*International Journal of Engineering Applied Sciences and Technology, 2022*
**Vol. 7, Issue 3, ISSN No. 2455-2143, Pages 11-19**
Published Online July 2022 in IJEAST (http://www.ijeast.com)

# DETECTION, TRACKING AND CLASSIFICATION OF ROGUE DRONES USING COMPUTER VISION

Dr. SN Omkar
Chief Research Scientist,
Indian Institute of Science,
Bengaluru, Karnataka, India

Nikhil Asogekar
Department of Aerospace,
Punjab Engineering College,
Chandigarh, India

Sudarshan Rathi
Department of Aerospace,
Indian Institute of Science,
Bengaluru, Karnataka, India

***Abstract*—The increase in the volume of UAVs has been rapid in the past few years. The utilization of drones has increased considerably in the military and commercial setups, with UAVs of all sizes, shapes, and types being used for various applications, from recreational flying to purpose-driven missions. This development has come with challenges and has been identified as a potential source of operational disruptions leading to various security complications, including threats to Critical Infrastructures (CI). Thus, the need for developing fully autonomous anti-UAV Defense Systems (AUDS) hasn't been more imminent than today. To attenuate and nullify the threat posed by the UAVs, either deliberately or otherwise, this paper presents the holistic design and operational prototype of drone detection technology based on visual detection using Digital Image Processing (DIP) and Machine Learning (ML) to detect, track and classify drones accurately. The proposed system uses a background-subtracted frame difference technique for detecting moving objects partnered with a Pan-Tilt tracking system powered by Raspberry Pi to track the moving object. The identification of moving objects is made by a Convolutional Neural Network (CNN) system called the YOLO v4-tiny ML algorithm. The novelty of the proposed system lies in its accuracy, effectiveness with low-cost sensing equipment, and better performance compared to other alternatives. Along with ease of operations, combining the system with other systems like RADAR could be a real game-changer in detection technology. The experimental validation of the proposed technology was justified in various tests in an uncontrolled outdoor environment (in the presence of clouds, birds, trees, rain, etc.), proving to be equally effective in all the situations yielding high-quality results.**

*Keywords*—**Anti-UAV Defense Systems, Digital Image Processing, Convolutional Neural Network, YOLO, POT.**

## I. INTRODUCTION

The development of Unmanned Aerial Vehicles (UAVs), also known as drones, has been rapid in recent years. Both commercials, as well as military applications, have increased considerably. Various companies like Uber, Amazon, etc., are pushing forward to use drones as service providers for packages and food in a commercial setup. At the same time, the military application includes warfare, identifying vulnerable areas prone to risks, and their mitigation.

Despite attracting wide attention in diverse civil and commercial applications, UAVs pose several threats to airspace safety that may endanger people and property. While such threats can be highly diverse regarding the attackers' intentions and sophistication, ranging from pilot unskillfulness to deliberate attacks with unmanned aerial vehicles, they all can produce severe disruption and cause menace.

The first-ever drone attack happened in the Indian Air Force base, Jammu, India, on 27th June 2021, with two drones dropped an IED packed with high explosives. The US did a lot of drone strikes in Pakistan, Yemen, and Somalia between the year 2010 to 2020. The reported data shows 14,040 minimum confirmed strikes with 8858 to 16901 total people killed.

**International Journal of Engineering Applied Sciences and Technology, 2022**
**Vol. 7, Issue 3, ISSN No. 2455-2143, Pages 11-19**
Published Online July 2022 in IJEAST (http://www.ijeast.com)

About 910 to 2200 civilians and 283 to 454 children were killed in these attacks. [1]
To mitigate and neutralize the threat posed by the misuse of UAVs against deliberate malicious and inadvertent activities, this paper presents a complete design of a Long-Range, Fully Autonomous Drone Detection Platform. Considering the requirements of a system that will automatically detect, track, and classify UAVs in a critical environment, the research in this paper presents a complete system that consists of both hardware and software components. The originality and elegance of the proposed approach lie in the convergence and confluence of hardware and software components efficiently and effectively for the accurate localization of intruder objects.

The rest of the paper is organized as follows. Literature Reviewis explained in section II. The proposed method is explained in section III. Experimental results are presented in section IV. Concluding remarks are given in section V.

