Traffic Sign Recognition using Deep Learning by Neural Networks

V Kranthi Kumar Reddy
vkkreddy45@gmail.com

K Tirupathi Reddy
thirupathikota75@gmail.com

R Pranav Kumar Raju
rajupranav1905@gmail.com

MS.V.Surya,
Assistant Professor
Computer Science and Engineering,
SRM Institute of Science &
Technology, Chennai

Abstract*— The automatic road traffic sign detection and recognition is very crucial and important for an intelligent driver assistance system . Due to rate of increase of population growth automobiles have been increased. Also every individual is fond of using their own vehicles. In this tough situation it becomes difficult for the application of self driving cars. Many of the individuals do not follow traffic signs and also they do not obey the traffic rules. This might lead to a problem for self driving car and accidents. The project aims for the self driving cars to detect and recognize the traffic sign and act according to that in real world. The paper focuses on correctness and efficiency in recognition of traffic sign's. The entire action was recorded by a video camera attached to the vehicle through a series of images in a video format. A deep neural network algorithm is used to train the model data. The training sets is trained by neural networks with a multiple convolution layers. So, efficient model comes as an outcome by identifying the most critical traffic sign's which makes us to achieve the purpose of identifying the traffic signs in the real Scenario. For an ITS, Traffic sign detection is crucial.

***Keywords*—TSR(Traffic Sign Recognition),Deep neural Network,Convolution Neural Network (CNN),ITS(Intelligent Transportation System).**

I. INTRODUCTION

For an Intelligent Transportation System Smart vehicles are must. Smart vehicles reduce number of accidents happening around the world. Not only that Smart vehicles resembles the driver safety and also helps in the automatic driving and passengers safety. The camera used for traffic sign detection captures an image with a 1236*968 pixels in a series of images just like a video format.

For a safer journey to happen, one must follow the road traffic signs and obey the traffic rules. As a rapid increase in rate of population ,number of motor vehicles used by individuals also increasing. As a result of that traffic is increased. For application of smart vehicles, traffic sign detection and recognition is must. Traffic signs are of several types, Out of them three are important and mostly used. They are regulatory signs, warning signs (used to warn the driver about upcoming hazards, lane changes/merges etc.,)Guide signs (used to guide the driver regarding a distance such as distance signs and mile markers). Detection of traffic sign is the first step for autonomous vehicles. The image was detected by a camera. Some of the factors that involve during the detection of traffic sign was:

1)Fading of colors: Due to exposure of traffic sign to continuous sunlight may cause color fading which leads to correlation of input data. Also due to the pollution ,causing fading of color of traffic sign which causes a problem for detection.

2)Weather condition: In some cases of rainy situations, detection of road signs might become difficult.

The same type of traffic sign may have different consistency in color, in the appearance.Since the camera mounted on the vehicle will not be at a 90 deg angle sight to the traffic signs. The paper comes up with a solution that works in real time recognition of traffic signs which is consistent in detection of road signs and classifies them instantly. The outlets of the paper are as follows: The system uses a RGB image as an input source for detection. A simplified filters are used to strengthen the edges of the image, also the size, the shape. It also reduces the noise in the darker or more brighter areas. Region of interest are founded abnormally by the use of maximally stable external regions algorithm.

Artificial neural network is made up of multiple hidden layers. The preprocessed image was given as input to the convolutional layer of the artificial neural network.This layer helps for the classification of traffic signs. Training of neural network algorithm takes a lot of time. Since we are using a training data set of over 5000 traffic sign images. More amount of input data makes the algorithm to provide an efficient output. Thus, the system enables for an efficient and correct prediction of traffic sign in the real time.

