Traffic Sign Recognition using Convolutional Neural Network

[1]Dr. Harish B G, [2]Mr. Chetan Kumar G S, [3]Ayesha Siddiqa K, [4]Manvitha M

[1]HOD Professor, [2]Assistant Professor, [3,4]Student
Department of Master of Computer Applications
UBDTCE, Davangere

Abstract: **The traffic signs on the road express a number of cautions. By assisting tourists achieving their final locations and informing them in ahead of time, leave, and turn locations, they maintain traffic flow. To guarantee the safety of drivers, road signs are set in certain locations. Additionally, they offer instructions on when and location cars should either turn or not. In this research, we developed a technique for taking a road sign out of a naturally complicated image, processing it, warning the motorist by voice order. We also proposed a technique for locating and recognizing traffic indicators. It is used in a way that makes quick decisions possible for driver. Traffic sign detection is difficult in actual time due to variables such changing weather, shifting light directions, and varied light intensity. The incomplete noise or complete substantial changes in color saturation, partial or complete underexposure, or full-on overexposure, a wide range of viewing angles, view depth, and traffic sign shape and color distortions (caused by light intensity) are just a few of the factors that can affect a machine's reliability. Three phases make up the suggested architecture. The first step is image pre-processing, where we decide on the learning input size, adjust the data for the learning stage, and size the input files for the dataset. In the course of the recognition process, the suggested algorithm sorts the observed symbol. This is accomplished in the second phase using a Convolutional Neural Network, and in the third step, text speech displacement is dealt with, the recognized evidence of the second phase delivered in audio file.**

Keywords- **Image Preprocessing, Feature Extraction, Segmentation, Data Augmentation, Convolution Neural Network convert text to speech.**

INTRODUCTION

400 traffic accidents are said to occur daily in India, according to official figures. In order to protect both drivers and pedestrians, traffic signs assist prevent accidents from occurring. Additionally, traffic lights lessen the likelihood of traffic violations by requiring road users to adhere to certain laws. Another way to make it easier to find your way is by using traffic signals. Everyone using the road, including vehicles and pedestrians, should put following traffic signals first. We miss traffic arrows for several reasons, including difficulty focusing, tiredness, and lack of sleep. Impaired vision, the effect of the outside world, and environmental conditions are some additional justifications for ignoring the indicators. Utilizing a program that can identify traffic lights and provide the driver with advice and warnings is considerably more crucial. Real-time picture analysis by a car front facing camera is used by image-based traffic-sign recognition algorithms to identify signals. They assist the motorist by issuing warnings. The main elements of a vision- The identification and recognition modules are the foundation of a traffic sign recognition system. The detecting module locates the sign region in the image or video while the recognition module identifies the sign. The sign areas with the highest likelihood are selected during the detection process and delivered to the recognition system for classification. To recognize traffic signs, many machine learning methods can be used, such as SVM, KNN, and Random Forest [6]. However, the fundamental disadvantage of these algorithms is that feature extraction must be done independently; CNN, in contrast, will handle feature extraction that is done alone [1]. So a CNN is used in the suggested way.

RELATED WORK

The most important step in every study is to do a literature review. By doing so, we will be able to spot any weaknesses or holes in the current framework and try to come up with a solution that will work around its drawbacks. In this segment, we briefly address related research on the detecting, understanding, and recognizing of traffic signs. Comparative evaluation of citations is shown Below in Table 1

Wasif Arman Haquea Samin Arefin b , A.S.M.

