TRAFFIC SIGN BOARD IDENTIFICATION AND VOICE VIGILANCE SYSTEM

M.Nikhitha[1], K.P.Swetha[2], L. Nagamani Swathi[3], A. Sushma[4] and E. Sai Santhoshini[5]

[1]*Department of Computer Science & Engineering, Santhiram Engineering College, India,*

[2, 3, 4, 5, 6]*Department of Computer Science & Engineering, Santhiram Engineering College, India*

[1]*E-mail: chaitu223@gmail.com*

ABSTRACT

Site visitors symptoms displayed at the roads play an vital role in our lives at the same time as riding. They deliver vital information, for the street users. This successively requires them to regulate their driving behaviour and make certain that they strictly comply with the street rules currently enforced without causing any problem to other drivers and pedestrians. Traffic sign class is hired to come across and classify site visitors signs to inform and warn a driving force beforehand to avoid violation of regulations. There are sure dangers of the prevailing structures, used for classification, like wrong predictions, hardware cost and maintenance, that are to a extraordinary quantity resolved via the proposed system. The proposed technique implements a visitors signs and symptoms category set of rules using a convolutional neural community. Also, it consists of the characteristic of internet cam detection of the traffic sign. This may help the driving force to look at the sign close to his / her eyes on the show screen and as a result keep his/her time in manually checking the traffic sign eachtime.

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

1. INTRODUCTION

Device gaining knowledge of algorithms have gained importance nowadays. Junk mail filtering, speech understanding, face popularity, street signal detection are only a few examples in which machine studying is deployed. In visitors zones, traffic sign recognition and class can be used to automatically become aware of traffic signs. That is carried out mechanically by means of the gadget because the traffic sign is detected and the signal call is displayed. So, even if any signal is missed with the aid of the motive force or has any lapse in attention, it will likely be detected. This allows to therefore warn the drivers and forbid sure moves like over speeding. It additionally disburdens the motive force and therefore, increases his/her consolation. As a result, making sure and maintaining a check on the traffic signs and symptoms and accordingly following them. Site visitors signs, certainly, offer us a large number of data and guide us for that reason in order that we can flow correctly. Visitors signal type may be very beneficial in automatic motive force help systems.

A convolutional neural network is a category of deep studying networks, used to study and check visible imagery. It is used to train the photo class and recognition version because of its high accuracy and precision.

2. LITERATURE REVIEW

In today's world, identity of traffic symptoms has emerge as an critical element of our lives. Looking on the growing site visitors, to ensure protection of all and for automatic using in the future, site visitors sign category is utmost essential. Substantial research has been performed round reputation of traffic and avenue signs. In 1987, the primary research on the subject "traffic signal recognition" become carried out through Akatsuka and Imai , wherein they tried to build a fundamental device that could understand visitors symptoms and alert the drivers and make certain his/her protection. But this turned into used to offer the automated reputation for just a few particular traffic signs.

Traffic signal reputation to start with regarded in the form of handiest velocity restriction popularity in 2008. Those symbols ought to handiest locate the circular speed restriction signs and symptoms. Alternatively, later, systems had been designed that accomplished detection on overtaking signs and symptoms. This technology was available within the Volkswagen Phaeton and within the 2012 in Volvo S80, V70 and lots of extra. But the foremost drawback of those systems became that they couldn't hit upon the town restriction signs and symptoms as they have been normally inside the shape of direction signs and symptoms. But in recent times, such structures are expected to be gift inside the future vehicles to help drivers at the same time as driving.

In [1], the authors used the coloration processing machine to lessen the impact of brightness and shadow at the images. This became the very first research performed on this topic by the authors, Akatsuka and Imai. In [2], the authors have performed a survey on traffic sign detection and popularity,

wherein HOG (Histogram of oriented Gradients) is used for classification cause. In [3], a entire study of various visitors sign recognition algorithms has been completed, wherein the best accuracy (99.46%) changed into obtained through MCDNN (Multi- column Deep Neural network). In [4], the authors have developed a version in which they're converting the pix to gray scale first and then filter out those pix the use of simplified Gabor wavelets. The gabor filters are used to extract functions. These wavelets are critical as they help to limit the product of its preferred deviation in both the time and frequency mapping. The authors extracted the regions of interest for reputation motive and labeled the signs the usage of "aid Vector gadget" (SVM).

