Jurnal Kejuruteraan 37(2) 2025: 609-616
https://doi.org/10.17576/jkukm-2025-37(2)-05

Drone Detection in Restricted Areas Using Deep Learning

Humaira Nishat[a,d*], Shakeel Ahmed[b,d] & Devarapally Akanksha[c,d]

[a]*Professor, Electronics and Communication Engineering,*

[b]*Sr. Assistant Professor, Electronics and Communication Engineering,*

[c]*Research Scholar, M.Tech (Embedded Systems) Electronics and Communication Engineering,*

CVR College of Engineering (CVRCE), Telangana, India

**Corresponding author: dr.humairanishat@gmail.com*

Received 2 September 2024, Received in revised form 14 November 2024
Accepted 14 Dicember 2024, Available online 30 March 2025

ABSTRACT

As the use of drones becomes more widespread, the corresponding rise in drone-related intrusions poses a growing threat to public safety and privacy. Traditional anti-drone systems typically rely on radio-frequency sensors for drone tracking. This paper investigates the fusion of deep learning-based detection algorithms with surveillance cameras within the framework of radio-frequency anti-drone systems. The primary aim is to assess the efficacy of contemporary models and training methodologies in achieving precise and real-time drone detection. One of the central challenges addressed in this paper is the detection of small drones at extended distances, coupled with the demand for real-time performance. To overcome the scarcity of small drone datasets, the research team constructed a real-world dataset for comprehensive evaluations. Various iterations of the YOLO (You Only Look Once) models were compared using this dataset, with specific modifications implemented to enhance small object detection. Additionally, the image sources are diversified for training, incorporating bird images to mitigate false positives and enhance model robustness. Among the YOLO models tested, YOLOv5 exhibited superior precision, recall, and F1 score. The work delves into the impact of additional detection layers on precision, recall, and F1 score, revealing trade-offs between these metrics. The inclusion of bird images in the background training process demonstrated improvements in accuracy and recall, underlining the importance of diverse training data. An intriguing finding was observed when excluding extremely small drones and birds from the analysis, resulting in heightened precision but diminished recall. This highlights the delicate balance required in optimizing detection algorithms for different scenarios. The paper also acknowledges the need for further investigation into the generalizability of the proposed approach across various drone types.

Keywords: Drone detection; deep learning; surveillance cameras; anti-drone systems; small object detection; YOLO Models

INTRODUCTION

The growing availability of Unmanned Aerial Vehicles (UAVs), sometimes known as drones, has raised worries regarding airspace security. Recent occurrences in Sweden, when drones intruded on nuclear reactors, government buildings, and other vital sites, have highlighted the necessity for robust countermeasures (Abualigah et al. 2021). Failure to detect and prevent drone incursions could lead to compromised privacy, aircraft collisions, or other security concerns (Jansen, B. USA Today 26 January 2015). Traditional means of detection, such as radar and radio-frequency detection, are utilized in drone tracking systems, although they can be imprecise in areas with signal interference. Skysense AB intends to improve the precision of their radio-frequency-based anti-UAV system by incorporating a Pan-Tilt-Zoom (PTZ) camera for real-time tracking and visual confirmation of drone identification. The public's interest has been piqued by the introduction of drones, which are remote-controlled flying machines

(Cruzan et al. 2016). However, due to their easier accessibility, there are now worries that they may be used improperly in restricted regions like delicate ones. This abuse may result in several problems, such as privacy violations and potential accidents. There have been attempts to use specialized sensors and cameras to find such unlicensed drones, but the results have been patchy. Imagine if computers were able to reliably recognize drones on their own. This idea of "deep learning," in which computers learn to grasp things on their own, is used (Chen et al. 2017). This method shows promise for accurate drone detection, especially in locations where they provide a danger.

