How to Cite:
Saraswathy, C., Nandhini, K., & Kalaivani, R. (2022). Traffic sign detection and recognition based on convolutional neural network. *International Journal of Health Sciences*, *6*(S8), 1224–1238. https://doi.org/10.53730/ijhs.v6nS8.9884

Traffic sign detection and recognition based on convolutional neural network

Saraswathy C
Associate Professor, Department of ECE, K.S.Rangasamy College of Technology, Tiruchengode, Tamil Nadu, India

Nandhini K
Under Graduate Students, Department of ECE, K.S.Rangasamy College of Technology, Tiruchengode, Tamil Nadu, India
Corresponding author email: nandhini1184@gmail.com

Kalaivani R
Under Graduate Students, Department of ECE, K.S.Rangasamy College of Technology, Tiruchengode, Tamil Nadu, India

***Abstract*---**Traffic sign detection is one of the critical technologies in the field of intelligent transportation systems (ITS). The difficulty of traffic sign detection mainly lies in detecting small objects in a wide and complex traffic scene quickly and accurately. To develop an efficient traffic sign detection and recognition that can detect and categorize traffic signal into different classes in real-time with the help of Deep Learning techniques. In this project, regard traffic sign detection as a region classification problem and propose a two-stage R- CNN-based approach to solve it. To build a deep neural Network model that can classify traffic sign present in images into different categories. The training process takes predefined traffic signals and has them 'learn' into a model. With this model, people are able to read and understand traffic signs which are a very important task for all autonomous vehicles. A visual traffic sign recognition system can be integrated into the automobile with the objective of detecting and recognizing all emerging traffic signs. In case the driver refuses to heed traffic signs, the system will trigger an alarm. People fail to understand that rules are made for their own safety, being unaware of the road sign is another reason for accident will happen. The traffic sign-red light will glow, if the bike is not stop in the traffic, the bike number plate is noted and send the message to the police station and the police an axes the bike.

***Keywords*---**traffic sign detection, recognition, deep learning, convolutional neural network.

Introduction

At present, the advanced driver assistance systems (ADAS) or autonomous driving vehicles have achieved fairly good results in simple environments such as highways and closed parks. To further improve the autonomous driving capability, dealing with more complex scenarios including the urban roads need to be addressed. In these cases, it is required to have the ability to sense the traffic conditions. One of the important issues is to detect the traffic lights and understand their states for driving instructions. Although the vehicle positioning by GPS combined with GIS mapping can provide the rough road junction information, the exact locations and states of the traffic lights are not guaranteed to be precisely marked in the HD maps. Thus, online detection and recognition of traffic signals are ease.

According to global road crash data, approximately 1.3 million people die in traffic accidents each year, averaging 3,287 deaths each day. Unfortunately, drunk driving, reckless driving, fatigue, and driver distraction continue to be the leading causes of road deaths. With today's traffic control technologies, there's a good chance the motorist will miss part of the traffic. A survey of Fatalities in 1970 to 2019 is shown the below survey on traffic signs.

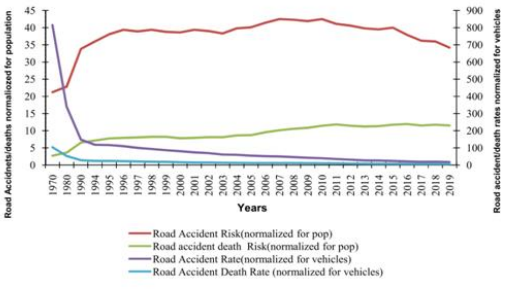


Figure 1: Survey on Traffic signals

The graphical representation of the above indicator. It will be noted from the chart that because of the steep increase in vehicle population over the years which has grown at CAGR of 11.15% during the period 1970 to 2019, the Road accident rate normalized for vehicles, through more than the Road accident death rate, shows a downward trend. On the other hand, Road accident Risk and Road death risk (normalized for population) both shows an upward trend with road accident risk higher than the death risk.

An on-board computer vision system that can detect and identify traffic signs could help drivers avoid accidents in a variety of ways. The on-board vision technology might supplement reality by displaying forthcoming warning signs ahead of time, or even keeping them shown on a screen after the sign has past. This would make it less likely that the driver would miss an important sign. Traffic sign recognition is a system that allows a vehicle to detect traffic signs

placed on the road, such as "speed limit," "children," or "turn ahead." It detects traffic signs using image processing techniques.

