3) to our results we calculate F1-Score

$$F1 - Score = \frac{19182}{19521} = 0.98$$

And the following table 4 shows system result:

Table 4: Result of system

Precision	Recall	F1-Score
0.97	0.99	0.98

And the following figures show different cases of training process:

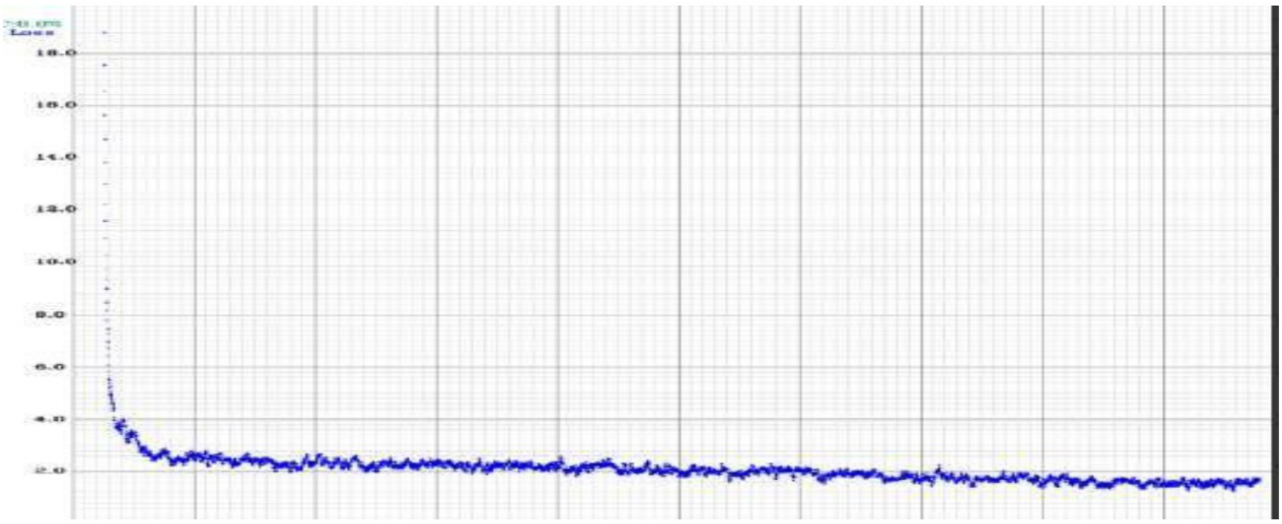


Figure 2: First stage - only drone class. Low accuracy with simple dataset.

8 comparison

Numerous researches on drone detection systems and their conclusions have been published in recent years. The proposed approach's conclusions are compared to those of many other methodologies reported in the literature in this section. The suggested method's detection measurement in the confusion matrix is compared to that achieved in past research in Table 5. According to the mentioned findings, our recommended technique looks to be superior to others.

