**Journal of Applied Informatics and Computing (JAIC)**
Vol.10, No.4, August 2026, pp. 3637~3646
e-ISSN: 2548-6861 

# Performance Analysis of YOLO26 in Pothole Detection on an Indonesian Road Dataset

**Mohammad Alwi Nanda Saputra [1], Farikh Al Zami [2*], Christy Atika Sari [3]**
Teknik Informatika, Universitas Dian Nuswantoro
111202315442@mhs.dinus.ac.id [1], alzami@dsn.dinus.ac.id [2], christy.atika.sari@dsn.dinus.ac.id [3]

**Article Info**

***Article history:***

Received 2026-07-06
Revised 2026-08-01
Accepted 2026-08-10

***Keyword:***

*YOLO26l,*
*Object Detection,*
*Deep Learning,*
*Road Damage Detection,*
*Road Damage Indonesia Dataset.*

**ABSTRACT**

Road damage is one of the infrastructure problems that can compromise safety, comfort, and the smooth flow of traffic. The road inspection process, which is still carried out manually, requires a relatively large amount of time, labor, and cost, making a more efficient method necessary. Advances in computer vision and deep learning technologies enable the automatic detection of road damage through an object detection approach. This study aims to analyze the performance of the YOLO26l model as a baseline model in detecting four categories of road damage such as potholes, alligator cracking, lateral cracking, and longitudinal cracking using the Road Damage Indonesia Dataset. The dataset was divided into 70% training data, 15% validation data, and 15% testing data. The training process was conducted using the pre-trained weights from yolo26l.pt via the Ultralytics framework without any architectural modifications or the application of image enhancement methods. Performance evaluation was conducted using the Precision, Recall, mAP@0.50 (mAP@0.50), and mAP@0.50:0.95 (mAP@0.50:0.95) metrics. The results of the study show that the YOLO26l model achieved a Precision of 0.7028, a Recall of 0.6492, a mAP@0.50 of 0.6890, and a mAP@0.50:0.95 of 0.3391. Analysis using a confusion matrix, precision–recall curve, and visualization of the detection results showed that the model was able to identify all four categories of road damage well, although there were still some objects that went undetected under poor lighting conditions, due to small object sizes, or complex road surface textures. Based on these results, it can be concluded that YOLO26l performs well as a baseline model for road damage detection on the Indonesian road dataset



## I. INTRODUCTION

Roads are a vital component of transportation infrastructure, playing a crucial role in facilitating public mobility, the distribution of goods, and economic growth. Adequate road infrastructure supports safety, comfort, and the smooth flow of transportation, thereby contributing to enhanced connectivity and the efficient movement of people and goods [1]. However, road damage particularly potholes remains a major challenge in the management and maintenance of road infrastructure in Indonesia, as it degrades service quality and necessitates periodic repair and upkeep [2], [3]. Such damage increases the risk of traffic accidents, accelerates vehicle component wear, compromises driving comfort, and drives up both infrastructure maintenance and vehicle repair costs [2]. Consequently, the rapid, accurate, and continuous assessment of road conditions is essential for effective road infrastructure monitoring and maintenance.

To date, the identification of potholes is largely carried out through manual field inspections by personnel. This method is time-consuming, labor-intensive, and entails high operational costs, particularly for road networks covering extensive areas [4]. Furthermore, the results of manual inspections rely heavily on the personnel's skills and experience in identifying the type and severity of road damage, potentially leading to subjective and inconsistent assessments [5]. With the advancement of Artificial Intelligence (AI) technology, the application of Computer Vision and Deep Learning has emerged as a widely used

approach to support automated object detection through digital image analysis. This approach enables systems to recognize and identify objects with greater accuracy and efficiency than conventional methods [6]. One widely implemented approach is Object Detectiona Computer Vision technique capable of identifying the presence of objects and determining their locations within an image using bounding boxes to represent their positions [7]. The use of Object Detection techniques is considered capable of enhancing the efficiency of road inspection processes by enabling faster, more consistent, and objective damage identification, thereby supporting more effective decision-making in infrastructure maintenance activities [4].

Developments in deep learning have driven significant progress across various computer vision tasks, notably object detection. Unlike image classification methods, which merely recognize the presence of an object, object detection identifies the object type and determines its location within an image using bounding boxes; this makes it better suited for applications requiring accurate positional information [7]. A widely used family of object detection models is You Only Look Once (YOLO), known for its fast inference speeds and competitive accuracy, leading to its adoption in diverse fields such as traffic surveillance, autonomous vehicles, security systems, and road damage detection [8]. As the architecture evolves, various YOLO variants continue to be developed to enhance detection capabilities while maintaining computational efficiency, establishing it as a prominent approach in current object detection research [9], [10].

Various studies have utilized object detection-based deep learning methods to automatically detect road damage. Previous research indicates that implementing deep learning models enhances the effectiveness of pothole identification compared to conventional inspection methods, while also delivering faster and more consistent detection results [11]. With the evolution of object detection architectures, various YOLO variants have been developed to improve detection accuracy while maintaining computational efficiency for real-time applications. Recent research demonstrates that YOLO variants such as YOLOv9, YOLOv10, and YOLOv11 offer improved detection performance and faster inference speeds, making them suitable for implementation in automated road inspection systems [12], [13], [14]. Furthermore, studies show that the YOLO family is widely applied to the detection of potholes and other types of road damage due to its balance between accuracy and processing speed, establishing it as a popular approach in the development of computer vision-based systems for road infrastructure maintenance [15].

