Traffic Sign Localization and Detection

Project Proposal

University of Toronto

Faculty of Applied Science and Engineering

Artificial Intelligence Fundamentals

Febrary 23rd, 2020

Team 02:

Siwei He

Xinru Li

Jiani Li

Qihan Liu

Word count: 1387

Introduction

Technology of traffic sign recognition and analysis is not only important to the development of advanced driver support systems, but also a crucial part of autonomous cars. The goal of this project is to develop a model using machine learning that is able to recognize the real-time traffic signs on the road, and further, to convert the information obtained from the sign into audio output in order to notify the drivers by audio when necessary. According to infrrd.ai, "Deep Learning models, with their multi-level structures, …, are very helpful in extracting complicated information from input images" [1]. Therefore, for this project, machine learning and possibly deep learning image processing techniques would be the best methods for us to develop a model that is able to learn and detect the patterns like colours, shapes, and letters in the sign.

Background & Related Work

The detection and recognition of traffic signs is a popular machine learning topic as it helps to improve driving safety and assist auto driving. We studied works related to this area to lay the foundation and improve our own project.

A research group from China focused on the detection and localization of traffic signs.
They developed an algorithm combined "white balance, color image enhancement, and affine transform correction" [2]. They applied this method and successfully detected traffic signs under real world conditions, including extreme weather and being obscured. We can study and develop our own method for detection and localization based on this previous work.

Another team in Slovak Republic designed different ANN models to recognize different traffic signs, and "the type of all used ANNs is a multilayer perceptron with one hidden layer" [3]. However the type of traffic signs that can be classified is limited and the accuracy is uncertain. As for our own project, we plan to focus on the classification of all types of traffic signs in Canada.

Data processing

There are several datasets of North America traffic sign annotations available online, take LISA traffic signs [4] and Mapillary [5] as examples. Both datasets can be used for localization and the classification model.

Preprocessing includes frame capture from driving recorders. We will then localize and crop the traffic signs from the preprocessed data by separating color channels, white balancing, detecting edges or shapes. To make sure our localization method works in Toronto, we will use real life data collected by ourselves as validation data.

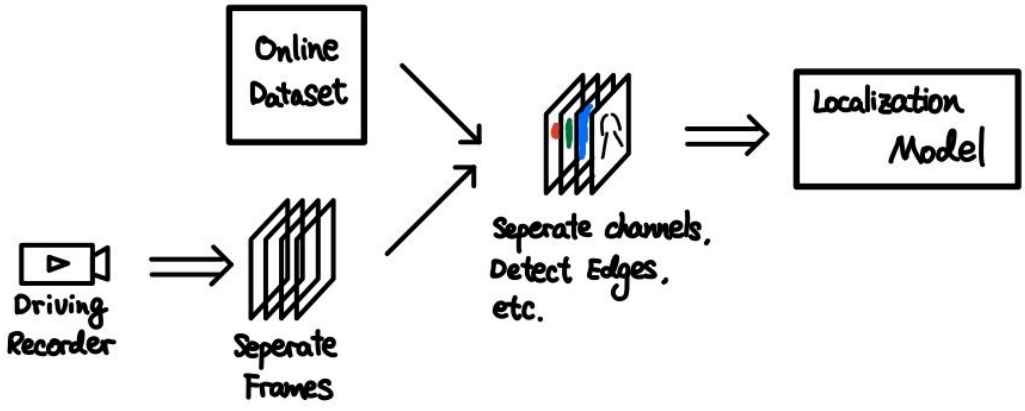

Figure 1: Preprocessing images for localization model

Then, we will use online datasets and the output from the localization model to train our classification model. To meet the dimension of the input, image preprocessing including rotation, resizing and padding blank pixels will be performed before feeding the data into the classification model.

