Enhancing road damage detection performance using the YOLOv9 model

Muhammad Farkhan Adhitama, Sutikno, Rismiyati
Department of Informatics, University of Diponegoro, Semarang, Indonesia

Article Info

Article history:

Received Aug 7, 2025
Revised Apr 3, 2026
Accepted Apr 15, 2026

Keywords:

Computer vision
Deep learning
Road damage
YOLOv9
You only look once

ABSTRACT

Roads are essential infrastructure that support community mobility, and their condition significantly impacts road user safety. However, manual road damage detection remains inefficient, time-consuming, costly, and prone to human error. To address this issue, this study proposed the YOLOv9 model for automated road damage detection and explored parameter combinations to optimize its performance. The proposed solution leverages the YOLOv9 model, which offers enhanced detection speed and accuracy compared to previous YOLO versions, due to its improved backbone and dynamic label assignment techniques. The method uses pre-trained weights and performs parameter tuning to adapt the model for identifying common road defects, including potholes, longitudinal, lateral, and alligator cracks. A publicly available dataset of road condition images was used for training and evaluation. Experimental results demonstrated that the optimized YOLOv9 model achieved a mean average precision (mAP) of 62.8%, indicating a promising ability to detect multiple types of road damage accurately. This study highlights the potential of YOLOv9 as an effective tool for road monitoring systems, contributing to proactive maintenance strategies and more efficient infrastructure management.

Corresponding Author:

Sutikno
Department of Informatics, Faculty of Science and Mathematics, University of Diponegoro
Prof. Jacob Rais Street, Tembalang, Banyumanik, Semarang, Indonesia
Email: sutikno@lecturer.undip.ac.id

1. INTRODUCTION

Roads play a crucial role in supporting social and economic activities, including mobility, goods distribution, and access to essential services. Well-maintained roads are essential for transportation safety and efficiency, making road maintenance a critical issue. Roads are a key piece of infrastructure that supports development but often suffer significant damage, requiring effective management [1]. Traditional road damage surveys have drawbacks, including high costs and low efficiency, as they involve manual inspections that are time-consuming, labor-intensive, and prone to human error. These limitations highlight the need for automated and efficient solutions. Computer vision technology offers a promising alternative by enabling the automatic detection of road damage at a lower cost, with higher efficiency and greater accuracy through image and video analysis [2].

Research on road damage detection employs image processing techniques, machine learning (ML), and deep learning (DL) methods. Threshold segmentation and edge detection identify crack features [2]. Sutikno *et al.* [3] employed edge detection techniques and thresholding in preprocessing, along with backpropagation neural networks, for classifying road damage. Arya *et al.* [4] used the gray-level co-occurrence matrix (GLCM) for feature extraction. To improve accuracy, Daneshvari *et al.* [5] combined local binary pattern (LBP) and GLCM feature extraction with the K-nearest neighbors (KNN) classifier. In ML, support vector machines (SVMs) and neural networks rely on hand-crafted features [6], whereas the light gradient boosting machine effectively categorizes crack patterns [7]. DL methods, particularly convolutional neural networks (CNNs), achieve high detection accuracy [8]. You only look once (YOLO) enables real-time damage classification [9], [10], and models like RetinaNet and elman neural networks (ENN) are used for high-resolution detection [11]. Generative adversarial networks (GANs) facilitate crack segmentation [12], enhancing overall detection systems.

The YOLO algorithm is widely used in research on road damage detection. Wan *et al.* [13] developed YOLO-LRDD, an optimized version of YOLOv5s, reducing the model weight by 28.8% and achieving a mean average precision (mAP) of 57.6%. Ruseruka *et al.* [14] used YOLOv5 for pothole detection, achieving 93% precision and an mAP of 96.3%. Meanwhile, Zhao *et al.* [15] combined a lightweight network and YOLOv5. Ye *et al.* [16] introduced YOLOv7-AMF for crack detection, achieving 91.3% precision and 94.1% mAP. Additionally, RDD-YOLO, an enhanced version of YOLOv8, was presented by Li *et al.* [17], achieving a mAP50 of 62.5%. While proposed enhancements to the YOLOv8n algorithm to achieve real-time performance while maintaining high detection accuracy and robustness, others have focused on alternative approaches [18], [19].

