

Traffic Sign Detection and Recognition

Anju C P

Computer Science and Engineering
Sahrdaya College of Engineering and
Technology,
Thrissur, Kerala, India

Andria Joy

Computer Science and Engineering
Sahrdaya College of Engineering and
Technology,
Thrissur, Kerala, India

Haritha Ashok

Computer Science and Engineering
Sahrdaya College of Engineering and
Technology,
Thrissur, Kerala, India

Joseph Ronald Pious

Computer Science and Engineering
Sahrdaya College of Engineering and Technology,
Thrissur, Kerala, India

Livya George

Assistant Professor
Computer Science and Engineering
Sahrdaya College of Engineering and Technology,
Thrissur, Kerala, India

Abstract:- As placement of traffic sign board do not follow any international standard, it may be difficult for non-local residents to recognize and infer the signs easily. So, this project mainly focuses on demonstrating a system that can help facilitate this inconvenience. This can be achieved by interpreting the traffic sign as a voice note in the user's preferred language. Therefore, the whole process involves detecting the traffic sign, detecting textual data if any with the help of available datasets and then processing it into an audio as the output to the user in his/her preferred language. The proposed system not only tackles the above-mentioned problem, but also to an extent ensures safer driving by reducing accidents through conveying the traffic signs properly. The techniques used to implement the system include digital image processing, natural language processing and machine learning concepts. The implementation of the system includes three major steps which are detection of traffic sign from a captured traffic scene, classification of traffic signs and finally conversion of classified traffic signs to audio message.

I. INTRODUCTION

This section gives an overview about the objective of the project. The objective of an automatic road signs recognition system is to detect, classify the road signs within the images captured by a camera and convert any text to audio if any into the user specified language. The basic steps involved are: Detection, Recognition and Natural language processing. Where detection implies the process in which we will be able to detect the traffic sign images from the scene image captured by the camera. The detector will check whether there present any objects which are most similar to the traffic signs which are given as an input dataset. The next process is the process of recognition which implies the classification of traffic sign and to identify which sign is present in the road. The detected image from the detector is given to the classifier for recognizing the traffic sign. Thus it produces the output which gives the details of the traffic sign that we have detected from the traffic scene image. This textual information is then given to the natural language processing section which will help us to convert the textual data into an audio message. This audio formatted content is given as output to the user as an audio message. The user can specify

the language that he wants to hear or that he can understand.

II. DRAWBACKS OF EXISTING SYSTEM

In order to solve the concerns over road and transportation safety, automatic traffic sign detection and recognition (TSDR) system has been introduced. An automatic Traffic Sign Detection and Recognition system can detect and recognize traffic signs from and within images captured by cameras or imaging sensors. In adverse traffic conditions, the driver may not notice traffic signs, which may cause accidents. In such scenarios, the Traffic Sign Detection and Recognition system comes into action. The main objective of the research on Traffic Sign Detection and Recognition is to improve the robustness and efficiency of the Traffic Sign Detection and Recognition system. To develop an automatic Traffic Sign Detection and Recognition system is a tedious job given the continuous changes in the environment and lighting conditions. There are diverse algorithms for traffic-sign recognition. Typical sign board shapes like hexagons, circles and rectangles define different signs, which can be used for classification. Other major algorithms for character recognition include Haar-like features, Freeman Chaincode, AdaBoost detection and deep learning neural networks methods. Haar-like features can be used to create cascaded classifiers which can then help detect the sign board characters.

Therefore, the major drawbacks of the existing system are

- Lack of efficiency
- Storage consuming
- Complicated algorithm

III. METHOD

The system we are introducing makes use of various domains namely digital image processing, deep learning, transfer learning, natural language processing. The whole system can be broadly categorized as two parts i.e., the detection part and the classification part. The detection part is used to detect the traffic sign from the video clip of either sides of the road captured with help of a camera while travelling. For detection we are using the German Traffic

Sign Detection Benchmark (GTSDB). The GTSDB is a single-image detection assessment with 900 images containing zero to six traffic signs. There is a general division of traffic signs into three categories namely prohibitive, danger and mandatory. The sizes of the traffic signs can vary from 16x16 to 128x128. Traffic signs may appear under any lighting condition. The classification part is used to classify the detected traffic sign as what it means. Traffic sign classification is the process of automatically recognizing traffic signs along the road, like speed limit signs, yield signs, merge signs, etc. The dataset we'll be using to train our own traffic sign classifier is the German Traffic Sign Recognition Benchmark (GTSRB). The GTSRB dataset consists of 43 traffic sign classes and 50,000 images. The traffic signs have been pre-cropped for us, implying that the dataset annotators have manually labelled the signs in the images and extracted the traffic sign Region of Interest (ROI) for us, thereby simplifying the project.

