Journal Homepage: - www.journalijar.com

INTERNATIONAL JOURNAL OF ADVANCED RESEARCH (IJAR)

Article DOI: 10.21474/IJAR01/14971
DOI URL: http://dx.doi.org/10.21474/IJAR01/14971

RESEARCH ARTICLE

DETECTION AND TRACKING OF DRONES USING ARTIFICIAL INTELLIGENCE TECHNIQUES - A SURVEY

Deepanshu Phartyal[1] and Mihir Ghooi[2]

1. Research Scholar, MCEME, Secunderabad.
2. Faculty, MCEME, Secunderabad.

……………………………………………………………………………………………………....

Manuscript Info

……………………….
Manuscript History
Received: 25 April 2022
Final Accepted: 27 May 2022
Published: June 2022

Key words:-
CNN, AI-Enabled Technology, ODS, Drone, Detection and Tracking

Abstract

……………………………………………………………….
Enormous mechanisms and mediums are being employed to threaten the defense system and civilians. Drone strikes are one of them, which may refer to the unloading of explosions and supervision as such. So, detecting and tracking the Drone could be a viable solution for any organization to tackle the aerial threat challenges and secure the environment from malicious activities. Thus the present research discusses the current paradigm of Drone strikes, challenges and solutions to deal with such security concerns. The present article aims to examine the current status of Drone Detection critically, and the applicability and advancement of Artificial Intelligence enabled technology. The present research article also explores the working mechanism of the Object detection system and Convolutional Neural Network (CNN) on the ground level to help future researchers gain knowledge. The paper also explains the maximum reach of accuracy achieved by the various model and algorithms so that a new benchmark can be defined.

……………………………………………………………………………………………………....

Introduction:-

Drone detection may be split into software that uses AI-enabled tools like machine learning and deep learning and the previous hardware-based group that uses detectors [1]. In this case, the recommended detection and tracking study focuses on the navy, one of the most important global defense systems. Air surveillance relies heavily on detection, which is much more important for military systems **(Fig. no-1).** Identifying sky-moving objects like drones seems to be one of the important difficulties in light of the nefarious maneuvers and drone strikes by several countries, which are not only convenient for the terrorists to murder but also to destroy military equipment and hurt people. Accordingly, the proposed endeavor would carefully examine and consider the technology and global information that can track moving objects at night, which may be the most hazardous and secure period for an attack [2].

Corresponding Author:- Deepanshu Phartyal
Address**:-** Research Scholar, MCEME, Secunderabad.