## II. LITERATURE REVIEW

With precise computer-controlled moments, variable speed, and maneuvering capabilities, it has become essential to predict the path of RPAS (Remotely piloted aircraft system) and APAS. With its small size (actual and apparent in-camera) and its resemblance and similarity to that of other aerial objects like aeroplanes, birds create challenges in automatic detection and accurate localization in real coordinates in space. Various methods have been proposed and implemented to solve this auto-detection and tracking problem, such as RF, GPS, Radar, Acoustic, Vision-based, etc. While these systems can identify moving objects, tracking and classification are acute problems they face. Since it is crucial to recognize the malicious and harmful drones, Vision-Based Systems are proposed and explored in this literature review.
Since the advent of fighter and commercial flights, detection has significantly improved safety and enemy intention. With the power of flight coming into the remotes and powerful UAVs becoming an ever-increasing part of the modern world, especially in the past five years, actively detecting, classifying, and tracking has become significant.
Various techniques have been employed in detecting Flying Aerial Vehicles, like RADAR, Acoustic, Computer Vision, and RF Based detection. But never has there been the need to implement these techniques localized and in mass than now. With the boom in the number of drones being produced and used, the need for robust, efficient, and cost-effective solutions for the Detection and Tracking of Drone is at its peak.
With the ever-improving sophistication in the camera hardware, the potential of the Vision-Based Detection system is immense. Better Algorithms paired with high-quality hardware have made Vision-Based Detection of objects one of the founding pillars of autonomous systems. From self-driving vehicles using sense and avoiding face detection, VBD has and will continue to become an inseparable part of technologies.
Much research is going on to enhance the performance of RADAR, RF, Visual, Acoustic, etc., methods [2]. Mukesh et al. [3] have proposed a review of other moving object detection and tracking methods. The significant challenges in detecting moving object are dynamic background, noise in the video, illumination changes, etc. Different classical models are proposed, such as Gaussian mixture background modelling [4] [5] [6], which performs well in some cases. Algorithms such as ViBe [7] are fast but prone to noise in the video. Different pixel-based [8], region-based [9] and texture-based [10] methods are used to model dynamic backgrounds.
Zhang et al. [11] have proposed a new algorithm using canny edge detection to detect camouflaged moving objects. They have used constant zoom to track an object detected in a frame using a PTZ camera. Huang et al. [12] have proposed an ANN (artificial neural networks) model to detect moving objects in dynamic backgrounds. Cao et al. [13] proposed an algorithm for dynamic background and irregular object movements. Bor-HorngSheu et al. [14] have proposed moving object detection using a frame difference algorithm and tracking with the conversion of the pixel into the required pan-tilt angle in 3-D space.
State-of-the-art methods such as Region-based CNN, fast R-CNN [15], and Faster R-CNN [16] are used as object detector and classifier methods in two stages. YOLO [17] and SSD [18] are one-stage object detectors and classifiers.

## III. PROPOSED ALGORITHM

### *A.* Proposed method

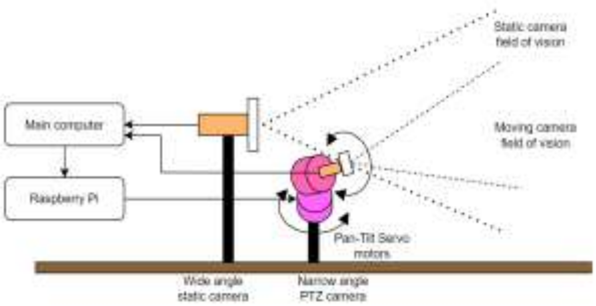


Fig. 1. Proposed System Architecture

Figure 1. shows the proposed drone defence method using two camera systems. A wide-angle static camera is mounted on a static frame. A pan-tilt setup is created using two servo motors. The pan servo motor is fixed on the frame, whereas the tilt servo motor is attached to the rotating head of the pan servo motor. One camera with zoom capability is attached to the rotating shaft of the tilt servo motor. This camera is referred to as a dynamic camera throughout the paper. Both motors are controlled with a raspberry pi. All primary algorithms process on the main computer showing the output results. The static camera act as a detection system, whereas

the dynamic camera tracks the detected object. The flowchart of the work is shown in figure 2.

The detection algorithm is used to detect the moving objects using modified background subtraction and frame difference method from the frames of the static camera. It returns the position of all detected objects in terms of a tuple (x, y), denoting the object's bounding box centre. This data is then transferred to the raspberry pi system using a LAN cable. This pixel information is converted into corresponding pan-tilt activation by the raspberry pi. Pan-tilt servo motors are controlled using this information by the raspberry pi, so the camera attached to it looks toward the object.

