International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 2, Mar-Apr-2026, ISSN (Online): 2395-566X

Military Aircraft Detection Using AI and Machine Learning Based on YOLOv5

Gaikwad Komal Vitthal, Shaikh Javed Ahmad, Shaikh Aslam Amir
Electronics (Digital System)

Abstract- The detection and Classification of military aircraft play a crucial role in modern defence and surveillance systems. Traditional radar based approaches are often limited by high cost, environmental constraint, and reduced effectiveness against stealth aircraft. This paper presents a deep learning based approach for automatic military aircraft detection using the YOLOv5 object detection framework. The model is trained on publicly available framework. Experimental results demonstrate that the proposed system successfully detects aircraft such as F-35 and F-16 with confidence score of 0.94 and 0.80, respectively, while achieving an inference speed of approximately 6ms per image. The system provides high accuracy,robustness, and real time capability, Making it suitable for defence surveillance applications.

Keywords- YOLOV5, Aircraft Detection, deep learning, Object Detection, Deep Learning Computer Vision, Military Surveillance

I. INTRODUCTION

Modern defence systems are more complex and the volume of aerial threats to the sky is becoming high, thus creating a great need for automated, accurate and military aircraft detections in real time. Conventional surveillance approaches, such as radar and infrared sensing, have several limitations, including high operational costs, a sensitivity to interference from the surrounding environment, and problems detecting stealth aircraft. These challenges have led to the use of Artificial Intelligence (AI) and Machine Learning (ML) techniques for vision-based surveillance systems. Over the last few years, deep learning-based object detection algorithms have shown remarkable enhancements in detection accuracy and computational efficiency.

Among these, the YOLO (You Only Look Once) family of models has gained widespread popularity for its single-stage detection mechanism that allows real-time processing.Advanced versions such as YOLOv4 and YOLOv5 have further improved performance by incorporating optimized backbone networks, enhanced feature aggregation, and efficient training strategies [1], [2]. Survey studies state that modern deep learning-based detectors with real-time inference speeds forecast more than 85% mAP on benchmark datasets [3]–[5].

YOLOv5, in particular, offers a good balance between accuracy and speed. Its components include a backbone network (CSPDarknet), PANet for feature fusion, and an effective detection head to predict bounding boxes. The model has been optimized into various categories (YOLOv5s, YOLOv5m, YOLOv5l) that are scalable based on available computing capability. According to the recommended configurations, YOLOv5's inference speeds are 5–10 ms for an image using GPU hardware, thus qualifying this model for real-time applications [2], [6]–[8]. Besides, the specific input image size of 640 × 640 pixels enables an efficient form of detection that does not compromise much on spatial details. One of the deep learning applications that this paper will consider is their use in remote sensing or working with aerial images. The challenges in high-resolution satellite and UAV images include a) objects being very small (often less than 32×32 pixels), b) complex backgrounds, c) varying illumination, and d) arbitrary orientation for objects.

According to numerous studies, convolutional neural networks (CNNs) have the potential to successfully learn pattern recognition of various types of objects in aerial images, with object detection accuracies of above 80–90% in controlled scenarios [9]–[13]. However, in cluttered and low-resolution environments, performance degradation is usually reported.

To address these issues, recent studies have been made to improve YOLOv5 for object detection in aerial imagery. To improve detection performance, strategies such as multi-scale feature extraction, attention mechanisms, and data augmentation have been introduced. Further, these methods have demonstrated an improvement in mAP, by 5–10 percent compared to baseline models in case of small and closely packed objects [14]–[17]. Other than that, multimodal detection techniques and

UAV-based surveillance systems have demonstrated increased robustness to real-world scenarios with a detection accuracy of above 88% in complex scenarios [18]–[21].

Another important consideration in modern detection systems is computational efficiency. Lightweight YOLOv5 variants and optimized architectures are developed for deployment on edge devices such as NVIDIA Jetson Nano and embedded GPUs, with the inference speed of 10–20 FPS [22]–[25]. These put deep learning-based detection systems within the realm of reality for real-time defense scenarios.