II. RELATED WORKS

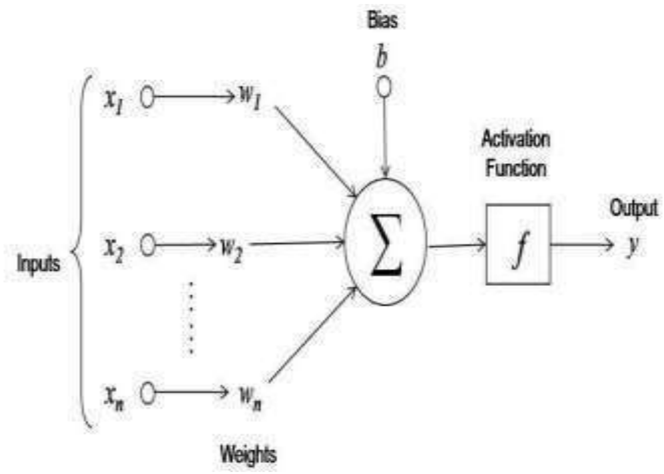
Mostly traffic sign recognition system involves detection and classification of images. Detection of images is performed by an onboard camera attached in front of vehicle. Due to the low light conditions, fading of color of road sign by exposure to sunlight and other some conditions causes a troubleshoot for the camera to get the input image. Basically the input image that the system gets in is RGB. So, for low level image processing this was overcome in [2] [3] by taking into account the single layer perception algorithm. The system has an accuracy of 93.25% in detection and recognition of traffic signs but the only thing is that it was HSV color segmentation process [Hue Saturation Value]. For some cases of conditions the color becomes too sensitive. Hue gets failed in this case. This becomes tough for some light colors to detect and classify the road sign. In case of neural networks algorithm, the system provides a precision rate of 95%. The paper [4][12] deals about the two camera system for obtaining the data continuously and simultaneously without any delay. The two cameras are communicated through the concept of OCC [Optional Camera Communication][6,7,13]. The system also uses computer vision[8][10] for detecting the type of images since the system uses neural networks it is stable for classifying the road signs. One of the drawback was found to be that in some of cases of multiple set of different size images it becomes hard for the two devices[cameras]. The paper [5] uses the same thing when compared to [4,10,11] but has only mono camera system supporting a high dynamic range mode for obtaining the image in poor light conditions.The average precision rate of the system found to be 88%[7]. The result with a faster version rate is found to be less accurate in terms of data accuracy. In case of machine learning algorithm applied for TSR, the paper uses color segmentation algorithms for shape classification, edge recognition, etc. The main advantage of the system is found to be the response time of a 64.2 m/s. Also,some certain achievement is less for outdoor conditions in detection of certain signs[6][7].Also delay in sending the recognized data to the system. The convolutional neural network is found to be the best algorithm for a higher precision rate[1][7]. The experimental results show that the decision rate is above 99%. Also the response time is about 51.5m/s. The system uses a 640*480 in the format of a video sequence. The system is proposed based on a faster rate.The one of the applications implemented in the advanced driver assistance systems ,is the use of sensors which hazards the driver/system to improve safety by the potential hazards[8]-[10]. The paper [12] describes the recognition of traffic signs based on the method of RCNN.Faster RCNN algorithm provides an accuracy rate of 96% which is almost similar to the CNN .

III. Traffic Sign Detection and Classification

For detection and classification of traffic sign images the paper uses a neural network structure for training and testing the model.The network structure is similar to the neurons present in the brain of a human . The system uses the concept of artificial neuron which works similar to the human brain.Each neuron recieves a certain set of inputs and pass it to the next layer through a weighted sum to form a bias function (summation of all inputs*weights).

