JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

# TRAFFIC SIGN RECOGNITION USING CNN WITH VOICE ASSISTANCE

**Shaik Mohammad Shareef[1], Kanusu Jyothi[2], Bammidi Hari Krishna[3], Kotapati Venu Madhav[4] , Dr.V.Balaji[5]**

[1,2,3,4] Student, Department of Computer Science Engineering, ANU College of Engineering and Technology, Acharya Nagarjuna University, Guntur

[5] Assistant Professor, Department of Computer Science Engineering, ANU College of Engineering and Technology, Acharya Nagarjuna University, Guntur

## Abstract

Traffic signs are vital for showing information to drivers, and they are fairly important for road safety. Failure to detect or understand these signs could pose risks, underscoring the robust detection systems importance. This study describes a voice-led traffic sign recognition system that operates in real-time to assist drivers. A pre-trained Convolutional Neural Network (CNN) on the back end handles detection and recognition, while a text-to-speech engine for the driver provides the narration. We propose an efficient traffic sign detection and classification technique that achieves state-of-the-art performance on the German traffic sign recognition benchmark GTSRB while needing minimal processing resources and working in real time, using a model trained on a large dataset.With this two-pronged approach, you should be protected even in the event that a motorist disregards a sign because the system will detect and transmit it. Through the use of Convolutional Neural Networks (CNNs) to narrate recognized signs to drivers in text format, this system not only advances the development of autonomous vehicles and intelligent transportation systems, but it also benefits drivers in general. The suggested algorithm's usefulness is demonstrated by the experimental findings, which also offer support for its potential usage in advanced driver assistance systems, traffic management, and autonomous driving.

**Keywords:** Classification, Recognition, Artificial Intelligence, Road signs, Autonomous Vehicles, Convolutional Neural Network (CNN), Voice Feedback.

## 1. INTRODUCTION

In today's world, where cars are an essential mode of transportation, ensuring road safety is paramount. Road signs play a vital role in this effort by providing basic information to drivers and communicating specific rules and regulations for each area. Designed to convey messages quickly and efficiently, road signs require minimal reading skills so they are universally understood. Despite their importance, issues such as carelessness, inattention and ignorance of road signs contribute to drivers ignoring or misinterpreting them, leading to accidents. In urbanized areas, the multitude of different road signs can overwhelm drivers, while in rural communities, unfamiliarity with specific signs presents additional challenges. Some drivers may also ignore certain signs and believe they are unnecessary, further increasing safety risks. Once recognized, the system uses voice assistance to verbally inform drivers of the identified signs, ensuring immediate understanding without distracting them from the road. This innovative approach not only aims to improve driver awareness and compliance with traffic laws, but also has potential applications in both vehicle assistance systems and autonomous vehicles. By improving sign recognition capabilities and promoting safer driving, our system aims to reduce accidents and save lives on the roads. This introduction highlights the key role of advanced technologies in improving road safety and highlights the potential of integrating road sign recognition with voice assistance to transform urban mobility and traffic safety measures.

## 2. RELATED WORK

Recent studies have witnessed huge advancements in integrating Convolutional Neural Networks (CNNs) with voice output talents for Traffic Sign Recognition (TSR) systems. Some researchers use CNN architectures inclusive of ResNet and DenseNet for strong characteristic extraction and type of traffic signs and symptoms from datasets like GTSRB. These fashions are skilled to acquire high accuracy in real-time signal detection and category responsibilities. In parallel, integrating voice output into TSR systems has won attention as a method to decorate person interaction and accessibility. Some researchers exemplify this method, where a CNN-based totally TSR model is coupled with a text-to-speech device to offer auditory remarks of recognized site visitors' signs and symptoms. This integration enables drivers to acquire actual-time indicators and guidance through spoken instructions, contributing to more secure riding practices. Such improvements spotlight the synergistic benefits of mixing CNNs for visual popularity responsibilities with voice output technology, paving the manner for extra intuitive and person-pleasant TSR structures able to support various person wishes and environments.

