Undergraduate Conference on Intelligent Computing and Systems (UCICS 2025)
26-27 February, 2025; Varendra University, Rajshahi, Bangladesh

Real-Time Pothole Detection Using YOLOv8-seg: A Deep Learning Approach to Smart Road Monitoring

MD. Jahid Hassan
Department of Information and Communication Technology
Islamic University, Kushtia-7003, Bangladesh
jahid.ict.iu@gmail.com

Romana Nourin Nipa
Department of Information and Communication Technology
Islamic University, Kushtia-7003, Bangladesh
romananourin.iu.ict@gmail.com

Md Faisal Ahammed
Department of Information and Communication Technology
Islamic University, Kushtia-7003, Bangladesh
faisal.ahammed.ict.iu@gmail.com

Ahsan Habib
Department of Information and Communication Technology
Islamic University, Kushtia-7003, Bangladesh
ahsan017533@gmail.com

MD. Alamgir Hossain
Department of Information and Communication Technology
Islamic University, Kushtia-7003, Bangladesh
hossain@iu.ac.bd

***Abstract*—Transportation systems are anchored by road infrastructure directly impacting economic growth, safety, and mobility. Despite this, the poor conditions on roads, particularly potholes, are impediments to road degradation which causes, increased vehicle operating costs, heightened accident risks, and decreased transportation flexibility. G e nerally, p o thole detection and road maintenance are expensive and time-consuming with conventional methods that depend on manual inspections or specialized sensors. This study addresses these challenges using an automated pothole detection system based on an advanced deep learning model optimized for real-time object detection and segmentation: You Only Look Once version 8 (YOLOv8 seg). The model was trained on a large dataset originating from Kaggle, with annotated images and video data, ensuring that it performs well in lit or unlit, rainy or sunny, conditions. Model generalization had been enhanced by data preparation, involving augmentation methods including flipping, r o tation, a n d brightness modifications. S t rong d e tection a b ilities w e re demonstrated by the proposed YOLOv8-seg model, which achieved a fitness score of 0.92 and a mean Average Precision (mAP) of 0.72 for both bounding boxes and masks. The system is appropriate for real-time road monitoring applications due to its consistent precision-recall curves, low latency, and effective segmentation accuracy, based on performance inquiry. In addition, this integrative method makes it easy to integrate with maintenance procedures, enabling prompt repairs as well as improving infrastructure management and road safety in general. The outcomes show how deep learning-based systems can revolutionize traditional road monitoring procedures into effective, scalable, and reasonably priced solutions.**

***Index Terms*—Pothole detection, YOLOv8, Deep learning, Real-time monitoring, Road infrastructure, Object segmentation.**

I. Introduction

Roads are the core of the transport system infrastructure since they are the essential links that bring people together and thus enable economic activities and mobility. The state of roads affects the day-to-day activities, the issue of safety, and on rate at which events unfold, thus the need to repair the roads as often as required. The problem with road monitoring and solving pothole issues, thus, remains unattended for extended periods because they are random. Cracks that are caused by environmental conditions, traffic load, or inadequate maintenance not only deteriorate infrastructure but also lead to increased usage costs of vehicles, traffic accidents, and reduced transportation productivity [1].

Pavement distress can generally be classified i nto three categories: pavement deformation (shoving, corrugation, and rutting), fatigue cracks (primary, secondary fatigue spalling and final), and Mechanism of pavement failure which includes: raveling and stripping. And while all these asphalt failings persist, potholes are still the largest problem since they occur randomly and are felt throughout the community. Officially, governments and agencies from different countries invest billions of dollars every year in road management and repair [2]. However, pothole-related incidence is still a worrying issue because global records show that on average, 1.25 million people die in road traffic a ccidents a nnually, 3 4% o f which is attributable to potholes [3].

To cope with these challenges, in this study, we propose a smart road monitoring system that incorporates real-time pothole detection using You Only Look Once version 8 - Segmentation (YOLOv8-seg) and maintenance integration.

Conventional pothole detection techniques, including manual inspection or specialized sensors, are costly, time-consuming, and ineffective for widespread use. The goal of this project is to automate the pothole identification process in real-time with high accuracy, low latency, and scalability using deep learning techniques.

The primary objectives of this research are as follows:

1) Design and implement a You Only Look Once version 8 - Segmentation (YOLOv8-seg) based system for the automated detection of potholes in real-time video streams acquired from road monitoring systems.
2) Evaluate the performance of the proposed model in terms of accuracy, detection speed, and reliability under varying real-world conditions, including diverse lighting, road textures, and weather scenarios.
3) Integrate the pothole detection system with road maintenance protocols to enable timely reporting, facilitate prompt repairs, and optimize infrastructure management processes.