The toolkit we are using for the detection of the traffic sign is the Tensorflow Object Detection API provided by Google. The method we are using is the transfer learning technique. It is a deep learning approach where trained models are used as the starting point on computer vision. The model we are using is the SSD (Single Shot Detector) MobileNet v2. This model is part of the Tensorflow object detection API. There are many models available with the Tensorflow Object Detection API toolkit. But we have selected this model as it can be easily executed in our system with a better accuracy. In order to use this model, we have to prepare our dataset which is in the form of XML and image formats. These are loaded and converted to as CSV file formats and as a mapping from XML to images. With this CSV we make the image and label into a TFRecord format which I going to be the data for transfer learning i.e., this TFRecord data is going to be the input to our model. The model will have 30000 epochs used which means the dataset will be passed through the model 30000 times. Before passing, the dataset will be divided into batches of 64. Every features/parameter are used same as that used in the model except for the image augmentation; the image is resized and contrast is varied. To get the accuracy, test.py file from the Object Detection API is executed. 80% of the dataset I used for training and the rest 20% I used for testing. After executing the test.py file we've got the accuracy of the model as 74.28%. After the completion of training of the model we get the trained model saved in the '.pb' file format and a label mapping.

Traffic sign detection and recognition are just one of the problems that computer vision and deep learning can solve. In the real-world, traffic sign recognition is a two-stage process, namely localization and recognition. In the localization process, the traffic sign is detected and localized from an input image/frame. In the recognition process the localized ROI is taken and actually recognized and classifies the traffic sign into any of the 43 classes.

Deep learning object detectors have a speciality that they can perform localization and recognition in a single forward-pass of the network. The challenges in the GTSRB dataset are, the first being that images are low resolution, and worse, have poor contrast. These images are pixelated, and in some cases, it's extremely challenging, if not impossible, for the human eye and brain to recognize the sign. The second challenge with the dataset is handling class skew i.e., the top class (Speed limit 50km/h) has over 2,000 examples while the least represented class (Speed limit 20km/h) has under 200 examples — that's an order of magnitude difference. In order to successfully train an accurate traffic sign classifier, we'll need to devise an experiment that can pre-process our input images to improve contrast and account for class label skew. Recognizing the traffic signs carried out using the Keras. To classify and recognize traffic signs, we need to implement a Convolutional Neural Network. Here the classifier is named as TrafficSignNet and also defined a Convolution-RELU-Batch Normalization-Pool layer set which uses a 5x5 kernel to learn larger features which in turn will help to distinguish between different traffic sign shapes and colour blobs on the traffic signs themselves. These sets of layers deepen the network by stacking two sets of Convolution-RELU-BatchNormalization layers before applying max-pooling to reduce volume dimensionality. The head of our network consists of two sets of fully connected layers and a softmax classifier. Dropout is applied as a form of regularization which aims to prevent overfitting. The result is often a more generalizable model. This concludes the TrafficSignNet architecture. Python training script is for loading our training and testing split from the GTSRB dataset, pre-processing the images, training our model, evaluating our model's accuracy and serializing the model to disk so we can later use it to make predictions on new traffic sign images. As mentioned earlier, one of the challenges with the GTSRB dataset was that many images have low contrast, making it challenging for the human eye to recognize a given sign. We can automatically enhance image contrast by applying Contrast Limited Adaptive Histogram Equalization (CLAHE). CLAHE is an algorithm which can be implemented using the scikit-image library. Once we have done with the loading the dataset, pre-processing, training and evaluating the accuracy, we obtained 95% accuracy. After the model is trained, we could easily implement prediction. All we need to do is to load the model and then sample images from disk, pre-process these sample images in the same manner as we did for training, pass our images through our traffic sign classifier and obtain our final output predictions. The voice output was interpreted with the help of an API, the Google Text to Speech API commonly known as the gTTS API.

Fig 1 shows the architectural diagram of the proposed method. In this initially a traffic sign is captured which will further undergo through the pre-processing stage, detection phase, classification phase and finally an audio message is generated which denotes the classified traffic sign.