Based on various prior studies, deep learning-based object detection methods particularly the YOLO family have demonstrated strong performance in detecting potholes and various types of road damage. Most research has focused on implementing and comparing different YOLO variants (such as YOLOv5, YOLOv8, YOLOv9, YOLOv10, and YOLOv11) to enhance detection accuracy and efficiency [15], [8], [13], [14]. However, research evaluating the performance of YOLO26 as a baseline model for pothole detection using Indonesian road datasets remains limited. Furthermore, while most studies emphasize model architecture development or performance optimization, there is a lack of reported analysis regarding the baseline performance of YOLO26 on Indonesian road datasets [16].

Although previous studies have demonstrated promising performance using object detection models such as YOLOv8, YOLOv9, YOLOv10, YOLOv11, and RT-DETR, these studies primarily focused on evaluating well-established architectures and improving detection accuracy or computational efficiency. However, the recently introduced YOLO26 architecture has not yet been systematically evaluated for multi-class road damage detection, particularly on the Road Damage Indonesia Dataset. Consequently, there is still limited empirical evidence regarding its detection capability and comparative performance against established object detection models under Indonesian road conditions.

Based on the identified research gap, this study aims to evaluate the detection performance of the recently introduced YOLO26l model and compare it with YOLOv8l using identical experimental settings on the Road Damage Indonesia Dataset. Both models are trained using the same dataset partition, image resolution, training parameters, and evaluation protocol to ensure a fair comparison. Their performance is assessed using Precision, Recall, mean Average Precision (mAP@0.50), and mean Average Precision (mAP@0.50:0.95) commonly employed in object detection model evaluation [17]. The findings are expected to provide insight into the baseline performance of YOLO26 for pothole detection and serve as a benchmark for future research aimed at evaluating or developing pothole detection models using Indonesian road datasets. In addition, this study contributes by providing a fair comparative evaluation between YOLO26l and YOLOv8l under identical experimental settings, enabling an objective assessment of the detection capability of the newly introduced architecture. The experimental results are expected to serve as a reference for future studies involving newly developed object detection models for road damage detection under Indonesian road conditions.

## II. METHOD

### A. Research Stages

This research was conducted through several systematically organized stages to analyze the performance of the YOLO26 model in detecting potholes within an Indonesian road dataset. The research process encompassed dataset collection, data preparation, model training, and performance evaluation using metrics such as Precision, Recall, mean Average Precision (mAP@0.50), and mean Average Precision (mAP@0.50:0.95). The research workflow was designed to ensure a structured approach at each stage, thereby allowing the evaluation results to objectively reflect the model's capabilities.

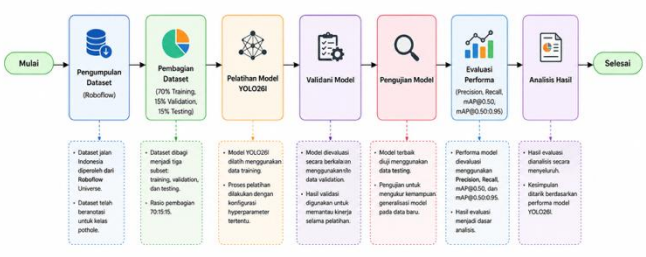


*Figure 1 Research Stages Flowchart*

## B. *Datasets*

This study utilizes the "Road Damage Indonesia" dataset obtained from Roboflow Universe https://universe.roboflow.com/detection-project-jvglg/road-damage-indonesia-srace. This object detection dataset comprises 3,321 images of road conditions featuring four damage categories potholes, alligator cracking, lateral cracking, and longitudinal cracking which serve as the detection targets. The study aims to analyze the performance of the YOLO26 model in detecting potholes under Indonesian road conditions. All images in the dataset are annotated with bounding boxes in YOLO format, enabling their immediate use in the object detection model training process.

Prior to the training process, the dataset is divided into three subsets: 70% for training, 15% for validation, and 15% for testing. This partitioning allows the model to learn data characteristics during training, undergo validation throughout the training process, and evaluate its generalization capabilities using previously unseen data. Dividing data into training, validation, and testing sets is a common practice in Deep Learning model development to ensure a more representative evaluation of the model. The specifications of the dataset used in this study are presented in Table 1.

TABLE 1
DATASET SPECIFICATIONS

| Parameter | Information |
|---|---|
| Dataset Name | Road Demage Indonesia |
| Dataset Source | Roboflow Universe |
| Task | Object Detection |
| Number of Images | 3,321 |
| Original Class | Pothole(2.657), Alligator Cracking(2.516), Lateral Cracking(1.222), Longitudinal Cracking(964) |
| Annotation Format | YOLO |
| Dataset Split Ratio | 70% Training, 15% Validation, 15% Testing |

Based on Table 1, the dataset used in this study consists of 3,321 images of road conditions featuring four categories of damage: potholes, alligator cracking, lateral cracking, and longitudinal cracking. However, this study focuses exclusively on the pothole class as the detection target, aligning with the research objective of analyzing the YOLO26 model's performance in detecting potholes. The dataset was partitioned into training, validation, and testing sets using a 70%:15%:15% ratio, enabling a systematic approach to model training and evaluation.

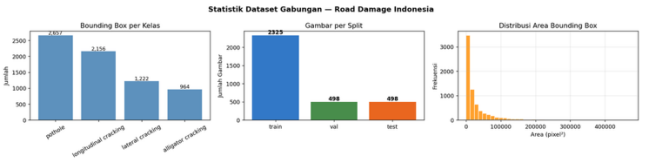


*Figure 2 presents the statistical distribution of the Road Damage.*

Figure 2 presents the statistical distribution of the annotated objects contained in the Road Damage Indonesia Dataset. The dataset exhibits a moderate class imbalance, with pothole representing the largest class (2,657 annotated instances), followed by longitudinal cracking (2,156 instances), lateral cracking (1,222 instances), and alligator cracking (964 instances). Although the distribution is not perfectly balanced, the differences among classes are not excessively large. Nevertheless, the relatively smaller number of annotated samples in the lateral cracking and alligator cracking classes may influence the model's learning process by providing fewer representative visual features compared with the majority classes. Therefore, the class distribution should be considered when interpreting the detection performance of each road damage category.