The remainder of this paper is organized as follows: Section 2 is a survey of pothole detection and monitoring systems as developed in prior work. Section 3 describes the source of data, model structure, and measures of performance. Section 4 displays experimental results and analysis. Last of all, concluding section 5 summarizes the study and discusses directions for further research.

II. Literature Review

Road monitoring and in particular detection of potholes play a major role in the safety and preservation of infrastructure. Conventional approaches to inspecting for defects are invasive and can be more inaccurate than automated ones [4]. Sharma et al. [5] employed automatic methods including techniques such as edge detection and morphological operations, but they were restricted by issues with lighting and even the weather. Glennie et al. [6] attempted to solve it using accelerometers, gyroscopes, and Laser Imaging Detection and Ranging (Lidar) systems. Even though these methods can identify road anomalies, these methods are costly and hardware-bound, thereby hindering their applicability. Many improvements are observed in the detection of potholes, especially after the integration of machine learning. Random Forest by Wu et al. [7] have been worked out. However, these approaches entail feature engineering, and may not be ideal for different road conditions as this study combines normal road and unpaved road conditions. Amid the state-of-the-art deep learning techniques, Convolutional Neural Networks or CNNs have greatly advanced the cause of pothole detection through their effective end-to-end solutions providing opportunities for automatic feature learning [8]. Object detection approaches like You Only Look Once (YOLO) [9], and Single Shot Detector (SSD) [10] are shown to provide great performance in pothole detection in terms of smart road monitoring. Of these, SSD is special due to its non-confined ability in real-time detections, and it can detect multiple objects in a single forward pass. Perhaps this makes SSD particularly appropriate for real-time pothole detection.

III. Methodology

In this section, we describe the approach to developing a real-time pothole detection system. We describe the dataset preparation, preprocessing techniques, and the model used (YOLOv8-seg) in detail. To ensure compatibility with the models, the dataset has been preprocessed by preloading road images and road videos with their annotations.

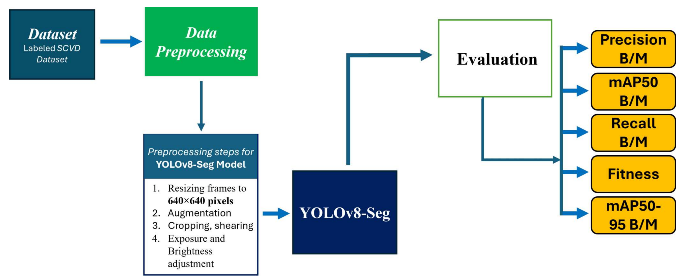

Fig. 1: Graphical Representation of our Overall Research Framework

A. Data Collection

The dataset used in this research was sourced from Kaggle's Pothole Segmentation for Road Damage Assessment dataset (https://www.kaggle.com/code/farzadnekouei/pothole-segmentation-for-road-damage-assessment/input). It consists of 1,680 annotated images and a single video file, with a total size of 62.72 MB. This dataset is structured into training and validation sets, along with a configuration file (data.yaml) that defines the classes and paths for the training and validation data. It also includes a sample video (sample-video.mp4) for testing and demonstration purposes. The images capture diverse road conditions, including variations in lighting, weather, and surface textures, ensuring robust training of the proposed model.

B. Data Preprocessing

The resolution of the dataset images was fine-tuned so they were properly aligned and uniformly resized to 640x640 pixels. To improve the model's learning capabilities we applied several augmentations to the training images. Flipping, cropping, rotation, shearing, brightness, and exposure adjustments were among these augmentations. Applying these preprocessing techniques, we sought to enhance the model's independence of model realizations in the face of a diversity of conditions and also its ability to automatically identify and segment potholes.

C. Model Selection and Training Process

This study develops a real-time pothole detection system using the YOLOv8-seg method, optimized for road monitoring. The system is evaluated using Precision, Recall, Mean Average Precision (mAP), and Fitness to assess its accuracy and effectiveness. YOLOv8-seg is chosen for its ability to detect and segment potholes of varying sizes and shapes in live video, enabling efficient road monitoring.

1) YOLOv8 (Proposed): The latest version of the YOLO series of object detection models, YOLOv8, serves as a strong benchmark for real-time pothole detection. The model is renowned for its fast processing speed and accuracy, achieved through advanced architecture and optimization techniques. Compared with SSD, which employs multiple feature maps to predict bounding boxes, YOLOv8 directly outputs bounding boxes and class probabilities on a single feature map. This single forward pass makes YOLOv8 one of the fastest object detectors available.

