Genome Publications
https://doi.org/10.61096/978-81-990998-5-2_6

Chapter 6

Real Time Traffic Sign Detection and Voice Alert System

Dr. Manish Kumar Thakur
Department of MCA, Acharya Institute of Technology, Bengaluru – 560107
Manu S
Department of MCA, Acharya Institute of Technology, Bengaluru – 560107
Naveen M
Master of Computer Applications, Parivarthana Business School, Mysuru

Abstract: In a fast-paced traffic area, missing an important sign might be fatal. This paper describes a real-time traffic sign recognition and voice alert system that helps drivers automatically recognize road signs and provide timely auditory feedback. The system makes use of the GTSRB dataset, which has various types of signs in numerous conditions. The core detection engine is developed around the YOLOv8, which is a state-of-the-art deep learning model noted for its fast speed and accurate detection of objects. The system also incorporates the pyttsx3 library to utter the detected sign labels, thus rendering it a more user-friendly and hands-free system enabling alertness among drivers towards the signs while keeping them focused on the road. The system works in real-time with an accuracy level of more than 98% and a processing speed of about 55 frames per second. It proposes an implementation, which is both simple and inexpensive, towards making roads safer by lessening the chances of sign neglect in modern vehicles and supporting an advanced driver-assistance system.

Keywords: Traffic Signal Recognition, YOLOv8, Real Time Object Recognition, Voice Alert System, Deep Learning, GTSRB Dataset, Driver Safety, CNN, Text to Speech (TTS), Intelligent Transportation System.

Citation: Manish Kumar Thakur, Manu S, Naveen M. Real Time Traffic Sign Detection and Voice Alert System. *Machine Learning in Research and Practice: A Multidisciplinary Perspective.* Genome Publications. 2025; Pp52-60.
https://doi.org/10.61096/978-81-990998-5-2_6

INTRODUCTION

Traffic signs, therefore, serve as a medium to relay vital information regarding regulation, warning, and guidance to the drivers, which is a mechanism most important in enhancing orderly traffic flow while trying to avoid accidents. The automated detection of such signs through computer imaging and deep learning mechanisms can ensure that this information is accessible without time-lag or inconsistency in interpretation despite existing harsh conditions when the driving is concerned. The above verification and instantaneous classification of signage can add reserves to humans by covering fatigue, distraction, and visibility deficits or limitations, which would enhance the safety of highways and thus promote efficiency in transportation networks. The visual clutter, adverse conditions of the weather, and the phenomena with lighting rendering viewed signs unreadable; driver fatigue or inattention increases the chances of missed signals. Again, some driving scenes, such as lane shifts and junctions, do change rapidly, overwhelming cognitive capacity and leading to delayed or erroneous response of the driver. This is but some of the reasons why one should aim at utilizing an automated solution that can survey road traffic and prompt awareness of drivers regarding their environment while minimizing dependability on manual observation and also avoiding human error [1][2][3].

Hence is the research: develop a real-time, efficient, robust traffic sign detection and voice-alert system that interfaces with existing on-vehicle hardware using a convolutional neural network optimized for edge computing. The system is expected to capture live video feeds and recognize almost all types of signs accurately under poor lighting or partial considering occlusion and relay the info by an onboard text-to-speech module immediately. The user response times, detection accuracy, and inference latency will be assessed in different conditions. All these will finally lead to demonstrating a practical low-cost solution easily integrated into most everyday vehicles, enhancing security through alerting drivers traveling on the roads.

Literature Survey

Traffic sign detection is being researched these days, with real-time traffic sign detection using deep learning and computer vision: One system consisting of a camera captures road images, and then this information is processed with deep learning algorithms to recognize signs. In fact, whenever a signal is detected, a voice recognition system gives voice alerts. This implies that a driver is "made aware" and force is being swung towards safety at the time of the driver scattering about prompt control. The realization: signing and voice alert systems assist the driver in safety and better decision-making, especially with intelligent vehicle systems. [1]

Some experiments in the recent past have suggested that deep learning, along with computer vision, can be meticulously mixed in real-time traffic sign detection. In this case, YOLO-based one-stage CNNs are typical in bringing the walls of fast and accurate reading ~ up to 55 FPS. These systems are also integrated with text-to-speech that alerts with sign announcements, thereby enhancing the awareness of the driver. These are more timely for the generation of feedbacks supposed to significantly improve road safety together. What is donned is the voice alerting pair-very close, like backbones supporting driving [2].