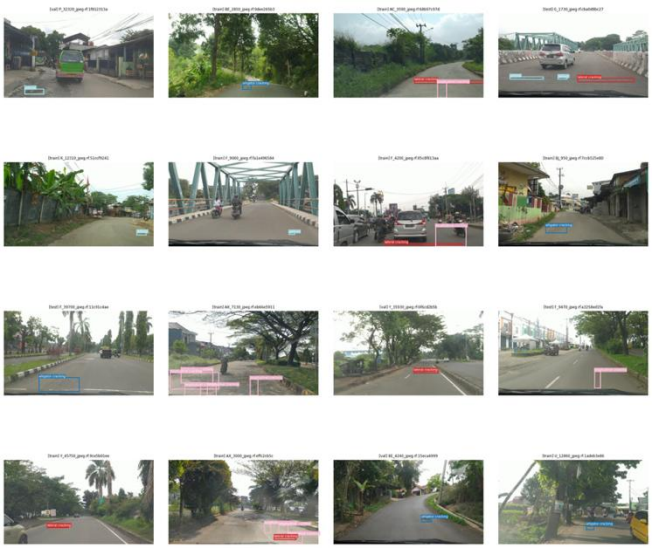

*Figure 3 Bounding Box Sample Dataset Result*

Figure 3 displays examples of images from the Road Damage Indonesia dataset used in this study. Each image includes bounding box annotations indicating the location of road damage objects, corresponding to the classes defined in the dataset. Although the dataset comprises four damage categories potholes, alligator cracking, lateral cracking, and longitudinal cracking This study utilizes annotations from all four road damage categories pothole, alligator cracking, lateral cracking, and longitudinal cracking as object detection targets. These annotations serve as ground truth during the

training and evaluation process, enabling the YOLO26l model to learn the visual characteristics and spatial locations of each road damage category.

### C. YOLO26

The YOLO26 (You Only Look Once 26) model is a state-of-the-art object detection model developed to perform real-time object detection while maintaining a balance between inference speed and detection accuracy. As a one-stage detector, YOLO26 performs object classification and object location prediction (bounding box) simultaneously in a single inference pass, making it more efficient than two-stage detector approaches. Furthermore, YOLO26 incorporates various architectural refinements to enhance feature extraction capabilities and prediction quality across diverse object detection scenarios [16].

In this study, the YOLO26 model was employed as a baseline model without architectural modifications or the application of image enhancement methods. This model was selected due to its capability for rapid and accurate object detection, making it suitable for evaluating baseline performance in the task of pothole detection using an Indonesian road dataset [16]. The entire training process was conducted using the official YOLO26 implementation via the Ultralytics framework [16], while model performance evaluation was based on the metrics of Precision, Recall, mean Average Precision (mAP@0.50), and mean Average Precision (mAP@0.50:0.95) [17].

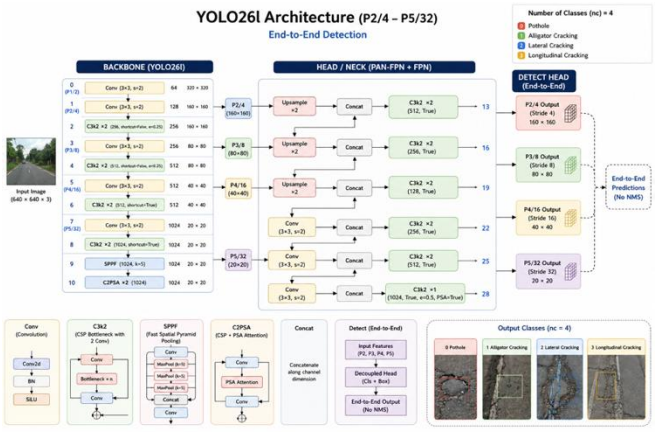


*Figure 4 YOLO26L Architecture*

Based on Figure 3, the YOLO26l architecture consists of three main components: the Backbone, the Neck, and the Detection Head. In the Backbone, the input image is processed through several convolutional (Conv) layers, C3k2 modules, SPPF (Spatial Pyramid Pooling Fast), and C2PSA (Cross Stage Partial with Position-Sensitive Attention) to extract features at various levels of representation [8]. The resulting features are then passed to the Neck via a feature fusion mechanism employing Upsampling and Concat operations, allowing information from different scales to be combined to enhance the capability to detect objects of varying sizes.

Next, in the Detection Head, the model generates end-to-end predictions across four feature scale levels: P2/4, P3/8, P4/16, and P5/32. Each scale is utilized to detect objects of varying sizes, enabling the model to more effectively detect both small and large objects. In this study, the Detection Head generates predictions for four categories of road damage: potholes, alligator cracking, lateral cracking, and longitudinal cracking. However, during the experimental phase, only the pothole category was used as the detection target to evaluate the baseline performance of the YOLO26l model [18].

### D. Experimental Configuration

The training process for the YOLO26l model in this study was conducted using the Ultralytics framework on the Kaggle Notebook platform, utilizing an NVIDIA Tesla T4 GPU for computation. Training was performed using a dataset split into training, validation, and testing sets at a ratio of 70%:15%:15%. A consistent configuration was applied throughout the training process to ensure a reliable baseline performance evaluation of the model. The training configuration used in this study is presented in Table 2.

