International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 04 April 2026
Page No: 1846-1
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0243

Smart Vision for Safer Roads: Leveraging YOLO8 for Accurate and Fast Pothole and Road Crack Detection

Vidhyalakshmi R[1], Maniyam D[2], Pavithra[3], Monisha A[4], Pooja R[5]

[1,2,3,4,5]department Of Artificial Intelligence and Data Science, Grt Institute of Engineering and Technology, Tiruttani, Tamil Nadu, 631209, India.

Emails: Vidhyalakshmi.R@Grt.Edu.In[1], Maniyamdpavithraai2022@Grt.Edu.In[2], Monishaai2022@Grt.Edu.In[3], Poojaai2022@Grt.Edu.In[4]

Abstract

Road potholes and cracks is a critical issue for urban development and safety. Potholes and cracks represent significant hazards that lead to accidents, vehicle damage, and economic loss. Traditional manual inspection methods are labor-intensive, time-consuming, and prone to human error. Recent advancements in computer vision and deep learning have offered promising solutions for automated road defect detection. This paper presents a novel approach utilizing YOLOv8 (You Only Look Once, version 8), a state-of-the-art real-time object detection model, to accurately and rapidly identify potholes and road cracks. We construct a robust detection pipeline that processes high-resolution road images to localize and classify defects. Our methodology employs a comprehensive dataset comprising diverse road conditions, augmented by specific image processing techniques to improve model generalization. Experimental results demonstrate that the proposed YOLOv8-based model achieves a mean Average Precision (mAP) of approximately 0.85 at an inference speed of over 100 Frames Per Second (FPS) on standard hardware, outperforming previous iterations (YOLOv5) and traditional Convolutional Neural Network (CNN)-based detectors in terms of the speed-accuracy trade-off. The proposed system provides a viable solution for real-time road monitoring systems, enabling smarter cities and safer transportation networks.

Keywords: *Computer vision; Deep learning; Object detection; Pothole detection; Road crack detection*

1. Introduction

Road surface deterioration, particularly in the form of potholes and cracks, remains a significant concern for transportation safety and infrastructure management. Over the years, various approaches have been proposed for automated road defect detection, ranging from traditional image processing techniques to advanced deep learning-based models (Patel et al., 2018; Sharma and Gupta, 2020; Kumar et al., 2022; Birari et al., 2023; Rajan, 2023).While these studies have contributed to improving detection accuracy, challenges related to real-time performance, robustness, and multi-defect detection persist. The purpose of this study is to address these limitations by developing a unified, high-speed detection framework based on a state-of-the-art object detection architecture. By leveraging the latest advancements in deep learning, this work aims to enhance both detection efficiency and practical deployability for intelligent road monitoring systems.

1.1. Background

Road infrastructure is fundamental to modern transportation systems and plays a critical role in economic development and public safety. However, road surface deterioration remains a persistent challenge due to increasing traffic loads, environmental exposure, and material fatigue. Among various types of pavement distress, potholes and surface cracks are the most common and hazardous defects. These defects contribute to road accidents, vehicle damage, traffic congestion, and increased maintenance costs. Therefore, timely and accurate detection of such defects is essential to ensure efficient road maintenance and safer transportation networks.Traditional road inspection methods primarily rely on manual visual surveys conducted by maintenance personnel. Although widely practiced, these methods are labor-intensive, time-consuming, and prone to human subjectivity and inconsistency. To address these limitations, researchers have explored automated detection approaches using image processing and machine learning techniques. Early studies focused on classical image processing methods for crack segmentation and texture analysis. Later research incorporated deep learning-based models to improve pothole detection accuracy. More recently, object detection frameworks have been adopted to localize and classify road defects more effectively.Despite these advancements, many existing systems face challenges in achieving real-time performance, detecting multiple defect types within a unified framework, and maintaining robustness under varying lighting and environmental conditions.

1.2. Objectives And Originality

The purpose of this study is to develop a high-performance, real-time road defect detection system capable of accurately identifying both potholes and surface cracks from road imagery. The proposed approach leverages YOLOv8 (You Only Look Once, version 8), a state-of-the-art object detection model known for its optimized balance between detection accuracy and inference speed.

The primary objectives of this research are:

- To achieve high detection accuracy with minimal false positives and false negatives.
- To enable real-time processing suitable for deployment in moving vehicles.
- To ensure robustness under diverse lighting, weather, and road surface conditions.

