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. 12, Issue 1, January 2025

DOI: 10.17148/IARJSET.2025.12120

# Road Safety with Deep Learning Voice Based Traffic Sign

**Mrs.J.Sarojini Premalatha.M.E[1], Annwin Remilka.R.C[2], Devicharan.M[3]**

Department of Computer Science and Technology, Sathyabama Institute of Science and Technology,

Jeppaiar Nagar, Rajivee Gandhi Salai, Chennai - 600 119[1,2,3]

**Abstract:** The research deals with deep learning-based, voice-operated traffic sign recognition to improve road safety. A deep learning-based model is presented for the recognition of traffic signs with CNN, trained on GTSRB Dataset with an identification and categorization precision of 95%. The developed system detects signs and, through audio warnings by speakers, assists drivers to make quick decisions. The system aims to mitigate accidents caused by missed or misinterpreted signage by alerting drivers to nearby traffic signs and rules. This approach has potential applications in both driver assistance systems and autonomous vehicles.

**Keywords:** CNN, Keras, Tensor flow, GTSRB Dataset, Traffic Signs.

## I . INTRODUCTION

Traffic signs are a crucial part of road safety because they convey important information to the driver. However, failure to obey or correctly interpret the signs is a major cause of traffic accidents. Due to the massive use of automobiles as a primary mode of transportation, countries worldwide have implemented safety measures through comprehensive road policies. In such policies, traffic signs are fundamental, informing drivers and communicating local laws. These signs are meant to communicate information rapidly and precisely, with minimal requirements in literacy. However, a myriad of factors make drivers disobey traffic signs, including carelessness, distraction, unfamiliarity, and intentional non-compliance, which are usually attended by accidents. In addition, drivers from rural areas may be confused by the innumerable number of signs used in urban areas due to lack of exposure. Some drivers may selectively disregard some signs, thinking that these are unnecessary. This lapse or unfamiliarity with traffic signs may lead to serious accidents and even loss of life.

To solve this problem, an intelligent voice-based traffic sign recognition system was proposed with the power of DL.It contains a CNN architecture for the identification and classification of traffic signs in images. Training of the CNN model is done using the GTSRB Dataset, which includes over 50,000 images of over 43 different classes of signs. Such a large and diverse dataset trains the model to identify a wide range of traffic signs in countless different conditions.

Integration of vocal messages is one of the major developments of this system. On the detection of a traffic sign, this system gives an audible announcement to alert the driver with the meaning of the sign and associated rules or actions, if any. In this way, drivers are assured of receiving important information in situations where visual alerts are not seen, thus improving road safety in general.

The proposed system has two main components:

•A traffic sign detection and classification CNN module.
•A text-to-speech engine to convert recognized signs into vocal messages.

Combining these elements obtains an overall driver assistance system by which accidents caused by recognizing or misinterpreting specific traffic signs maybe mitigated.

With all being said, it also has prospective applications in the development of self-driving cars, where the correct identification of signs is important for the safe movement of vehicles. The next sections discuss the research approach, which includes data pre-processing, model creation, training procedures, and system installation. This paper will also show our results, including information on the 95% accuracy and practicality of our model for voice-based traffic sign identification.

*A. Scope*

- Creating a CNN model to spot and group traffic signs.
- Adding voice feedback to warn drivers about signs the system notices.
- Using this in systems that help drivers and self-driving cars.

DOI: 10.17148/IARJSET.2025.12120

- Testing it on a set of traffic sign pictures that experts use as a standard.

*B. Aim and Objectives.*

The research work covers the following Objectives.

- Make a CNN model that can recognize traffic signs and set up voice alerts to inform the driver every time the system detects signs.
- Signaling can also be used to enhance drivers' perceptions for improving safety on the roads.
- Provide a system that works for both autonomous vehicles and human drivers.

*C. Problem Statement.*

Existing traffic sign recognition systems face challenges in.

- Needing lots of good-quality tagged data to train.
- Using too much computing power to work in real time.
- Keeping user data safe and private.
- The new system tries to fix these issues while giving drivers correct voice alerts when they need them.

