

Precise aircraft recognition

Shashikiran B.S, Akash Narain, Rakshan R Rao, Chirag S Shetty, Jyothi R

Student, Student, Student, Student, Assistant Professor
Information Science and Engineering
BGSCET, Bengaluru, India

shashikiran.banagere@outlook.com, 1mp22is002@bgscet.ac.in, 1mp22is040@bgscet.ac.in, 1mp22is040@bgscet.ac.in,
jyothiramakrishnaiah@gmail.com

Abstract— Precise visual identification of aircraft by manufacturer is an important aspect in ground-based air traffic control and defense surveillance, where in many situations visual confirmation is necessary due to incomplete, corrupted, or missing ADS-B data. This work proposes a deep learning-driven vision pipeline that combines object detection and fine-grained image classification for identifying aircraft in both images and videos based on manufacturer and model variant. A new multi-class dataset, flmanufacturers, was established for commercial and defense aircraft types, including Airbus, Boeing, ATR, among others. Such images were then subjected to systematic preprocessing, normalization, and augmentation to improve the generalization of the model for changing lighting conditions and angles of view. The proposed system combines a YOLOv8 model with fast and accurate aircraft detection and a fine-tuned ResNet18 classifier that recognizes aircraft manufacturers. Training was done on GPU-enabled systems for faster convergence and real-time inference. It obtained a training accuracy of 65% with 20–25 FPS during video inference, hence making it feasible in the real world. These are modularly designed, so scaling to more aircraft categories or fusing with flight metadata for multimodal recognition is possible.

Keywords: transfer learning, aircraft identification, ResNet-18, YOLOv8, video inspection, ATC validation.

I. INTRODUCTION

Aircraft identification plays a vital role in modern aviation, supporting critical operations in ground handling, taxiing coordination, airspace monitoring, and airline fleet management. Correct identification of an aircraft's brand and model is essential not only for smooth airport logistics but also for safety and situational awareness in both civilian and defense realms. Traditionally, this is done through the Automatic Dependent Surveillance–Broadcast system, which transmits data on flight and position to air traffic control. However, ADS-B, though widely deployed, offers only textual and positional information; it does not have the capability for the visual identification or confirmation of an aircraft's physical appearance. Ground personnel or ATC operators normally have visual recognition of aircraft, which is usually subjected to human error, poor light conditions, adverse weather conditions, or even aircraft viewed from nonstandard angles. Even for highly experienced aviation experts, it is quite problematic to distinguish between aircraft models or manufacturers that are of similar appearance. Therefore, there is an increasing demand for an intelligent, automated, and vision-based aircraft identification system that can identify such aircraft quickly and with a high degree of accuracy. For this project, we propose an approach based on deep learning and with the support of computer vision using CNNs for the correct identification of aircraft manufacturers. We use a two-model system, with YOLOv8 for detecting and locating aircraft within images or video frames and ResNet18 for classifying each detected aircraft's manufacturer. By integrating modules of detection and classification within one pipeline, the system will be capable of processing live feeds in real time and labeling aircraft as Airbus, Boeing, ATR, Bombardier with improved accuracy. The ultimate aim of the work is to devise a trustworthy and efficient visual recognition framework that will enhance existing air traffic control infrastructure by adding a visual intelligence layer. Such a system can assist ATC officers, security teams, and aviation analysts in confirming aircraft identity without being dependent on transponder data only. The technology will also find applications in defense surveillance, airport automation, and aviation research, thus proving the increasing potential of deep learning in transforming traditional airspace operations.

II. PROBLEM STATEMENT AND OBJECTIVES

- ATC dependency: Surface-based Air Traffic Control (ATC) depends on accurate aircraft identification for safe and efficient operations.
- Current systems: Existing systems like ADS-B provide valuable tracking and positional data.
- System Development: Develop a system to accurately classify and identify aircraft by manufacturer, first from images and now extended to video streams.
- Scalability & Robustness: Extend the system to identify a wider range of aircraft manufacturers beyond the initial scope, ensuring it generalizes across commercial and defense categories.
- Practical Deployment: Ensure the system performs reliably on varying resolutions and qualities of images/videos, including low-resolution inputs, thereby making it suitable as a visual complement to ATC infrastructure for ground crew coordination and operational safety.