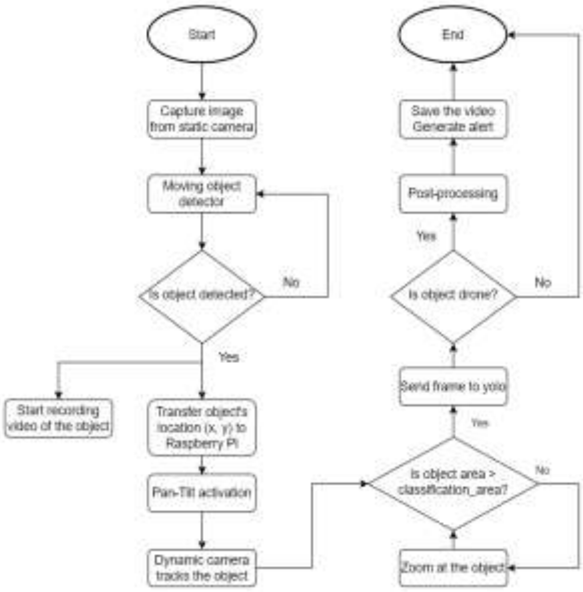


Fig. 2. Proposed System Flowchart

Once the dynamic camera (attached to the servo motors) looks toward the object, the frame difference and connected components analysis system are used to calculate the pixels on target (PoT). For better classification results, we have considered minimum pixels on the target should be 60x40 for a drone of size 30x20 cm. The system will compare the PoT with the required criterion. Zoom towards the object is applied until we get enough PoT.

With enough PoT, the classification algorithm YOLO uses the frame of the dynamic camera for classifying the object as either drone or no-drone. If the object is a drone, the system will keep tracking it, alerting the user. Post-processing will take place on the object to acquire more information such as payloads attached, a rough estimate of distance from the camera, etc. If the object is no-drone, the system will ignore it and reset it to the initial position.

### *B.* Detection

Detection of fast-moving objects in a complex background is a challenging task. A combination of background subtraction and frame difference is used to detect all moving objects present in the frame of the static camera.

When it comes to complex backgrounds containing objects such as trees, grass, etc., different algorithms fail to detect actual moving objects, as shown in figure 3. Here red circle denotes the moving objects detected by the corresponding algorithm.

Here figure 3a shows the original frame of the video. Considering the first video frame as a background frame, the background subtraction method is applied to detect the moving objects, as shown in figure 3b.Figure 3c shows the gaussian mixture model application to classify the pixels as foreground or background. Figure 3d shows the combination of simple background subtraction and gaussian mixture model methods. It reduces the false alarm up to a certain level but not enough.

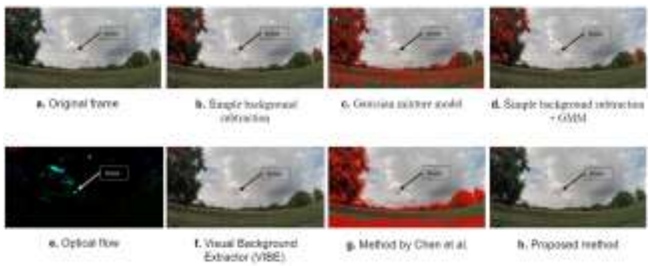


Fig. 3. Different moving object detection method comparison

Figure 3e shows the optical method fails to detect the drone signature alone when it's far away from the camera. As shown in figure 3f, a visual background extractor is applied, giving very few false alarms compared to previous methods. The result of the method proposed by Chen et al. [19] is shown in figure 3g. They have used edge detection and background extraction method to detect moving objects. Figure 3h shows the detected moving objects with the proposed method. As we can see, we get minimal false alarms from the proposed method compared to all other methods. The proposed detection method is shown in figure 4.

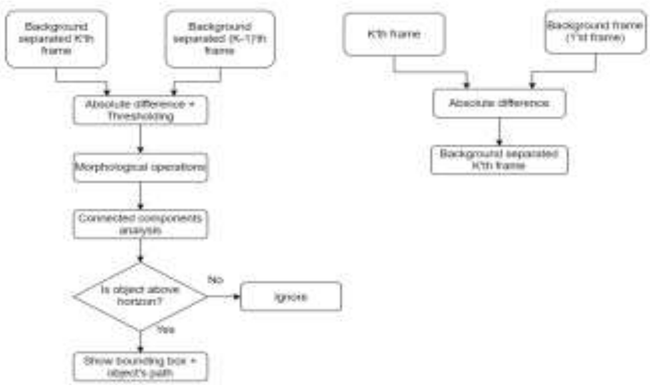


Fig. 4. Flow chart of proposed moving object detection method

**International Journal of Engineering Applied Sciences and Technology, 2022**
**Vol. 7, Issue 3, ISSN No. 2455-2143, Pages 11-19**
Published Online July 2022 in IJEAST (http://www.ijeast.com)

