AI-powered Road Sign Recognition with Voice Alert

DISSERTATION

Submitted in partial fulfillment of the requirements of the

Degree: M.Tech in Artificial Intelligence & Machine Learning

By

Mattaparthi Lakshmi Gowri
2023AA05178

Under the supervision of

Harsh Tiwary
(Principal Manager)

BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE
Pilani (Rajasthan) INDIA

(May,2025)

BIRLA INSTITUTE OF TECHNOLOGY & SCIENCE, PILANI
SECOND SEMESTER 2024-25

AIMLCZG628T DISSERTATION

Dissertation Title: AI Powered Road Sign Recognition with Voice Alert

Name of Supervisor: Harsh Tiwary

Name of Student : Mattaparthi Lakshmi Gowri

ID No. of Student : 2023AA05178

Courses Relevant for the Project & Corresponding Semester:

1. Deep Neural Networks
2. Computer Vision
3. Natural Language Processing
4. Conversational AI

Abstract

In recent years, the integration of Artificial Intelligence (AI) into intelligent transportation systems has paved the way for significant advancements in road safety and driver assistance technologies. This dissertation proposes the design and implementation of a Real-Time Road Sign Recognition and Voice Alert System that leverages deep learning, computer vision, and text-to-speech technologies to detect, interpret, and audibly alert drivers of road signs in real-world driving scenarios.

The objective of this project is to enhance driver awareness and reduce road accidents caused by missed or misinterpreted road signs. The system uses a real-time video feed, typically from a dashboard camera, to continuously monitor the road for standard road signs such as speed limits, sharp turns, pedestrian crossings, school zones, and speed breakers. Upon detecting a sign, the system classifies it using a convolutional neural network (CNN) model trained on benchmark datasets such as the German Road Sign Recognition Benchmark (GTSRB) and further fine-tuned on region-specific road signs relevant to Indian roads.

The deep learning backbone of the system is implemented using the YOLOv5 architecture, which is known for its balance of speed and accuracy in object detection tasks. The detected road signs are processed in real-time, and corresponding voice alerts are generated using a text-to-speech (TTS) engine. For example, when a left turn sign is detected, the system provides an audible message such as, "Left turn ahead. Please drive carefully." This approach not only supports visual cognition but also augments it with auditory feedback, making it particularly beneficial during low-visibility conditions or for elderly and distracted drivers.

An additional innovation of this project lies in its adaptability. With minor adjustments, it can be extended to support multilingual voice alerts, region-specific sign variants, and additional road conditions such as lane merges, road congestion warnings, or temporary construction zones. This extensibility makes the solution scalable and applicable to broader intelligent driving and navigation systems.

This dissertation also explores relevant challenges, such as variations in lighting, occlusion, motion blur, and non-standard signage, and presents strategies to address these using data augmentation, robust pre-processing, and model tuning. Comprehensive testing is conducted under simulated and real-time driving environments to validate accuracy, latency, and user effectiveness.

In conclusion, the Real-Time Road Sign Recognition and Voice Alert System serves as an important step towards intelligent driver assistance. By combining robust computer vision with real-time audio feedback, it provides a practical and scalable solution aimed at minimizing road-related risks and enhancing the overall driving experience. This project contributes significantly to the domain of smart mobility and AI-powered road safety systems.

Key Words: Road Sign Recognition, Real-Time Object Detection, Deep Learning, YOLOv5, Voice Alert System, Computer Vision, Intelligent Transportation, Road Safety, Edge AI, Text-to-Speech (TTS)

BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE, PILANI
II SEMESTER 24-25 AIMLCZG628T DISSERTATION
Dissertation Outline

BITS ID No. 2023AA05178 **Name of Student:** Mattaparthi Lakshmi Gowri

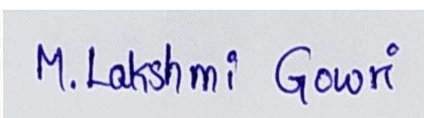
Name of Supervisor: Harsh Tiwary

Designation of Supervisor: Principal Manager

Qualification and Experience: B.Tech CSE and 13 years of experience in Software Development, Design and Architecture.

Official E- mail ID of Supervisor: htiwary@modeln.com

Topic of Dissertation: AI-powered Road Sign Recognition with Voice Alert



(Signature of Student)

Date: 21 May, 2025

(Signature of Supervisor)

Date: 21 May, 2025

Project Work Title

AI-powered Road Sign Recognition with Voice Alert

1.1. Purpose:

The purpose of this project is to develop a real-time road sign recognition system using deep learning, integrated with a voice alert module to assist drivers. By detecting road signs such as turns, speed limits, and warnings through a live camera feed and generating spoken alerts, the system aims to improve driver awareness and enhance road safety, especially in challenging driving conditions.