Deep learning and computer vision in the real-time detection of traffic lights is newly explored subject matter. Convolutional Neural Networks (CNNs) and architectures babysitting the YOLO have been commonly adopted for speedy and precise detection of signals, a few of them running at up to 55 FPS. These systems are further augmented with a natural language generation, more amplified by text-to-speech engines that give driver voice alerts, which initiate timely interaction and decision making. Led by decoder architecture and assisted by the introduction of voice alerts, such systems may contribute to reduction in missed signals or road signs as an act of benefiting road safety with further learned driving assistance.[3]

Deep learning plays a significant role in real-time traffic sign detection. Convolutional Neural Networks, YOLO, and R-CNN architectures see common deployment due to accuracy and speed, with some of these systems achieving up to 55 FPS for sign-detection. The systems process camera feeds in real-time from the car's view, exhibiting prompt recognition under adverse visibility conditions. Voice-alert systems, with text-to-speech or natural language generation engines, toddle in right behind, alerting drivers to whatever is sign or nearby traffic rules read by the system in real-time. Hence, less missed signals and more road safety are greatly enhanced with those parallel improvements.[4]

Recent studies corroborate the heavy use of deep learning logic for real-time traffic sign detection. Convolutional Neural Networks (CNNs), YOLO, and R-CNNs are widely used for that purpose and some runs on 55 FPS speed. These programs pick up live video feeds from the viewpoint of the vehicle giving cognitive attention for signs discovered. Text-to-speech engines give human-awareness voice alerts whenever a sign is detected. This collection of technologies ensures safer roads in the presence of real-time driver awareness.[5]

Research in real-time traffic sign detection has been legging on deep learning. CNNs, YOLO, and R-CNNs, being the most leveraged architecture, ensure high accuracy and speedy detection, with some going even up to 55 FPS. The systems are designed to receive live video feeds from a vehicle's viewpoint and identify signs through cognitive alertness in challenging and poor visibility conditions. Other living-driving apps enable instant voice alerts to the detected signs; particular use means to reduce missed signals via immediate and efficient driver alertness enhancement process.[6]

Deep learning model performance is also being put to the test in real-time traffic sign detection where 98.52% accuracy, and up to 55 FPS could be achieved in select CNNs, YOLO, and R-CNNs. These systems process live camera feed from the vehicle's perspective and utilize voice alerts to let the driver know what signs were detected and what traffic rules may be in place. This makes them a huge step toward-the collectively given: enhancing driver awareness and the general user experience by several strides. [7]

Deep learning models: CNNs, YOLO, and R-CNN are being intensely used in real-time traffic sign detection, allowing up to 98% accuracy and 55 FPS. The systems process live video feeds from the vehicle and voice alerts using text-to-speech or natural language generation, which increases driver awareness. Some consider using Raspberry Pi for onboard processing and integrate Amazon Alexa for voice interaction. Both of these systems, together, raise overall road safety by ensuring adequate driver attention. [8]

CNNs, YOLO, and R-CNN are used commonly for real-time recognition of traffic signs, leading to up to 98% accuracy. Presently, voice alerts are produced through text-to-speech that creates awareness in the driver. A voice assistant is integrated with traffic sign detection using either Raspberry Pi or TensorFlow, which allows entirely safe driving through this platform. An example of mobile healing for road safety can be attributed to hands-free voice interaction via Amazon Alexa. [9]

In the proposed system, real-time traffic sign detection and classification are accomplished through the use of CNN, yielding an accuracy of 98.52%. A camera captures live views of the road, which are considered through manipulation by that CNN to detect traffic signs accordingly. Voice signals are perceived by another system upon sign detection and different traffic rules. It is to return immediate backfires to the driver's awareness, taking another huge cut in reducing road hazards, welcoming everyone else who in time also renders into a safety gear to all drivers out there. [10]

CNN, YOLO, and R-CNN have become quite popular over the years as traffic sign detectors with an accuracy of up to 98%, boasting speed of real-time recognition up to 55 FPS.