All the frame operations are carried on grey frames to reduce the computation time. Figure 5b shows the colour image of the drone, which is converted into grey, as shown in figure 5c. The RGB frame of the static camera is first converted into the grey frame as per the following equation –

$$\text{Image[gray]} = 0.299 * R + 0.587 * G + 0.114 * B$$

The background frame is the static camera's first frame, as shown in figure 5a. Subtracting each k'th frame from this background frame gives a background-subtracted k'th frame. Such two consecutive background-subtracted frames are subtracted to get the actual moving object, as shown in figure 5d. This step reduces the false alarms generated by slight movements of grass, leaves of a tree, clods etc. since all these generate movement almost at the exact location in all frames. After this, thresholding is used to convert the image into a binary image, as shown in figure 5e.

Frame difference will leave broken gaps in the moving object detection, as shown in figure 5e. Morphological operation closing, a dilation followed by an erosion, can be used to fill these gaps, as shown in figure 5f.

The following equation defines morphological closing –

$$A \cdot B = (A \oplus B) \ominus B$$

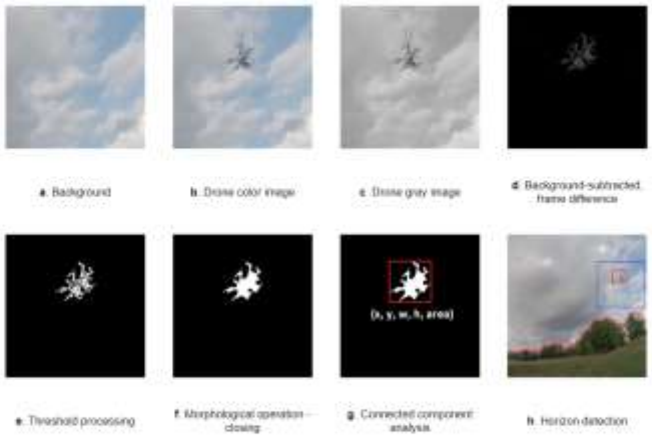


Fig. 5. Proposed Detection Steps

The connected component analysis is applied to get all the moving objects detected by the algorithm, as shown in figure 5g. A connected component is a set of pixels such that any pixel in this set is connected to any other pixel in the same set. The OpenCV function cv2.connectedComponentsWithStats is the most popular and gives the following information about a connected component –

1. The bounding box of the connected component
2. The area (in pixels) of the component
3. The centroid/centre (x, y)-coordinates of the component

This information is helpful for further processing of the components.

To remove the false alarm generated from objects present below the horizon, the horizon mask can be applied, as shown in the figure 5h. Only the moving object detected above the horizon will pass to the next step while ignoring the other. A bounding circle is denoted across the object detected for better visualization along with its path.

### *C.* Tracking

Object Tracking is the process of estimating or predicting the positions of moving objects in a video sequence using tracing algorithms. The object tracking program based on RaspberryPi takes in the initial set of object detection as a bounding box and makes a unique ID for each set of boxes. The algorithm then tracks the moving object as they move in subsequent frames in a video sequence. These tracking methods are readily applied to real-time video streams from the camera with the help of a USB or IP-based protocol. The video is then fed into the algorithm to perform object tracking. Each frame is fed into the tracking algorithm for subsequent detection and tracking, and as a result, high-performance tracking of the object of interest is obtained.

Object tracking is an important part of the overall system and is essential in the localization of the trespassing drone. The tracking method developed and used utilizes OpenCV-Tracking by Detection, which tracks the object with the help of detection. A bounding box is created around the object to be detected, showing the user where the object is in the frame. A self-fabricated model of a Pan-Tilt camera operated by Raspberry Pi and a Personal Computer (PC) Laptop for testing the algorithm is used for tracking the object of interest.

**Pixel to Angle conversion**

The pixel coordinates obtained from the master system are then processed by converting them into corresponding angles with respect to the field of vision of the setup using the equation of straight line approximating the field of vision as cartesian coordinates, as shown in figure 6.

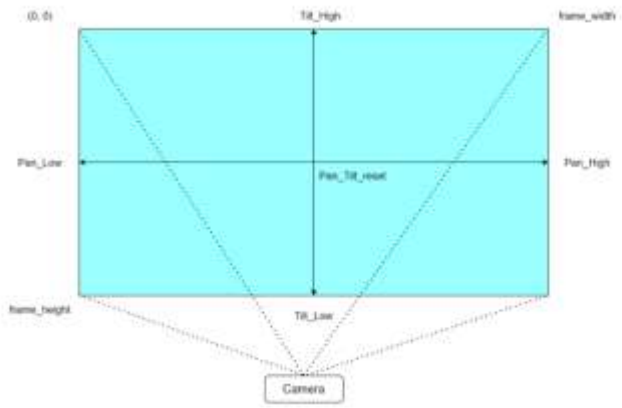


Fig. 6. Camera Field of Vision and its corresponding Pixel Image

The calculation for the pan angle is explained below. A similar procedure is applied for the tilt angle calculations.