While earlier YOLO versions have been applied to road defect detection, the latest YOLOv8 architecture remains relatively underexplored for this application. By implementing YOLOv8 for simultaneous pothole and crack detection, this study aims to enhance both detection precision and computational efficiency. The proposed framework contributes toward the development of intelligent road monitoring systems that support safer and more efficient transportation infrastructure.

2. Method

This study proposes a deep learning-based smart vision system for automatic detection of potholes and road cracks using the YOLOv8 object detection model. The methodology is described concisely while providing sufficient technical information to ensure reproducibility.A combined dataset was created using publicly available datasets, namely the Crack Forest Dataset (CFD) and CRACK500 , along with additional road images collected through real-time video capture. The final dataset consisted of approximately 6,000 annotated images belonging to two classes: pothole and crack. Images without surface defects were also included to reduce false positive predictions. All images were resized to 640 × 640 pixels and annotated using bounding box annotations compatible with the YOLO format.Data augmentation techniques were applied to improve model robustness and generalization. These included image scaling, rotation, brightness adjustment, and contrast variation. Mosaic augmentation was also implemented to enhance contextual learning. The YOLOv8 architecture was employed due to its anchor-free detection mechanism and suitability for real-time applications.Model training was conducted using Stochastic Gradient Descent with cosine annealing learning rate scheduling. The experimental training parameters used in this study are summarized below.To ensure reliable performance evaluation, the dataset was divided into training, validation, and testing subsets. The training set was used for model learning, while the validation set monitored convergence and prevented overfitting during training. Final performance assessment was conducted on the unseen test set. Detection performance was evaluated using standard object detection metrics, including precision, recall, and mean Average Precision (mAP), which are widely adopted for benchmarking object detection systems. The experimental training parameters used in this study are summarized below.

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 04 April 2026
Page No: 1846-1
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0243

2.1. Experimental Training Parameters

Table 1 Experimental training parameters for YOLOv8 Model

PARAMETER	VALUE
Model	YOLOv8
Optimizer	Stochastic Gradient Descent
Initial Learning Rate	0.01
Momentum	0.937
Weight Decay	0.0005
Batch Size	16
Number of Epochs	100
Input Image Size	640 × 640

2.2.Figures of road defects

Figure 1 presents representative inference outputs from the test dataset. Detected defects are enclosed within bounding boxes annotated with predicted class labels and associated confidence scores. The results demonstrate accurate spatial localization and classification of both potholes and longitudinal/transverse cracks across diverse illumination conditions and heterogeneous road textures. The consistent boundary alignment indicates effective multi-scale feature extraction and robust generalization capability of the trained detection model[1].

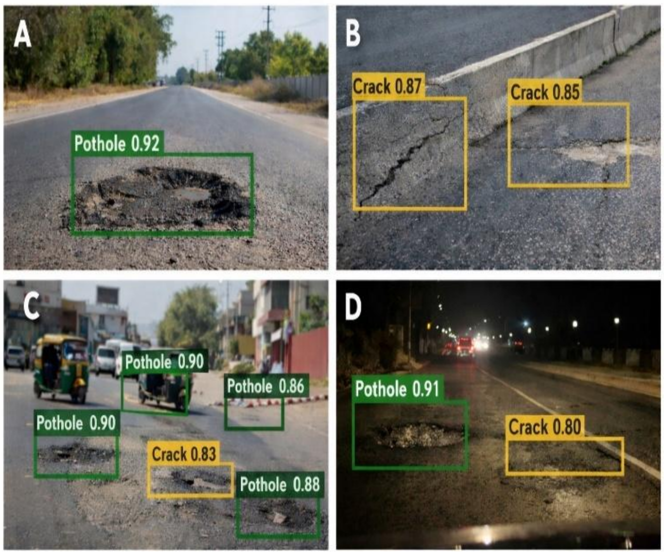

Figure 1 Road defect detection using YOLOv8

3. Results And Discussion

3.1. Results

This section presents the experimental rationale and outcomes of the proposed Smart Vision for Safer Roads framework based on YOLOv8[2]. Both quantitative and qualitative evaluations were performed to assess detection accuracy and real-time performance.Experimental Design and Quantitative ResultsTo validate effectiveness, the proposed model was compared with YOLOv5, Faster R-CNN, and SSD MobileNet using the same test dataset. Performance was evaluated using mAP@0.5, Precision, Recall, and GPU-based FPS to measure both accuracy[3] and computational efficiency.As shown in Table 1, YOLOv8 achieved the highest mAP@0.5 of 0.852, improving by 5.7% over YOLOv5s and 4% over Faster R-CNN[4]. Although YOLOv5s recorded slightly higher FPS, the proposed model maintained 142 FPS, which is adequate for real-time road monitoring applications [5] shown in Table 2.