This study explores the idea of using deep learning algorithms to detect drones in off-limits regions. The main goal is to create software that can analyze photos or videos taken by cameras and correctly detect the presence of a drone. The goal is to develop intelligent software that runs quickly and produces reliable results. This paper describes the investigation into the use of deep learning for drone detection in constrained environments. The papers explain how data is collected to train these programs, describe the types of complex computer models used, and assess how well they work. The difficulties experienced when working on this work, such as making sure the algorithms can distinguish between drones and other things while still being effective are also discussed. In summary, the goal is to use deep learning to build a whole system that protects limited regions from pesky drones. For crucial areas like airports and significant buildings, where ensuring safety and security is of the utmost importance, this development could be invaluable.

The work is based on YOLOv5 which enables an accurate drone detection system.

RELATED WORK

Detecting drones can be achieved through various methods, including radar, acoustic sensors, deep learning, and more. Enhanced deep learning architectures like Fast RCNN (Region-Based Convolutional Neural Network) have been employed for detection tasks (Girshick 2015). These experiments typically involve several stages where feature maps serve as input and yield object proposals as output. Fast RCNN evaluates these proposals in subsequent stages to achieve object detection. In a study by G. Chandan, et. Al 2018, the authors investigated the efficacy of SSD for object detection, achieving results with 99% accuracy and a significant confidence level. Despite its improved accuracy, SSD exhibits relatively slower detection speeds compared to other models, indicating the need for further research in object detection methodologies. Some experiments utilize cameras, particularly PTZ (pan-tilt-zoom) cameras, for real-time object detection. After conducting comparative analyses based on speed and accuracy using various models, Faster RCNN Inception ResNet emerged as superior in terms of the F1 score as given by J. Park et al. 2017. MobileNet and SSD demonstrate the fastest detection speeds, while YOLOv2 strikes the best balance between speed and accuracy among the models evaluated.

Different strategies can be employed for drone or object detection (Seidaliyeva et. Al 2020). For instance, MobileNet V2 CNN model with masked images has been utilized, where background subtraction is performed to isolate objects. This challenge was addressed through a two-stage process involving detection and classification. Notably, the detection speed is influenced by the rate of background change, prompting the inclusion of more bounding boxes into the classifier to enhance detection speed.

Integrating machine learning techniques (Svanström et al. 2021) with visible, thermal, and audio sensors has shown promising results in real-time drone detection. The dataset comprises various audio recordings capturing helicopters, drones, and background noises. To enhance the focus on moving objects, a fish-eye lens camera was utilized. Notably, the F1 scores between visible video detectors and infrared detectors showed similarities, while an audio classifier achieved a high F1 score in their experiments. Addressing these challenges, the researchers considered selected YOLO models for future work, anticipating that YOLOv3 would yield more efficient results in detecting small objects like birds and drones.

To address the challenge of detecting small objects, a single shot detector YOLO model was proposed. To mitigate the lack of specific data, an algorithm was devised to extract the dataset and integrate it with background-subtracted images, resulting in accurate bounding box predictions. However, the model exhibited some defects, particularly in cases where birds were in close proximity to drones, resulting in bounding boxes encompassing both the bird and the drone as given by Aker et al. 2017.Creating a reliable drone monitoring system poses several challenges (Chen et al. 2017) including the limited availability of drone images for training datasets. Augmentation techniques were employed to effectively manage this issue. To track small objects in clustered environments, the authors utilized residue information from images and trained a CNN-based tracking model. The MDNet (Multi-Domain Convolutional Neural Networks) object tracker was also employed to reduce complexity during testing in online environments. An integrated system capable of both detection and tracking was implemented, outperforming

individual models used for tracking and detection separately due to its ability to re-initialize tracking bounding boxes when objects are lost.