Unlike the traditional anchor box approach, YOLOv8 employs an anchor-free box prediction strategy, where bounding box regression is performed without predefined anchor boxes. This reduces the computational overhead associated with anchor box generation and refinement. Mathematically, the bounding box prediction is expressed as follows:

$$b = (\sigma(t_x) + c_x, \sigma(t_y) + c_y, w \cdot e^{t_w}, h \cdot e^{t_h}) \tag{1}$$

where b represents the predicted bounding box coordinates, (c_x, c_y) are the center offsets, (w, h) are the prior box dimensions, and (t_x, t_y, t_w, t_h) are the predicted offsets. The sigmoid function σ ensures that the predictions remain within the appropriate spatial range.

YOLOv8 is specifically designed to address real-world object detection variability. To further enhance the model's ability to focus on the critical features within the image, we employ Convolutional Block Attention Modules (CBAM).

IV. Result and Discussion:

This study focuses on the performance analysis of our proposed model YOLOv8-seg, conducting a comprehensive evaluation across several key aspects. The following analyses are performed to assess the model's effectiveness:

1) Learning Curve Analysis: The model's performance is evaluated through several metrics during training and validation. The box-loss metrics track the accuracy of bounding box predictions, while the seg-loss metrics reflect the model's segmentation accuracy. The cls-loss measures classification performance, and the dfl-loss assesses the model's ability to handle challenging instances. Precision and recall metrics for bounding boxes (B) and masks (M) gauge the model's prediction accuracy and detection capabilities. The mAP50 and mAP50-95 metrics summarize the model's accuracy for both bounding boxes and masks at different IOU thresholds.

2) Confidence Threshold Metrics Analysis: The Precision-Confidence Curve demonstrates near-perfect precision at all confidence levels, reflecting highly accurate predictions for both bounding boxes and masks. The Recall-Confidence Curve maintains a high recall across all confidence thresholds, indicating that the model consistently identifies true positives. The F1-Confidence Curve shows a stable and high F1 score, suggesting a well-balanced precision and recall, with the model performing effectively even at higher confidence thresholds.

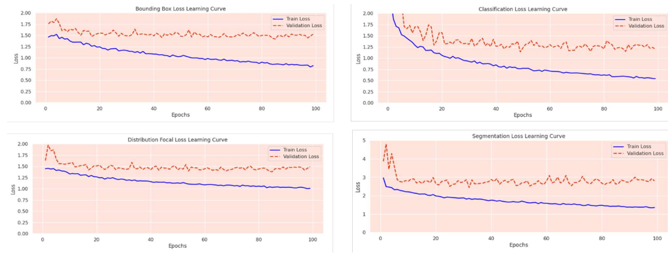

Fig. 2: YOLOv8-seg Training and Validation Loss over Epochs

3) Confusion Matrix Analysis: The normalized confusion matrix for our YOLOv8 pothole segmentation model provides valuable insights into its performance. The model has a 70% true positive rate, meaning it correctly identifies potholes 70% of the time. However, there is a 30% false negative rate, indicating that some potholes are missed during detection. Given the challenging nature of pothole segmentation, this performance is commendable.

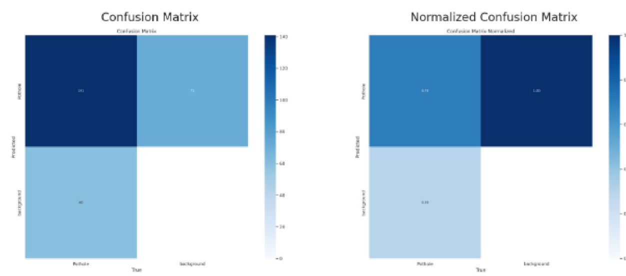

Fig. 3: Confusion Matrix and Normalized Confusion Matrix for Proposed Model

4) Validation Performance Metrics: The YOLOv8-seg model for pothole segmentation demonstrates promising performance in validation. The precision metrics for bounding boxes (B) and masks (M) are 0.75 and 0.71, respectively, indicating accurate predictions. While recall scores are slightly lower, with 0.61 for bounding boxes and 0.66 for masks, this suggests that some true potholes may be missed. The mAP50 scores are robust, achieving 0.72 for both bounding boxes and masks, indicating solid prediction accuracy. However, the mAP50-95 scores reveal that there is room for improvement in the consistency of predictions across different IoU thresholds. The overall fitness score of 0.92 reflects a well-trained model, with potential for further fine-tuning to optimize performance.

5) Comparison With Other Study: Table I compares object detection models based on mAP performance. YOLOv7 and Faster R-CNN achieved 55.6 and 64.12, respectively, while the proposed YOLOv8-seg outperformed both with 72.4 mAP. This improvement is due to anchor-free detection, segmentation capabilities, and CBAM integration, enabling precise pothole localization. Despite higher computational demands, YOLOv8-seg maintains real-time efficiency, making it ideal for road monitoring applications.

