Traffic Sign Classification - Better and Faster

Tejas R, Soham Patil, Darshan V, Siva Krishna P Mothukuri, Shashidhar G Koolagudi
Department of Computer Science and Engineering
NITK Surathkal
Surathkal, India
{tejas1908, sohampatil798, darshan.blh,msivakrish}@gmail.com, koolagudi@nitk.edu.in

***Abstract*—Traffic signs help drivers in vehicles think about the traffic principles and help the drivers for better and safe driving. It is also a key component of the Advanced Driver Assistance System (ADAS) and intelligent vision assistance applications that provide drivers with traffic sign information. It is required to use a recognition system that is accurate and has a high processing speed. Traffic signs in different orientations and different lighting conditions, make the task of traffic sign recognition a challenge. Our recognition system is based on Convolutional Neural Networks (CNN) with layers which can be trained and enhanced architecture. Our proposed model has a performance comparable to state-of-the-art methods in terms of both processing speed and accuracy.**

***Index Terms*—Softmax,Neural Network, fully connected layers**

I. Introduction

Due to the expansion in the amount of vehicles drivers experience perils while driving and this may in like manner cause setbacks. Many mishaps are happening each year throughout the world. These setbacks are overwhelmingly an aftereffect of the driver's weakness to process all the visual data that is accessible while driving. Traffic signs, which are present on the road to help drivers navigate, are not seen by the drivers. As indicated by a PRS Legislative Report, in 2015, there were around five lakh street mishaps in India, which claimed the lives of about 1.5 lakh individuals and harmed around five lakh individuals.

Besides, when constructing a traffic sign recognition framework, care must be taken that the outcomes given by the framework are precise because a traffic sign recognition framework must be dependable as it is utilized continuously. If not solid, a misrecognition may prompt mistaken tasks and lamentable outcomes. There are many degrees for misrecognition of signs, which must be dealt with. Variable size of traffic signals, obscuring of photographs, diverse lighting conditions and distinctive introduction of traffic signs are a couple of things which are dealt with.

The inspiration driving ADAS and vision help applications in vehicles are the automation of the vehicle structures for security and improving the experience of driving. Even though there have been numerous endeavors to improve the precision of traffic acknowledgment frameworks, they had not been contrasted with a standard dataset until the arrival of German Traffic Sign Recognition Benchmark (GTSRB) in 2011. We have utilized the datasets given by GTSRB to discover the exactness of our proposed model, which will enable us to contrast our model with others' models effectively.

This paper is sorted out as pursues: Section II gives a concise audit of the current work done in the field of traffic sign recognition, Section III portrays the approach utilized in our work, Section IV has the outcomes we acquired from our work and Section V has the conclusion.

II. Literature Survey

The work on Traffic sign recognition in real time systems were successfully implemented till like,2000s. Nonetheless due the difficulties like blurring,weather condition,faded color,unclear pictures due to vehicles motion,Scene Complexity etc.. its still a challenging problem.

Work on this has been mainly revolving around three methods: Color Based methods, Shape based methods and machine learning methods. Machine learning algorithms which have been used are mainly support vector machines and deep learning algorithms. Moreover, in deep learning algorithms, CNNs have really showed great results.

[1] uses transfer learning over the pre-trained model VGG-16 and has shown about 96% accuracy on GTSRB.In [4] false positives are removed by SVM classifier and then rest is classified based HOG feature and gave an accuracy of 97.75%.[2] uses Multi-task CNN on the detected images by there own model, and has produced upto 99% accuracy.[3]Traffic sign recognition with hinge loss trained convolutional neural networks gave 99.56%.[5]Spatial Transformer Networks have reached the accuracy of 99.%. Currently, the state of the art in Image detection is ResNet model. ResNet is a pre-trained model (Winner of ILSVRC 2015 in image classification, detection, and localization) which is intensively used in classification of Images.We have used transfer learning with model and compared the accuracy with the proposed model. Except [4] most of the papers didnt touch upon the time taken to classify the image. And [4] has shown that it outperforms the previous best [5] in time consumption. [4] justifies the tradeoff and shows that the simple model developed takes a time quantum of 3 ms with an accuracy of 97.7%.

