Advanced Automated Road Damage Detection Usingyolov9: Leveraging Machine Learning and Computer Vision for Enhanced Infrastructure Maintenance

Jonathan Estilong

College of Information Technology and Computer Science, University of the Cordilleras, Baguio City, Philippines

In this study, a method for detecting road damage is introduced by utilizing the YOLOv9 model along with machine learning and computer vision techniques. Conventional methods for inspecting roads, such as surveys and visual checks, are time-consuming, error-prone, and inefficient. The goal is to create a system capable of assessing road conditions accurately and consistently in real time, thereby enhancing the effectiveness of maintenance tasks.

A dataset of 10,343 high-quality images, captured in various scenarios and annotated using Roboflow's tool, was employed. The focus was on identifying types of road damage, including lengthwise and transverse cracks, patches, fatigue cracks, and potholes. By training the YOLOv9 model with features like Gradient Information (PGI) and the Generalized Efficient Layer Aggregation Network (GELAN), a significant boost in detection accuracy was achieved during validation.

The assessment, considering metrics such as accuracy (0.643), completeness (0.626), average precision (mAP) (0.615), and F1 measure (0.634), indicates that the model performs well in identifying patches (accuracy: 0.867, completeness: 0.961, F1 measure: 0.911, mAP: 0.944) and potholes (accuracy: 0.772, completeness: 0.814, F1 measure: 0.792, mAP: 0.835). However, difficulties were encountered with lengthwise cracks (accuracy: 0.465, completeness: 0.336, F1 measure: 0.391, mAP: 0.351) and transverse cracks (accuracy: 0.569, completeness: 0.444, F1 measure: 0.499, mAP: 0.484). These variations highlight the necessity for further efforts, especially through improved data enhancement techniques and sophisticated deep learning structures.

The results emphasize the potential of the YOLOv9 model to transform road maintenance systems by delivering precise damage assessments, leading to better-maintained road networks. Future studies will concentrate on tackling issues such as model overfitting and enhancing performance for more subtle forms of damage to ensure effective use in various real-world situations.
Keywords: Automated Road Damage Detection, YOLOv9 Model, Machine Learning, Computer Vision, Data Augmentation.

1. Introduction

Maintaining roads, in condition is crucial for ensuring smooth transportation. Proper road upkeep not reduces the chances of accidents. Also improves traffic flow and cuts down on vehicle maintenance costs leading to significant savings for both governments and taxpayers (Shim et al., 2021). Conventional inspection methods, such as surveys and visual checks are time consuming, labor intensive and prone to errors (Kopsida et al., 2015). These traditional approaches often fall short in providing the comprehensive and real time data needed for maintenance planning (Gopalakrishnan, 2018). Manual surveys and visual inspections are sluggish demanding in terms of labor and inconsistent causing delays in identifying and addressing road damages. These methods are expensive. Offer information resulting in inefficient decisions regarding maintenance activities (Mukherjee et al., 2021). Leveraging machine learning techniques alongside computer vision holds promise for automated real time detection of road damage. Advanced models like YOLO (You Look can swiftly analyze vast datasets pinpoint various types of damage, with high accuracy thereby boosting the efficiency and effectiveness of maintenance tasks (Redmon & Farhadi 2018).

The main objective is to create an automated system that can detect road damage by utilizing the YOLOv9 model. This system will make use of machine learning and computer vision technologies to deliver time precise and consistent evaluations of road conditions. The specific aims involve training and validating the YOLOv9 model with a dataset to identify types of road damage achieving high accuracy, in detecting damage while minimizing false alarms applying the model in a practical setting to showcase its effectiveness, in real world situations and improving road safety and maintenance efficiency by offering prompt and accurate damage assessments (Maeda et al., 2020; Zhang et al., 2018).

2. Review of Literature

The progress, in creating automated systems for detecting road damage has been impressive thanks to advancements in machine learning and learning techniques. Arya and colleagues (2021) assessed how well a Japanese road damage detection model could be applied to countries by analyzing a dataset from Japan, India and the Czech Republic. Their study showed that models trained on data from countries performed better than those trained on data from one country underscoring the importance of using local data to enhance model accuracy. Similarly, Naddaf Sh and team (2020) leveraged the EfficientDet models to achieve a balance between accuracy, scalability and real time performance while obtaining F1 scores and fast inference speeds in the IEEE BigData 2020 Road Damage Detection

Challenge dataset. They stressed the need for annotation practices and larger datasets. Doshi and Yilmaz (2020) created a model based on YOLO v4 that combined models trained with various iterations and resolutions to achieve competitive F1 scores indicating the significance of having diverse training data. Kulambayev et al. (2020) employed Mask R CNN for real time detection and segmentation of road damages with precision recall rates and F1 scores while recommending the expansion of datasets, for studies.