PROPOSED TECHNOLOGY

Drone detection in confined regions necessitates precision, speed, and adaptability. To overcome this issue, the incorporation of deep learning models has gained traction. These models, inspired by the neural networks of the human brain, excel in recognizing patterns and features in data, making them well-suited for object recognition tasks such as detecting drones in pictures or videos (Kong et al 2021). For picture or video processing, Convolutional Neural Networks (CNN) algorithm is used. A large collection of drones, photos are trained for distinguishing the characteristics of drones and other similar objects like birds in the sky. This input is given to deep learning algorithms to detect drones with noisy photographs of drones taken from various angles.

DEEP LEARNING MODELS

Introduction to Deep learning models, notably Convolutional Neural Networks (CNNs), have excelled at visual identification tasks. These models are made up of layers that gradually learn to recognize increasingly abstract elements in the input data. Because of this hierarchical learning, the models can catch subtle information and relationships inside photos, making them effective tools for drone identification. Model Selection: Several deep learning architectures can be used for drone detection, with YOLO (You Only Look Once) and Faster R-CNN (Region-based Convolutional Neural Network) being two popular options.

1. YOLO: YOLO is a real-time object detection system that predicts object classes and bounding box coordinates from input photos in real time. Its efficiency stems from the fact that it can process a whole image in a single pass, making it ideal for real-time applications.
2. Faster R-CNN: This model proposes regions of interest (possible item locations) first, and then classifies these regions. Although it is slower than YOLO, it is more accurate and is frequently used in complex settings.
3. Model Adaptation: Drone detection necessitates the following model modifications: Adjusting the input image scale allows the model to recognize small objects better because drones might appear small in photographs.
4. Anchor Boxes: These predefined forms assist the model in better localizing objects, which is necessary for effective drone identification. Transfer learning entails beginning with a pre-trained model and fine-tuning it for the unique objective of drone identification. This speeds up training and improves performance. Drone detection technology is critical in protecting restricted regions from unlawful incursions. A comprehensive strategy to properly recognize drones in real-time is provided by combining YOLO object detection with blob detection approaches.

UNDERSTANDING YOLO OBJECT DETECTION YOLO

YOLOv5 is a real-time object detection algorithm employing deep convolutional neural networks (CNNs) to detect and classify objects within images and video streams swiftly and accurately. Specifically designed for speed, accuracy, and versatility, it emerges as an optimal choice for drone detection. This iteration, the fifth in the series, operates within the PyTorch framework, distinguishing itself from previous versions. In addition to YOLO, there exist other models of one-stage detectors such as RefineNet, SSD, and MobileNet. YOLO, however, excels in certain scenarios, as discussed previously. Fast R-CNN encounters inefficiency due to its separate module for area generation, independent of CNN. YOLOv5 operates on the principle of dividing an input image into a grid of cells and utilizing CNNs to predict object presence and location within each cell. The algorithm leverages anchor boxes and anchor points to enhance object detection accuracy, especially in complex and cluttered environments. Anchor boxes represent pre-defined boxes of various shapes and sizes, approximating object location and size, while anchor points refine object locations by allowing adjustments to anchor box position and size to match objects within the image.

To utilize YOLOv5 for drone detection, the algorithm necessitates training on a substantial dataset comprising images containing drones and other objects. During training, the model learns distinctive drone features and discerns them from other scene elements. Post-training, the model analyzes video streams from cameras or sensors to detect drones in real-time.

Although YOLOv5 bears resemblance to YOLOv4 with minimal alterations, such as YOLOv4's use of the Darknet framework compared to YOLOv5's reliance on

PyTorch, and the differing configuration files (.cfg for YOLOv4, .yaml for YOLOv5), YOLOv5 stands out for utilizing a single neural network to process entire images.

The YOLOv5 architecture comprises three components: Backbone, Neck, and Head. The Backbone receives the input image or video stream and processes it to extract high-level features. The Neck refines these features and prepares them for the prediction layer, employing the Path Aggregation Network (PANet) and Spatial Pyramid Pooling (SPP) to enhance information flow and speed. The Head, also known as the prediction layer, utilizes anchor boxes and anchor points to predict object location, size, and category, producing the final output. Figure 1 gives the architecture of YOLOv5 and Figure 2 gives the drone detection attributes.

