CHALMERS
UNIVERSITY OF TECHNOLOGY

DAT295 - Autonomous and Cooperative Vehicular Systems

Machine Learning for Traffic Sign Recognition

Project Proposal

Fikri Farhan Witjaksono fikrif@student.chalmers.se

Myeongjin Jang jang@student.chalmers.se

June 10, 2022

1 Introduction

Due to the rapid development and demand in autonomous driving, image classification using deep learning has been important to assist autonomous driving. Nevertheless, image classification has been significantly developed and has shown considerably high accuracy, it is still important to maximize its performance and minimize its weakness.

2 Context

In the literature, we found that the implementation of You Only Look Once version 3 (YOLO v3) [1] has not been done on the dataset from the Swedish Traffic Sign Dataset (STSD). As of now, in what has been done by Sichkar et al. [2], they specifically implemented the algorithms on German Traffic Sign Detection Benchmark (GTSDB) [3] and German Traffic Sign Recognition Benchmark (GTSRB) datasets [4]. The writer used the GTSDB dataset for the localization stage and sequentially used the GTSRB dataset for the classification stage.

3 Goals and Challenges

Regarding the current implementation, we have not observed any similar image classification projects that use the STSD. This project aims to implement image classification on the STSD own and achieve reasonable results indicated by accuracy, loss and confusion matrix performance. We plan to utilize our own custom-made Convolutional Neural Network (CNN) to train.

4 Approach

4.1 Dataset

STSD provided by Fedrik Larsson at Linköping university [5] will be used. It contains more than 20,000 images with 20 per cent labeled and 3488 traffic signs. It is a set of records for highways and cities longer than 350km of Swedish roads. The dataset will be divided as training, validation, and test data set.

In the evaluation, the signs were classified into several class according to its categories as shown below in table 1. 1

Table 1: Classification of Traffic Signs

Categories	**Classes**
Warning signs	Uneven road, Slippery road, etc.
Priority signs	Pedestrian crossing, Obligation to stop, Give way, etc
Prohibitory signs	Maximum speed, No entry, No lorries, etc.
Mandatory signs	Direction to be followed Right only, Foot path, Roundabout, etc.
Instruction signs	Motorway, End of motorway, Clearway,etc.

4.2 Data preprocessing

We have not investigated how well distributed of classes in the dataset yet. Therefore, if the dataset is extremely imbalanced, then some balancing process will be required. Also, only 3488 images contain traffic signs among about 20,000 images in the whole dataset thus, extracting images will be required. In addition, the classes of signs are not digitalized yet, it is expected to vector classes in digital format. In our dataset, if we discover later on that the number of data that are useful for training to reach a reasonable accuracy are not enough, we are considering to utilize several different augmentation or image processing methods that are already available to create additional artificial training data to the existing dataset.

4.3 Tools

Google colab will be mainly used by both students. In the case that a solid GPU is required at the end of the project, Microsoft Azure can be utilized as well.

4.4 Training

The optimal number of training epochs will be determined after a few tests. After a certain number of epochs, the accuracy no longer significantly improves and might cause some overfit problems. Thus, one of the early points of that period will be selected as the optimal number of training epochs. The dataset will be divided into training, validation, and test sub-datasets. After each epoch of training, the model will be tested on the validation dataset and the best one, which shows the highest accuracy, will be saved. The saved best model will be tested on the test dataset as well. The performance metric will be accuracy and loss.

4.5 Expected results & Evaluation

This project is expected to produce results in terms of accuracy, loss, Receiver Operating Characteristics (ROC) Curve and confusion matrix. ROC [6] is a useful performance metric for multi-class classification tasks which is derived from the metrics in confusion matrix, where confusion matrix will give us Precision and Recall which could be used to construct ROC curve [7]. Precision and Recall could be expressed mathematically as the ratio between True Positives (TP) and sum of TP and False Positives (FP), whereas Recall could be expressed mathematically as ratio between True Positives (TP) and sum of TP and False Negatives (FN).

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

4.6 Network Architecture

For image classification, a simple CNN will be utilized. The specific parameters for the CNN will be determined during our training session. Preliminary model related functions and libraries are written in the following table 2

Table 2: Model related functions and libraries

Optimizer	Adam or SGD
Loss function	Cross-Entropy Loss
Activation function in the output layer	Softmax
Expected library	Either pytorch or Tensorflow. OpenCV, scikit-learn, numpy, PIL, pandas, etc.
Tools	Google colab (Mainly) and Azure (When it's necessary)

References

[1] J. Redmon and A. Farhadi, "Yolov3: An incremental improvement," 2018. [Online]. Available: https://arxiv.org/pdf/1804.02767.pdf

[2] V. Sichkar and S. Kolyubin, "Real time detection and classification of traffic signs based on yolo version 3 algorithm," *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, vol. 20, no. 3, p. 418–424, 2020.

[3] S. J. S. M. I. C. Houben S., Stallkamp J., "Detection of traffic signs in real-world images: The German traffic sign detection benchmark." *Proc. of the International Joint Conference on Neural Networks (IJCNN 2013)*, p. 6706807, 2013.

[4] S. J. I. C. Stallkamp J., Schlipsing M., "Man vs. computer: Benchmarking machine learning algorithms for traffic sign recognition." *Neural Networks*, vol. 32, p. 323–332, 2012.

[5] F. Larsson, M. Felsberg, and P.-E. Forssen, "Correlating Fourier descriptors of local patches for road sign recognition," *IET Computer Vision*, vol. 5, no. 4, pp. 244–254, 2011.

[6] T. Fawcett, "An introduction to roc analysis," *Pattern Recognition Letters*, vol. 27, p. 861–874, 2006. [Online]. Available: https://arxiv.org/pdf/1405.0312.pdf

[7] J. Davis and M. Goadrich, "The relationship between precision-recall and roc curves," *ACM International Conference Proceeding Series*, vol. 148, p. 233–240, 2006.