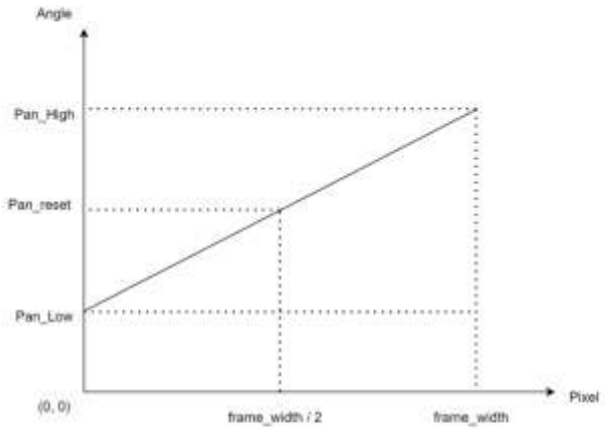


Fig. 7. Pixel to Pan angle conversion

Figure 7 shows the linear conversion of the pixel information into the corresponding Pan angle where the Pan-Tilt system will be facing the camera centre at reset.

$$\text{Pan\_angle} = \frac{\text{Pan\_High} - \text{Pan\_Low}}{\text{frame\_width}} * \text{Pixel\_x} + \text{Pan\_low}$$

**Details of the tracking module**

The structure of the system is such that the input in the form of (x, y) coordinates is passed through from the Server Computer (which computes the Object Detection and Tracking algorithm) into the Raspberry-Pi using Ethernet cable as a Server-Client model incorporated UDP (User Datagram Protocol) which is a lightweight data transport protocol that works on top of IP (Internet Protocol).

The raspberry pi then converts the input coordinates into an angle with the help of the above calculation, with is then converted into PWM that actuates the servo motors to the desired direction in the 3D space where the object of interest lies.

### *D.* Classification

Frame difference and connected component analysis are applied to the frames of the dynamic camera to get the pixels on target. Zooming is applied to the dynamic camera until we get enough PoT that is 60x40. With this condition's satisfaction, the dynamic camera's frame is sent to the classification algorithm YOLO. If the object is classified as a drone, then post-processing is carried out, such as alerting the user, recording the video of the target, etc. Else object will be ignored as a no-drone, and the system will reset to its initial position and start tracking the next object detected.

Different versions of YOLO are available such as v1, v2, v3, v4, and v5, along with their tiny models. The tiny models, modified versions of the main models, give more speed by retaining good classification accuracy. We have used the YOLO v4-tiny model trained on different sets of drone datasets with various backgrounds, illuminations, etc., for classification.

**Datasets**

Different drone datasets are available open-source. We have modified these datasets into the following categories –

**1. Drone close to the camera**– Images from different datasets available are combined so that all drones are very close to the camera.

**2. Drone far from the camera**– Dataset from the drone vs bird detection challenge [20] is used, including drone videos with a high distance between camera and drone.

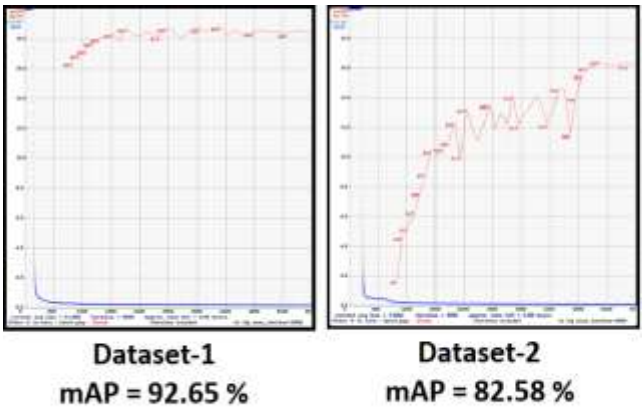


Fig. 8. YOLO v4-tiny training on datasets

As the distance between drone and camera increases, the model's classification accuracy decreases, as shown in figure 8. Apart from a large distance, different illuminations, complex backgrounds, etc., also affect the classification accuracy.

### *E.* System and code architecture

The operational prototype is carefully designed, keeping in mind the constraints and strengths of the algorithm for better overall performance.

