iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** II **Month of publication:** February 2022

DOI: https://doi.org/10.22214/ijraset.2022.40224

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue II Feb 2022- Available at www.ijraset.com

Traffic Sign Classification Using CNN

Preeti Bailke[1], Kunjal Agrawal[2]
[1, 2]*Department of Information Technology, Vishwakarma Institute of Technology, Pune, India*

Abstract: ***Traffic signs displayed on the roads play an important role in our lives while driving. They supply critical information, for the road users. This successively requires them to regulate their driving behaviour and ensure that they strictly follow the road regulations currently enforced without causing any trouble to other drivers and pedestrians. Traffic Sign Classification is employed to detect and classify traffic signs to inform and warn a driver beforehand to avoid violation of rules. There are certain disadvantages of the existing systems, used for classification, like incorrect predictions, hardware cost and maintenance, which are to a great extent resolved by the proposed system. The proposed approach implements a traffic signs classification algorithm employing a convolutional neural network. Also, it consists of the feature of web cam detection of the traffic sign. This will help the driver to observe the sign close to his / her eyes on the display screen and thus save his/her time in manually checking the traffic sign each time.***

Keywords: Convolutional Neural Network, Traffic Sign Classification, Real time Traffic Sign Recognition.

I. INTRODUCTION

Machine learning algorithms have gained importance nowadays. Spam filtering, speech understanding, face recognition, road sign detection are only a few examples where machine learning is deployed. In traffic zones, Traffic Sign Recognition and classification can be used to automatically identify traffic signs. This is done automatically by the system as the traffic sign is detected and the sign name is displayed. So, even if any sign is missed by the driver or has any lapse in concentration, it will be detected. This helps to accordingly warn the drivers and forbid certain actions like over speeding. It also disburdens the driver and hence, increases his/her comfort. Thus, ensuring and keeping a check on the traffic signs and accordingly following them. Traffic signs, indeed, provide us a multitude of information and guide us accordingly so that we can move safely. Traffic Sign Classification is very useful in Automatic Driver Assistance Systems.

A convolutional neural network is a class of deep learning networks, used to examine and check visual imagery. It is used to train the image classification and recognition model because of its high accuracy and precision.

II. LITERATURE REVIEW

In today's world, identification of traffic signs has become an important aspect of our lives. Looking at the increasing traffic, to ensure safety of all and for automatic driving in the future, traffic sign classification is utmost necessary. Considerable research has been done around recognition of traffic and road signs. In 1987, the first research on the topic "Traffic Sign Recognition" was done by Akatsuka and Imai , where they tried to build a fundamental system that could recognize traffic signs and alert the drivers and ensure his/her safety. But this was used to provide the automatic recognition for only some specific traffic signs.

Traffic sign recognition initially appeared in the form of only speed limit recognition in 2008. These symbols could only detect the circular speed limit signs. On the other hand, later, systems were designed that performed detection on overtaking signs. This technology was available in the Volkswagen Phaeton and in the 2012 in Volvo S80, V70 and many more. But the major drawback of these systems was that they could not detect the city limit signs as they were mostly in the form of direction signs. But nowadays, such systems are expected to be present in the future cars to help drivers while driving.

In [1], the authors used the colour processing system to reduce the effect of brightness and shadow on the images. This was the very first research done on this topic by the authors, Akatsuka and Imai. In [2], the authors have done a survey on traffic sign detection and recognition, where HOG (Histogram of Oriented Gradients) is used for classification purpose. In [3], a complete study of different traffic sign recognition algorithms has been done, where the highest accuracy (99.46%) was obtained by MCDNN (Multi-column Deep Neural Network). In [4], the authors have developed a model where they are converting the images to gray scale first and then filter those images using simplified Gabor wavelets. The gabor filters are used to extract features. These wavelets are important as they help to minimize the product of its standard deviation in both the time and frequency mapping. The authors extracted the regions of interest for recognition purpose and classified the signs using "Support Vector Machine" (SVM).

In [5], the authors have extracted the regions of interest in the detection stage and further examined the shapes of such regions. Here, in the classification system, they took the regions of interest and classified them into different classes.