TABLE 2
MODEL TRAINING CONFIGURATION

| Parameter | Information |
|---|---|
| Dataset Name | Road Demage Indonesia |
| Dataset Source | Roboflow Universe |
| Task | Object Detection |
| Number of Images | 3,321 |
| Original Class | Pothole, Alligator Cracking, Lateral Cracking, Longitudinal Cracking |
| Annotation Format | YOLO |
| Dataset Split Ratio | 70% Training, 15% Validation, 15% Testing |

As shown in Table 2, the training process utilized the YOLO26l model with the pre-trained weights (yolo26l.pt) provided by the Ultralytics framework. All input images were resized to 640 × 640 pixels, and training was conducted over 100 epochs with a batch size of 16. To prevent overfitting, an early stopping mechanism was implemented with a patience value of 20 epochs. Additionally, a seed of 42 was used to ensure the reproducibility of the training process and consistent evaluation results.

### E. Model Training

Once the dataset preparation and training configuration stages were complete, the training process was carried out using the YOLO26l model with the pre-trained weights (yolo26l.pt) provided by the Ultralytics framework. [16], [19]. The model is trained using training data and periodically evaluated using validation data to monitor performance progress during the training process. [19]. The entire training process was conducted without architectural modifications or

the application of image enhancement methods, allowing the model to serve as a baseline model for this study.

During the training process, the model learns the characteristics of pothole objects based on the bounding box annotations present in the dataset [18], [19]. At each epoch, model parameters are updated through an optimization process to minimize the loss value, while data validation is used to evaluate the model's ability to generalize to data not directly used in the training process [20]. In addition, an early stopping mechanism is implemented using the patience parameter to stop the training process if there is no performance improvement within a certain number of epochs, thereby reducing the risk of overfitting [19].


*Figure 5 YOLO26l model training process workflow*

### *F. YOLO26L Performance Evaluation*

Performance evaluation was conducted to assess the capability of the YOLO26l model in detecting potholes within the testing data. Testing utilized four metrics commonly employed in object detection research: Precision, Recall, mean Average Precision (mAP@0.50), and mean Average Precision (mAP@0.50:0.95). These metrics were used to evaluate prediction accuracy, the model's ability to detect all target objects, and the quality of the resulting bounding box localization. [17].

*1. Precision*

Precision is used to measure the model's accuracy in identifying objects predicted as potholes. The precision value is derived from the ratio of correct predictions (True Positives) to the total number of positive predictions generated by the model. The higher the precision value, the lower the likelihood of the model producing incorrect positive predictions (False Positives) [17].

$$\text{Precision} \frac{True\ Positive}{True\ Positive + False\ Positive} \quad (1)$$

*2. Recall*

Recall is used to measure the model's ability to detect all pothole objects present in the test data. The Recall value is calculated based on the ratio of correct predictions (True Positives) to the total number of actual objects (Ground Truth). The higher the Recall value, the fewer the objects that fail to be detected (False Negatives) [17].

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative} \quad (2)$$

*3. mean Average Precision (mAP@0.50)*

mAP@0.50 represents the mean Average Precision (AP) value at an Intersection over Union (IoU) threshold of 0.50. This metric is used to evaluate a model's ability to correctly classify and localize objects. A higher mAP@0.50 value indicates better model performance in object detection [17].

$$mAP@0.50 = \frac{1}{N}\sum_{i=1}^{N} APi \quad (3)$$

*4. mean Average Precision (mAP@0.50:0.95)*

mAP@0.50:0.95 is an evaluation metric based on the COCO evaluation standard that calculates the mean Average Precision across an IoU range of 0.50 to 0.95, with an interval of 0.05. Compared to mAP@0.50, this metric provides a stricter evaluation of bounding box localization accuracy, thereby offering a more comprehensive representation of the model's detection quality [17].

## III. RESULTS AND DISCUSSION

This section presents the results of training and evaluating the YOLO26l model for detecting potholes, alligator cracking, lateral cracking, and longitudinal cracking using the Road Damage Indonesia Dataset. The entire experimental process was conducted in accordance with the research methodology outlined in Chapter II. The results were analyzed using metrics such as Precision, Recall, mean Average Precision (mAP@0.50), and mean Average Precision (mAP@0.50:0.95) to evaluate the model's object detection accuracy. In addition to presenting evaluation results, this chapter discusses the model's performance based on training process visualizations and object detection outcomes, as well as their relationship to previous studies. [20].

### *A. Model Training Results*

Training of the YOLO26l model was conducted using the Road Damage Indonesia Dataset, which was split into training, validation, and testing sets with a ratio of 70%:15%:15%. The training process utilized the Ultralytics framework, employing the pre-trained weights `yolo26l.pt` and the training configuration detailed in Chapter II. During training, the model learned the characteristics of potholes based on bounding box annotations in the training data, while the validation data was used to monitor the model's performance progress across epochs.

Throughout the training process, the model iteratively updated its parameters via optimization to minimize the loss

value. Additionally, an early stopping mechanism was implemented to halt training if no performance improvement occurred over a specified number of epochs. This approach aimed to achieve a model with strong generalization capabilities while mitigating the risk of overfitting. Figure 5 illustrates the progress of the YOLO26l model's training process.

