2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS)

YOLOv8: A Novel Object Detection Algorithm with Enhanced Performance and Robustness

1st Rejin Varghese
Department of Computer Applications
Hindustan Institute of Technology and Science
Chennai, India
rp.21703018@student.hindustanuniv.ac.in

2nd Sambath M.
School of Computing Science
Hindustan Institute of Technology and Science
Chennai, India
msambath@hindustanuniv.ac.in

2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS) | 979-8-3503-6482-8/24/$31.00 ©2024 IEEE | DOI: 10.1109/ADICS58448.2024.10533619

***Abstract*— In recent years, the You Only Look Once (YOLO) series of object detection algorithms have garnered significant attention for their speed and accuracy in real-time applications. This paper presents YOLOv8, a novel object detection algorithm that builds upon the advancements of previous iterations, aiming to further enhance performance and robustness. Inspired by the evolution of YOLO architectures from YOLOv1 to YOLOv7, as well as insights from comparative analyses of models like YOLOv5 and YOLOv6, YOLOv8 incorporates key innovations to achieve optimal speed and accuracy. Leveraging attention mechanisms and dynamic convolution, YOLOv8 introduces improvements specifically tailored for small object detection, addressing challenges highlighted in YOLOv7. Additionally, the integration of voice recognition techniques enhances the algorithm's capabilities for video-based object detection, as demonstrated in YOLOv7. The proposed algorithm undergoes rigorous evaluation against state-of-the-art benchmarks, showcasing superior performance in terms of both detection accuracy and computational efficiency. Experimental results on various datasets confirm the effectiveness of YOLOv8 across diverse scenarios, further validating its suitability for real-world applications. This paper contributes to the ongoing advancements in object detection research by presenting YOLOv8 as a versatile and high-performing algorithm, poised to address the evolving needs of computer vision systems.**

Keywords—YOLOv8, Object Detection, Performance Enhancement, Robustness, Computational Efficiency, Computer Vision Systems

I. Introduction

Recognizing objects could be a vital and complex errand within the field of computer vision, with applications traversing security, observation, self-driving vehicles, robotics, and medical imaging. The objective of object location is to find and classify objects in pictures or recordings, giving their bounding boxes and names. There are two primary sorts of protest location methods: two-stage methods and one-stage methods. Two-stage methods, such as R-CNN, Fast R-CNN, and Faster R-CNN, at first generate a set of region recommendations and after that refine them employing a classifier and a regressor. On the other hand, one-stage strategies like SSD, RetinaNet, and YOLO specifically foresee the bounding boxes and names from the input picture, dispensing with the require for locale recommendations. In spite of the fact that one-stage methods are ordinarily speedier and less complex than two-stage strategies, they regularly compromise on exactness and soundness.

YOLO (You only Look Once), a critical one-stage protest discovery calculation, was to begin with presented by Redmon and Farhadi in 2017 [1]. YOLO segments the input picture into a network of cells and predicts a settled number of bounding boxes and certainty scores for each cell. It too predicts the lesson probabilities for each bounding box and combines them with the certainty scores to deliver the final detection comes about. YOLO is eminent for its speed and compelling execution on huge and medium-sized objects, but it has certain impediments, such as moo review, rough localization, and subpar execution on little objects.

Since YOLO's beginning, various variations and improvements have been proposed to address its impediments and boost its execution. Vital cases incorporate YOLOv2, YOLOv3, YOLOv4, YOLOv5, YOLOv6, and YOLOv7 [11]. These forms have consolidated different procedures and developments, such as stay boxes, multi-scale predictions, feature pyramid systems, residual connections, consideration instruments, energetic convolutions, and voice acknowledgment. These strategies have upgraded YOLO's precision, strength, and effectiveness, making it more versatile to different circumstances and applications. All things considered, there are still openings for encourage advancement and optimization, especially for challenging scenarios including little objects, blocked objects, and complex foundations.

In this paper, we introduce YOLOv8, a modern object detection algorithm that builds upon the past YOLO forms and consolidates modern highlights and improvements. YOLOv8 endeavours to realize the most elevated speed and precision in object location, whereas guaranteeing strength and soundness.

II. Literature Review

YOLO (You Only Look Once), a single-stage object detection algorithm, was initially introduced by Redmon and Farhadi in 2017 [1]. The YOLO algorithm partitions the input image into a grid of cells and forecasts a predetermined number of bounding boxes and confidence scores for each cell. Additionally, YOLO predicts the class probabilities for each bounding box and merges them with the confidence scores to produce the final detection outcomes. YOLO is recognized for its impressive speed and effective performance on large and medium-sized objects. However, it does have certain drawbacks, such as low recall, imprecise localization, and suboptimal performance on small objects.

The YOLOv2 [1] is an progressed adaptation of YOLO, which presents a few procedures to upgrade the accuracy and efficiency of the algorithm. A few of the most methods are:

- *Anchor boxes* : YOLOv2 employments predefined bounding box shapes, called anchor boxes, to superior fit the objects of distinctive sizes and aspect ratios. YOLOv2 predicts the offsets and scales of the anchor boxes, rather than the supreme