IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 11, Issue 3, March 2024

DOI: 10.17148/IARJSET.2024.11341

Traffic Sign Recognition Through Voice Assistance Using Convolutional Neural Network

B. Haritha[1], T. Venkata Sai Bhargavi[2], Y. Hemasri[3], N. Venkata Amrutha[4]

MTech, Computer science & Engineering, Bapatla women's Engineering College, Bapatla, INDIA[1]

BTech, Computer science & Engineering, Bapatla Women's Engineering College, Bapatla, INDIA [2-4]

Abstract: For self-driving cars and intelligent transportation systems, detecting and recognizing traffic signs is crucial. Real-time traffic sign detection and recognition from camera photos is the task at hand. Across a range of computer vision tasks, Convolutional Neural Networks (CNN) have demonstrated efficacy in achieving high accuracy. In this work, we provide a CNN-based method for identifying and detecting traffic signs. Our method makes use of a deep CNN architecture that is capable of simultaneous traffic sign detection and classification. We use a sizable dataset of photos of traffic signs to train the CNN model, and we assess its effectiveness using a dataset from real-world data. Our test findings show that the suggested method can identify traffic signs in real time with minimal processing overhead and high accuracy.

Keywords: Traffic sign detection, Traffic sign recognition, Deep Learning, Convolutional Neural Network (CNN).

I. INTRODUCTION

Since cars are now a necessary form of transportation, all nations have put in place appropriate road laws and regulations to ensure public safety. Traffic signs are one of them; they help to convey the laws that must be observed in that particular region and offer drivers useful information. An accurate and concise message should be sent by a traffic sign with the least amount of reading comprehension required. In addition, drivers in rural areas could have trouble understanding a particular traffic sign's message since they are unfamiliar with the abundance of signs found in urban areas. Certain traffic signs are often disregarded by drivers who feel they are superfluous. This ignorance is partly a result of the drivers' differing attitudes.

II. RELATED WORK

Through the smooth integration of voice commands with visual recognition systems, recent developments in voice-assisted traffic sign identification have concentrated on improving user experience and safety. The goals of these advancements are to lessen driver distraction and increase accessibility. Novel methods leverage natural language processing (NLP) to decode spoken commands associated with traffic signs, allowing for instantaneous reactions and feedback. Furthermore, deep neural networks—a type of machine learning algorithm—are used to improve the precision and speed of sign interpretation and detection. Voice assistance and traffic sign recognition systems work together to help drivers recognize and comprehend signs more rapidly. It also allows for hands-free engagement, which makes driving safer. In order to assure dependable performance, ongoing research focuses on improving these systems for wider deployment while taking different environmental circumstances into account.

III. METHODOLOGY

Information Pre-processing:
Images must be converted into NumPy clusters (i.e., numeric values) in order to handle them. The images are stacked and then reduced to 30 by 30 pixels. Following this, the image names are mapped to the picture, readying the dataset for preparation.

Model:
One Deep Learning algorithm is the Convolutional Neural Network (CNN) algorithm. CNN is able to receive an image as input, prioritize various things within the image, and distinguish between them. Compared to other classification algorithms, it requires a lot less preprocessing. When compared to manual filters used in primitive methods, Convolutional Networks can learn filters or features from the images. The design of a convolutional network is comparable to the neuronal connection pattern found in the human brain. The arrangement of neurons in the Visual

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 11, Issue 3, March 2024

DOI: 10.17148/IARJSET.2024.11341

Cortex of the human brain served as the model for the design. The Receptive Field is the precise area of the field of view where neurons exclusively react to inputs.