These results were further supported when it was demonstrated that incorporating data enhanced the models effectiveness in situations highlighting the benefits of using diverse datasets from various countries (Arya et al., 2021).

The development of YOLO models, from YOLOv1 to YOLOv10 has led to enhancements in both the accuracy and speed of object detection. YOLOv1 pioneered a grid based approach (Redmon et al., 2016) with versions like YOLOv2 incorporating features such as batch normalization, anchor boxes and multi scale training for improved performance (Redmon & Farhadi 2017). The evolution continued with YOLOv3 introducing backbones and residual connections for detection capabilities (Redmon & Farhadi 2018). YOLOv4 brought innovations like the CSPDarknet53 backbone and advanced data augmentation methods resulting in accuracy and faster processing speeds (Bochkovskiy et al., 2020). Implemented in PyTorch YOLOv5 introduced AutoAnchor and mosaic augmentation techniques to offer versions catering to performance requirements (Jocher, 2020). Meanwhile the focus of YOLOv6 and YOLOv7 shifted towards applications by enhancing model learning processes and scalability (Li et al., 2022; Wang et al., 2022). The introduction of an anchor design with YOLOv8 aimed at reducing overheads while achieving efficient results (Jocher, 2023). Noteworthy advancements were made with the release of YOLOv9 which introduced the Programmable Gradient Information framework followed by YOLOv10 that eliminated maximum suppression, for improved speed and accuracy (Alif & Hussain 2024; Hussain, 2024).

The use of YOLO models, in agriculture, such as identifying weeds monitoring crops and detecting diseases showcases their ability to provide time and accurate results (Alif & Hussain 2024). The ongoing advancements in YOLO models have expanded their range of applications making them essential tools across industries, including agriculture. Future studies are likely to focus on improving model efficiency and exploring uses. These research endeavors emphasize the importance of utilizing datasets on a scale implementing transfer learning techniques and employing ensemble methods to create reliable automated systems, for detecting road damage that can be applied in different regions. This contributes significantly to the effectiveness and cost efficiency of road maintenance and safety management.

3. Methodology

Data Collection and Preprocessing

The dataset contains high quality images taken from road settings showcasing a variety of lighting and weather scenarios well as road types. This diverse collection of images helps the model adapt to real world conditions effectively. The focus of these images is, on road issues like cracks, potholes and surface irregularities. To label objects in the images accurately and efficiently the datasets images were annotated using Roboflows annotation tool, which supports annotation formats for integration, with different machine learning systems.

The dataset contains types of road damage annotations, such, as cracks that run along the roads length transverse cracks perpendicular to the road patches indicating previous repairs fatigue cracks resulting from repeated stress and potholes causing depressions in the road surface. There are 10,343 annotated images in total with a representation of each class, for training purposes. The average image size is 12.19 megapixels (Workspace Pavement, 2021).

To improve the model’s ability to adapt and perform across scenarios we applied various methods to modify the dataset. These methods involved flipping images both vertically with a 90% chance, for each rotating images to create perspectives and resizing them to 416x416 pixels for consistency in input sizes. Data augmentation plays a role in training machine learning models for tasks like identifying objects and categorizing them. The key reasons behind using data augmentation in this research are to introduce diversity enhance adaptability strengthen stability and maintain a dataset. By expanding the dataset through augmentation, we expose the model to an array of variations that aid in recognizing road damages under varying conditions and viewpoints. Augmented data enables the model to perform better on data it hasn't seen before reducing the risk of memorizing examples and improving its accuracy in real world settings. Techniques such as rotation and flipping make the model more resilient against distortions found in real world images like angles and perspectives. Augmentation helps balance the dataset by creating samples for classes that're less represented ensuring that sufficient training data is available, for all types of road damage.

