Progress Report: YOLOv8 Model Training on VisDrone Dataset

Urjit Mehta, Krina Khakhariya, Brijesh Munjiyasara

Ahmedabad University

Team: Trinity (Github link)

WEEK NUMBER: 4

I. PROGRESS SUMMARY

This week, we focused on refining the YOLOv8-based object detection and tracking pipeline. Our primary objective was to address challenges observed in the previous experiments, particularly in tracking performance and detection robustness.

A. Work Completed:

- Deployed YOLOv8 for object detection on the VisDrone-MOT dataset and optimized its functionality for multi-object tracking.
- Converted the dataset annotations to YOLO format to make them compatible with the tracking framework.
- Implemented DeepSORT for tracking objects, improving Kalman filtering and IoU-based matching for more accurate object association.
- Conducted preliminary experiments to test tracking performance in complex environments, such as occlusion-rich scenes and abrupt motion changes.
- Recorded key performance metrics like precision, recall, and mean Average Precision (mAP) to evaluate detection and tracking performance.

B. Milestones achieved:

- Completed training YOLOv8 on VisDrone-MOT, hyperparameter tuning (batch size, learning rate, and augmentation methods) for better accuracy.
- Enhanced feature evolution analysis for reducing identity switches and improving object re-identification (ReID) across frames.
- Developed a Python-based evaluation framework to systematically benchmark tracking performance.
- Observed improvements in tracking robustness, particularly in dense and occlusion-heavy settings.

II. Challenges & Resolutions

A. Problems faced:

- Fixed tracking failures caused by occlusion by enhancing ReID-based feature matching for better object re-identification when there is object disappearance and reappearance.
- Corrected inconsistencies in the dataset, normalizing object annotations and repairing class mappings for consistent labeling.
- Enhanced computational efficiency by tuning YOLOv8's parameters and enabling CUDA acceleration, which improved real-time performance.
- Reduced false detections and ID switches by adjusting detection confidence thresholds and refining the feature association pipeline.

III. Upcoming Tasks

- Run additional analyses with more advanced tracking metrics such as Multiple Object Tracking Accuracy (MOTA) and ID switching analysis.
- Maximize model efficiency for real-time tracking by reducing computational burden without compromising accuracy.
- Test the system on different environmental circumstances (e.g., varying light, crowd density) to analyze its robustness.
- Develop complete documentation for the open-source release, including implementation guidelines and performance metrics.
- Perform further validation with more video sequences to establish consistency and reliability across different conditions.

IV. Conclusion

This week's progress has been significant in preparing the VisDrone dataset and building a YOLOv8 model for object detection. The implemented scripts have automated the preprocessing and annotation conversion, making the dataset ready for further model building. Merged and optimized model performance with DeepSORT algorithm and inference efficiency will be the primary focus.