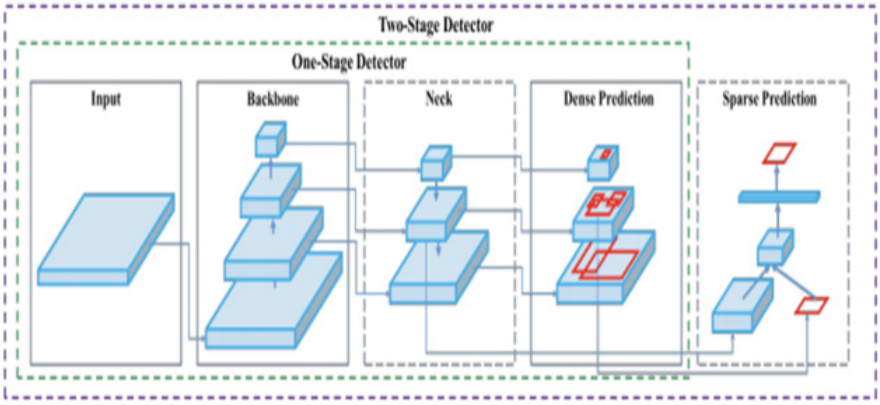

FIGURE 1. Architecture of YOLOv5

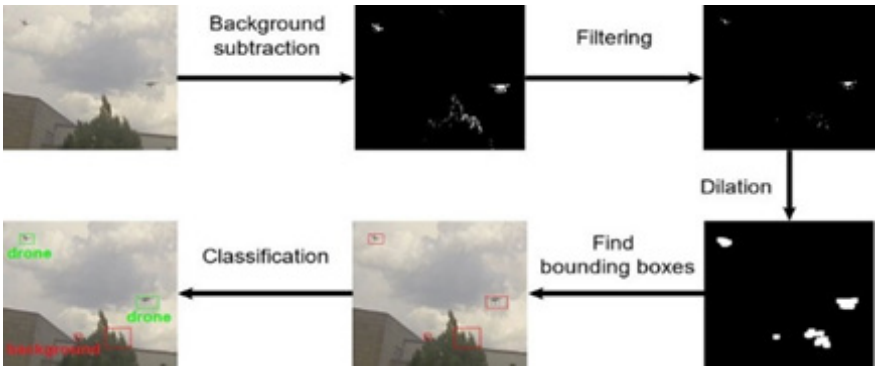

FIGURE 2. Drone detection attributes

DATASET

DATA COLLECTION

Data stands as a critical component for training models, with insufficient amounts leading to potential underfitting. To address this, we gathered 1000 unlabeled images from diverse sources including Google, Roboflow, Instagram, and open-source materials featuring drone imagery. The dataset comprises various drone images, each exhibiting distinct dimensions and multi-rotored structures, with images of varying sizes.

DATA LABELING

Utilizing the Labeling tool, we undertook the task of labeling the collected images. Following the acquisition of the LabelImg software, we initiated the labeling process by opening drone-containing images and creating bounding boxes in YOLO format, saving the annotations in text files. The LabelImg format follows "<object class-ID> <x_center> <y_center> <width> <height>", primarily facilitating bounding box assignments crucial for training purposes. Subsequently, we partitioned the dataset into train and test subsets.

TRAINING THE DATA

For training purposes, we employed the YOLOv5 model, a choice further elucidated in subsequent sections of this paper. Following the recommendations of the original YOLOv5 authors, we trained an optimized model, configuring parameters such as initial learning rate (0.05), final learning rate (0.005), weight_decay (0.0005), momentum (0.937), with a batch size of 8 and 100 epochs. The dataset was divided into training and validation sets, allocated with proportions of 80% and 20%, respectively.