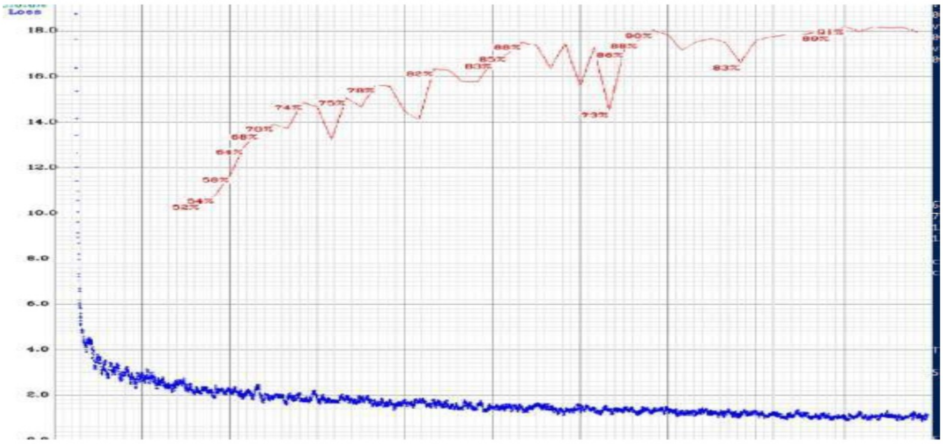


Figure 3: Second stage – model overfitting due to dataset distortion

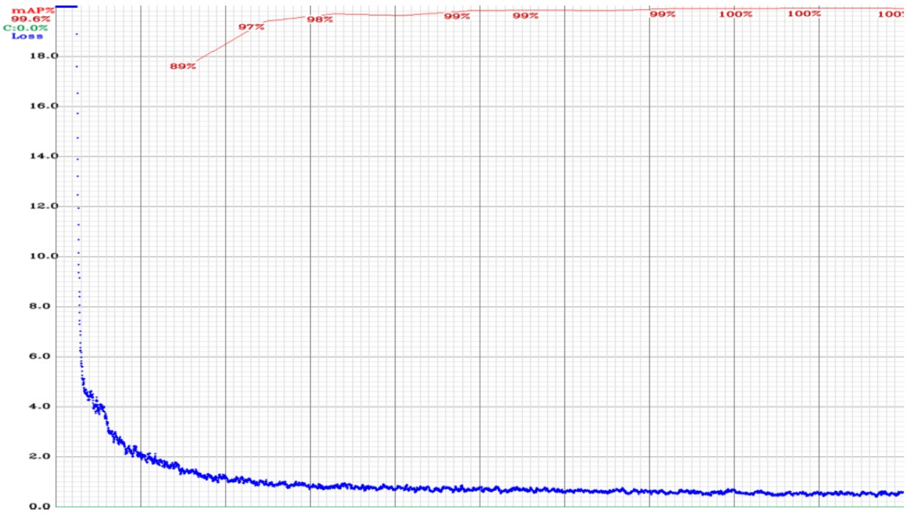


Figure 4: Third Stage – new dataset with letterbox applied.

Figure 5: Detection in a noisy environment.

Figure 6: Plane, Bird Detection.

Figure 7: Detection from distance.

Table 5: Comparison to Previous Studies

NO	Refers	Method	Dataset	work nature	percentage
1	Proposed system	YOLO	Dataset collected manually from standard datasets	Drone, bird and Plane detection using deep learning	0.97
2	2018 [27]	YOLO	Dataset collected manually	Real time Drone detection using deep learning	0.91
3	2019 [25]	YOLO	Dataset collected manually	Drone detection using machine learning in different angle	0.89
4	2020 [21]	YOLO	Dataset collected manually	Real time Drone detection using deep learning	0.89
5	2021 [22]	YOLO	Dataset collected manually	Drone, bird detection using deep learning	0.95

9 Conclusion

Drones will be increasingly employed for communications and security in future smart cities. Drones are commonly employed in military operations. Drone identification is crucial, just as security concerns are. In this study, we provided a unique approach and demonstrated how to deal with drone photographs and extract information from them in order to detect and identify drones from birds and planes. Long-range, where the goal is to detect and distinguish between the three categories in low-quality or high-quality images within Long or near range, with the main goal being to assist in the development of early detection systems that can provide high security for areas where drones are being used illegally. Concerning the three components (birds, planes and drones). In the future, drone detection will be based on increasingly complex models. We'll also focus on extending drone use in the communications industry, and future work will include training the model on night imagery, such as infrared cameras, to detect drones at night.

References

[1] K. Abbasi, A. Batool, M.A. Asghar, A. Saeed, M.J. Khan and M. ur Rehman, *A vision-based amateur drone detection algorithm for public safety applications*, 2019 UK/China Emerg. Technol. UCET **2019** (2019), 1–5.

[2] Y. He, I. Ahmad, L. Shi and KH. Chang, *SVM-based drone sound recognition using the combination of HLA and WPT techniques in practical noisy environment*, KSII Trans. Internet Inf. Syst. **13** (2019), no. 10, 5078–5094.

[3] M.A. Akhloufi, S. Arola and A. Bonnet, *Drones chasing drones: reinforcement learning and deep search area proposal*, Drones **3** (2019), no. 3, 1–14.

[4] M.S. Allahham, M.F. Al-Sa'd, A. Al-Ali, A. Mohamed, T. Khattab and A. Erbad, *DroneRF dataset: a dataset of drones for RF-based detection, classification and identification*, Data Br. **26** (2019).

[5] A. Bochkovskiy, C.-Y. Wang and H.-Y.M. Liao, *YOLOv4: optimal speed and accuracy of object detection*, arXiv preprint arXiv:2004.10934. (2020).

[6] E. Burger and G. Bordacchini, *Global space policies and programmes*, In Yearbook on Space Policy, Springer, Cham. 2019.

[7] Caltech Vision Lab, *C.-U. Birds-200-2011*, http://www.vision.caltech.edu/visipedia/CUB-200-2011.html, (2011).

[8] A. Coates, H. Lee and A.Y. Ng, *An analysis of single layer networks in unsupervised feature learning*, Proc. 14th Int. Conf. Artific. Intell. Statist. (AISTATS), Fort Lauderdale, FL, USA, 2011.