## II . RELATED WORKS

Extensive research has been conducted on the Road Traffic Sign Recognition (TSR) issue. Initial automated traffic sign detection work commenced in Japan, followed by various proposed methodologies for developing effective traffic sign recognition and detection (TSDR) systems. The following section examines literature related to traffic sign recognition. Yan Lai et al, presented a CNN-SVM-based traffic sign classification methodology, where the YCbCr color space was utilized for feature extraction; their method demonstrated better performances in practical applications compared with the previous approaches .[1]

J. Roopashree et al ,proposed a system which is meant to detect and classify road signs using neon indicators to identify and distance-based methodologies to classify, hence enabling driver-assistance functionalities.[2]

Giuseppe Guido et al, developed a neural network model based on GMDH for the prediction of vehicle involvement in accidents, which obtained high accuracy on urban and suburban accident data. [3]

Arman Sarraf, investigated various CNN architectures for binary image classification, reaching the best outcomes with a three-layer with dropout.[4]

Roshani S. Nage et al, proposed a weighted substructure-based image mining method, which was based on graph substructure for feature extraction and had obtained a good image analysis.[5]

Arun JB and Reshu Choudhary , improved image mining methods by following a content-based approach and achieved improved precision over the existing histogram-based approaches.[6]

Khandare et al, used the image mining domain for various domains by integrating textual and visual features to extract semantically relevant information and generating clusters for efficient retrieval. [7]

Wali et al, comprehensively reviewed the methodologies in detecting, tracking, and classifying traffic signs, including major challenges and technological improvement.[8]

Pradumn Kumar and Upasana Dugal , used the state-of-the-art CNN models with TensorFlow in plant classification, which showed a better performance using the CIFAR-10 dataset.[9]

Cao et al, presented the improvements in traffic sign detection and recognition by combining HSV color space, geometric feature analysis, and an improved LeNet-5 CNN for robust performance under different environmental conditions. [10]

Islam et al, Compared SVM, CNN, and ANN for face recognition using BoW, HOG, and IP features; ANN with IP achieved the highest accuracy on the AT&T face database. [11]

International Advanced Research Journal in Science, Engineering and Technology
Impact Factor 8.066 ⋇ Peer-reviewed / Refereed journal ⋇ Vol. 12, Issue 1, January 2025
DOI: 10.17148/IARJSET.2025.12120

Jung et al, designed a traffic sign recognition system by integrating deep learning with voice alert to improve drivers' safety and reduce their response time.[12]

Kim , studied YOLOv5 for efficient real-time traffic sign detection combined with strong SORT tracking and optimized for urban scenarios.[13]

Saadna & Behloul , reviewed methods of traffic sign detection and classification, which can be grouped into color-based, shape-based, and learning-based techniques.[14]

Lim et al, Designed a multi-modal traffic sign detection system fusing visual and auditory information, optimized for low-visibility and real-time operation.[15]

## III . PROPOSED WORKS

The proposed CNN models are implemented for traffic sign detection using the Python tool and, thus, should be allowed to validate functionality and performance using a traffic sign data-set.

*A. Data Collection and Analysis:*
Data-set Description It includes all forms, such as speed limits, no entry, signals for the flow of traffic, indicators, children crossing, restrictions of heavy vehicles, and many others. Traffic Sign Classification involves the classification of an input image into the class which it belongs to. The paper proposes a deep neural network model for classification of traffic signs captured in images. It can read the meaning of traffic signs and their representation-one of the central abilities of an autonomous vehicle. For the research collected a publicly available data-set from Kaggle. The data-set is diverse, with some categories having numerous images and others having fewer. The total dataset size is about 300 MB, including a train folder with class-sorted images and a test folder for model evaluation.

