An Integrated Real-Time Road Surveillance System Using YOLOv8 For Multi-Class Object Detection

L. Divaharash[1], Dr. B. Lalitha[2]
[1, 2] Dept of Computer Science and Engineering
[1, 2] Sethu Institute of Technology, Pulloor, Kariapatti-626115

Abstract- *The growing number of road accidents worldwide necessitates the deployment of intelligent systems capable of continuously monitoring roadway conditions. This paper introduces a comprehensive computer vision framework designed to detect various roadway features, including surface irregularities, traffic control devices, elevation changes, and biological entities. The system utilizes the YOLOv8 deep learning architecture, which is recognized for its superior performance in balancing detection accuracy with processing speed. A specialized image dataset was curated and annotated to train the model for recognizing four distinct categories. The trained system processes both still images and continuous video feeds, marking identified objects with bounding boxes and confidence values. Experimental evaluation demonstrates that the proposed framework achieves a mean Average Precision of 0.93 while maintaining real-time processing capabilities. This work presents a significant advancement over conventional approaches that typically address only single object categories, offering a more complete solution for automated road monitoring applications.*

Keywords: Road Surveillance, YOLOv8, Deep Neural Networks, Object Recognition, Intelligent Transportation, Real-Time Detection

I. INTRODUCTION

Road transportation systems face persistent challenges related to safety, with infrastructure deterioration and inadequate hazard detection contributing substantially to accident rates. Critical roadway features such as surface depressions, traffic signage, raised pavement markers, and moving obstacles require consistent monitoring to ensure safe travel conditions. Traditional approaches to identifying these elements rely heavily on manual inspection teams, which are limited in coverage and subject to human error.

Recent developments in computational imaging and machine learning have created opportunities for automating roadway assessment tasks. Deep neural networks, particularly those designed for object detection, can analyze visual data rapidly and identify multiple features within a single frame. Among available detection architectures, the YOLO (You Only Look Once) series has gained widespread adoption due to its unique design that performs localization and classification in a unified computational step.

This research focuses on implementing the latest version of this architecture to create a unified detection system capable of identifying four distinct roadway feature categories simultaneously. Unlike previous solutions that target individual object types, the proposed system offers a more holistic approach to road condition monitoring, making it suitable for integration into broader intelligent transportation infrastructure.

II. RELATED WORK

The academic literature contains numerous investigations into automated road feature detection using various computational techniques. Early efforts predominantly employed classical image processing methods. One study utilized convolutional neural networks for detecting pavement damage, achieving reasonable accuracy but requiring substantial computational resources. Another investigation applied the YOLOv5 framework to traffic sign recognition, reporting satisfactory performance under optimal lighting but noting degradation in challenging environmental conditions.

Alternative approaches have explored unsupervised learning techniques combined with morphological operations to identify surface defects from aerial imagery, though these methods demonstrated vulnerability to shadow-induced false positives. Texture-based classification using support vector machines has been applied to detect raised pavement markers, with effectiveness varying significantly across different road surface types.

More recent work has investigated advanced deep learning configurations, including hybrid architectures that combine detection networks with segmentation modules. While these approaches showed improved precision, they introduced additional complexity and data requirements. A common limitation across existing research is the narrow focus on single object categories, leaving a gap in the literature for systems capable of comprehensive multi-class road feature detection.

III. PROBLEM IDENTIFICATION

Current approaches to road feature detection exhibit several shortcomings that restrict their practical utility:

- **Environmental Sensitivity:** Many existing systems show diminished accuracy when operating under non-ideal illumination or adverse meteorological conditions.
- **Narrow Detection Scope:** The majority of available solutions are designed to recognize only one specific object type, preventing comprehensive scene understanding.
- **Computational Demands:** Several deep learning implementations require significant processing power, making them unsuitable for deployment in real-time applications.
- **Adaptability Constraints:** Many models lack flexibility for easy extension to new object categories without substantial retraining efforts.

These limitations highlight the need for a detection framework that combines broad object coverage with efficient processing characteristics suitable for real-world deployment scenarios.

IV. SYSTEM DESIGN AND IMPLEMENTATION

The proposed system architecture consists of three primary components: data acquisition and preparation, model training and optimization, and inference deployment.