Implementation:
Following training, the model is kept and utilized for prediction. A full stack web application developed using Node.js and Express Handlebars was predicted using this model. It combines a number of different logics to provide a product that can be used with some enhancements. This schematic represents the suggested system. The CNN model is applied in the first section where an image is the input. One of the 43 classes is produced as the output of the procedure. In the event that a given image is devoid of traffic signs, the user is presented with the message "No Sign Detected." Analysing the Python "model. pre2dict" function's return value of array allows for this. The "model.predict" function determines the class based on which value is highest after returning an array of values showing how closely each of the 43 classes matches the image.Following multiple repetitions, it was found that even in cases where an image does not fit into any of the specified classes, The model classifies it into one of the 43 classes even though it hasn't been trained for any additional classes; nonetheless, the value that the "model.predict" function predicts is relatively low. Thus, the threshold value of 0.68 is selected to discriminate between photos that do not have a traffic sign but are expected to have one.

IV. PROPOSED SYSTEM

We are developing a method that will enable us to explain traffic signs to drivers in real time and identify and understand them with greater accuracy. Both driver assistance systems and autonomous cars can make use of this type of technology. The system is built using the Convolutional Neural Network (CNN) paradigm, which provides faster detection rates.

ADVANTAGES:
1. The technology has the ability to recognize traffic signs in real time.
2. Drivers can benefit from the explanation of the message displayed on a certain traffic sign while they are driving. Through the voice-over narration.
3. There are solutions for problems such not seeing the traffic signs, not being familiar with them, and their intricacy.

V. ALGORITHMS AND TECHNIQUES

1. Data augmentation: To train machine learning models, data augmentation is the act of converting existing images into new ones. Modern machine learning models are highly powerfull, therefore this is a crucial phase in the dataset creation process. If these models are given too limited datasets, they may begin to "overfit," which is an issue where the models simply memorize the mappings between their inputs and expected outputs.

2. Convolutional Neural Network: One kind of deep learning algorithm that works especially well for tasks involving the recognition and processing of images is the convolutional neural network (CNN). Convolutional, pooling, and fully connected layers are some of the layers that make it up. The human brain's visual processing served as the inspiration for CNN architecture, which makes them ideal for identifying spatial connections and hierarchical patterns in images.

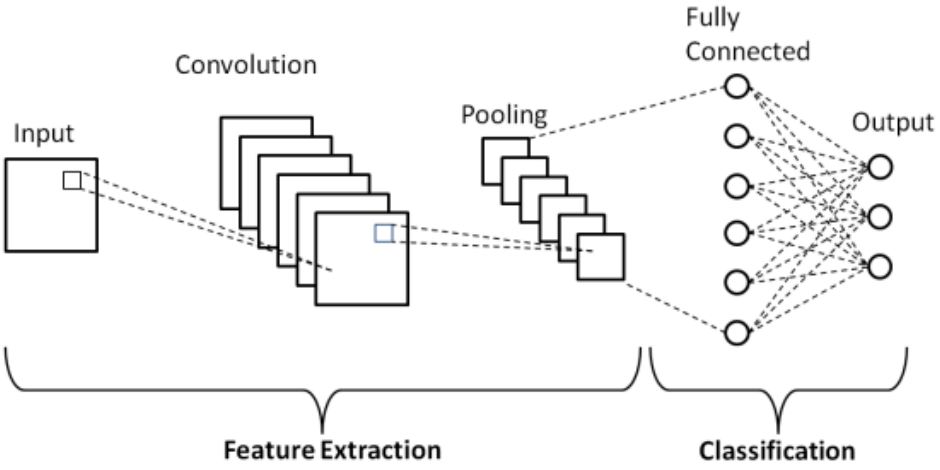

Figure1: CNN Architecture

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 11, Issue 3, March 2024

DOI: 10.17148/IARJSET.2024.11341

3. Text-to-Speech (TTS): The application turns recognized text into speech by using the pyttsx3 library, and it gives audible feedback on the recognition outcomes. By providing spoken output of the identified traffic sign labels, this feature improves the user experience and makes the application usable for users who are blind or in situations where visual feedback is not practical.

4. Using OpenCV for Image Processing: OpenCV is an effective library for image processing applications. It is utilized in the program to read, resize, and display photos. Furthermore, submitted images are preprocessed using OpenCV before being fed into the CNN model for recognition. The application guarantees effective handling and modification of image data, which is essential for precise traffic sign recognition, by utilizing OpenCV's functionalities.