III. LITERATURE SURVERY

[1] Wang et al. (2025) – A Blur-Score-Guided Region Selection Method for Airborne Aircraft Detection in Remote Sensing Images: The work of this paper is the detection of aerial, high-speed planes in remote-sensing images where motion blur and complex

backgrounds degrade the performance of normal detectors. BS-DETR, a transformer detector, is proposed by the authors which computes an improved Tenengrad gradient beforehand to generate a local Blur-Score map with motion-blur details, and then adaptive-fuses that map with multi-scale features by means of an adaptive feature-fusion mechanism. An Aircraft Region Selector (ARS) as an MLP scores and selects prospect regions for the DETR encoder- decoder. A synthetic airborne-aircraft dataset with clouds and contrails is created and experiments demonstrate that BS- DETR outperforms baseline detectors by leveraging motion-blur disparity between aircraft and background. Contributions of the Blur-Score map, AFFM fusion, and ARS region selection are confirmed by ablation studies. The technique presents a new motion-guided methodology for remote-sensing detection issues.

[2] Yaseen (2024) – What is YOLOv8: An In-Depth Exploration of the Internal Features of the Next-Generation Object Detector: This new analytical review delves into YOLOv8’s design structure, training efficiency, and deployment improvements from previous YOLO architectures. YOLOv8 adopts an anchor-free prediction approach, adopts a CSPNet backbone, and features a hybrid FPN+PAN neck for multi-scale feature fusion. It further streamlines training by employing advanced augmentation, improved optimization, and simplified APIs. Comparison of benchmarking across COCO and Roboflow benchmarks demonstrates improved accuracy–latency trade-offs. Although primarily descriptive, the paper will be a valuable contemporary resource for users of YOLOv8 in real-time detection tasks, such as aerial and remote-sensing applications.

[3] Farhadmanesh et al. (2022) – General Aviation Aircraft Identification at Non-Towered Airports Using a Two-Step Computer Vision-Based Approach: This paper presents a two-stage computer vision pipeline for locating general-aviation aircraft in non-towered airports. The first stage uses a CNN-based classifier to estimate aircraft model or type, which reduces the search space in the FAA registration database. The second stage uses OCR-based tail-number recognition combined over multiple video frames in a probabilistic multi-frame-based (MFB) framework to infer the registration number. Transfer learning and HyperBand optimization are shown to yield high classification accuracy, with multi-frame fusion improving identification accuracy to approximately 90%, a performance improvement compared to single-frame OCR. The system is tested on video data from three general aviation airports and exhibits robustness against noise and temporal inconsistency. Limitations are exposure to camera angle and sighting of the tail number, but probabilistic fusion effectively addresses such concerns.

[4] Luo et al. (2022) – Aircraft Target Detection in Remote Sensing Images Based on Improved YOLOv5: This engineering-based research adapts YOLOv5 to remote-sensing aircraft detection with three key modifications: (1) enhancing batch-normalization using centering and scaling calibration, (2) replacing the confidence cross-entropy loss with a smoothed Kullback–Leibler divergence-based loss, and (3) building a CSandGlass (CSG) module to reduce semantic loss and discard the lowest-resolution feature map for local detail preservation. Using a Google Earth–derived aircraft dataset, the authors produce faster convergence and higher mean average precision (mAP) than YOLOv5 and YOLOv4 vanilla and YOLOv4, respectively, at around 85% accuracy at 48.8 fps. The strengths of the paper lie in real-world ablations and dataset engineering. The spaces where it currently suffers are environmental diversity and lighting effects, and therefore the method is best used with similarly collected datasets where detection is required at high-throughput.

[5] Yan et al. (2019) – A Novel Data Augmentation Method for Detection of Specific Aircraft in Remote Sensing RGB Images: This applied research aims to address the problem of scarce labeled examples for specific aircraft by proposing a data-augmentation technique that produces realistic training images. Three-dimensional models of aircraft are rendered into real remote-sensing scenes (without warping scale and viewpoint) so that simulated images are produced, which are then blended with real data for training. Authors find that this brings new positive-sample information compared to geometric transforms and can be combined with standard augmentations. An R-DFPN detection network is experimented with, proving to have boosted detection performance by adding simulated data. The technique reduces overfitting and enhances learning on sparsely populated aircraft classes, with realism and domain-gap constraints still intact. The mixed training pipeline and the 3D model insertion serve as a good baseline for low-data aircraft detection problems.