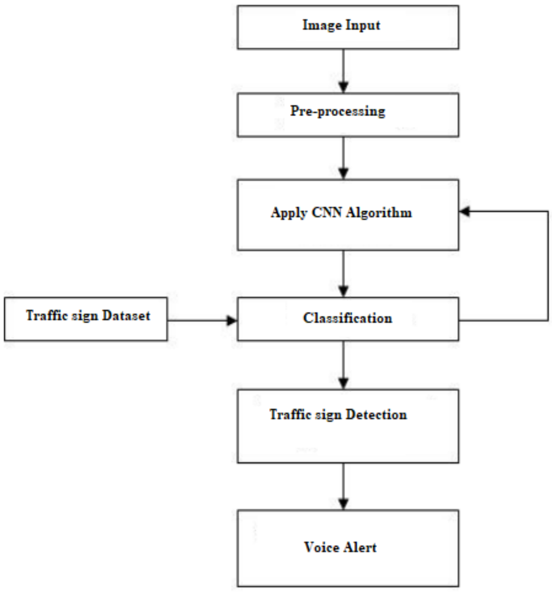


Fig. 1 Work Flow

*Modules*
*Data Collection*
The system requires an image with traffic signs in this module, and gather real-time photographs that include pictures of traffic signs. Using Keras's image processing utilities, the saved image is then loaded and re-sized according to the specific requirements of the machine learning models, ensuring consistency in input size. The image is transformed into an array format. This prepossessed image array is subsequently fed into deep learning models, each of which makes a prediction on the class of the traffic sign. The output results that detects the traffic sign.

© IARJSET This work is licensed under a Creative Commons Attribution 4.0 International License 

DOI: 10.17148/IARJSET.2025.12120

*Morphological Operations*

In the module, a picture containing a traffic sign should be prepossessed by carrying out a few procedures that will eliminate extraneous noise and prepare the image for additional processing. During the Pre-processing stage, a picture must undergo certain operations, such as opening, closing, dilation, and erosion.

*Sign Detection*

In the phase, after applying morphological operations and converting the image to binary format, features are identified based on shapes and colors. The Pre-processing step is iterated until new features emerge. Once an image is detected, it is forwarded for further processing, namely classification.

*Sign Classification*

The classification phase of the detected images, which is allowed with the help of any standard data-set. Here, the German Traffic Signs Data-set is used. Classification is allowed using CNN algorithms to classify the signs into three primary types: Circular, Triangular, and Rectangular, according to their functions.

*Voice Alert Message*

Following image classification, text conversion is performed. This system includes a module containing meanings for all traffic signs, as each sign has a specific interpretation. It generates a text message for the sign and subsequently produces a voice message based on the generated text.

*B. Deep Learning Technique*

The DL techniques are utilized in image classification of traffic sign without the requirement of pre-required process like extraction of feature and segmentation.

Deep learning is the rising region in an automated extraction of features in traffic images. Owing to an automated feature extraction capability, convolution neural network (CNN) is the developing for classification of image. It can be able to manage the applications of big data without utilizing image pre-processing. An area of image processing may be capable for providing the solution to challenges of agriculture sector. CNN will be employed for feature extraction of traffic images as it has rapid processing speed as well as highly accurate in classification. Therefore, DL methods have been shown favourable outcome in classification of images.

In the past few years, the implementation of deep learning-based approaches has markedly enhanced the precision of image classification.

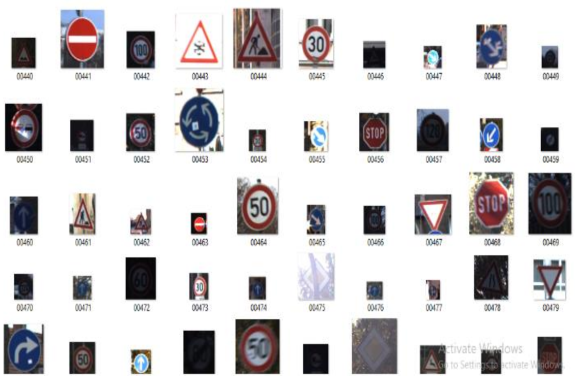

Fig.2 Traffic sign Dataset

*C. Deep Learning Methods*

CNN (Convolution Neural Network) The study implemented a Deep Learning approach based on Convolutional Neural Networks (CNN). This method processes image inputs, emphasizing various objects within the image and distinguishing between them. The deep learning model consists of multiple pre-processing layers that extract information from raw input to produce task-specific output. The model primarily comprises four layer types: convolution, max-pooling, fully-connected, and output, arranged in a stack. CNNs use extracted features as input, which substantially enhances classification accuracy. The model's output is the classified result. The following describes the CNN structure for model classification.