Paper	Technology/Algorithm	Advantages	Limitation
DeepThin: An innovative, light-weight CNN architecture that doesn't require a GPU for interpreting signals in traffic**Author**- Wasif Arman Haquea ,SaminArefin b , A.S.M. Shihavuddin c, Muhammad Abul Hasan **Year**-2021	Three modules— processing, learning, and forecasting of input make up the proposed Deep Thin architecture. Preprocessing is where images are resized. Learning is accomplished using two convolutional layer, two max-pooling layer, and one complete connectivity concealed layer, and class prediction. of CNN	It may be utilized on a low-end laptop, even without a GPU thanks to the architecture's modest weight. This network optimization provides green features in the solution by reducing the energy usage needs for deep learning testing..	During the detecting phase, just the sign's color characteristic is taken into account. They focused on the sign's RGB and grayscale values.
A powerful CNN for detecting minor traffic signs **Author**-Shijin Songa ,Zhiqiang Que b, JunjieHoua , Sen Dua , YuefengSonga **Year**-2019	Author evaluated R-CNN and Faster R-CNN accuracy while concentrating on small object detection challenges .CNN utilizing fully connected transformation, trimming superfluous layers, and convolution factorization optimize the example.	The model the past improved consuming less computational resources and GPU memory. There is no information on image preprocessing.	Information on image Preprocessing is absent
Author- AashrithVennelakanti, SmritiShreya, ResmiRajendran, Debasis Sarkar, Deepak Muddegowda, PhanishHanagal **Year**-2019	Instead of using RGB, Use of the Hue Saturation Value (HSV) color space, and the Then, shape-based detection is performed using the Douglas method.	To improve accuracy, CNN ensembles are used, and two data sets are evaluated..	However, only triangular and circular forms taken into account for detection, despite the high accuracy.
Large-Scale Traffic-Sign Detection and Recognition Using Deep Erudition **Author**-DomenTabernik; DanijelSkočaj **Year**-2020	Traffic signs are identified and analyzed using the CNN R-CNN mask. DFG traffic Traffic Sign Monitoring and Recognition is the name of the new data gathering they produced. achieving low inter-class and excellent intra-class variability by using a CNN Ensemble-sign.	There has been info augmentation, and Changes are made to segmented, real-world training samples to produce more fictitious traffic-sign instances. Two different kinds of distortions were applied: distortions in appearance (brightness changes) and geometric/shape distortions (perspective shifts, color shifts).	Miss detection of sign scenario.
Using Multi-Class SVM and the HougTransformation, speed limit road signs may be recognized. **Author**-Ivona Matoš; Zdravko Krpić; Krešimir Romić **Year**-2019	HOG descriptor is utilized for feature extraction, and SVM is used for classification.	Images that have a lot of noise handled successfully, and performance of up to 95% was attained.	The suggested method's domain only encompasses circular signs.

Shihavuddin c Muhammad Abul Hasan [1] describe the "a brand-new, light-weight CNN architecture that does not require a GPU for traffic sign recognition. The major challenges in identifying traffic signals in real-world circumstances that the author focused on were visual distortion, speed, motion effects, noise, and faded sign colors. Grayscale image-only training yields average accuracy. The writers then proposed the Deep Thin building, which consists of three modules for input preparation, research, and prediction. Architecture is both thick and shallow at the same time. Since only a few feature maps were employed per layer, they were both thin and deep due to the utilization of four layers. Furthermore, training without a GPU is now possible due to their consideration of small input images, a limited number of feature maps, and considerable convolutional advances. Utilizing stride convolution and overlapping max pooling judiciously sped up training and decreased overfitting problems. Data improvement is utilized to achieve resilience. They applied augmentation techniques such original random shearing of practice photos, moving horizontally and vertically, as well as zooming in and out. The tests employ the German language the Traffic Sign Recognition Benchmark and the Belgian Traffic Sign Classification dataset. Modifications are made to the hyperparameters, feature map, and kernel size.

1. **AashrithVennelakanti, Smriti Shreya, ResmiRajendran, Debasis Sarkar, Deepak Muddegowda, PhanishHanagal [3]** give a description of "Traffic Sign Detection and Recognition Using a CNN Ensemble". Using data from Belgium and the German Traffic Sign Benchmark, the method that is being offered in this study is tested in two modules: detection and recognition.. CNN ensemble is used to name the discovered sign after the detection stage, which involves photographing the traffic sign and identifying the object from the image, has been completed. HSV model is used in the initial phase rather than RGB color space since it has a

broader variety of colors and is more in accordance with how the human eye analyzes images. Following that, color-based detection and shape-based detection are implemented. In color-based detection, the red values of the sign are checked to see whether they go below a predetermined threshold; if they do, that portion is analyzed to determine whether the sign is there or not. The Douglas conduct is then used to accomplish shape-based detection. The circle and triangle were the only two shapes on which the authors concentrated. After this method selected a region based on the number of edges present in the image, ROI was divided using bounding boxes. An inversion filter and picture thresholding are now verify the sign inside the bounding box. In the second stage, a feed-forward CNN network with six convolutional layers is used to classify the detected sign. When using

2. **Do men Tuberin; DanijelSkočaj [4]** describe the "Deep Learning for Large-Scale The circle and triangle were the only two shapes on which the authors concentrated.After this method selected a region based on the number of edges present in the image, ROI was divided using bounding boxes.The sign inside the bounding box is now confirmed using picture thresholding and an inversion filter. In the second stage, a forward-looking To categorize the detected sign, a CNN network with six convolutional layers is utilized. Utilizing the combined output of three CNNs, they applied the ensemble approach. For circles, they received a score of 99.18%, and for triangles, they received a score of 98.11%.

