

Project Proposal for DriveVision

Naz Saral

1010184408, naz.saral@mail.utoronto.ca, <https://github.com/nazsaral5/project-aps360.git>

Introduction

Traffic sign recognition systems play a fundamental role in developing technologies such as autonomous driving and driving assistance systems used in vehicles today. It is exciting to witness that vehicles today can interpret road rules and respond actively to their environment. This semester I'll be working on the traffic sign recognition tool called DriveVision. The goal of my project is to design and prototype an image classification tool that will recognize traffic signs around the globe, but specifically evaluated on our very own signs by the Ministry of Transportation (MTO) in Ontario. I will scope my project to identify five types of signage; the red and white octagonal Stop sign, the upside down triangular Yield sign, the rectangular and circular Speed Limit sign, the circular No-entry sign, and the rectangular One-way sign. This subset of signage influences driving, considering they directly indicate when to stop, speed up, slow down which is fundamental to the activity of operating a vehicle. The project will make use of deep learning concepts, as taught in this course, particularly for the training, validation, and testing of data. The dataset I will use to train the model is the The German Traffic Sign Recognition Benchmark (GTSRB), in addition to this I will also incorporate a collection of data from Ontario, including the rectangular speed limit and one-way signs to overall train then evaluate the system, to ensure I account for the differences in signage. On the backend, the overall image classification will be done using convolutional neural networks (CNNs) to recognize and classify the five traffic signs. For the speed limit signs especially, in addition, I will classify the limit value itself, as this is critical. Model performance is evaluated using classification accuracy and confusion matrices on both the benchmark dataset and the Ontario evaluation dataset.

Illustration

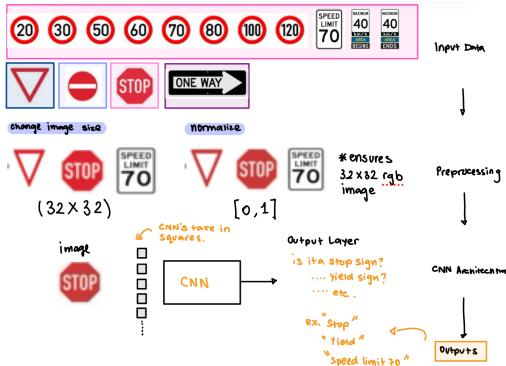


Figure 1: High-level overview of the DriveVision traffic sign recognition pipeline.

Background & Related Work

For over 40 years researchers have iteratively developed feature classification systems for road signage based on edges, colors, shapes using image processing. Current research evolves within applications in daily driving such as autonomous vehicles, driver assistance systems and commercial vision pipelines. In recent years, related work that motivates this project for me are current technological advancements. In general traffic sign recognition is performed by feature matching, a first article mentions first preprocessing by converting RGB color space into HSV color space, then using transformation

techniques to detect shapes (circles, triangles, squares), then applying feature matching methods to generate an output [1]. Another article explores the use of the German Traffic Sign Recognition Benchmark for traffic sign classification and recognition, using an improved *LeNet-5 CNN* model which is a highly complex network with seven layers, which include two convolutional layers, two pooling layers, two fully-connected layers and one output layer [2]. I will be drawing inspiration from both of these works to build out my model and establish a procedure as highlighted below for my DriveVision model. In addition to exploring the technical side of research in this field I also looked into specifics which motivate my own research such as Tesla Autopilot, where in the early 2020's they updated their autopilot equipment to read speed limit signs and detect green lights being able to stop and slow down as needed, pioneering the commercialization of vision-based driver assistance systems for cars, initially launching the software back in 2019 [3]. Another popular application is Waymo's self-driving technology, offering users in West Coast and Southern American states to enjoy a driverless rideshare. Benchmarking rideshare experiences with their own advancements in search and image recognition training, valuable in serving a public ridership [4]. Google Maps services make use of AI and imagery to identify speed limits, building on models trained on hundreds of different types of signs from all over the world [5].

Data Processing

This project is based on the GTSRB dataset but will also include a few additional photos of signage publicly available from the MTO, which lends DriveVision to be relevant for use locally. The data to be processed is labeled images of traffic signs, of varying picture quality. All images that are inputted into the model will undergo the same preprocessing steps to ensure consistency. To begin, as shown in Figure 1, all images will be turned to 32×32 pixels and maintain a consistent RGB format to match the CNN input size and reduce computational complexity and keep consistent color information. Pixel values will then be normalized to $[0,1]$ range by dividing by 255 to ensure consistent input data, faster model convergence, and better numerical stability during training and compatibility with PyTorch.

Architecture

For the image recognition, I have opted to use a Convolutional Neural Network setup, which will follow a sequential, layered, architecture that acts as an image processing pipeline. It will be capable of detecting edges, shapes as needed to classify the signs and decide which traffic sign it is. If needed, it will also detect the value of the speed limit sign on both a circular and rectangular sign.

Baseline Model

For the baseline model, I will use a support vector machine (SVM) classifier. The SVM baseline will be trained and evaluated using the same training, validation, and test sets as the CNN. Since I have yet to learn about this classifier in lectures, I will go more into its use in the following project deliverables.

Ethical Considerations

As for ethical considerations for DriveVision, it is important to note that the tool's goal is to decipher photos of signs into five possible categories. With such a task at hand some important considerations are as follows: **1. Misclassification:** This presents as a limitation to the model, considering traffic sign recognition is critical in applications of autonomous driving and driver's assistance technology. If a sign is misclassified, there may pose risks to the driver and/or vehicle in operation while in traffic. It is important that in evaluating my model, that there is reasonable error, to ensure its safety. **2. Dataset Representation and Generalization:** This is a limitation with respect to the data. Since the dataset I have chosen to use is primarily composed of European traffic signs. A model trained on one region may not generalize to another. Since I would like this model to be evaluated on Ontario traffic signs, which may differ in fonts, color schemes and layouts. To overcome this, I will add a subset of relevant MTO signs that would be found within the province to strengthen my model's use here in Ontario. **3. Data Collection:** As a general ethical consideration, as mentioned, I will be including a subset of my own personal data collection in the sets. All traffic sign images collected by me will be sourced from publicly available MTO data only, to ensure reliability and confidence.

References

- [1] @inproceedingsren2009traffic, author = Ren, Fei-Xiang and Huang, Jinsheng and Jiang, Ruyi and Klette, Reinhard, title = General Traffic Sign Recognition by Feature Matching, booktitle = Proceedings of the 24th International Conference on Image and Vision Computing New Zealand, pages = 409–414, year = 2009, doi = 10.1109/IVCNZ.2009.5378370
- [2] @articlecao2019traffic, title = Improved Traffic Sign Detection and Recognition Algorithm for Intelligent Vehicles, author = Cao, J. and Song, C. and Peng, S. and Xiao, F. and Song, S., journal = Sensors, volume = 19, number = 18, pages = 4021, year = 2019, doi = 10.3390/s19184021
- [3] @onlinebbc2020tesla, title = Tesla Updates Autopilot to Read Speed Signs, author = BBC News, organization = BBC, year = 2020, url = <https://www.bbc.com/news/technology-53973511>
- [4] @onlinewaymo2020perception, title = How Waymo Builds a Generalizable Perception System, author = Waymo, organization = Waymo, year = 2020, url = <https://waymo.com/blog/2020/02/content-search>
- [5] @onlinegooglemaps2020speedlimits, title = How AI and Imagery Keep Speed Limits on Google Maps Updated, author = Google Maps, organization = Google, year = 2020, url = <https://blog.google/products-and-platforms/products/maps/how-ai-and-imagery-keep-speed-limits-on-google-maps-updated/>