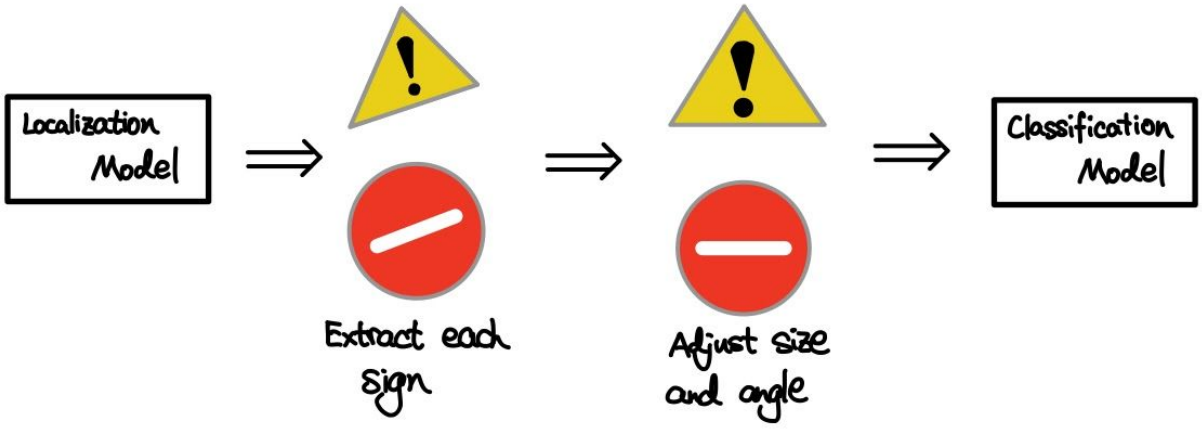

Figure 2: Preprocessing images for classification model

Project Architecture

The architecture consists of the following four stages:

1) Fetch real-time image.

2) Localize and extract traffic signs.

3) Identify each traffic sign.

4) Appropriate audio output.

Machine learning will be used in stage 3.

In stage 2, image processing will be used to extract and adjust each sign detected by analyzing the color and/or edges of the image. It is desired for the process to be fast to meet the need for real time traffic sign identification. However, we will use machine learning in this stage if it turns out that the image processing model yields low accuracy.

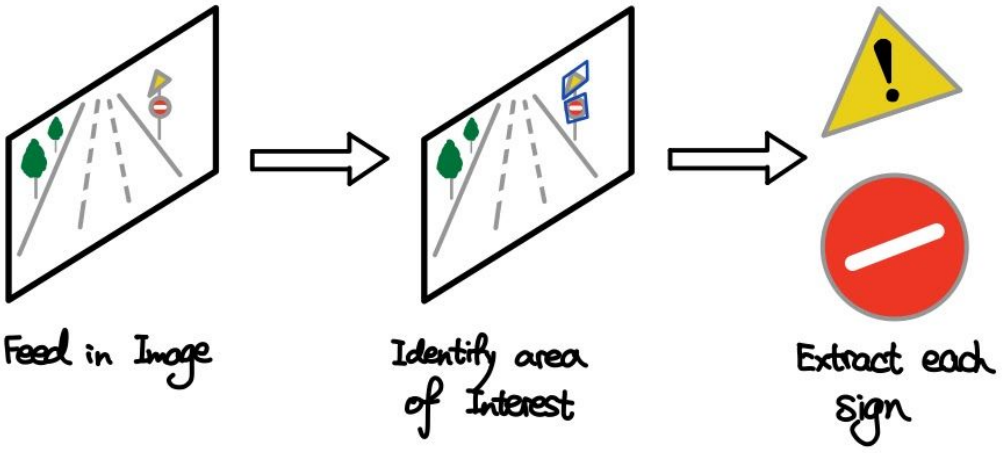

Figure 3: Stage 2 - mage processing model for localization

Our plan is to use two submodels (i.e. hierarchical architecture) to improve the accuracy of the performance. The first stage is to capture distinct features such as colors and shapes of the signs in order to categorize the type of the traffic signs among the five types: regulation signs (red background colour), regulation signs(green background colour), warning signs, information sign(white background), and others. After that, more detailed information including words in the sign will be analyzed from the sign to identify the exact sign from its type. All the models will use CNN as their main structure in order to detect spatial features in the image.