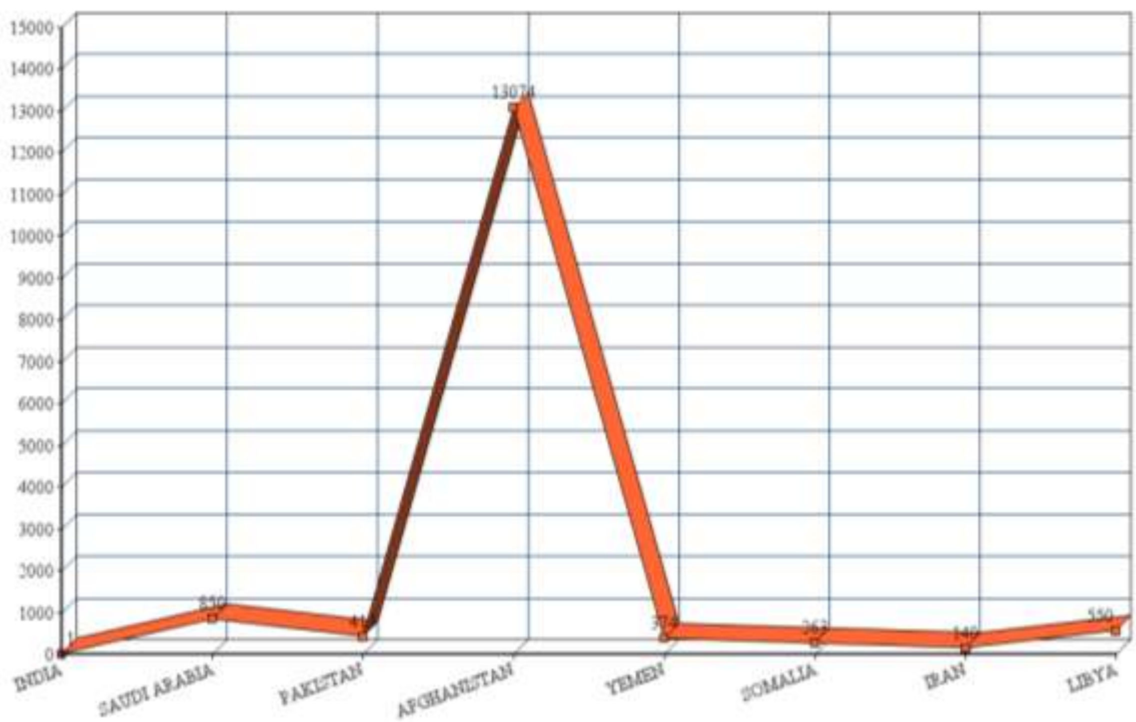

Fig. No-1:- Global statistics of total Drone strikes in the recent year over different countries.

Methods Used To Review Recent Works

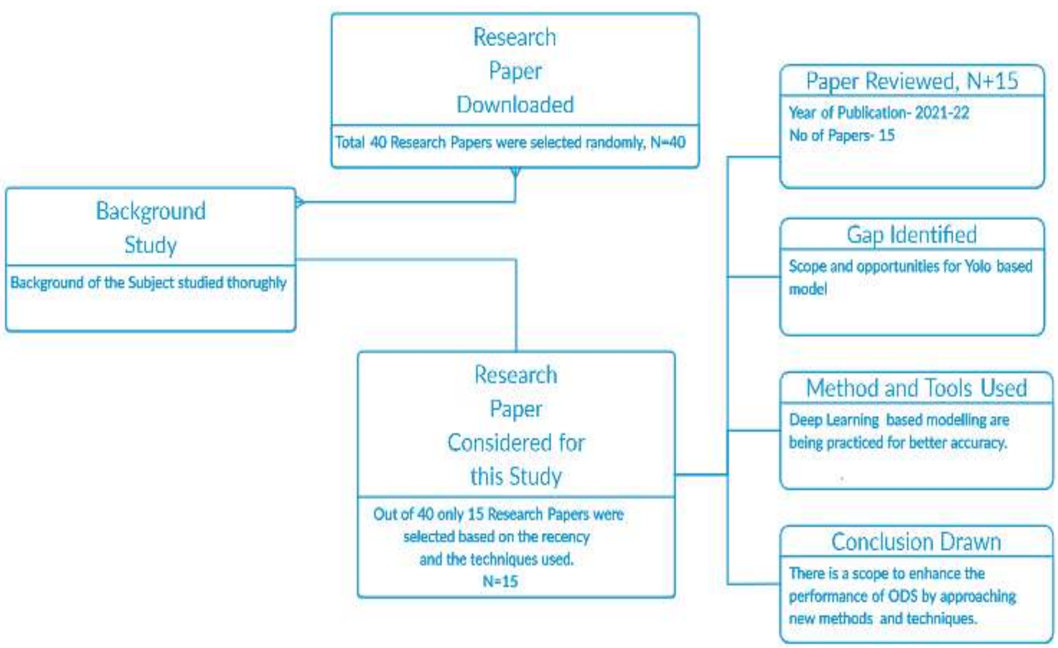

Fig. No-2:- Methodological approach was adopted to critically review the recent technological advancement in Drone Detection and Tracking.