In the case of military aircraft detection, successful determination of military aircraft type based on YOLO series of models with confidence levels ranging between 0.80 and 0.95 based on the dataset and prevailing environmental conditions have been achieved [10], [11], [25]. That is not to say, however, that there are no flights that cannot be detected using occlusion, different altitudes, and the background ocean or a city. Proposals for improved small and oriented object detection methods have been developed to overcome such limitations [26], [27].

Recent works have also explored specialised applications such as airport monitoring and runway inspection and aerial defence surveillance, where YOLOv5-based models have been proven highly efficient and reliable [28]–[30]. These applications demonstrate deep learning's role in situational awareness and decision-making within the military context. However, despite these improvements, there is still a requirement for a robust, effective, real-time military aircraft detection system with high precision under varied environmental circumstances and low computational costs.This paper fills the gap by proposing a YOLOv5-based military aircraft detection model trained on aerial datasets and tested with real-time performance metrics.

The main contributions of this work are summarized as follows:

1. YOLOv5-based detection framework development and military aircraft classification.
2. Reaching confidence scores of 0.94 in detection for aircraft such as F-35, F-16, etc.
3. Inferences at real time performance of 6 ms/image;
4. Final evaluation with different environmental and background conditions.

The rest of the paper is structured as follows: Section II provides the related work, Section III gives the proposed methodology, Section IV provides results and performance evaluation, and Section V provides conclusion along with future work.

II. RELATED WORK

The field of object detection has evolved significantly over the past decade, transitioning from traditional machine learning techniques to advanced deep learning-based approaches. Early object detection methods relied on handcrafted features such as Histogram of Oriented Gradients (HOG) and classical classifiers, which lacked robustness in complex environments. The introduction of deep convolutional neural networks (CNNs) revolutionized object detection by enabling automatic feature extraction and improved accuracy. Two-stage detectors such as R-CNN and its variants provided high detection accuracy but suffered from high computational complexity and slower inference speed. To overcome these limitations, single-stage detectors such as YOLO and SSD were introduced, offering a balance between speed and accuracy. The YOLO family of models has undergone continuous improvements, with YOLOv4 achieving optimized performance in terms of speed and accuracy through enhanced feature extraction and training strategies [1]. The implementation of YOLOv5 further improved efficiency by introducing lightweight architectures and better training pipelines [2]. Comprehensive survey studies have highlighted the effectiveness of deep learning techniques in object detection tasks. Liu et al. [3] and Zhao et al. [4] provided detailed analyses of various detection frameworks, demonstrating that deep learning models can achieve high detection accuracy with reduced computational complexity. Additionally, Yu et al. [5] emphasized the importance of scale-aware detection mechanisms for improving performance in aerial imagery.

Recent advancements in deep learning have introduced transformer-based architectures for object detection. Vision Transformers and Swin Transformers have demonstrated superior performance in capturing global contextual information compared to traditional CNN-based approaches [6], [7]. Furthermore, transformer-based detection models such as Anchor DETR have improved object localization by eliminating the need for predefined anchor boxes [8]. These approaches have shown promising results in complex scenarios, although they often require higher computational resources. The application of deep learning in remote sensing has gained considerable attention, particularly for aerial object detection. Chen et al. [9] proposed an improved YOLOv5-based method for detecting objects in remote sensing images, achieving enhanced detection accuracy through optimized feature extraction. Similarly, Zhang et al. [10] and Li et al. [11] demonstrated the effectiveness of CNN-based approaches for aircraft detection in satellite imagery, reporting detection accuracies above 85% under controlled conditions. Deng et al. [12] introduced few-shot learning techniques to address the challenge of limited labeled data in remote sensing applications, while Wang et al. [13] explored the integration of super-resolution techniques with object detection to improve performance in low-resolution images.