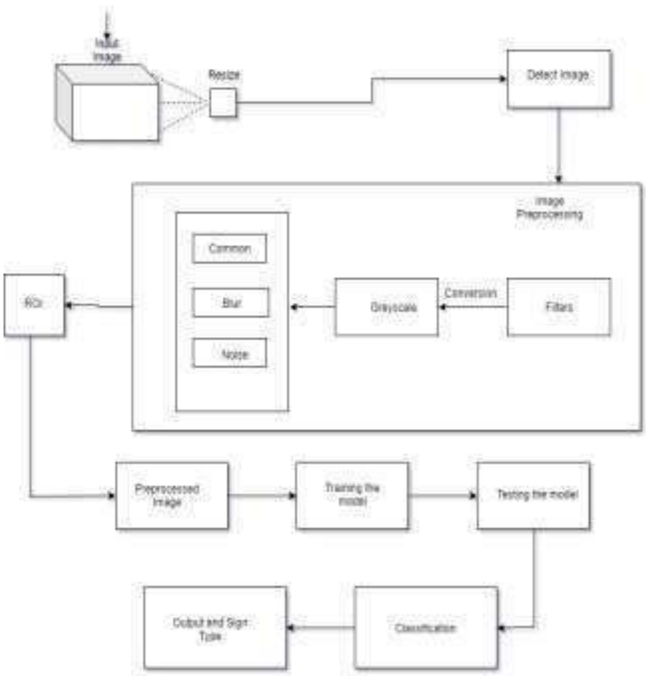
The images related to traffic sign are been taken from a dataset named German Traffic sign Recognition Benchmark (GTSRB) for training, testing and deploying the model.

Bias In Neural Networks



The model takes set of input images for processing and training the model.
The bias function in neural networks takes the set of input values from the matrix format of the image and assigns certain weights where the summation of all weights and inputs form a bias function .The bias function is later converted in to activation function for linear model of the outputs on the graph.
The mathematical expression can be given as
=>Y=F(x)= Σxiwi
=>Output=sum (weights * inputs) + Bias function

The architecture of the model looks like:



Proposed Architecture model

In Proposed architecture model the system takes a set of input images which are adjusted,resized in terms of shape and configuration.Images are converted in to greyscale format for the computer vision called preprocessing stage of the image.
The grayscaled image is converted in to pixel values ranging from 0 to 255.Pixel values for a greyscaled image are put in a matrix form of representaion.The values are arranged based upon the regions of interest.The process was done for all the images which are passed as input.Now,the entire set of preprocessed images are passed in to the convolutional layer of the network model for training and testing the model for output and classification of images.

A. DATASET

The deep neural network has set a benchmark in recognition of images and voices. The Current paper uses a convolutional layer of neural network architecture to detect and classify the road traffic signs .

GTSRB Dataset of traffic signs

The German Traffic sign Benchmarks dataset used for the training of the model.The GTSRB dataset was splitted in to training set and testing set of which training set has a 4167 images in a 32*32 format for training of model and for testing and validation of the trained model,the paper uses a 1994 images to get a desired output. The images are of 1236*968 pixels with a measured ground truth value.

The training and testing datasets may have different set of images .One with an uneven color while the other with sign tilt or distortion etc., The problem statements of the above kind are come across the preprocessing stat of the image

B. IMAGE PREPROCESSING

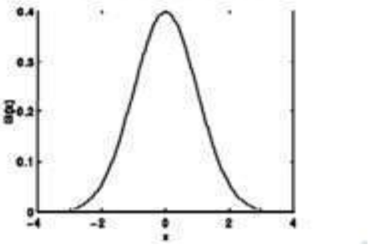
The training set provides the input image for the model. For training the neural network, the input image must be of a normalized form with a pixel value ranging from 0 to 255. Input images are adjusted and reshaped by sharpening the edges and by removing any disturbance further if any such as high saturation .For filtering the images the system uses Gaussian Filter algorithm which helps in reduction of noise in blur areas of the image.

$$G(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}}$$

Where σ is the standard deviation of the distribution. The distribution is assumed to have a mean of 0.

Shown graphically, we see the familiar bell shaped Gaussian distribution.