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. 12, Issue 1, January 2025
DOI: 10.17148/IARJSET.2025.12120

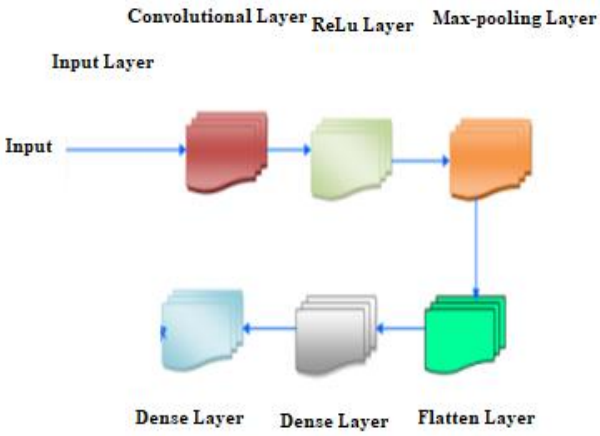


Fig. 3 CNN Architecture

The CNN architecture from Fig. 6 includes a set of two convolutional layers followed by a ReLU and supported by two max-pooling layers, followed by a flatten layer and two dense, or fully connected, layers. The last fully connected layer serves as an output layer using the Softmax activation function.The process requires an intermediate part that makes the image one-dimensional, enhancing efficiency by making it easier to handle; hence, the first two layers serve as a feature extractor, while the third one will serve the role of a classifier. Image features are extracted using convolution and pooling techniques, which are then input into the initial fully-connected layer. The final fully connected layer's output is then fed into the output layer for classifying traffic signs.

*D. Technologies Used*

*Python*

Python is a high-level, general-purpose language used for extensive programming. It was initiated by Guido van Rossum in 1991 and developed further by the Python Software Foundation. The focus of development has been on enhancing code readability, with a syntax that enables programmers to express concepts using minimal lines of code.Python targets, in many ways, being an advanced and multi-faceted language, is interpretive and further generalized. It has a dynamic data type with automatic memory management. It allows multiple-programming paradigms to be performe object-oriented, imperative, functional, and procedural. It also has an extensive set of standard libraries offering basic tools and functionalities.

TensorFlow and Keras (Deep Learning Frameworks)

Both TensorFlow and Keras are used in loading, pre-processing, and making predictions from pre-trained machine learning models.

## IV . RESULT AND DISCUSSION

The traffic signs data-set has been used in testing the proposed CNN model, and the experimental result is given in terms of some performance metrics. Further, the present section explains the obtained results as well as comparative discussions about models with existing methods considering metrics for performance evaluation. A description is provided here relating to experimental setup and experimental results as compared to their performance.The results and evaluation suggested that a neural network architecture comprising 4 convolution layers, 2 max pooling layers, along with dropout, flatten, and dense layers, outperformed other CNN structures.

Voice-based traffic sign recognition systems using deep learning represent the new breed of technology innovations to improve road safety and driver assistance. Such systems provide voice feedback on visual traffic sign information to support drivers. They depend on deep learning algorithms, such as CNN, which can detect a wide variety of traffic signs with high accuracy from images or video streams in real time.Conclusion The active-voice-controlled traffic sign recognition system presents the motorists with an easily accessible and user-friendly interface of safe driving and thus an informed driving experience.

A. *Experimental Approach*

The performance evaluation of the suggested model was done with the implementation of software written in Python on

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. 12, Issue 1, January 2025
DOI: 10.17148/IARJSET.2025.12120

the traffic sign dataset. The case study was done on the "Intel processor, 8 GB of RAM, and a 64-bit Windows 11 operating system."