Object Detection System- An Introduction

According to (**Fig. No.3)**, the standard item detection framework is divided into four steps. This is the typical method of object detection. using a sliding window to construct candidate areas on a given image, then using these regions to extract powerful capabilities, categorizing and identifying the regions with the help of an expert classifier, and finally updating and improving the NMS detection results (Non most Suppression)

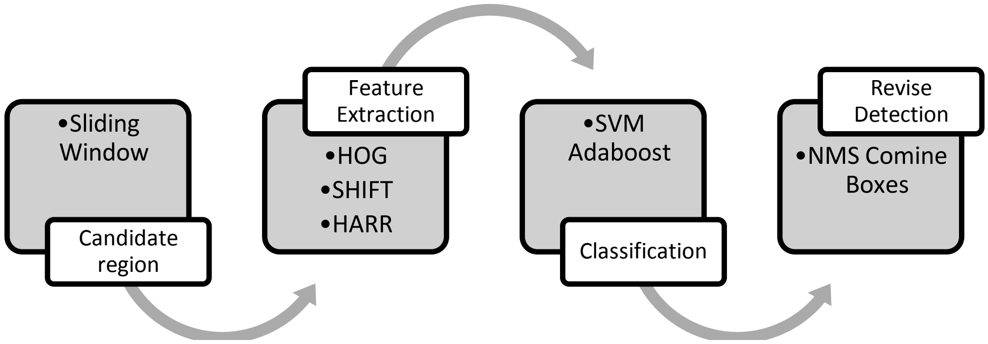

Fig. No-3:- Working mechanism of ODS **(Wang Zhiqiang1 and Liu Jun 2017).**

1. **Candidate areas -** This stage establishes the item's territory. On the other hand, the item can be visible across the whole image. Since the item's size and aspect ratio are both unknown, a sliding window technique was utilized to browse the whole picture using a succession of scale and component ratio sliding windows **(see Fig. No-3).** This thorough approach covers most of the object's likely locations but has three noticeable flaws. First, there are far too many redundancy windows because of the complicated time. Using this might significantly influence the efficiency and speed of characteristic extraction.
2. **Function extraction:** The performance and design of the classifier are instantly impacted by the characteristic extraction approach. However, it is difficult to create a continuous function because of various external factors, such as moving devices, changing lighting, and an ever-changing environment. Scale-Invariant feature remodels (SIFT), Histograms of Oriented Gradients (HOG), and Binary neighborhood patterns are some of the current hot features (LBP).
3. **Classification:** It is standard practice to classify the acquired features at this point using SVM or AdaBoost classifiers.
4. **Refine Detection**: After type, the detection results still include many redundant windows, so removing them and improving the detection outcomes by merging overlapping Bounding bins and using Non-most Suppression (NMS) is important. Traditional object identification has two major problems: first, the sliding window approach is not robust enough for various changes, has an excessive level of temporal complexity, and has an excessive number of duplicate home windows.

Convolutional Neural Network

Characteristic extraction and classification have long been important research topics in computer vision. The recovered functions in traditional image processing tasks are often pre-designed functions entirely dependent on statistical regularities or prior data. For this reason, it cannot accurately and completely duplicate the information from the original shot. The gradient descent method may be used to train the parameters of the CNN quit-to-give-up learning version. A well-trained CNN may also learn more about the features of the image, improving its ability to extract information from the dark field. CNN lessens the complexity of the community version and the schooling settings by expanding the idea of shared weights and receptive fields. By sharing the weight of the convolutional kernel, each layer's capabilities are derived from the adjacent area of the previous layer (receptive discipline).

In addition to helping CNNs learn and create visible data more effectively than other neural networks, these characteristics also help them maintain translation and scale invariance. The first few layers of a typical CNN are made up of alternative convolution and pooling layers, and the last levels, which are often fully-connected networks, go on the output layer. To learn the layer-connection weights, bias, and other parameters, CNN uses forward propagation and BP (again Propagation) algorithms.

