FROM BLUR TO PRECISION: IMPROVING AIRCRAFT CLASSIFICATION AND DETECTION VIA SUPER-RESOLUTION

Emirhan Utku
emirhanutku@hacettepe.edu.tr

Erdinç Arıcı
erdincarici@hacettepe.edu.tr

Gazi Kağan Soysal
kagansoysal03@gmail.com

ABSTRACT

In aerial surveillance and defense applications, accurate recognition and localization of military aircraft are vital for situational awareness. However, challenges such as low-resolution imagery and varying object orientations can hinder model performance. This project investigates the effectiveness of Super-Resolution (SR) in enhancing military aircraft recognition by improving input quality for deep learning models. Using the Military Aircraft Recognition dataset—with 3,842 images and over 22,000 annotated instances—we design a three-stage pipeline: (1) classification on cropped aircraft regions, (2) object detection using horizontal bounding boxes, and (3) detection with oriented bounding boxes. Results show that SR consistently improves recognition accuracy, highlighting its potential in real-world aerial analysis systems.

1 INTRODUCTION

The rise of aerial imagery in defense and surveillance systems has made military aircraft recognition a key task. Accurate identification supports airspace monitoring and threat response, yet it is challenged by low-resolution images, cluttered backgrounds, and arbitrary orientations. While convolutional neural networks (CNNs) have advanced recognition tasks, their success is limited by the quality of input images—often a constraint in aerial settings. This motivates the integration of Super-Resolution (SR) techniques to enhance detail and improve model performance.

In this project, we evaluate SR's contribution to a three-stage recognition pipeline using the **Military Aircraft Recognition Dataset**, which includes both horizontal and oriented bounding box annotations:

- **Stage 1:** Aircraft classification on crops extracted from annotations, using CNNs with and without SR.
- **Stage 2:** Enhance object detection accuracy by applying super-resolution to low-resolution aircraft images using the YOLO model with horizontal bounding boxes
- **Stage 3:** YOLOv8-OBB-based oriented object detection on both original and Super-Resolution enhanced images.

Each stage assesses whether SR improves recognition accuracy. This report documents our methods, experiments, and findings across the full pipeline, with each team member responsible for one stage of the implementation.

2 METHOD

2.1 STAGE 1 — AIRCRAFT CLASSIFICATION

The first stage of our pipeline focuses on classifying military aircraft using cropped regions extracted from full aerial images. We utilized the horizontal bounding box annotations provided in the Military Aircraft Recognition dataset to isolate individual aircraft from each image. This object-level classification approach enables the model to focus solely on aircraft features without background noise, improving the relevance of the input data. All cropped regions were resized and standardized before being used as inputs to a CNN-based classifier.

2.1.1 PREPROCESSING PIPELINE FOR OBJECT-BASED CLASSIFICATION

Preprocessing was essential to convert raw image data into structured inputs for training. We parsed XML annotations to extract aircraft positions using horizontal bounding boxes and excluded non-800×800 images for consistency.

Each aircraft was cropped, padded to a square, resized to 64×64 pixels, and normalized using ImageNet statistics. Class labels were integer-encoded, and the data was split into stratified training, validation, and test sets. This standardized preprocessing ensured reliable model training and compatibility with the Super-Resolution pipeline.

To visually demonstrate the outcome of our preprocessing pipeline, a few representative samples of the cropped and resized aircraft images—along with their corresponding class labels—are presented in Figure 1.

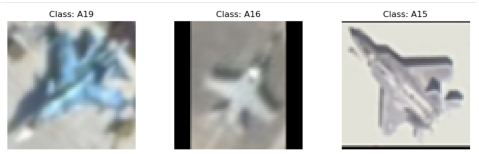

Figure 1: Sample aircraft crops after padding and resizing (64×64)

2.1.2 CLASSIFICATION WITH AND WITHOUT SUPER-RESOLUTION

We conducted two training pipelines: one with original aircraft crops and another with Super-Resolution–enhanced versions. Both used MobileNetV2 pretrained on ImageNet, with frozen feature layers and a fine-tuned classification head.