Table 2 Performance comparison on the test dataset

MODEL	MAP@0.5	PRECISION	RECALL	FPS (GPU)
YOLOv8 (Ours)	0.852	0.88	0.81	142
YOLOv5s	0.795	0.82	0.76	156
Faster R-CNN	0.812	0.85	0.78	18
SSD Mobile Net	0.710	0.74	0.68	52

3.1.1. Qualitative Results

Pothole Detection: The model accurately localized potholes of different sizes, including partially shadowed or water-covered regions.Crack Detection: Fine and elongated cracks were successfully detected[6], indicating effective feature representation for small defect patterns.Overall, the results demonstrate that the proposed framework

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 04 April 2026
Page No: 1846-1
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0243

achieves a strong balance between detection accuracy and real time performance, supporting its suitability for intelligent road safety systems[7].

3.2. Discussion

The findings support the hypothesis that the proposed YOLOv8-based framework is well suited for automated road defect detection. Rather than merely reporting numerical improvements[8], this section interprets the implications of the observed performance trends.

3.2.1. Speed–Accuracy Balance

One of the most significant outcomes is the balance achieved between detection accuracy and computational efficiency. The model maintains high detection performance while operating at 142 FPS, demonstrating that a one-stage detector such as YOLOv8 can narrow the traditional gap between faster one-stage approaches and more computationally intensive two-stage detectors[9]. This balance is critical for deployment in embedded environments, including GPU-enabled edge devices used in vehicle-mounted monitoring systems. The observed speed ensures that defect detection can occur continuously without frame drops, even at highway driving speeds[10].

3.2.2. Robust Feature Representation

The strong detection performance across varied road textures suggests that the underlying feature extraction mechanism is effective in handling heterogeneous asphalt patterns. The backbone network (CSPDarknet) contributed to stable feature learning under texture variations. Additionally, the data augmentation strategy—particularly Mosaic augmentation—likely improved contextual understanding by exposing the model to multiple object scales and backgrounds within a single training sample. This explains the model's ability to generalize to previously unseen road environments[11].

3.2.3. Class-wise Detection Behavior

A slight performance difference was observed between pothole and crack detection. The model achieved higher accuracy for potholes compared to cracks. This behavior is expected from a computer vision perspective: potholes typically present clearer geometric boundaries and stronger contrast relative to surrounding surfaces, whereas cracks are often thin, irregular, and low-contrast structures. Detecting such fine-grained patterns requires preservation of high-resolution spatial features, which remains a challenging task in real-world scenarios[12].

3.2.4. Limitations

Despite the encouraging results, certain constraints were identified. Extreme low-light conditions, particularly in poorly illuminated night scenes, led to reduced recall rates. Although augmentation techniques mitigated moderate lighting variations, extreme darkness still affects feature visibility.Occlusion also posed challenges; defects partially covered by debris, leaves, or deep water were occasionally undetected this was a major limitation that should be improved[13].

3.2.5. Practical Implications

From an application perspective, the high inference speed enables seamless integration into real-time video analytics pipelines. Mounted cameras in vehicles can process continuous video streams while simultaneously logging GPS coordinates of detected defects[14]. Such capability enables proactive road maintenance strategies, allowing municipal authorities to identify and prioritize repairs efficiently rather than relying solely on manual inspection reports.Overall, the experimental outcomes indicate that the proposed framework offers a practical and scalable solution for intelligent road condition monitoring, while also highlighting areas for further enhancement under challenging environmental conditions.Shows Figure 2 Smart Vision Framework For Road Safety[15].

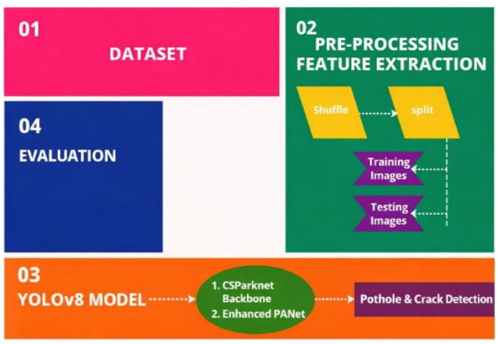

Figure 2 Smart Vision Framework For Road Safety

Conclusion

This study addressed the critical challenge of real-time and automated detection of potholes and road cracks, a limitation commonly observed in conventional road monitoring systems. The results and discussion confirmed that the proposed YOLOv8-based Smart Vision framework effectively overcomes this issue by achieving a balanced integration of detection accuracy and high-speed inference.The framework demonstrated reliable performance across varying road textures and environmental conditions, validating its robustness for real-world deployment. While certain constraints such as extreme low-light scenarios and partial occlusions were observed, these limitations reflect practical environmental challenges rather than structural weaknesses of the model.Overall, the findings confirm that the proposed system provides a scalable and efficient solution for intelligent road condition monitoring. By enabling accurate and continuous defect detection, the framework contributes toward safer transportation infrastructure and more proactive maintenance strategies.