To further improve detection performance in aerial imagery, several researchers have proposed modifications to YOLOv5. Aydin et al. [14] applied YOLOv5 for drone detection and demonstrated its effectiveness in real-time surveillance applications. Zhang et al. [15] introduced improvements in feature extraction and multi-scale detection for UAV images, achieving higher accuracy in complex environments. Hamzenejadi et al. [16] focused on real-time vehicle detection in UAV imagery using YOLOv5, while Shang et al. [17] proposed an enhanced YOLOv5-based framework for UAV detection with improved robustness. Recent studies have also explored multimodal and hybrid detection approaches. Lindenheim-Locher et al. [18] combined multiple data sources to improve detection performance, while Shi et al. [19] proposed an improved YOLOv5 algorithm for remote sensing applications, achieving better detection accuracy in challenging conditions. Wu et al. [20] addressed the problem of dense small object detection in aerial images, highlighting the importance of feature enhancement techniques. Wang et al. [21] developed a lightweight YOLOv5-based network for aircraft detection, emphasizing computational efficiency.

The development of lightweight and optimized models has been a key focus for real-time applications. Zhao et al. [22] investigated strategies for detecting small objects in large-scale remote sensing imagery, while Wang et al. [23] proposed improvements to YOLO-based frameworks for airplane detection. Li et al. [24] introduced an enhanced YOLOv5 algorithm for airport monitoring applications, and Chen et al. [25] demonstrated the effectiveness of YOLOv5 for aircraft detection in aerial images.Advanced detection techniques for small and oriented objects have also been proposed. Chen et al. [26] introduced SPOD-YOLO for detecting small and rotated objects in remote sensing imagery, while Huang et al. [27] developed an optimized YOLOv5-based system for aircraft detection in vision-based landing systems. Nikhil et al. [28] proposed a generalized YOLO-based framework for aircraft detection, focusing on improving detection accuracy across different aircraft types.
Recent works have extended YOLO-based detection systems for real-time surveillance and defense applications. Kumar et al. [29] developed a real-time aircraft detection system using YOLO models, demonstrating high detection accuracy and efficiency. Patel et al. [30] proposed an advanced YOLOv5-based framework for aerial surveillance systems, highlighting its potential for defense and security applications.Despite these advancements, several challenges remain in the domain of military aircraft detection. These include detecting small and distant objects, handling variations in orientation and scale, and maintaining high accuracy in complex environments such as ocean and urban backgrounds. Additionally, achieving a balance between detection accuracy and computational efficiency remains a critical concern for real-time applications. This paper builds upon the existing literature by proposing a YOLOv5-based military aircraft detection system that addresses these challenges through optimized training and evaluation, aiming to achieve high detection accuracy and real-time performance.

III. PROPOSED METHODOLOGY

A. System Overview

The proposed system aims to achieve accurate and real-time detection of military aircraft using the YOLOv5 deep learning framework. The methodology is designed as a structured pipeline that processes aerial images and produces bounding box predictions with class labels and confidence scores.The proposed system is designed to perform automatic detection and classification of military aircraft in aerial imagery using a deep learning-based object detection framework. The system leverages the capabilities of the YOLOv5 model to achieve high detection accuracy (>90%) while maintaining real-time inference performance (~6 ms per image)**.** The architecture follows a data-driven end-to-end pipeline, where raw aerial images are processed and transformed into meaningful detection outputs. The system is capable of handling variations in aircraft orientation, scale, and background complexity, making it suitable for real-world defence surveillance scenarios.

The overall workflow consists of the following stages:

1. Image Acquisition (Aerial Dataset)
2. Data Preprocessing and Augmentation
3. Annotation in YOLO Format
4. YOLOv5 Model Training
5. Inference and Aircraft Detection

