Evaluating YOLO V8 Models for Efficient Drone Detection in Anti-UAV Systems

Field - Computer Vision, Machine Learning, Artificial Intelligence, Computer Science

1st Kushaj Pavankumar Solanki
Department of Computer Science
Vellore Institute of Technology
Amaravati, India
kushajsolanki0007@gmail.com

2nd Puneet Verma
Department of Computer Science
Vellore Institute of Technology
Amaravati, India
puneet.22bce9208@vitapstudent.ac.in

3rd Prof. Suresh Dara
Department of AI, ML
Vellore Institute of Technology
Amaravati, India
dara001@gmail.com

***Abstract*—The rising UAV air presence in areas of restricted airspace now raises concerns in terms of security, safety, and privacy issues. An efficient anti-UAV detection system has gained a high demand. The paper delves into the presentation of a comparative analysis between two variants of YOLOv8, YOLOv8-S and YOLOv8-L, which may be used for the purpose of UAV detection. The performance of both models was checked based on the following parameters: accuracy, pre-cision, F1-score, and mean average precision (mAP). YOLOv8 brings some new innovations compared to previous YOLO variants, such as improved feature extraction, better backbones, improved prediction heads, and so on, resulting in enhanced detection performance. These comparisons indicate that both the variants, YOLOv8-S and YOLOv8-L, are doing outstandingly well at the detection of small, fast-flying drones and this is with huge differences in their precision-speed trade-off. Some degree of distinction between the different variants of YOLOv8 can be observed in terms of size and complexity of the model. Although larger models tend to be more accurate, they are also slower, while the small, less complex models are faster, albeit less accurate. The comparison helps in identifying the trade-offs and guides selection based on optimal models for a range of anti-UAV applications such as airspace security, wildlife protection, border surveillance,and public event monitoring.**

***Index Terms*—YOLOV8, AntiUAV, Computer Vision, Comparative analysis, Evaluation**

SECTION 1. INTRODUCTION

- In recent times, throughout the globe, UAVs have increasingly been used for leisure activities and aerial photography. Despite their widespread usage, personal as well as commercial, the military is still atop the list of UAV development. Military UAVs are indeed from the very drawing board designed according to conventional aircraft but add a great number of enhancements to complement flights at long ranges and greater altitudes. The UAV industry has grown meteorically during the last decade, and therefore, demonstrated its applicability across governmental, commercial, and civilian sectors.
- It is pretty evident that UAV technology assumes strategic importance due to its role in modern conflicts. For instance, during the current Russia-Ukraine war, anti-UAV systems are playing a critical role in defense operations. The need for Ukraine has grown to detect and destroy increasingly used Russian reconnaissance, surveillance, and combat UAVs. While effective for protection of high-value assets, increased battlefield situational awareness, and aerial threat denial, the requirements of advanced anti-UAV systems will become even more pronounced based on such militarily-centered applications but extend far beyond the realm of military use as well as for national security and public safety using better YOLOV models by comparing them with different variants as in [1].
- Deep learning techniques have continued to proliferate in the recent past. Models such as YOLO (You Only Look Once) have revolutionized real-time object detection due to their accuracy and high efficiency. The latest in the YOLO series, YOLOv8, exhibited superior performance in various domains and applications, including agricultural domains and surveillance, which require the rapid and accurate detection of objects as in [2]. These suggest that there is hope in developing powerful anti-UAV systems to work in real-time UAV identification and tracking in dynamic environments.
- This paper presents an anti-UAV system in two phases: it is designed as a short-range detection and identification of UAVs. To conduct a holistic analysis, two datasets were considered: one solely made up of images of drones to test in a controlled environment, and another consisting of bird and drone images, where the system's efficiency in distinguishing between highly similar objects is tested. For measuring performance between YOLOv8-S and YOLOv8-L models, the metrics accuracy, precision, recall, and F1 score were used.
- The YOLOv8-S model is lightweight and resource-efficient, making it a solid candidate for real-time detection in highly constrained resources. On the other hand, the computationally intensive architecture of YOLOv8-L makes it a candidate for applications where more precision is needed. The authors seek, in this comparative study, balance between speed and accuracy to provide insights into adjustment of anti-UAV systems with op-

erational needs. Further scope of ensembling benefits about the models is realized based on the advantages of YOLOv8-S and YOLOv8-L in enhancing detection and identification capabilities.

