A Literature Review of TSR System with A Visual Display and Audio (Speech) Alerts

Mr. Mayureshwar Shendre[1], Dr. Shirish Pattalwar[2], Ms. Neha Wadi[3], Ms. Shreya Adsod[4], Mr. Kartik Rathod[5]
[1,2,3,4,5] *Dept. of Electronics and Telecommunication Engineering, PRMIT&R, Badnera*

***Abstract*—The Traffic Signs Recognition System with Visual Display and Audio Alerts is an intelligent system designed to enhance road safety by detecting and recognizing traffic signs in real-time. The system employs advanced machine learning techniques, such as the YOLO (You Only Look Once) algorithm, to accurately identify various traffic signs captured through an ESP32 camera. Once a traffic sign is detected, the system provides immediate feedback to the driver through visual displays and audio alerts using text-to-speech technology, ensuring timely awareness and response. The back-end processing is handled on a laptop system, enabling efficient computation and real-time performance. This system aims to reduce traffic violations, enhance driver assistance, and contribute to safer road environments by integrating machine learning, IoT, and speech-based alert mechanisms.**

***Index Terms*—CNN, ESP32 Camera, IoT, Machine Learning (ML), Text-to-Speech Technology, TSR System, YOLO**

I. INTRODUCTION

Road safety has become one of the most critical concerns in today's fast-growing transportation systems. With the rapid increase in the number of vehicles, drivers are often challenged to remain attentive to traffic rules and signs, which can sometimes be overlooked, leading to accidents or violations. To address this challenge, intelligent systems that combine computer vision and machine learning are being developed to assist drivers in real-time.

The proposed paper, Traffic Signs Recognition System with Visual Display and Audio Alerts, is designed to automatically detect and recognize traffic signs using an ESP32 camera and Machine Learning algorithms such as YOLO. Once a traffic sign is identified, the system provides instant feedback by displaying the recognized sign and delivering an audio alert through text-to-speech technology. This dual feedback mechanism ensures that drivers are promptly informed without diverting their attention from the road. The backend computation is managed on a laptop system to achieve accurate and real-time performance.

By integrating IoT devices, machine learning, and speech technology, this paper aims to create a reliable driver-assistance solution that improves road safety, minimizes the chances of human error, and contributes to the development of smart transportation systems.

II. LITERATURE REVIEW

1] Krish Sukhani et al., In this paper they stated the critical issue of road safety by developing a system that utilizes Convolutional Neural Networks (CNNs) for real-time traffic sign detection and classification. The system is trained on the German Traffic Sign Recognition Benchmark (GTSRB) dataset, which comprises approximately 43 categories and 51,900 images of traffic signs. The proposed CNN model achieves an accuracy of about 98.52%, effectively identifying various traffic signs. Upon detection, the system provides voice alerts to the driver, enhancing situational awareness and reducing the likelihood of accidents due to overlooked or misinterpreted signs. This approach integrates computer vision and audio feedback to create an intelligent driver assistance system aimed at improving road safety for drivers, passengers, and pedestrians.

2] Yogesh C.M et al., In this paper they present a system designed to enhance road safety through real-

time traffic sign recognition and voice alerts. The system employs a Convolutional Neural Network (CNN) trained on the German Traffic Sign Recognition Benchmark (GTSRB) dataset, which comprises 51,900 images across 43 categories. Achieving an accuracy of approximately 98.52%, the system effectively identifies various traffic signs. Upon detection, it provides voice alerts to the driver, thereby improving situational awareness and reducing the likelihood of accidents due to overlooked or misinterpreted signs. This integration of computer vision and audio feedback aims to ensure the safety of drivers, passengers, and pedestrians.

3] Atharva Jadhav et al., In this paper they stated a system that uses Convolutional Neural Networks (CNNs) to detect traffic signs in real time and provide voice alerts to drivers. Trained on a diverse dataset of traffic signs, the system improves driver awareness and road safety by accurately recognizing signs such as speed limits and stop signs. The study highlights the effectiveness of CNNs, the need for preprocessing techniques, and addresses challenges like varying lighting and sign designs, emphasizing the potential of AI-based driver assistance systems.