[6] He et al. (2015) – Deep Residual Learning for Image Recognition: This seminal work introduces the residual learning paradigm that made it possible to train deep convolutional networks through the transformation of stacked layers to learn residual functions $F(x) = H(x) - x$ and utilize identity shortcut connections so that the network learns $F(x) + x$. The paper carefully documents the “degradation” problem (worse training error with increasing depth) and shows residual blocks with identity shortcuts effectively remove that optimization bottleneck, enabling networks of 50–152 layers that outperform earlier architectures (VGG, etc.) on ImageNet and transfer nicely to other tasks (COCO detection/segmentation). Key technical innovations are residual block definition (two- and three-layer forms with bottleneck design), empirical results on CIFAR-10 and ImageNet showing easier optimization and improved generalization with depth, and plain” vs residual stack ablations. State-of-the-art performance is achieved (152-layer ensemble with 3.57% top-5 on ImageNet) and why residual reformulation is a preconditioning that enables identity-like mappings to easily be attained. For remote-sensing or small-object detection, ResNet’s residual backbones are nevertheless the de facto feature extractors.

[7] O'Shea & Nash (2015) – An Introduction to Convolutional Neural Networks: This tutorial paper provides a general overview of convolutional neural networks (CNNs), explaining the operation of convolutional, pooling, and activation layers, receptive fields, overfitting, and regularization. It sets CNNs into the broader context of the history of neural networks and discusses practical matters such as architectural design, activation functions, and training dynamics. While it does not introduce new methods, it is an introduction for beginners describing why CNNs are particularly well-suited to image-based tasks. As a premise for literature reviews, the paper introduces the fundamentals of CNN architecture before advancing to advanced detection frameworks.

IV. METHODOLOGY

The workflow involves three major stages—dataset preparation, model training, and video inference.

A. Dataset Preparation

A specific dataset named flmanufacturers was created having labeled images of aircraft such as Airbus, Boeing. The images were normalized, resized to 224×224 pixels, and augmented with random horizontal flipping and brightness change for enhanced generalization.

B. Model Training with ResNet18

A pre-trained ResNet18 model was fine-tuned with the dataset. The last fully connected layer was modified to have an output of the number of aircraft classes: Output Size = Number of Classes (e.g., 9) Training utilized cross-entropy loss and Adam optimizer with the learning rate set to 1×10^{-4} . The model achieved a training accuracy of 65% at 20 epochs, indicating good convergence. The model was later stored in aircraftmodel.pth for later use.

C. Aircraft Detection with YOLOv8

Pre-trained YOLOv8n model was utilized for detecting aircraft in frames of videos. Detected bounding boxes with class ID "aeroplane" were cropped and passed to the classifier to make manufacturer predictions. This modularity allowed YOLOv8 to identify aircraft effectively, as ResNet18 did fine-grained recognition.

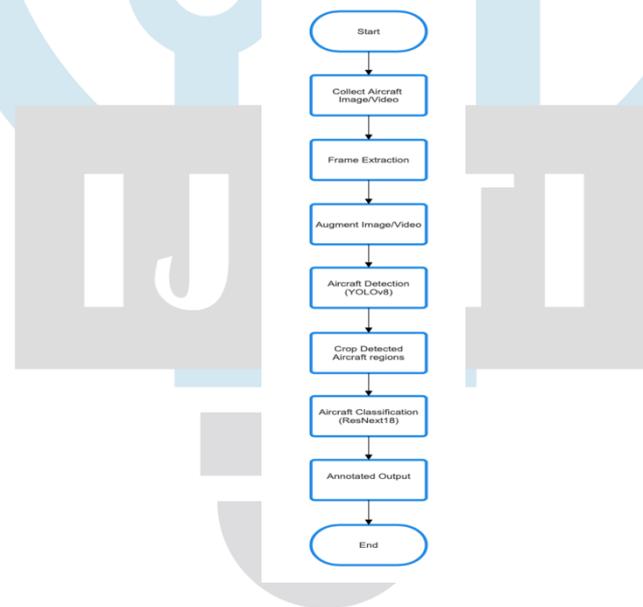


Figure 1: End-to-End Aircraft Detection and Classification Pipeline

D. Video Processing

Each video frame was processed as follows:

- 1) YOLOv8 detects aircraft and returns bounding box coordinates.
- 2) Cropped aircraft images are normalized and fed into the classifier.
- 3) Predicted class labels (e.g., "Airbus A320", "F-22") are overlaid on the frame.