A. Organization of Paper

This paper will be structured to support the development, implementation, and evaluation of an anti-UAV detection system using YOLOv8 variants. In the introduction, it presents increasing relevance of UAVs, applications, and their critical role in modern warfare and national security, as well as the rationale for using YOLOv8-S and YOLOv8-L models.

- The Methodology section explains the detailed approach adapted by this research, which include steps on the choice of YOLOv8 variants, characteristics of the datasets, preprocessing techniques, and the training process. The importance of dataset design is also explained here, citing how it enhances the robustness of the system as it deals with various operational scenarios.
- In the Experiments, Results, and Discussion section, the experimental setup with hardware and software specifications is presented. The paper compares the detection capabilities of YOLOv8-S and YOLOv8-L on different GPUs, discussing the speed to accuracy trade-offs for different model variants. The use of confusion matrices in the evaluation of detection capability is explained in detail as an insight into the strengths and limitations of each model variant.
- Finally, the Conclusions section summarizes the study's findings, focusing on the potential for application of each variant of YOLOv8 in real-life scenarios based upon operational requirements. This section will also discuss potential avenues for future work, such as enhancements of the architecture and additional experiments in challenging environments.

SECTION 2. LITERATURE STUDY

The detection and tracking of unmanned aerial vehicles (UAVs) have emerged as critical tasks in various fields, ranging from defense and security to commercial airspace management. With the proliferation of drones, the need for robust, real-time detection systems has led to the exploration of advanced object detection models, particularly in the realm of deep learning. This section reviews existing work in the field and highlights how our research builds upon and diverges from previous methodologies.

In the study conducted as in [5], a comparative evaluation of YOLOv5, YOLOv8, and YOLOv8 DeepSORT was performed to enhance UAV detection and tracking systems. The research focused on improving real-time detection and tracking, a key requirement in aerial surveillance applications, through the assessment of key performance metrics such as precision, recall, and mean Average Precision (mAP). Their findings demonstrated the potential of YOLOv8, particularly in its tracking capabilities when paired with DeepSORT. Our study aligns with their focus on real-time detection but moves further by conducting an in-depth comparison between two variations of the YOLOv8 model: YOLOv8-S and YOLOv8-L. Additionally, [5] evaluated model performance primarily on datasets featuring UAVs, our research incorporates two distinct datasets: one focusing solely on UAVs and another that includes both UAVs and birds. This dual dataset approach allows us to address practical obstacles that arise in real-world scenarios, such as distinguishing between drones and other airborne objects like birds, which can lead to false detections.

Similarly, [4] investigated the efficacy of the YOLOv5 and YOLOv8 architectures for instance segmentation in corrosion detection, using datasets enhanced by data augmentation techniques like motion blur, rotations, and exposure changes to simulate challenging real-world conditions. Their goal was to improve model robustness in detecting corrosion patterns. Although our work shares the broader goal of enhancing model performance under varying conditions, we differ in both the application and methodological approach. While work as in [4] focused on corrosion detection and utilized specific augmentations (blur, lighting variations), our research applies a more focused approach to UAV detection without the inclusion of motion blur, lighting variations, or noise augmentations. Instead, we prioritize model performance comparisons between YOLOv8-S and YOLOv8-L across two distinct datasets, evaluating metrics like F1 score, precision, recall, mAP, mAP@0.5, and accuracy to determine their effectiveness in UAV detection scenarios.

In addition, the foundational work by Ultralytics introduced significant advancements in the YOLOv8 architecture, particularly the incorporation of the C2f module and an anchor-free detection head, both of which improve computational efficiency and detection accuracy. These improvements are integral to our study, as we seek to assess how the architectural differences between YOLOv8-S and YOLOv8-L translate into performance disparities, especially in environments with fluctuating drone speeds, lighting conditions, and altitude changes. While the advancements introduced by Ultralytics have been explored for general object detection tasks, our research narrows the focus to UAV detection, incorporating real-world datasets that include challenging variables like object occlusion and environmental interference.

A. Comparisons and Similar Approaches

- The object detection algorithm is part of an intelligent surveillance system and the basic one in the identity identification domain, which is important in practical terms. YOLO series algorithms are of high speed and accuracy, and every version of YOLO since its appearance has continued to set a benchmark in the domain of object detection. We will conduct experiments with three very popular versions of the YOLO model: YOLOv3, YOLOv4, and YOLOv5, including YOLOv5l, YOLOv5m, YOLOv5s, YOLOv5x. The performance of the above YOLO models is analyzed and summarized by training and predicting on the public VOC dataset. Results show that although YOLOv4 eclipses YOLOv3 in mAP values but lags a bit behind in speed, whereas the YOLOv5 series outperforms YOLOv3 and YOLOv4 in both mAP values and speed as in [3].

