Traffic Sign Classification and Detection of Indian Traffic Signs using Deep Learning

Manjiri Bichkar, Suyasha Bobhate, Prof. Sonal Chaudhari

Department of Computer Engineering, Datta Meghe College of Engineering Airoli, Navi Mumbai, Maharashtra, India

ABSTRACT

This paper presents an effective solution to detecting traffic signs on road by first classifying the traffic sign images us-ing Convolutional Neural Network (CNN) on the German Traffic Sign Recognition Benchmark (GTSRB)[1] and then detecting the images of Indian Traffic Signs using the Indian Dataset which will be used as testing dataset while building classification model. Therefore this system helps electric cars or self driving cars to recognise the traffic signs efficiently and correctly. The system involves two parts, detection of traffic signs from the environment and classification based on CNN thereby recognising the traffic sign. The classification involves building a CNN model of different filters of dimensions 3×3, 5×5, 9×9, 13×13, 15×15, 19×19, 23×23, 25×25 and 31×31 from which the most efficient filter is chosen for further classifying the image detected. The detection involves detecting the traffic sign using YOLO v3-v4 and BLOB detection. Transfer Learning is used for using the trained model for detecting Indian traffic sign images.

Keywords : CNN, BLOB, EDAS, GTSRB

I. INTRODUCTION

Developing an automated driver guidance system is very important in the context of Indian road conditions. Be it a human driver driving a car or a self driving car, following traffic signs on Indian roads is necessary. A lot of traffic sign violations are seen and heard of ,mainly the reasons for which are over speeding, broken traffic signs, distorted signs, night time conditions, etc. In India there are multiple roads where traffic signs do not survive and are mostly broken are covered in trees or any hindrance. Traffic signs are mostly damaged and distorted by the civilians itself which causes more traffic sign violations. In case of self driving cars, if such conditions prevail it will hardly be able to recognise the traffic sign and may cause serious problems for all vehicles surrounding it. Road safety is always a problem everywhere, especially in developed countries like the US, India etc. Thus it is necessary to develop a system which will help recognise traffic signs without being affected by these anomalies. Ultimate aim of such Intelligent Transport Systems (ITS)[2] is to realize fully autonomous vehicles. Many systems have been proposed and implemented to achieve different features of the ITS. An important field in IT'S is driver assistance systems (DAS)[3]. Electronic driver assistance systems (EDAS) are installed in vehicles to alert drivers of upcoming road signs. To identify road boundaries and barriers, as

well as recognise road signs, computer vision-based methods with high resolution can be used. The manual control of road signs becomes more difficult as vehicle speeds increase. As a result, a vision-based road sign detection and recognition device is desirable for catching a driver's attention and avoiding traffic hazards. For autonomous driving or driver assistance systems, automatic identification of traffic signals is important. There are some advantages to the issue of traffic sign identification. For starters, traffic sign designs are unique, so object variations are minimal. Further-more, sign colours often contrast well with the surrounding area. Furthermore, signs are rigidly arranged in relation to the environment, and are often placed in plain sight of the driver. The form and colour of traffic signs are often used in traffic sign detection algorithms. Shape-based methods detect signs by comparing them to a collection of predefined models, making them susceptible to complete or partial occlusion as well as target rotation. Color-based approaches use the pixel intensity in RGB or HSI colour spaces to detect signs in a scene[4]. Since it provides various pieces of information in each component, HSI is the most widely used colour space. The HSI colour space also has the advantage of being insensitive to brightness and shadows. Furthermore, the HSI colour space is ideal for extracting colour features under difficult conditions such as extreme weather and damaged road signs. Using colour information and CNN, this paper describes a general method for detecting and classifying traffic signs from image sequences. Colour based segmentation techniques are employed for traffic sign segmentation. Red, blue, yellow and white colours are the most commonly used signs in road traffic.

II. LITERATURE REVIEW

1) : [5]J. Stallkamp, M. Schlipsing, J. Salmen, and C. Igel, "The German traffic sign recognition benchmark: a multi-class classification competition," in Proc. IEEE IJCNN, 2011, pp. 1453–1460. This paper proposes the design and analysis of the "German Traffic Sign Recognition Benchmark" dataset and competition. The results of the competition show that state of-the-art machine learning algorithms perform very good in the challenging task of traffic sign recognition. The participants achieved a very high performance of up to 98.98% correct recognition rate which is similar to human performance on this dataset.

2) : [6]Valentyn Sichkar, Sergey A. Kolyubin, "Effect of various dimension convolutional layer filters on traffic sign classification accuracy": This paper proposes a classification model for traffic sign detection together with carefully chosen evaluation metrics and baseline results. In their evaluation, they separate sign detection from classification, also measured the performance on relevant categories of signs to allow for benchmarking specialized solutions. Further, they use different filters of CNN using GTSRB dataset.

