Progress Report: YOLOv8 Model Training on VisDrone Dataset

Urjit Mehta, Krina Khakhariya, Brijesh Munjiyasara

Ahmedabad University

Team: Trinity (Github link)

WEEK NUMBER: 3

I. OVERVIEW

This week, we worked extensively on preparing and processing the VisDrone2019-MOT dataset for object detection using YOLOv8. The core focus was ensuring the dataset was properly formatted, annotations were correctly converted, and the model was trained effectively.

II. TASKS COMPLETED

A. Image Processing

- Implemented a resizing function to adjust all images to 640x640 pixels while preserving the aspect ratio.
- Applied padding where necessary to maintain consistency.
- Processed the dataset and replaced original images with the updated versions.

B. Dataset Preparation

- Created a configuration file ('dataset.yaml') for YOLOv8.
- Defined training and validation directories.
- Assigned a single object detection class for simplified annotation.

C. Annotation Conversion

- Transformed VisDrone's original annotation format into YOLO-compatible format.
- Normalized bounding box coordinates as per YOLO requirements.
- Ensured all images had corresponding annotation files.

D. Automation of Dataset Processing

- Developed a script to automate dataset conversion and organization.
- Identified and resolved missing or corrupted data.

E. Training YOLOv8 Model

- Used 'yolov8n.pt' as a base model for training.
- Configured model parameters:
  - Epochs: 5
  - Image size: 640x640
  - Batch size: 16
  - Workers: 8
- Evaluated model performance on validation data.

III. CHALLENGES FACED

- Encountered missing images in the dataset.
- Optimization of annotation conversion for efficiency.
- Ensuring bounding box normalization was accurate.

IV. NEXT STEPS

- Focusing on merging YOLO and SORT algorithm for tracking and object detection together
- Increasing training epochs to enhance model accuracy.
- Performing hyperparameter tuning for improved generalization.
- Testing the trained model on real-world drone footage.
- Developing an automated pipeline for video-based object detection.

V. 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. Moving forward, merging and optimizing model performance with SORT algorithm and inference efficiency will be the primary focus.