The proposed operative model is a two-camera system with a raspberry pi powering the tracking module, as shown in figure 9. The first camera is a Static wide-angle lens being used as an input for the moving object detector algorithm. The second camera acts as a PTZ as it is mounted on the actuation module. The two systems act as a server-client model, with the central computer used as the System server on which all the detection algorithms are compiled and run. In contrast, the raspberry pi is the system client powering the tracking module with control signals.

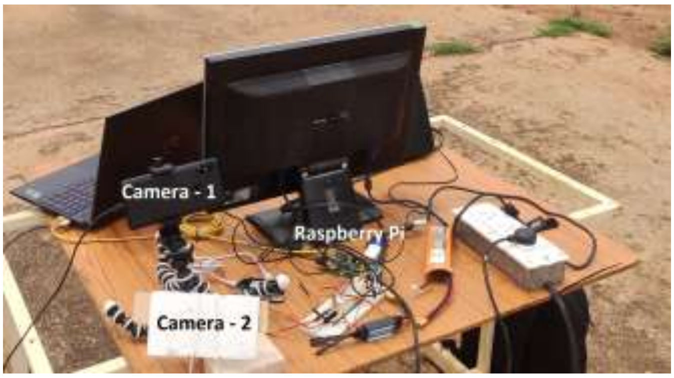

Fig. 9. Experimental Setup

The algorithm takes input from the Static camera to process the live video frame by frame to detect any moving object in the video feed. If any moving object is detected, a bounding box is created around it and is tracked over the feed using tracking by detection. The pixel coordinate of the bounding box's centroid is calculated for each frame and sent to the raspberry pi using an ethernet cable under the UDP protocol. These pixel coordinates are readily received by the raspberry pi as raw data and are processed into the angle and subsequently into signals fed in the form of pulse width through GPIO (General Input Output)pins to the servo motors to actuate the PT module.


Fig. 10. Fabricated PT-Module

The custom-made PT module is powered by two motors and metal moving parts, as shown in figure10. It actuates in real-time to point towards the object detected in the static camera by the system server. PT module enables moving camera with better zoom qualities and tracks the object and zooms into for more clear features for classification. The output of this camera is received on the second Computer as raw footage for input into CNN YOLO for classification as a drone or no drone.

**Code Structure**
The code is made entirely in Python-3, is consistent with convention, and is easy to understand for anyone with prior knowledge of Image Processing and Python. All the executions are made as functions for faster executions, better performance, and a stable system.

As shown in figure 11, the primary detection and classification processing happens on the main computer. Raspberry pi is used to convert pixels into Pan-Tilt angles using the data from the main computer. Visualization and post-processing of the tracked object happen on the secondary computer.

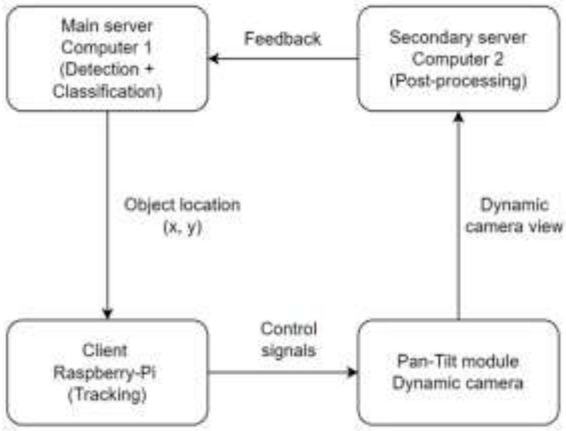

Fig. 11. Code Structure

IV. EXPERIMENT AND RESULT

Initial testing of the detection algorithm was done on the "Bird vs drone detection challenge" dataset [20]. Real-time experiments are carried out in different environments to test the entire system's performance. Different types of testing that are carried out are as follows –

**A. Results oftesting on the dataset [20]**
The detection algorithm is tested on different videos of drones with complex backgrounds, as shown in figures 12 and 13. The green line indicates the ground truth object's path, and the blue line indicates the object's path generated by the algorithm.