[9] A. Coluccia, A. Fascista, A. Schumann, L. Sommer, A. Dimou, D. Zarpalas, F.C. Akyon, O. Eryuksel, K.A. Ozfuttu, S.O. Altinuc and F. Dadboud, *Drone-vs-bird detection challenge at IEEE AVSS2019*, 16th IEEE Int. Conf. Adv. Video Signal Based Surveillance, AVSS, 2019, pp. 1–8.

[10] A.R. Eldosouky, A. Ferdowsi and W. Saad, *Drones in distress: A game-theoretic countermeasure for protecting UAVs against GPS spoofing*, IEEE Internet Things J. **7** (2020), no. 4, 2840–2854.

[11] J. Gao, C.-D. Lü, Y.-L. Shen, Y.-M. Wang and Y.-B. Wei, *Precision calculations of $B \longrightarrow V$ form factors in QCD*, arXiv preprint arXiv:1907.11092. (2019).

[12] Š. Grác, P. Beňo, F. Duchoň, M. Dekan and M. Tölgyessy, *Automated detection of multi-rotor UAVs using a machine-learning approach*, Appl. Syst. Innov. **3** (2020), no. 3, 1–23.

[13] D. Kinaneva, G. Hristov, J. Raychev and P. Zahariev, *Early forest fire detection using drones and artificial intelligence*, 42nd Int. Conv. Inf. Commun. Technol. Electron. Microelectron. MIPRO 2019 Proc., 2019, pp. 1060–1065.

[14] S. Marathe, *Leveraging drone based imaging technology for pipeline and RoU monitoring survey*, Soc. Pet. Eng. - SPE Symp. Asia Pacific Heal. Safety, Secur. Environ. Soc. Responsib. **2019** (2019).

[15] S. Mayer, L. Lischke and P.W. Woźniak, *Drones for search and rescue*, 1st Int. Workshop on Human-Drone Interaction, 2019.

[16] D. Misra, *Mish: a self regularized non-monotonic activation function*, arXiv preprint arXiv:1908.08681. (2019).

[17] PASCAL 2, *Visual Object Classes Challenge 2011 (VOC2011)*, http://host.robots.ox.ac.uk:8080/pascal/VOC/voc2011/inde 2011.

[18] M. Pawełczyk and M. Wojtyra, *Real world object detection dataset for quadcopter unmanned aerial vehicle detection*, IEEE Access **8** (2020), 174394–174409.

[19] T. Preethi Latha, K. Naga Sundari, S. Cherukuri and M.V.V.S.V. Prasad, *Remote sensing uav/drone technology as a tool for urban development measures in apcrda*, Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. ISPRS Arch. **42** (2019), no. 2/W13, 525–529.

[20] J. Redmon and A. Farhadi, *YOLOv3: an incremental improvement*, arXiv preprint arXiv:1804.02767. (2018).

[21] Q Shi and, J Li , *Object detection of UAV for anti-UAV based on YOLOv4*, IEEE 2nd Int. Conf. Civil Aviation Safety Inf. Technol.(ICCASIT. IEEE, 2020, pp. 1048–1052.

[22] S. Singha and B. Aydin, *Automated drone detection using YOLOv4*, Drones **5** (2021), no. 3.

[23] D. Tezza and M. Andujar, *The state-of-the-art of human-drone interaction: A survey*, IEEE Access **7** (2019), 167438–167454.

[24] E. Unlu, E. Zenou, N. Riviere and P.E. Dupouy, *Deep learning-based strategies for the detection and tracking of drones using several cameras*, IPSJ Trans. Comput. Vis. Appl. **11** (2019), no. 1.

[25] E. Unlu, E. Zenou, N. Riviere and P.E. Dupouy, *An autonomous drone surveillance and tracking architecture*, Electronic Imag. **2019** (2019), no. 15, 1–35.

[26] J.P. Winkler, J. Grönberg and A. Vogelsang, *Optimizing for recall in automatic requirements classification: an empirical study*, IEEE 27th Int. Requir. Engin. Conf., IEEE, 2019, pp. 40–50.

[27] M. Wu, W. Xie, X. Shi, P. Shao and Z. Shi, *Real-time drone detection using deep learning approach*, Lect. Notes Inst. Comput. Sci. Soc. Telecommun. Eng. LNICST **251** (2018), 22–32.

[28] J. Xu, L. Guo, J. Jiang, B. Ge and M. Li, *A deep learning methodology for automatic extraction and discovery of technical intelligence*, Technol. Forecast. Soc. Change **146** (2019), 339–351.

[29] Z. Zhang, T. He, H. Zhang, Z. Zhang, J. Xie and M. Li, *Bag of freebies for training object detection neural networks*, arXiv preprint arXiv:1902.04103. (2019).