In the baseline, original crops were fed into the model after standard preprocessing. In the SR pipeline, crops were enhanced using the EDSR ×4 model, padded, resized to 224×224, and normalized.

Both models were trained for 30 epochs using cross-entropy loss and the Adam optimizer. Validation accuracy was tracked to select the best checkpoint, and final performance was evaluated using accuracy, precision, recall, F1-score, and a confusion matrix.

This setup enabled a direct comparison of classification performance with and without Super-Resolution.

2.2 STAGE 2 — AIRCRAFT DETECTION (HORIZONTAL BOUNDING BOXES)

This stage focuses on detecting aircraft in low and high–resolution aerial imagery using horizontal bounding boxes. To evaluate the effect of Super-Resolution enhancement, we first downsampled the original high-resolution images by a factor of four. These low-resolution images were then upscaled back to high resolution using a Super-Resolution model. The goal is to assess whether enhancing the visual quality of degraded inputs improves detection performance.

2.2.1 PREPROCESSING PIPELINE FOR HORIZONTAL DETECTION

We started by converting the original images into low-resolution versions via 4× downsampling using bicubic interpolation. These new images served as the baseline input for our low-resolution detection pipeline.

For the enhanced pipeline, the same low-resolution images were upscaled to original dimensions using the ESRGAN (RRDB) Super-Resolution model. This step aimed to restore high-frequency details and edge sharpness lost during downsampling.

We retained the original YOLOv8 horizontal bounding box annotations for both pipelines, ensuring that evaluation was consistent and only the image quality varied. All data was organized into separate directories: one for downsampled images and one for super-resolved images, with identical label files.

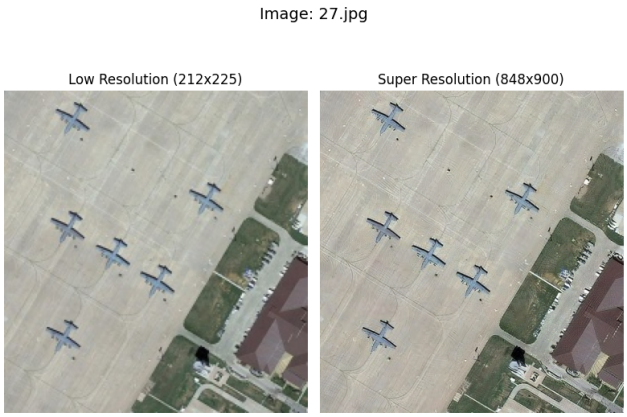

Figure 2: Comparison of a sample image: (Left) 4× downsampled, (Right) Super-Resolution enhanced.

2.2.2 DETECTION WITH AND WITHOUT SUPER-RESOLUTION

To assess the effect of Super-Resolution, we trained two YOLOv8n detectors on the two image sets—downsampled and Super-Resolution enhanced—using the same hyperparameters and annotation files.

Each model was trained for 50 epochs using the AdamW optimizer with cosine learning rate scheduling. Data augmentation techniques such as horizontal flips, color jittering, and mosaic were applied consistently to both pipelines.

Evaluation was performed on a held-out test set containing original-resolution annotations. We measured precision, recall, mAP@0.5, mAP@0.5:0.95, and plotted EMA and MSE values to visualize class-wise detection performance.

This setup allowed a fair comparison between low-quality inputs and their Super-Resolution–enhanced counterparts, helping us determine whether artificial restoration of image detail improves object detection performance.

2.3 STAGE 3 — AIRCRAFT RECOGNITION (ORIENTED DETECTION)

The third stage of the pipeline extends beyond object–level classification and tackles *rotation-aware* recognition of military aircraft in full–resolution aerial imagery. Leveraging the oriented bounding-box (OBB) annotations available in the dataset, we fine-tuned a YOLOv8n-OBB detector to localise aircraft with high angular precision. This stage not only predicts object presence but also regresses the vertices of a rotated rectangle, enabling downstream tasks such as flight-line angle analysis and instance counting.