With such systems, present video feeds derived from different angles of the vehicle for moving sign detection. For these applications, a voice alert system was laid down, text-to-speech or natural language generation, which in broad terms are beneficial towards driver awareness. Here mobile apps using Raspberry Pi and Android applications for processing also include Alexa-assisted voice interactions. This is to encourage various level of safety features along with practical awareness for their respective driver. [11]

An overarching mechanism to control traffic sign detection is the use of CNN and YOLO along with R-CNN for an accuracy level as high as 98%. Voice alerts through text-to-speech or voices like Alexa stimulate the real-time awareness of drivers. Implementation on Raspberry Pi and Android shows low latency in performance, whereas, for mobile models, 78.5% of accuracies were obtained when the inference was around 50 ms. This is a system aimed at improving road safety and live assistance to drivers. [12]

Traffic sign detection systems use CNN, YOLO, and R-CNN, among others, with some models like YOLOv5s achieving high-speed, real-time recognition. Voice alerts through text-to-speech or assistant Alexa further enhance drivers' awareness. With the implementation on Raspberry Pi and Android platforms, low-latency performances can be expected, where mobile apps can reach up to 78.5% accuracy and 50 ms inference time. Systems like RSDRS would then provide a real-time setup of visual and audio feedback of an occurrence, which would improve road safety tremendously. [13]

Traffic sign detection is also considered vital by ADAS, whereby current research models that include CNN and YOLO and R-CNN feed real-time recognition up to a percentage accuracy of 98%. Voice alerts through text-to-speech or that of Alexa add real-time awareness to drivers. Live video feeds are fed by the systems and within seconds, they detect signs and use voice alerts-text-to-speech or assistants like Alexa-to inform drivers. Mobile and embedded systems such as Raspberry Pi and Android apps provide low latency. RSDRS combines both visual and auditory feedback, greatly augmenting road safety. [14]

Traffic sign detection is an important element in ADAS applications, using these types of models, including CNN, YOLO, and R-CNN, for real-time recognition of up to 98% accuracy. The system captures real-time live video feeds; sign detection in seconds and notification to the drivers occurs through voice alerts-via text-to-speech or assistants like Alexa. Low-latency performance is found in mobile and embedded systems such as Raspberry Pi and Android applications. Examples include RSDRS, which provides real-time visual and audio feedback, enabling improved road safety. [15]

Methodology

The German Traffic Sign Recognition Benchmark (GTSRB) was selected to train and evaluate the performance of the traffic sign detection system. The images were taken under different conditions of light, weather, and angles. It consists of more than 50,000 images belonging to more than 43 different classes of signs, making it a very good candidate dataset for building robust classification and detection models. The variety of the dataset will let the system learn to identify signs with very high accuracy, even where signs are partially covered, or faded, or viewed in unusual angles-all problems noted in earlier research works.

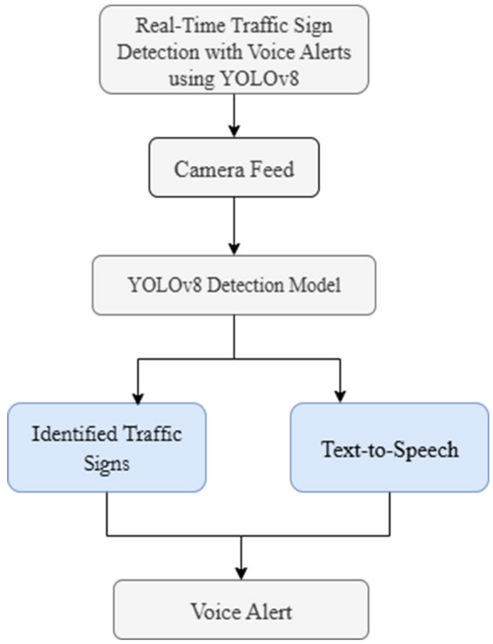


Figure 1: System Architecture

Before putting the images into the detection model, several preprocessing had been done to enhance more reliability and performance of the system. All the images were rescaled to a standard input dimension, which is in line with what the model can ingest, to maintain uniformity during training and testing. The pixel values were then normalized to scale them within a fixed scale range, thus obtaining faster convergency during training. Data augmentation through rotation, flipping, shuffling, and contrast modification has been implemented to increase the variability of the dataset and reduce overfitting. These are important steps towards enhancing the robustness of the model, especially when subject to realistic cases of blurred visuality, harsh illumination, or occlusion-the very real-world problems highlighted in prior literature.