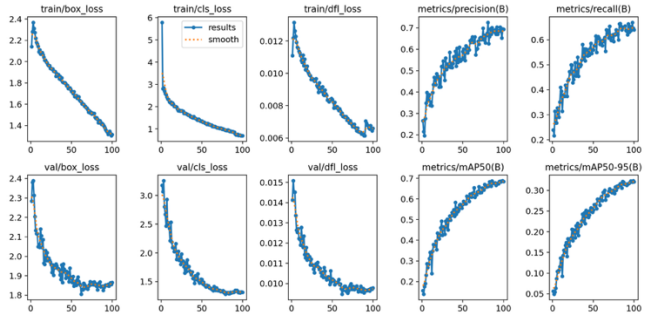


*Figure 6 YOLO26l model training result curve*

Based on Figure 5, the training process demonstrates that the training loss values comprising train/box_loss, train/cls_loss, and train/dfl_loss consistently decreased as the number of epochs increased. This decline indicates that the model became increasingly adept at learning the characteristics of potholes from the training data. Regarding validation, the validation loss values (val/box_loss, val/cls_loss, and val/dfl_loss) also exhibited a downward trend, eventually reaching a relatively stable state by the end of the training process. This suggests that the model possesses good generalization capabilities regarding data not utilized during training.

Furthermore, evaluation metrics specifically Precision, Recall, mAP@0.50, and mAP@0.50:0.95 showed gradual improvement across epochs. Precision rose to nearly 0.70, while Recall reached approximately 0.66 by the end of training. Meanwhile, mAP@0.50 increased to nearly 0.69, and mAP@0.50:0.95 reached approximately 0.32. These upward trends indicate that the model progressively improved its ability to identify potholes while simultaneously enhancing the quality of bounding box localization throughout the training process.

Overall, the training curves indicate that the YOLO26l model achieved convergence by the conclusion of the training process. This is evidenced by the stabilization of loss values and the leveling off of evaluation metrics in the final epochs, suggesting that the resulting model is suitable for evaluation using testing data.

### *B. Model Evaluation Results*

After the training process was completed, the YOLO26l model with the best weights (best.pt) was used to perform testing on the test dataset. The evaluation phase aimed to measure the model's ability to detect road damage objects using data that had not been utilized during the training or validation processes. Evaluation results were obtained using the metrics of Precision, Recall, mean Average Precision (mAP@0.50), and mean Average Precision.

TABLE 3
MATRIX RESULTS

| Matrix | Mark |
|---|---|
| Precision | 0.7028 |
| Recall | 0.6492 |
| mAP@0.50 | 0.6890 |
| mAP@0.50:0.95 | 0.3391 |

Based on Table 3, the YOLO26l model achieved a Precision score of 0.7028, indicating that the majority of objects predicted as potholes were correctly identified. This value suggests a relatively low False Positive rate, enabling the model to generate reasonably accurate predictions on the test data.

A Recall score of 0.6492 indicates that the model successfully detected approximately 64.92% of the potholes present in the test data. However, there remained a number of undetected objects (False Negatives). This outcome may be influenced by factors such as variations in pothole size, lighting conditions, road surface texture, and background complexity within the images.

Regarding the mAP@0.50 metric, the model achieved a score of 0.6890, demonstrating a strong capability for both classification and object localization via bounding boxes at an Intersection over Union (IoU) threshold of 0.50. Meanwhile, the mAP@0.50:0.95 score of 0.3391 indicates a decline in performance when evaluated across a stricter range of IoU values. This suggests that while the model effectively detects the presence of potholes, the precision of the bounding box positioning could be further improved to better align with the actual object locations.

Overall, the evaluation results demonstrate that the YOLO26l model delivers solid detection performance on the Indonesian road dataset. The obtained Precision, Recall, mAP@0.50, and mAP@0.50:0.95 scores indicate that the model possesses adequate capability for pothole detection, although there is still room to enhance bounding box localization accuracy. To gain a more detailed understanding of the prediction result distribution, a subsequent analysis was conducted using the Confusion Matrix presented in Figure 6.

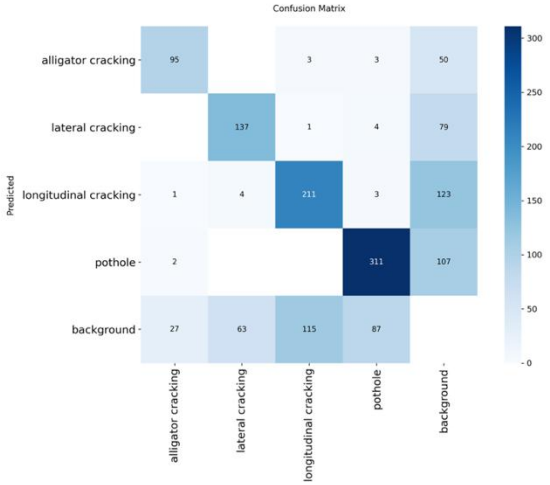


*Figure 7 Confusion Matrix YOLO26l models*

Based on Figure 6, the majority of the model's predictions lie along the main diagonal of the confusion matrix, indicating that the model is capable of classifying most objects according to their actual classes. The pothole class achieved the highest number of correct predictions (311 objects), followed by longitudinal cracking (211 objects), lateral cracking (137 objects), and alligator cracking (95 objects). These results demonstrate that the model possesses a strong ability to distinguish between the four types of road damage examined in this study.

Nevertheless, some prediction errors are evident in the values outside the main diagonal. Certain objects failed to be detected and were instead predicted as background: 27 objects for the alligator cracking class, 63 for lateral cracking, 115 for longitudinal cracking, and 87 for potholes. This indicates that some objects remain difficult for the model to recognize, particularly in images featuring small objects, complex road surface textures, or suboptimal lighting conditions.

A more in-depth analysis reveals that several environmental factors and object characteristics contribute to these detection failures. First, poor lighting reduces the visibility of road damage features, making it difficult for the model to distinguish damaged areas from the surrounding road surface. Second, small-scale road defects, particularly hairline cracks, have fewer distinctive visual features, resulting in lower localization accuracy and more false negatives. Furthermore, complex road textures and irregular pavement patterns sometimes exhibit visual characteristics resembling actual road damage, causing the model to generate lower confidence scores or even fail to detect certain objects entirely. These observations explain the false-negative predictions in the confusion matrix and align with the recall values obtained during model evaluation.

