Vision System for Traffic Violation Detection

Huandong Chang, Raphael Krief

Motivation & Project Objectives

Motivation

- In US, over 40,000 people died in motor vehicle traffic crashes in 2023[1]
- A lot of people are unaware of their bad driving behaviors
- Current literature on traffic violation detection rely on CCTVs

Project Objectives

- Given a dashcam setup and a telematic sensor, we detect traffic rule violations
- Enhance driver safety by our vision-based system
  - Help new drivers form good driving habits
  - Collaborate with insurance company to offer discounts to safe drivers

1 https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/813561

Image & Video Dataset

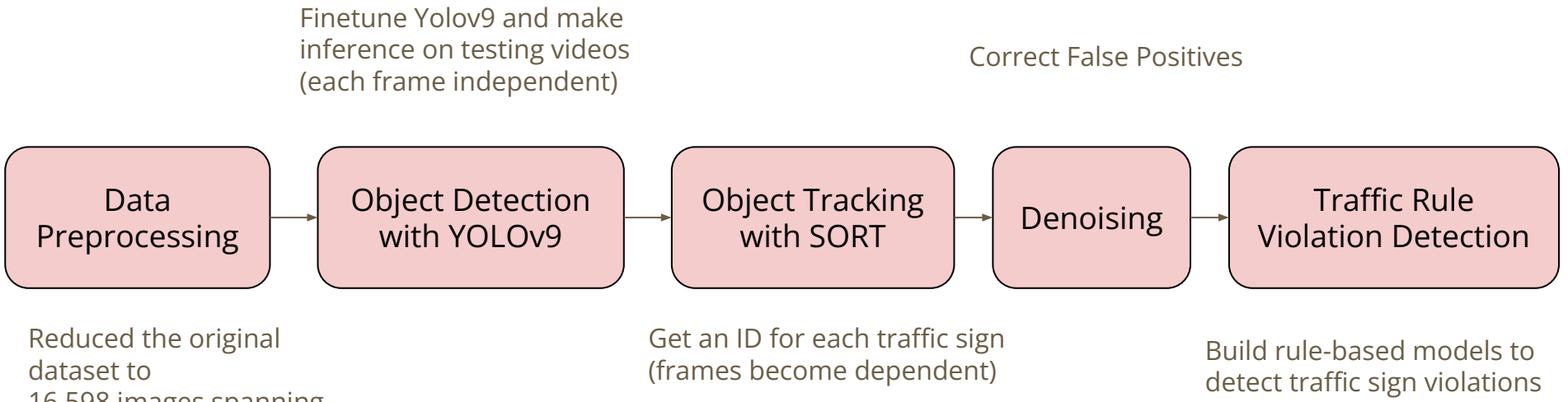
Training set : Mapillary Traffic Sign Dataset (MTSD)

- Comprehensive Coverage: Over 250,000 annotated traffic signs from diverse geographic locations.
- Detail Rich: Features 313 unique traffic sign classes across 52,453 high-quality annotated images.
- Research Utility: Designed to reflect real-world variability, enhancing the robustness of detection algorithms.

Test set : Different driving videos

- Video Excerpts: Contains traffic signs from our dataset, ensuring relevant test scenarios.
- Diverse Locations: Includes one video captured in Cambridge and another in the countryside, providing varied environmental contexts.

Algorithm Pipeline



Object Detection with YOLOv9

Why Yolo (You Only Look Once) Architecture?

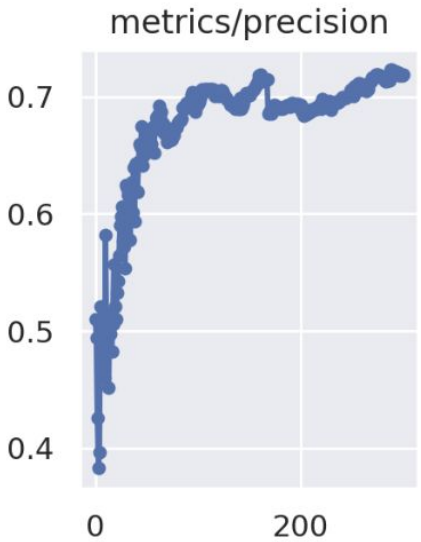
Fast: Only process the image once to predict multiple bounding boxes and class probabilities

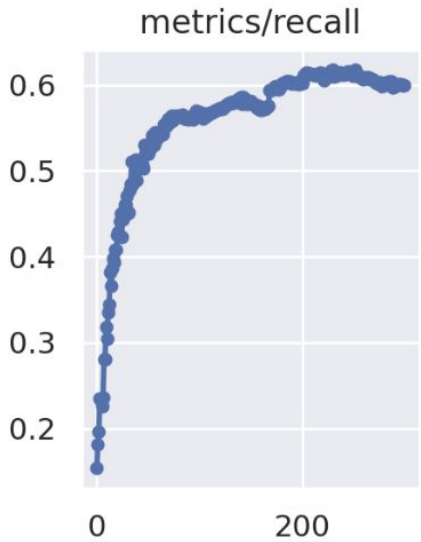
Finetune Yolo

- Batch Size: 16
- Epoch: 300
- Image Size: 640*640
- Validation Precision: 0.719
- Validation Recall: 0.599

Inference on video (Demo)

- Confidence Threshold: 0.5





Object Tracking with SORT

SORT (Simple Online and Realtime Tracking)

- The Kalman filter predicts the future state of each object based on its current state
- Then use Hungarian Algorithm (a combinatorial optimization algorithm) to assign detections to tracks (assign detections to object IDs)

Denoising

- Only keep traffic signs that appear more than 20 frames (60 frames/second, so 0.33 seconds)
- If an ID has >1 class predictions, take the mode

Video Demo

Traffic Sign Violation Detection

Stop Sign Violation Detection

- Stability-Based Detection: The algorithm identifies when the associated bounding box of a stop sign remains nearly identical (99% IoU similarity) for more than 30 consecutive frames (0.5 second)
- Stop Time Tracking
- Visual Indication

Video Demo

Limitations

- Cannot handle complex scenarios
- Real-Time Processing

Thanks!