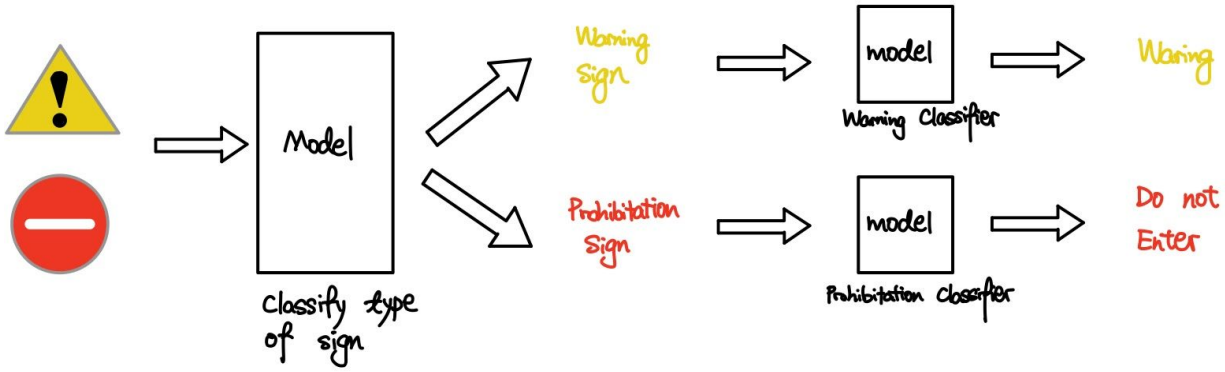

Figure 4: Stage 3 - Identify each traffic sign using hierarchical model.

Baseline model

We plan to employ a one stage Convolutional Neural Network(CNN) as our baseline model for this project. CNN is an efficient way to classify images. It works well for 2D images as it uses 2D convolutional layers. We will create a sliding window to scan the whole image and then apply our CNN model. It consists of input layers, hidden layers and output layers. The hidden layers include convolutional layers, fully-connected layers and pooling layers. In addition, CNN requires a little preprocessing and is suitable for real time analysis.

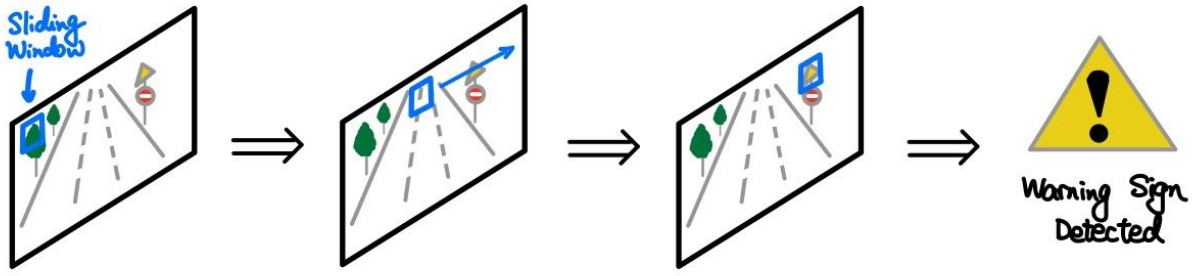

Figure 5: Baseline model using sliding window and CNN

Ethical consideration

As an essential part of developing driver assistant and autonomous cars systems, satisfying the speed of image process and the detection accuracy at the same time becomes the biggest challenge for our project. Either large delay of detection or a large detection loss would possibly mislead the driver and therefore results in receiving a fine, adding demerit points, or even worse, causing traffic incidents. Another possible issue that may arise for our project is that some drivers may rely too much on our traffic sign system and if there are any bugs for our design, the consequence would also be severe.