Gaussian distribution with mean 0 and σ = 1



Gaussian function in 1D

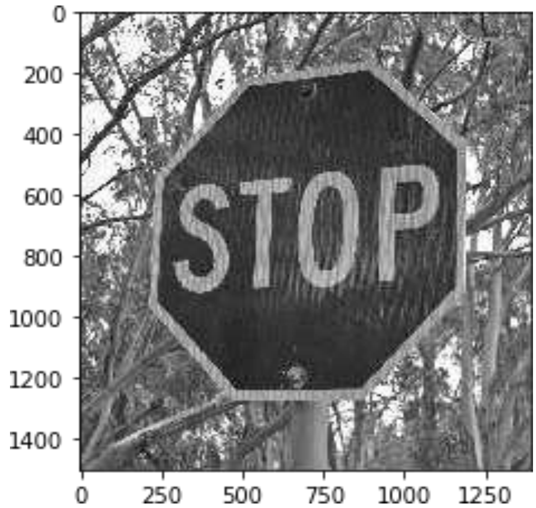
If any white balance was present it was removed using gray world algorithm. The algorithm converts a more white balanced factor of the images to a normal form. Simply it converts an RGB image to a grayscale image and reduces noise and blur from the obtained gray scale from the obtained grayscale images. The algorithm also adjusts hue, contrast and saturation of the images.

Greyscaled image is assigned with a certain pixel vales in a matrix format ranging from 0 to 255. The pixel values assigned for an image are represented in matrix format.

Preprocessed output can be shown as

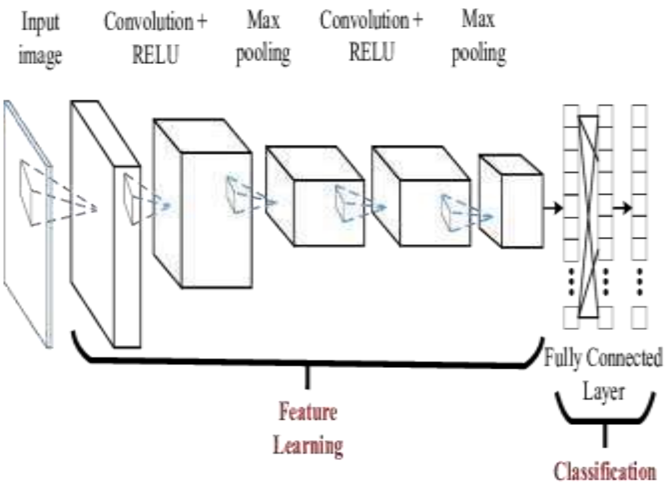
RGB Image

GREYSCALED Image



C. NETWORK STRUCTURE

The paper uses a convolution neural network structure. The convolution layer of the network structure extract features by choosing region of interest from an input image.Also,extracts the image features by forming small squares on the input image of the matrix format, Basically, it is a mathematical operation that performs the task by considering two inputs, one is image matrix and other is filter or kernel.



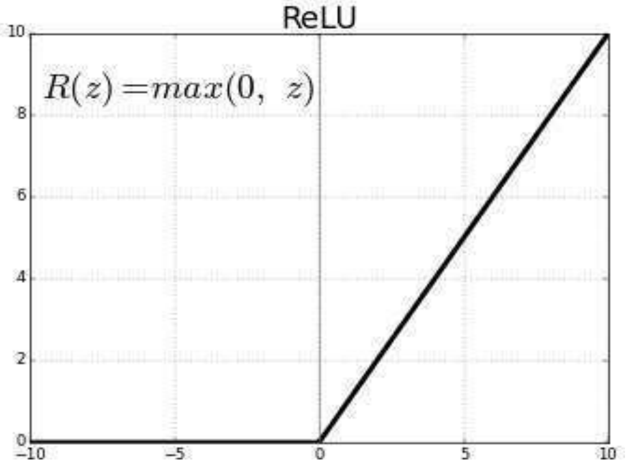
Convolutional Network Layer for Future Learning and Image Classification

=>Image matrix of dimension in terms of volume(h*w*d).
=>Kernal of filter fh *fs *fd.
=>Output (h-fh+1)*(w-fw+1)*1.
The next layer of the structure are hidden layer and the output layer. The paper uses 3 convolution layers with 2*2 max pooling. Max pooling is a technique used to reduce the dimensions of an image by considering the maximum amount of pixel value in the grid. Max pooling helps to reduce over fitting and makes the model to look more generic.A flatten layer was used to flatten the output of convolution layer before it is sent to the output layer.finally the sotmax layer for probablity analysis of the output.