- Our study enhances earlier approaches found in related research works and brings in new additions to the present work, which enhance UAV detection in real-time applications. For instance, advanced data augmentation techniques, such as motion blur and illumination changes, have been used in the work as in [4] for superior robustness in corrosion detection. Motivated by that, our work takes an away step toward exploring YOLOv8 variants in greater depth: namely YOLOv8-S and YOLOv8-L and its unique challenge of UAV detection. We rather highlight fine-tuned models than a dependence on augmentations, implementing two datasets: one purely drones-only and another with both drones and birds. The design enables us to tackle false positives, which form a major problem in distinguishing UAVs from similar-sized airborne objects like birds, through comprehensive evaluation metrics such as precision, recall, F1 score, mAP, mAP@0.5, and accuracy.
- Similarly, [5] works with UAV tracking using YOLOv5, YOLOv8, and YOLOv8 DeepSORT. Although their work is a demonstration of the usefulness of tracking in the integration, our work goes beyond and into more detail for comparing YOLOv8-S and YOLOv8-L performance, focusing on real-world environmental conditions. To make it complete, we must include a second dataset which includes drones and birds, fill the critical gap between UAV and non-target objects, improve the training of models with respect to detection reliability in operational conditions.
- Unlike previous works, our work looks at the architectural differences and dataset-specific effects of YOLOv8 variants rather than depending on types of augmentations or tracking methods to make models adaptable to environmental changes. This study assesses the performance of YOLOv8-S and YOLOv8-L across scenarios with changing lighting, varying drone speeds, and altitudes, providing a nuanced understanding of size-performance trade-offs. These findings will aid improvements to UAV detection systems by revealing model configurations that are most suited to real-time deployment under specific conditions.

B. Contributions of Our Research

In summary, our research highlights the following key contributions:

- Our work builds upon the research as in [5], and [4], and Ultralytics by focusing on a deeper comparison of YOLOv8-S and YOLOv8-L models in the context of UAV detection.
- We introduce a novel approach by incorporating two datasets:
  - One dataset exclusively for UAVs.
  - Another dataset combining UAVs and birds to address practical challenges not explored in detail in previous studies.
- Our findings contribute to the development of robust anti-UAV systems by providing:
  - Comprehensive insights into the performance of YOLOv8 models under real-world conditions.
  - Results obtained without relying on extensive data augmentations such as motion blur or lighting changes.
- This research advances the fine-tuning of YOLOv8 models for UAV detection in dynamic, real-time scenarios.
- Our work enhances the accuracy, reliability, and applicability of these systems in practical surveillance environments.

SECTION 3. METHODOLOGY

YOLOv8 is a major leap in the YOLO family of object detection algorithms, especially designed for higher accuracy and real-time detection performance. In this paper, we utilize two variants of YOLOv8: YOLOv8-S and YOLOv8-L. We have picked up these models since they vary in complexity, and with one model being quite larger than the other, it is possible for us to investigate the trade-off between model size, detection accuracy, and inference speed.

A. Dataset and Split

The dataset used in this study was sourced from Roboflow and contained a total of 778 images for Drone model training and 7469 images for Combined Bird-Drone model training. Dataset referred as in [7]. These images were split into three subsets:

- **Training Set**:
  - Used for model training.
  - Drone model training - 545 images (70%)
  - Combined Bird-Drone model training - 6480 images (87%)
- **Validation Set**:
  - Used for hyperparameter tuning and performance validation during training.
  - Drone model training - 155 images (20%)
  - Combined Bird-Drone model training - 348 images (5%)
- **Test Set**:
  - Reserved for evaluating the final performance of the models.
  - Drone model training - 78 images (10%)
  - Combined Bird-Drone model training - 647 images (9%)

The dataset includes annotated images of drones and birds, which were used to assess the ability of YOLOv8-S and YOLOv8-L to discriminate between UAVs and other similar objects in complex environments. Data augmentation techniques, as described earlier, were applied to the training set to increase variability and enhance model robustness.