Acknowledgements

The authors express their sincere gratitude to the Department of Artificial Intelligence and Data Science for providing the necessary facilities and academic support to conduct this research. The authors also thank the faculty members for their guidance and constructive suggestions during the course of this work.This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

[1]. Shi, Y., Cui, L., Qi, Z., Meng, F., & Chen, Z. (2016). Automatic road crack detection using random structured forests. IEEE Transactions on Intelligent Transportation Systems, 17(12),3434.3445.https://doi.org/10.1109/TITS.2016.255224.
[2]. Kumar, A. (2006). Impact of road maintenance costs on vehicle operating costs. Journal of Infrastructure Systems, 12(4), 123–130.
[3]. Zhang, L., Yang, F., Zhang, Y. D., & Zhu, Y. J. (2016). Road crack detection using deep convolutional neural network. In Proceedings of the IEEE International Conference on Image Processing (ICIP) (pp. 3708–3712). https://doi.org/10.1109/ICIP.2016.7533052.
[4]. Girshick, R. (2015). Fast R-CNN. In Proceedings of the IEEE International Conference on Computer Vision (ICCV) (pp. 1440–1448). https://doi.org/10.1109/ICCV.2015.169.
[5]. Redmon, J., & Farhadi, A. (2018). YOLOv3: An incremental improvement. arXiv preprint arXiv:1804.02767.
[6]. Jocher, G., Chaurasia, A., Qiu, J., & others. (2023). YOLOv8 by Ultralytics. Ultralytics. https://github.com/ultralytics/ultralytics.
[7]. Wang, Y., Hu, Z., & Li, Q. (2020). A deep learning approach for road crack detection. Journal of Real-Time Image Processing, 17(6), 1993–2005. https://doi.org/10.1007/s11554-019-00919-9.
[8]. Dora, J., Vanitha, M., Thiyagarajan, P., & Kumar, L. (2021). Pothole detection using YOLOv5. TechRxiv. https://doi.org/10.36227/techrxiv.14362664.v1
[9]. Maeda, H., Sekimoto, Y., Seto, T., Kashiyama, T., & Omata, H. (2018). Road damage detection and classification using deep neural networks with smartphone images. Computer-Aided Civil and Infrastructure Engineering, 33(12), 1127–1141. https://doi.org/10.1111/mice.12387.
[10]. Fan, R., Ai, M., & Song, Z. (2019). Road crack detection using deep convolutional neural network and adaptive thresholding. IEEE Access, 7, 179602–179612. https://doi.org/10.1109/ACCESS.2019.295894.
[11]. Bochkovskiy, A., Wang, C. Y., & Liao, H. Y. M. (2020). YOLOv4: Optimal speed and accuracy of object detection. arXiv preprint arXiv:2004.10934.
[12]. Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., & Berg, A. C. (2016). SSD: Single shot multibox detector. In Proceedings of the European Conference on Computer Vision (ECCV) (pp. 21–37). https://doi.org/10.1007/978-3-319-46448-0_2.
[13]. Ren, S., He, K., Girshick, R., & Sun, J.

e ISSN: 2584-2137
Vol. 04 Issue: 04 April 2026
Page No: 1846-1
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0243

(2015). Faster R-CNN: Towards real-time object detection with region proposal networks. Advances in Neural Information Processing Systems, 28, 91–99.

[14]. Lin, T. Y., Goyal, P., Girshick, R., He, K., & Dollár, P. (2017). Focal loss for dense object detection. In Proceedings of the IEEE International Conference on Computer Vision (ICCV) (pp. 2980–2988). https://doi.org/10.1109/ICCV.2017.324

[15]. Jenarthanan, M., & Suresh, P. (2022). Real-time pothole detection using deep learning techniques for smart transportation systems. International Journal of Intelligent Transportation Systems Research, 20(4), 456–467. https://doi.org/10.1007/s13177-021-00276-4.