©IJRASET: All Rights are Reserved

In [6], the authors have created a module which consists of numerous convolutions. They combined the 1*3 kernel and 3*1 kernel and finally linked it with the 1*1 kernel to attain the 3*3 kernel. This was used to extract more features and thus reduce the number of parameters. In [7], the author has reviewed the traffic sign detection methods and divided them into 3 types of methods: color, shape and learning based methods. In [8], the author used the number of peaks algorithm to detect and recognize circular shaped traffic signs.

In [9], the authors have tried creating a classification model using the Enhanced LeNet-5 architecture, which consists of two consecutive convolution layers (before the MaxPooling layer) to extract high level features from the image. Also, they have used the data augmentation technique to make the dataset stable. In [10], the authors have used the technique of colour segmentation and the RGB based detection which is used to identify the traffic signs on the road. The optimizer used was "Stochastic Gradient Descent" with Nesterov Momentum. The text to speech system was implemented to alert the driver about the traffic sign. Also, they utilized the GPU (graphical processing unit) system, as part of hardware. In [11], the authors have tried generating a dataset for the Arabic Road signs and thus develop a CNN model for Arabic sign recognition.

III. METHODOLOGY

A. Traffic Sign Dataset

Before moving on to detection or classification, the most important part is the availability of a generalized dataset. A prediction model is trained using this dataset and predictions are done for test dataset. Table I below shows sample datasets:

TABLE I Dataset Information

Dataset	Information
GTSRB	Total traffic sign images more than 50,000 and classes = 43
GTSDB	Total traffic sign images = 900
BTSCB	Total traffic sign images = 10,000 and classes = 62
BTSDB	Total traffic sign images = 7000

Among these, the most common dataset is the GTSRB (German Traffic Sign Recognition Benchmark) dataset. The reason for its popularity is:

1) It consists of large number of images
2) The traffic signs are of different variety, background, and colour variation which in turn will help the model to perform accurately.

As the GTSRB dataset can be used for both detection as well as classification, the proposed system makes use of the same. The dataset is further split into training, testing and validation dataset. The training dataset is the one which is used to train the model. The validation dataset, in general, is used to evaluate the model and update the hyper parameters. Hyper parameters are used to control the learning process and improve the accuracy, for example, number of epochs, the choice of activation function. The test dataset is only used once the model is trained. It is used to check whether the model can make correct predictions or not.

Further, histogram graphs are plotted (as shown in fig. 1), to show the number of images in each class, for the training, testing and validation data sets respectively, where the X label denotes the "Class ID", and the Y label denotes the number of images. Plotting the graph helps to visualize the dataset.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue II Feb 2022- Available at www.ijraset.com

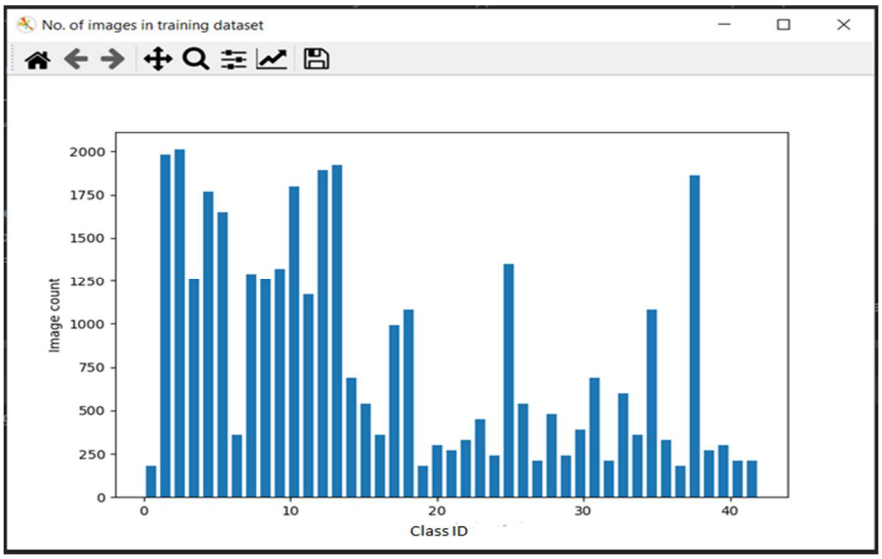


Fig.1: Class ID vs No. of images graph

B. Flowchart

Structure of the model:

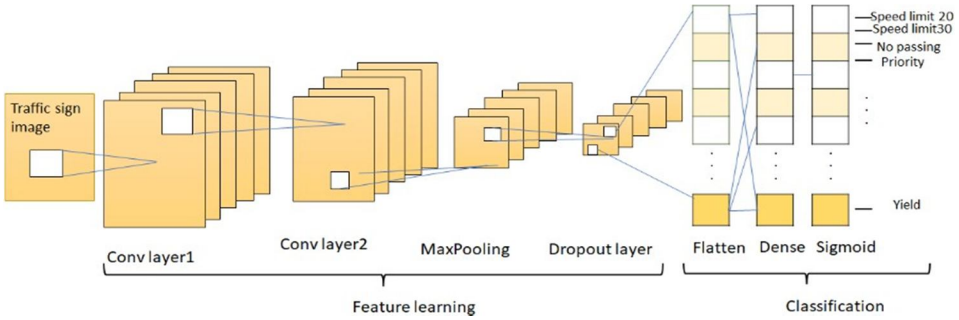


Fig.2: Model Architecture

Initially, the CNN model architecture is built (as seen in the fig. 2). The following steps are followed :

1) Sequentially add the layers in the order: two convolutional layers, one pooling layer, dropout layer, flattening layer, dense layer, again a dropout layer and finally the dense layer.

2) In the convolutional layer, number of filters is specified. It performs the convolution operation on the original image and generates a feature map.

3) The ReLU performs the maximum function to convert the negative values to zero without changing the positive ones and generate a rectified feature map. The Pooling layer takes the rectified feature map and performs a down-sampling operation (like Max Pooling or average pooling) and thus reduces the dimensionality of the image.

4) The flattening layer is used to convert the input feature map to a 1-dimensional array.

5) The dropout layer is used to avoid over fitting by setting some of the input neurons to 0 during the training process. The dense layer, on the other hand, feeds all the outputs from the preceding layer to all its neurons and perform the matrix- vector multiplication (the row vector of the output from the preceding layer should be equal to the column vector of the dense layer), to generate a m-dimensional vector.

6) After addition of the layers, the model is to be compiled (final step in the creation of model to define the loss function and apply optimization techniques) and assign the loss function as “sparse_categorical_crossentropy” and use the “Adam optimizer”. The reason for specifying this loss function is that the proposed system is a multiclass classification problem, where multiple classes are considered but one image belongs to exactly one class.

7) Next, the model is trained using the training dataset, by passing the pre-processed images from the training dataset.

8) Finally, the predictions on the test data are done using the trained model and the traffic sign name along with the class Id is shown as an output.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

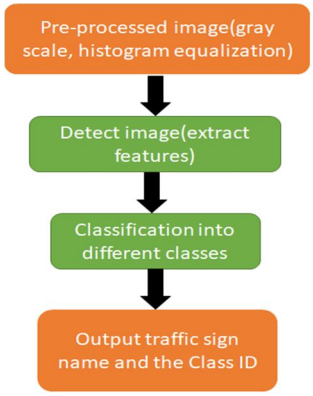


Fig. 3: Basic flow of the CNN model

C. Design/Method

Traffic Sign Classification is one of those rare topics of discussion. Most of the existing systems focus on detection only. Detection is mainly the extraction of features and find out the important coordinates in the image. Classification is the categorization of image into different classes.

The most common dataset used for the purpose is GTSRB which consists of 43 classes. In the proposed system, a prediction model is trained using this dataset. It performs best for image classification. Lately, Convolutional Neural Network has been adopted in object recognition for its high accuracy and less computational cost.

In the proposed system, the primary focus is towards the traffic sign classification which also prints the traffic sign name once the detection of the image is done. There is a csv file which consists of the pairs of traffic sign name and the class ID. This file helps to load the labeled data.

D. Gray Scale

Converting the RGB data set into gray scale is one of the important steps before classification using CNN (as shown in the fig. 3). This has several advantages like:

1) Images after converting to gray scale, help the neural network to process them easily as the unwanted biases are removed.
2) Gray scaling the images helps to reduce the number of computations, as the number of channels will get reduced after conversion.
3) This in turn helps to improve the model accuracy.

Before gray scaling, image data shape (of training data set) is: (34799, 32, 32, 3). This means that the images were of 32x32 size and colored in rgb format (3 channels).

After gray scaling, image data shape (of training data set) becomes: (34799, 32, 32, 1). After gray scaling the size of the image remains the same (32x32) but the number of channels is reduced to 1.