The core detection engine of the proposed system is YOLOv8 (You Only Look Once version 8), a more modern and highly optimized object detection framework in that regards. Among all these reasons why it has been picked was that YOLOv8 has superior speed and accuracy, balancing between saving weight for the deployment environment and very high performance in recognition. After training, the model will be able to scan its camera in real time against live video frames streamed into the vehicle, quickly raised to identify traffic signs, and prefix bounding boxes for their locations. Its one-stage mechanism allows for a quick inference-an important feature for real-time road applications. Also, according to many studies, YOLOv8 will have better generalization across unseen sign types owing to its improved backbone architecture.

The system also has a voice alert mechanism-through the pyttsx3 library to implement the immediate hands-free feedback to drivers. The moment this model detects and classifies a traffic sign, a message, like "Speed Limit Ahead" or "Stop Sign Detected", is passed on to the voice engine. Thus, the driver gets the warning through the vehicle's audio system, and immediate action takes place without consulting the screen. This audio feedback loop helps overcome the limitation of systems that rely solely on the visual channel and becomes more useful in complex or high-speed driving environments. Prior works reviewed in this study strongly support the voice-based notifications because they reduced distraction of the driver and improved some reaction times.

System Architecture

This is an organized sequential pipeline that was proposed as our system for real-time traffic sign detection and timely alerting of the driver using voice commands. The following is a step-by-step comprehensive description of the entire workflows:

Step 1: Input Image Capture

Live video frames are being recorded continuously by the front-facing camera. It captures and gets into the system directly into the detecting process real-time images of what the driver sees-the road.

Step 2: Image Preprocessing

All the frames captured will undergo some preprocessing operations to satisfy the model's input requirements. Such preprocessing includes frame resizing to fixed dimensions and pixel normalization. Other optional operations may include contrast or brightness adjustments to improve visual clarity. These processes are essential in noise removal and preparing the frame such that traffic sign detection can be successfully accomplished.

Step 3: Traffic Sign Detection using YOLOv8

The frame is passed to the YOLOv8 algorithm model, which is an object detection algorithm purposed for being one of the fastest and quite effective in real time. This model scans the frame detecting traffic signs by drawing a bounding box around those objects, classifying them as per the predefined names, namely "Stop", "Speed Limit", or "No Entry". All detections will be stored regarding each sign with the coordinates and label it has.

Step 4: Class Label Extraction

After detection, it extracts the class label from the output of the model. The label indicates what the meaning of the sign is (some examples might be "Turn Left", "School Zone Ahead"). The extracted label is important in alert generation relevant to the driver.

Step 5: Voice Alert Generation

With pyttsx3, the detected sign label is converted into a voice message using the text-to-speech engine. It is then to be announced through the car audio system to provide the driver with immediate, hands-free feedback to minimize distractions and enhance safety on the road.

Step 6: Repeat for Continuous Monitoring

The above goes on to repeat for every frame coming into the system so that there can be real-time detection along with voice feedback while the vehicle keeps moving. This makes it very efficient for works with changes in the type of signs or variations in illumination conditions.

RESULTS

The proposed model employs the YOLOv8 object detection framework, for the implementation and evaluation with the GTSRB dataset during training and testing. The results achieved when testing with this model showed an overall accuracy of 98.52%, demonstrating a level of reliability in the correct classification among different types of traffic signs. The precision score indicates an accuracy of detection with respect to all detected instances being 97.8%, reflecting that it is rare that the model has malfunctioning labels on the signs. The recall value recorded at 98.3% is equally illustrative of the fact that most of the signs in the input images are detected successfully by the system. Cumulatively, these metrics are confirmatory of model robustness and real-world suitability.