Through optimizing community parameters and using annotated picture data, the training is a supervised learning approach that creates an optimized-weight model. The practical layer shape of CNN has several levels **(see Fig. No-4)**. Convolutional, pooling, and fully-linked layers are all in a conventional CNN. Only a few of the additional layers that CNN has provided as part of its development and enhancement strategy include the SPP-layer from the SPP-internet, the ROI (area of interest)-pooling layer from fast R-CNN, and the RPN (area proposal network) layer from faster R-CNN. The typical CNN shape might be larger to provide superior performance depending on the specific challenges. This section offers a conventional CNN's BP technique and basic community topology.

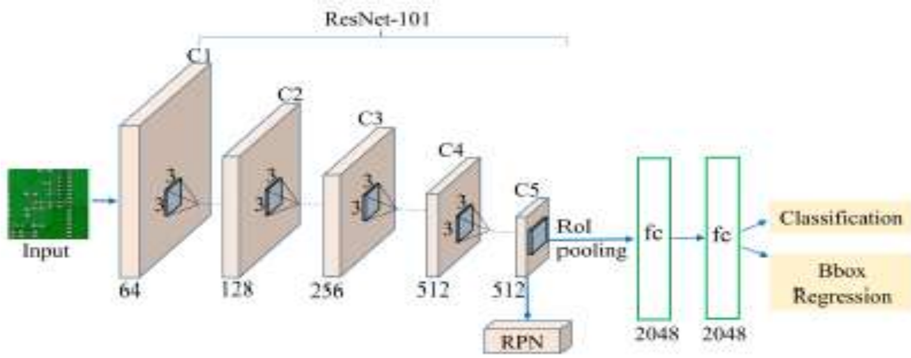

Fig. No-4:- Basic Structure of CNN **(Runwei Ding et al., 2019).**

Review Of Literature:-

According to **Farhad et al. (2022)**, an attempt was made in this research to locate and identify birds and drones more accurately. The authors used Deep learning to develop the model to achieve this study's purpose. The recommended model, which assists in separating birds and drones, considers this. The physical features of birds and drones are one of the main issues that cause erroneous impressions. For the ten thousand images in the data set used for this examination, the cautious version provides 83% accuracy. According to **Xiang Ren et al. (2022)**, an analysis of the drone identification field's limitations was conducted, and it had previously been recommended to include other factors, including processing speed, storage capacity, and many others. The proposed work was preferable since it would hasten to process and need less drone storage space (**see Fig. No-2 for Review Process)**.

The masks-RCNN technique and MobileNetV3 were combined to fix the problems with the UAV TIR video flow dataset. The authors of this research study by **Lizhi Yang et al. (2022)** have seriously handled the dataset issue and developed a solution to postpone such problems, which need much attention. Artificial infrared images have been produced using simulation techniques like Sim Simulation Engine and CycleGAN. Additionally investigated were the fusion techniques used for the hybrid RGB-IR dataset and any issues that may arise. Later, the built setup was assessed using NVIDIA Jetson Xavier. Researchers **Fredrik Svanstrom et al. (2021)** have developed a novel drone detecting approach to reduce mistaken detection. System mastery and fusion sensor techniques have been used to achieve the aim using a video flow dataset gathered from Sweden's three airports. The gathering includes equipment for training with drones, birds, helicopters, and airplanes. This study's dataset, including audio and video, contributes to the machine's sophisticated design. The authors suggest using cloud and bug as two more sorts of things in order to improve false detection. In this study by **Brian et al. (2021)**, a DNN-based technique was used to recognize and track the UAV for IR and visible pictures. The dataset used in this study comprises a variety of ambient occurrences that help the system adjust and interpret things more precisely. With a 98 percent visibility cost for visible photos and an 82 percent IR visibility fee, the mAP yields the greatest results. This result points to the ideal design that would provide the best performance within the DNNs area. **Sara Al-Emadi (2021)** In this study, an attempt was made to improve the Drone's acoustic properties to find and track it. Several Deep learning systems have been used to understand and detect drones. The authors developed a hybrid drone acoustic dataset to meet the current requirement and future research. It is due to the ebook's discussion of the scarcity of acoustic datasets.

