INTI JOURNAL | Vol.2024:08
eISSN:2600-7320

Traffic Sign Board Recognition and Voice Alert System using CNN

Yogesh C.M[1], Usha Sree R.[2], Hushalictmy.P.[3]

[1,2]Dayananda Sagar Academy of Technology and Management, Karnataka, India
[2]Faculty of Data Science and IT, INTI International University, 71800 Nilai, Negeri Sembilan, Malaysia

***Email:** [1]yogeshyogimca@gmail.com, [2]ushashree-mca@dsatm.edu.in, [3]husha.palliyanny@newinti.edu.my

Abstract

Street to guarantee a secure and efficient flow of traffic. Street accidents sometimes occur on account of carelessness in reading traffic signs incorrectly. The suggestedframework aids in recognizing the stop sign and giving a voice warning to the motorist for the speaker to make their point, and crucial decisions. The proposed framework is prepared Using a Convolutional Neural Network (CNN), which aids with the recognition and arranging of rush hour congestion sign pictures. To increase precision, a number are of classes generatedand characterized on a particular dataset. Utilized was the German Traffic Sign Benchmarks Dataset, which includes 51,900 pictures of road signage in 43 classifications. Around 98.52 percent during execution was precise. After the framework recognizes the sign, the driver is informed through a voice alarm issued through the speaker. The suggested framework also includes a section where drivers are warned about nearby traffic signs so they can keep track of which laws to follow while on a highway. The system's goal is to protect the driver, passengers, and pedestrians from harm.

Keywords

Convolutional Neural Network, Object Detection, Object Classification Traffic, Traffic Signs, Voice Alert

Introduction

There have been several technical advancements and vehicles that auto-pilot mode have come up. Independent vehicles have appeared. There has been a blast in oneself driving vehicle industry. Be that as it may, these features are only available in a small number of expensive, high-quality cars that are out of reach for most people. We needed to create a framework to make the process of marginally leading easier. Managing a review we observed that the extentof street mishaps in India is disturbing. Reports propose that there are around 53 an hour. setbacks occurring on the streets. Additionally, consistently more than 16 passings happen because of these disasters [1]. Whenever somebody forgets to comply with traffic signs whiledriving, they are putting their life also the existence of different drivers, their travelers and those out and about in danger. Subsequently, we thought of this framework in which traffic signs are naturally distinguished utilizing the live video transfer and are perused out resoundingly to the driver who

Submission: 4 May 2024; **Acceptance:** 26 June 2024

then might make the appropriate decision. Getting the area thatthe client is utilizing is another area on which our system focuses.

GPS. Likewise, every traffic sign will be kept in an information base next to their designated locations so that drivers are aware of the upcoming traffic signs. Coming up next isa breakdown of the paper's design: Section II sums up the writing survey; Section III makes sense of the procedure and how the models work; and Section IV shows the Results and Analysis. The paper's decision is in Section V, while the Future Scope is in Section VI.