Road and traffic signs must be properly installed in the necessary locations and an inventory of them is ideally needed to help ensure adequate updating and maintenance. Automatic means of detecting and recognising traffic signs can make a significant contribution to this goal by providing a fast method of detecting, classifying and logging signs. This method helps to develop the inventory accurately and consistently. Once this is done, the detection of disfigured or obscured signs becomes easier for human operator. Different Types of Traffic signals are shown in below. The total 43 classes in German traffic sign region benchmark (GTSRB), there are three categories: Warning, Prohibitory & Mandatory signs.

Figure2: Different Types of Traffic signals

Literature review

[1] Lijing Wei et.al., (2020) has proposed traffic sign detection is one of the critical technologies in the field of intelligent transportation systems (ITS). The difficulty of traffic sign detection mainly lies in detecting small objects in a wide and complex traffic scene quickly and accurately. In this paper, the regard traffic sign detection as a region classification problem and propose a two-stage CNN-based approach to solve it. At the first stage, design an efficient network which is built with improved fire-modules to generate object proposals quickly. The network up-samples and merges the feature maps of different scales to attain a high-resolution fused feature map which contains semantically strong features of multi-scale objects. Specially, the prediction is made on the fuse feature map and based on the novel center-point estimation. With the overall designs, in this region proposal network can achieve high recall value while using low-resolution images. At the second stage, a separate classification network is proposed.

[2]Prajwal M.J et.al.,(2019) has proposed **i**n current situation, the come across various problems in traffic regulations in India which can be solved with different ideas. Riding motorcycle/mopeds without wearing helmet is a traffic violation which has resulted in increase in number of accidents and deaths in India. Existing system monitors the traffic violations primarily through CCTV recordings, where the traffic police have to look into the frame where the traffic

violation is happening, zoom into the license plate in case rider is not wearing helmet. But this requires lot of manpower and time as the traffic violations frequently and the number of people using motorcycles is increasing day-by-day. What if there is a system, which would automatically look for traffic violation of not wearing helmet while riding motorcycle/moped and if so, would automatically extract the vehicles' license plate number. Recent research have successfully done this work based on CNN, R-CNN, LBP, Hog, Hear features,etc. But these works are limited with respect to efficiency, accuracy or the speed with which object detection and classification is done.

[3]Jian-he et.al., (2017) has proposed based on image processing, bilateral Chinese transform, and vertex and bisector transform techniques. The images captured from the dash cam are processed with the histogram of oriented gradients to form feature vectors, followed by support vector machines to detect the traffic signs. The bilateral Chinese transform and vertex and bisector transform are used to extract the area of traffic sign from images. Finally, a neural network is adopted to identify the traffic sign information. In this work, test the algorithms using the images captured from the camera mounted behind the front windshield. The experiments are evaluated with real traffic scenes and the results have demonstrated the effectiveness of the proposed system.

[4]C Krishna Mohan et.al., (2016) has proposed an approach for automatic detection of bike-riders without helmet using surveillance videos in real time. The proposed approach first detects bike riders from surveillance video using background subtraction and object segmentation. Then it determines whether bike-rider is using a helmet or not using visual features and binary classifier. Also, present a consolidation approach for violation reporting which helps in improving reliability of the proposed approach. In order to evaluate an approach, they have provided a performance comparison of three widely used feature representations namely histogram of oriented gradients (HOG), scale-invariant feature transform (SIFT), and local binary patterns (LBP) for classification. The experimental results show detection accuracy of 93.80% on the real world surveillance data. It has also been shown that proposed approach is computationally less expensive and performs in real-time with a processing time of 11.58 ms per frame.

[5] Tarequal Islam et.al., (2017) has proposed the traffic sign detection and recognition topic are one of the most popular topics of computer vision and image processing in recent years, as they play an important role in autonomous driving and traffic safety. In this proposes a system that will detect and classify different types of traffic signs from images. This paper differs from other papers as it uses signs that are globally recognized and isn't limited to very few signs like many other papers. The number of signs used in this paper for classification is 28, which are used all around the globe. Two separate neural networks have been used for detection and recognition purpose; one classifies the sign and other the shape. Image augmentation has been used to create the training and validation dataset. 40,000 images have been used to train the first classifier with 28000 positive images (images that contain the traffic signs) and 12000 negative images(images that do not contain any traffic signs) and 3600 images were used to train the second classifier with 2400 positive images and 1200 negative images.