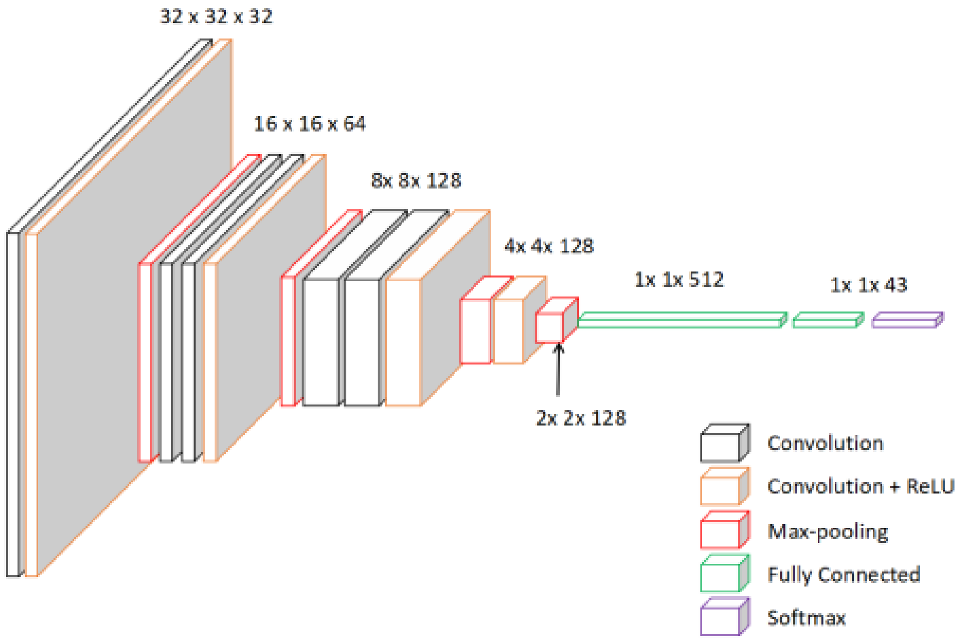

Fig. 1. CNN Architecture

III. Methodology

A. Dataset

Convolutional neural networks have proved to be exceptional for image classification tasks. But deep convolutional neural networks need a large amount of data to train on. Because of the nonattendance of a comprehensively accessible dataset with enough picture information, execution of various methodologies couldn't be looked at for traffic sign recognition until the arrival of German Traffic Sign Recognition Benchmark (GTSRB) and the German Traffic Sign Detection Benchmark (GTSDB) which are two openly accessible broad datasets for the correlation of methodologies for Traffic sign recognition and detection tasks individually.

In this paper we have used the GTSRB dataset to compare the performance of the state of the art ResNet 50 model and our own model. The GTSRB dataset has 43 different classes of images and more than 50000 images in total. It is a large lifelike database suited for single-image multi-class classification problem. Physical traffic sign instances are unique within the dataset. It includes many types of complex traffic signs such as sign tilt , uneven lighting , occlusions and distortions .Hence this is a good choice for training models which can be deployed for practical purposes.

B. Image preprocessing

The images in the GTSRB dataset is of size ranging from 15 x 15 to 250 x 250. This is first resized to 32 x 32.For the first experiment , the proposed model was trained on grayscale images . The colored images were converted to grayscale by taking the weighted average of the 3 channels R. G and B. Then these grayscale images were normalized so that the pixel values was between 0.1 and 0.9. In our second experiment , we trained the model directly on the colored images which were resized to 32 x 32. The intuition behind this is that the human eye classifies the signs based on the color and shape. Hence, retaining the color information in the data gets the trained neural network closer to that of the human brain. As expected , colored data produced better results as compared to grayscale data.

C. Network Structure

The neural network model that we have used for traffic sign classification has the following structure as shown in Fig 1. It contains four stacked convolution layers pursued by two fully connected layers and Softmax. Every convolution stack has an arrangement of convolution layers pursued by a maximum pooling layer. The last convolution layer of each stack is trailed by a batch normalization layer and activated using ReLU function. This is followed by two fully connected layers whose output is given to the Softmax layer to obtain the final output probabilities of all the image classes.