3. **Ivona Matoš; Zdravko Krpić; Krešimir Romić**

[5] Describe the "Speed limit road signs can be identified using Multi-Class SVM and the Hough Transformation. The preprocessing stage of this work uses hue, saturation, and lightness (HSL) variables to boost contrast in dataset photographs and facilitate detection. The Hough Circle feature was used in the detecting process. The Hough transformation is used to find circles in photographs. uses the HOG descriptor for edge detection and an SVM classifier to train and assess it, On the MASTIF and GTSRB data sets, the suggested model is put to the test.

EXISTING SYSTEM

The detection and recognition of traffic signs has been the subject of a lot of study. Numerous studies have concentrated on the color and form of the image for detection as two common features of traffic signals. Over a large number of frames, these properties can be utilized to track and identify moving objects. When the object that needs to be recognized has a distinctive color that shines out against the backdrop color, this technique is helpful. An object's edges, corners, and curves can be used to determine its shape. However, authors only paid attention to the detection and recognition techniques, ignoring the voice element, a key driver warning system. The tuning of hyperparameters has also received scant attention.

PROPOSED SOLUTION

In the proposed system, CNN's algorithm detects and recognizes traffic signs. Preprocessing the input is completed prior to categorization in order to minimize complexity, increase accuracy, and remove noise from the implemented method. Since we can't create a unique method for every circumstance in which a picture is captured, we frequently convert photos into a format.

Image Preprocessing :

Gray Scale Conversion: In order to save space or make computations easier, it might occasionally be beneficial to remove unnecessary information from images. Grayscale conversion of colorful images, for example. This is because color isn't always employed to differentiate between and see a picture in different objects. Grayscale may be sufficient to locate these items [1][3]. Color photos could be more sophisticated and require more memory because they have more detail than black and white photos. Because color images are represented by three channels, grayscale conversion reduces the number of pixels that need to be processed. Grayscale values work well for identifying traffic signs.

Traffic sign recognition:

Convolutional Neural Networks are a component of the Machine Learning subfield known as Deep Learning. Similar to how albeit on a far smaller scale than the human brain, deep learning algorithms do the same. To determine patterns in a collection, image classification requires extracting characteristics from an image. Given that CNN excels at feature extraction, we use it to recognize traffic signs [1][2].We employ filters at CNN. Depending on the purpose they are meant for, filters come in a range of sizes and shapes. By establishing a local communication pattern between neurons, filters enable us to benefit from the spatial localization of a particular image. During convolution, two variables are multiplied logically to add a fresh function. Our image pixel matrix is one function, and our filter is another. The To get the dot product of the two matrices, move the filter over the image. "Feature Map" or "Activation Map" matrix. The output layer consists of a number of convolutional layers that extract characteristics from the image. In order to enhance CNN, one can employ hyper parameter optimization. It determines the machine learning algorithm's hyperparameters that, when compared to a validation set, yield the best results. Hyper parameters must be specified before learning [1]. The learning rate and total number of units for a dense layer are reported. Our system will take into account the optimizer hyperparameter, kernel size, learning rate, and dropout rate.

Convolutional Neural Network Architecture

1. **Convolution Layer**

A key step in the convolution process is this layer. To extract several features from a given image, convolution is used[1]. Basically, it scans the entire pixel grid before performing a dot product. A feature from a number of features that we wish to identify from the input image makes up the filter or kernel. For instance, in the case of edge detection, a distinct filter for curves, blur, and image sharpening may be used. We can identify more sophisticated traits as we go further into the network.

Pooling Layer

Using this layer, the features are down sampled. The huge image's dimensionality is decreased yet important characteristics are kept. It helps reduce calculating and weighing. You can select either maximum Depending on the circumstance, pooling or average pooling is used. Max pooling collects the most value from the feature map, whereas average takes the average of all the pixels.

2. Activation Function

This layer adds non-linear properties to the network. Decision making about whether details has to be processed further and which does not is aided by this. Activating a function with the weighted sum of the inputs produces one signal output. This step is important since the output signal would be a simple linear function without it, which has limited potential for complicated learning. Sigmoid, Tan H, ReLU, Identity, and Binary Step functions are a few examples of activation function types. TanH's range is -1 to 0, whereas the sigmoid function's range is 0 to 1. This function is simple to optimize for. ReLU has a range of 0 to infinity and is one of the most common activation functions.