The images are processed to find the region of interest, which is then fed to two CNN classifiers for classification.

Conclusion from the literaturereview

Traffic sign recognition system (TSRS) is a significant portion of intelligent transportation system (ITS). Being able to identify traffic signs accurately and effectively can improve the driving safety. The traffic sign recognition technique on the strength of deep learning, which mainly aims at the detection and classification of circular signs. Firstly, an image is preprocessed to highlight important information. Secondly, Hough Transform is used for detecting and locating areas. Finally, the detected road traffic signs are classified based on deep learning. In this article, a traffic sign detection and identification method on account of the image processing is proposed, which is combined with convolutional neural network (CNN) to sort traffic signs. On account of its high recognition rate, CNN can be used to realize various computer vision tasks. TensorFlow is used to implement CNN. In the German data sets, they are able to identify the circular symbol with more than 98.2% accuracy.

Methodology

Training dataset

The dataset contains more than 50,000 images of different traffic signs. It is further classified into 43 different classes. The dataset is quite varying; some of the classes have many images while some classes have few images. Traffic images of each category are divided into the training dataset and test dataset, where the training dataset is contains of classes, and every category contains various image. Most of these images are collected from the internet and selected according to the specific requirements. The below 3 different super classes of prohibitory, mandatory and warning sign.

(a) (b) (c)

Figure 3: Sample images of prohibitory, mandatory and warning sign

Existing system

The early development of image-based traffic sign detection is mostly based on conventional computer vision techniques. The input images are converted to different colour spaces and various features such as colour, shape, edge and gray level intensity are used for detection. In the later machine learning based approaches, image features such as HOG or Harr-like operators are adopted for SVM or AdaBoost classification techniques.

There also exist techniques using multiple sensor inputs. Another method was to integrate the depth information obtained using a stereo camera for traffic sign detection Alternatively, they used a dual camera system to increase the range of traffic sign detection with different focal length settings .They used a long focal length camera to detect the far away traffic signs, while a wide-angle lens camera is adopted to detect the close by traffic signs.

Proposed system

In the proposed method make a CNN block where predictions are directly preformed across multiple feature levels. For this project, they are using the general public dataset available at Kaggle i.e. German traffic sign recognition benchmark (GTSRB). The procedure for traffic sign steps as follows
Step 1: Explore the dataset the 'train' folders contain 43 folders each representing a special class. The range of the input dataset is from 0, 1, 2, up to 42. With the assistance of the OS module, iterate over all the classes and append images and their respective labels within the data and labels list.
Step 2: Build a CNN model to classify the pictures into their respective categories; it will build a CNN model (Convolutional Neural Network). CNN is best for image classification purposes.
Steps 3: Train and validate the model after building the model architecture, then teach the model using model.fit ().
Step 4: Test the model with test dataset contains a test folder and during a test.csv file, the small print associated with the image path and their respective class labels. Now they are getting to build a graphical interface for in traffic signs classifier with Tkinter. Tkinter is nothing but a GUI toolkit within the standard python library. Here we upload the pictures and classify the image.

Block diagarm of proposed method

To give the input image that will include the preprocessing Image is an important part of the TSDR system that shows the below diagram. Whose main idea is to remove low-frequency background noise, normalizing the intensity of the individual particles images, removing reflections, and masking portions of images. Below is a description of selected image preprocessing techniques. The input image is divided into gray scale image separately. The selected Dataset is properly pre-processed and then renamed into the traffic dataset. The model is properly trained using CNN and then classification takes place. The input test image is acquired, pre-processed and then converted into array form for comparison. The comparison of test image and trained model takes place. The main heart of this traffic system is microcontroller. IR sensor are connected to the PIN2 of the microcontroller and LCD Display are connected to 8 to 13 pin.IR sensor output becomes logic 0 otherwise logic 1. By receiving these IR sensor outputs, Buzzer connected to pin 6 and motor pin 7.we have to write the program to monitoring traffic signs. The ESP8266 Wifi module is a complete Wifi network where you can easily connect as a serving Wi- Fi adapter, wireless internet access interface to Ardiuno on its simple connectivity through Serial Communication or UART interface.

