Shahid Beheshti University
Mathematical Sciences Faculty

Report *of*
Deep Learning Course Project

German Traffic Sign Classification

by **Zahra Oruji**

Masters Student of Data Science

Under Supervision *of*
Dr. Hossein Hajiabolhassan

February 2021

Introduction

Traffic signs are an integral part of our road infrastructure. They provide critical information, sometimes compelling recommendations, for road users, which in turn requires them to adjust their driving behavior to make sure they adhere with whatever road regulation currently enforced. For Autonomous Driving, the vehicle must know about the surrounding infrastructure like maximum allowed speed, stop signs, and yield signs. These require them to be able to perceive traffic signs and understand the surroundings. Traffic Sign Classification (TSC) plays a major role in deciding about the behavior of self-driving cars and helps them prepare for events like pedestrian crossings in advance.

Sign recognition is a multi-category classification problem with unbalanced class frequencies. Traffic signs show a wide range of variations between classes in terms of color, shape, and the presence of pictograms or text. However, there exist subsets of classes (e.g., speed limit signs) that are very similar to each other. The classifier has to cope with large variations in visual appearances due to illumination changes, partial occlusions, rotations, weather conditions, scaling, etc.

In recent years, the deep learning methods for solving classification problem have become extremely popular. Due to its high recognition rate and fast execution, the convolutional neural networks have enhanced most of computer vision tasks.

In this project the **German Traffic Sign Recognition Benchmark (GTSRB)** is presented, which is a large, lifelike dataset of more than 50,000 traffic sign images in 43 classes. Deep neural networks and convolutional neural networks will be used to classify traffic signs.

History

Several approaches to traffic sign recognition have been published. An integrated system for speed limit detection, tracking, and recognition is presented [1]. The classifier is trained using 4,000 samples of 23 classes, with samples per class ranging from 30 to 600. The individual performance of the classification component is evaluated on a training set of 1,700 traffic sign images with a correct classification rate of 94 %. Moutarde et al. present a system for recognition of European and U. S. speed limit signs based on single digit recognition using a neural network [2]. A drawback of their work is that, the individual classification results are not provided. The overall system including detection and tracking achieves a performance of 89 % for U. S. and 90 % for European speed limits, respectively, on 281 traffic signs. Broggi et al. [3] used several neural networks to classify different traffic signs. Shape and color information from the detection stage is used to select the appropriate neural network. In this work, only, qualitative results are provided. In another study, a number-based speed limit classifier is trained on 2,880 images [4]. It achieved a correct classification rate of 92.4 % on 1,233 images. However, it is not clear whether the images of the same traffic sign instance are shared between sets or not. Various approaches are compared a dataset containing 1300 preprocessed examples from 6 classes (5 speed limits and 1 noise class) [5]. The best classification performance observed was 97%.

Dataset

The dataset was created from 10 hours of video that was recorded while driving on different road types in Germany during daytime.

The dataset contains more than 50,000 traffic sign images in 43 classes and images of more than 1,700 traffic sign instances. The size of the traffic signs varies between 32 × 32 and 128 × 128 pixel.

The dataset is splitted into three subsets. Subset I was published as training data, subset II as validation data and subset III as test data.

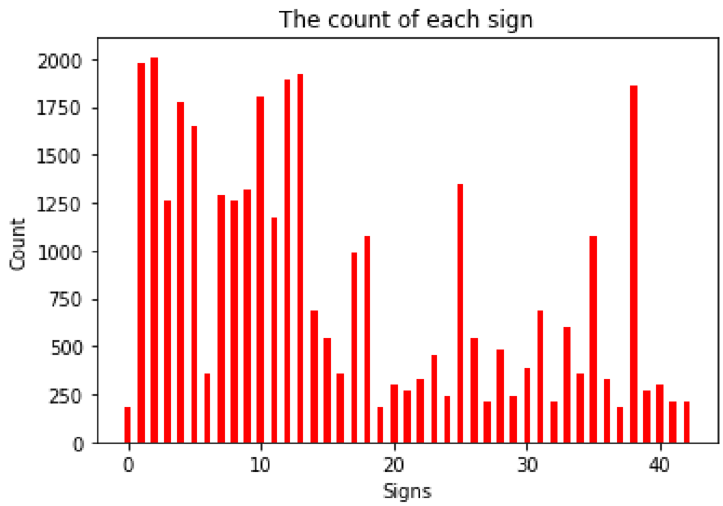
Training subset contains **34799** Images, test subset contains **12630** images and the validation subset contains **4410** images.