YOLOv9s has been utilized for highway damage detection, with a focus on highways in China [20]. This study proposes YOLOv9 for road damage detection, leveraging a combination of parameters, and selects it due to its superior capabilities in visual analysis. The dataset used in this study was collected from roadways in Indonesia and includes four types of road damage: alligator cracking, lateral cracking, longitudinal cracking, and potholes. Compared to classical ML approaches that rely on hand-crafted features and multi-stage pipelines, YOLOv9 offers an end-to-end DL framework that automatically learns discriminative visual features and performs real-time object detection with higher accuracy and robustness in complex visual environments. YOLOv9 was chosen for its advantages in image detection and better performance compared to previous YOLO algorithms, such as YOLOv7 and YOLOv8, on object detection using the MS COCO 2017 dataset [21]. YOLOv9 has also proven reliable across several detection tasks, including dam crack detection [22] and vehicle damage detection [23], [24]. An automatic road-damage detection approach is expected to facilitate decision-making in road maintenance. Therefore, this study contributes to improving the accuracy of detecting specific types of road damage, such as potholes, longitudinal cracks, and alligator cracks, using the YOLOv9 model.

2. METHOD

The proposed method for road defect detection was divided into several steps, as shown in Figure 1. The input of this method was a highway image dataset, and the result was the detected image. The proposed method was divided into five processes: dividing data, image resizing, model training, model validation, and model testing.

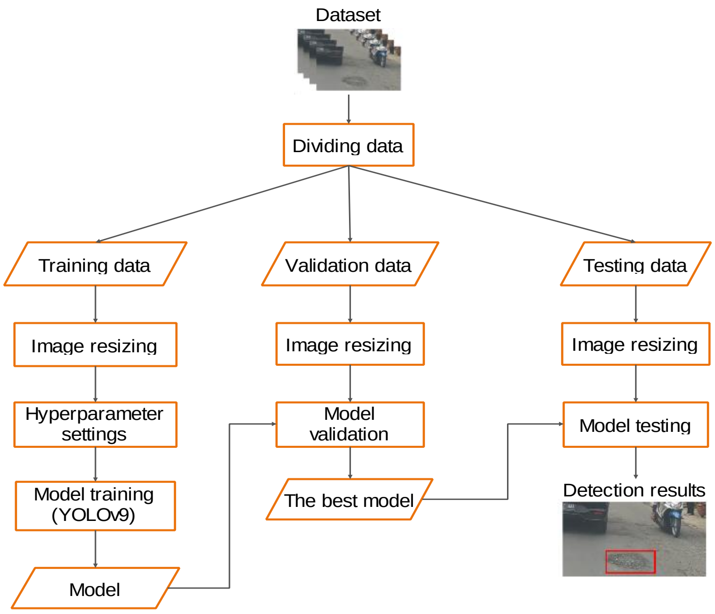
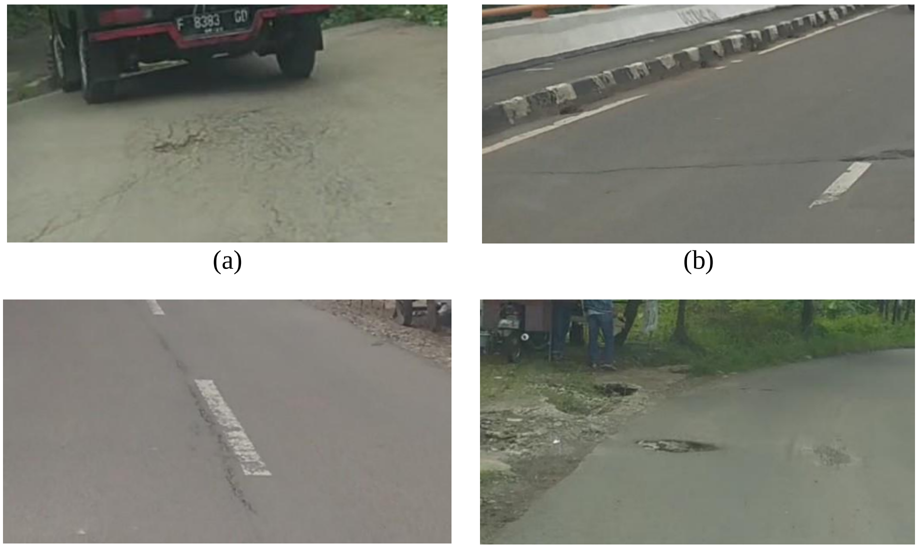


Figure 1. Research method for road damage detection using YOLOv9

2.1. Dataset

This research used a road-damage dataset consisting of images of paved roads with surface damage. The dataset is publicly available on Kaggle [25]. This dataset contains images of paved roads with surface damage measuring 1080 × 1920 pixels. Figure 2 shows sample images from the road-damage dataset used in this study, covering four damage classes. Figure 2(a) shows alligator cracking, characterized by interconnected cracks resembling a reptile's skin. Figure 2(b) presents lateral cracking, which runs perpendicular to the direction of traffic. Figure 2(c) depicts longitudinal cracking that follows the direction of traffic flow. Figure 2(d) displays a pothole, a surface depression caused by water infiltration and repeated traffic loads.