These days, the development of self-driving car (SDC) systems for transportation may be booming. One of the difficult issues that researchers and developers must tackle in those systems is the recognition and popularity of traffic signs. This problem is tackled as an ongoing task that involves using computer vision to detect, recognize, and classify objects (traffic symptoms). This research focuses on the reputation of traffic symptoms without taking the detecting stage into account. This section addresses the best relevant works from this angle in order to further this cause. Sign recognition and capacity extraction are two factors that contribute to the popularity of traffic symptoms.Using the Multi-layer Perceptron Neural Network can yield better outcomes. Currently, convolutional networks are gradually taking the place of traditional PC vision algorithms for a number of packages, including pattern popularity and object kind. It is used to extract and learn the intensity description of the symptoms that users of the website encounter. This method gets around the descriptor extraction step, which can be highly sensitive to a number of variables. This community uses convolution methods to manipulate 2D images. It is capable of analyzing a photo's representative description.

## 3. METHODOLOGY

**Data Collection:** Gathering a sizable dataset of traffic sign photos is the initial stage in creating the traffic sign detection and recognition system. We collected the GTSRB dataset from its official website. The dataset includes 43 types of traffic sign images. This dataset includes images of traffic signs from various countries and regions, with different shapes, colours, and sizes

**Data Preprocessing:** The collected dataset needs to be preprocessed to remove any noise or artefacts that might interfere with the performance of the deep learning model. This entails actions like scaling the pictures, standardizing the pixel values, and expanding the collection to include more examples.

**Network Architecture:** The next step is to design the deep learning architecture that will be used for traffic sign detection and recognition. This involves selecting a pre-existing CNN architecture. The network should be capable of detecting traffic signs of various shapes and sizes, as well as recognizing them accurately.

**Training:** Once the network architecture has been defined, the preprocessed dataset will be used to train the model in the following step. This entails configuring the settings for the training, including the learning rate, batch size, and number of epochs. By modifying its internal parameters in response to input photos and matching output labels, the network gains the ability to identify and detect traffic signs during training.

The GUI component would allow users to input an image and have the CNN model detect and recognize any traffic signs present in the image. The GUI could also display the model's output, such as the type of traffic sign detected and any relevant information (e.g., speed limit, warning, etc.). This would make the application more user-friendly and accessible to individuals who may not have a technical background in machine learning.

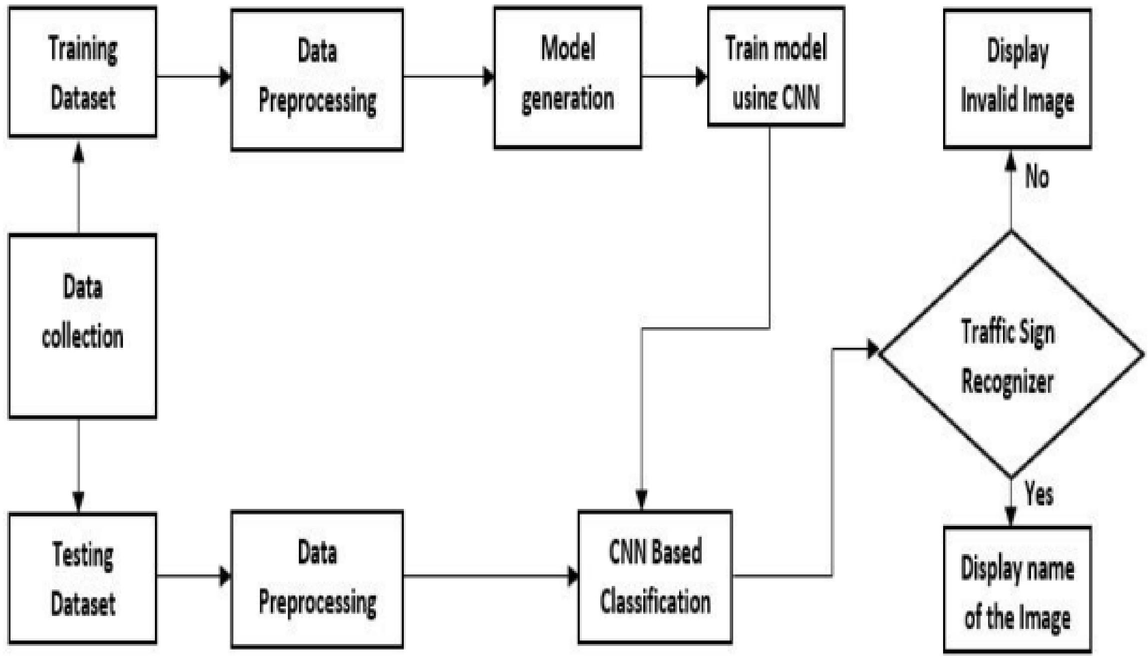



**Fig -1: Steps involved in the methodology**

## 4. SYSTEM ARCHITECTURE

Our deep CNN architecture is the basis of our traffic sign detection and recognition system. It consists of several convolutional layers, max-pooling layers, and fully linked layers. The goal of the architecture is to locate and identify traffic signs in a given dataset of images.