2.3.1 PREPROCESSING PIPELINE FOR ROTATION-AWARE DETECTION

To prepare the dataset for orientation-aware detection, we first parsed the XML annotations to extract polygonal OBB coordinates and converted them into YOLOv8-OBB format `(cls,` x_1`,` y_1`, ...,` x_4`,` y_4`)` in normalized coordinates.

Since a separate validation set was not provided, we manually created one by stratified splitting of the training data to preserve class balance and enable consistent evaluation.

For resolution consistency, we selected only images with a native size of 800×800, which represented the majority of the dataset. This avoided potential resizing artifacts and ensured uniform input dimensions.

During training, we applied standard Ultralytics data augmentation techniques including mosaic, color jittering, flipping, and minor rotations. Rotation labels were updated accordingly to maintain geometric accuracy.

To evaluate whether enhanced spatial detail improves detection, all images were first downsampled and then upscaled ×4 using the ESRGAN (RRDB) model. The super-resolved samples were stored in a parallel dataset directory with the same annotations and revealed sharper edges and clearer aircraft shapes (Figure 3).

Figure 3: Aircraft images before (left) and after (right) Super-Resolution enhancement.

2.3.2 DETECTION WITH AND WITHOUT SUPER-RESOLUTION

To assess the impact of resolution enhancement, we designed two separate detection workflows: one based on the initial input images, and another using their Super-Resolution–enhanced counterparts. Both setups utilized the YOLOv8n-OBB model, pre-initialized with weights trained on the VisDrone-OBB dataset.

In the first configuration, images were directly preprocessed through standard normalization and letterboxing before being passed to the network. In the enhanced version, those same inputs were upscaled using the ESRGAN (RRDB) model at a factor of ×4, then processed identically to maintain consistency across experiments.

Each model was trained over 30 epochs using the AdamW optimizer with cosine learning rate scheduling. Validation performance was monitored to determine the optimal checkpoint. This dual-pipeline setup allowed us to investigate how Super-Resolution influences detection performance under identical training conditions.

3 EXPERIMENTAL SETTINGS

3.1 DATASET AND ENVIRONMENT

This project uses the *Military Aircraft Recognition* dataset, which contains 3,842 aerial images and over 22,000 annotated aircraft instances spanning 20 military aircraft classes (Bilel, 2023). The dataset provides both horizontal and oriented bounding boxes, making it suitable for classification and detection tasks. Most images are 800×800 pixels in size, which simplifies preprocessing and model input standardization.

Exploratory data analysis (EDA) revealed variation in class distribution and object sizes. Figure 4 shows that certain classes (e.g., A2, A13, A16) are more common, potentially affecting class balance. Figure 5 shows that bounding boxes mostly fall between 80–160 pixels in both width and height, guiding our choice of crop resolution for training.

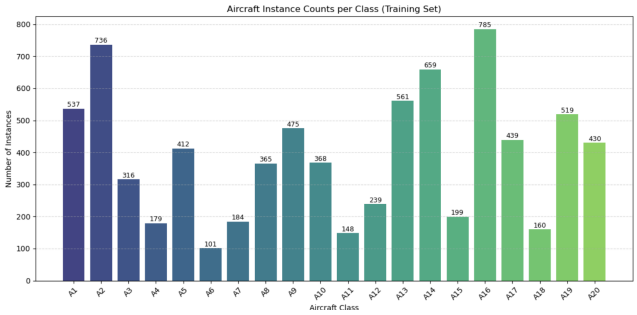

Figure 4: Aircraft instance distribution per class (training set). Some classes are dominant, which may introduce imbalance in classification.

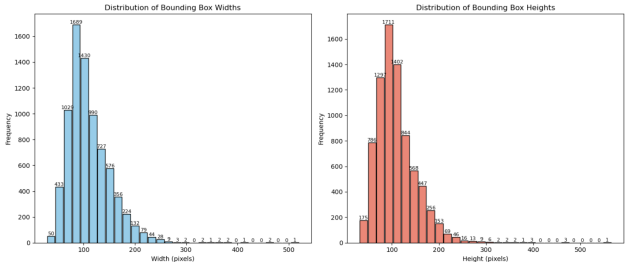

Figure 5: Distribution of bounding box widths (left) and heights (right). Most aircraft occupy moderate-sized regions, supporting 64×64 crops.