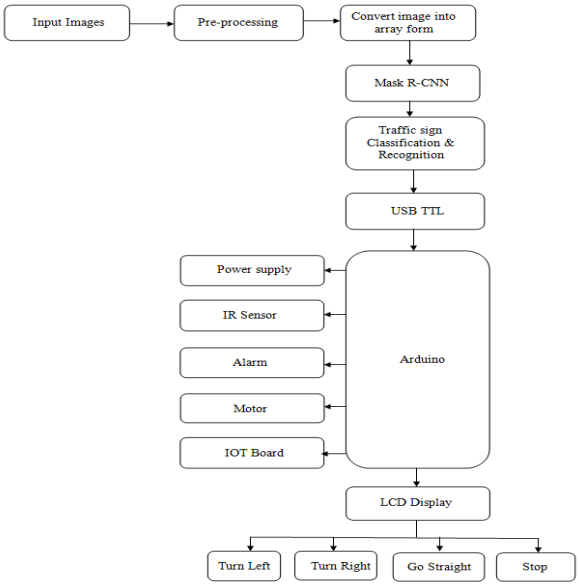


Figure 4: Block diagram of proposed method

By using Software Serial Library here, we have allowed serial communication on pin 10 and 11, and made them Rx and TX respectively and left the Rx pin of IOT Module open. By default Pin 0 and 1 of Arduino are used for serial communication but by using the Software Serial library, we can allow serial communication on other digital pins of the Arduino. 12 Volt supply is used to power the iot Module.

Here in this project, they are going to build an Arduino based traffic sign. Stop sign detects the sudden change in the IOT module sends the alert message on your Mobile Phone with the location of the sign recognize place is sent in the form of Google Map link, derived from the latitude and longitude from android app.

Description of block diagram

Modules

- Data Acquisition
- Pre-processing
- Convolutional neural network
  - Data Classifier
  - Parametric evaluation
  - Arduino
  - IR sensor
  - LCD Display
  - Buzzer
  - Wifi module
  - Motor

Data Acquisition

Input traffic data will be acquired by the system, the image taken will be proposed per frame by the system. Some testing process will make use of traffic images. For, the detection phase, a set of local traffic images from Google images containing actual road and traffic images with different traffic signs to be detected will be used to test the models. Another set of data is composed of road traffic images from dataset is containing actual road with traffic signs to be detected captured in different distances. The last set of data is actual image from local to test the full performance of the traffic sign detection and recognition system.

Pre-processing

Preprocessing is one of the most important steps in text classification. The detection phase is composed of pre-processing, color based segmentation, shape-based detection and object localization. This study will evaluate four different pre-processing and color based segmentation methods that solved the problem of lighting variations affecting the detection phase.

Convolutional Neural Network (CNN)

CNNs are a class of Deep Neural Networks that can recognize and classify particular features from images and are widely used for analysing visual images. Their applications range from image and image recognition, image classification, medical image analysis, computer vision and natural language processing. The term 'Convolution" in CNN denotes the mathematical function of convolution which is a special kind of linear operation wherein two functions are multiplied to produce a third function which expresses how the shape of one function is modified by the other. In simple terms, two images which can be represented as matrices are multiplied to give an output that is used to extract features from the image.

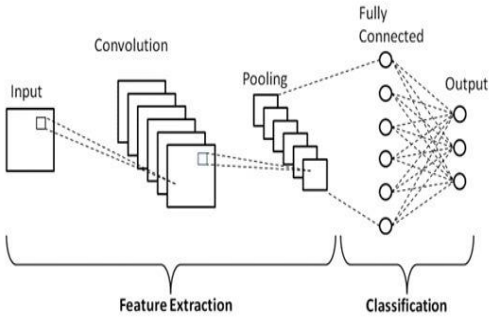


Figure 5: Architecture of convolutional neural network

In this case, the decision function is

$K(x,z)=(x-z+1)^p$ (1)

Traffic Signs are captured and detected and then pre-processed by CNN. After the Detection of traffic signs then Feature Extraction is taken place. In Feature

Extraction, the traffic sign is classified and segmented into multiple pixels. This feature extraction results in the extraction of multiple features of the image. Then the training process is carried on by input image pixels which can differentiate from each other. Clustering is a process that segregates different objects in such a way similar objects are placed into similar groups. The Classification of traffic signs is filtered based upon the size and shape of the image. Filtration is a process to detect the edge of image blur image and a shaded image is detected with the kernel sliding in the extracted pixel of the image. Normally a Convolution Neural Network recognizes this field and learns everything by the basic shapes by evolving the many features in the training process. The CNN learns everything in such a way that this can distinguish from one sign to the other sign. This Max-pooling is a technique that decreases the density and to classify the respective sign. In the detection of traffic Sign image, the Gaussian technique is introduced to reduce the noise.

