Traffic Signs Recognition

Anna Antończak
Bruno Jamuła

Wroclaw University of Science and Technology

Abstract. The objective of this paper is to do a machine learning model to recognize traffic signs. All images were reprocessed with methods grayscale and histogram equalization to extract features needed to perform model training. In this paper was done a comparison of three classifiers: Random Forest, SGD Classifier and SVC Classifier.

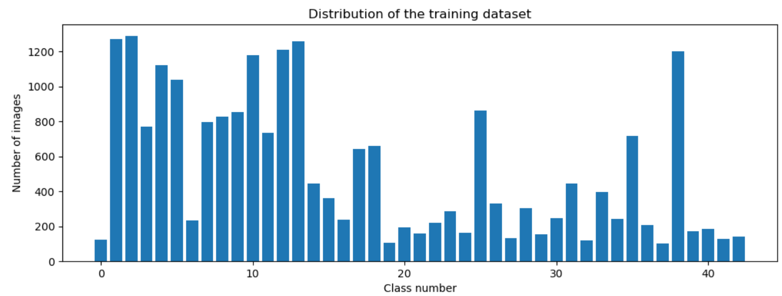
Keywords: image recognition · trafic signs recognition · machine learing · random forest

1 Introduction

In recent years intelligent cars are becoming a standard. Mainly interest of drivers caught driver support system. This systems can prevent from blocking wheels, monitor tiredness of driver, active cruise control or call emergency service in case of car accident. In this work we want to focus on systems that use image recognition, namely traffic sign recognition. Traffic sign recognition is a multi-category classification problem with unbalanced class frequencies. In today's world signs that are mainly recognized by cars are speed limits signs. Systems based on image recognition are dependent on weather and light conditions. Car companies does not share many information about their traffic sign recognition system and about precision of this systems. System from Opel, named as Opel eye recognizes speed limits and no entry signs and informs driver about end of speed limit [4]. In this paper we described machine learning experiment of recognizing traffic signs based on paper [5].
The paper is organized as follows: Section 2 describes used dataset and how it was prepared to processing. Section 3 presents problem statement. In section 4 is described project of experiment and listed used learning algorithms. Results are presented and described in section 5. In section 6 shows conclusions.

2 Dataset

We used dataset provided by [5]. It was created from 10 hours of video recorded while driving in German during daytime in March October and November 2010. Dataset contains 51 840 images of the 43 classes. Classes distribution are shown in fig 5. It contains images of the traffic sign in size between 15 x15 and 222x193 pixels.



Dataset classes distribution

Dataset was devided for two subsets - training set which consists of 75% of the whole dataset and test set with consisting 25%.

2.1 Image processing

Image processing enables exacting information and characteristics from image. It uses specific operations to achieve that. To perform digital processing image has to be transformed into the matrix of pixels.
Some of image processing methods are grayscale and histogram equalization and they are used in our experiment.

Grayscale Grayscale image is and image which is a one channel, in contrast to for example RGB where there are three channels. Each pixel is represents a sample of intensity information what results that image composed of shade of gray. It can be used when less information of a pixel is needed or when some other processing operations can not be done on other scales. In grayscale intensity is stored as integer in scale from 0 to 256 (from black to white).

Histogram equalization Histogram represents density distribution of an image, which means that for each intensity value is assigned number of pixels from the image. Histogram equalization is a technique which enables to improve image contrast without loss of information by spreading out the most frequent intensity values. However it can not be performed on Red, Green, Blue scale, but image has to be converted to another scale first[7].

3 Problem statement

The goal of this experiment is to create a machine learning models that recognizes traffic signs and then compare them. We used following classifiers to train the model:

– Random Forest Classifier,

– SGD Classifier,
– SVC Classifier.

To measure a performance of each classifier we used three metrics: **accuracy score** which is ratio of correctly predicted observation to the total observations, **F1 score** which is weighted average of the precision and recall and **The Hamming Loss** - the fraction of labels that are incorrectly predicted.