B. Data Preprocessing

During the pre-training processing of the dataset, it was optimized to be done for better performance:

- **Auto-Orientation**: An automatic orientation adjustment was applied to correct the alignment of images, ensuring that all objects were consistently oriented for effective detection.
- **Resize**: All the images were resized to 640x640 pixels for consistency in the dimension of input into the YOLO model. The size is a good trade-off between detail and computational efficiency.
- **Augmentations**: There were no additional data augmentations done on the dataset since, in this case, the focus was on how well the model performed on the original data distribution.

The following preprocessing steps improved the data quality and prepared it for training and testing of the YOLOv8 models.

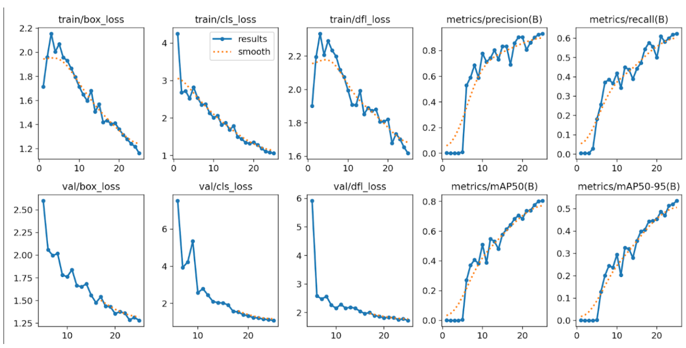

Fig. 1. Epoch based training of Drone model using YOLOv8-L.

C. YOLOv8-S Architecture

A smaller and lighter version of YOLOv8 models that can be optimized toward efficiency in real-time, low-end, or resource-constrained device application. Features are characterized by:

- **Backbone**: Generally, in YOLOv8-S, an improved version of (Convolutional Neural Network) CNN forms part of the backbone for feature extraction. The backbone is shallow enough to still achieve the desired accuracy on object detection tasks; in particular, even on detection of small objects such as drones.
- **Anchor-Free Detection Head**: The YOLOv8-S uses an anchor-free detection head that includes the direct prediction of center coordinates of the object. This removes the commitment for predefined anchor boxes, which saves computation time and the accuracy of the model can be improved in the case of small and fast-moving objects.
- **Residual Connections**: YOLOv8-S includes residual connections. This implies that the network is capable of learning deeper feature representations without suffering from the vanishing gradient problem, and hence trains faster.
- **Activation Functions**: There will be the usage of Leaky ReLU activation function in early layers so that characteristics are well learned in low-light conditions, which are usual in the case of a drone.

D. YOLOv8-L Architecture

YOLOv8-L - the much larger and more complex variant for those applications where detection accuracy is desired but comes with much higher computational cost. Key features of this architecture include:

- **Deeper Backbone**: Compared to YOLOv8-S, the backbone is deeper with more layers in YOLOv8-L, meaning it can learn more complex features hence improving its ability to make more accurate detections in cluttered or challenging environments.
- **Advanced Feature Pyramid Network (FPN)**: The YOLOv8-L uses an advanced FPN in order to have multi-scale feature extraction. This makes the model pretty accurate at picking up objects at different scales, very useful in particular in UAVs detecting flights where the drones appear at different sizes due to the difference in their altitudes.
- **Prediction Head**: YOLOv8-L combines the anchor-free approach with larger receptive fields to further finely localize UAVs, even in the case of distant or occluded UAVs.
- **Optimized Inference**: YOLOv8-L is optimized for high-performance computing environments, especially those on GPUs. Although larger in size and more complex than other versions, it's a good choice when the case is something where speed is de-emphasized but detection accuracy takes precedence.

Section 4. Experiments, Results and Discussions

A. Experiment Setup

Detailed experimental preparation was done to evaluate the performance of YOLOv8-S and YOLOv8-L models on drone and mixed drone-bird detection functions with different computation settings. To analyze the impact of differences in hardware capability on the execution of models, the experiments were done using two high-performance GPUs: NVIDIA Tesla T4 and NVIDIA RTX 4050. The framework provided by the Ultralytics YOLOv8, Python, was used in implementing the models for smooth experimentation and evaluation. A carefully curated dataset was utilized, encompassing diverse scenarios to ensure robustness. The setup facilitated systematic evaluation using metrics like precision, recall, mAP@0.5, and accuracy, with detailed analysis supported by confusion matrices. This ensured comprehensive insights into model efficiency and reliability across different tasks.