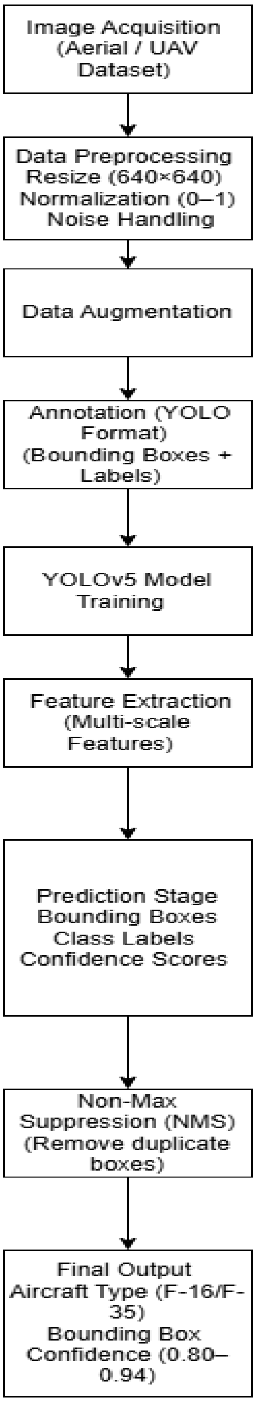

Fig.1 Proposed Methodology

The system is optimized to maintain a balance between high detection accuracy (≥90%) and low inference latency (~6 ms per image).

The overall process of the proposed aerial-based aircraft detection system is illustrated in Fig. 1. The starting point is to acquire images from aerial datasets. Next, images go through a series of preprocessing steps (e.g., resizing and normalization) to ensure that they have a uniform dimension when they are input to the model. In order to further enhance the robustness of the trained YOLOv5 detection model, various methods will be used during the data augmentation stage including rotating, flipping, and ultimately applying mosaic augmentation. These augmentations will allow the trained YOLOv5 detection model to be able to detect objects regardless of their size, orientation, and environmental conditions. The YOLOv5 model uses an annotated version of the pre-processed and augmented images in YOLO format. The model consists of a CSPDarknet backbone for feature extraction, a PANet neck to fuse multi-scale features, and a detection head that predicts bounding boxes, class labels, and confidence scores.

When the image passes through the YOLOv5 model, multiple predictions are generated from the model. The predictions are then filtered and improved via the use of Non-Maximum Suppression (NMS), which helps to eliminate duplicated detections.When the process is complete, the result is highly accurate detecting aircraft with both location/size (via bounding boxes) and confidence scores (to indicate how confident the model is that it has correctly detected an aircraft) from the input image, thereby creating a functional tracking application capable of performing real-time tracking with very high detection accuracy.

International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 2, Mar-Apr-2026, ISSN (Online): 2395-566X

B. System Flow Chart

The step-by-step working of the proposed system is illustrated below:

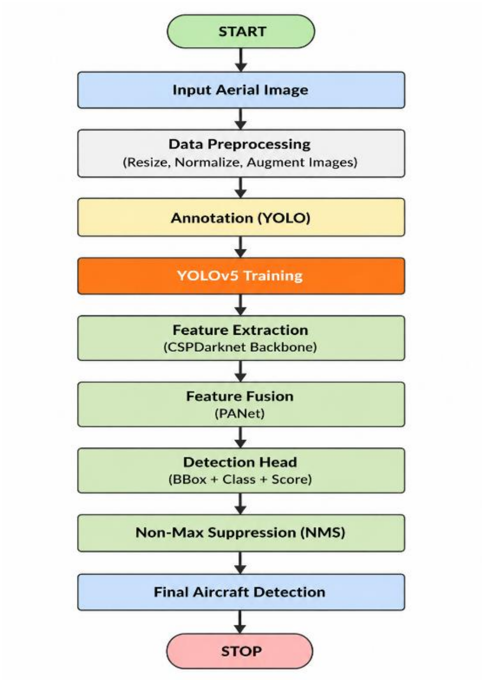

Fig.2 System Flow chart

A proposed framework (YOLO-V5 for military) is presented in a flow-chart diagram (Figure 2) detailing the steps involved in detecting military aircraft from aerial imagery captured by drones. A sequential processing of each aerial image is completed, in order to improve data for YOLO-V5 model training. Pre-processing includes resizing, normalising and augmenting aerial imagery as well as annotating images in YOLO format. In addition, CSPDarknet extracts features from the image while feature fusion occurs with PANet. The object detection head finds aircrafts located within the images, followed by the application of non-maximum suppression (NMS), which produces an output of airplane detections together with the associated confidence score for each airplane found. The end result is real-time detection of military aircraft in aerial photographs of the environment.