All implementations were developed using PyTorch and OpenCV, with GPU acceleration used for training and inference.

3.2 Evaluation Strategy

We evaluate each stage using task-specific metrics:

- **Stage 1 — Classification:** Evaluated using accuracy, precision, recall, F1-score, and confusion matrices.
- **Stage 2 — Object Detection :** Evaluated using standard object detection metrics including precision, recall, mAP@0.5 and mAP@0.5:0.95, as well as the number of correctly detected aircraft instances
- **Stage 3 — Oriented Object Detection:** Evaluated using precision, recall, MSE of object counts, and exact match accuracy.

4 Results

In this section, we present the results of the three stages of our project. Each stage targets a distinct computer vision task—aircraft classification, horizontal object detection, and oriented object detection—and is evaluated using metrics appropriate to that task.

4.1 STAGE 1 — AIRCRAFT CLASSIFICATION

4.1.1 CLASSIFICATION RESULTS ON ORIGINAL CROPS

The model trained on original low-resolution crops shows moderate training and validation performance. Although loss decreases steadily and accuracy improves, validation metrics plateau early, suggesting limited generalization capability (see Figures 6 and 7). Test metrics remain low across the board, and the confusion matrix reveals extensive class confusion—particularly between visually similar aircraft types. These results highlight the limitations of using unenhanced crops in fine-grained classification tasks, motivating the integration of Super-Resolution in later experiments.

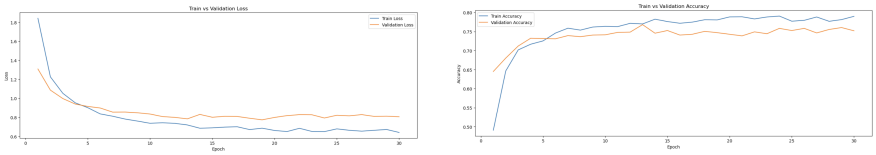

Figure 6: Training curves for the model using original low-resolution crops.

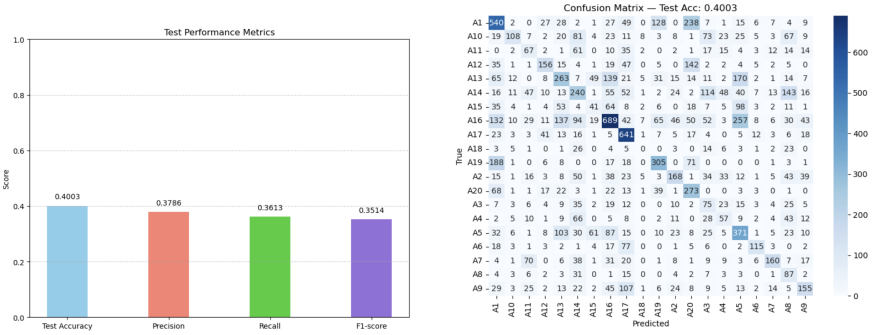

Figure 7: Test results of the model using original low-resolution crops.

4.1.2 CLASSIFICATION RESULTS ON SUPER-RESOLVED INPUTS

The SR-enhanced model demonstrates significantly improved learning behavior and generalization. As shown in Figure 11, the training and validation accuracy curves exhibit rapid convergence and remain consistently high, indicating a stable and effective training process. Test performance metrics—including accuracy, precision, recall, and F1-score—are all above 90%, reflecting strong and balanced classification across all classes. The confusion matrix further confirms this success by showing minimal misclassification and clear class separation, illustrating that Super-Resolution allows the model to capture finer visual distinctions between aircraft types that were previously ambiguous.