B. Hardware and Software Specifications

The experiments were conducted on the following hardware and software platforms:

- **Hardware Specifications:**
  - GPU 1: NVIDIA Tesla T4 (16 GB-VRAM)
  - GPU 2: NVIDIA RTX 4050 (6 GB-VRAM)

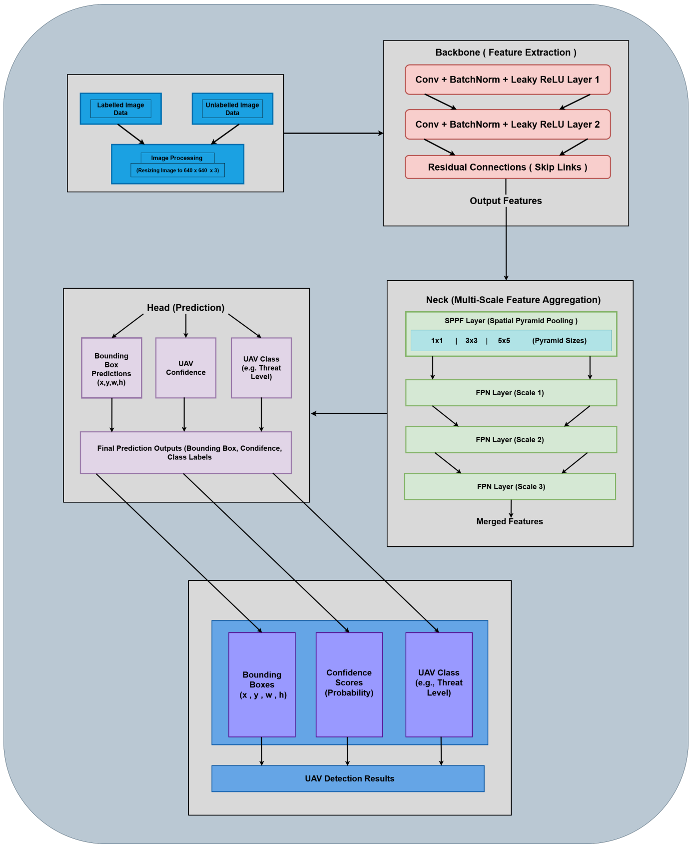

Fig. 2. YOLOv8 Architecture Diagram as given in [6].

- **Software Specifications:**
  - Programming Language: Python
  - Framework: Ultralytics, YOLOv8S, YOLOv8L

C. Confusion Matrix

The confusion matrix summarizes the performance of the YOLOv8 models in terms of true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN). It allows for a detailed assessment of how well each model is able to differentiate between drones and other similar objects.

- **Using GPU 1 :**

Drone Model				
Class	**TP**	**FP**	**FN**	**TN**
YOLOv8-S	188	24	22	0
YOLOv8-L	142	27	68	0

Combined Bird-Drone Model				
Class	**TP**	**FP**	**FN**	**TN**
YOLOv8-S	0.84	0.08	0.135	0.43
YOLOv8-L	0.79	0.10	0.08	0.46

TABLE I
CONFUSION MATRIX TABLE FOR DRONE MODEL AND NORMALIZED CONFUSION MATRIX TABLE FOR COMBINED MODEL USING GPU 1

- **Using GPU 2:**

Drone Model				
Class	**TP**	**FP**	**FN**	**TN**
YOLOv8-S	188	13	22	0
YOLOv8-L	185	31	25	0

Combined Bird-Drone Model				
Class	**TP**	**FP**	**FN**	**TN**
YOLOv8-S	0.96	0.02	0.34	0.33
YOLOv8-L	0.93	0.03	0.33	0.33

TABLE II
CONFUSION MATRIX TABLE FOR DRONE MODEL AND NORMALIZED CONFUSION MATRIX TABLE FOR COMBINED MODEL USING GPU 2

D. Evaluation Metrics

We will be comparing the performance of the YOLOv8-S and YOLOv8-L models using the following evaluation metrics: precision, recall, mean Per Class Accuracy, mean Average Precision (mAP) at IoU thresholds (0.5), and accuracy. Below are the definitions and formulas for each one.

- **Precision**: Precision is the number of correctly classified positive observations divided by the total number of predicted positive observations. It measures how many of the predicted positive examples are truly positive.