Deep learning GAN methods were applied to create audio clips of both recorded drones and artificially created noises. Several algorithms, including CNN, RNN, and CRNN, were used to assess the utility of the compiled dataset. The innovation also emphasizes the advantages of the suggested technology over its competitors in terms of utility and benefits. In this study by **Tazo Hong et al.**, real-time drone surveillance and detection were examined (2021). YOLOv3, with Deep learning, provided the best channel to carry out the advised concept with the help of a Drone monitoring dataset that comprises images of varied environments and four amazing types of drones.

The implementation's actual results show that tracking accuracy has reached 94%. The results also include the setup's overall economic expenses. The efficiency of several YOLO technologies is compared in this study by **Shuo Wang (2021)**. Pascal VOC with mAP and FPS were assessed and examined using XTDrone UAV simulation software. YOLO v3, YOLO v3 small, YOLO v3-SPP3, YOLO v4, and YOLO v4-tiny were compared. The test results indicate that YOLOv4 offers a higher percentage of accuracy when compared to the rest of the options, which is 87.48. When the performance of the premium styles was compared in terms of test duration, it was shown that YOLOv3-tiny performs best. According to **Daniel Tan Wei Xun et al. (2021)**, switch learning techniques have been employed with the YOLOv3 detector. The results show that YOLOv3 had a mean accuracy of 88.9 percent, in contrast to the device learning a technique. The correctness of the cautioned model was previously examined using an NVIDIA Jetson TX2 for an actual-time detection test. The YOLOvs Deep Learning-based Detector was primarily trained for Drone Detection utilizing Transfer Learning. According to **Syed Ali Hassan et al. (2019)**, the scientists produced an actual dataset captured to identify drones. Given that expanding the dataset is one of the essential elements of this sort of implementation, the researchers employed YOLO versions 2 and 3 to achieve their objective. The performance of the upgraded model was evaluated in terms of MAP and accuracy. The results indicate that the YOLOv3 performs better than the comparison. According to research by **Behera and Raj (2020)**, Convolution Neural Network (CNN) is the most accurate method for identifying objects in photographs and extracting features from images.

The recommended work, which also uses laptop vision, is seen to be the most truthful option. 150 epochs were used due to YOLOv3's complex design and lack of class. Researchers should approach RF sign detection, according to the authors. **Yuanyuan Hu et al. (2019)**, the authors of this reference document, strongly advise utilizing YOLOv3 since it provides the highest performance for detection with the best accuracy and velocity by gathering the deep and high-level features. The proposed version for anti-UAV uses the most current four-scale feature maps to anticipate bounding containers. The size of all four scales was determined from the input data. The authors suggest that the approved technique may be changed to enhance performance in the future. The book by **Haipeng Zhao et al. (2020)** provides an evolved version for lightweight real-time object detection using YOLOv3-LITE.

The cautioned version may be used on both GPU and mobile devices.

The mechanism of the proposed model comprises complement residual block, parallel high-to-low resolution, and growing network intensity. According to the extracted output from the study, the size of the created version is 20.5MB, much smaller than YOLOv3-Tiny, YOLOv3, and SlimYOLOv3spp3 in terms of size by 91.70%, 38.07%, and 74.25%. According to **Adarsh and Rathi (2020)**, the suggested strategy uses degree detection systems with exceptional dreams. The one-degree detection method includes YOLO v1, v2, v3, and SSD, which primarily focus on speed, while the two-degree detection approach comprises RCNN, rapid-RCNN, and faster RCNN. The results indicate that the YOLOv3-tiny may help increase speed, but there is no optimum instrument for accuracy since it depends on the situation. The study by **Mingjie Liu et al. (2020)** offers a novel method for model advent from a UAV detecting perspective. The model that is advised is the YOLOv3 Resblock. ResNet devices were interconnected to improve performance on the darknet.

Noteworthy Contributions

Table No 1:- An overview on the maximum Accuracy achieved in the past work.