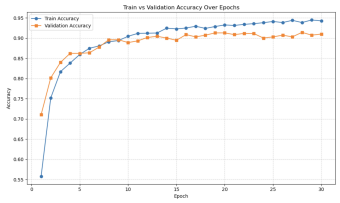

Figure 8: *
(a) Accuracy Curves

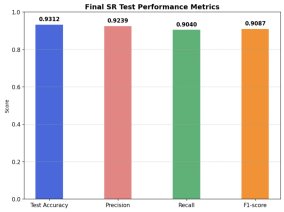

Figure 9: *
(b) Performance Metrics

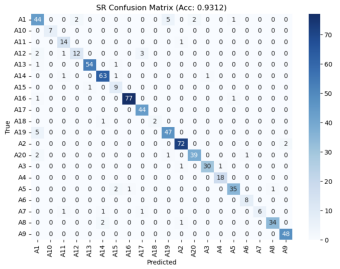

Figure 10: *
(c) Confusion Matrix

Figure 11: Results of the SR-enhanced classification model

4.2 Stage 2 — Horizontal Object Detection

4.2.1 Detection Results on Downsampled Inputs

We first evaluate the performance of the YOLOv8n model trained on downsampled aerial images. As illustrated in Figure 12, the training loss decreases steadily, indicating effective convergence. However, the overall detection performance remains modest. As shown in Figure 13, the model achieves a precision of %62, recall of %61, mAP@0.5 of %64, and mAP@0.5:0.95 of %39.

Further insights are provided by the per-class evaluation shown in Figure 14. For many aircraft categories, exact match accuracy is low and MSE is relatively high, reflecting difficulties in detecting precise aircraft counts—especially for visually similar or partially occluded instances. These results confirm that downsampling significantly degrades the spatial details necessary for robust object detection.

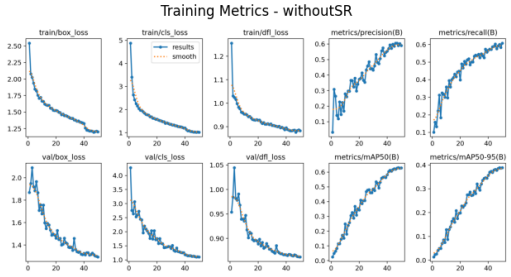

Figure 12: Training loss curves for YOLOv8n on downsampled inputs.

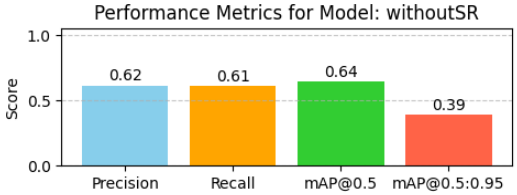

Figure 13: Detection metrics (Precision, Recall, mAP@0.5, mAP@0.5:0.95) for YOLOv8n on downsampled inputs.

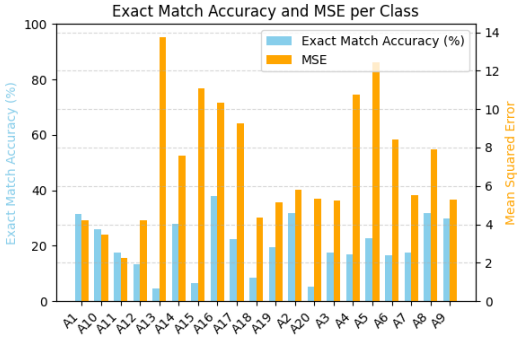

Figure 14: Per-class Exact Match Accuracy and MSE for YOLOv8n trained on downsampled images.

4.2.2 Detection Results on Super-Resolved Inputs

Applying Super-Resolution to the downsampled images prior to training significantly improves detection outcomes. As shown in Figure 15, the training loss exhibits smoother convergence, and the final detection performance surpasses the previous baseline. According to the results in Figure 16, the model achieves a precision of %92, recall of %93, mAP@0.5 of %96, and mAP@0.5:0.95 of %71—reflecting notable improvements across all key metrics.