Literature Survey

In this time frame of a fast-paced life, individuals for the most part will generally pass up being aware of the stop sign, and consequently defy the guidelines. A ton of exploration has been done in this space to diminish the quantity of mishaps. Specialists have utilized an assortment of arrangement calculations and various CNN designs to group the traffic signs and caution the driver. Our framework means to improve the course of acknowledgment and simultaneously give different advantages like early awareness of the driver. The discovery of traffic signs has beendone in an assortment of procedures in various examinations. [8] One of the cycles utilizes theSupport Vector Machine procedure. For reasons of preparation and testing, the dataset was divided into 90/10 groups, and it utilizes direct characterization. To accomplish the ideal outcome, a progression of stages called Following color segmentation was shape classification and recognition. Raspberry Pi is utilized in distinguishing and perceiving TrafficSigns with considerably less coding [2]. Be that as it may, it requires the Raspberry Pi board atone's talk for execution which is very exorbitant. One more method of Traffic sign acknowledgment is picture serious [3]. A video is gained and separated into outlines. Picture preprocessing is done which incorporates isolating the frontal area and the foundation, diminishing and contrast enhancement. The signs are hexagonal three-sided, or round healthy, and sent for layout matching after these activities. The articles with some unequivocal shapes are matched from the pre-trained calculation. Caffe is an open-source framework, that assists with distinguishing and perceiving street traffic signs with high precisionand efficiency [4]. A CNN approach is proposed for preparing traffic sign preparation sets andgetting a model that can arrange traffic signs. One more technique for utilizing the CNN plot isproposed in [8], in which the genuine line of the objective sign is assessed by extending the limit of a comparing layout sign picture into the info picture plane. The technique advances tobecome start-to-finish teachable when we change the limit assessment issue into a posture andshape forecast work given CNN. It is more impervious to impediments and limited targetsthan other limited assessment methods that emphasize form assessment or picture division. [6] proposes a multi-goal including combination network engineering for sign identification, whichhelps with the partition of various little items from sign sheets. An upward spatial arrangementconsideration (VSSA) module can likewise be utilized to accumulate additional setting data forfurther developed recognition. Utilizing GPS-based following, Augmented Reality innovationis consolidated in portable applications [5]. It utilizes the directions of a client's cell phone as a pointer to help individuals progressively find potential assets in the prompt area in light of the course of the client's camera view. In [7] AlexNet construction of CNN is utilized in which the design contains eight layers. The initial convolutional layers

total five and the last three are completely associated layers. The exactness of this design emerges to be 92.63%. Additionally, the Google Net design is carried out in [7] which helps in working with huge information and countless boundaries. Anyway, it accompanies an issue that the huge information makes network overfitting diminishing the exactness by simply 80.5%. VGGCNN is proposed in [8]

Methodology

Dataset

In the proposed system, the German Traffic Sign Benchmarks (GTSRB) Dataset is used. Fig. 1 shows the 43 different traffic signs that are considered to train the model. It has 51,900 single images distributed among the 43 classes including the training and the test dataset. The count of the number of photos per class is shown in Fig. 2. There is no ambiguity as the images are just focused on the traffic signs and each of them is unique. The training dataset has different folders for each of the present classes. A CSV file is also present wherein the path of each image and its class and other details such as width, and height are mentioned.

Fig.1. Traffic Signs Taken into consideration

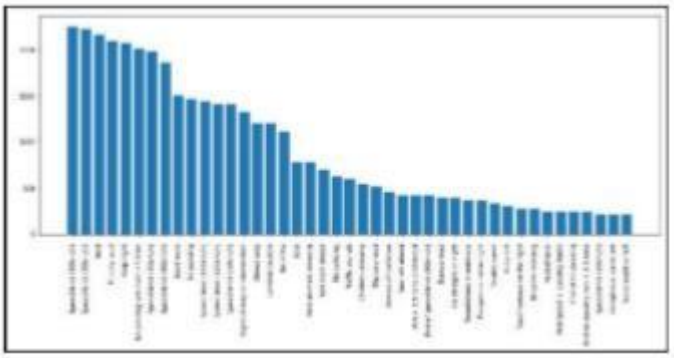

Fig.2. Number of images per class in the dataset

Information Preprocessing:

To perform picture handling, pictures should be changed over into number exhibits (for example numeric qualities). After stacking the pictures, they are resized to 30*30 pixels. Post this, the marks of the picture are planned with the picture and subsequently, the dataset is fit tobe prepared.