S.NO	AUTHOR	YEAR	NOTEWORTHY CONTRIBUTION
1	Farhad et al.	2022	83%
2	Xiang Ren et al.	2022	50% (S)
3	Lizhi Yang et al.	2022	90%
4	Fredrik Svanstrom¨et al.	2021	0.9623% (P)
5	Brian et al.	2021	82%
6	Sara Al-Emadi et al.	2021	92%
7	Tao Hong et al.	2021	94%
8	Shuo Wang.	2021	87.48%
9	Daniel Tan Wei Xun et al.	2021	88.9%
10	Syed Ali Hassan et al.	2019	25 % (mAP)
11	Behera & Raj	2020	150 epoch
12	Yuanyuan Hu et al.	2019	89%
13	Haipeng Zhao et al.	2020	48.25 (mAP)
14	Adarsh & Rathi	2020	57% (mAP)
15	Mingjie Liu et al.	2020	72.21% (mAP)

Due to the use of DL-CNN, which can only increase accuracy by 83 percent, there was a need for a version with improved accuracy **(Table no-1)**. The proposed version attempted to address these issues with the help of which incorporates YOLO Darknet 53. Moreover, the suggested model has appropriately advanced the overall detecting performance with a noticeably higher accuracy stage **Adarsh and Rathi (2020)**,. Detection accuracy was one of the crucial areas that this study seemed to have skipped over. The recommended version was thus developed with accuracy testing in mind, and an effort was made to determine the accuracy's maximum range. This analysis looks for the most important factors for dataset availability and quality **Haipeng Zhao et al. (2020)**. For these paintings, no models were made. The suggested model aims to both find and monitor. Another variation uses a dataset with unique classifications. Fake detection was attempted. However, the recommended model's degree of accuracy was not discussed. The suggested version improves by adding more classes to the dataset and using the most current version of the DL algorithm, allowing one to discover with a higher level of accuracy.

This version's maximum accuracy of 82% when utilizing an identical dataset is insufficient to assess ODS's overall effectiveness in identifying small moving objects. The recommended approach offers 98 percent accuracy in detecting and tracking identical IR photos **Mingjie Liu et al. (2020)**. The authors have contrasted the models' accuracy with many alternatives that could perform better; consequently, this can be assessed using various technologies and techniques. Consequently, while using the most current version of CNN, the proposed model outperforms and delivers noticeably more accuracy than the previously analyzed version. While numerous drones were tested in diverse environmental circumstances, this may rise to 94%.

The cutting-edge research for the same environmental nation has provided numerous types of goods. This will be a particular test strategy that the suggested model attempted. Many YOLO class technologies have been tried to discover improved performance. Additionally, it was advised that YOLOv3 choose enhanced output. It was implied in the intended analysis that the YOLOv3 class set of rules would be used in order to obtain the maximum accuracy for detection and tracking. The transfer learning strategy, which might increase accuracy by 88.9 percent, was studied in the current study **Yuanyuan Hu et al. (2019)**,. There is still room for advancement. In order to increase the accuracy percentage, the recently presented model employed the same generation as YOLOv3's most recent new release. In the previous examination, the two algorithms, Yolo v2 and Yolov3, were compared. The highest attainment is found in both approaches. The proposed version employed Yolov3 and the transfer mastering technique in order to improve the version's correctness, which has yet not been accomplished **Behera and Raj (2020)**,. In this analysis, CNN and Yolov3 used epochs with only one item elegance. A test involving many devices may be carried out to lower the number of epochs. Therefore, the current research challenge used a dataset with four different item classifications. This will make an epochs discount easier. In this examination, Yolo was used with the ok-approach, which had comparatively lesser accuracy.

The suggested attempt integrated the equal Yolo and Darknet to enhance the model's overall performance. A small-scale prototype was made. However, it performed badly and lacked the accuracy that was expected. The suggested variant is the ideal one since it has the potential to improve accuracy while being somewhat bigger **Syed Ali Hassan et al. (2019)**,. Even when striving for speed and accuracy, the suggested two-level detecting device could not reach the highest precision required to fulfill the criteria. The most recent version of the YOLO v3 algorithm was used in the present analysis, which specialized in the detection and achieved the highest level of accuracy when compared to the previous model **Farhad et al. (2022)**, the version is excellent at detecting when tracking work is being done. As a result, the current effort aimed at both detection and tracking may be completed with a 98% accuracy rate.