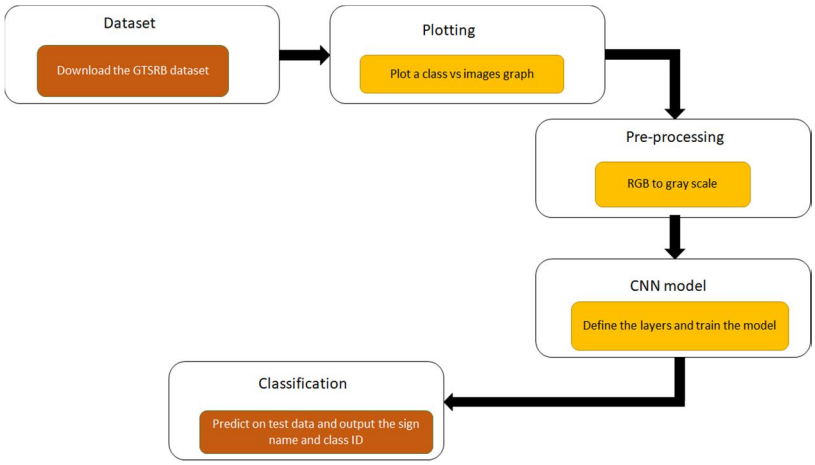


Fig.4: Flowchart of the proposed system

The flow of the proposed system is mentioned in the fig. 4. The proposed system consists of different functions corresponding to each operation.

1) Building of the model: This primarily focuses on converting the images to gray scale, normalizing the images (normalization is done to accelerate the training process and improve the model performance), histogram equalization (to improve image contrast), addition of the layers to the model, train the model, get predictions on the test data set, and finally show some sample images with their traffic sign name and class Id as the output. The train, test and validation split percentage is 65%, 25% and 10% respectively for the proposed system.

2) One of the main functionalities which are implemented in this work, is prediction of unknown images. Here, a small dataset was generated gathering images from different sources. This was the most crucial part as this dataset includes some different images with different colour and structure. Although there are several existing datasets available, a small dataset (consisting of 13 images) is built. The dataset includes some speed limit symbols, yield sign, caution signs (like stop and no entry), informatory signs (like pedestrians, ahead only, no passing, roundabout mandatory, and right-of-way at the next intersection). Extracting features from these images is not easy for the model. The reasons being, these images are enlarged, having different background colours and reduced clarity. Despite all the issues, the model successfully predicted around 9 images out of 13. Only the images which are curvilinear or in circular format, are not predicted accurately. The model predicts the closest traffic sign name found for such images.

The activation function used is “ReLU”. ReLU is one of those non-linear activation functions that is used in multi-layer neural networks. The reason being, ReLU applies a function f(x) = max (0, x) to all the values given as input. The ReLU layer just changes the negative values to 0 and keeps the positive values as it is. ReLU has become the default activation function to be used in the hidden layers of the neural network. The ReLU function is simple and less computationally expensive since there are no complex mathematical calculations involved and thus makes the model learn and train faster.

The model successfully predicts majority of the images with 93% accuracy as is shown in fig. 5. The only images which are not accurately predicted are the ones including circular shapes in it. This can be resolved by augmenting the images and using one hot encoding technique.

IV. RESULTS AND DISCUSSIONS

Traffic sign classification is the process of automatically recognizing traffic signs (like speed limit, yield, and caution signs, etc.) and accordingly classifies them as to which class they belong to. The project has two main functionalities: Prediction on the newly generated dataset (fig. 6) and live web cam traffic sign detection.

Accuracy is the ratio of number of correct predictions to the number of total predictions (eqn. 1). Accuracy = (The Number of Correct Predictions) / (Total predictions).

A. Equation 1: Formula for accuracy

The accuracy achieved on the test dataset is 93%. The accuracy on the GTSRB dataset and the built dataset are shown in Table II.