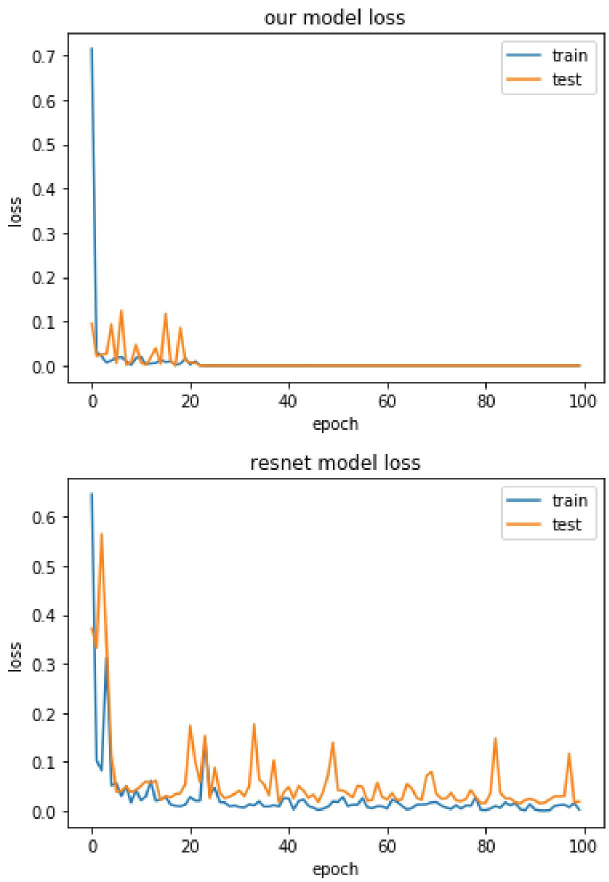

Fig. 2. Comparison of training and testing loss of the two models

D. *Feature Extraction and Model training*

The GTSRB dataset is trained on the above model. For each input image , the first convolution stack extracts the basic edge features using 32 convolution kernels .ReLU operation is then performed on these output filter maps to introduce non-linearity .

This information is further filtered and only the maximum pixel values are passed onto the next stack using the max-pooling operation.This is then used to extract 64 feature maps in the second convolution stack and so on. After passing through the last convolution stack and max-pooling layers, the feature maps are flattened out and given as input to the FC layers.

IV. EXPERIMENTATION AND RESULTS

The images from the GTSRB dataset were first resized to 32x32. The dataset has 39209 colored images in the training set and 12630 colored images in the testing set. The images in the testing set are unseen images and serve the purpose of testing the model. The training set is further randomly split into the training and the validation set in the ratio 80:20 . The training set is used to tune the weights of the neural network and the validation set is used to calculate the loss of the network after every iteration.

TABLE I
COMPARISON OF ACCURACIES OF RESNET50 AND THE PROPOSED MODEL

CNN Model	*Training Accuracy (%)*	*Testing Accuracy (%)*
ResNet 50	99.4	95
Proposed Model	100	98.66

A. *Proposed Model*

We performed experiments on the above mentioned model. In the first experiment , we converted the colored images to grayscale using weighted mean of R,G and B channels and then normalized them so that the values of pixels is between 0.1 and 0.9.The accuracy obtained by our model was 97.48% . In the second experiment, we trained the same model on colored images. The loss function used is categorical cross entropy and the optimizer is Adam. The number of training images used for training per batch is 128. The model is trained for 100 epochs and an accuracy of 98.66% is obtained.

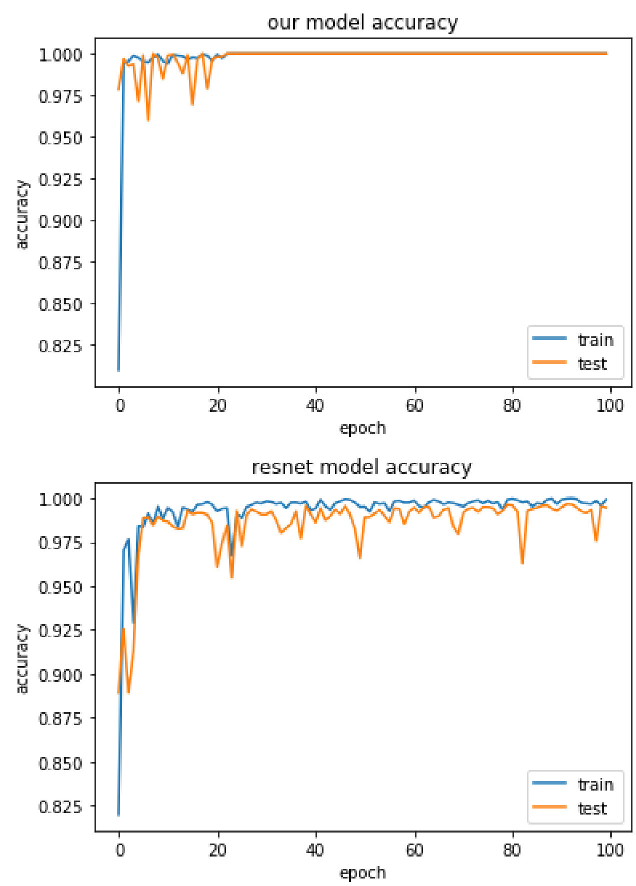

Fig. 3. Comparison of training and testing accuracies of the two models

TABLE II
COMPARISON OF TEST ACCURACIES AND CLASSIFICATION TIME FOR THE TWO MODELS

CNN Model	*Testing Accuracy (%)*	*Classification time per image (ms)*
Proposed Model	98.66	0.296
ResNet 50	95	0.554