EXPERIMENTAL SETUP

The entirety of our experiment was conducted on a Linux PC via Google Colab, harnessing its GPU (Tesla) capabilities. The operating system utilized was Ubuntu OS. The YOLOv5 model underwent training for approximately 500 iterations. A prediction is deemed a true positive if the Intersection over Union (IOU) value surpasses 0.5. From data collection and model training through real-time deployment and response actions, this flowchart given in Figure 3 shows each stage of the drone detection process utilizing deep learning.

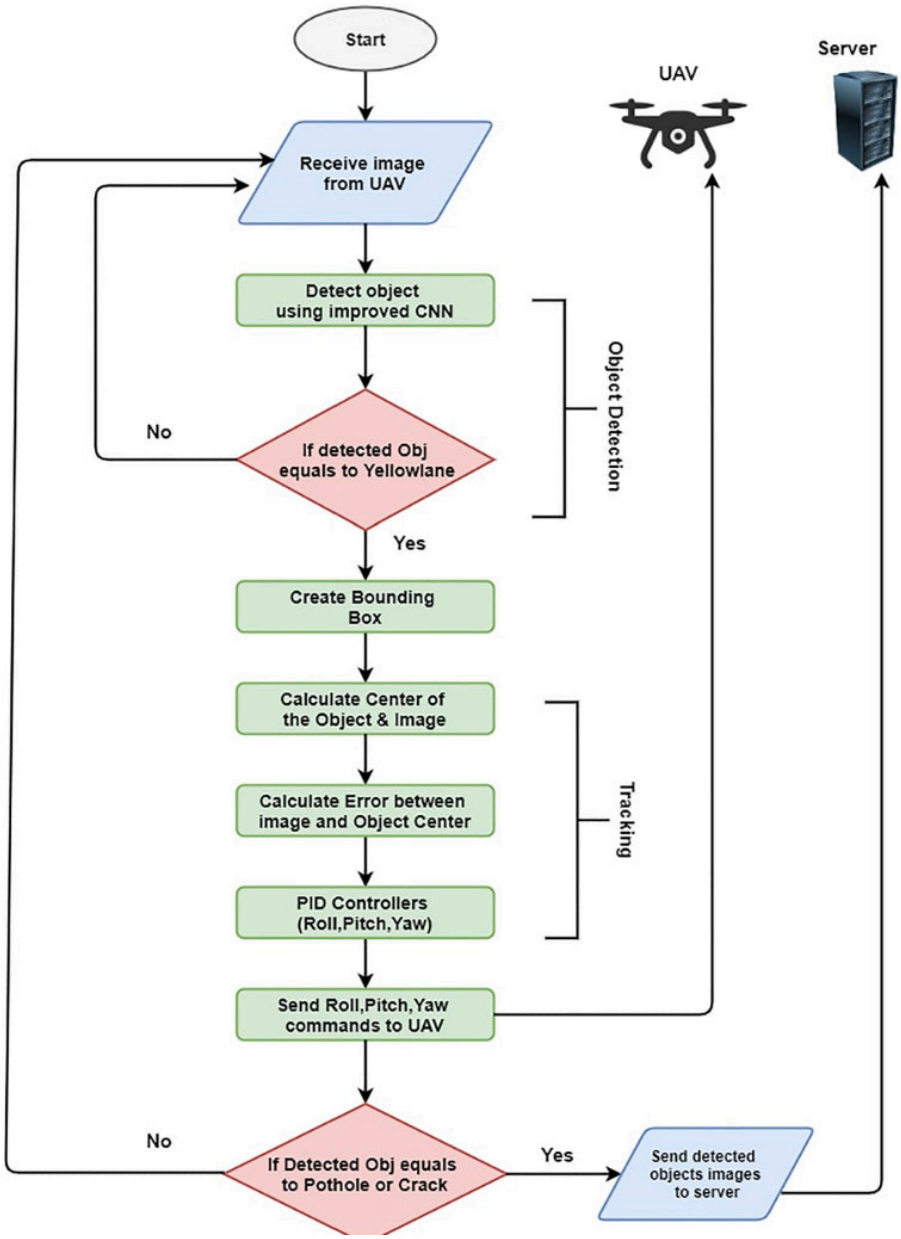

FIGURE 3. Flowchart

EVALUATION METRICS

Evaluation metrics for YOLOv5 include several key concepts such as IOU (Intersection over Union), ROC curves, F1 scores, Precision, and Recall.