(a) (b) (c) (d)

Figure 2. Dataset example: (a) alligator cracking class, (b) lateral cracking class, (c) longitudinal cracking class, and (d) pathole cracking class

2.2. Dividing data

The dataset was split into training, validation, and test sets. The training data accounted for 80% of the total dataset. Meanwhile, the validation and testing data each comprised 10% of the total. This split ratio is widely adopted in ML research to ensure sufficient data for model training while retaining representative subsets for hyperparameter tuning and unbiased evaluation on unseen data [26].

2.3. Image resizing

Image resizing in the road damage image dataset does not change the image's aspect ratio. The size of the image dataset before resizing is 1080 × 1920 pixels. After the image resizing process, the image size changes to 288 × 512 pixels and 360 × 640 pixels. Resizing reduces computational burden, memory usage, and training time, while retaining sufficient visual information for effective feature learning.

2.4. Hyperparameter setting

Hyperparameters will be configured before training the YOLOv9 model. The optimal configuration of hyperparameters plays a crucial role in accelerating training, reducing memory usage, and improving model performance. The hyperparameters that will be modified are the learning rate, batch size, image size, and number of epochs. Other hyperparameter values are fixed. Tuning these hyperparameters helps balance training efficiency, computational cost, and model generalization. Other hyperparameters are fixed to ensure training stability and enable a fair evaluation of the effects of the selected parameters.

2.5. Training and validation

Training the road damage detection model begins by feeding the training data into the model. Training was conducted for 150 epochs with a batch size of 16. The image size and learning rate were adjusted based on the scenario to be run. In YOLOv9, the model is trained using pre-trained weights from the

MS COCO dataset. The model validation process will also run at each completed epoch during training. The validation process produces several values: precision, recall, and mAP. The best pavement damage detection model is selected based on the highest mAP value obtained during validation.

The YOLOv9 model training process will pass through blocks in the YOLOv9 network, which are divided into three parts: the backbone, neck, and head. The backbone extracts initial features from the input image and progressively improves the feature representation [27]. The neck combines and refines the features extracted by the backbone to produce better object detection [27]. The head is the part that detects objects by predicting bounding box coordinates and class probabilities. The architecture diagram of the YOLOv9 model training process is shown in Figure 3.