4] Dr. T.R. Arunkumar, et al., In this paper they present a system that employs CNNs to detect and classify traffic signs in real time. Trained on the GTSRB dataset, the model achieves high accuracy in recognizing various traffic signs and provides voice alerts to drivers, enhancing situational awareness and reducing the risk of accidents. The system combines computer vision and audio feedback to improve road safety for drivers, passengers, and pedestrians.

5] Mayura Manawadu et al., In this paper the author presents a system that uses CNNs to detect and classify traffic signs in real time and provides voice alerts to drivers. By integrating computer vision with a text-to-speech engine, the system helps drivers notice signs they might otherwise miss, enhancing road safety and supporting the development of autonomous vehicle technologies.

6] S. Venkataramana et al., In this paper the author states that CNNs use CNNs to detect traffic signs in real time and provide voice alerts to drivers. It also integrates location-based services to help users find nearby amenities like hospitals, restaurants, and gas stations, displaying routes on an interactive map. Combining machine learning with navigation features, the system enhances driver awareness and offers a user-friendly solution for safer and more informed travel.

7] T.Yoshitha et al., In this paper the author presents a CNN-based system that accurately classifies 43 road sign categories and provides real-time audio feedback. The system includes multilingual support for 11 Indian languages and a user-friendly interface, enhancing accessibility for diverse users, including the visually impaired, while improving road safety and regulatory compliance in autonomous vehicles.

8] M.Rama Krishna Murthy et al., In this paper the author presents a CNN-based system that detects traffic signs in real time and provides voice alerts to drivers. Using the HSV color model to identify regions of interest, the system achieves 98% accuracy on the GTSRB dataset. By integrating deep learning with audio feedback, it enhances driver awareness and aims to reduce the risk of road accidents.

9] Sarvesh Burli et al., In this paper the author presents a CNN-based system that detects and classifies traffic signs in real time and provides voice alerts to drivers. By leveraging AI, the system enhances driver awareness, helps ensure compliance with traffic regulations, and contributes to safer driving experiences.

10] Yadav et al., In this paper the authors propose a system to recognize Indian traffic signboards from live video streams and issue driver alerts using voice. Their approach involves five stages: collecting a dataset of signboard images, processing and enhancing image quality, extracting relevant features (such as edges) and suppressing non-informative parts, classifying signs using a Support Vector Machine (SVM), and finally triggering a voice alert whenever a detected sign matches a trained class. The system targets three broad categories of Indian signs (Regulatory, Cautionary, Informational), and is trained on more than a thousand samples for each of ten sign types. Their work demonstrates that classical image processing plus an SVM classifier can achieve

reliable recognition in real road scenarios, though limited to a small number of sign types and under somewhat controlled conditions.

III. METHODOLOGY

The proposed system is designed to recognize traffic signs in real-time and provide instant feedback to drivers through both visual and audio alerts. The methodology begins with live image acquisition using an ESP32 camera which is connected to ESP8266 Microcontroller. ESP32 Camera continuously captures road scenes during vehicle movement. The captured images are transmitted to a laptop system where the YOLO (You Only Look Once) algorithm or other machine learning techniques are employed for traffic sign detection and recognition. YOLO is chosen for its high speed and accuracy in real-time object detection, making it suitable for dynamic driving environments.

Once a traffic sign is recognized, the system generates a visual display of the detected sign on the screen and simultaneously activates a text-to-speech (TTS) module to produce an audio alert. This dual feedback ensures that the driver is informed immediately without diverting attention from the road. The integration of IoT (ESP32 camera and NODEMCU 8266), machine learning (YOLO), and speech technology creates a robust driver-assistance system.

The proposed methodology ensures a cost-effective, scalable, and efficient solution that can be implemented in general vehicles to improve road safety, reduce driver errors, and support the development of smart transportation systems.