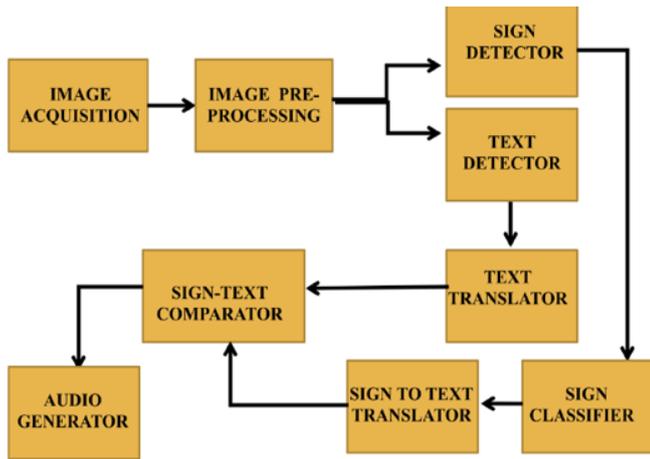


Fig 1:- Architectural Diagram

The different modules of traffic sign detection system are image acquisition, image pre-processing, sign detector, text detector, text translator, sign-text comparator, sign classifier, sign to text translator, audio generator. The brief descriptions about these modules are given below:

A. Image Acquisition

Image acquisition can be defined as the process of retrieving/obtaining an image using some source usually it may be hardware based source. In this stage with the help of camera the traffic scenes are obtained.

B. Image Pre-processing

Digital image processing is the process of using algorithms to perform image processing on digital images. As a subfield of digital signal processing, digital image processing has many advantages over analogue image processing. It allows a much wider range of algorithms to be applied to the input data. The aim of digital image processing is to improve the image data by suppressing unwanted distortions and/or enhancement of some important image features so that our AI-Computer Vision models can benefit from this improved data. In this proposed system CLAHE algorithm is used for the pre-processing of image. It will help to improve the contrast in images.

C. Sign Detector

The pre-processed traffic scene image is given to the sign detector algorithm which will detect whether any traffic signs are present in the captured image or not. If it detects any traffic sign then the sign will be fed into the classifier for further classification process. It is implemented using transfer learning method in google's object detection API toolkit.

D. Sign Classifier

The sign classifier will classify the traffic signs into appropriate classes. There are 43 different classes in the dataset we are using, i.e., GTSRB for classification. In this proposed method a convolution neural network (CNN) is used for the purpose of classification of traffic signs.

E. Text Detector

Text detector uses the concept of OCR (optical character recognition). OCR is a reader which recognizes the text character of the computer which may be in printed or written form. Along with the sign detector we use text detector for obtaining the text written on the sign board, if any.

F. Text Translator

It converts the text to natural language specified by the user. This also uses the concept of OCR. OCR is a reader which recognizes the text character of the computer which may be in printed or written form.

G. Sign to Text Translator

It is used to translate the classified sign to corresponding textual data. In this system we are planning to take the dataset that has labels as textual representation of the classified images. Input is obtained from the sign classifier and the output is fed to the sign text comparator.

H. Sign-Text Comparator

The sign-text comparator module of the proposed system architecture is intended for the purpose of comparing the text obtained from the text translator module and the text obtained from the sign translator module. This helps in attaining better accuracy in our proposed system for traffic sign detection, its recognition and conversion of the same into natural language audio output. Hence it can be used as an end-to-end application.

I. Audio Generator

It uses the concept of Natural language processing. The user can start the proposed system by saying "START" and "END" to stop the system. NLP is a subset of linguistics, computer science and information science, information engineering, and artificial intelligence which include how to program computers and analyze large amounts of data.

Fig 2 shows the use case diagram of the proposed system in which the user can specify the language based on their wish. Such that user will be able to get the audio message in their own language. The rest of the processes such as image processing, detection of traffic sign, classification and conversion of the text to audio message are performed by the proposed machine. The image can be captured and given to the detector with the help of a high quality camera.