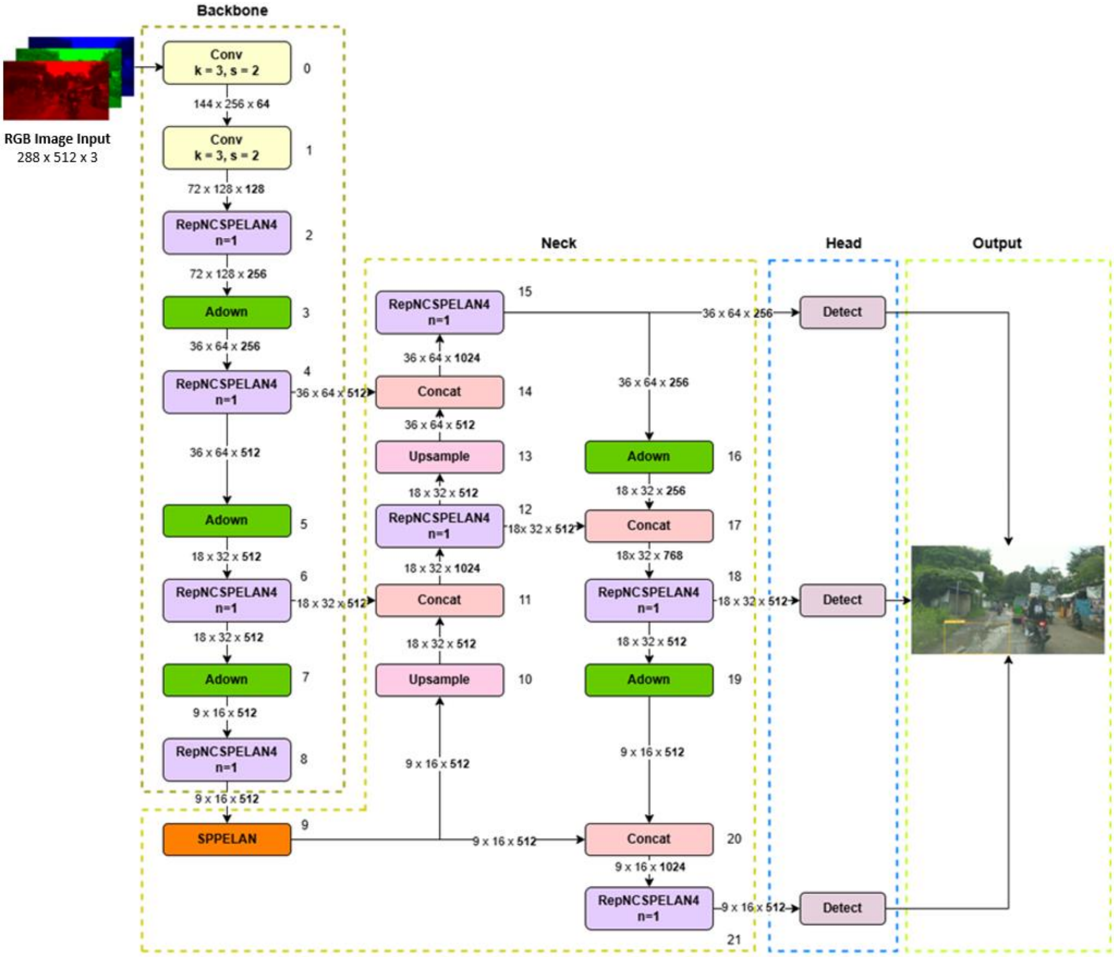


Figure 3. Architecture diagram of the YOLOv9 model for road damage detection

The YOLOv9 detection model training involves several key building blocks and operations. The building blocks include the convolution block, RepNCSPELAN4 block, Adown block, and spatial pyramid pooling enhanced with local attention network (SPPELAN) block. Operations performed on these blocks include concatenation, upsampling, and partitioning. The convolution block extracts key features from the input image using a 2D convolutional layer, batch normalization, and the SiLU activation function [27]. RepNCSPELAN4 is a block of the YOLOv9 network architecture that extracts and enhances the feature representation [28]. The Adown (asymmetric downsampling) block reduces the dimensionality of feature maps by downsampling the input while retaining important information. SPPELAN incorporates spatial pyramid pooling into the ELAN architecture, starting with a convolutional layer, followed by max pooling to capture information across scales, and advanced convolution for more detailed feature extraction [27]. During the training of YOLOv9, advanced data augmentation methods such as Mosaic (combining four images), MixUp (blending two images with label interpolation), flipping, random rotation, brightness and contrast adjustment, and random cropping were applied to enrich dataset variability and improve model generalization against diverse visual conditions [29].

2.6. Evaluation metrics

The evaluation metrics used in this study include precision, recall, and mAP [30]. Precision can be calculated using (1). Meanwhile, the recall value can be calculated using (2). The mAP value can be calculated using (3). True positive (TP) detections refer to instances where objects are correctly identified, while false positive (FP) detections indicate cases where objects are incorrectly identified as present. False negatives (FN), on the other hand, represent ground-truth objects that the model failed to detect.

$$P = \frac{TP}{TP+FP} \quad (1)$$

$$R = \frac{TP}{TP+FN} \quad (2)$$

$$mAP = \frac{1}{N}\sum_{i=1}^{N} AP_i \quad (3)$$

The average precision for class i, AP_i, is calculated using (4), where N represents the total number of classes. The value of $P_i(R_{n+1})$ is determined through (5). Additionally, precision is represented as $P(\tilde{R})$, is evaluated at a given recall level $\tilde{R}$.

$$AP = \sum_n (R_{n+1} - R_n) P_i(R_{n+1}) \quad (4)$$

$$P_i(R_{n+1}) = \max_{\tilde{R}:\tilde{R}\geq R_{n+1}} P(\tilde{R}) \quad (5)$$

3. RESULTS AND DISCUSSION

Testing is done with variations in learning rate and image size. The learning rate variations tested were 0.1, 0.01, 0.001, and 0.0001. The tested image sizes were 288 × 512 and 360 × 640 pixels. All tests were conducted on Google Colab using a 15.0 GB Tesla T4 with 12.7 GB of memory. The results of testing using this validation data are shown in Table 1. Based on this table, the highest mAP is achieved with a learning rate of 0.001 for an input size of 288 × 512 pixels. The model achieved 71.8% precision, 54.3% recall, and 59.7% mAP. The best model, with an input size of 360 × 640 pixels, uses a learning rate of 0.001. The model achieved a precision of 69.3%, a recall of 63.0%, and a mAP of 65.4%. The mAP at 360 × 640 pixels outperformed that at 288 × 512 pixels. Despite having inputs of different sizes, both achieved the best performance using a learning rate of 0.001. Therefore, the best model chosen is the one with a learning rate of 0.001 and an input image size of 360 × 640 pixels.

