Step 4: Project Proposal - Video Velocity Finder

Problem Statement:
Finding the velocity of moving objects from a camera video feed is a challenging problem. There are several ways to detect motion and object velocities in video imagery but require to run algorithms like optic flow to find how elements move in respect to the scenery in the image. The approach taken in this project is to leverage machine learning and train a model to be able to detect motion in video imagery and then based on the motion, predict the velocity of the moving object in the feed.
Theoretically, the trained model should be able to detect static and non-static objects and indicate the velocity of the moving features based on knowing the video frame rate. The intention is to augment the video and overlay the predicted velocity on top of the imagery.

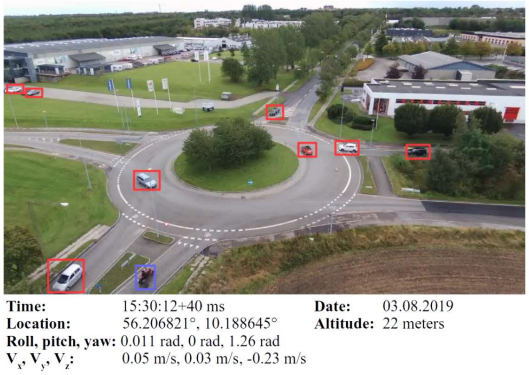
The Dataset:
The main page for the dataset used is located here: https://viratdata.org/

Videos are found here:
https://data.kitware.com/#collection/56f56db28d777f753209ba9f/folder/56f581ce8d777f753209ca43

Latest most up to date annotations are the DIVA Annotation. This is a public repository of the VIRAT video data annotations as annotated by the IARPA DIVA program. Annotations can be found here:
https://gitlab.kitware.com/viratdata/viratannotations

Annotations are in the format per:
https://gitlab.kitware.com/meva/meva-data-repo/-/blob/master/documents/KPF-specification-v4.pdf

An additional dataset that was looked at was:
AU-AIR: Multi-modal UAV Dataset for Low Altitude Traffic Surveillance found at: https://bozcani.github.io/auairdataset



The AU-AIR dataset has several features:

- Object detection in aerial images
- >2 hours raw videos
- 32,823 labelled frames
- 132,034 object instances
- 8 object categories related to traffic surveillance
- Frames are also labelled with time, GPS, IMU, altitude, linear velocities of the UAV

This dataset is maintained by:
Bozcan, Ilker, and Erdal Kayaan. "AU-AIR: A Multi-modal Unmanned Aerial Vehicle Dataset for Low Altitude Traffic Surveillance." IEEE International Conference on Robotics and Automation (ICRA), 2020

The approach:

- Leverage the datasets and use Opencv to construct videos containing annotated data.
- Format and clean dataset as needed
- Compute velocities on dataset and annotate information on video

- Train the machine learning model and select algorithm that produces best results
- Test the ML/AI algoriments and improve as needed
- Run the ML algorithms on new video streams in a “real world” environment.
- Change algorithms and enable the ML to run in a streaming configuration. Goal is to be able to connect to a live video stream and add annotated velocities to objects moving in it.