Data classifier

To classify the candidate traffic signs images, learning algorithms will be used. For this study, learning algorithms are implemented using in python. The classifiers to be evaluated based on accuracy and processing speed for traffic sign recognition. The fig 6 shows the classification of phase.

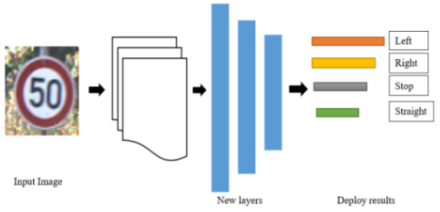


Figure 6: Classification of phase

Parametric Evaluation

For the system level testing, the best detection and recognition models based on the evaluation will be integrated to create the entire model of the system. Accuracy and speed of detection will be tested using actual image taken in local road. System level accuracy will be computed to the product of detection accuracy and recognition accuracy using equation 2.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \qquad (2)$$

A successful recognition means that the system has provided a voice alert about the detected traffic sign. A text-to-speech python module will be used for this function and the corresponding voice alert for each traffic sign detected is shows in table 1. A text annotation will also be generated to the display.

Traffic sign	Text Annotation
Left Turn Prohibition	No Left Turn
Right Turn Prohibition	No Right Turn

U- Turn Prohibition	No U-Turn
Overtaking Prohibition	No Overtaking
30kph speed limit	30kph
60kph speed limit	60kph

Table 1: Text Annotation for Target Traffic Signs

Arduino

The Arduino Uno is a microcontroller board based on the ATmega328. It has 14 digital input/output pins (of which 6 can be used as PWM outputs), 6 analog inputs, a 16 MHz crystal oscillator, a USB connection, a power jack, an ICSP header, and a reset button. It contains everything needed to support the microcontroller; simply connect it to a computer with a USB cable or power it with a AC-to-DC adapter or battery to get started.

IR sensor

An infrared sensor is an electronic device that emits in order to sense some aspects of the surroundings. An IR sensor can measure the heat of an object as well as detects the motion.

LCD Display

Liquid Crystal Display (LCD) screen is an electronic display module. A 16x2 LCD have two display lines each capable of displaying 16 characters. This LCD has Command and Data registers. The command register stores command instructions given to the LCD while the Data register stores the data to be displayed by the LCD.

Buzzer

The click, beep or ring can indicate that a button has been pressed. A buzzer takes some sort of input and emits a sound in response to it. They may use various means to produce the sound; everything from metal clappers to electromechanical devices.

Wifi module

To transmit the measured value or monitored from remote location to the another location a network of internet is required. To enable this a Wi-Fi module is used in the system. There are different modules available but here we are using the module ESP8266. For the interaction of internet with the system using a Wi-Fi module which is connected to the microcontroller through the serial port. So, the measured data is sent from the module to any location with its range from the data can be fetched using a laptop /mobile. A tested this system at various places.

Gear Motor

A gear motor is a specific type of electrical motor that is designed to produce high torque while maintaining a low horsepower, or low speed, motor output. Gear motors are commonly used in devices such as can openers, garage door openers and even electric alarm clocks. Gear motors are primarily used to reduce speed in a series of gears, which in turn creates more torque.

Material And Method

Mask R-CNN

Mask R-CNN is basically an extension of faster R-CNN. Faster R-CNN is an object detection algorithm. Mask RCNN, is a Convolutional Neural Network (CNN) and state-of-the-art in terms of image segmentation and instance segmentation. Mask R-CNN was developed on top of Faster R-CNN, a Region-Based Convolutional Neural Network. In the training dataset, if pixel-level positions of object are also labeled on images, the *mask R-CNN* can effectively leverage such detailed labels to further improve the accuracy of object detection