Table 1. Test results with varying learning rates and image sizes

Learning rate	Image size (pixel)	mAP (%)	Precision (%)	Recall (%)
0.1	288 × 512	33.0	52.8	30.8
0.01	288 × 512	56.1	63.3	54.1
0.001	288 × 512	59.7	71.8	54.3
0.0001	288 × 512	50.7	61.4	44.6
0.1	360 × 640	48.5	51.9	47.2
0.01	360 × 640	62.6	71.1	58.7
0.001	360 × 640	65.4	69.3	63.0
0.0001	360 × 640	56.2	63.1	53.4

The results of the model testing process, which achieved the best performance using the test dataset, are shown in Table 2. The model achieved an average precision of 76.2%, a recall of 49.2%, and a mAP of 62.8% across all classes. The model performed best at detecting the pothole class, achieving a precision of 84.9%, a recall of 63.5%, and a mean mAP of 73.6%. Then, the longitudinal cracking class had a precision value of 76.2%, a recall of 54.9%, and an mAP of 66.3%. Furthermore, in the alligator cracking class, the model achieves 74.6% precision, 48.5% recall, and 62.2% mAP. The lateral cracking class achieves the lowest performance, with a precision of 69.2%, recall of 29.8%, and mAP of 49.3%. Figure 4 presents the Precision-Recall curve for the proposed method for road damage detection. The test results show that the model detects well in the pothole, alligator, and longitudinal cracking classes. However, the model still needs improvement to detect the lateral cracking class. The lower detection performance of the lateral cracking class is mainly due to its thin, low-contrast, and linear structure, which visually resembles road textures and markings. Additionally, lateral cracks often appear near road edges and are more susceptible to perspective distortion and resolution reduction, leading the model to miss critical features during detection.

Figure 5 presents examples of road-damage detection results from the YOLOv9 model across four damage classes. In Figure 5(a), the model successfully detects both alligator cracking and longitudinal cracking on the road surface. Figure 5(b) shows a correctly identified lateral crack, with a high confidence score of 0.87. In Figure 5(c), a longitudinal crack is detected along the driving path. Figure 5(d) illustrates a detected pothole, accurately localized in a high-traffic area. These results demonstrate the model's ability to identify multiple types of road damage in real-world conditions, supporting its potential for practical road monitoring applications. Furthermore, the model's lightweight, efficient design makes it suitable for deployment on mobile and edge computing devices with real-time processing capabilities.

The proposed method (YOLOv9) was then compared with prior research and other YOLO variants. The results of the comparison are shown in Table 3. YOLOv5 has been used for road defect detection by Ruseruka *et al.* [14]. Additionally, we compared YOLOv7 and YOLOv8. The highest mAP is achieved by the YOLOv9 detection model, with 65.4%. The YOLOv9 detection model is the best among several previous YOLO algorithms. Therefore, YOLOv9 is a strong choice for building a road-damage detection model.

The dataset is limited by the use of a single smartphone camera and constrained environmental variations, which may introduce bias. Additionally, high image resolution, resizing effects, and class imbalance can reduce detection accuracy and model generalization, particularly for subtle damage such as lateral cracks. Future work may explore multi-modal inputs such as infrared imagery, ensemble models to enhance detection robustness, and domain adaptation techniques to improve generalization across different environments and data sources.

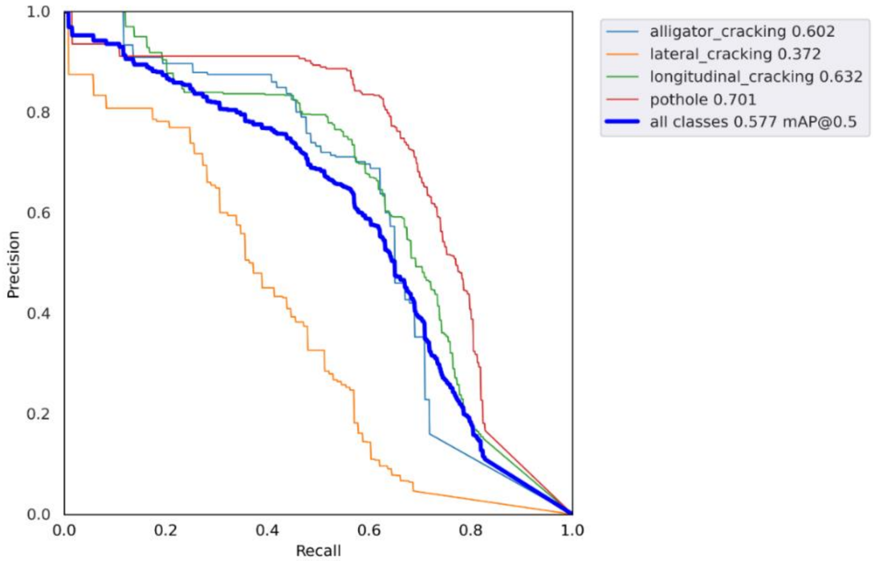


Figure 4. Precision-recall curve of the proposed method for road damage detection