Additionally, one of the challenges for our model is that it should work in any environment conditions, including night time and extreme weather such as storming or snowstorm. Furthermore, the training dataset for the situations mentioned above is rare online, so we might have to collect those ourselves.

Project plan

Teamwork is a significant part of our project. We would have weekly off-line meetings during the Wednesday lab section. During the meeting, we will update our progress and assign tasks for the following week. If we get into trouble, we would help each other out and could also ask for the professor's and TAs' help if necessary. Besides, regular online meetings will be made during evenings and weekends so that new information can be shared and updated quickly. Multiple communication methods will be employed, and calling group members is only for emergencies.

Since we will use Google Colab and Github to do our project, we can avoid overlapping our tasks and trying to programme during different time. Also, backups of the current project versions will be made frequently in case of unexpected events.

We would divide work evenly inside the group. For example, each of us completed equal parts in this proposal, and proofread is done together before submission. In the coding part, we would divide work into parts into milestones, and assign internal deadlines to make sure the project is moving forward.

Our project consists of multiple stages, and therefore we have the following internal deadlines at this point to make sure the progress of the project.

Task	Internal Due Date
Traffic sign localization	March 1, 2020
Traffic sign extraction	March 8, 2020
Traffic sign classification	March 15, 2020
Audio output	March 22, 2020
Project Completion	March 28, 2020

Table 1: Internal deadline assigned for the project

The internal due date might change according to the progress from time to time.

Risk Register

One of the potential risks for our project is that we may not have enough time devoted to the project. As workload from all other courses becomes heavy after the reading week, it may be hard for all the team members to spare adequate amounts of time for the project. Therefore, work efficiency would be significant throughout our project. We will have our internal deadlines to ensure the project is at a good pace.

In addition, we may face the problem of stucking in one of our stages and can not move forward. For example, we might spend too much time on the traffic signs localization and extraction but still fail to work it out. In that case, to avoid making no progress, we would skip this stage temporarily and work on the next one, meanwhile, we would ask for external help and keep looking back on the problem.

An unsatisfactory final accuracy is also a risk to our project. In this situation, we would keep doing hyperparameter tuning, including tuning the training hyperparameters and model structure

hyperparameters, to improve the result. If the hyperparameter tuning does not work well, we would change to another type of model and redo the training if time allows.

Besides, the time our model takes to make the prediction could be a problem. If the model takes too long to identify the sign and then gives the audio output, it would cause problems in practice. In case of that, we would reduce the complexity of our model and meanwhile try to maintain the current accuracy.

Link to Github and Colab Notebook

Colab Link: https://colab.research.google.com/drive/1k1MeodC3BQhEBct4RtRXZ_L836czf1SB
Github Repository: https://github.com/Louis-He/APS360_Project

Reference

[1] Medium. (2018). *Image Processing With Deep Learning- A Quick Start Guide*. [online] Available at: https://becominghuman.ai/image-processing-with-deep-learning-a-quick-start-guide-38e166340200 [Accessed 13 Feb. 2020].

[2] X. Han, C. Chen, K. Wang, Y. Yuan and Y. Song, "Study on Detection and Localization Algorithm of Traffic Signs from Natural Scenes", *Advances in Mechanical Engineering*, vol. 6, p. 318635, 2014. Available: 10.1155/2014/318635.

[3] E. Kršák and Š. Toth, "Traffic Sign Recognition and Localization for Databases of Traffic Signs", *Acta Electrotechnica et Informatica*, vol. 11, no. 4, 2011. Available: 10.2478/v10198-011-0039-2.

[4] Andreas Møgelmose, Mohan M. Trivedi, and Thomas B. Moeslund, *Vision based Traffic Sign Detection and Analysis for Intelligent Driver Assistance Systems: Perspectives and Survey*, IEEE Transactions on Intelligent Transportation Systems, 2012.

[5] Mapillary.com. (2020). Mapillary - *Street-level imagery, powered by collaboration and computer vision*. [online] Available at: https://www.mapillary.com/dataset/trafficsign.