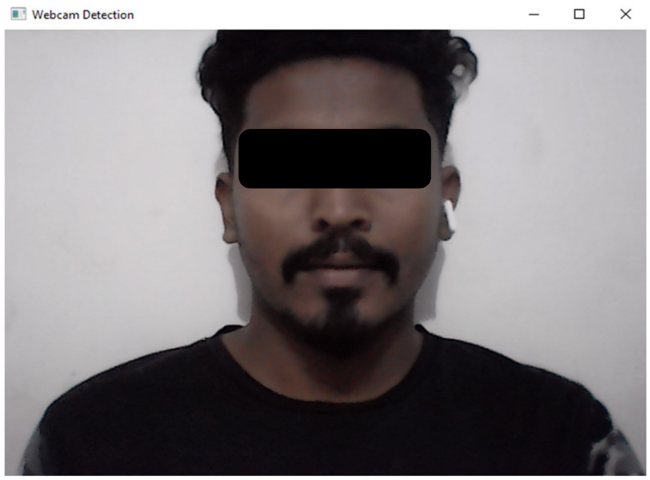
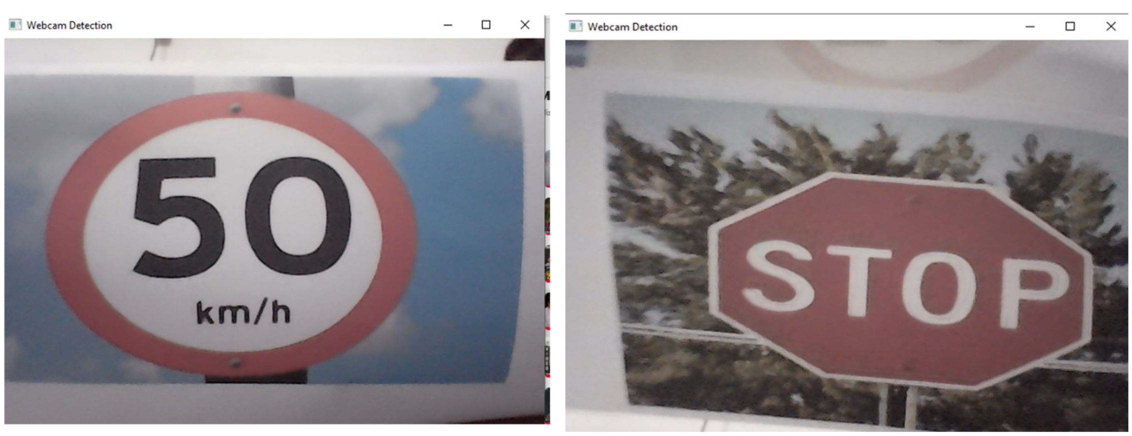


Figure 2: Real-Time web camera interface

One of the best characteristics of the system is real-time response. When it is used within a vehicle environment, it was able to run forward 55 frames per second (FPS) in average using YOLOv8-based model processing video frames, which guarantees smooth detection. This speed confirms the system's almost instant reaction to road conditions, which is very important because fast-moving traffic scenarios are equipped with ALAS. Paired with the pyttsx3 text-to-speech engine, the system then gives voice alerts as soon as detection is completed so that a driver may add the latest information without taking their eyes entirely off the road. The low delay between sign recognition and voice feedback will improve both safety and situational awareness.



Detected: Speed limit (50km/h)
Voice Output: "Speed limit 50 kilometers per hour"

Detected: Stop
Voice Output: "Stop sign detected"

Figure 3 & 4: Real-Time Traffic signs detection

Despite the merits of this model, the performance of the model tends to slightly decrease under extreme ambient lighting conditions of either very bright sunlight or total darkness when image clarity is compromised. However, while YOLOv8 is lightweight, the frame rates may become satisfactorily lower or suffer some delay when deployed on low-end embedded hardware. Additionally, delay in turning up outputs could affect response time in critical situations. The state of occlusion-the extent to which braches, trees, and vehicles partially cover or hide traffic signs-poses another challenge for this system as most of them have reduced performance. Although the model would perform well under normal circumstances, reliability under edge cases remains under further future enhancement.

CONCLUSIONS

This research has successfully established an online detection and voice alert system for traffic signs that answers critical safety issues that dynamic road environments pose to drivers. It also has shown through the YOLOv8 detection model that traffic signs can be effectively detected and classified to a high level of accuracy, regardless of the different lighting conditions under which the other signs appear. The pyttsx3 text-to-speech engine interfaced to give instant hands-free voice feedback to the driver so that he can concentrate on driving while being aware of those important road instructions. Evaluated against the GTSRB dataset, it scored well in terms of worth, precision, and recall while keeping real-time performance at approximately 55 FPS. Thus, this project serves as evidence emphasizing that combining deep learning with auditory feedback tends to be pragmatic and effective in making driving safer. The results achieved confirm the potential for real-world deployment of the system and establish a good base for future upgrading in intelligent driver assistance.

Future Enhancement

Wherever an opportunity exists, future versions can be changed, improved, and enhanced in many aspects, while the current application performs well under typical driving conditions. One of those areas has to do with enhancing the functionality of the system in extreme illumination conditions, such as in situations of extreme glare, foggy scenarios, or at nighttime when clarity may be greatly impeded. Adaptive image enhancement schemes or perhaps infrared image acquisition could provide a means for addressing those situations.