Apart from errors involving the background, misclassifications between the damage classes themselves were relatively infrequent. For instance, only 4 lateral cracking objects were predicted as longitudinal cracking, 3 pothole objects as longitudinal cracking, and 3 alligator cracking objects as longitudinal cracking. This suggests that the model effectively distinguishes the visual characteristics of each class; prediction errors stem more from a failure to detect the objects than from confusing them with other types of damage.

In addition to environmental factors, the distribution of classes within the training dataset may also influence detection performance for each road damage category. Dataset statistics reveal that the pothole and longitudinal cracking classes have a higher number of annotated samples compared to the lateral cracking and alligator cracking classes. A larger pool of training samples enables the model to learn visual patterns that are more representative of majority classes, whereas the relatively limited number of samples for minority classes can hinder the model's ability to generalize their characteristics. This situation may partly explain why the pothole class achieved the highest number of correct detections, while the alligator cracking and lateral cracking classes yielded relatively fewer correct predictions.

Overall, the confusion matrix results indicate that the YOLO26l model performs well in classifying the four categories of road damage. However, further improvements in detection capabilities are needed for objects with complex visual characteristics to reduce the number of undetected objects (false negatives). While the confusion matrix provides an overview of the distribution of correct predictions and misclassifications for each class, it does not illustrate the relationship between Precision and Recall across various confidence threshold levels. Therefore, a Precision–Recall (PR) curve analysis was conducted to evaluate the consistency of the model's performance in detecting each road damage class. The Precision–Recall curve for the YOLO26l model is shown in Figure 7.

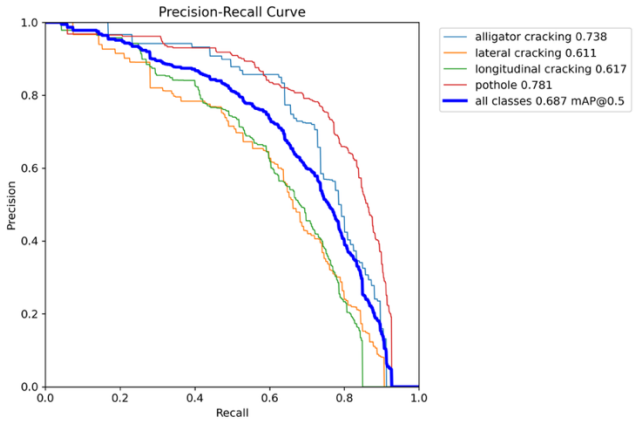


*Figure 8 Precision-Recall Curve YOLO26l models*

Based on Figure 7, the Precision–Recall curve illustrates the relationship between Precision and Recall values for each road damage class. A larger Area Under the Curve (AUC) indicates a superior ability of the model to maintain a balance between Precision and Recall. In this study, the pothole class demonstrated the best performance with an Average Precision

(AP) value of 0.781, followed by alligator cracking (0.738), longitudinal cracking (0.617), and lateral cracking (0.611).

Overall, the model achieved an mAP@0.50 score of 0.687, demonstrating a strong capability to detect all four road damage categories at an Intersection over Union (IoU) threshold of 0.50. This result aligns with the evaluation data presented in Table 3, confirming the model's ability to strike a favourable balance between prediction accuracy (Precision) and the capacity to detect all target objects (Recall).

The curves reveal that the model performs better at detecting the pothole class compared to the other classes. Conversely, the lateral cracking and longitudinal cracking classes exhibited lower Average Precision values; this indicates that these classes possess more varied visual characteristics, making them more difficult for the model to identify consistently.

### *C. Comparative Evaluation Between YOLO26l and YOLOv8l*

#### *1. Experimental Setup*

To ensure a fair comparison, the YOLO26l model was compared with YOLOv8l under identical experimental conditions. Both models were trained using the Road Damage Indonesia Dataset, with the data split into 70% for training, 15% for validation, and 15% for testing. Furthermore, both models utilized identical input image sizes, epoch counts, training configurations, and evaluation metrics—specifically Precision, Recall, mean Average Precision (mAP@0.50), and mean Average Precision (mAP@0.50:0.95). Consequently, any observed performance differences are expected to stem from the architectural characteristics of the respective models rather than variations in experimental configuration.

#### *2. Performance Comparison of YOLO26l and YOLOv8l*

TABLE 4
PERFORMANCE COMPARISON OF YOLO26L AND YOLOV8L

| Model | YOLO26l | YOLOv8l |
|---|---|---|
| Preprocessing | Original | Original |
| Precision | 0.7028 | 0.6720 |
| Recall | 0.6492 | 0.6125 |
| mAP@0.50 | 0.6890 | 0.6513 |
| mAP@0.50:0.95 | 0.3391 | 0.3062 |
| Training Schedule | 233.39m | 248.02m |

Based on Table 4, the YOLO26l model demonstrates superior performance compared to YOLOv8l across all evaluation metrics used. YOLO26l achieved a Precision score of 0.7028, surpassing YOLOv8l's score of 0.6720. These results indicate that YOLO26l is capable of generating more accurate positive predictions with fewer false positives.

Regarding the Recall metric, YOLO26l also outperformed YOLOv8l, achieving a score of 0.6492 compared to YOLOv8l's 0.6125. This indicates that YOLO26l successfully detected a greater number of road damage objects in the test data, resulting in fewer undetected objects (false negatives).