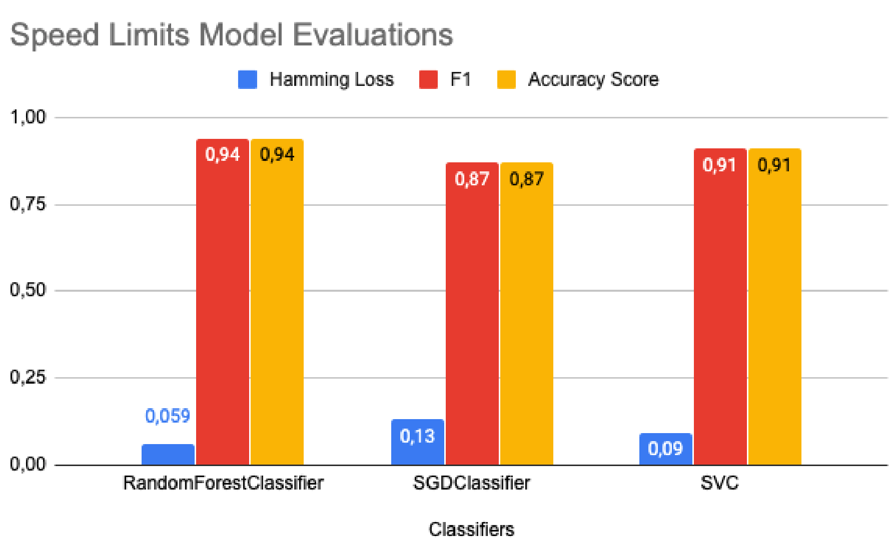
4 Project of experiment

As was mention in previous section, we used three classifiers: Random forest, SGD and SVC in context of traffic signs recognition.

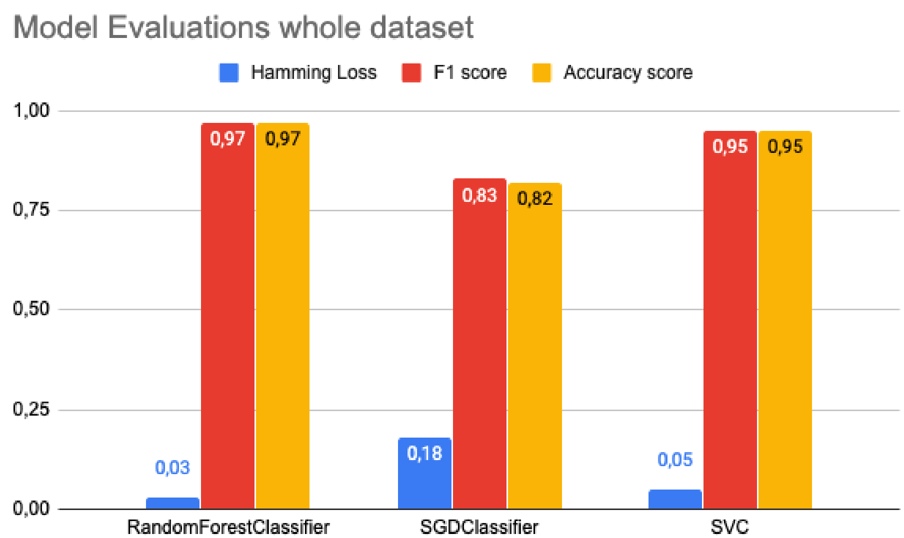
The psuedo code of algorithm looks as follows:

1. Loading dataset
2. Preprocessing data
   2.1. Grayscale images
   2.2. Equalizes the histogram of a grayscale images
   2.3. Resize the image to a fixed size, then flatten the image into one dimensional array
3. Initializing classifiers
4. Fitting classifiers with the training set
5. Checking performance with the testing set
6. Analyzing the result

5 Experiments results



Classifiers comparison trained on speed limits signs



Classifiers comparision trained on whole dataset

In the figure 1 we can see that Random Forest Classifier achieved the best result according to F1 metric and accuracy score. We achieved almost same accuracy as researchers in paper [5]. We can also see that our model would preform better if we test speed limits signs with SGD classifier but worse with other classifiers. In general we can assume that that there is no difference between using all kind of signs or just speed limits.

6 Conclusion

This project allowed us to better understand image processing. We learned that there is need to preprocess images before training the classifiers. From our results we can also see that different classifiers can achieve different results. In our specific problem Random Forrest Classifier achieved best results.

References

[1] G. Bradski. "The OpenCV Library". In: *Dr. Dobb's Journal of Software Tools* (2000).

[2] Arturo De la Escalera, J Ma Armingol, and Mario Mata. "Traffic sign recognition and analysis for intelligent vehicles". In: *Image and vision computing* 21.3 (2003), pp. 247–258.

[3] Markus Mathias et al. "Traffic sign recognition—How far are we from the solution?" In: *The 2013 international joint conference on Neural networks (IJCNN)*. IEEE. 2013, pp. 1–8.

[4] Jan Matuszewski. “System inteligentny rozpoznawania znaków drogowych”. In: *Przeglad Elektrotechniczny* 92.1 (2016), pp. 41–44.

[5] Johannes Stallkamp et al. “Man vs. computer: Benchmarking machine learning algorithms for traffic sign recognition”. In: *Neural networks* 32 (2012), pp. 323–332.

[6] Johannes Stallkamp et al. “The German traffic sign recognition benchmark: a multi-class classification competition”. In: *The 2011 international joint conference on neural networks*. IEEE. 2011, pp. 1453–1460.

[7] Shreenidhi Sudhakar. *Histogram Equalization*. 2017. URL: `https://towardsdatascience.com/histogram-equalization-5d1013626e64`.