1.2. Expected Outcome:

- A real-time road sign recognition system using state-of-the-art object detection algorithms (e.g., YOLOv5).
- A functional voice alert module that uses detected road signs to generate audio prompts for drivers.
- A trained and validated deep learning model capable of detecting and classifying key road signs with high accuracy.
- Enhanced driver assistance functionality that supports situational awareness and road safety.
- An open-source, modular architecture that can be extended to support additional signs and languages (e.g., regional Indian languages).

2. Literature Review:

[1] Road Sign Recognition with Deep Learning: Research in road sign recognition has grown significantly with the introduction of Convolutional Neural Networks (CNNs). The German Road Sign Recognition Benchmark (GTSRB) is a widely used dataset, and studies using models such as LeNet, VGGNet, and YOLO have demonstrated high accuracy in classifying road signs in real-time driving environments.

[2] YOLO for Real-Time Object Detection: YOLO (You Only Look Once) has revolutionized object detection by enabling end-to-end detection in a single pass through the network. Studies indicate YOLOv5 is particularly well-suited for embedded and edge devices, balancing speed and accuracy for road applications.

[3] Voice Assistance in Driver Support Systems: Research shows that voice-based guidance systems significantly reduce cognitive load compared to visual interfaces alone. Integration of TTS (Text-to-Speech) in navigation aids and alert systems has been shown to improve reaction time and driver compliance.

[4] Intelligent Driver Assistance Systems (ADAS): Advanced Driver Assistance Systems are increasingly incorporating AI-based modules for pedestrian detection, sign recognition, and blind spot monitoring. These systems aim to reduce road accidents through early warning mechanisms and decision support systems.

[5] Edge Computing in Real-Time Vision Applications: Studies have emphasized the advantages of deploying AI models on edge devices like Jetson Nano and Raspberry Pi, citing reduced latency, improved data privacy, and cost-effectiveness for real-time vision applications.

3.1. Existing Process:

Currently, road sign recognition and driver assistance in most vehicles are either non-existent or limited to high-end Advanced Driver Assistance Systems (ADAS) integrated into premium vehicles. In the absence of such systems, drivers must manually observe and interpret road signs, which can be challenging due to:

- Poor visibility (e.g., at night, during rain or fog),
- Driver distraction or fatigue,
- Inconsistent or obscured signage.

Some existing ADAS systems offer visual warnings on dashboards but lack auditory feedback, making them less effective in situations where the driver's attention is divided. Moreover, such systems are typically proprietary, expensive, and not available for gener al or low-cost implementation.

3.2. Limitations:

- No Real-Time Assistance in Most Vehicles: Budget and mid-range vehicles often lack real-time road sign recognition or alert systems.
- No Audio Feedback: Existing systems, where present, often provide only visual cues, which may be missed when the driver is not actively looking at the display.
- Limited Accessibility: Premium driver assistance systems are expensive and not accessible to mass-market users or in developing regions.
- Poor Visibility Conditions: Road signs may not be visible due to weather, darkness, or occlusions like tree branches, reducing their effectiveness

4. Justification for Methodology:

- Use of CNN-based Detection Models: Deep learning models such as YOLOv5 are selected due to their real-time detection capabilities and lightweight architecture, suitable for both training and deployment phases.
- Text-to-Speech Integration: TTS is chosen to deliver immediate and intuitive feedback to the driver without requiring visual attention.
- Edge Deployment: The system is optimized for deployment on embedded hardware to ensure minimal latency and portability.
- Custom Dataset Augmentation: To improve performance under Indian road conditions, the model is trained on both standard datasets and augmented with Indian road sign imagery.
- Open-source Stack: The entire solution stack is based on open-source libraries and frameworks (PyTorch, OpenCV, gTTS), making it replicable, scalable, and cost-effective.

5. Project Work Methodology:

- Phase 1: Data Collection and Annotation
  Collect road sign images from datasets like GTSRB and local sources. Annotate them using tools like Labelling for object detection.

- Phase 2: Model Training and Evaluation
  Train YOLOv5 using custom-configured datasets. Evaluate performance using metrics such as mAP (mean Average Precision) and F1-score.
- Phase 3: Voice Alert Integration
  Use Python-based libraries like pyttsx3 or gTTS to convert detected sign classifications into voice alerts.
- Phase 4: Real-Time Video Feed Integration
  Integrate the model with OpenCV to process real-time video feeds from a webcam or dash camera.
- Phase 5: Testing & Evaluation in Simulated Environments
  Test system performance in varied lighting and movement conditions. Conduct latency and accuracy benchmarks.