3.Flattening Layer

We must send data to the fully connected layer in the form of a 1D feature map because the output of the pooling layer is a 3D feature map.

This layer converts a 3*3 matrix into a one-dimensional list as a result.

4.Fully connected Layer

The real classification process happens in this layer. The results of convolution or polling applied layer by flattened layer are used to determine classification. Each input is connected to each output in this case using weights. In order to improve class prediction, it incorporates the information into additional properties.

Output of recognized sign in audio format

Currently, Although the motorist must reading the word on the designated sign, enhanced comfort is assured use a speech module. When a sign is identified, a text-to-speech module. They are numerous APIs for text to voice conversion that are accessible in Python. One of these APIs is the Google Text to Speech API, or gTTS API. A straightforward program called gTTS converts typed text into audio that may be saved in the MP# format. Multiple languages are supported via the gTTS API, and audio delivery speeds can be adjusted.

Conclusion

In this research, we reviewed the literature on traffic sign identification using machine learning approaches and conducted a comparative analysis of these methods. With the help of hyper parameter adjustment, CNN works well for recognition, and It is possible to boost accuracy or recognition rates. Therefore, We used CNN identification in the suggested architecture for a driver-warning traffic sign detection system. Preprocessing complete, the algorithm will be used to identify the items in the photos, which will have been taken using a car-mounted camera used for the picture acquisition phase. When it recognizes a traffic sign, the device speaks an alert. during the image acquisition phase, a camera installed on the vehicle. When a traffic sign is detected, the device delivers an alarm.

REFERENCES:

[1] W. Haque, S. Arefin, A. Shihavuddin and M. Hasan,"DeepThin: A novel lightweight CNN architecture for traffic sign recognition without GPU requirements", Expert Systems with Applications, vol. 168, p. 114481, 2021.

[2] S. Song, Z. Que, J. Hou, S. Du and Y. Song, "An efficient convolutional neural network for small traffic sign detection", Journal of Systems Architecture, vol. 97, pp. 269- 277, 2019. Available: 10.1016/j.sysarc.2019.01.012.

[3] A. Vennelakanti, S. Shreya, R. Rajendran, D. Sarkar, D. Muddegowda and P. Hanagal, "Traffic Sign Detection and Recognition using a CNN Ensemble," 2019 IEEE International Conference on Consumer Electronics (ICCE), 2019, pp. 1-4

[4] D. Tabernik and D. Skočaj, "Deep Learning for Large-Scale Traffic-Sign Detection and Recognition," in IEEE Transactions on Intelligent Transportation Systems, vol. 21, no. 4, pp. 1427- 1440, April 2020

[5] I. Matoš, Z. Krpić and K. Romić, "The Speed Limit Road Signs Recognition Using Hough Transformation and Multi-Class Svm," 2019 International Conference on Systems, Signals and Image Processing (IWSSIP), 2019, pp. 89-94.Degui Xiao, Liang Liu," Super-resolution-based traffic prohibitory sign recognition ",2019

[6] W. Haque, S. Arefin, A. Shihavuddin and M. Hasan, "DeepThin: A novel lightweight CNN architecture for traffic sign recognition without GPU requirements", Expert Systems with Applications, vol. 168, p. 114481, 2021.

[7] S. Song, Z. Que, J. Hou, S. Du and Y. Song, "An efficient convolutional neural network for small traffic sign detection", Journal of Systems Architecture, vol. 97, pp. 269- 277, 2019. Available: 10.1016/j.sysarc.2019.01.012.

[8] A. Vennelakanti, S. Shreya, R. Rajendran, D. Sarkar, D. Muddegowda and P. Hanagal, "Traffic Sign Detection and Recognition using a CNN Ensemble," 2019 IEEE International Conference on Consumer Electronics (ICCE), 2019, pp. 1-4

[9] D. Tabernik and D. Skočaj, "Deep Learning for Large-Scale Traffic-Sign Detection and Recognition," in IEEE Transactions on Intelligent Transportation Systems, vol. 21, no. 4, pp. 1427- 1440, April 2020

[10] I. Matoš, Z. Krpić and K. Romić, "The Speed Limit Road Signs Recognition Using Hough Transformation and Multi-Class Svm," 2019 International Conference on Systems, Signals and Image Processing (IWSSIP), 2019, pp. 89-94.

[11] Degui Xiao, Liang Liu," Super-resolution-based traffic prohibitory sign recognition ",2019.