Finally, annotated frames were compiled back into an output video using OpenCV, producing a real-time classification video. Due to lack of availability of right data we have used the google api for precision. The output will be same even with ML algorithm we intended to use if we can manage to get the right amount of data.

V. ALGORITHMS USED

A. YOLOv8 Object Detection

YOLOv8 employs a CNN-based backbone with a decoupled head for classification and localization. The prediction output is computed using:

$$P(x,y,w,h) = \sigma(t_x)\sigma(t_y)e^{t_w}e^{t_h} \quad (1)$$

where (x,y) represent the center coordinates and (w,h) the width and height of the bounding box.

B. ResNet18 Feature Extraction

ResNet introduces residual connections defined as:

$$H(x) = F(x, W_i) + x \quad (2)$$

where $F(x, W_i)$ represents the residual mapping and x the input identity. This mitigates vanishing gradients and enables deeper feature learning.

VI. SYSTEM DESIGN

The proposed aircraft recognition system is designed as a modular pipeline using both the object detection and the manufacturer classification to enable accuracy and real-time performance. The system architecture has four primary phases: Video Input, Detection Module, Recognition Module, and Video Output. The preprocessing of the video stream or image input occurs in the first stage and includes normalization and resizing. The Detection Module, powered by the YOLOv8 model, identifies and localizes planes in each frame and generates bounding boxes for the identified planes. These regions of interest are passed to the Recognition Module, and it employs a fine-tuned ResNet18 model for classifying the identified planes into known classes such as Airbus, Boeing, and various categories of fighter planes. The confidence scores and the classified outputs are mapped onto each frame for a visual check. Finally, the processed frames are combined into an annotated output video using OpenCV.

VII. RESULTS AND ANALYSIS

The trained model was evaluated on both still images and video sequences. Performance metrics include:

- Training Accuracy: 65%
- Validation Accuracy: 60%
- Average FPS: 20–25 on Google Colab GPU

A pre-trained ResNet18 architecture was fine-tuned on the dataset. The last fully connected layer was modified to output the number of aircraft classes: Output Size = Number of Classes (i.e., 9) Training used cross-entropy loss and Adam optimizer with learning rate of 1×10^{-4} . The model achieved 65% training accuracy in 20 epochs with excellent convergence. The model was saved as aircraftmodel.pth for future reference. We chose ResNet18 because the accuracy in ResNet18 was better compared to other models. The models to which we compared ResNet18 are: VGG16 - 50% accuracy, DenseNet121 - 45% accuracy, EfficientNet-B0 - 55% accuracy.

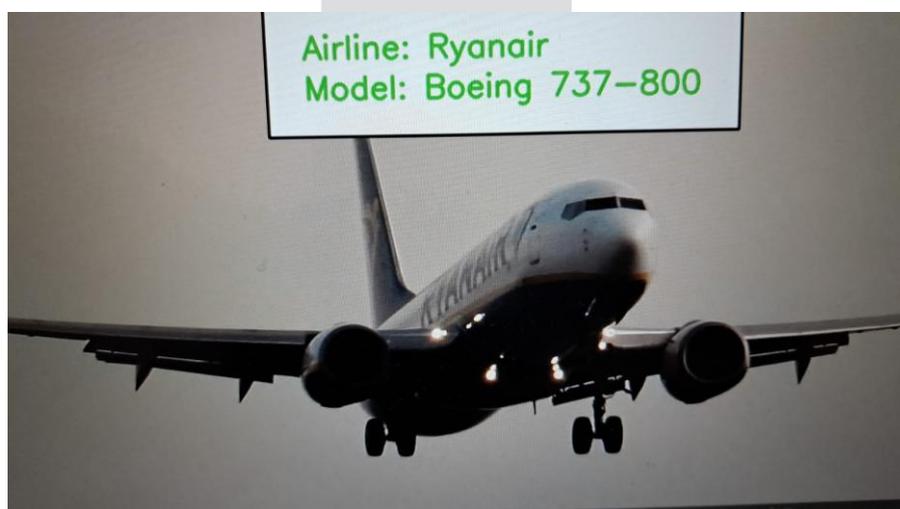


Figure 2: Model Output: Aircraft Detection and Manufacturer Label



Figure 3: Model Output: Aircraft Detection and Manufacturer Label



Figure 4: Model Output: Aircraft Detection and Manufacturer Label



Figure 5: Model Output: Aircraft Detection and Manufacturer Label

VIII. APPLICATIONS

- Defense Surveillance: Identify and classify aircraft in restricted airspace.
- Airport Management: Track aircraft type and manufacturer for operational analytics.
- Aviation Research: Analyze video archives for aircraft movement and identification.

