

**VOICE ASSISTED REAL TIME TRAFFIC SIGN DETECTION AND  
RECOGNITION SYSTEM UNDER ADVERSE ENVIRONMENTAL  
CONDITIONS BASED ON YOLOV8**

P. H. S. Kumari<sup>a</sup> and H. M. M. Naleer<sup>a\*</sup>

<sup>a</sup>*Department of Computer Science, Faculty of Applied Sciences, South Eastern University of Sri Lanka, Sammanthurai, Sri Lanka.*

<sup>\*</sup>*hmmnaleer@seu.ac.lk*

**Abstract**

The world is currently in an era of AI and automation. Therefore, the development of AI technologies is increasing rapidly. Automated navigation technology is one aspect of rapidly developing AI technology. In consequence, many researchers focus on TSDR systems. A key factor affecting the accuracy of the TSDR system is the clarity of traffic signs. Haze, low light, and other atmospheric conditions can significantly degrade the visibility of signs. Traffic accidents often occur due to the inability of drivers to recognize road signs while driving. Therefore, clearly and accurately recognizing signs is important for both drivers and pedestrians, thereby ensuring their safety. Drivers can lose focus while driving for various reasons. In such cases, providing a voice alert about a traffic sign can help bring the driver's attention back to driving. However, existing methods fail to perform haze removal, TSDR, and voice warning simultaneously. Therefore, in this work, a TSDR system has been developed with a deep learning-based HRU-Net algorithm with a voice assistant. According to the proposed pipeline, the HRU-Net model takes haze images as input and produces a dehazed image as output. The TSDR model then uses this haze-free image as input. After detecting and classifying that image, the traffic sign is fed into a gTTS. It generates a concise, real-time voice alert. It enables drivers to receive critical sign information without diverting their attention from the road. The proposed system was evaluated using the CURE-TSD dataset, which contains roughly 45,000 traffic sign instances based on 43 categories. Those images were captured under a wide range of environmental conditions. In the dehazing stage, the model achieved a Mean Absolute Error (MAE) of 0.0526, a Structural Similarity Index Measure (SSIM) of 0.8442, and a Peak Signal-to-Noise Ratio (PSNR) of 20.55 dB, around 50 epochs. In the YOLOv8 detection and classification stage, enhanced images are used for training, which results from the dehazing step. In this step, the model reached 99.07% precision, 99.13% recall, mAP@0.5 of 99.38%, and mAP@0.5:0.95 of 85.69% at the optimal 40th training epoch. The voice alert module achieved an average latency of ~230 ms between detection and audio playback. This voice alert module provides clear, concise feedback. When compared to existing methods, the proposed system provides superior accuracy and responsiveness. This model gives a robust and practical solution for advanced driver-assistance systems in adverse visual environmental conditions.

**Keywords:** *Traffic Sign Detection and Recognition (TSDR), HRU-Net, You Only Look Once (YOLOv8), Convolutional Neural Network (CNN), Voice Feedback*