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

  where TP is the number of true positives and FP is the number of false positives.

- **Recall**: Recall is the ratio of correctly predicted positive observations to the total actual positives. It reflects how well the model captures all the positive examples.

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

- **Mean Per Class Accuracy**: The accuracy of all classes is averaged by this metric. It also helps to understand how far it deviates less well than in some other tasks as expected for the multi-class classification problem. It is defined as:

$$\text{Mean Per Class Accuracy} = \frac{1}{C}\sum_{i=1}^{C}\frac{TP_i}{TP_i + FP_i + FN_i}$$

  where C is the number of classes, and TP_i, FP_i, and FN_i are the true positives, false positives, and false negatives for class i.

- **Mean Average Precision (mAP) at IoU 0.5**: mAP is the mean of the Average Precision (AP) values at various Intersection over Union (IoU) thresholds. For this evaluation, mAP is calculated at an IoU threshold of 0.5. The formula for AP is:

$$\text{AP} = \frac{1}{N}\sum_{k=1}^{N} P(k)\Delta R(k)$$

where $P(k)$ is the precision at each threshold, and $\Delta R(k)$ is the change in recall. mAP is the average of AP across all classes.

- **Accuracy**: Accuracy is the ratio of correctly predicted instances (both positive and negative) to the total instances.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

The following tables summarize the performance metrics for drone detection and the combined detection of drones and birds.

1) Performance: The tables below presents the evaluation metrics for the YOLOv8-S and YOLOv8-L models in detecting drones and combined drone-bird using GPU 1 and 2 respectively.

- **Using GPU 1:**

Model	Precision	Recall	MPA	mAP 0.5 (%)	Accuracy
v8-S Drone	0.887	0.895	0.803	0.886	0.804
v8-L Drone	0.840	0.676	0.599	0.842	0.599
v8-S Combined	0.913	0.861	0.856	0.930	0.857
v8-L Combined	0.886	0.908	0.874	0.880	0.874

TABLE III
PERFORMANCE METRICS FOR DRONE DETECTION USING GPU 1 (YOLOV8-S VS YOLOV8-L)

- **Using GPU 2:**

Model	Precision	Recall	MPA	mAP 0.5	Accuracy
v8-S Drone	0.935	0.895	0.447	0.442	0.843
v8-L Drone	0.856	0.880	0.440	0.383	0.767
v8-S Combined	0.979	0.738	0.840	0.602	0.780
v8-L Combined	0.969	0.738	0.827	0.811	0.778

TABLE IV
PERFORMANCE METRICS FOR DRONE DETECTION USING GPU 2 (YOLOV8-S VS YOLOV8-L)

E. Discussions

The results of our experiment show notable differences between the YOLOv8-S and YOLOv8-L models across various metrics when deployed for UAV detection, highlighting the unique strengths of each model variant in terms of precision, recall, mean pixel accuracy (MPA), mean Average Precision (mAP), and mAP 0.5.

F. Analysis on GPU 1

On GPU 1, YOLOv8-S consistently outperformed YOLOv8-L in both the drone-only and combined datasets. Specifically, for the drone-only dataset, YOLOv8-S achieved a precision of 88.7% and recall of 89.5%, while YOLOv8-L showed lower values at 84.0% and 67.6%, respectively. This suggests that YOLOv8-S is more structurally optimized for smaller, faster objects, allowing it to maintain high precision and recall in controlled environments.

When evaluated on the combined dataset of drone and bird images, YOLOv8-S again demonstrated superior performance, achieving a precision of 91.3%, recall of 86.1%, and mAP 0.5 of 93.0%. These results indicate that YOLOv8-S performs better than YOLOv8-L in complex scenarios with similar-looking objects, likely due to its improved anchor-free detection mechanism that enhances its ability to differentiate between drones and visually similar objects, such as birds.

In contrast, YOLOv8-L performed better over the entire combined dataset, with a recall rate of 90.8% and MPA of 87.4%. This may be attributed to the deeper backbone of YOLOv8-L, which is better suited for handling noisy scenarios where multiple object classes are present. However, this improvement comes at the cost of slower inference, indicating that YOLOv8-L may be more suitable for applications where accuracy is prioritized over speed.