In [5], the authors have extracted the areas of interest in the detection level and in addition examined the shapes of such areas. Here, inside the classification machine, they took the regions of hobby and categorised them into one of a kind classes.

In [6], the authors have created a module which includes numerous convolutions. They blended the 1*3 kernel and three*1 kernel and eventually connected it with the 1*1 kernel to acquire the 3*3 kernel. This become used to extract greater features and for this reason lessen the quantity of parameters. In [7], the writer has reviewed the visitors signal detection techniques and divided them into 3 types of techniques: color, form and gaining knowledge of primarily based methods. In [8], the author used the quantity of peaks algorithm to hit upon and understand round shaped site visitors signs and symptoms.

In [9], the authors have attempted developing a classification version the usage of the improved LeNet-five architecture, which includes two consecutive convolution layers (before the MaxPooling layer) to extract excessive stage functions from the picture. Additionally, they have used the facts augmentation approach to make the dataset stable. In [10], the authors have used the method of shade segmentation and the RGB based detection that's used to perceive the traffic signs and symptoms on the road. The optimizer used changed into "Stochastic Gradient Descent" with Nesterov Momentum. The text to speech device became carried out to alert the driver about the visitors sign. Also, they utilized the GPU (graphical processing unit) system, as a part of hardware. In [11], the authors have tried producing a dataset for the Arabic road signs and symptoms and for this reason increase a CNN model for Arabic signal reputation.

3. 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

Amongst those, the maximum not unusual dataset is the GTSRB (German traffic sign reputation Benchmark) dataset. The motive for its reputation is:

1) It includes massive number of images

2) The visitors' signs and symptoms are of various range, historical past, and coloration variant which in flip will assist the version to perform correctly.

Because the GTSRB dataset may be used for both detection in addition to class, the proposed gadget makes use of the equal. The dataset is in addition cut up into training, testing and validation dataset. The schooling dataset is the one which is used to teach the version. The validation dataset, in fashionable, is used to assess the version and update the hyper parameters. Hyper parameters are used to control the getting to know method and improve the accuracy, as an example, variety of epochs, the selection of activation characteristic. The test dataset is handiest used once the version is educated. It's miles used to check whether the model can make accurate predictions or no longer.

Similarly, histogram graphs are plotted (as proven in fig. 1), to expose the quantity of images in each class, for the training, checking out and validation facts sets respectively, where the X label denotes the "magnificence identification", and the Y label denotes the quantity of pictures. Plotting the graph enables to visualize the dataset.
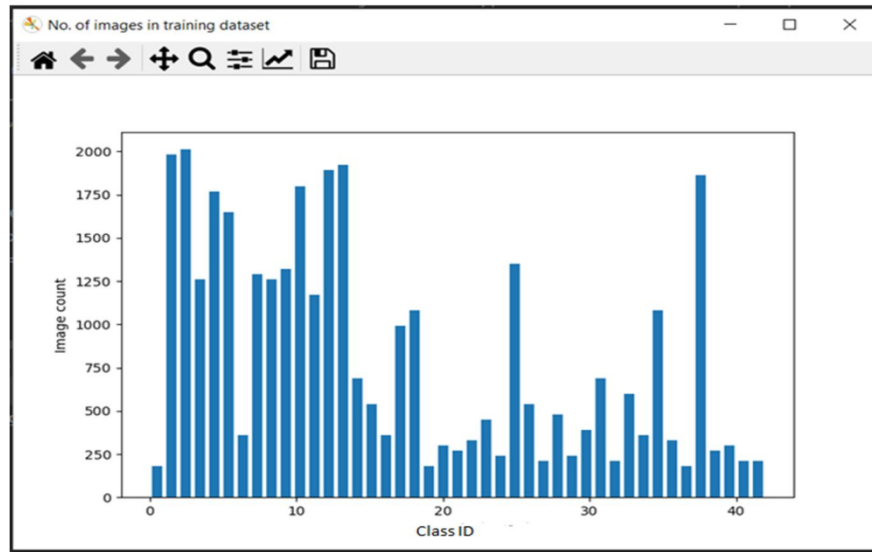