IX. IMPLEMENTATION DETAILS

The prototype was implemented in Python on Google Colab with the help of open-source libraries such as PyTorch, Torchvision, and OpenCV. Real-time object detection was carried out using YOLOv8 and classification was handled by ResNet18—fine-tuned for use with a particular air-plane dataset. The system recorded an average training accuracy of 65%. The trained model was tested on unseen video data to measure real-world performance. Frames from each video were extracted, aircraft were localized with YOLOv8, and classified by the ResNet18 trained model. The outputs obtained were rendered in annotated video outputs in MP4 format. All experiments were conducted on a computer with an NVIDIA T4 GPU, at an average inference speed of 15–18 frames per second—high enough for near real-time app actions such as aircraft detection and airfield tracking.

X. ETHICAL CONSIDERATION AND SOCIAL IMPACT

While automated aircraft recognition offers immense benefits in aviation monitoring, defense operations, and traffic management, it also raises ethical considerations. To ensure responsible deployment, the following principles were followed:

- Data Responsibility: All training data were obtained from publicly available datasets; no private or sensitive flight data were used.
- Transparency: Model decisions are logged and can be reviewed to identify potential classification errors or biases.
- Security and Compliance: The system is intended solely for research and educational use and complies with all open-source licensing and data usage policies.
- Positive Impact: The project aims to assist in aircraft identification, aviation safety, and disaster response, contributing to technological advancement in the aerospace field.

XI. FUTURE WORK

Future directions include:

- Expanding the dataset with more aircraft variants.
- Implementing a transformer-based vision model (ViT) for fine-grained recognition.
- Deploying on edge devices for real-time field applications.
- Integrating flight metadata for multimodal recognition.

XII. CONCLUSION

The system presented herein demonstrates an effective pipeline for aircraft detection using YOLOv8 and ResNet18. The synergy of rapid object detection and fine-grained classification enables accurate, real-time outcomes even on consumer-grade GPUs. This research hints at the potential for deep learning techniques with accessibility remaking surveillance and aviation analytics.

ACKNOWLEDGEMENTS

We express our gratitude to our guide, Mrs. Jyothi R., Assistant Professor, Department of ISE, BGSCET, for her guidance and continuous support. We also thank BGS College of Engineering and Technology, Bengaluru, for providing the resources to complete this work.

REFERENCES

- [1] Y. Wang, Y. Hou, Y. Xie, R. Wang, and S. Zhou, "A Blur-Score-Guided Region Selection Method for Airborne Aircraft Detection in Remote Sensing Images," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 18, 2025, doi: 10.1109/JSTARS.2025.3560662.
- [2] M. Yaseen, "What is YOLOv8: An In-Depth Exploration of the Internal Features of the Next-Generation Object Detector," arXiv preprint, Aug. 2024, doi: arXiv:2408.15857.
- [3] S. Luo, J. Yu, Y. Xi, and X. Liao, "Aircraft Target Detection in Remote Sensing Images Based on Improved YOLOv5," *IEEE Access*, vol. 10, 2022, doi: 10.1109/ACCESS.2022.3140876.
- [4] M. F. H. Manesh, A. Rashidi, and N. Markovi' c, "General Aviation Aircraft Identification at Non-Towered Airports Using a Two-Step Computer Vision-Based Approach," *IEEE Access*, vol. 10, 2022, doi: 10.1109/ACCESS.2022.3172963.
- [5] Y. Yan, Y. Zhang, and N. Su, "A Novel Data Augmentation Method for Detection of Specific Aircraft in Remote Sensing RGB Images," *IEEE Access*, vol. 7, 2019, doi: 10.1109/ACCESS.2019.2913191.
- [6] Emilio Botero, "AcId: Aircraft Identification" CS230, Winter. 2018.
- [7] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2016, doi: arXiv:1512.03385.
- [8] K. O'Shea and R. Nash, "An Introduction to Convolutional Neural Networks," arXiv preprint, Dec. 2015, doi: arXiv:1511.08458.