B. ResNet 50 model

We used transfer learning for training the ResNet 50 model. A pre-trained ResNet 50 model which was trained on the ImageNet dataset consisting of 1000 different classes of images was used. The last FC layer of this model was pruned and a new FC layer of 43 nodes was added to suit our application. An additional Softmax layer was added after the FC layer for obtaining the output probabilities.This model was trained on the colored images using transfer learning. The loss function used was categorical cross entropy and the optimizer was Adam . The batch size used was 128 and the training was trained for 100 epochs . The accuracy obtained from this model is 95%.

Fig. 2 is a comparison of the plots for training and testing loss. We observe that the training and testing accuracy for proposed model obtain stability after 25 epochs whereas ResNet 50 model has some inconsistencies in the testing loss and accuracies even after epoch 80.

Table 1 summarises the training and testing accuracies obtained using the two models. We see that the proposed model outperforms the state of the art approaches in our application.

Fig. 3 compares the plots of train and test accuracy for the proposed model and the ResNet 50 model. Further , a comparison of the testing time for the proposed model and the ResNet 50 model is performed. Testing is performed on 12630 images and the total time taken is divided by the number of images . We observe that the proposed model is almost twice as fast as the ResNet 50 model during classification . Hence it is closer to a real time system.

Table 2 summarises the testing accuracies for the two models and the time taken for classification per image.

Hence we see that the proposed model has outperformed ResNet 50 trained on the Imagenet dataset both in terms of classification time and classification accuracy. This model can further be fine tuned for better accuracy and speed which can then be deployed in a real time system.

V. CONCLUSION

In this article, we have proposed a traffic sign recognition system which can be used in Advanced Driver Assistance Systems using Convolutional Neural Networks that is on par with the state of the art methods, being both accurate and fast, and hence more reliable than other traffic recognition systems in academia.

REFERENCES

[1] C. Wang, "Research and Application of Traffic Sign Detection and Recognition Based on Deep Learning," 2018 International Conference on Robots Intelligent System (ICRIS), Changsha, 2018, pp. 150-152.

[2] H. Luo, Y. Yang, B. Tong, F. Wu and B. Fan, "Traffic Sign Recognition Using a Multi-Task Convolutional Neural Network," in IEEE Transactions on Intelligent Transportation Systems, vol. 19, no. 4, pp. 1100-1111, April 2018.

[3] J. Jin, K. Fu and C. Zhang, "Traffic Sign Recognition With Hinge Loss Trained Convolutional Neural Networks," in IEEE Transactions on Intelligent Transportation Systems, vol. 15, no. 5, pp. 1991-2000, Oct. 2014.

[4] Y. Yang, H. Luo, H. Xu and F. Wu, "Towards Real-Time Traffic Sign Detection and Classification," in IEEE Transactions on Intelligent Transportation Systems, vol. 17, no. 7, pp. 2022-2031, July 2016.

[5] Nick Schneider, Lukas Schneider, Uwe Franke, Thomas Brox, Andreas Geiger Sparsity Invariant CNNs, September 2017

[6] D. Ciresan, U. Meier, J. Masci, and J. Schmidhuber, Multi-column deep neural network for traffic sign classification, Neural Netw., vol. 32, pp. 333338, Aug. 2012.

[7] S. Houben, J. Stallkamp, J. Salmen, M. Schlipsing, and C. Igel, Detection of traffic signs in real-world images: The German traffic sign detection benchmark, in Proc. IEEE Int. Joint Conf. Neural Netw. (IJCNN), Aug. 2013, pp. 18.

[8] 8. Oruklu, E., Pesty, D., Neveux, J., Guebey, J. (2012). Real-time traffic sign detection and recognition for in-car driver assistance systems. 2012 IEEE 55th International Midwest Symposium on Circuits and Systems (MWSCAS), 976-979.