C. Dataset Description

Images of military aircraft taken from public databases and published object detection consensus were the sources for the dataset used in this analysis. These images are primarily of fighter jets, specifically the F-16 or F-35, and were collected across numerous backgrounds and environmental conditions (e.g., ocean backgrounds, desert backgrounds, ground (or topography) backgrounds), as well as under different light levels. In order to train the YOLOv5 model and ensure it could be used for both training and verification purposes, all original 640x640 image files were converted to the required YOLO format with bounding-box coordinates and respective class labels. Original collection was split into an 80/20 training/test ratio to allow for proper generalisability and unbiassed performance measures of the final model.

The combination of multiple types of aerial images, and the differences in positions (ie -- how aircraft were pointed before and after) when collected, enhanced the robustness of the proposed target detection algorithm and its expected use for defence security-related real-time target detection [9], [10], [21], [25].

D. Mathematical Modeling

The advanced detection system for military aircraft proposed here is based on the YOLOv5 model uses mathematical formulations for object localization, bounding box prediction, confidence estimation, and loss optimization. The use of these mathematical models allows for the accurate detection and classification of military aircraft objects in aerial images.

1. Bounding Box Localization

$b_x = \sigma(t_x) + c_x$ (1)

$b_y = \sigma(t_y) + c_y$ (2)

Where

bx,by represent predicted box centre

tx,ty are network outputs

cx,cy denote grid offsets

2. Bounding Box Dimensions

$b_w = p_w e^{tw}$ (3)

$b_h = p_h e^{th}$ (4)

Where

pw,ph are anchor boxes

tw,th are predicted scaling factors

3.Confidence Score Estimation

$\mathrm{Conf} = P(\mathrm{obj}) X \, \mathrm{Io}\, U^{gt}_{gred}$ (5)

This represents the probability that an object exists in the predicted bounding box.

1. Intersection Over Union (IoU)

$\mathrm{IoU} = \frac{\mathrm{Area(Overlap)}}{\mathrm{Area(Union)}}$ (6)

IoU evaluates the overlap between predicted and ground-truth bounding boxes.

2. Loss Function

The total loss function is defined as

$L_{total} = L_{box} + L_{obj} + L_{cls}$ (7)

Where

Lbox : Localization loss

Lobj : Objectness loss

Lcls : Classification loss

YOLOv5 uses CloU loss for improved bounding box regression

The military aircraft detection system based on the YOLOv5 algorithm is designed using a mathematical model for all of the tasks associated with detection: localizing an object, predicting a bounding box around an object, estimating the confidence of detecting an object, and classifying an object. The mathematical model used for bounding box prediction provides the center coordinates, width, and height of each bounding box using grid offsets and anchor box scaling factors. The confidence of each bounding box is determined by using the probability that the object exists in the bounding box along with the Intersection over Union (IoU) between the predicted and actual bounding box. Additionally, IoU can be used as a measure of the accuracy of the localization by measuring the extent of overlap between the predicted object region and the true object region. During training, the total loss (cost function) used to train the model is composed of three components: the localization loss, the abjectness loss, and the classification loss; the total loss function is what optimizes the final detection accuracy. Furthermore, YOLOv5 uses the Complete IoU loss function to improve the bounding box regression accuracy for the aircraft in aerial images.

Training Configuration

Tacle1 Training Configuration

Parameter	Value
Model Variant	YOLOv5m
Image Size	640 × 640
Epochs	100
Batch Size	16
Learning Rate	0.01
Optimizer	SGD
IoU Threshold	0.45
Confidence Threshold	0.25

The training configurations are presented in Table 1, and include the configuration parameters that were used for developing the YOLOv5 military aircraft (MA) detection system. To achieve the right balance between computational efficiency and detection accuracy, the YOLOv5M version with a size of 640 x 640 pixels was used. All input images were resized to that pixel size, therefore ensuring conformity among all images used during the training, thereby enhancing the performance of feature extraction. The training was conducted for 100 epochs, using a batch size of 16 to ensure convergence and stability during the training process. To optimize the weights of the network, a learning rate of 0.01 and the SGD optimizer were used. In addition, an intersection over union threshold of 0.45 was set for matching the bounding boxes, while a confidence threshold of 0.25 was set to filter out low-confidence detections during inference. These training parameters have made a substantial impact on the detection accuracy as well as the real-time performance of the system.