❖ BLOCK DIAGRAM

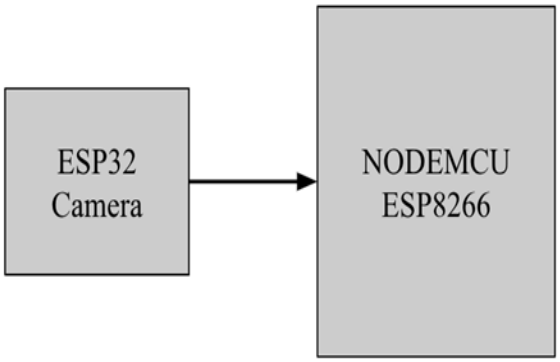

Fig. 1 Shows the Block Diagram of the System

DESCRIPTION

In this block diagram NODEMCU ESP8266 is used as a microcontroller and ESP32 Camera is used as an input device and it is connected to NODEMCU ESP8266.

FLOW CHART

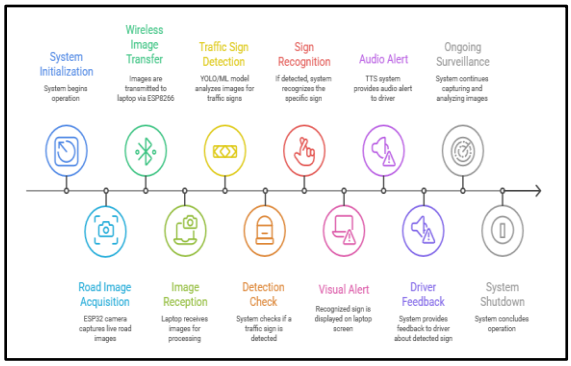

Fig. 2 Shows the Flow Chart of the System

WORKING

The system begins with initialization, after which the ESP32 camera starts capturing live road images. These images are wirelessly transferred to the laptop through the ESP8266 module. Once received, the laptop processes the images using the YOLO or other machine learning models to detect any traffic signs. The system then checks whether a valid traffic sign is present in the frame. If a sign is detected, it is recognized and displayed visually on the laptop screen while the text-to-speech module simultaneously generates an audio alert for the driver. This provides immediate feedback to ensure timely awareness. The system continues to monitor and analyze incoming images in a loop for ongoing surveillance. Finally, when the operation is complete, the system undergoes a shutdown process, concluding the detection cycle.