Fig.1: Class ID vs No. of images graph

B) **Flowchart**

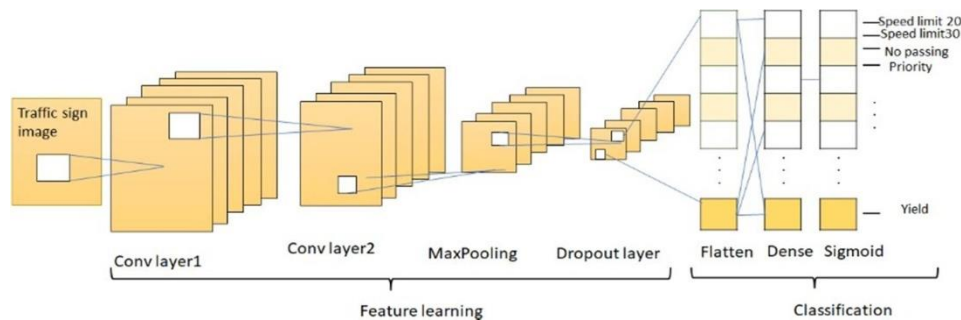**STRUCTURE OF THE MODEL**

Fig.2: Model Architecture

Initially, the CNN version architecture is constructed (as seen within the fig. 2). The subsequent steps are observed:

1) Sequentially upload the layers inside the order: two convolutional layers, one pooling layer, dropout layer, pulling down layer, dense layer, once more a dropout layer and finally the dense layer.

2) Inside the convolutional layer, variety of filters is distinct. It plays the convolution operation at the authentic photograph and generates a function map.

3) The ReLU performs the most feature to transform the poor values to 0 without converting the fine ones and generate a rectified function map. The Pooling layer takes the rectified characteristic map and plays a down-sampling operation (like Max Pooling or average pooling) and as a consequence reduces the dimensionality of the photograph.

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

5) The dropout layer is used to keep away from over fitting through putting a number of the input neurons to zero at some point of the education procedure. The dense layer, then again, feeds all the outputs from the preceding layer to all its neurons and carry out 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 (very last step inside the introduction of model to outline the loss characteristic and observe optimization strategies) and assign the loss characteristic as "sparse_categorical_crossentropy" and use the "Adam optimizer". The purpose for specifying this loss feature is that the proposed device is a multiclass class problem, in which more than one instructions are taken into consideration however one image belongs to precisely one class.

7) Next, the model is trained using the education dataset, by using passing the pre-processed photos from the training dataset.Eight) eventually, the predictions at the take a look at information are executed the use of the skilled version and the traffic signal call at the side of the class id is proven as an output.
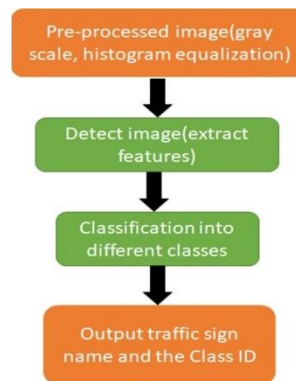
Fig. 3: Basic flow of the CNN model

C) Design/Method

Site visitors sign class is one of these uncommon topics of discussion. Maximum of the prevailing systems consciousness on detection most effective. Detection is in particular the extraction of features and find out the vital coordinates inside the image. Class is the categorization of picture into exclusive lessons.

The most not unusual dataset used for the cause is GTSRB which consists of forty three classes. Inside the proposed system, a prediction version is trained the use of this dataset. It performs excellent for image classification. Recently, Convolutional Neural network has been adopted in object reputation for its high accuracy and much less computational value.

Within the proposed gadget, the number one focus is in the direction of the visitors signal class which additionally prints the visitors sign call as soon as the detection of the picture is executed. There is a csv report which consists of the pairs of visitors sign name and the class id. This report enables to load the categorised records.

D) Gray Scale

Converting the RGB facts set into gray scale is one of the vital steps earlier than category the usage of CNN (as shown within the fig. 3). This has numerous blessings like:

1) Photos after converting to grey scale, assist the neural network to system them without problems as the unwanted biases are removed.

2) Grey scaling the pix allows to reduce the variety of computations, because the number of channels will get decreased after conversion.

3) This in flip allows to enhance the version accuracy.