G. Analysis on GPU 2

On GPU 2, the models exhibited varying patterns, particularly in precision and mAP 0.5. YOLOv8-S achieved the highest precision of 93.5% and a recall of 89.5% on the drone-only dataset, making it suitable for applications where high precision is a priority. However, MPA and mAP 0.5 were lower compared to GPU 1, suggesting that hardware differences may impact YOLOv8-S's ability to maintain consistent performance across different computational environments.

For the combined dataset on GPU 2, YOLOv8-S again surpassed YOLOv8-L in precision, achieving an impressive 97.9%, though with a lower recall of 73.8%. This suggests that while YOLOv8-S is highly precise in detecting drones, it may miss some objects when processing a more diverse dataset on GPU 2, potentially due to hardware limitations affecting model efficiency.

YOLOv8-L on GPU 2 showed balanced results, achieving a high precision of 96.9% and a stable mAP 0.5 of 81.1% on the combined dataset. This indicates that YOLOv8-L may benefit from high-end hardware, maintaining consistent accuracy when different hardware configurations are used. Its performance is further emphasized by higher MPA and accuracy on the combined dataset, although it detects drones slightly slower in complex scenarios.

H. General Observations

In summary, the YOLOv8 variants exhibit a clear trade-off between speed and accuracy. YOLOv8-S demonstrates higher precision and better efficiency, making it ideal for real-time applications, particularly on lower-powered GPUs. On the other hand, YOLOv8-L, with its deeper architecture, is better suited for applications that prioritize accuracy, especially in challenging environments with mixed object types.

These findings suggest practical applications for each model variant:

- **YOLOv8-S** is well-suited for real-time drone detection on resource-constrained devices, where quick and precise identification is essential.
- **YOLOv8-L** is ideal for more complex, accuracy-critical applications, such as security monitoring in environments with high object diversity, where precise differentiation between similar objects is critical.

This discussion underscores the potential of each YOLOv8 variant to address specific needs in anti-UAV systems, depending on the operational requirements of speed and accuracy.

Section 5. Conclusion

This study demonstrates the effectiveness of YOLOv8 variants, particularly YOLOv8-S and YOLOv8-L, for UAV detection in real-time applications. Our findings indicate that YOLOv8-S outperforms YOLOv8-L in terms of precision and recall, especially on resource-constrained hardware, making it well-suited for real-time detection tasks that prioritize speed. In contrast, YOLOv8-L, with its deeper architecture, shows stronger performance in more complex detection environments, making it preferable for applications where accuracy is prioritized over inference speed.

The performance differences observed align with prior studies on small object detection in UAV applications. For example, Wang et al. (2023) [10] supports the findings of YOLOv8-S having advantages on accuracy for the detection of small objects as it evokes that YOLOv8-S excelled in localization ability with the smaller object, which may account for why it performed better in some cases of detecting drones. Similarly, Xu et al. (2024) [11] emphasize the balance between speed and accuracy in UAV detection, reflecting our observations that YOLOv8-S is better suited for real-time applications, while YOLOv8-L provides accuracy advantages in more computationally intensive settings [11].

In conclusion, YOLOv8 offers a versatile approach for UAV detection, with each variant presenting unique strengths adaptable to different operational needs. Future work may involve further optimization of these models for specific UAV detection tasks, potentially incorporating enhancements from recent research to improve the balance between precision, speed, and resource efficiency.

A. Future Work

Future attempts to enhance UAV detection capabilities will consider the ensembling of various YOLOv8 model variants. In particular, this means integrating the models YOLOv8-S, YOLOv8-M, and YOLOv8-L, combining robust ensemble learning, thus obtaining the power from all the models. We expect better performance in detection by training these models on different subsets of our datasets, that is focusing just on UAVs and the other UAVs paired with birds. Because the resolution of images can be very crucial and distinguishing between airborne objects of similar sizes requires attention.

Following the ensemble model developed by Sohaib et al. [8] for crack detection in concrete structures that contains YOLOv8s, YOLOv8m, and YOLOv8x, our approach will be drawn. We train each of the models on different subsets of the images; this will allow us to expose our ensemble to different UAV scenarios, minimize overfitting, and thus enhance the generalization capabilities of our ensemble. Furthermore, we will adapt the quantization techniques applied in that study toward optimizing the inference speed and memory usage of the UAV detection system we are implementing.