Performance differences were also evident in the mAP@0.50 scores; YOLO26l achieved 0.6890, whereas YOLOv8l scored 0.6413. Furthermore, under the stricter mAP@0.50:0.95 evaluation, YOLO26l maintained a higher score of 0.3391 compared to YOLOv8l's 0.3062. These results demonstrate that YOLO26l possesses superior capabilities in both object classification and localization across various Intersection over Union (IoU) thresholds.

In terms of training efficiency, YOLO26l required slightly less training time—233.39 minutes—compared to the 248.02 minutes required by YOLOv8l. This difference indicates that YOLO26l not only delivers superior detection performance but also offers higher training efficiency under the experimental configuration used in this study.

#### *3. Discussion of Comparison Results*

Comparison results indicate that the YOLO26l architecture outperforms YOLOv8l across all evaluation metrics: Precision, Recall, mAP@0.50, and mAP@0.50:0.95. The improved Precision demonstrates YOLO26l's superior ability to minimize false positive predictions, while the higher Recall indicates a greater capacity to detect road damage objects present in the images.

YOLO26l's superior performance in mAP@0.50 and mAP@0.50:0.95 suggests that the model not only recognizes object classes effectively but also achieves more accurate bounding box localization compared to YOLOv8l. These results demonstrate that the YOLO26l architecture possesses superior feature representation capabilities for the road damage dataset used in this study.

Although the performance gap between the two models is not substantial, experimental results show that YOLO26l consistently delivers better performance across all evaluation metrics. Consequently, YOLO26l can be considered more suitable than YOLOv8l for road damage detection tasks using the Road Damage Indonesia Dataset under the same experimental configuration.

#### *4. Comparison Conclusion*

Based on the comparative evaluation results, the YOLO26l model outperformed YOLOv8l across all evaluation metrics—namely Precision, Recall, mAP@0.50, and mAP@0.50:0.95. In addition to delivering higher detection accuracy, YOLO26l also demonstrated shorter training times compared to YOLOv8l. These results confirm that YOLO26l is a more effective model for detecting the four categories of road damage in the Road Damage Indonesia Dataset, given the experimental configuration used in this study.

### *D. Visualization of Detection Results*

Following the quantitative evaluation using Precision, Recall, mAP@0.50, and mAP@0.50:0.95 metrics, the next step involves a visual analysis of the model's detection results. Visualizing these results aims to demonstrate the YOLO26l

model's ability to identify the locations of road damage using bounding boxes, while also providing insight into the quality of predictions across various road image conditions. This visual analysis complements the quantitative evaluation, allowing for a more comprehensive assessment of the model's performance.


*Figure 9 YOLO26l Detection Result*

Based on Figure 8, the YOLO26l model is capable of detecting road damage objects by displaying bounding boxes, class names, and confidence scores for each successfully identified object. In general, the predicted object locations align with the actual positions of road damage in the images, demonstrating the model's effective object localization capability.

For several samples, the model successfully detected potholes with a high level of confidence. Furthermore, alligator cracking, lateral cracking, and longitudinal cracking were also effectively identified, despite their distinct visual characteristics. This indicates that the model successfully learned the texture and shape patterns associated with each type of road damage during the training process. Nevertheless, certain conditions resulted in suboptimal detection performance. Objects that were relatively small, poor lighting conditions, or complex road surface textures could lead to reduced confidence scores or even result in objects going undetected [20]. These findings align with the Confusion Matrix results, which revealed that some objects were misclassified as background, thereby contributing to a lower Recall value.

Overall, the visualization of detection results demonstrates that the YOLO26l model produces reasonably accurate predictions for the four road damage categories within the dataset. These visual results are consistent with the previously obtained quantitative evaluation metrics, reinforcing the conclusion that the model performs well in the task of object-detection-based road damage identification.

The experimental results demonstrate that YOLO26l has the potential to support automated road inspection by providing consistent and objective detection of multiple road damage categories. Such capability may assist transportation agencies in reducing the reliance on manual visual inspections, thereby improving inspection efficiency and reducing operational costs. However, the current performance also indicates that further improvements are necessary before deployment in real-world environments. Variations in illumination, complex road textures, occlusions, weather conditions, and diverse road characteristics remain significant challenges that may affect detection reliability. Therefore, future work should focus on improving the robustness and generalization capability of the model through more diverse datasets, image enhancement techniques, and advanced training strategies.

## IV. CONCLUSION

This study aims to analyze the performance of the YOLO26l model as a baseline for detecting four categories of road damage potholes, alligator cracking, lateral cracking, and longitudinal cracking using the Road Damage Indonesia Dataset. Based on training and evaluation results, the YOLO26l model achieved a Precision of 0.7028, Recall of 0.6492, mAP@0.50 of 0.6890, and mAP@0.50:0.95 of 0.3391. These results indicate that the model demonstrates good capability in detecting various types of road damage under Indonesian road conditions.

Analysis using a confusion matrix, Precision-Recall curve, and visualization of detection results reveals that the model successfully identifies the majority of objects according to their actual classes. The pothole class yielded the best detection performance compared to the other classes; however, prediction errors were still observed with relatively small objects, poor lighting conditions, and complex road surface textures, leading some objects to be misclassified as background. Nevertheless, inter-class misclassification rates remained relatively low, demonstrating the model's ability to effectively distinguish the visual characteristics of each road damage category.

Overall, this study demonstrates that YOLO26l can serve as a solid baseline for object detection-based road damage detection tasks. Future research could explore more advanced data augmentation strategies, image enhancement techniques such as Contrast Limited Adaptive Histogram Equalization (CLAHE) and Histogram Equalization (HE), the use of higher image resolutions, the implementation of attention mechanisms, multi-scale feature fusion, and lightweight model architectures, as well as evaluation on other road damage datasets to improve detection accuracy, computational efficiency, and the model's generalization capability regarding road conditions [12], [13].

