Object Detection from UAV Perspective using YOLOv8 and VisDrone Dataset

M. B. Shamrai[1]

[1]Institute of Mathematics of NAS of Ukraine, Kyiv, Ukraine

m.shamrai@imath.kiev.ua

Unmanned Aerial Vehicles (UAVs) have gained increasing popularity in various applications, including surveillance, monitoring, and mapping. One key task in UAV-based applications is object detection, which involves identifying and localizing objects of interest in images captured by UAVs. In this work, we fine-tune and evaluate a new version of the popular YOLOv8 (You Only Look Once, version 8) framework for UAV-perspective object detection on the VisDrone 2019 dataset [1].

YOLOv8 is a state-of-the-art real-time object detection algorithm that has shown outstanding performance in various computer vision tasks. It uses a single neural network to predict object classes and bounding box coordinates directly from the input image, making it efficient and suitable for real-time applications. The YOLOv8 framework was released in early 2023, but no official paper, to the author's best knowledge, has been published yet. The VisDrone dataset is a large-scale benchmark for object detection, containing thousands of images with diverse scenes and object categories, making it ideal for training and evaluating the object detection algorithms for UAV applications.

(a) Source image. (b) Image with predicted annotations.

Figure 1: YOLOv8 prediction example on the VisDrone 2019 dataset.

The proposed approach leverages the power of YOLOv8 and the richness of the VisDrone dataset to achieve accurate and efficient object detection from the UAV perspective. We train the model on the VisDrone dataset, fine-tuning the YOLOv8 architecture to adapt to unique challenges of the UAV imagery, such as small object sizes, varying lighting conditions, and motion blur. The data augmentation techniques are adopted to enhance the model's robustness to real-world variations. For the training, the Google Colab Pro with Nvidia A100-SXM4-40Gb GPU and 12-core CPU is used. Fine-tuning is performed over 100 epochs.

$$\mathrm{IoU}(A,B) = \frac{A \cap B}{A \cup B}, \qquad \mathrm{mAP} = \frac{1}{|\text{ classes }|} \sum_{c \in \text{classes}} \frac{\#TP(c)}{\#TP(c) + \#FP(c)}.$$

The models are evaluated using the mAP^{50-95} metric, which is commonly used in Computer Vision. It indicates the mean average precision over a range of intersections over union (IoU)

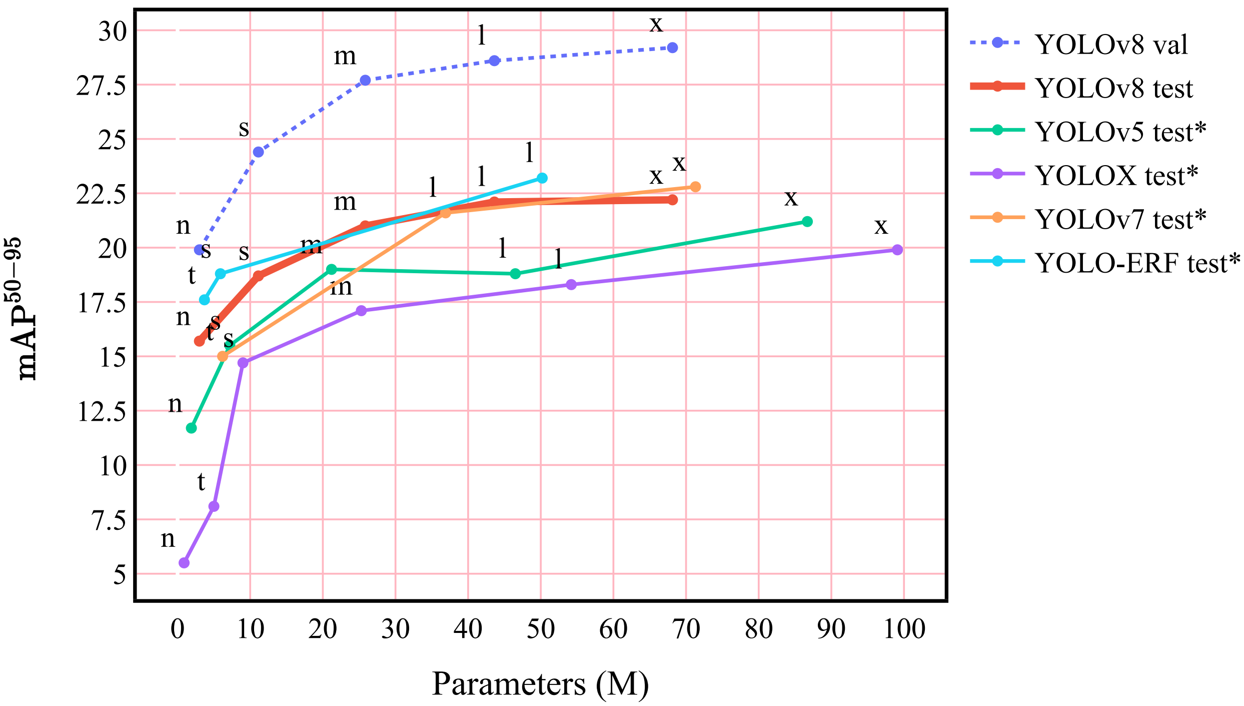

Figure 2: A number of the trainable parameters of different YOLO models versus the evaluation score on the VisDrone 2019 dataset using the mAP^{50-95} metric. "*" represents the evaluation scores taken from [2].

thresholds from 0.5 to 0.95 with a 0.05 step. The threshold value in IoU refers to the minimum required overlap between the predicted bounding box and the ground truth bounding box to consider the detection as a true positive. For example, if the threshold value is 0.5, the predicted bounding box must have an overlap of at least 50% with the ground truth bounding box to be considered a true positive.

The fig. 2 shows experimental results on the VisDrone dataset and demonstrates the effectiveness of our approach in comparison to the YOLO's older versions in the object detection from the UAV perspective, but still loses to YOLO-ERF-L model according to the given scores in [2]. Nevertheless, the YOLOv8-M fine-tuned model achieves high precision rates, outperforming the existing methods and proposing a trade-off between size and performance.

In conclusion, an evaluation of the YOLOv8 models for object detection from the UAV perspective, leveraging the VisDrone dataset for training is reported. We show that the YOLOv8 gives better results than its predecessors in object detection from the UAV images but in comparison with the UAV-oriented modified versions it shows a slightly weaker result. Our findings contribute to the field of computer vision and applied mathematics, providing valuable insight for practitioners working on UAV-based object detection.

1. Du, Dawei, et al. VisDrone-DET2019: The vision meets drone object detection in image challenge results. Proceedings of the IEEE/CVF international conference on computer vision workshops, 2019.
2. Wang, Xin, et al. YOLO-ERF: Lightweight Object Detector for UAV Aerial Images. 2023.