Ahmedabad University

CSE 641 Computer Vision

Weekly Report - Week 4

Project Title: Evaluate Performance of YOLO Family Models in Small Object Detection (HBB)

Section - 1
Group - 02

Name	Enrollment No.
Namra Maniar	AU2240080
Romir Bedekar	AU2240254
Hetav Shah	AU2240253

Introduction

This week was centered on sharing our mid-term achievements and comparing five YOLO family models for small object detection on two datasets: VisDrone and COCO. The goal was to determine the best models at small object detection and areas for future improvement. Comparison was made based on mAP values and other performance measures to inform future improvements.

Performance Evaluation and Key Findings

In our comparison, we used YOLO NAS, YOLO World, PP-YOLO, YOLOR, and YOLOX. The mAP scores showed that YOLO NAS and YOLO World performed better in detecting small objects compared to the others. PP-YOLO and YOLOR had lower mAP scores, which are possible areas for enhancement.

- YOLO NAS: Performed exceptionally well because of Neural Architecture Search (NAS), anchor-free detection, and dynamic heads, providing greater flexibility in small object detection.
- YOLO World: Did well because of its generalization power on different object categories and capability to detect new, unseen objects.
- PP-YOLO and YOLOR: Performed worse due to weaker feature extraction mechanisms for tiny objects. These models can improve by including methods such as multi-scale prediction and enhanced anchor assignments.

TABLE I
PERFORMANCE COMPARISON OF YOLO MODELS ON DIFFERENT DATASETS

Model	Dataset	mAP@50
YOLO-X	COCO	40.5
PP-YOLO	COCO	44.4
YOLOR	COCO	35.08
YOLO-NAS	COCO	47.03
YOLO-World	COCO	51.0
YOLO-X	VisDrone	30.6
PP-YOLO	VisDrone	27.4
YOLOR	VisDrone	20.5
YOLO-NAS	VisDrone	30.6
YOLO-World	VisDrone	33.4

Proposed Enhancements

According to the results, we intend to improve the poorer models by adopting good methods from other frameworks, including:

- Decoupled Head: Disables classification and localization tasks to achieve accuracy.
- SIMOTA: Trains label assignment to achieve improved detection of smaller objects.
- Mosaic Augmentation: Merges fragments of various images to enhance object diversity in training.
- IoU Aware Loss: Improves accuracy of localization by emphasizing overlap between the predicted and ground-truth boxes.
- Anchor-Free Detection: Streamlines detection by avoiding the use of anchor boxes, enhancing performance on smaller objects.

Conclusion

The table of mAP values presents insightful information about the strengths and weaknesses of different YOLO models for small object detection. YOLO NAS and YOLO World were the top performers, while PP-YOLO and YOLOR needed room for improvement. In the future, we plan to adopt specialized techniques to enhance small object detection performance, leveraging the best practices from multiple architectures.