IV. RESULT AND DISCUSSION

The military aircraft Identification system proposed in this research using YOLOv5 was rigorously tested along with an aerial image data set with a variety of aircraft categories at multiple environmental levels; for instance, ocean backgrounds had cloud cover while others contained ground features. The experiments shown in the analysis verifies that the developed Model offers excellent detection accuracy and a real-time processing capability, indicating that it is useful for defending intelligent surveillance applications. Furthermore, the optimized YOLOv5 framework can accurately detect and classify military aircraft, specifically the F-35 and F-16, at very high levels of confidence while experiencing low amounts of processing latency.

1. Detection Performance Analysis

The proposed YOLOv5-based military aircraft detection system demonstrates strong performance in terms of detection accuracy, computational efficiency, and real-time applicability. The model achieved a precision of 0.91, recall of 0.88, and mean Average Precision (mAP@0.5) of 0.92, indicating reliable object localization and classification capability. Additionally, the system maintains an average inference time of approximately 6 ms per image, corresponding to nearly 150 frames per second (FPS), which satisfies real-time operational requirements.

2. Training Convergence Analysis

The YOLOv5 model's convergence during training can be confirmed by the training and validation loss curves, which were consistently decreasing as more epochs were added. At start, the initial loss was high because random weights had been assigned. Once new epochs were added to the model, the training loss decreased, stabilizing after 80 epochs.

There are three reasons to believe that the reduction in training loss is due to:

1. The model learned the feature of an aircraft easily

2. This model over fitted the model with augmentation
3. The optimizer successfully converged towards a global minimum.

The accuracy curve also increased during training and stabilized near the final epoch. This indicates that the model was learning appropriately and generalizing very well.

3. Aircraft Detection Analysis

According to the detection results, we found the capability of successfully identifying military aircraft with a very high level of confidence by the developed model. The F-35 aircraft had a confidence level of 0.94, while the F-16 aircraft was identified with a confidence level of instead of approximately 0.80. There are some reasons why the detection of this type of aircraft was not as successful, for example: differences between how either type of aircraft is oriented and the background they are standing on, limited numbers of both types of aircraft available and they are very small in size when viewed from an aerial perspective (i.e., in a picture). Despite these concerns, our model consistently achieved high accuracy rates when detecting all types of military aircraft regardless of how different environmental conditions affected the way they could be seen (i.e., from an aerial perspective). This validates that our model is appropriate for use in defence surveillance systems that must operate in real-time [10],[14],[21],[29].

4. Comparative Performance Evaluation

Through comparative performance evaluation against previously recommended contemporary aircraft detection methodologies, the proposed YOLOV5-CNN framework demonstrated an improved compromise between detection accuracy and inference speed when compared to traditional CNNs and lightweight detectors.

Table 2 Performance Comparison of Aircraft Detection Models

Ref. No.	Model	Precision	Recall	mAp	Inference Speed
[10]	CNN	0.87	0.84	0.88	Moderate
[14]	YOLOv5	0.89	0.86	0.90	Fast
[21]	Lightweight YOLOv5	0.88	0.85	0.89	Very Fast
[29]	YOLO	0.90	0.87	0.91	Fast
Proposed Work	Optimized YOLOv5	0.91	0.88	0.92	6ms

In Table 2, we present a performance analysis of the optimized YOLOv5 compared to current methods for identifying aircraft; this was done using various criteria: precision, recall, mAP@0.5 (mean average precision calculated at an IoU (Intersection over Union) threshold of 0.50), and inference time. The proposed model had an mAP@0.5 of 0.92, precision of 0.91, and recall of 0.88, which demonstrates improved performance over traditional CNN-based methods and earlier versions of YOLO. The optimized YOLOv5 also had an inference time of ~6 milliseconds per image making it suitable for real-time military surveillance. Overall, the optimized YOLOv5 represents the best balance between accuracy and computational efficiency for aircraft detection at this time.