INTERSECTION OVER UNION (IOU)

The Intersection Over Union measures the accuracy of the model by evaluating the overlap between the Ground Truth and Prediction regions. A higher IOU indicates better model performance. Additionally, IOU aids in eliminating duplicate bounding boxes for the same object. The IOU values range from 0 to 1, where 0 signifies no overlap and 1 indicates perfect overlap. By setting an IOU threshold, we can determine whether a prediction is True Positive, True Negative, False Positive, or False Negative.

For instance, if the prediction IOU is 0.7 and the threshold is set at 0.98, the detection is classified as a False Positive.

True Positive (TP): The model correctly predicts the positive class.
True Negative (TN): The model correctly predicts the negative class.
False Positive (FP): The model incorrectly predicts the positive class.
False Negative (FN): The model incorrectly predicts the negative class.

IOU = TP/ (TP+FP+FN)

THE ROC CURVE

It is the short for Receiver Operator Characteristic, illustrates the relationship between True Positive Rate and False Positive Rate. It offers insight into the performance of a classification model across various classification thresholds.

True Positive Rate (TPR) = TP/(TP+FN)
False Positive Rate (FPR) = FP/(FP+TN)
Accuracy = (TP+TN)/Total Samples

PRECISION

It reflects how effectively the model identifies a specific object, particularly useful when the cost of false positives is high. For instance, in the case of detecting a fruit like a mango, precision measures how accurately the model identifies mangoes.

Precision = TP/(TP+FP)

Recall: It assesses the model's performance in capturing all instances of a particular class. Using the mango detection example, recall evaluates how well the model identifies the entire mango class.

Recall = TP/(TP+FN)

In comparison, recall tends to be a better measure than precision.

F1 SCORE

It is an overall accuracy metric, amalgamates precision and recall. A high F1 score signifies minimal false positives and false negatives, indicating accurate identification of serious threats while minimizing false alarms. An F1 score of 1 denotes a perfect method, while 0 indicates a complete failure of the model.

F1 Score=2*(Precision*Recall)/(Precision+ Recall)

RESULTS

The precise identification and localization of drones within the designated region is the outcome of drone detection in restricted zones using deep learning.

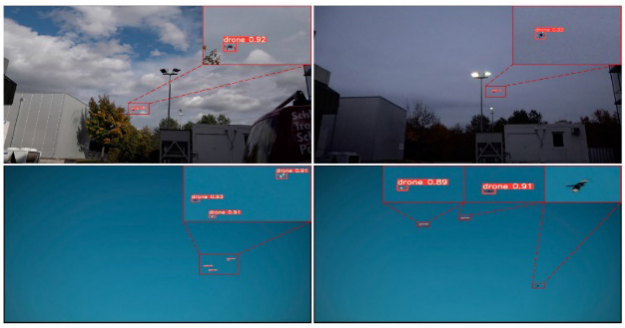

FIGURE 4. Drone Detection

To assess the performance of the models, several metrics such as precision, recall, and F1-score are utilized. The evaluation of these metrics is conducted using a test dataset, as the model has been divided into training and testing sets. Fig. 4 gives the images of accurate drone detection.The maximum F1 value achieved is 0.92, aligning with a confidence value that maximizes recall up to 0.94. Typically, a higher F1 score and confidence value are preferred.Table.1 gives a comparison of models with different number of images and the accuracy percentage. Numerous authors have conducted experiments addressing

drone detection challenges utilizing multiple models, as discussed by A. Pansare et al 2022. In one instance, drone detection was performed using MobileNetSSD with 1800 images, resulting in an accuracy of 0.788 and recall of 0.524. Another study by A. J. Garcia et al. 2020, employed Faster R-CNN (resnet101) trained on a dataset from the SafeShore project, achieving an accuracy of 0.934.