Table II Accuracy Statistics

Sr No.	Dataset used	Accuracy
1.	GTSRB	93%
2.	Generated dataset	69%

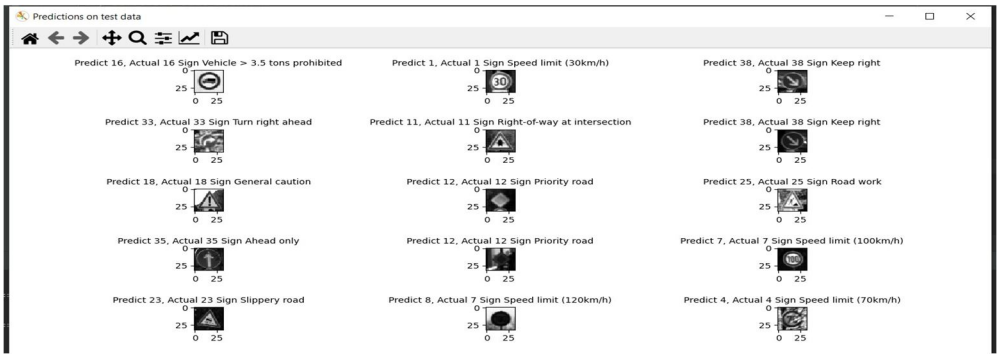


Fig.5: Predictions on Test dataset

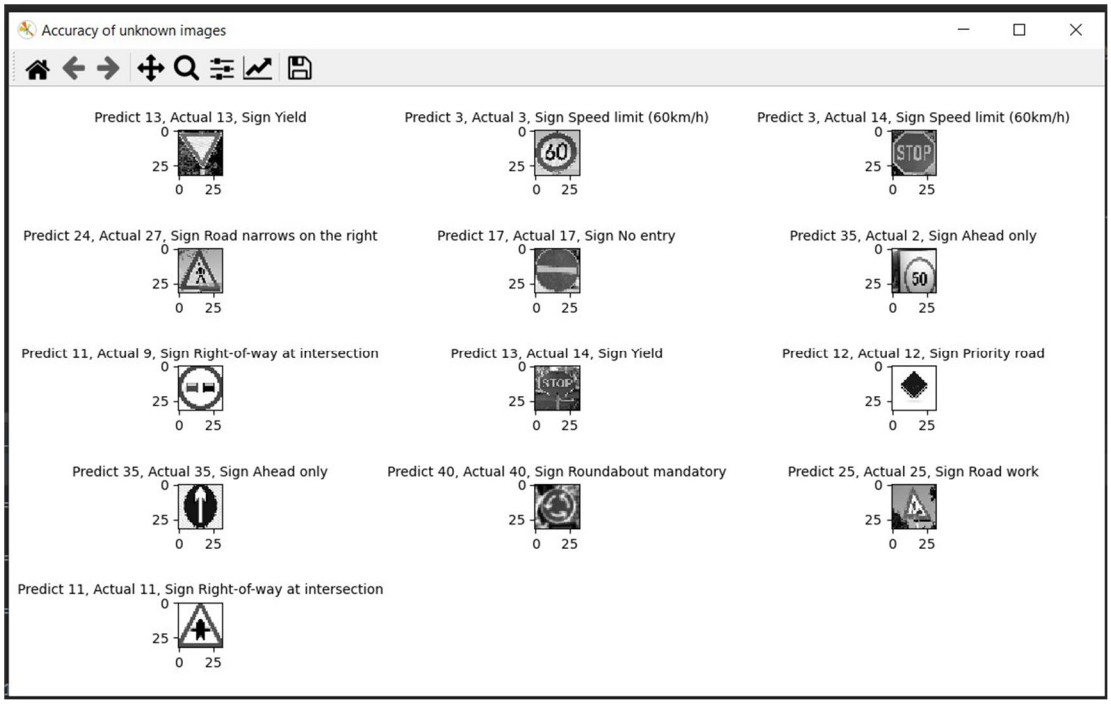


Fig.6: Predictions on Generated dataset

The Training accuracy and Validation accuracy graph is as follows:

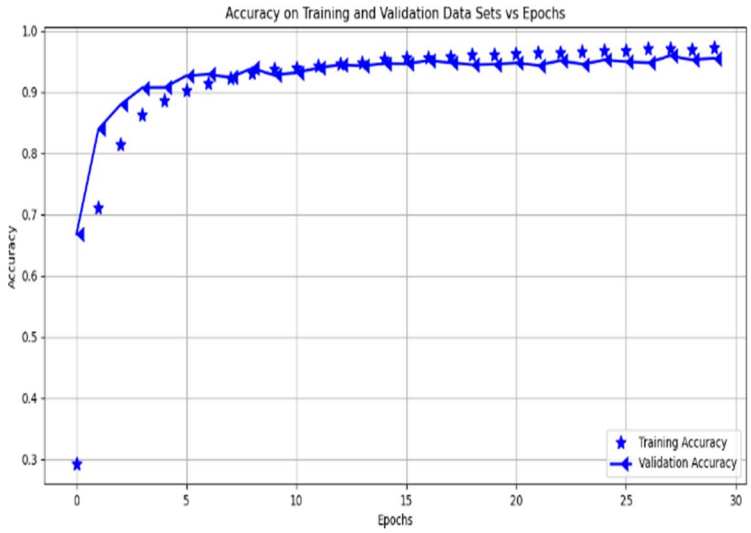


Fig.7: Accuracy vs Epoch graph

In fig. 7, the training accuracy is the accuracy generated from the predictions on the training dataset. Similarly, validation accuracy depicts the one using the validation dataset. The number of epochs is a hyper parameter that defines the number times that the learning algorithm will work through the entire training dataset. It is the number of iterations the training dataset will go through each time during training of the model. As seen from fig. 7, at around 30 epochs, the training and validation accuracy match and is a straight line. Here, the training and validation accuracy are maximum. If these lines start separating consistently, then we should stop the training process at an earlier epoch by visualizing the graph. This shows that 30 epochs are enough for the model to extract features.

The loss on training and validation dataset vs epochs graph is shown in fig. 8:

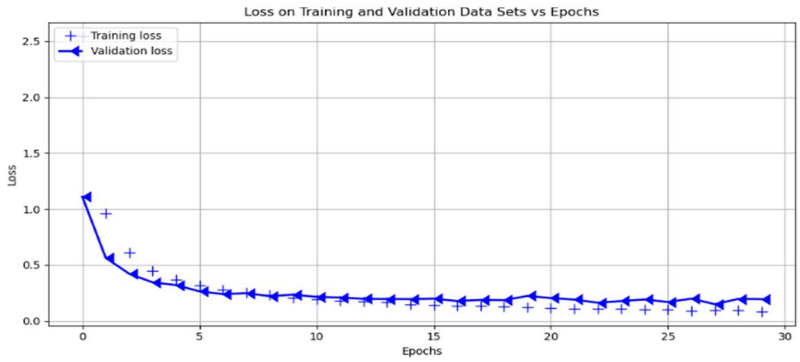


Fig.8: Loss vs Epoch graph

Table III
Loss Statistics After 30 Epochs

Sr No.	Dataset	Loss
1.	Training Loss	0.0929
2.	Validation Loss	0.1648

The loss statistics are shown in the Table III . Fig. 8 shows that the loss is minimum (nearly 0.2) at around 30 epochs. The mean square loss is calculated as the average of the squared differences between the predicted and actual values. In the Cross- Entropy loss, each predicted value is compared to the actual class output value and a score is calculated that penalizes the probability based on the distance from the expected value. Cross Entropy calculates the difference between the actual and the predicted probabilities. The formula is mentioned in the eqn. 2:

Cross Entropy Loss = $- (y_i \log(y'_i) + (1 - y_i) \log(1 - y'_i))$

B. Equation 2: Mathematical formula

where,

i is the i^{th} training example in the data set,

y_i - ground truth label for i^{th} training example, y'_i - prediction for i^{th} training example.

The traffic sign when shown from the web cam, automatically classifies the symbol and shows the corresponding class ID and the sign name.

Below are some of the images which are classified through the live web cam capture:

1) When the yield sign is shown to the web cam, it gets classifies accurately as the “Yield” sign along with the class ID. As seen in the fig. 9, even though the image is not completely visible, the model classified it accurately, saying it is the “Yield” symbol along with the class ID. Similarly, when the No Passing sign is shown to the web cam, the model identifies it and classifies the traffic sign name as “No Passing” and the corresponding class ID is displayed as 9 (as seen in the fig. 10).



Fig. 9: Yield sign detected through the web cam



Fig. 10: No Passing sign detected through web cam

2) When the General Caution symbol is shown to the web cam, the model identifies it and classifies the traffic sign name as "General caution" and the corresponding class ID is displayed as 18 (as seen in the fig. 11). When the No Entry symbol is shown to the web cam, the model identifies it and classifies the traffic sign name as "No entry" and the corresponding class ID is displayed as 17 (as seen in the fig. 12)



Fig. 11: General Caution sign detected through web cam



Fig. 12: No Entry sign detected through web cam

The predictions are done in a very less time (near real-time) which is a boon for the drivers. Also, the hardware used in this project for the model creation and classification purpose is very minimal as compared to other existing systems. Thus, the hardware cost along with the maintenance is drastically reduced.