*B. Performance metrics*

*Accuracy :*

The number of correct identifications, both cases and no-cases, in relation to the total number examined.

$$\text{Accuracy} = \frac{TP + FN}{TP + TN + FP + FN}$$

$$\text{Accuracy} = \frac{950 + 940}{950 + 940 + 10 + 10} = 0.9869$$

*Precision :*

Precision is defined as the number of true positives divided by the total number of cases identified as positive.

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Precision} = \frac{950}{950 + 10} = 0.9891$$

*Recall :*

Recall calculates the ratio of correctly identified real positive cases.

$$\text{Sensitivity} = \frac{TP}{TP + FN}$$

$$\text{Sensitivity} = \frac{950}{950 + 13} = 0.9869$$

*F1 Score :*

The F1 Score combines Precision and Recall using their harmonic mean, offering a single metric that balances both measures.

$$\text{F1-Score} = 2* \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}$$

$$\text{F1-Score} = 2* \frac{0.9891 * 0.9869}{0.9891 + 0.9869} = 0.9880$$

TABLE 1: OVERALL MODEL ACCURACY

| Model | Accuracy | Precision | F1-score | Recall |
|---|---|---|---|---|
| CNN | 0.9869 | 0.9891 | 0.9880 | 0.9869 |

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. 12, Issue 1, January 2025
DOI: 10.17148/IARJSET.2025.12120

*C. Results*

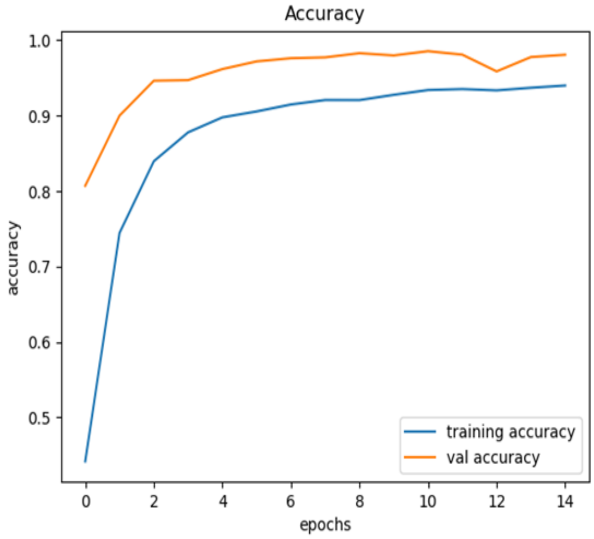


Fig. 4 Training loss Accuracy

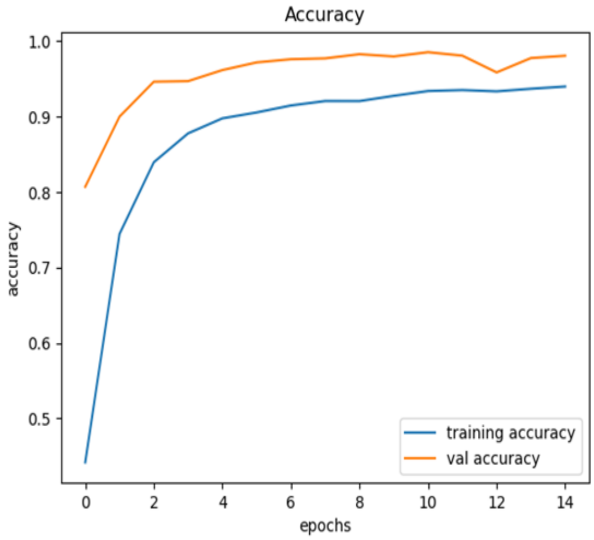


Fig. 5 Training Validate Accuracy


Fig. 6 Stop Sign Output Detected

 This work is licensed under a Creative Commons Attribution 4.0 International License 

