EE599 Deep Learning - Initial Project Proposal

March 31, 2019

Project Title: Collision warning system for self-driving vehicle.

Project Team: Jingbang Zhong, Weizhong Jin, Xinyi Tao.

Project Summary: In this project we propose to construct a collision warning system for driverless vehicle to detect dynamic objects in the drivable region. Through the input video from driving recorder, the system can estimate the drivable space in 3D, provide category of obstacles present in lane, and eventually output "Go" and "Stop" command. To implement these function, we need to train convolutional neural networks for object detection and semantic segmentation.

Data Needs and Acquisition Plan: We collect video generated by driving recorder as our raw data, including day and night view. To reduce memory requirement, we will preprocess those videos and resize them into smaller size. Approximately, 15 hours of video will be used for model training.

Primary References and Codebase:

- Ross Girshick. "Fast R-CNN", IEEE International Conference on Computer Vision. 2015.
- S. Ren, K. He, R. Girshick and J. Sun. "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks", 2017 IEEE Transactions on Pattern Analysis and Machine Intelligence.
- Wieland Brendel, Matthias Bethge. **"**Approximating CNNs with Bag-of-local-Features models works surprisingly well on ImageNet**",** ICLR 2019 Conference Blind Submission
- Blog Post: Berkeley DeepDrive**,** coursera Visual Perception for Self-Driving Cars
- GITHub codebase: Faster-R-CNN for Pytorch

Architecture Investigation Plan: We plan to first utilize the architecture used in the above reference. Then, we will explore replacing the Traditional CNN with a BagNet.

Estimated Compute Needs: Based on the data set size in the above paper and the benchmarks in this Fast Simultaneous Object Detection and Segmentation, we estimate that one training run for our initial GAN architecture will take 15 hours on a single Nvidia V100 GPU, which is the GPU resource in the AWS p3.2xlage instance. With spot pricing, which is roughly $1 per hour, we expect $16 per training run. We expect to do a number of provisional runs to tune hyper-parameters. We estimate that this will cost approximate $50. In addition, we expect to do approximately 3 full runs for training other traditional object detection network like Fast-R-

CNN, Faster-R-CNN， YOlOv2 and SSD which brings our total estimated computing cost to roughly $200.

Team Roles: The following is the rough breakdown of roles and responsibilities we plan for our team:

- Jingbang Zhong:, Build and train the required model.
- Weizhong JIn: Get existing codebase running on an AWS instance.
- Xinyi Tao: Data collection and cleaning.

All team members will work on the final presentation, slides, and report.

Requested Mentor with Rationale: We request Keith Michael to be our team mentor because of his expertise in signal processing, detection, estimation. Jiali is our second choice because of his expertise in GANs. We have a good idea of what we want to do and have a good starting point from the paper and codebase, so we are flexible regarding our mentor assignment.