V. LIMITATIONS

Although, there are many advantages of traffic sign classification, there are certain difficulties as well. It may happen that the traffic sign is hidden behind the trees or any board at the road side which may cause the inaccurate detection and classification of traffic sign. Sometimes it may happen that the vehicle went so fast, that it did not detect the traffic sign. This may be dangerous and can lead to accidents. There is a need for further research to deal with these issues

VI. FUTURE SCOPE

Traffic Signs are useful to all the individuals who are driving a vehicle on the road. Traffic Signs guide the drivers for following all the traffic rules and avoid any disruption to the pedestrians. The environmental constraints including lighting, shadow , distance (sign is quite far), air pollution, weather conditions in addition to motion blur, and vehicle vibration which are common in any real time system may affect the detection and thus the classification. Hence, there is a need for further research and advancements to deal with these issues. Also, there are certain traffic signs that may not be predicted accurately. For this, augmentation and one hot encoding techniques can be used. Augmentation involves shifting of the image, zoom in and rotate the images (if required).

This system helps the driver to observe the sign close to his / her eyes on the screen. This saves the time and efforts in manually checking whether any traffic sign board is there, identifying what type of sign it is and act accordingly. Traffic Sign Classification, thus, has a wide application in building smarter cars like automatic driving cars, where the system automatically detects, recognizes a traffic sign and displays it.

VII. CONCLUSION

The proposed system is simple and does the classification quite accurately on the GTSRB dataset as well as the newly generated one (consisting of truly existing images of all type), and finally the model can successfully capture images and predict them accurately even if the background of the image is not much clear.

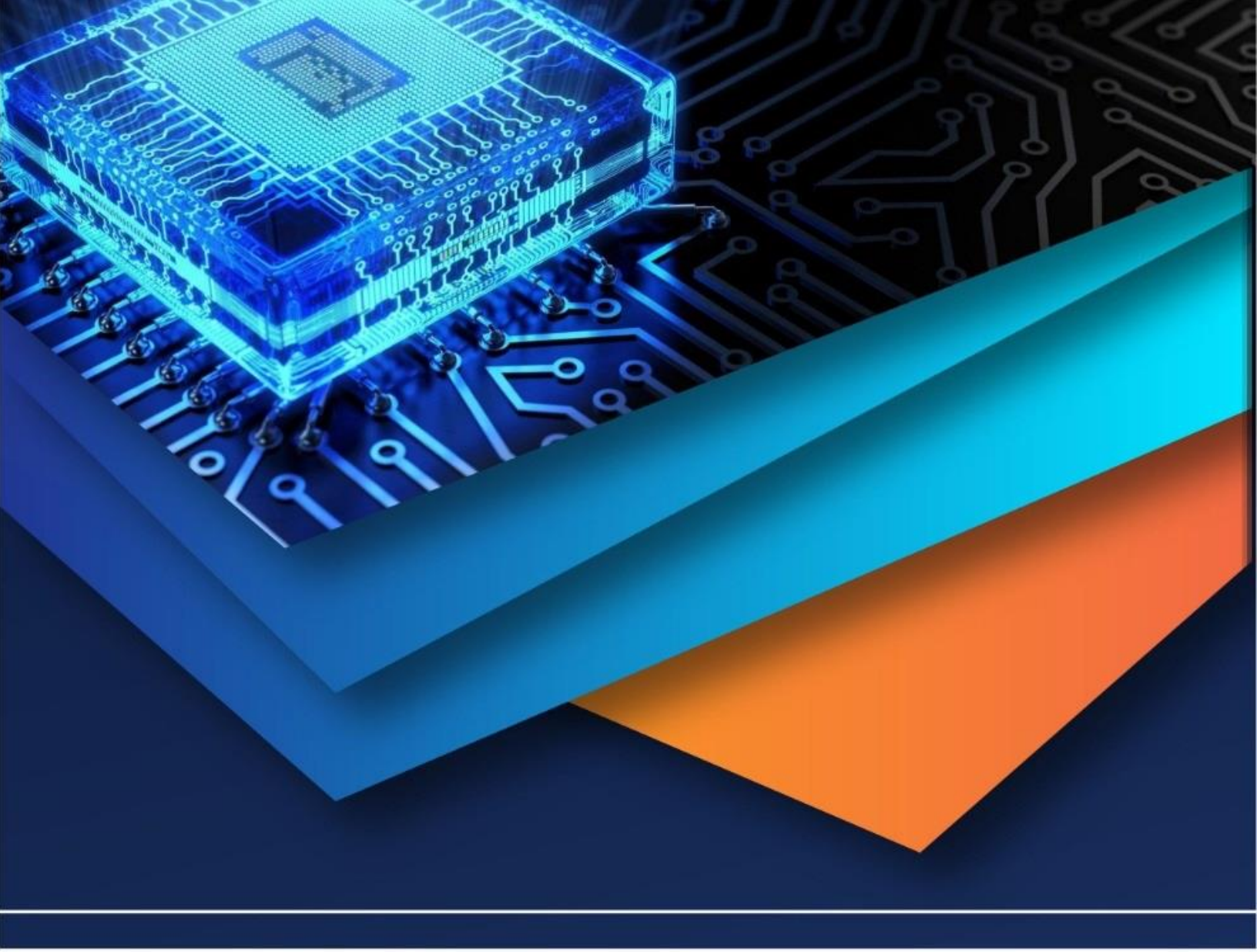
The proposed system uses Convolutional Neural Network (CNN) to train the model. The images are pre-processed, and histogram equalization is done to enhance the image contrast. The final accuracy on the test dataset is 93% and on the built dataset is 69%. The web cam predictions done by the model are also accurate and take very less time.