5. Model Serialization: CNN models that have been trained are serialized and kept in the HDF5 format as model weights. This makes it simple to save and get the model parameters, which makes it easier to deploy and infer within the application. When the serialized model is loaded during recognition, traffic sign classification based on learnt features can be done in real-time. The trained CNN model can be effectively used in the web application without requiring retraining thanks to model serialization, which guarantees the model's portability and permanence.

6. Flask File Upload: This feature enables users to upload photos to the web application from their local system. Users are allowed to supply input data for the recognition of traffic signs. This functionality, which makes use of Django's FileSystemStorage, guarantees safe and effective management of uploaded files within the web.

VI. RESULTS AND ANALYSIS

Figure2 : Traffic Signs in different weather conditions

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 11, Issue 3, March 2024

DOI: 10.17148/IARJSET.2024.11341

Figure3: Different kinds of traffic signs

VII. CONCLUSION

Utilizing a convolution neural network, the traffic sign board detection and voice alert system is put into action. The CNN model that produced the best results on the GTSRB dataset was put into use after several other models were examined. A potential way to increase traffic safety is the suggested voice alerting and CNN-based traffic sign recognition system. To provide real-time traffic sign detection and alerts, the system combines machine learning, image processing, and voice-based alerting. Using a camera installed on the car, the system takes pictures of the road ahead, pre-processes them to reduce noise and improve the contrast of the traffic signs, and then uses CNNs to classify the images and predict the names of the signs.

REFERENCES

[1] Shrivastava, A., & Gupta, M. (2020). Real-Time Traffic Sign Detection and Recognition using Deep Learning: A Comprehensive Review. Journal of Ambient Intelligence and Humanized Computing, 12(9), 8915-8945.

[2] Ghiasi, G., Dumanli, S., & Le, Q. V. (2016). NAS-FPN: Learning Scalable Feature Pyramid Architecture for Object Detection.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 11, Issue 3, March 2024

DOI: 10.17148/IARJSET.2024.11341

[3] Wang, T., Liu, X., & Tao, D. (2019). Crowd counting by adaptively fusing predictions from an image pyramid. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 7185-7194.

[4] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 779-788.

[5] Lin, T. Y., Dollár, P., Girshick, R., He, K., Hariharan, B., & Belongie, S. (2017). Feature pyramid networks for object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2117-2125.

[6] Su, H., Qi, C. R., Li, X., & Guibas, L. J. (2015). Render for CNN: Viewpoint estimation in images using CNNs trained with rendered 3D model views. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), pp. 2686-2694.

[7] Xie, S., Girshick, R., Dollár, P., Tu, Z., & He, K. (2017). Aggregated residual transformations for deep neural networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1492-1500.

[8] Yu, L., Zhang, G., and Ma, B. "Traffic sign detection and recognition in the wild." In Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2017.

[9] Zhang, S., Wen, L., Bian, X., Lei, Z., Li, S. Z. "Single-shot refinement neural network for object detection." In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018.

[10] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., & Berg, A. C. (2016). SSD: Single shot multibox detector. In European conference on computer vision, pp. 21-37.

[11] Cho, S., Lee, S., Kim, J., Kim, T., & Choy, B. (2019). Inverse graphics GAN: Learning to generate 3D shapes from natural images. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 9243-9252

BIOGRAPHIES

Bethapudi Haritha M. Tech,
Asst. Professor,
Department of Computer Science and Engineering,
BWEC, Andhra Pradesh,
India.

Thirumalareddy Venkata Sai Bhargavi
[B. Tech], Student, Department of Computer Science and Engineering, BWEC, Andhra Pradesh, India.

Yandrapu Hemasri
[B. Tech], Student, Department of Computer Science and Engineering, BWEC, Andhra Pradesh, India.

Nalam Venkata Amrutha
[B. Tech], Student,
Department of Computer Science and Engineering,
BWEC, Andhra Pradesh,
India.