## 5. FLOW DIAGRAM

The dataset is first collected from the Kaggle website, and then it is preprocessed to get rid of the noise in the images. Training data and testing data are two files that contain the complex dataset. The images are recognized using a CNN model. The model is now trained using the training data, and it is then saved. The stored model is then given a test image that has been preprocessed. The classifier then names the image and assigns it a classification

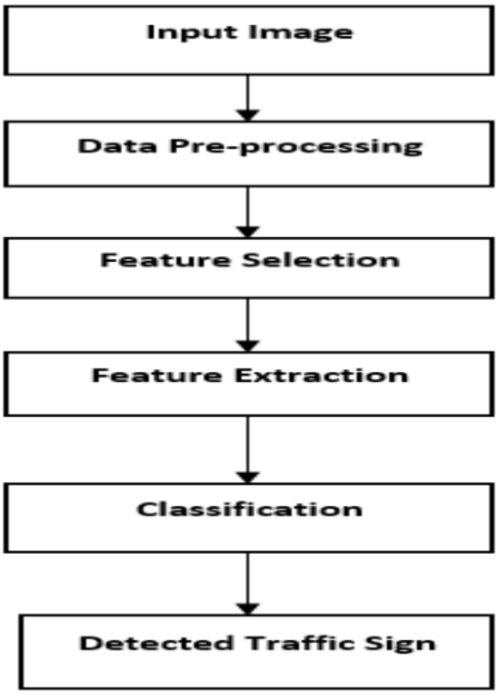



**Fig -2**: System Architecture

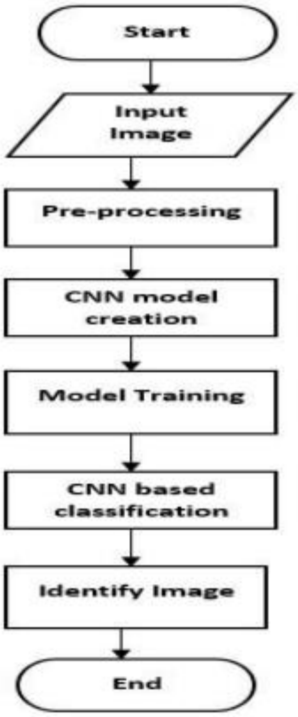



**Fig -3:** Data Flow

## 6. IMPLEMENTATION

Traditional TSR methods rely on a series of handcrafted algorithms for tasks such as feature extraction, feature selection, and classification, which require extensive knowledge of computer vision and signal processing. However, CNNs, or artificial neural networks, have set records in a variety of computer vision applications, including segmentation, object identification, and picture classification. CNNs are highly skilled at processing visual input. Research utilizing the German Traffic Sign Recognition Benchmark (GTSRB) dataset has shown how effective CNN models are in reducing crashes and improving traffic safety. These studies have employed CNNs to achieve precise feature extraction and high accuracy in classifying traffic signs. Building upon these findings, researchers have developed systems that combine CNN-based TSR with text-to-speech technology, enabling real-time voice feedback for detected traffic signs. This auditory output helps drivers stay focused on the road by providing instant alerts and instructions, thereby enhancing driving safety. The combination of CNNs for visual recognition and voice output technologies demonstrates the potential for creating more intuitive and user-friendly TSR systems that cater to diverse user needs and promote safer driving practices.

## 6.1 DATA COLLECTION

In this project, we are going to train and classify traffic signs using Convolutional Neural Networks (CNN), it will be done using Opencv in real-time using our webcam, and in this project, we will train traffic science with over 35000 images of 43 different classes with the help of TensorFlow and Keras. The dataset can be utilized for supervised learning tasks like object detection and picture classification because each image is labelled with the appropriate traffic sign class. Of the more than 35,000 photos in the training set, 70% are utilized for training, 20% are used for validation, and 10% are used for testing.