Table 2. Best performance model testing results

Class	mAP (%)	Precision (%)	Recall (%)
Alligator cracking	62.2	74.6	48.5
Lateral cracking	49.3	69.2	29.8
Longitudinal cracking	66.3	76.2	54.9
Pothole	73.6	84.9	63.5
All	62.8	76.2	49.2

Table 3. Comparison of the proposed method and other YOLO models

Model	mAP (%)	Precision (%)	Recall (%)
YOLOv5 [14]	55.8	60.1	55.2
YOLOv7	62.9	65.5	61.9
YOLOv8	58.4	71.6	53.4
YOLOv9 (proposed method)	65.4	69.3	63.0

(a) (b) (c) (d)

Figure 5. Example of road damage detection results using the YOLOv9-based model for multiple damage classes: (a) alligator cracking class, (b) lateral cracking class, (c) longitudinal cracking class, (d) pathole cracking class

4. CONCLUSION

This study proposed a road damage detection method in four categories: alligator cracking, lateral cracking, longitudinal cracking, and pothole usage. The primary contribution of this study is the demonstrated capability of the YOLOv9 model to detect specific types of road damage with optimized hyperparameters accurately. The test results achieved an average precision value of 76.2%, a recall of 49.2%, and a mAP of 62.8% for all classes. The highest detection performance was achieved for the pothole class, with a precision of 84.9%, recall of 63.5%, and an mAP of 73.6%. Then, the longitudinal cracking class achieved 76.2% precision, 54.9% recall, and an mAP of 66.3%. The alligator cracking class achieved a precision of 74.6%, a recall of 48.5%, and an mAP of 62.2%. Then, the lateral cracking class gets a precision value of 69.2%, a recall of 29.8%, and an mAP of 49.3%. The YOLOv9 model with the best performance was trained with a learning rate of 0.001 and an input image size of 360 × 640 pixels. The model's validation results yielded an average precision of 69.3%, a recall of 63%, and an mAP of 65.4% across all classes. The model showed strong performance, particularly in detecting potholes and longitudinal cracks, which are among the most critical defects affecting road safety and comfort. These findings demonstrate the potential of YOLOv9 as an effective tool for supporting automated road condition assessments. In practical terms, this approach can help decision-makers and road maintenance agencies prioritize repairs, improve the efficiency of inspection processes, and enhance infrastructure management. However, the relatively low recall and mAP for lateral cracking indicate a limitation in detecting specific damage patterns, possibly due to variations in crack shape and size or dataset imbalance. Future research could focus on further optimizing the model, expanding the training dataset to include more diverse examples of lateral cracks, and validating the approach on larger, more varied road-condition datasets to improve generalizability and robustness.

FUNDING INFORMATION

Authors state no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Muhammad Farkhan Adhitama	✓	✓	✓		✓	✓	✓	✓	✓		✓		✓	
Sutikno	✓	✓		✓	✓				✓	✓		✓	✓	✓
Rismiyati		✓		✓	✓				✓			✓	✓	

C : **C**onceptualization	I : **I**nvestigation	Vi : **Vi**sualization
M : **M**ethodology	R : **R**esources	Su : **Su**pervision
So : **So**ftware	D : **D**ata Curation	P : **P**roject administration
Va : **Va**lidation	O : Writing - **O**riginal Draft	Fu : **Fu**nding acquisition
Fo : **Fo**rmal analysis	E : Writing - Review & **E**diting	

CONFLICT OF INTEREST STATEMENT

The authors state no conflict of interest.

DATA AVAILABILITY

This study used a publicly available road damage image dataset containing photographs of paved roads exhibiting various types of surface damage. The dataset is accessible through Kaggle at https://www.kaggle.com/datasets/alvarobasily/road-damage/data.

REFERENCES

[1] S. D. Prihapsasti *et al.*, *Informasi Statistik Infrastruktur PUPR 2022*, vol. 11, no. 1. 2022.

[2] W. Cao, Q. Liu, and Z. He, "Review of pavement defect detection methods," *IEEE Access*, vol. 8, pp. 14531–14544, 2020, doi: 10.1109/aCCESS.2020.2966881.