IV. SYSTEM REQUIREMENT

HARDWARE REQUIREMENT

- NodeMCU ESP8266
- ESP32 Camera Module

SOFTWARE REQUIREMENT

- Python
- Embedded C

MODULE USED

- Flask

V. EXPERIMENTAL SETUP & RESULT

EXPERIMENTAL SETUP

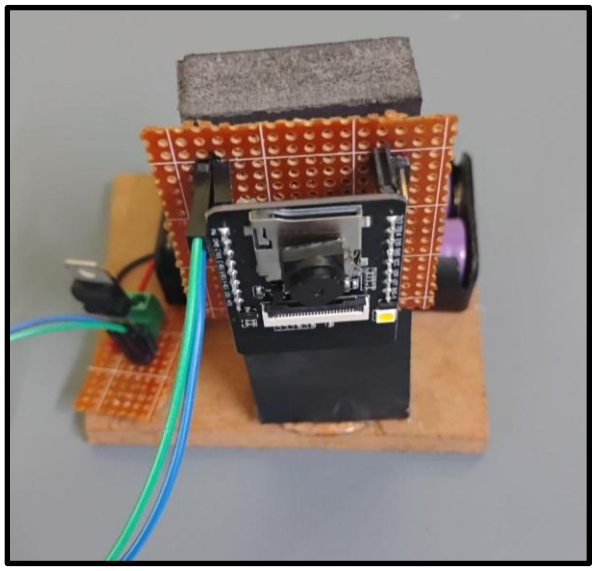

Fig. 3 Show Top View of Experimental Setup of System

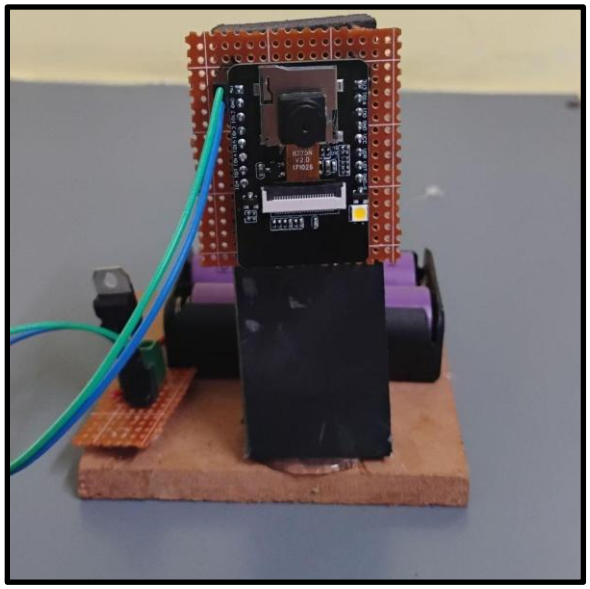

Fig. 4 Shows Front View of Experimental Setup of System

The ESP32 camera starts capturing live road images. These images are wirelessly transferred to the laptop through the ESP8266 module. Once received, the laptop processes the images using the YOLO or other machine learning models to detect any traffic signs.

IMPLEMENTATION

The ESP32 camera starts capturing live road images. These images are wirelessly transferred to the laptop through the ESP8266 module. Once received, the laptop processes the images using the YOLO or other machine learning models to detect any traffic signs.

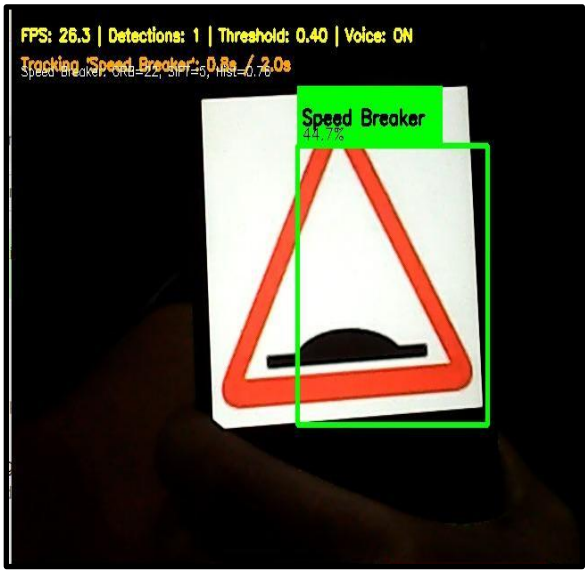

Fig. 5 Shows the Alert Message for Speed Breaker Detection Along with the Confidence Value Of 44.7%

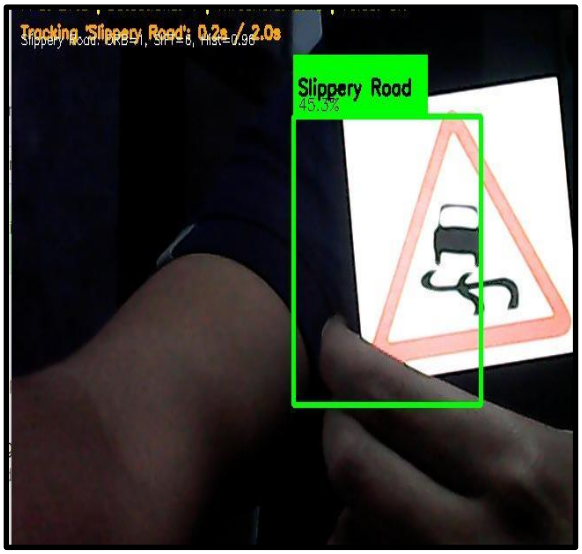

Fig. 6 Shows Alert Message for Slippery Road Along with Confidential Value of 45.3 %