Conclusion And Future Scope:-

The study focuses on object detection using CNN, and it introduces the structure of CNN, the framework for object detection using CNN, and approaches for enhancing detection performance. Because CNN is so good at extracting features, it can compensate for the flaws in hand-crafted features. CNN also outperforms traditional methods in terms of real-time, accuracy, and adaptability, but it still has much space for development. Improving the structure of CNN can reduce feature information loss while properly exploiting object and context relationships, and constructing fuzzy inference can help the computer deal with issues like occlusion and low resolution. Future studies will focus on improving intelligence and the practicality of object detection using CNN. There is still considerable work to be done in the future; for example, one can focus on the application of identifying UAVs by developing models such as SSD, Fast-RCNN, and others. Our next steps will be to create a big data collection, compare the outcomes using advanced models, and follow the UAVs.

References:-

[1] Farhad Samadzadegan et al.,(2022), "Detection and Recognition of Drones Based on a DeepConvolutional Neural Network Using Visible Imagery," aerospace,9.
[2] Xiang Ren et al.,(2022)," An improved mask-RCNN algorithm for UAV TIR video stream target detection," International Journal of Applied Earth Observations and Geoinformation, v-106.
[3] Lizhi Yang, et al.,(2022), "Drone Object Detection Using RGB/IR Fusion," arxiv.
[4] Fredrik Svanstrom et al.,(2020), "Real-Time Drone Detection and Tracking WithVisible, Thermal and Acoustic Sensors," International Conference on Pattern Recognition (ICPR).
[5] Brian K. S. Isaac-Medina et al.,(2021), "Unmanned Aerial Vehicle Visual Detection and Tracking using Deep Neural Networks: A Performance Benchmark," arxiv.
[6] Sara Al-Emadiet al.,(2021), "Audio-Based Drone Detection and Identification Using Deep Learning Techniques with Dataset Enhancement through generative Adversarial Networks," Sensors, v-21.
[7] Tao Hong et al.,(2021), "Multitarget Real-Time Tracking Algorithm for UAV IoT," Wireless Communications and Mobile Computing,v-2021.
[8] Shuo Wang (2021), "Research Towards Yolo-Series Algorithms: Comparison and Analysis of Object Detection Models for Real-Time UAV Applications," Journal of Physics: Conference Series-2021.
[9] Daniel Tan Wei Xunet al.,(2021), "Drone detection using YOLOv3 with transfer learning on NVIDIA Jetson TX2", Second International Symposium on Instrumentation, Control, Artificial Intelligence, and Robotics (ICA-SYMP).
[10] Syed Ali Hassan et al.,(2019), "Real-time UAV Detection based on Deep Learning Network," ICTC, IEEE.
[11] Dinesh Kumar Behera et al.,(2020), "Drone Detection and Classification using Deep Learning," Proceedings of the International Conference on Intelligent Computing and Control Systems (ICICCS 2020).
[12] Yuanyuan Hu et al.,(2019), "Object Detection of UAV for Anti-UAV Based on Improved YOLO v3", Proceedings of the 38th Chinese Control Conference, China.
[13] Haipeng Zhao et al.,(2020), "Mixed YOLOv3-LITE: A Lightweight Real-Time Object Detection Method", Sensors,v-20.
[14] Pranav Adarsh et al.,(2020), "YOLO v3-Tiny: Object Detection and Recognition using one stage improved model",6th International Conference on Advanced Computing & Communication Systems (ICACCS).
[15] Mingjie Liu et al.,(2020), "UAV-YOLO: Small Object Detection on Unmanned Aerial Vehicle Perspective," Sensors,v-20.