D. Feature extraction and training of model:

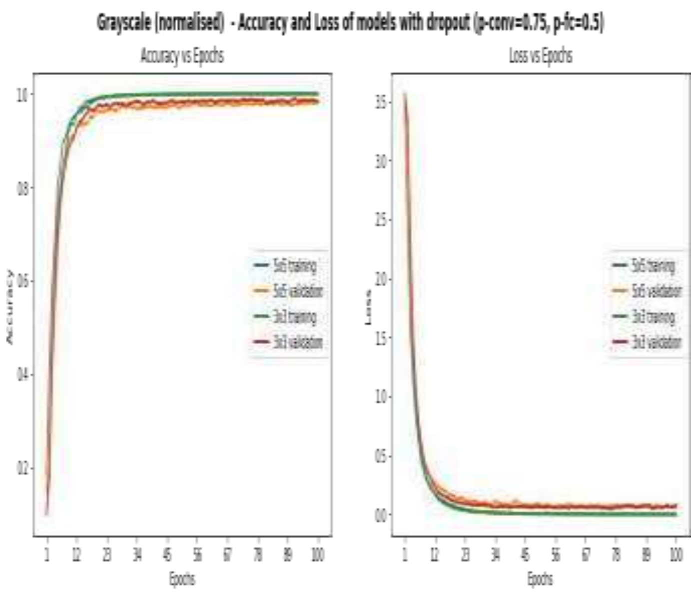
A 3 convolution layers with 2*2 max pooling was used for training the model. Max pooling is a technique used to reduce the dimensions of an image by considering the maximum amount of pixel value in the grid. Max pooling helps to reduce over fitting and makes the model to look more generic. A flattened layer is connected in between the two convolution layers for more accuracy of the data. The input given for fully connected layer was two dimensional layer of matrix format and the output produced is four dimensional. By the end of the fully connected layers are got transferred to a soft max layer. The activation function RELU was used to find the gradient descent of the model. RELU stands for rectified linear unit responsible for linear representation of values on the graph. Mathematically, expressed as

Y=Max(0,X).



Activation Function RELU

In order to improve the model reliability and correctness, we turned to dropout, which is a form of regularisation in which weights are kept standard with probablity values between 0 and 1 The dropout function is used for overfitting the data in the normalization stage of the neural network structure.



The entire network structure model starting from the input layer and ending with the output layer is implemented through keras deep learning software along with python as the base programming language for the neural network structure. Kera's a deep learning software developed by Google which supports fast prototyping of the data.

The coding part is done by setting three convolutional layers followed by a max pooling layer, followed by an activation Relu layer and finally attached to the output layer with an SoftMax function layer for probability predictions of the output data(image). The entire action was done in Jupyter notebook server.

IV. *Experimental verification*

Several adjustments were made for training the data. The paper uses a batch size of 32 since the images were around from 2000-3000. The training set uses epoch of 40. After iteration of 20,000 times, the precision accuracy was found to be 96%. Batch size of 32 and epoch of 20 /50 are used for testing set. The model is capable of recognizing the images with an accuracy of 98% for an epoch value of 50. For increase in epoch Value can produce accuracy rate of 98.1%, but time complexity of algorithm increases.