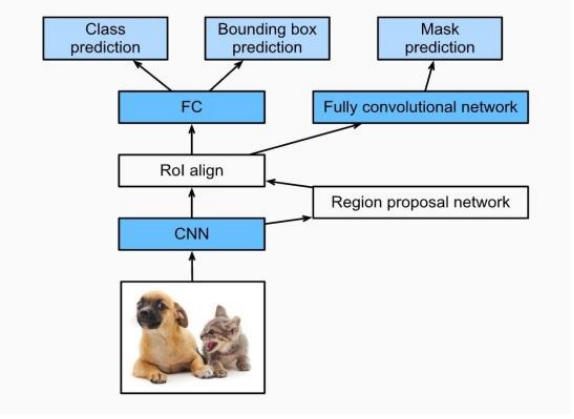
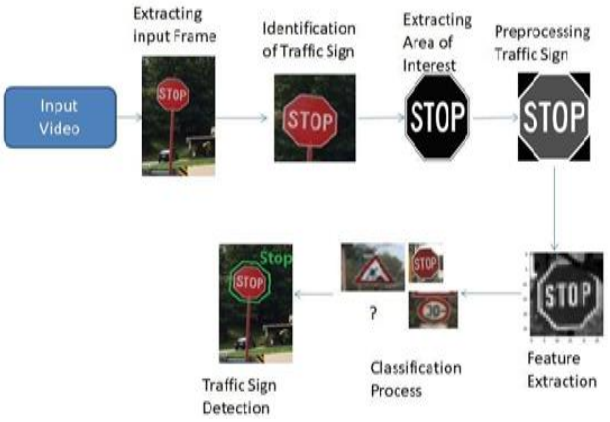


Figure 7: Architecture of Mask R-CNN

The mask R-CNN is modified based on the faster R-CNN. Specifically, the mask R-CNN replaces the region of interest pooling layer with the *region of interest (RoI) alignment* layer. This region of interest alignment layer uses bilinear interpolation to preserve the spatial information on the feature maps, which is more suitable for pixel-level prediction. The output of this layer contains feature maps of the same shape for all the regions of interest. They are used to predict not only the class and bounding box for each region of interest, but also the pixel-level position of the object through an additional fully convolutional network. More details on using a fully convolutional network to predict pixel-level semantics of an image will be provided in subsequent sections.

Sample output of traffic sign detection

Traffic Sign detection is the detection of traffic signs that takes place within the autonomous smart vehicle itself. The detection of traffic sign is taken place as below.



Sample output image of traffic

The smart vehicle can be equipped with the software. The above diagram indicates the proposed work of the Traffic sign detection system.
The procedure of this work is implemented as follows:

- First, take a sample image as an input here, then the first detection of traffic sign in a image is taken place with the help of object detection Even though it's in any shape and size no matter it detects and this fed to Again traffic sign classification Process
- The detect traffic sign is fed to the Feature Engineering Process that divides the whole image into a huge number of pixels and then extracts the Features and then clustering and Filtration operation process is also performed.
- By the above process, the Image is then extracted and detected and then the image is compared with the testing process which is performed in the traffic sign Classification.
- The label map is done to check with the testing feature to verify the predicted result is correct and after verification, this displays the result.
- The Traffic Sign Detection Process is performed with the help of a dataset that gets trained and tested by the CNN in the Traffic sign Classification Process.

Results and Discussion

Traffic sign recognition system recognizes those signs placed in front of the camera by bounding it with a black box and displaying the name of the give sign it. It also provides voice output of each recognized sign. Traffic signs using convolutional neural network this will be done using OpenCV in real time using a simple webcam. CNNs have been gaining popularity in the past of years due to

their ability to generalize and classify the data with high accuracy. In this image train the traffic signs with the images of 43 different classes with the help of TensorFlow.

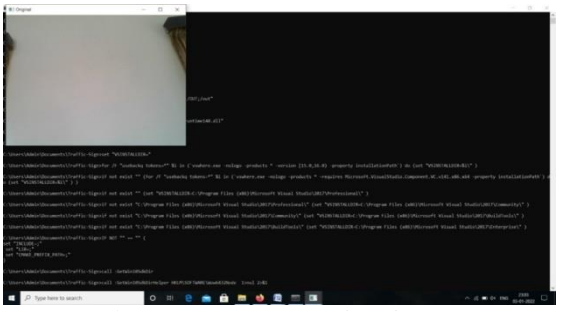


Figure 8: Output of webcam

The above figure 8 shows the Output of Traffic sign Detection and recognition in webcam is entered to the command page. The traffic sign is recognition is performed from a real time webcam camera with the display the word.

Figure 9: Speed limit 40km

The proposed project consists of 16x2 LCD display to showcase the speed of each vehicle passing through. A buzzer will been when a vehicle is passed and the speed of each vehicle will be displayed on the LCD display. When a vehicle is going beyond the speed limit the buzzer will been continuously and speed of vehicle will be shown on the display.