[3] Sutikno, H. A. Wibawa, and P. Y. Budiarto, "Classification of road damage from digital image using backpropagation neural network," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 6, no. 4, p. 159, 2017, doi: 10.11591/ijai.v6.i4.pp159-165.

[4] D. Arya, S. K. Ghosh, and D. Toshniwal, "Automatic recognition of road cracks using gray-level co-occurrence matrix and machine learning," in *Lecture Notes in Electrical Engineering*, vol. 858, 2022, pp. 443-452. doi: 10.1007/978-981-19-0840-8_33.

[5] M. H. Daneshvari, E. Nourmohammadi, M. Ameri, and B. Mojaradi, "Efficient LBP-GLCM texture analysis for asphalt pavement raveling detection using eXtreme gradient boost," *Construction and Building Materials*, vol. 401, p. 132731, Oct. 2023, doi: 10.1016/j.conbuildmat.2023.132731.

[6] R. Ahila Priyadharshini, S. Arivazhagan, and M. Arun, "Crack recognition on concrete structures based on machine crafted and hand crafted features," *Expert Systems with Applications*, vol. 228, p. 120447, Oct. 2023, doi: 10.1016/j.eswa.2023.120447.

[7] N.-D. Hoang and Q.-L. Nguyen, "Computer vision-based recognition of pavement crack patterns using light gradient boosting machine, deep neural network, and convolutional neural network," *Journal of Soft Computing in Civil Engineering*, vol. 7, no. 3, pp. 21–51, 2023, doi: 10.22115/SCCE.2023.367276.1547.

[8] A. K. Pandey, R. Iqbal, T. Maniak, C. Karyotis, S. Akuma, and V. Palade, "Convolution neural networks for pothole detection of critical road infrastructure," *Computers and Electrical Engineering*, vol. 99, pp. 107725, 2022, doi: 10.1016/j.compeleceng.2022.107725.

[9] Z. Zhang, W. Cui, Y. Tao, and T. Shi, "Road damage detection algorithm based on multi-scale feature extraction," *Engineering Letters*, vol. 32, no. 1, pp. 151–159, 2024.

[10] H. Zhou, S. Chen, G. Yan, Y. Huang, and Z. Lu, "An intelligent road damage detection method based on YOLOv7-tiny framework," in *International Conference on Communication Technology Proceedings, ICCT*, 2023, pp. 286-290, doi: 10.1109/ICCT59356.2023.10419388.

[11] M. A. Duhayyim *et al.*, "Road damage detection using the hunger games search with elman neural network on high-resolution remote sensing images," *Remote Sensing*, vol. 14, no. 24, 2022, doi: 10.3390/rs14246222.

[12] S. Shim, J. Kim, S.-W. Lee, and G.-C. Cho, "Road damage detection using super-resolution and semi-supervised learning with generative adversarial network," *Automation in Construction*, vol. 135, 2022, doi: 10.1016/j.autcon.2022.104139.

[13] F. Wan, C. Sun, H. He, G. Lei, L. Xu, and T. Xiao, "YOLO-LRDD: a lightweight method for road damage detection based on improved YOLOv5s," *Eurasip Journal on Advances in Signal Processing*, vol. 2022, no. 1, pp. 1-18, 2022, doi: 10.1186/s13634-022-00931-x.

[14] C. Ruseruka, J. Mwakalonge, G. Comert, S. Siuhi, F. Ngeni, and Q. Anderson, "Augmenting roadway safety with machine learning and deep learning: Pothole detection and dimension estimation using in-vehicle technologies," *Machine Learning with Applications*, vol. 16, no. March, p. 100547, 2024, doi: 10.1016/j.mlwa.2024.100547.

[15] J. Zhao, Y. Tao, Z. Zhang, C. Huang, and W. Cui, "Lightweight road damage detection network based on YOLOv5," *Engineering Letters*, vol. 32, no. 8, pp. 1708–1720, 2024.

[16] G. Ye, J. Qu, J. Tao, W. Dai, Y. Mao, and Q. Jin, "Autonomous surface crack identification of concrete structures based on the YOLOv7 algorithm," *Journal of Building Engineering*, vol. 73, no. March, p. 106688, 2023, doi: 10.1016/j.jobe.2023.106688.

[17] Y. Li, C. Yin, Y. Lei, J. Zhang, and Y. Yan, "RDD-YOLO: road damage detection algorithm based on improved you only look once version 8," *Applied Sciences (Switzerland)*, vol. 14, no. 8, 2024, doi: 10.3390/app14083360.