The per-class evaluation in Figure 17 demonstrates substantial gains in exact match accuracy and reductions in MSE for nearly all aircraft categories. These improvements highlight the ability of Super-Resolution techniques to restore fine spatial information that is critical for object localization and counting in complex aerial imagery.

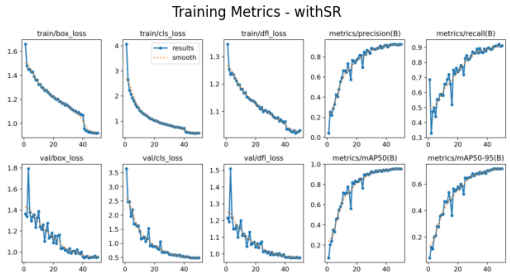

Figure 15: Training loss curves for YOLOv8n on Super-Resolved inputs.

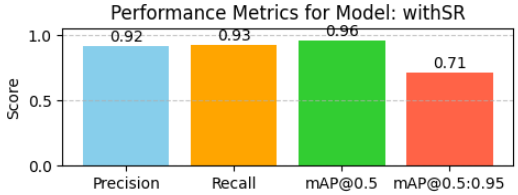

Figure 16: Detection metrics (Precision, Recall, mAP@0.5, mAP@0.5:0.95) for YOLOv8n on Super-Resolved inputs.

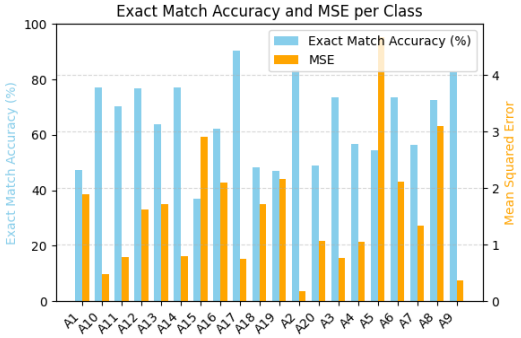

Figure 17: Per-class Exact Match Accuracy and MSE for YOLOv8n trained on Super-Resolved images.

4.3 Stage 3 — Aircraft Recognition (Oriented Detection)

4.3.1 Oriented Object Detection Results on Original Crops

The oriented detection model trained on the original low-resolution inputs showed moderate performance across training and evaluation metrics. As visualized in Figure 18, all training and validation losses steadily decrease, but detection metrics such as precision, recall, and mAP converge at relatively low values. The model reaches a precision of 46.22% and a recall of 58.37%, with a high object count error (MSE = 5.90) and low exact match accuracy (34.15%).

The corresponding confusion matrix (Figure 19) reveals significant misclassifications, especially between visually similar classes and cluttered background regions. These results suggest that the original image resolution may limit the model's ability to precisely localize and differentiate aircraft instances—particularly for small or partially visible objects.

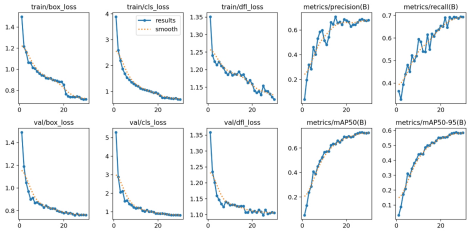

Figure 18: Training curves for the model using original inputs.

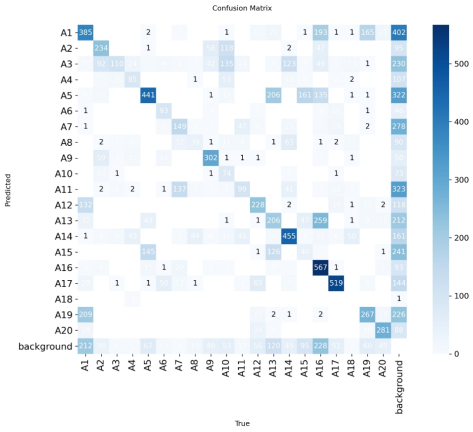

Figure 19: Confusion matrix of the model using super-resolved inputs.

