Traffic Sign Classification using Convolutional Neural Networks[1]

Hesam Damghanian
SBU, Computer Science Dept.
h.damghanian@mail.sbu.ac.ir

***Abstract*— I apply Convolutional Networks (ConvNets) to the task of traffic sign classification on GTSRB dataset consisting 43 different class of traffic signs. Converting given input color images to grayscale does not affect the accuracy negatively and just removes redundant the computational burden from the model. The Best accuracy yielded after rounds of experiments was 96.28% on test dataset.**

Keywords—CNN, GTSRB, histogram equalization

I. Dataset

The given dataset is consisted of 34799, 32*32 color images of German traffic signs taken with different distance, lighting conditions. Class populations are not the same and some classes have more images than others, Fig. 1. demonstrates the variations of images along each class. As it shows the dataset is not balanced.

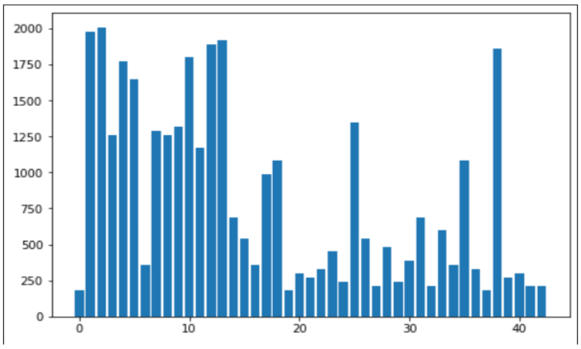

Fig. 1.

A simple discovery in classes shows that the dataset is heavily augmented and there are images in a class that are repeated multiple times but with different rotations and translations which leaves no place for further augmentation.

II. Preprocessing

Experiments were conducted on two variations of the given dataset. One with cropped version of 3 channel images Fig. 2a. and the other with their gray scaled equalized histogram, Fig. 2b. (also cropped images). Equalized histogram method is useful in this case that images are not in the same lighting conditions. This results in significantly better accuracy as we can see in Results section.

Fig. 2a.

Fig. 2b.

Equalized histogram

This method usually increases the global contrast of images, especially when the usable data of the image is represented by close contrast values. Through this adjustment, the intensities can be better distributed on the histogram. This allows for areas of lower local contrast to gain a higher contrast. Histogram equalization accomplishes this by effectively spreading out the most frequent intensity values. [Wikipedia]

III. Experiments

As a practice I used transfer learning method to use pretrained models such as Xception and MobileNet_V2 (weights from ImageNet dataset) with two aforementioned versions of GTSRB dataset. Then multiple versions of ConvNets were implemented, also with two mentioned versions of dataset, as their summary can be seen in attached notebook.

IV. Results

Table 1 shows the results of the implemented models. An important finding is that results on the second dataset are significantly better because of the method used to equalize images respect to their lighting conditions. Also, as we can see, as the model goes deeper and wider, its accuracy in classifying the signs increases. Another result is that pretrained models that has been trained on color images from imagenet dataset have lower accuracy in comparison to a simple ConvNet.

TABLE I. ACCURACY ON TEST DATASET

Dataset	**Models**				
	One trainable layer (last layer)		**All layers trainable**		
	Xception	***MobileNet_v2***	***CNN1***	***CNN2***	***CNN3***
First dataset (only cropped)	75.84 %	77.76 %	**77.90%**	-	-
Second dataset (equalized histogram)	73.24 %	-	86.17 %	92.22 %	**96.27%**

*Models from cnn1 to cnn3 grow deeper and wider. Xception and MobileNet models are loaded with pretrained weights from imagenet dataset.

In Fig. 3. We can see a sample from the first Conv layer's output in CNN3 model. Output of other conv layers and related technical details are available in attached notebook.

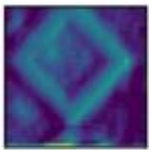

Fig. 3.

1. Technical notes are in attached notebook

V. DISCUSSION

A quick review on the train, test and validation datasets shows that images are already augmented and further augmentation may have negative impact on generalization of the model. As [1] suggests color images are not necessary in classification tasks like this one and also may have negative impact on accuracy, in contrast to [2]. Signs with similar shapes like circles, triangles etc. are more likely to be misclassified. (e.g., speed limit (20 km/h) and speed limit (70 km/h). a simple way to avoid this misclassification is to acquire more data with higher resolution to feed the model. Due to the unbalanced dataset the model is biased to classes with more populations. One way to avoid this is to reduce more populated classes to same size as the average population but this method also would not be so great because there is a tradeoff between dataset size and accuracy of the model and techniques to generate more data like augmentation would not be so effective because, as mentioned, images that we already have are not unique.

REFERENCES

[1] Pierre Sermanet, Yann LeCun. Traffic Sign Recognition with Multi-Scale Convolutional Networks. In International Joint Conference on Neural Networks, 2011.

[2] Paclk, P and Novovicov, J. Road sign classification without color information. In Proceedings of the 6th Conference of Advanced School of Imaging and Computing, 2000.