Vehicle Speed Estimation from Video Using YOLOv8 and Speed Formula

By Nihal Gireesh
nihal.gireesh@btech.christuniversity.in

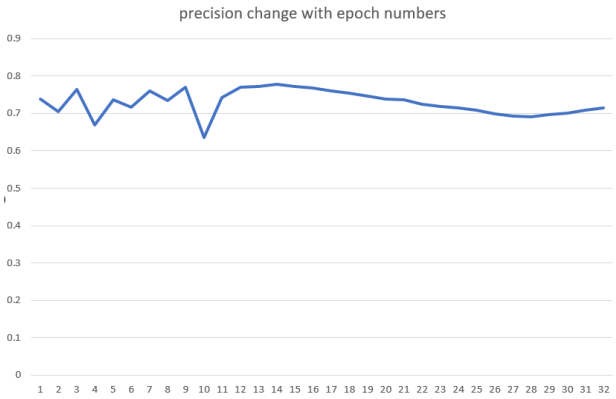
I. Introduction

Vehicle speed estimation has its significance in areas like traffic management, road safety and autonomous driving. However, speed detection is a challenging problem as it requires accuracy and precision. This program demonstrates a working model of speed detection from videos. Our approach consists of two main steps: first detecting vehicles that are present in a video and then extracting information such as vehicle type, bounding box coordinates and confidence score. In the second step, we use a speed detection algorithm that works on the principle of speed formula; we also use the frame rate to estimate the time taken. With all these inputs we estimate the speed that the detected vehicle is travelling in and annotate it in the video.

We used yolov8 to train the model. yolov8 is a state-of-the-art deep learning model for object detection with the capability of detecting vehicles of different sizes and categories in various scenarios and situations.

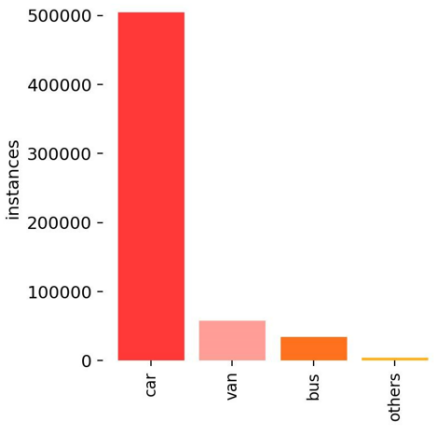
II. Developing an AI Model

The AI model has been trained on a server that consisted of **4 T4 GPU**. We have used a **multi-GPU train** approach. The batch size was set to 64, workers to 16 and cache set to ram mode. We have used X model of yolov8 as the base model, as the previous sizes were giving lower accuracy values that failed to identify many vehicles. With the new model, we are able to accurately identify every vehicle within the scope of the frame.



III. Training Dataset

The dataset used for training was the DETRAC dataset (https://detrac-db.rit.albany.edu/). The dataset includes ten hours of footage shots from 24 different places in Beijing and Tianjin in China by using a Cannon EOS 550D camera. The videos have a resolution of 960x540 pixels and are captured at 25 frames per second (fps). The UA-DETRAC dataset has around 140 thousand frames and 8250 manually annotated cars, resulting in a total of 1.21 million labeled bounding boxes of objects.



It consists of 500,000 images of cars, around 80,000 images of vans, 60,000 buses and few thousand other categories.

IV. Conclusion

This working model is capable of detecting the speed of different types of vehicles in various scenarios from their live video footages. Hence, this program is going to be of immense value for different stakeholders such as traffic-law enforcement agencies, road safety authorities and for the applications in autonomous driving vehicles.