[18] X. Li and Y. Zhang, "Improved road damage detection algorithm based on YOLOv8n," *IAENG International Journal of Computer Science*, vol. 51, no. 11, pp. 1720–1730, 2024.

[19] H. Maeda, Y. Sekimoto, T. Seto, T. Kashiyama, and H. Omata, "Road damage detection and classification using deep neural networks with smartphone images," *Computer-Aided Civil and Infrastructure Engineering*, vol. 33, no. 12, pp. 1127–1141, 2018, doi: 10.1111/mice.12387.
[20] M. W. Khan, M. S. Obaidat, K. Mahmood, B. Sadoun, H. M. S. Badar, and W. Gao, "Real-time road damage detection using an optimized YOLOv9s-fusion in IoT infrastructure," *IEEE Internet of Things Journal*, vol. 12, no. 11, pp. 17649–17660, 2025, doi: 10.1109/JIOT.2025.3537640.
[21] C.-Y. Wang, I.-H. Yeh, and H.-Y. Mark Liao, "YOLOv9: learning what you want to learn using programmable gradient information," in *Computer Vision*, A. Leonardis, E. Ricci, S. Roth, O. Russakovsky, T. Sattler, and G. Varol, Eds. Cham: Springer Nature Switzerland, 2025, pp. 1–21.
[22] H. Zhang, X. Jiang, Y. Liu, J. Qian, and L. Ni, "Improvement of dam crack detection algorithm for YOLOv9," *IET Image Processing*, vol. 19, no. 1, 2025, doi: 10.1049/ipr2.70124.
[23] S. Dolhopolov, T. Honcharenko, D. Chernyshev, O. Panina, and A. Makhynia, "Advancing automated quality control in automotive manufacturing: A comparative analysis of YOLOv8, YOLOv9, and YOLOv10 for vehicle damage detection," in *CEUR Workshop Proceedings*, 2025, vol. 3966.
[24] S. Dolhpolov, T. Honcharenko, V. Hots, P. Kruk, and I. Porokhovnichenko, "YOLOv8, YOLOv9, and YOLOv10: a study in automated vehicle damage detection," in *CEUR Workshop Proceedings*, 2024, vol. 3896, pp. 25 – 44.
[25] A. Basily, "Road damage," *kaggle.com*, 2020.
[26] J. N. Prome, F. Sultana, and S. Anika, "Skin cancer detection: leveraging hybrid deep learning models and traditional machine learning classifiers," 2024, doi: 10.1145/3723178.3723280.
[27] A. Imran, M. S. Hulikal, and H. A. A. Gardi, "Real time american sign language detection using YOLO-v9," *arXiv preprint arXiv:2407.17950*, 2024.
[28] A. Febriani, "Improved hybrid machine and deep learning model for optimization of smart egg incubator," *Journal of Applied Data Sciences*, vol. 5, no. 3, pp. 1052–1068, 2024, doi: 10.47738/jads.v5i3.304.
[29] Ultralytics, "Data augmentation using ultralytics YOLO," *2025*.
[30] R. Padilla, S. L. Netto, and E. A. B. Da Silva, "A survey on performance metrics for object-detection algorithms," in *International Conference on Systems, Signals, and Image Processing*, 2020, pp. 237–242, doi: 10.1109/IWSSIP48289.2020.9145130.

BIOGRAPHIES OF AUTHORS

Muhammad Farkhan Adhitama is a graduate of informatics from Universitas Diponegoro, Indonesia, in 2024, with a GPA of 3.86. His main interests include machine learning, data science, natural language processing (NLP), and computer vision. He has completed several projects involving deep learning, large language models (LLM), and recommendation systems. He can be contacted at email: muhammadfarkhanadhit@alumni.undip.ac.id.

Dr. Sutikno received a Doctor of Philosophy (Ph.D.) degree in computer science from the faculty of mathematics and natural sciences at the University of Gadjah Mada, Indonesia. Now, he is an Assistant Professor at the Department of Informatics, Faculty of Sciences and Mathematics, University of Diponegoro. His research interests include machine learning, computer vision, and artificial intelligence. He can be contacted at email: sutikno@lecturer.undip.ac.id.

Rismiyati received a Bachelor of Engineering (B.Eng) from the School of Electrical and Electronics Engineering, Nanyang Technological University(NTU), in 2007 and a Master of Computer Science (MCs) from the Universitas Gadjah Mada, Indonesia, in 2016. She has been a lecturer in the Department of Informatics at Universitas Diponegoro since 2017. Her main research interests are artificial intelligence, image processing, and machine learning. She can be contacted at email: rismiyati@live.undip.ac.id.