Before grey scaling, photograph information shape (of training statistics set) is: (34799, 32, 32, three). Which means the photographs had been of 32x32 size and coloured in rgb format (three channels).After grey scaling, photo data form (of schooling statistics set) will become: (34799, 32, 32, 1). After grey scaling the scale of the photograph stays the equal (32x32) however the wide variety of channels is reduced to one.
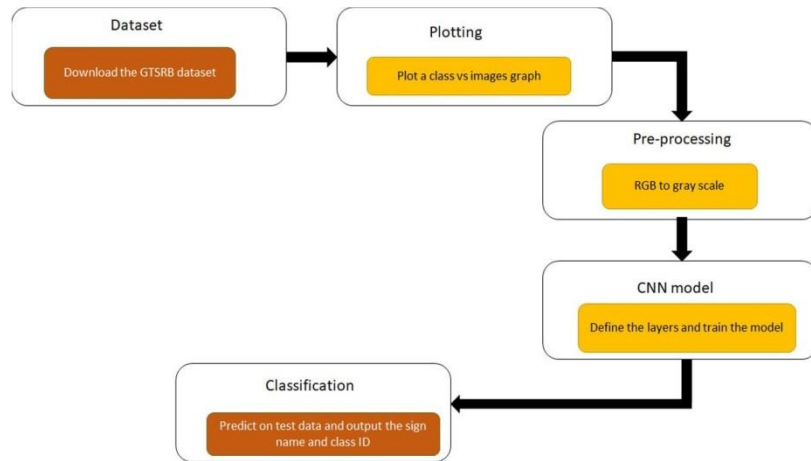
Fig.4: Flowchart of the proposed system

The drift of the proposed gadget is stated in the fig. 4. The proposed system is composed of various features corresponding to every operation.

1)building of the version: This in most cases makes a speciality of changing the pictures to grey scale, normalizing the pix (normalization is completed to boost up the education system and enhance the version overall performance), histogram equalization (to enhance image comparison), addition of the layers to the version, train the version, get predictions at the test statistics set, and subsequently display a few pattern photos with their site visitors signal name and sophistication identity because the output. The train, test and validation split percent is sixty five%, 25% and 10% respectively for the proposed system.

2)one of the major functionalities which can be applied on this paintings, is prediction of unknown pictures. Here, a small dataset was generated gathering pix from exclusive assets. This changed into the most important component as this dataset consists of some one of a kind images with special shade and structure. Even though there are several present datasets to be had, a small dataset (inclusive of 13 images) is constructed. The dataset consists of some speed restriction symbols, yield signal, caution signs (like prevent and no access), informatory signs (like pedestrians, in advance simplest, no passing, roundabout obligatory, and proper-of-manner at the subsequent intersection). Extracting functions from those pics isn't easy for the model. The reasons being, those snap shots are enlarged, having distinctive history shades and reduced clarity. Notwithstanding all of the troubles, the version efficiently anticipated around nine pictures out of 13. Simplest the pix that are curvilinear or in circular format, are not expected appropriately. The version predicts the closest traffic signal name determined for such snap shots.

The activation characteristic used is "ReLU". ReLU is one of these non-linear activation capabilities that is utilized in multi-layer neural networks. The motive being, ReLU applies a function $f(x) = max\ (zero, x)$ to all the values given as input. The ReLU layer simply modifications the negative values to 0 and maintains the advantageous values as it is. ReLU has emerge as the default activation characteristic to be used within the hidden layers of the neural network. The ReLU function is simple and much less computationally highly-priced because there aren't any complicated mathematical calculations involved and thus makes the version examine and teach quicker.

The model efficaciously predicts majority of the pictures with 93% accuracy as is proven in fig. Five. The best snap shots which aren't correctly anticipated are those which includes round shapes in it. This will be resolved by means of augmenting the pix and the usage of one warm encoding technique.

4.   RESULTS AND DISCUSSIONS

A)   Traffic signal type is the procedure of routinely recognizing traffic signs (like speed restriction, yield, and caution symptoms, and many others.) and for that reason classifies them as to which magnificence they belong to. The undertaking has predominant functionalities: Prediction on the newly generated dataset (fig. 6) and live internet cam site visitors sign detection.