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. 12, Issue 1, January 2025

DOI: 10.17148/IARJSET.2025.12120

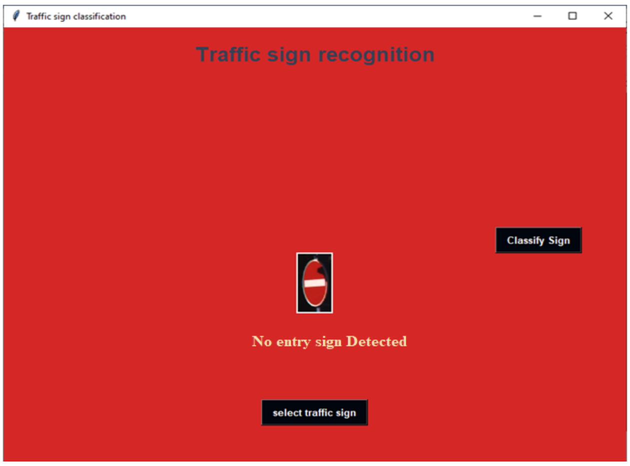


Fig. 7 No Entry Sign Output Detected

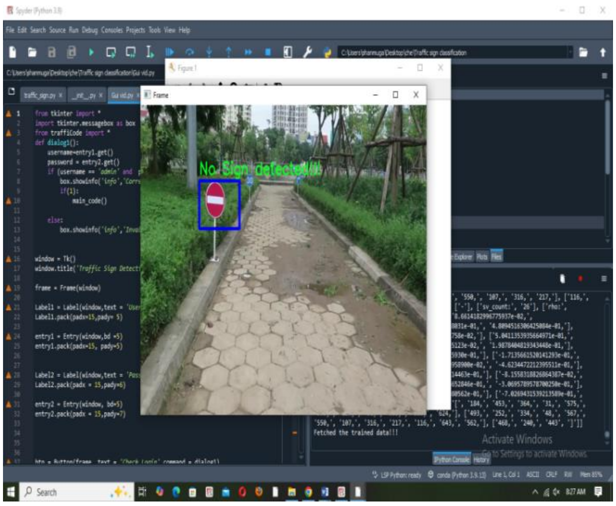

Fig. 8 No Sign Output Detected

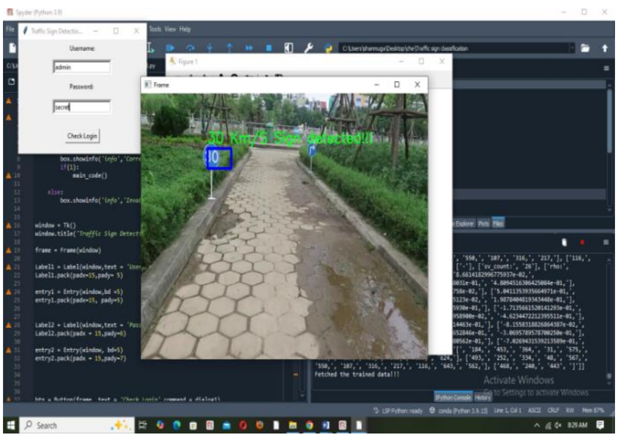

Fig. 9 Speed 30km Sign Output Detected

DOI: 10.17148/IARJSET.2025.12120

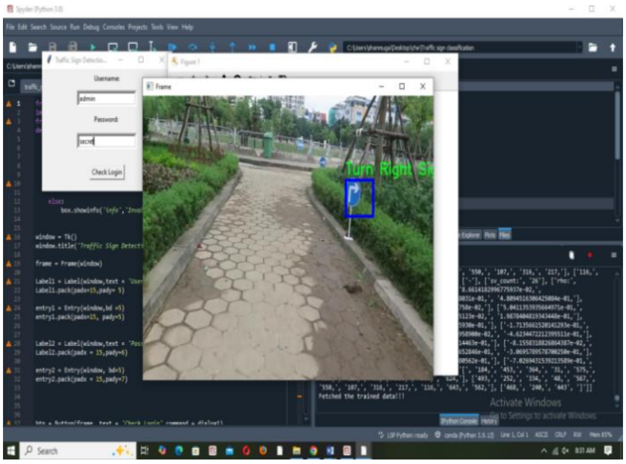


Fig. 10 Right Turn Sign Output Detected

## V.CONCLUSION

It proposed the developed Traffic Sign Recognition and Voice alert System with CNN architecture that could achieve 98.69% on GTSRB. This intelligently integrates the process of detection of traffic signs with voice alerts for more enhanced driver and road safety. This particular research demonstrates the potential of utilizing deep learning techniques for real-time recognition of traffic signs and voice-based applications that assist drivers.