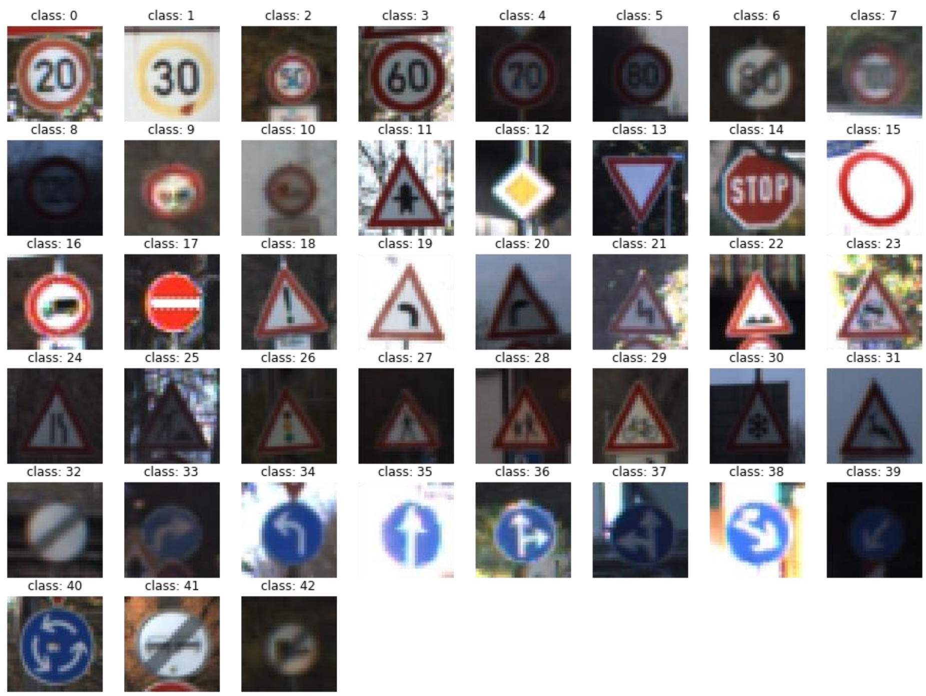
Project plan

In this project the **German Traffic Sign Recognition Benchmark (GTSRB)** is presented, which is a large, lifelike dataset of more than 50,000 traffic sign images in 43 classes.

First, we call dataset and perform preprocessing operation using pytorch libraries.

We draw diagrams for different classes.





There is also a significant imbalance across classes in the training set, as shown in the histogram below. Some classes have less than 200 images, while others have over 2000. This means that our model could be biased towards over-represented classes, especially when it is unsure in its predictions. We will see later how we can mitigate this discrepancy using data augmentation.

Pre-Processing Steps

We initially apply two pre-processing steps to our images:

Grayscale

We convert our 3 channel image to a single grayscale image.

Image Normalisation

We center the distribution of the image dataset by subtracting each image by the dataset mean and divide by its standard deviation. This helps our model treating images uniformly.

Model Architecture

- We mainly tried 5x5 and 3x3 filter (kernel) sizes, and start with depth of 30 for our first convolutional layer.
- The network is composed of 2 convolutional layers. kernel size is 5*5, Using RELU as the activation function, each followed by a 2x2 max pooling operation. And we use Cross Entropy Loss function and stochastic gradient descent optimizer. We set learning rate equal to 0.001 and momentum = 0.9

- After running the model, we reach the following result:

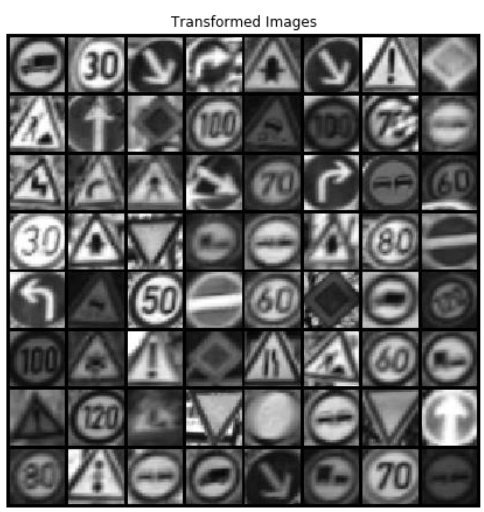
  Test loss: 0.969987 Test accuracy : 84.014%

- By running the model on the gray data, we get Test accuracy : 88.036%.

- After apply Flipping on the model we achieve Test accuracy : 88.187%

- And After applying data augmentation we achieve Test accuracy : 93.793%
- Finally, we run a 3-layer CNN model, with using Batch Normalization in 2 layer and Max pooling 2*2 and Dropout in the last layer. Also we use early stopping and model stopped at 39th epoch and achieve Test accuracy : 98.884%
- Also by running Spatial transformer network, we achieve Test accuracy : 99.327%

Visualizing the STN results:



Conclusion

By running the above models on the pre-processed data set, we conclude that STN Model with accuracy=99.327% has better results.

References

[1] Moutarde, Fabien, et al. "Robust on-vehicle real-time visual detection of American and European speed limit signs, with a modular Traffic Signs Recognition system." *2007 IEEE Intelligent Vehicles Symposium*. IEEE, 2007.

[2] Broggi, Alberto, et al. "Real time road signs recognition." 2007 *IEEE Intelligent Vehicles Symposium*. IEEE, 2007.

[3] Keller, Christoph Gustav, et al. "Real-time recognition of US speed signs." *2008 IEEE Intelligent Vehicles Symposium*. IEEE, 2008.

[4] Muhammad, Azam Sheikh, et al. "Analysis of speed sign classification algorithms using shape based segmentation of binary images." *International Conference on Computer Analysis of Images and Patterns*. Springer, Berlin, Heidelberg, 2009.

[5] Bascón, S. Maldonado, et al. "An optimization on pictogram identification for the road-sign recognition task using SVMs." *Computer Vision and Image Understanding* 114.3 (2010): 373-383.