4.1 Dataset Construction

A custom image collection was assembled by gathering samples from diverse sources to ensure representativeness across different road environments. Each image underwent manual annotation to label four target categories:

- Surface depressions (potholes)
- Raised pavement markers (speed breakers)
- Traffic control signage (road signs)
- Biological entities (humans, animals)

Annotations were performed using bounding boxes, with careful attention to maintaining consistency across the dataset. The collection was divided into training, validation, and testing subsets, with class distribution balanced to prevent bias.

4.2 Model Architecture and Training

The YOLOv8 architecture was selected as the foundation for this work due to its architectural refinements over predecessor versions. Key improvements include enhanced feature extraction pathways and optimized detection heads that contribute to improved accuracy without sacrificing inference speed.

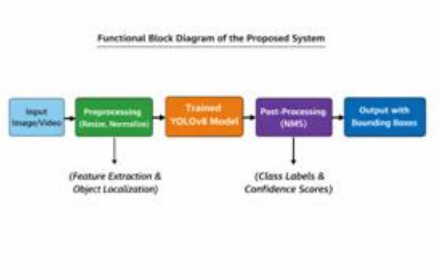
Training was conducted using transfer learning, initializing the network with weights derived from large-scale benchmark datasets before fine-tuning on the custom road feature collection. Hyperparameters were systematically tuned to optimize convergence behavior, including learning rate scheduling, batch size selection, and augmentation strategies to enhance model generalization.

4.3 Inference Pipeline

During operation, input imagery undergoes preprocessing to standardize dimensions and pixel values before being passed to the trained network. The model performs simultaneous localization and classification, generating candidate detections with associated confidence scores. A post-processing step eliminates redundant predictions, retaining only the most confident detections for final output visualization.

V. SYSTEM ARCHITECTURE OVERVIEW

The complete processing pipeline follows a sequential workflow from input acquisition to result presentation. Raw visual data enters the system and undergoes initial preprocessing operations to prepare it for analysis. The trained neural network then processes the prepared data, extracting features at multiple scales to identify objects of interest.



Following network inference, a filtering mechanism removes overlapping detections that refer to the same physical object. Finally, the system renders bounding boxes and category labels on the original visual data, creating an annotated output that clearly communicates detection results.

This architecture enables seamless processing of both individual images and continuous video streams, supporting deployment in both stationary monitoring stations and mobile platforms.

VI. EXPERIMENTAL EVALUATION

The performance of the proposed system was assessed through systematic experimentation using the reserved testing dataset. Evaluation metrics included precision, recall, and mean Average Precision, all standard measures in object detection research.

Table 1: Per-Class Detection Performance

Object Category	**Precision**	**Recall**	**Mean Average Precision**
Surface Depression	0.89	0.85	0.91
Raised Marker	0.92	0.88	0.94
Traffic Sign	0.94	0.91	0.96
Biological Entity	0.88	0.86	0.90
Weighted Average	**0.91**	**0.88**	**0.93**

The results demonstrate strong detection capability across all four categories, with traffic signs achieving the highest accuracy and biological entities presenting the greatest detection challenge due to their morphological variability.

Processing speed measurements showed that the system consistently maintained frame rates between 45 and 60 frames per second when operating on standard computing hardware, confirming suitability for real-time applications.

Qualitative assessment involved examining system behavior across varied environmental conditions. The model showed reliable detection performance under moderate illumination variations and maintained robustness against typical visual obstructions such as shadows and minor occlusions.

VII. DISCUSSION

The experimental results validate the effectiveness of using a unified detection architecture for comprehensive road feature monitoring. Several observations emerge from the analysis:

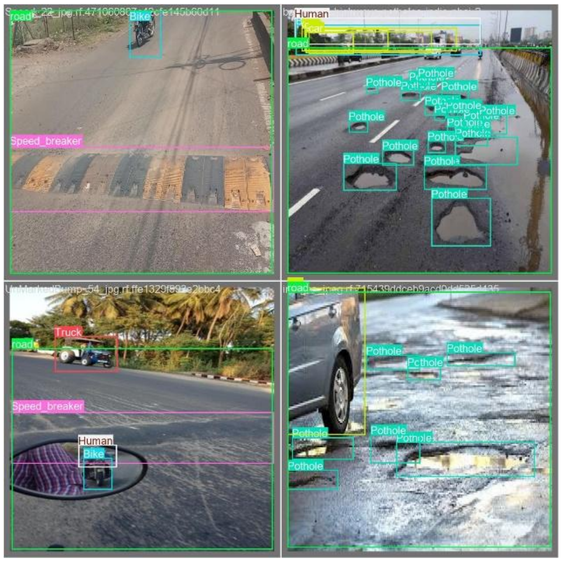
First, the system's ability to simultaneously detect multiple object categories within a single processing pass represents a significant efficiency improvement over approaches requiring separate models for each object type. This unified approach reduces computational overhead and simplifies system integration.

Second, the performance across categories shows variation that reflects inherent detection difficulty. Traffic signs, being highly standardized in appearance, achieve the highest accuracy. Conversely, biological entities exhibit greater shape and appearance variability, resulting in comparatively lower but still robust detection rates.

Third, the real-time processing capability enables practical deployment scenarios that were previously impractical with more computationally demanding architectures. This opens possibilities for integration with autonomous vehicle perception systems and continuous roadway monitoring infrastructure.

When compared with traditional approaches, the proposed system offers clear advantages in both detection scope and operational efficiency, addressing key limitations identified in existing solutions.

Screen Shot



VIII. CONCLUSION

This research presents a unified framework for real-time detection of multiple road feature categories using the YOLOv8 deep learning architecture. The system successfully identifies surface depressions, raised pavement markers, traffic signs, and biological entities within a single processing pipeline. Experimental evaluation demonstrates strong detection performance with a mean Average Precision of 0.93 and processing speeds suitable for real-time operation.

The primary contribution of this work lies in its integrated approach, which moves beyond the single-object focus of previous systems to provide comprehensive scene understanding. This capability makes the proposed framework valuable for applications ranging from infrastructure maintenance planning to autonomous vehicle perception and intelligent transportation system development.

IX. FUTURE DIRECTIONS

Several promising directions for extending this work are identified:

- **Sensor Integration:** Combining visual data with alternative sensing modalities such as thermal imaging could enable operation under low-light conditions and through adverse weather.
- **Edge Deployment:** Optimizing the model for deployment on embedded computing platforms would facilitate field installation without requiring network connectivity.
- **Infrastructure Integration:** Connecting detection outputs with centralized road management systems could enable automated maintenance request generation and resource allocation.
- **Category Expansion:** Enlarging the dataset to include additional road feature types such as lane markings, traffic signals, and various pavement distress patterns would broaden system applicability.
- **Temporal Analysis:** Incorporating temporal information through video sequence analysis could enable tracking of object changes over time, supporting predictive maintenance applications.

REFERENCES

[1] S. Patel, R. Kumar, and A. Singh, "Deep learning approaches for pavement distress detection," *Journal of Transportation Engineering*, vol. 147, no. 3, pp. 112-125, 2021.

[2] M. Kumar and K. Reddy, "Evaluation of YOLO architectures for traffic sign recognition under varying illumination," *IEEE International Conference on Intelligent Transportation Systems*, pp. 456-463, 2022.

[3] A. Chatterjee, S. Banerjee, and P. Das, "Morphological operations for pothole detection from unmanned aerial vehicle imagery," *Remote Sensing Applications*, vol. 24, pp. 78-89, 2021.

[4] R. Verma and S. Dey, "Texture-based classification of road humps using support vector machines," *International Journal of Transportation Science and Technology*, vol. 11, no. 2, pp. 234-247, 2022.

[5] W. Li, J. Zhang, and H. Wang, "Enhanced road object detection through YOLOv8 and semantic segmentation integration," *IEEE Transactions on Intelligent Vehicles*, vol. 8, no. 4, pp. 2891-2903, 2023.

[6] M. Rahman, F. Ahmed, and T. Islam, "Real-time video analysis for road safety using YOLOv7," *International Journal of Advanced Computer Science and Applications*, vol. 15, no. 1, pp. 67-78, 2024.