We will also apply sophisticated post-inference techniques following the ensemble framework, as suggested by Özbılge et al.(2024). [9], for the detection of malaria parasites. Our proposed ensemble will employ multiple object detection models; we will combine the detection boxes and confidence scores using different methods such as Non-Maximum Suppression and Weighted Boxes Fusion to further enhance the reliability of our UAV detection system. This will further help us to reduce false positives and improve the detection correctness, especially for complex environments. These tests will allow us to fine-tune the ensemble method and consolidate its efficacy within real-world UAV detection scenarios.

Fig. 3. Drone Detection using YOLOV8-L Combined Model (Drone tag - 0)

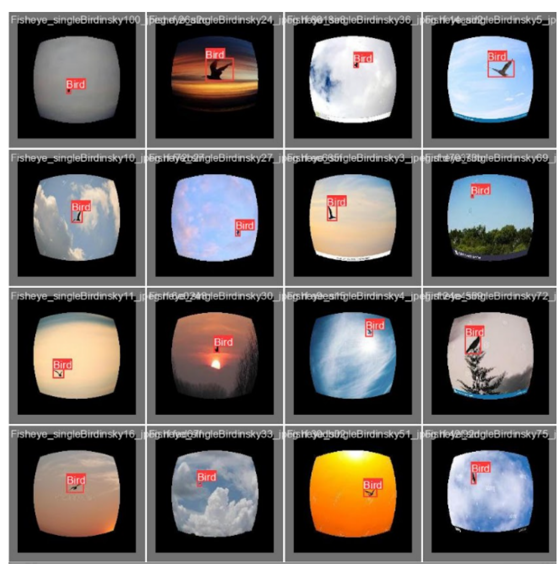

Fig. 4. Bird Detection using YOLOV8-L Combined Model

References

[1] Hussain, Muhammad. "Yolov1 to v8: Unveiling each variant–a comprehensive review of yolo." IEEE Access 12 (2024): 42816-42833.

[2] Gope, Hira Lal, Hidekazu Fukai, Fahim Mahafuz Ruhad, and Shohag Barman. ”Comparative analysis of YOLO models for green coffee bean detection and defect classification.” Scientific Reports 14, no. 1 (2024): 28946.

[3] Liu, Kaiyue, Haitong Tang, Shuang He, Qin Yu, Yulong Xiong, and Nizhuan Wang. ”Performance validation of YOLO variants for object detection.” In Proceedings of the 2021 International Conference on bioinformatics and intelligent computing, pp. 239-243. 2021.

[4] Casas, Edmundo, Leo Ramos, Cristian Romero, and Francklin Rivas-Echeverría. ”A comparative study of YOLOv5 and YOLOv8 for corrosion segmentation tasks in metal surfaces.” Array (2024): 100351.

[5] Suewongsuwan, Kamphon, Natchanun Angsuseranee, Prasatporn Wongkamchang, and Khongdet Phasinam. ”Comparative analysis of UAV detection and tracking performance: Evaluating YOLOv5, YOLOv8, and YOLOv8 DeepSORT for enhancing anti-UAV systems.” Edelweiss Applied Science and Technology 8, no. 5 (2024): 708-726.

[6] Terven, Juan, Diana-Margarita Córdova-Esparza, and Julio-Alejandro Romero-González. ”A comprehensive review of yolo architectures in computer vision: From yolov1 to yolov8 and yolo-nas.” Machine Learning and Knowledge Extraction 5, no. 4 (2023): 1680-1716.

[7] Lagari Team, *Drone Detection Dataset*, 2024. Available at: https://universe.roboflow.com/lagari-team/drone-tespit/dataset/3. Accessed: 2024-11-24.

[8] Sohaib, Muhammad, Saima Jamil, and Jong-Myon Kim. ”An ensemble approach for robust automated crack detection and segmentation in concrete structures.” *Sensors* 24.1 (2024): 257.

[9] Özbılge, Emre, Emrah Gìler, and Ebru Ozbilge. ”Ensembling object detection models for robust and reliable malaria parasite detection in thin blood smear microscopic images.” *IEEE Access* (2024).

[10] Wang, G., Chen, Y., An, P., Hong, H., Hu, J., & Huang, T. (2023). *UAV-YOLOv8: A Small-Object-Detection Model Based on Improved YOLOv8 for UAV Aerial Photography Scenarios*. Sensors, 23(19), 7190.

[11] Xu, W., Cui, C., Ji, Y., Li, X., & Li, S. (2024). *YOLOv8-MPEB Small Target Detection Algorithm Based on UAV Images*. Heliyon, 10, e29501.