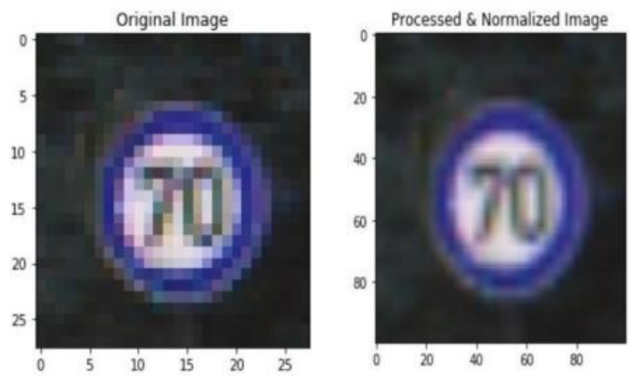




**Fig -4:** Sample Data **Fig -5:** Image before and after Pre-processing

## 6.2 DATA PREPROCESSING

When working with the GTSRB dataset, I've noticed that the images come in different sizes. To make sure my model works correctly, I need to resize these images to a common size before training. I can do this by either cropping or scaling the images, ensuring that their aspect ratio is maintained. Additionally, normalizing the pixel values of the images is crucial. This step helps improve my model's performance by reducing the impact of variations in lighting and colour. A common approach I use is scaling the pixel values to a range of 0 to 1. Alternatively, I might subtract the mean and divide it by the standard deviation. By taking these preprocessing steps, I can ensure that the data I feed into my model is consistent, which leads to better and more reliable training results.

## 6.3 MODEL ARCHITECTURE

One type of deep learning neural network that is frequently used for photo identification and classification tasks is the convolutional neural network (CNN). It is meant to automatically and adaptively learn spatial hierarchies of features from raw input data, such as photos, by applying convolutional filters. Each layer in a typical convolutional neural network (CNN) has a distinct purpose in the picture analysis process. An outline of the most often utilized layers in CNNs is provided below:

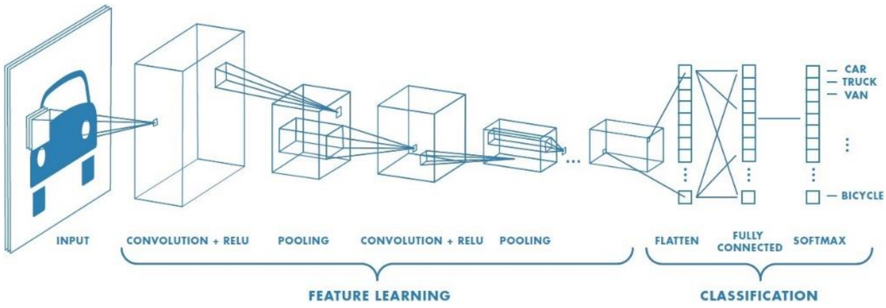



**Fig -6**: Understanding of CNN Model

**Input layer**: This layer represents the input image, which is typically a 2D or 3D matrix of pixel values.

**Convolutional layer:** To extract features from the input picture, a collection of convolutional filters is used in this layer. Each filter creates a feature map by performing element-wise multiplication and summing on the input picture using a tiny matrix of weights.

**Activation layer:** To add non-linearity to the model and enhance its capacity to capture intricate patterns, this layer applies a non-linear activation function, such as ReLU, to the feature map produced by the convolutional layer.

**Layer for pooling**: To reduce the size of the feature map, this layer downsamples it. The most common pooling technique is max-pooling, which selects the maximum value inside a rectangular region of the feature map.

**Fully connected layer**: To complete the final classification or regression process, this layer adds a set of weights to the flattened feature vector created by the preceding layers.

The outcome of the network, which is often a probability distribution over the several classes involved in the classification operation, is shown in the output layer.

**Dropout layer:** In deep learning neural networks, such as convolutional neural networks (CNNs), dropout layer is a regularization technique. The Dropout layer's job is to stop overfitting, which happens when a model gets too good at fitting the training set and performs poorly when applied to fresh, untested data. During training, the Dropout layer eliminates a predetermined proportion of its neurons at random. Because it cannot rely too much on any one trait, this drives the network to acquire more robust and generalizable features.

**ReLU (Rectified Linear Unit):** ReLU is a basic activation function that preserves positive values while setting all negative values to zero. Because of its efficiency and simplicity, CNNs use it extensively.

**Softmax Function:** A popular activation function in neural networks, particularly convolutional neural networks, is the Softmax function (CNNs). In a multi-class classification issue, it is used to generate probability distributions over a number of classes. A vector of real numbers is fed into the Softmax function, which outputs a probability distribution over the classes with a total probability of 1 and a value for each probability between 0 and 1.