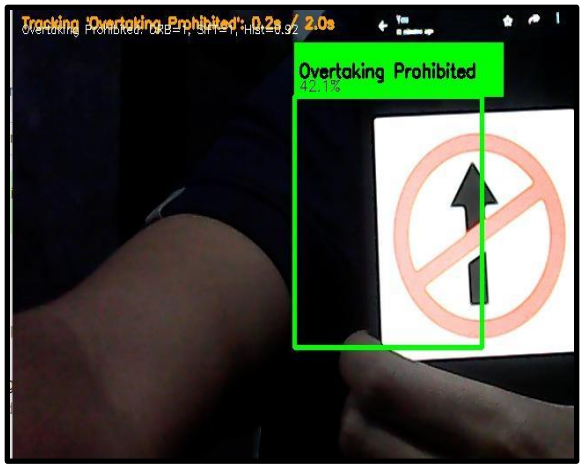

Fig. 7 Displayed the Message for Overtaking Prohibited Along with Confidential Value of 42.1%

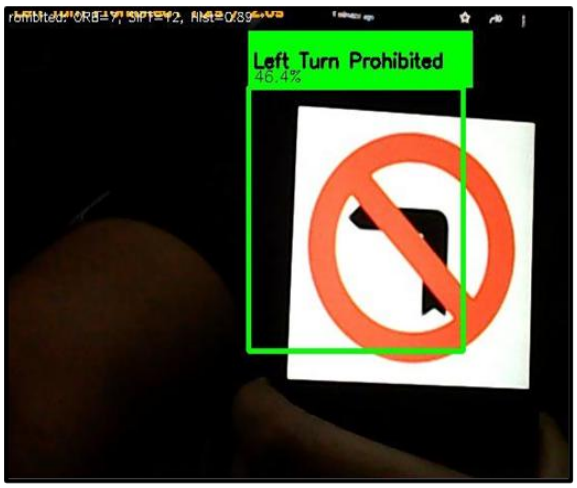

Fig. 10 Shows the Warning Message that Left Turn is Prohibited Along with Confidential Value of 46.4%

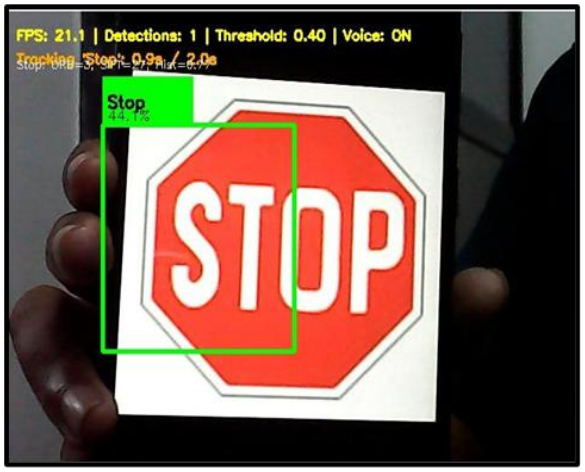

Fig. 8 Shows the displaying stop sign along with confidential value of 44.1%

Fig. 11 Displaying the Message about Upcoming Construction Work Along with Confidential Value of 59.1%

Fig. 9 Warns About there is Cyclist Ahead Along with Confidential Value of 51.2%

Fig. 12 Displaying Warning about Vehicles Prohibited in Both Directions Along with Confidential Value of 48.6%

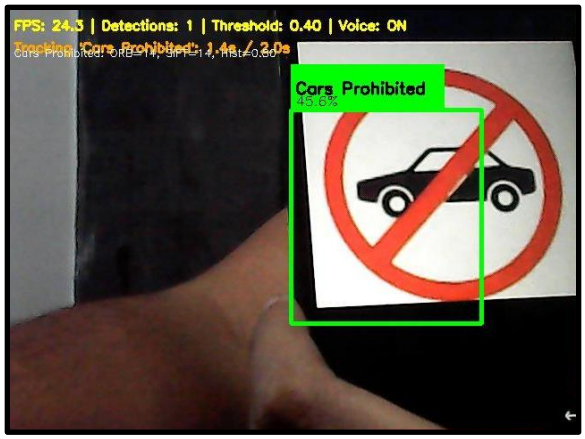

Fig. 13 Shows the Alert Message for Cars are Prohibited Along with Confidential Value of 45.6%

Fig. 14 Shows the Warning About there is Tunnel Ahead Along with Confidential Value of 58.9%