## REFERENCES


[1] Y. Zhang and L. Cheng, "The role of transport infrastructure in economic growth: Empirical evidence in the UK," *Transp. Policy (Oxf).*, vol. 133, pp. 223–233, Mar. 2023, doi: 10.1016/j.tranpol.2023.01.017.

[2] M. A. Rasee *et al.*, "AI-driven Vision-based Pothole Detection for Improved Road Safety," *Pertanika J. Sci. Technol.*, vol. 33, no. 3, Apr. 2025, doi: 10.47836/pjst.33.3.20.

[3] R. A. Gumelar and A. Susetyaningsih, "Pengaruh Kerusakan Jalan Terhadap Kenyamanan Pengguna Jalan di Jalan Raya," *Jurnal Konstruksi*, vol. 21, no. 2, pp. 265–274, Oct. 2023, doi: 10.33364/konstruksi/v.21-2.1416.

[4] X. Yang, J. Zhang, W. Liu, J. Jing, H. Zheng, and W. Xu, "Automation in road distress detection, diagnosis and treatment," Mar. 01, 2024, *KeAi Publishing Communications Ltd.* doi: 10.1016/j.jreng.2024.01.005.

[5] R. Fan, S. Guo, L. Wang, and M. J. Bocus, "Computer-Aided Road Inspection: Systems and Algorithms," Mar. 2022, doi: 10.48550/arXiv.2203.02355.

[6] A. B. Amjoud and M. Amrouch, "Object Detection Using Deep Learning, CNNs and Vision Transformers: A Review," *IEEE Access*, vol. 11, pp. 35479–35516, 2023, doi: 10.1109/ACCESS.2023.3266093.

[7] Z. Zou, K. Chen, Z. Shi, Y. Guo, and J. Ye, "Object Detection in 20 Years: A Survey," *Proceedings of the IEEE*, vol. 111, no. 3, pp. 257–276, Mar. 2023, doi: 10.1109/JPROC.2023.3238524.

[8] M. L. Ali and Z. Zhang, "The YOLO Framework: A Comprehensive Review of Evolution, Applications, and Benchmarks in Object Detection," Dec. 01, 2024, *Multidisciplinary Digital Publishing Institute (MDPI)*. doi: 10.3390/computers13120336.

[9] B. Fan and X. Song, "A Review of Deep Learning Based Object Detection Algorithms," *Journal of Engineering Research and Reports*, vol. 26, no. 11, pp. 88–99, Oct. 2024, doi: 10.9734/jerr/2024/v26i111316.

[10] M. N. Andrean *et al.*, "Comparing Haar Cascade and YOLOFACE for Region of Interest Classification in Drowsiness Detection," *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 8, no. 1, p. 272, Jan. 2024, doi: 10.30865/mib.v8i1.7167.

[11] H. Kusumah, M. R. Nurholik, C. P. Riani, and I. R. Nur Rahman, "Deep Learning for Pothole Detection on Indonesian Roadways," *Journal Sensi*, vol. 9, no. 2, pp. 175–186, Aug. 2023, doi: 10.33050/sensi.v9i2.2911.

[12] L. V. Fortin and O. E. Llantos, "Performance Analysis of YOLO versions for Real-time Pothole Detection," in *Procedia Computer Science*, Elsevier B.V., 2025, pp. 77–84. doi: 10.1016/j.procs.2025.03.013.

[13] R. Febriyanti Puspita, M. Naufal, and F. Al Zami, "Improving YOLO Performance with Advanced Data Augmentation for Soccer Object Detection," 2025. [Online]. Available: http://jurnal.polibatam.ac.id/index.php/JAIC

[14] R. C. Subianto, M. Naufal, and F. Alzami, "Improving YOLO12 Performance Using Efficient Channel Attention For Ship Object Detection," 2026. [Online]. Available: http://jurnal.polibatam.ac.id/index.php/JAIC

[15] S. S. Park, V. T. Tran, and D. E. Lee, "Application of various yolo models for computer vision-based real-time pothole detection," *Applied Sciences (Switzerland)*, vol. 11, no. 23, Dec. 2021, doi: 10.3390/app112311229.

[16] G. Jocher, J. Qiu, M. Liu, S. Lyu, F. C. Akyon, and M. E. Kalfaoglu, "Ultralytics YOLO26: Unified Real-Time End-to-End Vision Models," Jun. 2026, [Online]. Available: http://arxiv.org/abs/2606.03748

[17] R. Padilla, W. L. Passos, T. L. B. Dias, S. L. Netto, and E. A. B. Da Silva, "A comparative analysis of object detection metrics with a companion open-source toolkit," *Electronics (Switzerland)*, vol. 10, no. 3, pp. 1–28, Feb. 2021, doi: 10.3390/electronics10030279.

[18] T. Diwan, G. Anirudh, and J. V. Tembhurne, "Object detection using YOLO: challenges, architectural successors, datasets and applications," *Multimed. Tools Appl.*, vol. 82, no. 6, pp. 9243–9275, Mar. 2023, doi: 10.1007/s11042-022-13644-y.

[19] B. M. Hussein and S. M. Shareef, "An Empirical Study on the Correlation between Early Stopping Patience and Epochs in Deep Learning," *ITM Web of Conferences*, vol. 64, p. 01003, 2024, doi: 10.1051/itmconf/20246401003.

[20] S. S. A. Zaidi, M. S. Ansari, A. Aslam, N. Kanwal, M. Asghar, and B. Lee, "A Survey of Modern Deep Learning based Object Detection Models," May 2021, [Online]. Available: http://arxiv.org/abs/2104.11892.