Model

Convolutional Neural Network (CNN) is a calculation falling in Deep Learning. CNN can snap a photo as info, dole out various things in the image, and recognize themfrom each other. It requires considerably less preprocessing when compared with other order calculations. Convolutional Networks can become familiar with the channels or attributes in thepictures instead of the crude techniques channels where they are done physically. The engineering of a Convolutional Network can measure up to the available example of Neuronsin the Human Brain. The actual plan was propelled by the association of neurons as present inthe Visual Cortex of the human cerebrum. The neurons answer upgrades just in a specific areaof the field of view which is known as the Receptive Field. The visual region is an assortmentof various responsive fields that help us in surveying objects. When the model is preparedover a progression of ages for example cycles, it fosters the capacity to recognize the ruling elements and certain low-level highlights in the pictures. Given this preparation, the modelorders them utilizing the SoftMax Classification procedure. Fig. 3 addresses the quantity of layers utilized in the model. There are 4 convolution layers and 2 max-pooling layers alongsidedropout, straighten, and thick layers. Adam analyzer is utilized in brain organization. The info size of the picture is 30*30*1. The model utilizes the ReLU initiation work. We acquire acompletely associated layer after the Flatten layer. lastly the not entirely settled by utilizing thesoftmax enactment work.

Proposed Solution

Fig. 4 shows the exactness of the prepared organization. This model ended up giving thebest exactness when contrasted with the different models that we examined.

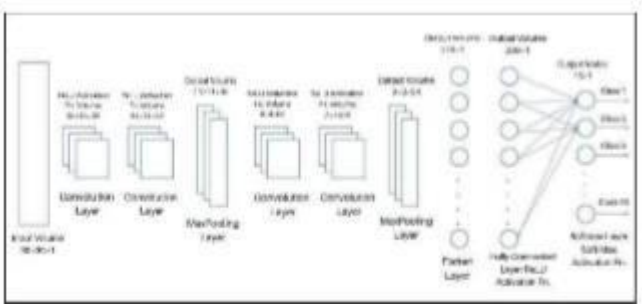

Fig.3. Neural Network Representation

Implementation

After training the model, it is saved and then the saved model is used for prediction. A full-stack web application with NodeJs and Express Handlebars has been developed using this model for prediction. It incorporates different logic to make it a product that can be used with certain improvements in place. Fig. 5 depicts the Flow. The suggested system is depicted in this diagram. The CNN model is applied in the first part wherein the input is an image. After

being processed, one of the classes out of the 43 classes is obtained as the output. If certain image does not contain a traffic sign, then the user gets a prompt of “No Sign Detected”. This is done by analyzing the output array of the "model. predict" function in Python. The "model. predict" function returns an array of values representing how closely the image falls under each of the 43 classes and finally predicts the class based on the highest value.

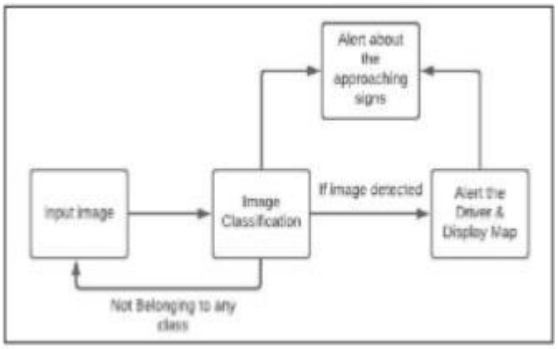

Fig.5. Flow Diagram

After multiple iterations, it was found that even if an image does not fall in any of the given classes, the model, not being trained for an extra other class, classifies it into one of the 43 classes, but the value predicted by the “model. predict” function is pretty low. So the threshold value for separating the images that do not have a traffic sign in them but are predicted as one is taken as 0.68. The value of the classes in "model. predict" is in the range 0 to 1 and hence if the model classifies it in a particular class with a value less than 0.68 it will be identified as none of the above, else it will be assigned a class. Once the image has been classified, the metadata is fetched from the image using “Exif-parser” and then the sign text along with the GPS coordinates are stored in the database.

All this data is then available to theuser on a map. The map has markers containing the latitude and longitude along with the nameof the traffic sign. Another important feature that needs to be highlighted here is that the proposed system aims to alert the drivers. Therefore, rather than just alerting about the signthat the car is approaching i.e. the sign which has been detected, an algorithm in which the traffic signs that are in proximity i.e. the ones that will be approached within the next 5 minutes(or 1 km) are also to be alerted to the driver, is implemented. The computation of this is done by taking into account the locations of the signs which were stored by extracting the metadata. Fig. 6 is a sample test case given to the model and Fig. 7 represents the predicted output whichwill be voiced out to the driver. It also contains a map depicting the location of the various traffic signs in the database.