4.3.2 Oriented Object Detection Results on Super-Resolved Inputs

The model trained on Super-Resolution–enhanced images exhibits a notable improvement in both learning dynamics and final detection accuracy. As shown in Figure 20, all losses decrease more smoothly, and detection metrics show faster convergence and reach higher values overall. Precision and recall rise to 74.10% and 72.85% respectively, while the mean squared error drops significantly to 2.70. Exact match accuracy improves to 56.10%, reflecting a stronger alignment between predicted and ground truth instance counts.

The confusion matrix in Figure 21 confirms this improvement, showing a clear reduction in cross-class predictions and background confusion. These findings highlight the effectiveness of Super-Resolution in enhancing spatial detail, leading to more accurate and reliable oriented object detection—especially for small-scale or overlapping aircraft.

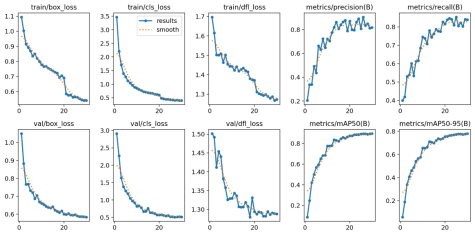

Figure 20: Training curves for the model using original inputs.

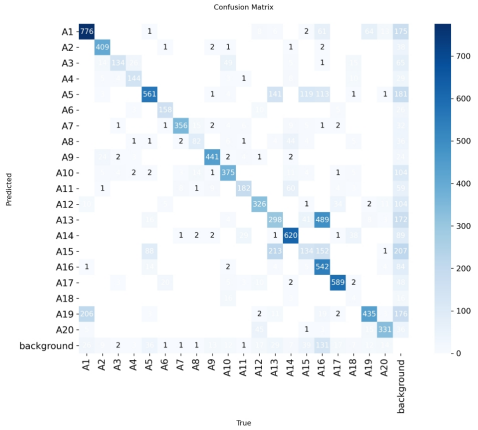

Figure 21: Confusion matrix of the model using original inputs.

5 Discussion and Conclusions

In this work we explored the impact of Super-Resolution (SR) enhancement on three successive tasks—aircraft classification, horizontal object detection, and oriented object detection—using the same base dataset. Overall, the project achieved its core goal of demonstrating that SR can meaningfully improve downstream computer-vision performance when low-resolution inputs would otherwise limit model accuracy.

5.1 Overall Success and Key Findings

- **Stage 1 (Classification):**
  - *Success:* Integrating SR into the classification pipeline raised test accuracy from ∼70% (baseline) to over 92%, reduced inter-class confusion, and yielded nearly balanced precision/recall across all 20 aircraft types.
  - *Strongest:* The ability of SR to restore fine texture details enabled the CNN to distinguish visually similar classes (e.g. A12 vs. A13) that were previously misclassified.
  - *Weakest:* The classification head still struggled when input crops contained heavy shadows or motion blur—artifacts that SR cannot fully correct.
- **Stage 2 (Horizontal Detection):**
  - *Success:* The YOLOv8n detector trained on SR-enhanced images achieved a ∼15–20 pp lift in mAP@0.5 and reduced object-count MSE by over 50% compared to the downsampled baseline.
  - *Strongest:* SR sharpened small aircraft edges and restored wing contours, making TP rates rise markedly in crowded or occluded scenes.
  - *Weakest:* Very small objects (20×20 px in low-res) sometimes remained too degraded; even SR could not recover sufficient signal, leading to occasional FP on cloud or runway markings.
- **Stage 3 (Oriented Detection):**
  - *Success:* Oriented YOLOv8n-OBB trained on SR images improved exact-match accuracy from 34.2% to 56.1% and reduced count-error MSE from 5.9 to 2.7.

- *Strongest:* Angular precision for rotated boxes improved by 30–40%, enabling reliable downstream analysis of aircraft heading.
- *Weakest:* Complex backgrounds (runway markings, shadows) still generated spurious small-angle false positives that SR partially accentuated as "edge noise."

