Team 28 ProjF Proposal: Traffic Sign Detection with CNN Architectures

Naga Usha Mahathi Lanka
email: nlanka@ncsu.edu

Dewang Tara
email:dtara@ncsu.edu

I. Motivation

Scene understanding is the ultimate goal of computer vision; detecting and classifying objects of various sizes in the scene is an important sub-task. Recently, deep learning methods have shown superior performance for many tasks such as image classification. Autonomous vehicle technology is being implemented across the automotive industry.

With the self-driving cars, road safety is a major concern that this technology should address. Due to this, machine learning engineers have been working on improving both the accuracy of neural networks that are used for image recognition and speed of the aforementioned algorithms.

The objective of this study is to implement various neural networks such as CNNs for the problem of traffic sign recognition to achieve a real-time efficient prototype for detection. With the increase in the number of object detection algorithms in the past five years, there is a lot of scope to improve both accuracy and speed of detection of traffic signs.

II. Data description

The dataset to be examined contains 100,000 Tencent street view panoramas, and 100,000 images containing 30,000 traffic sign instances.[1] These images cover a large variation in illuminations and weather conditions. Each traffic sign in the benchmark is annotated with a class label and its bounding box. The dataset is called the Tsinghua - Tencent 100K.

We plan to partition the dataset into two parts: a training set with labelled data and a test set without labels. The training set will be composed of 80% of the total dataset and will be chosen at random and the remaining part of the dataset will be allocated to the test set.

III. Methodology

Traffic sign recognition contains two technologies, namely, traffic sign classification (TSC) and traffic sign detection (TSD). The challenge of TSR is to ensure its efficiency, which means adequate accuracy, generalization, and speed in real-time by a computationally limited platform.

1) Baseline: In our baseline model, Traffic sign recognition is performed using a custom made Convolutional Neural network called *Overfeat* Framework. Their network is fully convolutional, and the last layer is branched into three streams: a pixel layer, a bounding box layer and a label layer. This allows the model to simultaneously detect and classify traffic signs. If the label layer is removed, it can be used as a traffic sign detector. Branching the layer helped the model perform better with a more balanced speed and accuracy.

2) Proposed: Our objective in the project is to build on our baseline model by first implementing the R-CNN with the last layer split into separate layers. Then, we plan on using the custom model in the baseline and evaluate the model performance and monitor the accuracy. We will finally implement YOLOv4, the state-of-the-art object detection algorithm for the traffic sign recognition system. We plan on drawing a comparitive conclusion between the three models.

We also want to build on our YOLOv4 based Traffic sign recognition system by splitting the last layer into 3 layers as it was done in the Custom CNN model in [1] to speed up the detection and deliver better accuracy.

IV. Evaluation

Our study serves to find the performance gap between these CNN architecture in our baseline model and YOLOv4 in the context of traffic sign recognition. A part of our evaluation will also test if the model detects the traffic signs with available generic datasets. Performance in this study is a measure of accuracy and precision in determining traffic signs as well as the speed of the detection. Accuracy is the RMSE measure whereas speed can be quantified by the runtime of the algorithm.

References

[1] Zhu, Zhe and Liang, Dun and Zhang, Songhai and Huang, Xiaolei and Li, Baoli and Hu, Shimin, "*Traffic-Sign Detection and Classification in the Wild,*" The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016 doi:10.1109/CVPR.2016.232

[2] J. Redmon, S. Divvala, R. Girshick and A. Farhadi, *"You Only Look Once: Unified, Real-Time Object Detection,"* 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 779-788, doi: 10.1109/CVPR.2016.91.

[3] V. Sichkar, Sergey A. Kolyubin *"Real time detection and classification of traffic signs based on YOLO version 3 algorithm"* 2020 Scientific and technical journal of information technologies, mechanics and optics. doi: 10.17586/2226-1494-2020-20-3-418-424