Fig.6. No Entry Sign (Input)

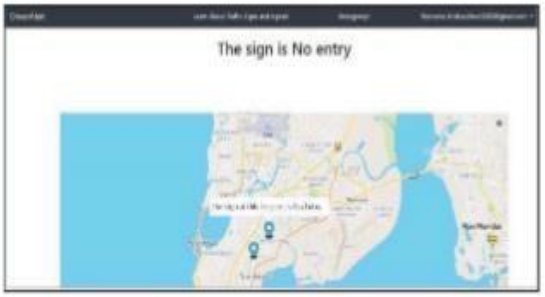

Fig.7. Sign Prediction and Display of Signs on the Map

Results And Discussion

The trained neural network which contains 4 convolution layers and 2 maximum pooling layers as well as dropout, flattened and dense layers proved to give a better result as compared to the other CNN Architectures used in AlexNet, GoogleNet, VSSANet, and VGGNet. As mentioned in Table 1, the precision of the trained network is 98.52%.

Table 1. Accuracy of various models available

Method	Accuracy
AlexNet	92.63%
GoogleNet	80.5%
VSSANet	94.42%
VGGNet	98.03%
Trained Neural Network	98.52%

Acknowledgments

The authors are thankful to Dr Radha Shankarmani for mentoring us throughout the research and development of the Traffic Sign Board Recognition and Voice Alert System using Convolutional Neural Network. It is because to her assistance that the paper has been able to achieve its goal of serving the intended purpose. The authors would like to take this opportunity to thank Sardar Patel Institute of Technology for supplying all of the necessary resources for this study's completion.

References

Bi, Z., Yu, L., Gao, H. et al. Improved VGG model-based efficient traffic sign recognition for safe driving in 5G scenarios. Int. J. Mach. Learn. & Cyber. (2020

Anushree.A., S., Kumar, H., Iram, I., &Divyam, K. (2019). Automatic Signboard Detection System by the Vehicles.

S. Harini, V. Abhiram, R. Hegde, B. D. D. Samarth, S. A. Shreyas and K. H. Gowranga, "A smart driver alert system for vehicle traffic using image detection and recognition technique," 2017 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT), Bangalore, India, 2017, pp. 1540-1543, doi: 10.1109/RTEICT.2017.8256856.

C. Wang, "Research and Application of Traffic Sign Detection and Recognition Based on Deep Learning," 2018 International Conference on Robots & Intelligent System (ICRIS), Changsha, China, 2018, pp. 150-152, doi: 10.1109/ICRIS.2018.00047.

M A Muchtar etal2017J. Phys.: Conf. Ser. 801 012010 [6] Y. Yuan, Z. Xiong and Q. Wang, "VSSA-NET: Vertical Spatial Sequence Attention Network for Traffic Sign Detection," in IEEE Transactions on Image Processing, vol. 28, no. 7, pp. 3423-3434, July 2019, doi: 10.1109/TIP.2019.2896952.

S. Huang, H. Lin and C. Chang, "An in-car camera system for traffic sign detection and recognition," 2017 Joint 17th World Congress of International Fuzzy Systems Association and 9th International Conference on Soft Computing and Intelligent Systems (IFSA-SCIS), Otsu, Japan, 2017,pp. 1-6, doi: 10.1109/IFSA-SCIS.2017.8023239.

Yadav, Shubham & Patwa, Anuj & Rane, Saiprasad&Narvekar, Chhaya. (2019). Indian Traffic Sign Board Recognition and Driver Alert System Using Machine Learning. International Journal of Applied Sciences and Smart Technologies. 1. 1-10. 10.24071/ijasst.v1i1.1843.