B)   Accuracy is the ratio of range of correct predictions to the range of general predictions (eqn. 1). Accuracy = (The number of correct Predictions) / (total predictions).
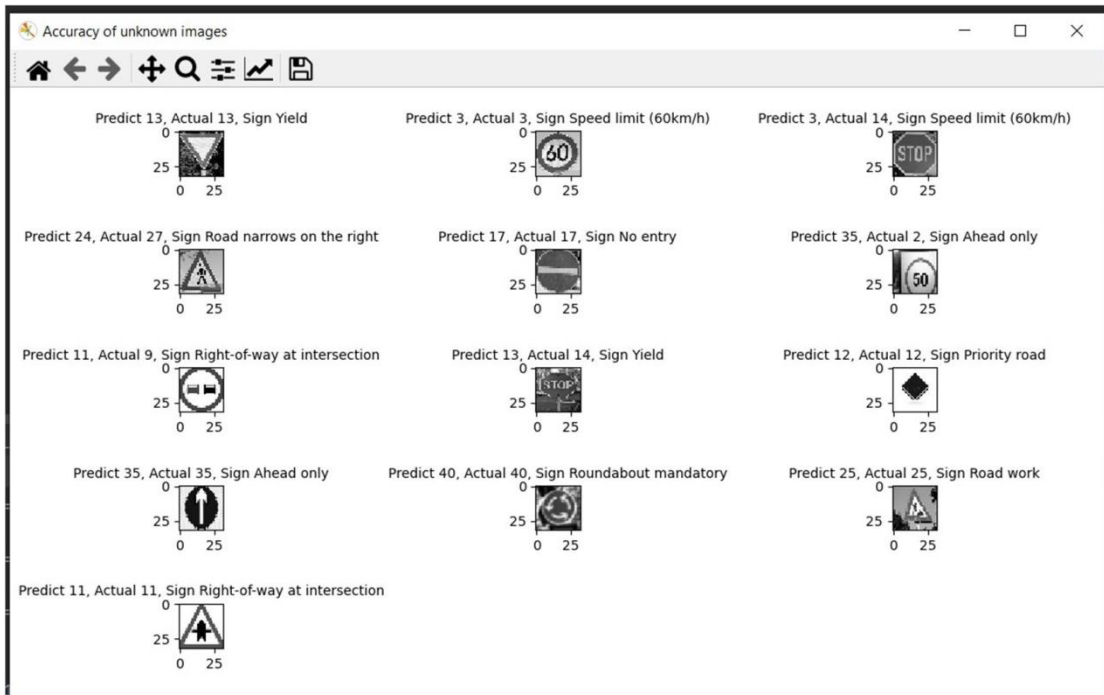
C)   Equation 1: Formula for accuracy

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

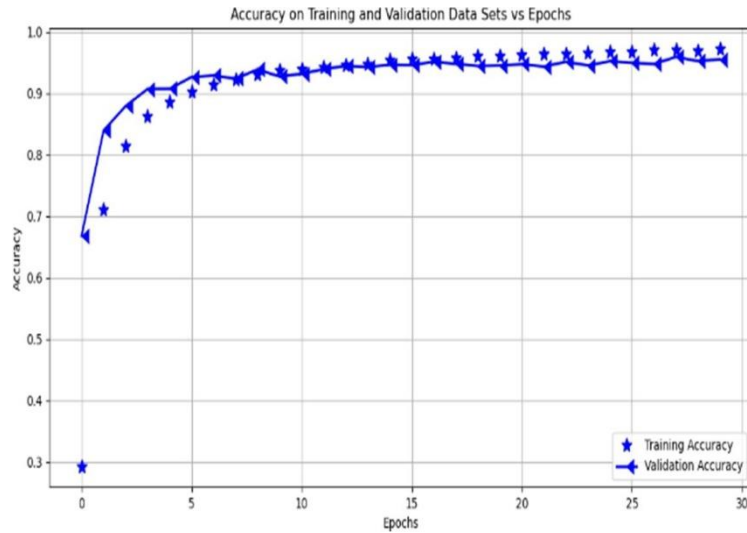| Sr No.	Dataset used	Accuracy
1.	GTSRB	93%
2.	Generated dataset	69%
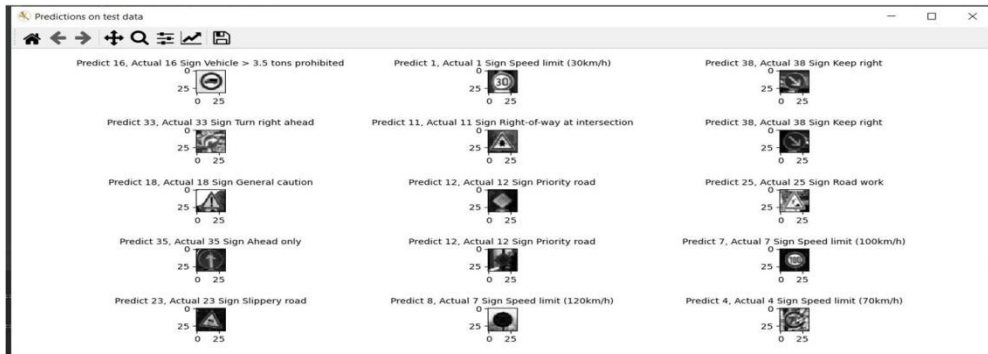
Fig.5: Predictions on Test dataset