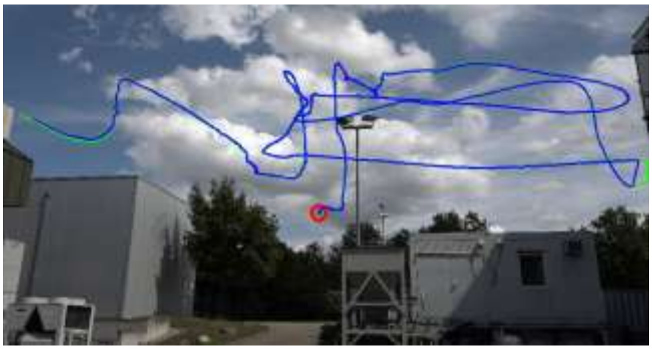
Fig. 12. Results of the dataset video 2019_08_19_C0001_5319_phantom

Fig. 13. Results of the dataset video GOPR5844_002

**Quantitative Comparisons**

The detection algorithm is quantitatively compared with different parameters on different videos of the dataset [20]. The indexes for evaluating the performance of these techniques are as follows.

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

Here,
TP - True positives,
TN - True negatives,
FP - False positives,
FN - False negatives.

| Video name [20] | Binarization threshold = 50 | | Binarization threshold = 20 | |
|---|---|---|---|---|
| | Precision | Recall | Precision | Recall |
| GOPR5842_005 | 0.796 | 0.865 | 0.597 | 0.983 |
| GOPR5844_002 | 0.80 | 0.987 | 0.585 | 1 |
| GOPR5847_003 | 0.753 | 0.976 | 0.821 | 0.997 |
| GOPR5848_004 | 0.760 | 0.924 | 0.364 | 1 |

Table -1 Quantitative comparisons with binarization threshold

The results show that to detect an object at a far distance, we must reduce the binarization threshold, reducing false negatives and improve the recall. But this increases the false positives making the precision worse in some cases.

| Method | Precision | Recall |
|---|---|---|
| Method proposed by Chen et al. [19] | 0.275 | 1 |
| Proposedmethod | 0.796 | 0.865 |

Table -2 Quantitative comparisons with method proposed by Chen et al. [19]

The method proposed by Chen et al. [19] gives no false-negative making recall 1, but it gives many false positives making the precision worse. The proposed algorithm gives both precisions and recall high, as shown in table 2.

**B. Two cameras (static and dynamic (PTZ)) system**

Two camera system with one static and one dynamic camera is considered for a single target. The view of the static camera is shown on the left of the figure as 14a, and the view from the dynamic camera is shown on the right of the figure as 14b. The path of the drone is shown in green colour.

Fig. 14. Result-Two cameras detection and tracking

As soon as the static camera detects a moving object, it sends the object's positional coordinates to the raspberry pi, which controls the PT servo motors so that the dynamic camera can look toward the object.

Fig. 15. Result-Two cameras classification

Zooming towards the object is implemented to get the required pixels on target (60x40 pixels). This image from the dynamic camera is then fed to the YOLO algorithm for classification. The fame number and pixels on target are denoted in figure 15. Fame number 408, we get 2484 PoT; this frame will be sent to YOLO for classification. If the object is a drone, it is denoted on the frame in red colour.

**C. Single (PTZ) camera system**

A single-camera detection, tracking and classification system is made using a camera attached to the PT system. Initially, the system works in detection mode, which uses the frame of the dynamic camera. Once the object is detected, it sends the object's location to raspberry pi so that it can drive the PT servo motors to keep the object at the centre of the camera, as shown in figure 16.

Fig. 16. Result-Single camera detection and tracking

Zooming is applied to get enough PoT, after which the frame of the dynamic camera is sent to the YOLO algorithm for classification. If the object is a drone, it is denoted in the red circle, as shown in figure 17.


Fig. 17. Result-Single camera classification

**D. Multiple objects detection, tracking and classification**

When a static camera detects multiple moving objects, unique IDs are assigned to each object individually. Unique IDs are assigned to each object detected in the first frame. Euclidean distance between objects in two consecutive frames is calculated. The same ID is assigned to the closest pair of objects, whereas new objects get the following ID. If an object moves out of the frame, its ID is erased from the system. Figure 18 shows four objects detected by the static camera, two birds with ID 0 and 4 and two drones with ID 1 and 2. The dynamic camera will track only one ID for which the PoTs are higher, assuming it is nearest to the system. Here drone with ID2 is being followed by the dynamic camera, and other objects are ignored until the classification of the ID2 object.

Fig. 18. Result-Multiple objects detection and tracking

## V.CONCLUSION

This paper illustrates a holistic design and operational prototype of drone detection technology based on image processing to solve the problem of illicit rogue drones trespassing in classified and protected locations. The proposed system accurately detects any moving object in the complex background and varied environments up to but not limited to a distance of 300 m, tracking it in real-time with near-zero latency and classifying it as a potential threat (i.e., drone) or otherwise. The seamless confluence and convergence of smartly designed hardware combined with tailored software (proprietary) have resulted in a high-performance, low-cost alternative to other detection methods. This system aims toward fully automated surveillance and protection solutions for critical infrastructure, emphasizing precision, accuracy, and reduced response time. The culmination of Background subtraction and Frame Difference has improved the standards considerably compared to other similar works, as outlined, and validated in the experimental result section.