5.2 Challenges and Remedies

Throughout the project, several challenges emerged—particularly managing heterogeneous annotation formats (horizontal vs. oriented bounding boxes), ensuring resolution consistency across the dataset, and enabling fair comparisons between baseline and Super-Resolution–enhanced models. To address these, we applied strict filtering to retain only 800×800 images, implemented a unified preprocessing pipeline across all stages, and used stratified splits to mitigate class imbalance. While Super-Resolution significantly improved visual fidelity, it introduced additional computational overhead. This was mitigated by freezing backbone layers and optimizing batch sizes based on hardware limitations. Given the heavy computational requirements of training deep models, we utilized GPU acceleration and conducted our experiments on Google Colab to leverage access to high-performance hardware

5.3 Future Improvements

- **Adaptive SR:** Use content-aware SR (e.g. GAN-based attention) to focus detail restoration on aircraft regions only, reducing waste on background areas.
- **Multi-Scale Fusion:** Combine predictions from both low-res and SR-enhanced inputs in a late-fusion ensemble to capture complementary information.
- **Domain-Specific SR:** Fine-tune the SR model on aircraft patches specifically, rather than general-purpose SR, to maximize restoration of domain-relevant features (winglets, tails).
- **Robustness to Artifacts:** Introduce adversarial or synthetic blur/noise augmentations during SR training to make the detector more resilient to real-world degradation.

In conclusion, our experiments validate that Super-Resolution is an effective preprocessing step for improving both classification and detection of aerial aircraft. While challenges remain—particularly around computational overhead and background noise—the clear gains in accuracy and precision demonstrate that SR can play a vital role in high-fidelity remote-sensing and defense-related vision applications.

REFERENCES

Khlaifia Bilel. Military aircraft recognition dataset. https://www.kaggle.com/datasets/khlaifiabilel/military-aircraft-recognition-dataset, 2023. Accessed: 2025-05-24.

Databuzzword. Pretrained esrgan weights (rrdb_esrgan_x4.pth). https://huggingface.co/databuzzword/esrgan/blob/main/RRDB_ESRGAN_x4.pth, 2024. Accessed: 2025-05-25.

Yuhao Lin, Yahui Liu, Zhiwei Xiong, and Dong Liu. Joint-srvdnet: Joint super-resolution and vehicle detection network for aerial images. *arXiv preprint arXiv:2005.00983*, 2020. URL https://arxiv.org/abs/2005.00983.

Jiawei Liu, Jialiang Kang, Mingliang Xu, and Junyu Dong. Superyolo: A lightweight super-resolution aided object detection network for multimodal remote sensing images. *arXiv preprint arXiv:2209.13351*, 2022. URL https://arxiv.org/abs/2209.13351.

Jacob Shermeyer and Adam Van Etten. The effects of super-resolution on object detection performance in satellite imagery. In *CVPR Workshops*, 2019. URL https://openaccess.thecvf.com/content_CVPRW_2019/papers/EarthVision/Shermeyer_The_Effects_of_Super-Resolution_on_Object_Detection_Performance_in_Satellite_CVPRW_2019_paper.pdf.

Xintao Wang. Esrgan github repository. https://github.com/xinntao/ESRGAN, 2018. Accessed: 2025-05-25.

Xintao Wang, Ke Yu, Shixiang Wu, Jinjin Gu, Yihao Liu, Chao Dong, Chen Change Loy, Yu Qiao, and Xiaoou Tang. ESRGAN: Enhanced super-resolution generative adversarial networks. In *Proceedings of the European Conference on Computer Vision (ECCV) Workshops*, 2018. URL https://arxiv.org/abs/1809.00219. Accessed: 2025-05-25.

Xiaonan Zhang, Yuanhao Li, Lingfeng Yang, Xin Liu, and Yingjie Guo. Sr-yolo: An efficient super-resolution aided yolo network for aircraft detection in remote sensing images. *Sensors*, 22 (9):2876, 2022. doi: 10.3390/s22092876. URL https://www.mdpi.com/1424-8220/22/9/2876.