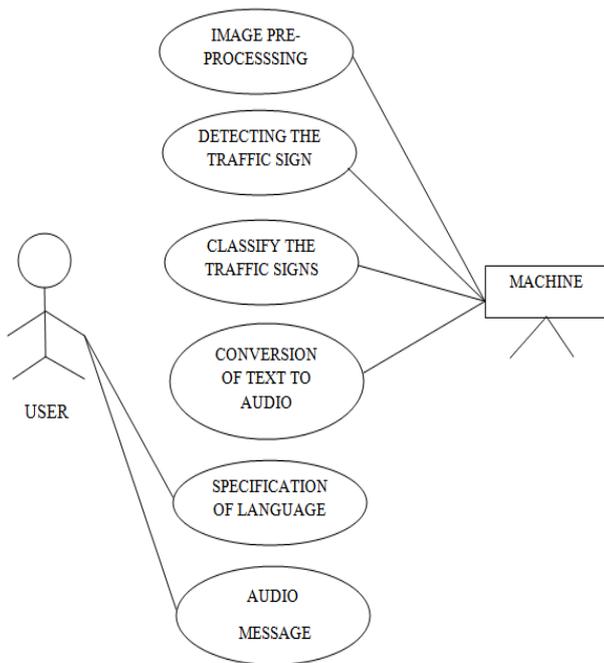


Fig 2:- Use Case Diagram

The use case provides us with a clear look on how the users perform different tasks. As you can see here, the user is only involved with providing the language preference. And all the other tasks don't have anything to do with the user. Image processing, traffic sign detection, classifying the detected sign, conversion of text to audio are all done by the system. And the output will be produced by the system to the user's speaker.

IV. RESULTS

The traffic sign detection and recognition system solve the problem of identifying and interpreting the traffic sign set on either side of the road. Through this project we were able to successfully carry out the process of identifying and interpreting the traffic sign and depicting it as a voice message while driving.

While traveling, the image was successfully captured by the attached camera module. This captured image was then fed to the traffic sign detection algorithm. The detection algorithm successfully detected the traffic sign present, which was done using the TensorFlow API. The frame which contains this was then fed to the classification algorithm. The classifier algorithm then classified the traffic sign present in the frame passed into one of the 43 classes and the name of the traffic sign was interpreted as a voice using the google text to speech API. For detection part we got an accuracy of 74.8% and the classification portion showed a 94.5% accuracy.

V. CONCLUSION

Accidents, due to disregarding the traffic sign are increasing in an alarming rate. Hence its control has to ensure in order to save the human lives. We produce a system that could reduce the accidents caused due to negligence of the traffic sign boards. The system consists of an image processor, classifier, detector and finally to convert the text to audio form. The image is captured by the camera, the traffic sign is being extracted, it is detected from the datasets, and if the traffic sign boards contain any text, it being converted into the audio according to the user's specification. The machine developed for the detection and classification of the traffic signs board have given encouraging results. German datasets are used for the classification via the faster R-CNN which fine tunes the image. We believe that our proposed system can reduce the rate of deaths that are caused due to the traffic sign boards and thereby make it easier in the understanding of the sign boards irrespective of the location.

ACKNOWLEDGEMENT

We would like to express our immense gratitude and profound thanks to all those who helped us to make this project a great success. We express our gratitude to the almighty God for all the blessings endowed on us. We express our thanks to our Executive Director REV. FR. GEORGE PAREMAN, Director Dr. ELIZABETH ELIAS Principal Dr. NIXON KURUVILA for providing us with such a great opportunity. We are thankful for the help and appreciation we received from head of the department Dr. RAJESWARI M., project coordinators Ms. DEEPA DEVASSY, Ms. ANLY ANTONY and Mr. WILLSON JOSEPH. We would also extend our deep sense of gratitude to our project guide Ms. LIVYA GEORGE for their guidance and advice. We would like to express our gratitude towards our parents for their timely co-operation and encouragement. Every project is successful due to the effort of many people. Our thanks and appreciations go to all our peers who had given us their valuable advice and support and pushed us into successfully completing this project.

REFERENCES

- [1]. S. Houben, J. Stallkamp, J. Salmen, M. Schlipfing, and C. Igel, "Detection of traffic signs in real-world images: The German traffic sign detection benchmark," in Proc. IEEE Int. Joint Conf. Neural Netw., Aug. 2013, pp. 1–8.
- [2]. J. Stallkamp, M. Schlipfing, J. Salmen, and C. Igel, "The German traffic sign recognition benchmark: A multi-class classification competition," in Proc. IEEE Int. Joint Conf. Neural Netw., Aug. 2011, pp. 1453–1460.
- [3]. Y. Yang, H. Luo, H. Xu, and F. Wu, "Towards real-time traffic sign detection and classification," IEEE Trans. Intell. Transp. Syst., vol. 17, no. 7, pp. 2022–2031, Jul. 2016.