3) : [7]Towards Real-Time Traffic Sign Detection and Clas-sification by Yi Yang, Hengliang Luo, Huarong Xu and Fuchao Wu,2014, IEEE. This paper points to deal with real-time traffic sign recognition, i.e. localizing what type of traffic sign appears in which area of an input image at a fast processing time. To achieve this objective, a two-module framework (detection module and classification module) is proposed. In detection module, the input color image is transformed to probability maps by using color probability model. Then the road sign proposals are extracted by finding maximally stable extremal regions on these maps. Lastly , an SVM classifier which prepare with color HOG features is used to further filter out the false positives and classify the present proposals to their super classes. In classification module, they used CNN to classify the detected traffic signs to their subclasses within each super class.

4) : [8]"Traffic indication symbols recognition with shape context" by Kai Li, Weiyao Lan, Department of Automation Xiamen University, China, 2011, IEEE In this paper to detect the traffic sign, HIS color model followed by circle detection is used. The regions detected by color detection cannot be determined to

the exact sign region. In this method the edge of interested regions is traced to get their contours after morphologic operations. Then to find the target region, Hough circle transform is applied. The object have been detected and extracted after the previous two steps.The image is preprocessed to remove noise. To obtain a clear silhouette boundary of the traffic indication symbol Edge detection and segmentation are used specifically to the image. Shape context is based on the contour of the object.

III. MOTIVATION

Though driverless AI is advancing rapidly with the support of American giants like Google and Tesla; the technology is not yet mainstream. In the Indian scenario, however, technical problems persist. This exclusive Indian data provide much-needed impetus to the autonomous industry not only in India but across the world. For a densely populated country like India, leading automotive companies are looking to capture a major chunk of the market cap and would certainly use up resources such as this IDD dataset. "Autonomy will come step by step. We'll see semi-automatic systems, driving assistants, interactive systems and safety features that are AI enabled before that," observed Jawahar, who led the team behind this year-long prestigious project. But, in India, things don't look so smooth. Earlier this year, Nitin Gadkari, Minister of Road Transport and Highways, had blatantly opposed the arrival of driverless cars. On the contrary, the Traffic Amendment Bill encourages exploring new technologies to improve road transport in India. India is on track to become the world's third largest car manufacturer. With such high numbers, policymakers need to be flexible about the stance they take and imbibe solutions which might seem technologically advanced but will soon become a norm in the near future.

IV. PROPOSED SYSTEM

A. Classification

In this system, a CNN model was built with different dimensional filters like 3×3, 5×5, 9×9, 13×13, 15×15, 19×19, 23×23, 25×25 and 31×31. Based on that, the model with the best accuracy was further used in the detection model. In order to classify traffic signs we used the GTSRB dataset which consists of 50000 images of traffic signs which are basically divided into 43 different classes. This dataset was splitted into training and validation sets. The Indian Dataset is used as well in the test sets for building the model using Transfer Learning. The size of split sets are

xtest : (12630; 32; 32; 3)

ytest : (12630;)

yvalidation : (4410; 43)

xvalidation : (4410; 32; 32; 3)

xtrain : (86989; 32; 32; 3)

ytrain : (86989; 43)

labels: 43

All the images and their labels were stored into lists (data and labels).The list was then converted into numpy arrays for feeding to the model. The shape of data is (86989) which means that there are 86,989 images of size 32×32 pixels and the last 3 means the data contains colored images (RGB value). While importing images total classes related to that image is detected and every image has its own corresponding id which is stored in class number. Data related to images was stored in xtrain , xtest , xvalidation and corresponding id in each class is stored in ytrain , ytest , yvalidation. The CNN model has one convolutional layer with 32 filters, ReLU (Rectified Linear Unit) activation function, one downsampling layer with 2×2 maximum factor, and hidden affine layer with 500 neurons that is followed by the output layer with 43 neurons as number of classes. It uses mAP (mean Average Precision) metric to evaluate

accuracy every 1000 iterations during training. After splitting, the model was built based on the above filters and 9 such models were generated having different accuracy. Out of those filters model with 3x3 filters was the most accurate. Therefore we considered this trained model for further detection.

B. Detection

The detection model was trained on GTSRB that has 900 RGB images. The dataset was divided into sub-datasets for training and validation in proportion 85% and 15% respec-tively. The total amount of images for training and validation is 630 and 111. The number of excluded images without traffic signs is 159. Using DarkNet Framework[9], the network size (input). Before feeding to the network, the input images were resized to this spatial dimension by the framework without keeping aspect ratio. We selected only detection YOLO v3 layers that are 82, 94 and 106. Annotations of bounding boxes in GTSRB are in a single text file for all images. Originally, coordinates of bounding boxes are defined as top left corner and bottom right corner. As a result, the following coordinates were translated to YOLO: centre of bounding box in x, centre of bounding box in y, object width, and object height. Each image has an annotation file that contains the class number and bounding box coordinates. Each line represents a single bounding box.Images themselves were converted from PPM to JPG format in order to make possible training of the model in the Darknet framework.