6. Benefits Derivable from the Work:

Benefits:

- Enhanced Road Safety through real-time voice alerts
- Assistance to Drivers in poor visibility or distraction scenarios
- Real-Time Inference using efficient object detection algorithms
- Low-Cost, Scalable Solution with open-source and edge compatibility
- High Localization Potential by supporting Indian road signs and languages

7. Additional Details:

- Future Enhancements: While the current implementation is developed and tested on a high-performance computing system, a key future enhancement is to deploy the solution on embedded edge devices such as Raspberry Pi or NVIDIA Jetson Nano. This will allow for on-board real-time processing in vehicles, reducing dependency on cloud services and ensuring low-latency alerts. Deployment on edge hardware will also make the system more portable, cost-effective, and scalable for integration into commercial or personal vehicle dashboards.

1. Broad Area of Work: Computer Vision and Deep Learning for Intelligent Transportation Systems (ITS)

2. Objectives

The objectives of my project are as follows:

- To develop a real-time road sign detection and classification system using deep learning techniques, particularly YOLOv5.
- To create a voice alert module that generates audio messages corresponding to the detected road signs to assist drivers.
- To enhance the model's accuracy by training it on a combination of global (e.g., GTSRB) and Indian road sign datasets.
- To evaluate the system's real-time performance in various environmental conditions using webcam or video feeds.

3. **Scope of Work**

- Collect and preprocess road sign images from standard datasets and local Indian road environments.
- Annotate the images using bounding boxes for object detection training.
- Train and fine-tune a YOLOv5 model for accurate detection and classification of road signs.
- Develop a pipeline to process real-time video streams from a webcam or camera module.
- Integrate a Text-to-Speech (TTS) engine (e.g., gTTS, pyttsx3) to deliver voice-based alerts corresponding to detected signs.
- Test the complete system in various scenarios (e.g., different lighting and motion conditions) using recorded or live video.
- Analyze the model's accuracy, latency, and usability to assess its effectiveness in improving driver awareness and safety.

4. Detailed Plan of Work

Sno.	Tasks/Subtasks	Start Date - End Date	Planned Duration (weeks)	Specific Deliverable
1	Literature Survey and Market Research	Week 1 - Week 2	2	Summary of related works in road sign recognition and driver alert systems
2	Data Collection, & Processing	Week 3 - Week 4	2	Curated and annotated dataset of road signs (e.g., GTSRB + Indian signs), normalized images
3	Model Selection and YOLOv5 Setup	Week 5	1	Configured YOLOv5 environment for object detection, model parameters finalized
4	Model Training and Validation	Week 6 - Week 7	2	Trained YOLOv5 model with accuracy and loss metrics, visualizations of detection results
5	Model Testing with Real-Time Video Feed	Week 8 - Week 9	2	Real-time sign detection on webcam or dashcam footage using OpenCV
6	Integration of Voice Alert System (TTS)	Week 10	1	Functional module that triggers audio alerts based on recognized road signs
7	System Integration and GUI	Week 11 - Week 12	2	Fully integrated Python script or GUI application
8	Testing and Performance Evaluation	Week 13-Week - 14	2	Evaluation report with detection accuracy, latency, and use-case demos
9	Documentation and Report Preparation	Week 15 – Week 16	2	Complete dissertation report, code documentation, user guide, final presentation deck

5.Literature References:

The following are referred journals from the preliminary literature review.

[1]: Manawadu, M., & Wijenayake, U. (2024). Voice-Assisted Real-Time Traffic Sign Recognition System Using Convolutional Neural Network. *arXiv preprint arXiv:2404.07807* https://arxiv.org/abs/2404.07807

[2]: Beevi, S. L., Krishna, S. V. S., Sai, S. C., & Palagummi, S. (2022). Traffic Sign Board Recognition using Convolution Neural Network and Voice Alerting System. *International Journal of Engineering and Computer Science*, 14(2), 4179–4183. https://www.researchgate.net/publication/352668818_Traffic_Sign_Board_Recognition_and_Voice_Alert_System_using_Convolutional_Neural_Network

[3]: Jadhav, A., Samgir, S., Jagtap, A., Jadhav, A., & Bhadre, S. P. (2023). Traffic Sign Detection and Voice Alert. *International Journal for Research in Applied Science and Engineering Technology*, 11(5), 1234–1238. https://www.ijraset.com/research-paper/traffic-sign-detection-and-voice-alert

Supervisor's Rating of the Technical Quality of this Dissertation Outline

EXCELLENT / GOOD / FAIR/ POOR (Please specify): EXCELLENT

Harsh Tiwary

Date: 21 May, 2025

(Signature of Supervisor)

Name of the supervisor: Harsh Tiwary

Email Id of Supervisor: htiwary@modeln.com

Mob # of supervisor: 8147592747