TABLE I: Comparison With Other Study

Ref	Method	Result (mAP)
[11]	YOLOv7	55.6
[12]	Faster R-CNN	64.12
This Work	YOLOv8-seg	**72.4**

A. *Real-Time Road Damage Assessment*

In this study, we utilized evolutionary algorithms for both image and video inference to enhance the performance of the YOLOv8-seg model. These algorithms simulate natural selection processes, iteratively refining the model by selecting the best-performing solutions from a population of candidates. This process not only improved the accuracy of the model's predictions but also contributed to better generalization and robustness across different inference tasks. [Figure 4] illustrates the model's performance in detecting potholes on static images, demonstrating high accuracy across various lighting and road conditions. Minor challenges include slight performance drops in extreme lighting and occasional false positives. [Figure 5] showcases real-time video inference. The model effectively detects potholes despite challenges like motion blur and occlusions.

Metric	Value
precision(B)	0.751
recall(B)	0.617
mAP50(B)	0.724
mAP50-95(B)	0.440
precision(M)	0.708
recall(M)	0.657
mAP50(M)	0.721
mAP50-95(M)	0.427
fitness	0.924

TABLE II: Performance Evaluation of Proposed Model Across Different Evaluation Metrics

Fig. 4: Inference Validation On Images

Fig. 5: Inference Validation On Real-Time

V. Conclusion

In this study, a robust real-time pothole detection system was developed using YOLOv8-seg deep learning model. The model managed to achieve promising precision and recall metrics with a mean Average Precision (mAP) of 0.72 both for bounding boxes and masks, through the use of diverse datasets and advanced preprocessing techniques. The model exhibited high detection accuracy under different roads and weather conditions, but some limitations like missed detections (false negatives) do show areas for future improvements. Finally, we provide real-world validation and integration with maintenance workflows, demonstrating that the system can enhance road monitoring processes, increase safety, and reduce infrastructure costs. Future work should apply evolutionary algorithms as well as further model fine-tuning to improve consistency and performance in broader scenarios.

References

[1] N. G. Sorum, T. Guite, and N. Martina, "Pavement distress: a case study," *International Journal of Innovative Research in Science, Engineering and Technology*, vol. 3, no. 4, pp. 274–284, 2014.

[2] I. G. Heggie and P. Vickers, *Commercial management and financing of roads*. World Bank Publications, 1998, vol. 23.

[3] A. M. Legreid, "Potholes and strategies on the road to campus internationalization." *International Research and Review*, vol. 6, no. 1, pp. 1–9, 2016.

[4] J. Yu, J. Jiang, S. Fichera, P. Paoletti, L. Layzell, D. Mehta, and S. Luo, "Road surface defect detection—from image-based to non-image-based: A survey," *IEEE Transactions on Intelligent Transportation Systems*, 2024.

[5] H. Sharma and N. Kanwal, "Video surveillance in smart cities: current status, challenges & future directions," *Multimedia Tools and Applications*, pp. 1–46, 2024.

[6] V. Pellegrino, V. Godard, and J. Fleury, "Investigation of the use of topographic data derived from pléiades imagery for high-resolution hillslope-scale morphometry," *Geomorphology*, vol. 454, p. 109162, 2024.

[7] P. Singh, A. E. Kamal, A. Bansal, and S. Kumar, "Road pothole detection from smartphone sensor data using improved lstm," *Multimedia Tools and Applications*, vol. 83, no. 9, pp. 26 009–26 030, 2024.

[8] Y. Safyari, M. Mahdianpari, and H. Shiri, "A review of vision-based pothole detection methods using computer vision and machine learning," *Sensors*, vol. 24, no. 17, p. 5652, 2024.

[9] N. Bhavana, M. M. Kodabagi, B. M. Kumar, P. Ajay, N. Muthukumaran, and A. Ahilan, "Pot-yolo: Real-time road potholes detection using edge segmentation based yolo v8 network," *IEEE Sensors Journal*, 2024.

[10] A. Paramarthalingam, J. Sivaraman, P. Theerthagiri, B. Vijayakumar, and V. Baskaran, "A deep learning model to assist visually impaired in pothole detection using computer vision," *Decision Analytics Journal*, vol. 12, p. 100507, 2024.

[11] M. Jakubec, E. Lieskovská, B. Bučko, and K. Zábovská, "Comparison of cnn-based models for pothole detection in real-world adverse conditions: overview and evaluation," *Applied Sciences*, vol. 13, no. 9, p. 5810, 2023.

[12] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," *IEEE transactions on pattern analysis and machine intelligence*, vol. 39, no. 6, pp. 1137–1149, 2016.