## 6.5 TRAINING AND TESTING THE MODEL

Due to their greater representation in terms of quantity, certain classes are given preference over others during the training phase of the German Traffic Sign Benchmark due to the uneven distribution of images. In order to ensure optimal network learning, several classes' data is augmented by performing geometric modifications (shear mapping, translation, and rotation) on a large number of their photos.

We split our data in the training and testing phase as follow: 70 per cent for the training, 10 per cent for the testing and 20 per cent for the validation. We have done our training data after creating our convolutional neural network, using TensorFlow we train it and print the result. Training our model took about 60 minutes because we set the variable epoch_val to 10 and each epoch was about 6.1 minutes.

It is important to know that we do not have the same number of images for each class

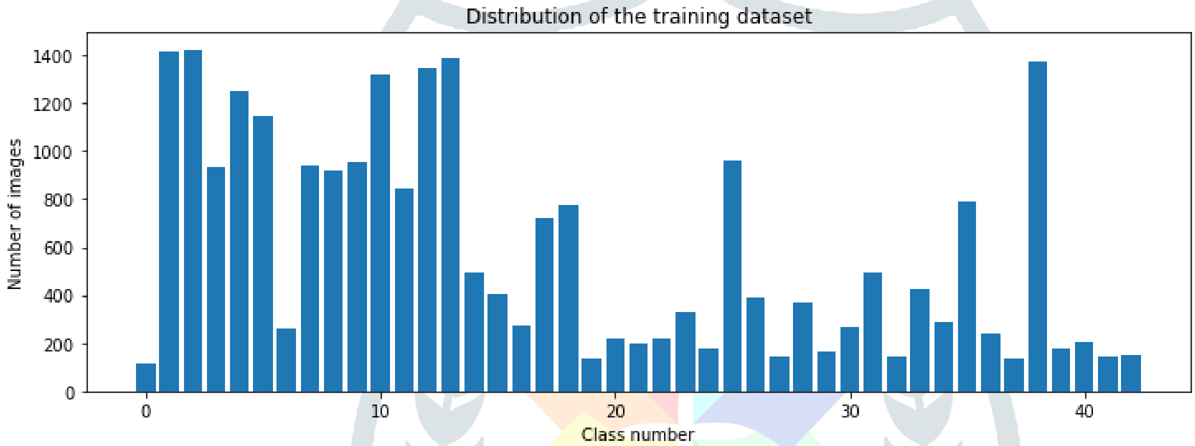
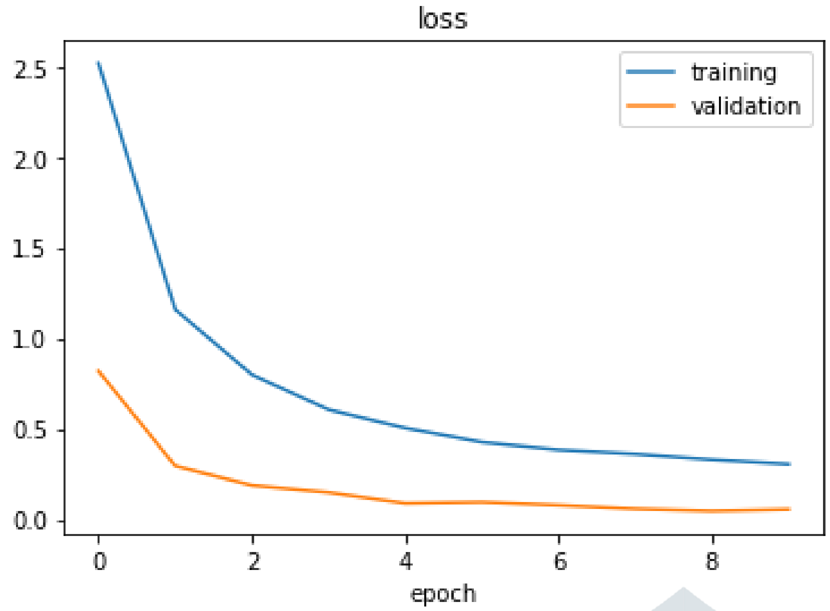