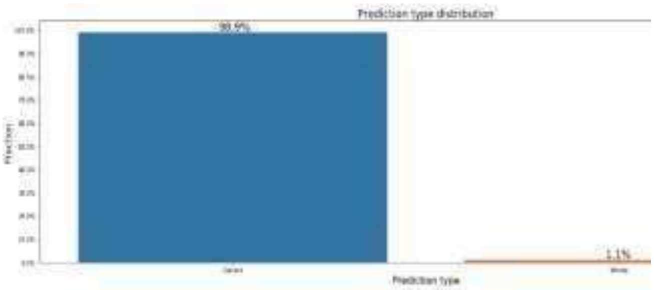
Testing of the Trained Model

V. *Experimental improvement*

Addition of more data to the training data set would help in accurate decision for traffic sign, for example some traffic signs may have high color contrast, dimming of colors due to sunshade can result an error in detection or takes a longer time for detection.

VI. *Result*

Since, after training the model over 50,000 iterations with a epoch of 100 the prediction accuracy was found to be 98.9%
which is a satisfactory value for recognition of traffic signs.



Probability of Prediction rate

VII. *Conclusion*

Focusing on the need of driver's safety and safe journey to happen. The paper concludes a model which aims in recognition of road traffic signs in the real world supporting for the development of the driverless cars to reduce the no.of accidents happening around the world. The network model was developed based on artificial neural networks. Comparisons of each stage by training and validating the dataset makes our model look good fit and helps in the output prediction. Thus, by concluding that real time accuracy of 98% can be achieved for any situation in the real-world scenario.

References

[1] Traffic Sign Recognition Algorithm Based on Deep Convolutional Neural Network. Authors:Wang Cong, Ma Weixin and Xiong Changzhen,published in 2016.

[2] Traffic Sign Detection and Recognition by SVM.Authors: V.Prabhu,Abdu Rahiman V. Kiran C.G and Rajeev K. published in 2009.

[3] Detection and Recognition of Traffic Signs based on RGB to RED Conversion Authors:Mr.MohithBhairav Mahatme,Mrs.Sonia Kuwelkar,copyright and Pulblished on May ,2017.

[4] Simultaneous Traffic Sign Detection and Real time communication using Dual camera in ITS.MohdKhalid Hasan, Md. Shahjahal,Mostafa Zaman Chowdary, Nam Tuan Le and Yeong Min Jang.Dept of Electronics Engineering ,Kookmin University ,Seoul,South Korea. 2019.

[5] Real time Large Scale Traffic Detection .Aleksej Avramovic,Domen Tabernik, and Danijel Skocaj.April 2018.

[6] Traffic Sign Recognition for Computer Vision Project Based Learning.Authors:David Geronimo, Joan serrat, AntonioM Lopez and Raman Baldrich.cc2013.

[7] Vision-Based Traffic Sign Detection and Recognition Systems: Current Trends and Challenges. Safat B. Wali , Majid A. Abdullah , Mahammad A. Hannan,Aini Hussain, Salina A,Samad, Pin J. Ker and Muhamad Bin Manson.

[8] Image segmentation and shape analysis for road-sign Detection,J. Khan,Adhami.K and S. Bhuiyan, IEEE Transactions, Mar. 2011.

[9] Road-sign detection and recognition based on machine learning by support vector machines, P.Gil-Jimenez, H.Gomez Moreno, S.Maldonado Bascon,S.Lafuente-Arroyoand and F.Lopez-Ferreras,IEEE Trans. Intell. Transport., Jun. 2007.

[10] Traffic sign recognition and analysis for intelligence vehicles based on Image Vision, J. Armingol, M. Mata and A. de la Escalera,vol.21, pp. 247–258, 2003.

[11] Detection and classification of speed limit traffic signs, M.Mostakim, H.Fleyeh, Biswas.R, Proc. Computer Applications and Information Systems(WCCAIS), Jan. 2014.

[12] Faster R-CNN: Towards Real Time Object Detection with Region Proposal Networks, R.Girshick, J.Sun,S.Ren and K. Hen.2015.

[13] Performance analysis and improvement of optical camera communincation,Mr K.Hasan Applied sources ,Dec 2018.

[14] Encoding Rich Frequencies for Classification of Stroke Patients EEG Signals. Authors: Fawaz Sammani, Kok Swee Sim , and Shing Chiang Tan.