Title: YOLOv8: A State-of-the-Art Object Detector
Author: Nahid Hasan
Date: 2/5/2025

1. Introduction

Object detection is a critical task in computer vision, with applications in autonomous vehicles, security, and medical imaging. YOLO (You Only Look Once) models have been widely used due to their balance between speed and accuracy. The latest version, YOLOv8, introduces several improvements in architecture and efficiency, making it one of the most advanced object detection frameworks.

2. Key Contributions

The paper introduces the following improvements over previous YOLO models:

- Anchor-Free Detection: Removes the need for predefined anchor boxes, reducing computational cost and improving flexibility in detecting objects of varying sizes.
- Improved Backbone Network: Uses a modified CSPDarkNet to enhance feature extraction and improve performance on smaller objects.
- Dynamic Head Architecture: Adapts based on image complexity, improving localization accuracy and reducing false positives.
- Enhanced Data Augmentation: Implements Mosaic and MixUp augmentation for better generalization, increasing model robustness in detecting objects under varying conditions.
- Improved Post-Processing: Introduces soft non-maximum suppression (Soft-NMS) to better handle overlapping detections, leading to improved accuracy.

3. Methodology

- The model was trained on the COCO dataset (Common Objects in Context) with 640×640 image resolution, ensuring high-quality detections across diverse categories.
- It optimizes with the SGD optimizer and a learning rate scheduler to achieve better convergence and reduce overfitting.
- Uses IoU-aware object assignment to improve bounding box accuracy and better distinguish overlapping objects.
- The architecture follows a convolutional-based detection pipeline, reducing computational cost while maintaining state-of-the-art accuracy.

4. Results and Performance

The table below shows YOLOv8's performance compared to YOLOv5 and Faster R-CNN:

Model	FPS (Speed)	mAP (Accuracy)	Parameters
YOLOv5	55 FPS	50.2%	7.5M
YOLOv8	75 FPS	52.8%	6.2M
Faster R-CNN	12 FPS	48.5%	42M

YOLOv8 achieves a higher mean Average Precision (mAP) while maintaining a smaller model size, making it efficient for real-time deployment.

5. Applications

- Autonomous Vehicles: Faster object detection for real-time driving decisions, helping self-driving cars navigate safely.
- Security Surveillance: Detecting people and suspicious activities in video feeds to enhance public safety.
- Medical Imaging: Identifying tumors and abnormalities in X-ray and MRI images, assisting radiologists in faster diagnosis.
- Retail & Inventory Management: Automatically identifying and tracking products in stores to optimize stock management.
- Robotics & Industrial Automation: Assisting robots in identifying objects in manufacturing environments, reducing human intervention.

6. Conclusion

YOLOv8 provides an efficient and accurate object detection model, making it suitable for real-time applications across various industries. Its anchor-free approach, dynamic head architecture, and advanced post-processing techniques offer significant improvements over previous models. The combination of speed, accuracy, and low computational cost makes it a valuable tool in modern computer vision tasks. Future improvements could focus on further reducing model size while maintaining high accuracy for edge-device deployments.