VI. RESULT

The system then checks whether a valid traffic sign is present in the frame. If a sign is detected, it is recognized and displayed visually on the laptop screen while the text-to-speech module simultaneously generates an audio alert for the driver. The detected sign is highlighted using bounding boxes along with its class label and confidence score to ensure clear identification. The audio alert provides real-time verbal feedback, enabling the driver to respond promptly without constantly monitoring the display. This dual-mode alert mechanism enhances driver awareness and reduces the chances of missing critical traffic signs, especially in high-speed or low-visibility conditions. The process is continuously repeated for each incoming frame, ensuring smooth, real-time traffic sign monitoring and assistance.

VII. CONCLUSION

The proposed Traffic Signs Recognition System with Visual Display and Audio Alerts is an innovative approach to enhancing road safety through the integration of IoT, machine learning, and speech technology. By utilizing an ESP32 camera for real-time image acquisition and applying YOLO-based recognition, the system ensures accurate detection of traffic signs even in dynamic driving conditions. The dual feedback mechanism, combining visual display and audio alerts, provides immediate assistance to drivers without causing distractions. This cost-effective and efficient solution not only reduces the risk of accidents and traffic violations but also supports the development of intelligent driver-assistance technologies for smart transportation systems.

REFERENCE

[1] Sukhani, Krish, Radha Shankarmani, Jay Shah, and Krushna Shah. "Traffic sign board recognition and voice alert system using convolutional neural network." In 2021 2nd International Conference for Emerging Technology (INCET), pp. 1-5. IEEE, 2021.

[2] Yogesh C.M, Usha Sree R., Hushalictmy.P., “Traffic Sign Board Recognition and Voice Alert System using CNN”, (INTI JOURNAL), 26 June 2024.

[3] Atharva Jadhav, Sidhhesh Samgir, Anant Jagtap, Akanksha Jadhav, S. P. Bhadre, “Traffic Sign Detection and Voice Alert”, (International Journal for Research in Applied Science & Engineering Technology), Volume 11 Issue V May 2023.

[4] Dr. T.R. Arunkumar, Miss. Aishwary M, Mr. Malasidda Madari, “Traffic Sign Recognition and Voice Alert System Using Convolutional Neural Network”, (International Journal of Research in Engineering and Science), Volume 12 Issue 8, August 2024.

[5] Mayura Manawadu, Udaya Wijenayake, “Voice-Assisted Real-Time Traffic Sign Recognition

System Using Convolutional Neural Network", (International Conference on Advanced Research in Computing), 2021.

[6] S. Venkataramana, G. V. Satya Sri Sai, J. D. S. Preetham, J. Mohan Sai, and K. Jashwanth Sree, "Voice Alert System for Traffic Sign Recognition and Nearby Places Finder", (International Journal of Innovative Research in Computer Science and Technology), Volume-13, Issue-2, March 2025.

[7] T.Yoshitha, N.Vasanthi, M.Pavani, K.Yuvatej, Dr.S.Sai Kumar, "Intelligent Road Sign Recognition System with Multi-Language and Audio Capabilities for Self-Driving Vehicles", (International Journal of Creative Research Thoughts), Volume 13, Issue 2 February 2025.

[8] M.Rama Krishna Murthy, P. Rakesh, S. Akriti, S. Vamsi Kiran, L. Vamsi Krishna, "Traffic Sign Board Recognition and Alerting System", (International Journal of Advances in Engineering and Management), Volume 4, Issue 5 May 2022.

[9] Sarvesh Burli, Sheetal Limbigidad, Abhishek Nejakar, Arihant Patil, Prof. Dattatreya Choudhari, "AI-POWERED TRAFFIC SIGN REC0GNITION WITH VOICE ALERTS", (International Research Journal of Modernization in Engineering Technology and Science), Volume:07/Issue:05/May-2025.

[10] Yadav, Shubham & Patwa, Anuj & Rane, Saiprasad & Narvekar, Chhaya. (2022). Indian Traffic Sign Board Recognition and Driver Alert System Using Machine Learning. International Journal of Applied Sciences and 10.24071/ijasst. v1i1.1843.