V. EXPERIMENTAL RESULTS

After performing the proposed steps 9 such models were generated from generating filters with different dimensions, having different accuracy. Out of those filters, model with 3x3 filters was the most accurate with an accuracy of training accuracy = 0.98850, validation accuracy = 0.88594, testing ac-curacy = 0.86880 and classification time = 0.07510. Therefore

we considered this trained model for further detection.
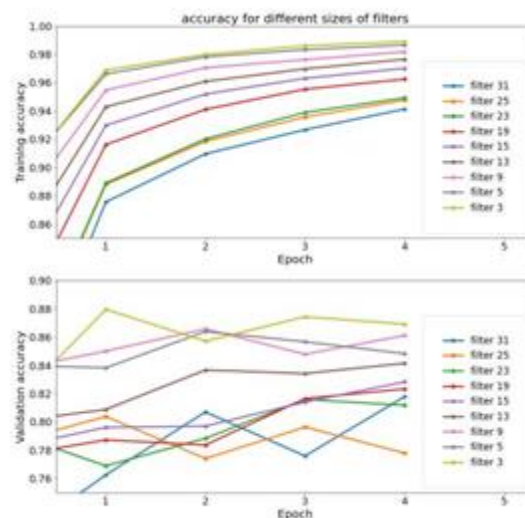
Graph showing accuracy of each filter

The above chart shows the accuracy for different sizes of filters against the training accuracy. We can see that, the model with 3x3 size is the most accurate after which lies model with 9x9 size.

Also this shows the chart for different sizes of filters against the validation accuracy. We can see that, the model with 3x3 size is the most accurate after which lies model with 9x9 size. So we conclude that the best trained and validated model is 3x3 and after that 9x9.

ClassId: 13
Label: Yield

Output of classifier

After calculating the best accurate model we moved on to detection part. Using DarkNet Framework, the

network size (input). Before feeding to the network, the input images were resized to this spatial dimension by the framework without keeping aspect ratio.We selected only detection YOLO v3 layers that are 82, 94 and 106.For detection, in order classify the detected image based on our trained model and naming the traffic sign image we used csv file which contained names of all 43 different classes of signs from the dataset

```
ClassId                SignName
0        0    Speed limit (20km/h)
1        1    Speed limit (30km/h)
2        2    Speed limit (50km/h)
3        3    Speed limit (60km/h)
4        4    Speed limit (70km/h)
```

Classes of the dataset

The framework resized the input images to this spatial dimension without holding the aspect ratio before feeding them to the network. We used YOLO v3 weights in our project and set the DNN backend to OpenCV and the goal to CPU. We took a video as input to our detection model and split the video into different no of frames according to the size of the video. In our case we took a video of Indian traffic signs of size 42 MB with duration of 1 minute and 12 seconds which was divided into 151 frames.

Total number of frames 151

Total amount of time 171.54610 seconds

FPS: 0.9

Input video frame

Output video frame

Each frame was evaluated and a video was generated which shows the input video with markings on each frame containing the predicted traffic sign classification.

VI. CONCLUSION

The classification of traffic signs using CNN with different filters helped improve the system by selecting the most accurate model with an accuracy of 87% and train the model using GTRSB dataset and test it with Indian Dataset. The detection model used YOLO and BLOB analysis to detect the traffic signs from the environment and classify the image according to the class it belongs to. Transfer Learning was used to classify and detect Indian Traffic Signs specifically.

VII. FUTURE SCOPE

Another direction for further research is to develop a real time traffic sign recognition system which captures a video by a camera mounted on the vehicle, detects and recognises the traffic signs in real time and gives the result to the driver within a sufficient time frame in order to take the right action.More research can be done to make an AI that can verify reports ,images and videos automatically without human intervention and make the entire process smooth and fast.

VIII. REFERENCES

[1]. German Traffic Sign Recognition Benchmark https://www.kaggle.com/meowmeowmeowmeo wmeow/gtsrb-german-traffic-sign

[2]. Intelligent Transpostation Systems ITS https://en.wikipedia.org/wiki/Intelligent_ transportation_system

[3]. Advanced Driver Asssistance System ADAS https://en.wikipedia.org/wiki/Advanced_ driver-assistance_systems

[4]. RGB HSI https://www.vocal.com/video/rgb-and-hsvhsihsl-color-space-conversion/

[5]. J. Stallkamp, M. Schlipsing, J. Salmen, and C. Igel, "The German traffic sign recognition benchmark: a multi-class classification competition,". in Proc. IEEE IJCNN, 2011, pp. 1453–1460.

[6]. Valentyn Sichkar, Sergey A. Kolyubin, ""Effect of various dimension convolutional layer filters on traffic sign classification accuracy".

[7]. Yi Yang, Hengliang Luo, Huarong Xu and Fuchao Wu, Towards Real-Time Traffic Sign Detection and Classification 2014, IEEE.

[8]. Kai Li, Weiyao Lan, "Traffic indication symbols recognition with shape context". Department of Automation Xiamen University, China, 2011, IEEE

[9]. DarkNet Framework https://pjreddie.com/darknet

Cite this article as :