TABLE 1. Comparison of Models

S.No	Model	F1 Score	Recall	Accuracy
1	SSD	0.73	0.524	0.788
2	Faster R-CNN	0.62	0.64	0.67
3	Mask R-CNN	0.7	.642	0.812
4	YOLOv3	0.62	0.63	0.61
5	YOLOv4	0.79	0.68	0.74
6	YOLOv5	0.91	0.94	0.92

In a different investigation, Mask R-CNN with MobileNet Architecture emerged as the preferred model for drone identification. Training conducted on 1000 drone images yielded an accuracy of 0.812 and recall of 0.642 as given by Mubarak et Al 2022. Additionally, utilizing YOLOv4 (Singha et al. 2021) with 2395 images (479 birds, 1996 drones) produced an accuracy of 0.743, recall of 0.68, and F1-score of 0.79. Furthermore, YOLOv5 attained commendable accuracy, F1 Score and recall of 0.92, 0.91 and 0.94, respectively.

CONCLUSIONS

Comparing object detector models entails striking a balance between speed and accuracy, posing a challenge. Understanding the priority of the application—whether it values speed or accuracy more—is crucial. While accuracy serves as the primary metric for assessing a model's performance on a dataset, the significance of speed in detection cannot be overlooked. Object detection models like SSD exhibit significantly faster speeds compared to R-CNN, albeit with lower performance in detecting small objects. Faster R-CNN and YOLO, on the other hand, both leverage CNN at their core but differ in their frameworks. Faster R-CNN employs ROI pooling for detection, which is notably faster than its predecessors. Conversely, YOLO conducts detection and classification simultaneously, rendering it a more efficient model for identifying drones from images or videos.

FUTURE SCOPE

In the near future, drone detection technology will see significant advancements to ensure the safety and security of various environments. Enhanced tools and systems will swiftly identify unauthorized drones, seamlessly integrating with existing infrastructure like airports and security networks. Machine learning and computer algorithms will become pivotal in accurately distinguishing drones from other objects, minimizing false alarms. Adherence to regulations and privacy standards will be paramount, balancing security with individual rights. To counter rogue drones, safe and effective measures will be developed, potentially with autonomous response capabilities.

Further investigation is required to implement the proposed algorithms on the different types of drones as per their sizes- very small drones (nano with dimensions between 1 and 50 cm for military applications), small drones (indoor for inspections, recreation, and photography with dimensions between 50cm to 2 m). As per the range- very close range (upto 5 km), close range (644 km), short range (utpto 150 km) and large range (more than 644 km), also, on various types of drones based on power sources, weight, applications, equipment, and motors used.

DECLARATION OF COMPETING INTEREST

None

REFERENCES

Abualigah, L., Diabat, A., Sumari, P., & Gandomi, A. H. 2021. Applications, deployments, and integration of Internet of Drones (IoD): A review. *IEEE Sensors Journal* 21(22): 25532-25546. https://doi.org/10.1109/JSEN.2021.3114266

Jansen, B. 2015. Drone crash at white house reveals security risks. *USA Today,* 26 January. https://www.usatoday.com/story/news/2015/01/26/drone-crash-secret-service-faa/22352857/.

Cruzan, M. B., Weinstein, B. G., Grasty, M. R., Kohrn, B. F., Hendrickson, E. C., Arredondo, T. M. & Thompson, P. G. 2016. Small unmanned aerial vehicles (micro-UAVs, drones) in plant ecology. *Appl Plant Sci.* 4(9): 1600041. doi: 10.3732/apps.1600041. PMID: 27672518; PMCID: PMC5033362.