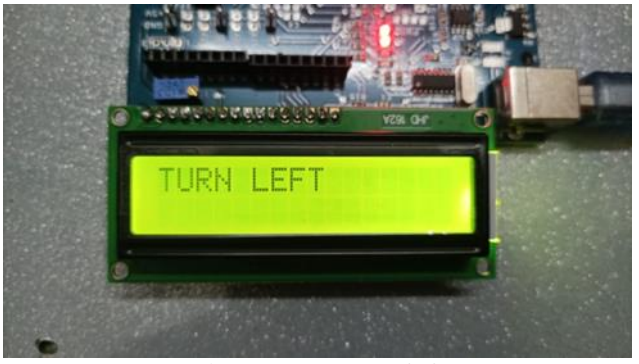


Figure 10: Turn left

The Arduino based traffic sign shows the LCD display with sign of turn left.

Conclusion

A neural network can learn independently as well as generate results that aren't limited to the input. Since it stores input in its own network, rather than a database, data loss has no effect on its operation. Instead of using a pre-defined neural network model, created the own. That model was more efficient because of the torch library instead of any other library. Using the torch technique are detecting live streams, the recognition process is faster and easier. They have used the neural network for traffic sign recognition due to its state-of-the-art accuracy. Therefore, combining computer vision and deep learning to develop a real time traffic sign recognition system. In this model, a voice alert signals the driver when the sign is detected. With this system, the driver never misses a traffic signal because a person receives an alert before crossing the sign. Thus, the driver is more likely to drive on the road safely. Furthermore, it allows the driver to stay within the traffic laws.

Future work

The traffic signs center around decrease of the traffic load on existing street organization through different travel interest the board measures. Traffic signs should eliminate the infringements, clog and get to the next level the traffic light, street condition and geometrics highlights at crossing points. The traffic signs ought to be as a direction or on the other hand speaker on a street organization. Traffic sign diminish the gridlock along the street and furthermore give offices for the street clients. Street signs inform street clients of guidelines also give cautioning and direction required for protected, uniform and productive activity.

References

1. KunalDahiya, Dinesh Singh, C Krishna Mohan , July (2016) 'Automatic detection of bike riders without helmet using surveillance videos in real-time', International joint conference on neural networks (IJCNN), pp.345-453
2. H. N. Do, M.-T. Vo, H. Q. Luong, A. H. Nguyen, K. Trang, and L. T. Vu,(2017) "Speed limit traffic sign detection and recognition based on support vector machines," in 2017 International Conference on Advanced Technologies for Communications (ATC). IEEE , pp.274–278.
3. Sri Uthra V, Sariga Devi V, Vaishali K S, Padma Priya S, Feb (2020) 'Helmet Violation using Deep Learning', International Research Journal of Engineering and Technology (IRJET), vol 07, pp3091-3095,
4. Prajwal M J, Tejas K B, VarshadV, Mahesh MadivalappaMurgod, Shashidhar R, 'Detection of NonHelmet Riders and Extraction of License Plate Number using Yolo v2 and OCR Method', Dec (2019) International Journal of Innovative Technology and Exploring Engineering (IJITEE), vol 09, pp 5167-5172, Dec2019
5. K C Dharma Raj, AphinyaChairat, VasanTimtong, Matthew N Dailey, MongkolEkpanyapong, Jan (2018) 'Helmet Violation processing using Deep Learning', International Workshop on Advanced Image Technology(IWAIT),

6. Edison A. Roxas, Joshua N. Acilo, Ryan Rhay P. Vicerra, Elmer P. Dadios, Argel A. Bandala, "Vision-based traffic sign compliance evaluation using convolutional neural network", Applied System Invention (ICASI) 2018 IEEE International Conference on, pp. 120-123, 2018.R.
7. Susilo, C. B., Jayanto, I., & Kusumawaty, I. (2021). Understanding digital technology trends in healthcare and preventive strategy. *International Journal of Health & Medical Sciences*, *4*(3), 347-354. https://doi.org/10.31295/ijhms.v4n3.1769
8. Widana, I.K., Sumetri, N.W., Sutapa, I.K., Suryasa, W. (2021). Anthropometric measures for better cardiovascular and musculoskeletal health. *Computer Applications in Engineering Education*, *29*(3), 550–561. https://doi.org/10.1002/cae.22202