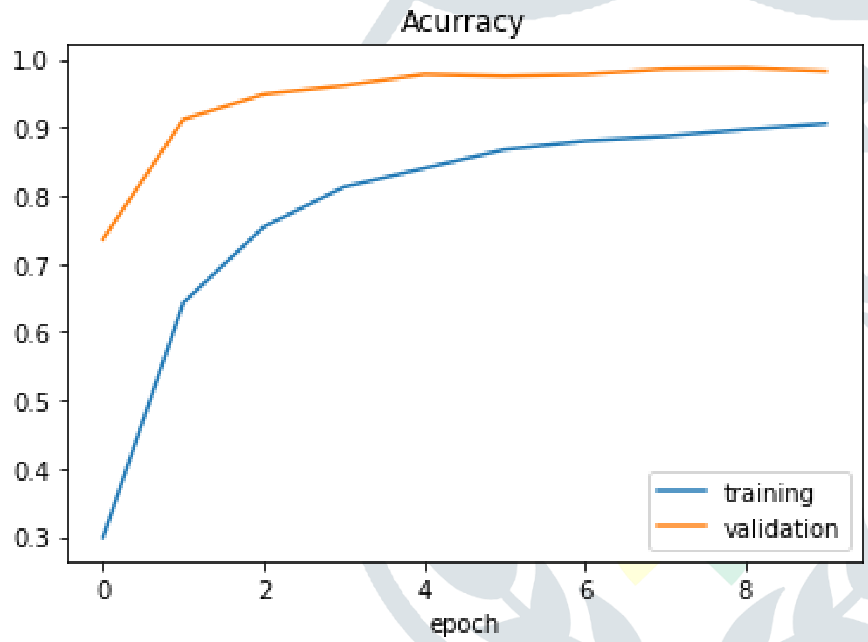

**Fig -7:** Distribution of Training Dataset

Among the deep learning models, convolutional neural networks (CNNs) can be evaluated by evaluating how well they perform on an alternative test set of data that wasn't used for training or validation. This is done to ensure that the model has learned to generalize to new data adequately and is not overfitting. Furthermore, a common method of evaluating the model's performance on test data is to create a graphical user interface (GUI) that allows users to interact with the model and evaluate its predictions in real-time. This is useful for a variety of applications, including as picture segmentation, object detection, and facial recognition.

Loss and Accuracy plots on training and validation datasets for iterations (epochs).

These two images represent our lost accuracy on the training and validation datasets:




We can see that we are getting a fairly good result, we can see that after about 6 epoch it is going at the same level so probably 6 to 8 epochs will be a good estimation of where we want to learn.

## 7. CONCLUSION

We discussed how to utilize a convolutional neural network with OpenCV in real-time to classify traffic signs using a basic webcam with good accuracy and experiment with alternative model topologies. We created extremely adaptable code and devised a versatile method for assessing several architectures. On the test set, our model achieved 98% accuracy, while on the validation set, it approached 90% accuracy. In this study, we examined the use of convolutional neural networks (CNN) for traffic sign identification and recognition using the German Traffic Sign Recognition Benchmark (GTSRB) dataset.

# REFERENCES

[1] Sunitha. A, Shanthi. S, "Traffic Sign Classification and Detection Using Deep Learning" International Journal Of Innovative Research In Technology, Volume 6, Issue 11, ISSN: 2349-6002, April 2020.

[2] Opencv documentation https://opencv.org/

[3] Python Project on Traffic Signs Recognition with 95% Accuracy using CNN&Keras https://data-flair.training/blogs/python-project-traffic-signs-recognition/

[4] Mohit Singh, Manish Kumar Pandey, Lakshya Malik, "Traffic Sign Detection and Recognition for Autonomous Vehicles" International Journal of Advance Research, Ideas and Innovations in Technology, Volume 4, Issue 2, ISSN: 2454-132X, 2018.

[5] Deepali Patil, Ashika Poojari, Jayesh Choudhary, Siddhath Gaglani, "CNN-based Traffic Sign Detection and Recognition on Real Time Video" International Journal of Engineering Research & Technology, Volume 9, Issue 03, ISSN: 2278-0181, February 2021.

[6] Glory Reuben Maxwell, Dr. Dinesh D. Patil, "A Review on Traffic Sign Detection and Recognition System" International Research Journal of Engineering and Technology, Volume: 07, Issue: 05, e-ISSN: 2395-0056, p-ISSN: 2395-0072, May 2020.

[7] P. Dewan, R. Vig, N. Shukla and B. K. Das, "An Overview of Traffic Signs Recognition Methods," International Journal of Computer Applications, Vol. 168 – N..11, June 2017