Chen Y, Aggarwal P, Choi J and Kuo C. J, A deep learning approach to drone monitoring, 2017 *Asia-Pacific Signal and Information Processing Association Annual Summit and Conference* (APSIPA ASC), Kuala Lumpur, Malaysia, 2017, pp. 686–691, doi: https:// doi.org/10.1109/APSIPA.2017.8282120.

Girshick, R. 2015. Fast r-CNN, *Proceedings of the IEEE International Conference on Computer*
Vision, pp. 1440–1448, 2015.

Chandan, G., Jain, A., Jain, H. & Mohana. 2018. Real time object detection and tracking using deep learning and OpenCV, 2018. *International Conference on Inventive Research in Computing Applications* (ICIRCA), Coimbatore, India, 2018, pp. 1305–1308. doi: https://doi.org/10.1109/ICIRCA.2018.8597276.

Park, J., Kim, D. H., Shin, Y. S. & Lee, S. -h. 2017. A comparison of convolutional object detectors for real-time drone tracking using a PTZ camera, 2017. *17th International Conference on Control, Automation and Systems* (ICCAS), Jeju, Korea (South), 2017, pp. 696–699, doi: https://doi.org/10.23919/ICCAS.2017.8204

Seidaliyeva, U., Akhmetov, D., Ilipbayeva, L. & Matson, E.T. 2020. Real-time and accurate drone detection in a video with a static background. *Sensors* 20: 3856. doi: https://doi.org/10.3390/s20143856.

Svanström, F. 2020. Englund C and Alonso-Fernandez F, real-time drone detection and tracking with visible, thermal and acoustic sensors, 2020. *25th International Conference on Pattern Recognition* (ICPR), Milan, Italy, 2021, pp. 7265–7272, doi: https://doi.org/10.1109/ICPR48806.2021.9413241.

Aker, C. & Kalkan, S. 2017. 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, doi: https://doi.org/10.1109/AVSS.2017.807853

Chen, Y., Aggarwal, P., Choi, J. & Kuo, C. -C. J. 2017. A deep learning approach to drone monitoring, 2017 *Asia-Pacific Signal and Information Processing Association Annual Summit and Conference* (APSIPA ASC), Kuala Lumpur, Malaysia, 2017, pp. 686–691, doi: https://doi.org/10.1109/APSIPA.2017.82820

Kong, P. Y. 2021. A survey of cyberattack countermeasures for unmanned aerial vehicles. *IEEE Access* 9: 148244–148263.

Pansare, A. et al. 2022. Drone detection using YOLO and SSD a comparative study, 2022. *International Conference on Signal and Information Processing* (IConSIP), Pune, India, 2022, pp. 1–6, doi: https://doi.org/10.1109/ICoNSIP49665.2022.10007489.

Garcia, A. J. Min Lee, J. & Kim, D. S. 2020. Anti-Drone system: A visual-based drone detection using neural networks. *2020 International Conference on Information and Communication Technology Convergence* (ICTC), Jeju, Korea (South), 2020, pp. 559–561, doi: https://doi.org/10.1109/ICTC49870.2020.9289397.

Mubarak, A. S., Vubangsi, M., Al-Turjman, F., Ameen, Z. S., Mahfudh, A. S. & Alturjman, S. 2022. Computer vision based drone detection using mask R-CNN. *2022 International Conference on Artificial Intelligence in Everything* (AIE), Lefkosa, Cyprus, 2022, pp. 540–543, doi: https://doi.org/10.1109/AIE57029.2022.00108.

Singha, S. & Aydin, B. 2021. Automated drone detection using YOLOv4. *Drones* 5(3): 95. September 2021, doi: https://doi.org/10.3390/drones5030095.