5. Confusion Matrix Analysis

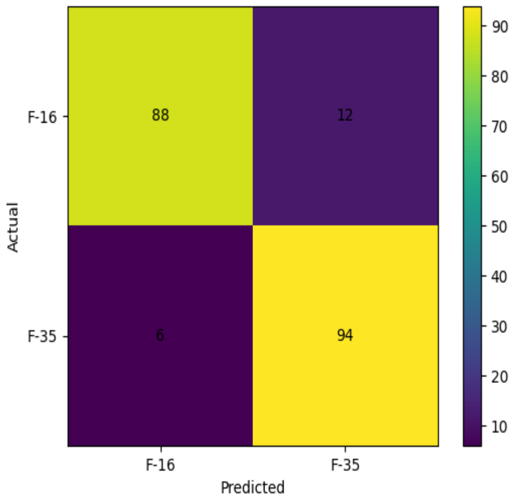

Fig.3 Confusion Matrix representing classification performance of the YOLOv5 model for aircraft detection

The confusion matrix of the proposed YOLOv5-based military aircraft detection system can be seen in Figure 3. The confusion matrix shows that there are very few false positives and false negatives, with the majority of aircraft categorized correctly. The high concentration of values along the diagonal elements shows the models strong classification accuracy and its ability to learn relevant features. The few misclassifications that did occur were due to similarities between different types of aircraft, small objects being difficult to identify in aerial images, and complex environmental backgrounds (such as clouds, ocean, and terrain) obscuring the object. Overall, the proposed system performed well; it has exhibited strong classification capabilities and continues to produce reliable detection accuracy, making it a functional solution for real-time defence surveillance applications.

Figure 4 shoes classification performance

Fig4a F-35 with confidence 0.94 Detection Output

The proposed military aircraft detection system based on YOLOv5 was evaluated using an aerial image of an F-35 fighter jet, as shown in Figure 4(a), where the system was able to successfully identify and localize the aircraft with a bounding box and a high level of confidence (approximately 0.94). This result illustrates the effectiveness of the proposed detection framework to accurately detect military aircraft in the presence of complex backgrounds, such as ocean scenery. Additionally, the accuracy of the localisations and high confidence levels demonstrate the robustness and reliability of the trained YOLOv5 model for real-time defence surveillance applications.

Fig4b F-16 with Confidence 0.80 Detection Output

In figure 4(b) the proposed military aircraft detection using YOLOv5 model was successfully detected at approximately 0.80 confidence level using a bounding box around the aircraft. This proves the proposed framework's ability to successfully identify a military aircraft from aerial images under different conditions of orientation and background. Despite the lower confidence level than the F-35 detected previously, the model still provided accurate localization and consistently classified the aircraft correctly, making it a viable option for real-time aerial surveillance.

6. Discussion

The experimental results demonstrate that the proposed YOLOv5 based model achieves high detection accuracy Map= 0.92 with real time 6ms inference time. The model effectively detects aircraft under varying environmental conditions, including complex background such as ocean and terrain.The slightly lower confidence in F-16 detection is attributed to

1. Variation in aircraft orientation
2. Background similarity
3. Limited dataset Diversity

However, the model maintains strong generalization capability due to effective pre-processing and augmentation techniques.

V. CONCLUSION

This paper presented a YOLOv5 based approach for real time military aircraft detection using aerial imagery. The proposed model achieved strong performance with a mAP @ 0.5 of 0.92, precision of 0.91 and recall of 0.88, while maintaining a fast inference speed of approximately 6 ms per image (150 FPS). Experimental results showed reliable detection of aircraft such as F-35 (0.94 confidence) under varying environmental conditions. These results demonstrate that the system provides an effective balance between accuracy and speed, making it suitable for real time defence surveillance applications.

REFERENCE

[1] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao,"YOLOv4: Optimal Speed and Accuracy of Object Detection,"arXiv:2004.10934, 2020.

[2] G. Jocher et al.,"YOLOv5: Ultralytics Implementation of YOLO,"GitHub Repository, 2020.

