Sam Hoang Hong (UCID: 30088823)
Tin Le (UCID: 30089727)
Nam Nguyen Vu (UCID: 30154892)

CPSC 599 Project Proposal

Our main goal for this project is to create a program that can detect 3 types of objects: cars, pedestrians, and traffic lights with high accuracy (aiming to 85% and up). Such a model can be the starting point for the software of an autonomous vehicle. With self-driving cars becoming more and more popular nowadays, exploring this field is a relevant project that could one day further enhance driver safety and efficiency.

We will use one of our team member's dashcams to record video data, which will then be converted to a dataset of images for model building. Since there were no other dashcam datasets online, we chose to develop our own dataset. This avoids potential copyright issues, increases accuracy, and allows us to tailor our model to Calgary. With regards to privacy, we will take measures to keep the dataset protected and secure.

During our preliminary research, we identified YOLO (You Only Look Once) as one of the most robust pre-trained models for object detection tasks. First, we plan on capturing about 3000 images total from the dashcam. We will then label our images using LabelMe. Then, we will fine-tune a pre-trained YOLOv5 model to make it work better with our dataset, while using Torch as the main library. The target is for the model to draw distinguishing boxes around cars, pedestrians, and traffic lights in a given image. After the first model is deployed, we will collect more edge-case images (such as false positive cases from the old model) to further tune and optimize the model. We will use 70% of the dataset for training, 20% for validation, and 10% for testing.

For now, we aim to have about 3000 images of traffic, which should be sufficient to train the model. We will then label them manually for our own use case. The images will be captured in different contexts (day/night, winter/summer, crowded traffic/less traffic). Our team has used LabelMe in the past, and one of us once labeled 5000 railway gate images in 1 week. Therefore, we do not anticipate any issues with labeling all of the images once the work is divided among the three of us.

"Object detection using YOLO: challenges, architectural successors, datasets and applications"[1] describes the architectural design of prior versions of YOLO (up to YOLOv4). It explains the loss function relating to the boxes on the detected objects. Although we will refer to this paper to understand the mechanics of YOLO, we will also use other tutorial articles[2] to help us with coding.

Timeline:

- Phase 1: By Oct 31, collect 3000+ images with objects labeled
- Phase 2: By Nov 10, deploy the first model
- Phase 3: By Nov 20, test the model on new video/images and collect false positives and/or new cases and add to the current dataset
- Phase 4: By Nov 30, test on the second model and write documentation, explanation, next step (further use case after this project closed)

---

[1] Diwan, T., Anirudh, G. & Tembhurne, J.V. Object detection using YOLO: challenges, architectural successors, datasets and applications. *Multimed Tools Appl* **82**, 9243–9275 (2023). https://doi.org/10.1007/s11042-022-13644-y

[2] https://towardsdatascience.com/guide-to-car-detection-using-yolo-48caac8e4ded