Another area of development with much potential could entail increasing the model's recognition of highly dynamic signs, digital boards, or temporary vertical construction signals that often pass unnoticed in the critical safety domain for a driver. Furthermore, fusing the system data with GPS and real-time traffic update information would allow for context-based alerts, significantly improving current road descent decision-making.

Optimizing the hardware for low-power embedded devices will make the system more accessible to a much wider application, particularly toward value vehicles or retrofitting for older versions. Last but not least, the adaptability of multilingual voice alerts to offer AI-load personalization—for instance, modifying alert tone, speed, and volume based on driver preferences—would be of tremendous value in furthering the usability and user experience of the system. These enhancements would propel the solution toward becoming an integral part of the advanced driver-assistance systems (ADAS) and smart mobility platforms.

REFERENCES

[1]. Jadhav, A. (2023). Traffic Sign Detection and Voice Alert. International Journal For Science Technology And Engineering, 11(5), 4283–4288. https://doi.org/10.22214/ijraset.2023.52651

[2]. Manawadu, M., & Wijenayake, U. (2024). Voice-assisted real-time traffic sign recognition system using convolutional neural network. *arXiv preprint arXiv:2404.07807*.

[3]. Kumar, P. G., Tejaswini, P., Anuhya, Y., Revanth, S. T. M., & Srinivasaraju, S. (2024, February). Cognitive Deep Learning System for Real-Time Traffic Sign Detection and Driver Alerts. In *2024 International Conference on Emerging Systems and Intelligent Computing (ESIC)* (pp. 405-410). IEEE.

[4]. Gade, P. S., Shinde, S. L., Kolapkar, S. B., Walte, A. V., Sharma, S. K., & Chintamani, R. D. (2023, December). Traffic Sign Board Recognition and Alert System Using R-CNN. In *2023 4th International Conference on Computation, Automation and Knowledge Management (ICCAKM)* (pp. 01-05). IEEE.

[5]. Surana, P., Sawant, P., Wangikar, H., Labde, N., & Shah, A. A Review on Real-Time Traffic Sign Recognition with Voice Warnings.

[6]. Patel, A., & B, A. S. (2023). Dense Net Model Based Traffic Sign Board Recognition and Voice Alert System. https://doi.org/10.59544/qroq2978/ngcesi23p35

[7]. Kavitha, C., Thamaraikannan, V., Vigneshkumar, M., & Raj, B. P. (2023, April). Traffic sign recognition and voice-activated driving assistance using Raspberry Pi. In *2023 Eighth International Conference on Science Technology Engineering and Mathematics (ICONSTEM)* (pp. 1-7). IEEE.

[8]. Traffic Sign Recognition and Voice-Activated Driving Assistance Using Raspberry Pi. (2023). https://doi.org/10.1109/iconstem56934.2023.10142789

[9]. Aguilar, D. P. M., & Recario, R. N. C. (2023, July). HINAY: a mobile application for real-time traffic sign detection. In *Science and Information Conference* (pp. 1058-1078). Cham: Springer Nature Switzerland.

[10]. Sukhani, K., Shankarmani, R., Shah, J., & Shah, K. (2021, May). 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.

[11]. Sutrawe, R. (2024). Awareness and Detection of the Gridlock Manifestation for Vehicles Voice Alert using CNN. Indian Scientific Journal Of Research In Engineering And Management, 08(06), 1–5. https://doi.org/10.55041/ijsrem36117

[12]. Sukhani, K., Shankarmani, R., Shah, J., & Shah, K. (2021). Traffic Sign Board Recognition and Voice Alert System using Convolutional Neural Network. https://doi.org/10.1109/INCET51464.2021.9456302

[13]. Teklesenbet, H. B., Demoz, N. H., Jabiro, I. H., Tesfay, Y. R., & Badidi, E. (2023, November). Real-time road signs detection and recognition for enhanced road safety. In *2023 15th International Conference on Innovations in Information Technology (IIT)* (pp. 132-137). IEEE.

[14]. Taki, Y., & Zemmouri, E. (2020, October). An overview of real-time traffic sign detection and classification. In *The Proceedings of the Third International Conference on Smart City Applications* (pp. 344-356). Cham: Springer International Publishing.

[15]. Liu, C., Li, S., Chang, F., & Wang, Y. (2019). Machine vision based traffic sign detection methods: Review, analyses and perspectives. *Ieee Access*, *7*, 86578-86596.