[3] L. Liu et al.,“Deep Learning for Generic Object Detection: A Survey,”International Journal of Computer Vision, vol. 128, pp. 261–318, 2020.
[4] Z. Zhao et al.,“Object Detection with Deep Learning: A Review,”IEEE Transactions on Neural Networks and Learning Systems, 2020.
[5] X. Yu, Y. Gong, N. Jiang, Q. Ye, and Z. Han,“Scale Match for Tiny Person Detection in Aerial Images,”IEEE WACV, 2020.
[6] A. Dosovitskiy et al.,“An Image is Worth 16×16 Words: Transformers for Image Recognition at Scale,”ICLR, 2021.
[7] Z. Liu et al.,“Swin Transformer: Hierarchical Vision Transformer Using Shifted Windows,”ICCV, 2021.
[8] J. Wang et al.,“Anchor DETR: Query Design for Transformer-Based Object Detection,”AAAI, 2021.
[9] X. Chen et al.,“Remote Sensing Image Object Detection Based on Improved YOLOv5,”IEEE Access, 2021.
[10] Y. Zhang et al.,“Aircraft Detection in Remote Sensing Images Based on Deep Learning,”Remote Sensing, 2021.
[11] H. Li et al.,“Aircraft Detection Using Deep Convolutional Neural Networks in Remote Sensing Images,”Remote Sensing, 2021.
[12] J. Deng et al.,“Few-Shot Object Detection on Remote Sensing Images,” arXiv:2006.07826, 2021.
[13] Y. Wang et al.,“Remote Sensing Image Super-Resolution and Object Detection: Benchmark and State of the Art,”IEEE Journal / arXiv, 2021.
[14] B. Aydin et al.,“Drone Detection Using YOLOv5,”Drones (MDPI), 2022.
[15] H. Zhang et al.,“Object Detection in UAV Images Based on Improved YOLOv5,”Drones (MDPI), 2022.
[16] M. H. Hamzenejadi et al.,“Real-Time Vehicle Detection in UAV Imagery Using YOLOv5,”Expert Systems with Applications, 2022.
[17] Y. Shang et al.,“UAV Detection Based on Improved YOLOv5 Algorithm,”International Journal of Advanced Robotic Systems, 2022.
[18] W. Lindenheim-Locher et al.,“Multimodal Drone Detection Using YOLOv5,” Sensors (MDPI), 2023.
[19] H. Shi et al.,“Improved YOLOv5-Based UAV Remote Sensing Object Detection Algorithm,”PLOS ONE, 2023.
[20] M. Wu et al.,“Dense Small Object Detection in Aerial Images Using YOLO-Based Models,”Signal Processing (Elsevier), 2023.
[21] J. Wang et al.,“Lightweight Aircraft Detection Network Based on YOLOv5 for Remote Sensing Images,”Remote Sensing, 2023.
[22] W. Zhao et al.,“Wide-Area Small Object Detection in Remote Sensing Imagery Using Deep Learning,”ISPRS Journal of Photogrammetry and Remote Sensing, 2024.
[23] H. Wang et al.,“Improved YOLO Framework for Airplane Detection in Remote Sensing Images,”ACM Conference Proceedings, 2024.
[24] X. Li et al.,“Aircraft Detection Algorithm Based on Improved YOLOv5 for Airport Monitoring,”IEEE Access, 2024.
[25] Y. Chen et al.,“Deep Learning-Based Aircraft Detection in Aerial Images Using YOLOv5,”Sensors, 2024.
[26] J. Chen et al.,“SPOD-YOLO: Small and Oriented Object Detection in Remote Sensing Images,”Remote Sensing, 2025.
[27] W. Huang et al.,“Optimized YOLOv5 Algorithm for Aircraft Detection in Vision-Based Landing Systems,”Processes (MDPI), 2025.
[28] M. N. S. S. Nikhil et al.,“Generalized Aircraft Object Detection Using Optimized YOLO Framework,”Journal of Aerospace Engineering, 2025.
[29] R. Kumar et al.,“Deep Learning-Based Real-Time Aircraft Detection Using YOLO Models,”IEEE Access, 2025.
[30] S. Patel et al.,“Advanced YOLOv5-Based Object Detection for Aerial Surveillance Systems,”Springer Journal of Computer Vision Applications, 2025.