The benefits of "Traffic Sign classification and detection system" are generally focused on driver convenience. Despite the advantages of traffic sign classification, there are drawbacks. There can be times when the traffic signs are covered or not visible clearly. This can be dangerous as the driver won't be able to keep a check on his vehicle speed and can lead to accidents, endangering other motorists or pedestrians, demanding further research.

REFERENCES

[1] Akatsuka, H., & Imai, S. (1987). Road signposts recognition system (No. 870239). SAE Technical Paper.
[2] Albert Keerimolel, Sharifa Galsulkar, Brandon Gowray, "A SURVEY ON TRAFFIC SIGN RECOGNITION AND DETECTION", Xavier Institute of Engineering, Mumbai, India, International Journal of Trendy Research in Engineering and Technology Volume 7 Issue 2 April 2021.
[3] Aditya, A.M., & Moharir, S. (2016). Study of Traffic Sign Detection and Recognition Algorithms.
[4] Shao, F., Wang, X., Meng, F., Rui, T., Wang, D., & Tang, J. (2018). Real-time traffic sign detection and recognition method based on simplified Gabor wavelets and CNNs. Sensors, 18(10), 3192.
[5] SADAT, S. O., PAL, V. K., & JASSAL, K. RECOGNIZATION OF TRAFFIC SIGN.
[6] Li W., Li D., & Zeng S. (2019, November). Traffic Sign Recognition with a small convolutional neural network. In IOP conference series: Materials science and engineering (Vol. 688, No. 4, p. 044034). IOP Publishing.
[7] Brkic, K. (2010). An overview of traffic sign detection methods. Department of Electronics, Microelectronics, Computer and Intelligent Systems Faculty of Electrical Engineering and Computing Unska, 3, 10000.
[8] Almustafa, K. M. (2014). Circular traffic signs recognition using the number of peaks algorithm. Int J Image Process (IJIP), 8(6),514.
[9] Zaibi A., Ladgham A., & Sakly A. (2021). A Lightweight Model for Traffic Sign Classification Based on Enhanced LeNet-5 Network. Journal of Sensors, 2021.
[10] G. Bharath Kumar, N. Anupama Rani, "TRAFFIC SIGN DETECTION USING CONVOLUTION NEURAL NETWORK A NOVEL DEEP LEARNING APPROACH", International Journal of Creative Research Thoughts (IJCRT), ISSN:2320-2882, Vol.8, Issue 5, May 2020.
[11] Alghmgham, D. A., Latif, G., Alghazo, J., & Alzubaidi, L. (2019). Autonomous traffic sign (ATSR) detection and recognition using deep CNN. Procedia Computer Science, 163, 266-274.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)