The proposed system is an excellent initiative for future work to continue the investigation in the same direction to yield an industry-level robust system. A fusion of this technology with existing technology, viz. RADAR or RF Detection in a certain way would result in a sound and robust system with a high range- high accuracy capability that would be practically impenetrable.

## IV. REFERENCE

[1] Retrieved from The Bureau of Investigative Journalism: https://www.thebureauinvestigates.com/ (accessed on January 2022)

[2] BilalT., and Shoufan A.(2019). Machine learning-based drone detection and classification: State-of-the-art in research.IEEE access 7 (2019): 138669-138682.

[3] Tiwari M., and Singhai R.(2017). A review of detection and tracking of object from image and video sequences.Int. J. Comput. Intell. Res 13.5 (2017): 745-765.

[4] Lee D., Hull J., and Erol B.(2003). A Bayesian framework for Gaussian mixture background modeling.Proceedings 2003 international conference on image processing (Cat. No. 03CH37429). Vol. 3. IEEE, 2003.

[5] Reynolds, Douglas A., Thomas F. Quatieri, and Robert B. Dunn(2000). "Speaker verification using adapted

Gaussian mixture models." Digital signal processing 10.1-3 (2000): 19-41.

[6] Zivkovic Z.(2004). Improved adaptive Gaussian mixture model for background subtraction.Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004.. Vol. 2. IEEE, 2004.

[7] Olivier B., and Droogenbroeck M.(2009). ViBe: a powerful random technique to estimate the background in video sequences.2009 IEEE international conference on acoustics, speech and signal processing. IEEE, 2009.

[8] Michael H.(2002). A framework for high-level feedback to adaptive, per-pixel, mixture-of-gaussian background models. European Conference on Computer Vision. Springer, Berlin, Heidelberg, 2002.

[9] How-Lung E., and Junxian W., (2004). "Novel region-based modeling for human detection within highly dynamic aquatic environment." Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004.. Vol. 2. IEEE, 2004.

[10] Jing Z.(2003). Segmenting foreground objects from a dynamic textured background via a robust kalman filter. Proceedings Ninth IEEE International Conference on Computer Vision. IEEE, 2003.

[11] Xindi Z., and Kusrini K.(2021). Autonomous long-range drone detection system for critical infrastructure safety. Multimedia Tools and Applications 80.15 (2021): 23723-23743.

[12] Shih-Chia H., and Ben-Hsiang D.(2013). Radial basis function based neural network for motion detection in dynamic scenes.IEEE transactions on cybernetics 44.1 (2013): 114-125.

[13] Xiaochun C., Yang L., and Guo X.(2015). Total variation regularized RPCA for irregularly moving object detection under dynamic background. IEEE transactions on cybernetics 46.4 (2015): 1014-1027.

[14] Sheu B., Chiu C., Lu W., Huang C., and Chen W.(2019). Development of UAV tracing and coordinate detection method using a dual-axis rotary platform for an anti-UAV system. Applied Sciences 9.13 (2019): 2583.

[15] Ross G. (2015). Fast R-CNN.IEEE International Conference on Computer Vision (ICCV), 2015, pp. 1440-1448, doi: 10.1109/ICCV.2015.169.

[16] Ren S., He K., Girshick R., and Sun J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. Advances in neural information processing systems 28 (2015).

[17] Redmon J., Divvala S., Girshick R., Farhadi A. (2016). You only look once: Unified, real-time object detection. Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.

[18] Liu W.,Anguelov D., Erhan D., Szegedy C., Reed S., Fu C., and Berg A. (2016). Ssd: Single shot multibox detector. uropean conference on computer vision. Springer, Cham, 2016.

[19] Chen S., Xu T., Li D.,Zhang J., and . Jiang1 S (2016). Moving object detection using scanning camera on a high-precision intelligent holder. Sensors 16.10 (2016): 1758.

[20] Coluccia, A.; Fascista, A.; Schumann, A.; Sommer, L.; Dimou, A.; Zarpalas, D.; Méndez, M.; de la Iglesia, D.; González, I.; Mercier, J.-P.; Gagné, G.; Mitra, A.; Rajashekar, S. Drone vs. Bird Detection: Deep Learning Algorithms and Results from a Grand Challenge. Sensors **2021**, 21, 2824. https://doi.org/10.3390/s21082824