## REFERENCES

[1] Yan Lai, Nanxin Wang, Yusi Yang and Lan Lin, "Traffic Signs Recognition and Classification based on Deep Feature Learning", Proceedings of the 7th International Conference on Pattern Recognition Applications and Methods, pp. 622-629, 2018.

[2] J. Roopashree, Roshan Pandey, Vinuta T Pawaskar and R. Vaishnavi, "Road Sign Detection and Recognition System", International Research Journal of Engineering and Technology, Volume 5, Issue 7, pp. 69-71, 2018.

[3] Giuseppe Guido, Sina Shaffiee Haghshenas, Sami Shaffiee Haghshenas, Alessandro Vitale, Vincenzo Gallelli and Vittorio Astarita, "Development of a Binary Classification Model to Assess Safety in Transportation Systems Using GMDH-Type Neural Network Algorithm", Sustainability, Volume 12, Issue 17, pp. 1-19, 2020.

[4] Arman Sarraf, "Binary Image Classification through an Optimal Topology for Convolutional Neural Networks", American Scientific Research Journal for Engineering, Technology, and Sciences, Volume 68, Issue 1, pp. 181-192, 2020.

[5] Roshani S. Nage, Vikramsingh R. Parihar and Atul S. Dahane, "A Novel Graph-based Image Mining Technique Using Weighted Substructure", Image Segmentation: A Guide to Image Mining, An ICSES Book published by ITIPPR, Volume 4, Issue 4, pp. 16-25, 2018.

[6] Arun J.B. and Reshu Choudhary, "Image Mining for Retrieval of Image using Histogram Technique", International Journal of Current Engineering and Technology, Volume 7, Issue 6, pp. 1-6, 2017.

[7] Kiran P. Khandare, Swati V. Kamble, Shivani R. Jaiswal and Nutan P. Khelkar, "Retrieval of Images & Text using Data Mining Techniques", International Research Journal of Engineering and Technology, Volume 5, Issue 3, pp. 2605-2608, 2018.

[8] Safat B. Wali, Majid A. Abdullah, Mahammad A. Hannan, AiniHussain, Salina A. Samad, Pin J. Ker and Muhamad Bin Mansor "Vision-Based Traffic Sign Detection and Recognition Systems: Current Trends and Challenges", Sensors, Volume 19 , Issue 9, pp. 1-28, 2019.

[9] Pradumn Kumar and Upasana Dugal, "Tensorflow Based Image Classification using Advanced Convolutional Neural Network", International Journal of Recent Technology and Engineering, Volume 8 Issue 6, pp. 994-998, 2020.

[10] Jingwei Cao, Chuanxue Song, Silun Peng, Feng Xiao and Shixin Song, "Improved Traffic Sign Detection and Recognition Algorithm for Intelligent Vehicles", Sensors, Volume 19, Issue 18, pp. 1-21, 2019.

[11] Islam et al., "Comparison of SVM, CNN, and ANN with Feature Extraction Techniques like BoW, HOG, and IP for Face Recognition," AT&T Face Database, achieving the best accuracy with ANN and IP.

[12] Jung et al., "Deep Learning Integration with Voice Alert for Traffic Sign Recognition to Enhance Drivers' Reaction Time and Safety," 2021.

DOI: 10.17148/IARJSET.2025.12120

[13] Kim, "YOLOv5 for Real-Time Fast and Accurate Object Detection of Traffic Signs Combined with StrongSORT for Tracking, Optimized for Urban Environments," 2022.

[14] Saadna & Behloul, "A Review of Traffic Sign Detection and Classification Approaches Categorized as Color-Based, Shape-Based, and Learning-Based Methodologies," 2021.

[15] Lim et al., "A Bimodal Traffic Sign Recognition System with Visual and Auditory Inputs Optimized for Low-Visibility Conditions and Real-Time Applications," 2023.