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 education dataset. In addition, validation accuracy depicts the one the use of the validation dataset. The wide variety of epochs is a hyper parameter that defines the number instances that the studying algorithm will work thru the whole education dataset. It's far the number of iterations the education dataset will go through every time in the course of training of the version. As seen from fig. 7, at around 30 epochs, the education and validation accuracy in shape and is a directly line. Here, the training and validation accuracy are most. If those strains start keeping apart consistently, then we have to forestall the schooling procedure at an in advance epoch via visualizing the graph. This indicates that 30 epochs are sufficient for the version to extract features.

The loss on schooling 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 = - (yi log(y'i) + (1 – yi) log(1 – y'i))

D)   Equation 2: Mathematical formula

Where,

i is the ith training example in the data set,

yi - 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.

5.   LIMITATIONS

Although, there are many benefits of traffic signal class, there are certain problems as nicely. It can manifest that the traffic signal is hidden behind the timber or any board at the street side which may additionally purpose the wrong detection and classification of traffic sign. Every so often it can happen that the automobile went so speedy, that it did no longer detect the visitors sign. This can be risky and might lead to accidents. There's a want for in addition research to address these issues

6.   FUTURE SCOPE

Visitor's signs are beneficial to all of the individuals who are riding a vehicle on the street. Visitors signs guide the drivers for following all the traffic policies and keep away from any disruption to the pedestrians. The environmental constraints such as lights, shadow , distance (sign is pretty a ways), air pollutants, climate situations similarly to motion blur, and automobile vibration that are commonplace in any real time device might also have an effect on the detection and for that reason the type. As a result, there is a need for in addition studies and improvements to cope with those issues. Also, there are positive site visitors signs that won't be anticipated appropriately. For this, augmentation and one hot encoding strategies can be used. Augmentation entails transferring of the picture, zoom in and rotate the images (if required).

This device enables the driving force to take a look at the signal near his / her eyes on the screen. This saves the time and efforts in manually checking whether or not any traffic signal board is there, figuring out what sort of signal it is and act for that reason. Traffic sign class, for this reason, has a huge software in building smarter cars like automated using cars, in which the system mechanically detects, acknowledges a traffic signal and presentations it.

7.　CONCLUSION

The proposed gadget is simple and does the category pretty correctly on the GTSRB dataset in addition to the newly generated one (consisting of truly present pics of all kind), and ultimately the version can efficiently seize pix and are expecting them as it should be even supposing the history of the image isn't lots clear.

The proposed gadget uses Convolutional Neural network (CNN) to teach the model. The photos are pre-processed, and histogram equalization is executed to beautify the photograph contrast. The final accuracy at the check dataset is ninety three% and on the built dataset is sixty nine%. The net cam predictions executed by the version also are accurate and take very much less time.

The advantages of "site visitors signal category and detection system" are commonly focused on driving force comfort. Notwithstanding the benefits of site visitors sign type, there are drawbacks. There may be times while the visitors signs are included or now not visible genuinely. This will be dangerous because the motive force received't be able to maintain a check on his automobile pace and might cause 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.