These techniques have proven to be effective, in research studies. For instance, manipulating aspects like rotation and translation has been shown to enhance the resilience of models against transformations a crucial aspect in creating strong computer vision algorithms (Hao et al. 2023). Moreover, employing tailored data augmentation strategies has been noted to enhance the performance of object detection models. For example, experiments conducted on the COCO dataset suggest that optimized data augmentation techniques can boost detection accuracy by 2.3 mAP (Perez & Wang 2017).

YOLOv9 Model for Road Damage Detection

The YOLOv9 model used for detecting road damage is known for its design and effective training methods that lead to accuracy rates. By leveraging a backbone feature extractor made up of layers the model can identify types of road damage such, as cracks and potholes by capturing intricate patterns. Its multi scale detection heads enable the detection of both large damages effectively. The utilization of pretrained weights from a dataset streamlines the process through transfer learning accelerating training speeds and enhancing performance.

To address information bottlenecks and enhance detection accuracy the model incorporates Gradient Information (PGI) which focuses on training objectives. Additionally, the Generalized Efficient Layer Aggregation Network (GELAN) optimizes feature extraction by aggregating features from layers keeping the model lightweight yet robust enough for deployment in diverse environments including mobile devices for real time monitoring.

During training images are resized, converted to RGB format and augmented with rotations, flips and scaling to increase robustness. Hyperparameter tuning plays a role, in optimizing performance with training conducted over a number of epochs using suitable batch sizes and learning rates.

Evaluation of performance involves examining Precision, Recall and mean Average Precision (mAP) offering information, on accuracy and adaptability as mentioned by Wu et al. (2023).

By utilizing the capabilities of YOLOv9 the detection of road damage can attain levels of accuracy and efficiency. The incorporation of PGI and GELAN ensures that the model is robust yet lightweight enabling its suitability, for real time applications. This significantly contributes to enhancing automated road inspection systems and maintenance tasks according to Youwai et al. (2024).

Training Setup

The training took place in a Google Colab setting using an L4 GPU to speed up computations. The model employed was yolov9s.pt, the version of YOLOv9 selected for its blend of speed and accuracy. Training spanned 200 epochs to cover the dataset thoroughly. A batch size of 16 was chosen to balance efficiency and model generalization. Image input size was fixed at 640 pixels, for detection accuracy. This configuration enhances YOLOv9s ability to efficiently identify and categorize road damages.

Evaluation Metrics

The model, for detecting road damage underwent testing with metrics like precision, recall, mean Average Precision (mAP) and the F1 score. By conducting experiments on validation and test datasets valuable insights were gained regarding the accuracy and flexibility of the model. Furthermore, a visual assessment comparing the models' predictions, to annotations was carried out to enhance understanding of its performance and pinpoint areas for improvement.

Recall is a metric that assesses the percentage of identified instances compared to all the actual positive instances, in the dataset. The calculation of recall follows the formula depicted in Equation 1;

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \quad \text{(Equation 1)}$$

In contrast precision measures the accuracy of identified cases, among all instances labeled as positive by the model. Equation 2 illustrates the formula, for precision.

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \quad \text{(Equation 2)}$$

The Mean Average Precision (mAP) gives an evaluation of how the model performs across various categories by averaging the Average Precision (AP) scores, for each category. Equation 3 outlines the calculation, for mAP.

$\text{mAP} = \frac{1}{N}\sum_{i=1}^{N} \text{APi}$ (Equation 3)

In this scenario N represents the number of classes. The mean Average Precision (mAP) was computed by utilizing the Intersection, over Union (IoU) metric.

The F1 score, a metric that combines precision and recall through the mean offers an evaluation. It proves beneficial, in scenarios with imbalanced datasets accounting for both positives and false negatives. Equation 4 displays the formula, for calculating the F1 score.

$\text{F1 Score} = 2 \times \frac{\text{Precision}\times\text{Recall}}{\text{Precision}+\text{Recall}}$ (Equation 4)

When the model is assessed using these measures, insights are gained into its ability to detect road damages and its performance across different types of damage. Precision and recall indicate the model's accuracy in spotting road damages and its capability to locate all instances. The mAP metric provides a comprehensive view of performance across all categories, while the F1 score balances precision and recall, offering a reliable measure for assessment.

Performing these evaluations helps identify the points of the model and areas that need improvement guaranteeing its efficiency in conditions. This comprehensive approach, to assessing performance allows for adjustments to be made to the model ultimately enhancing its accuracy and dependability in detecting road damage, in real world situations.

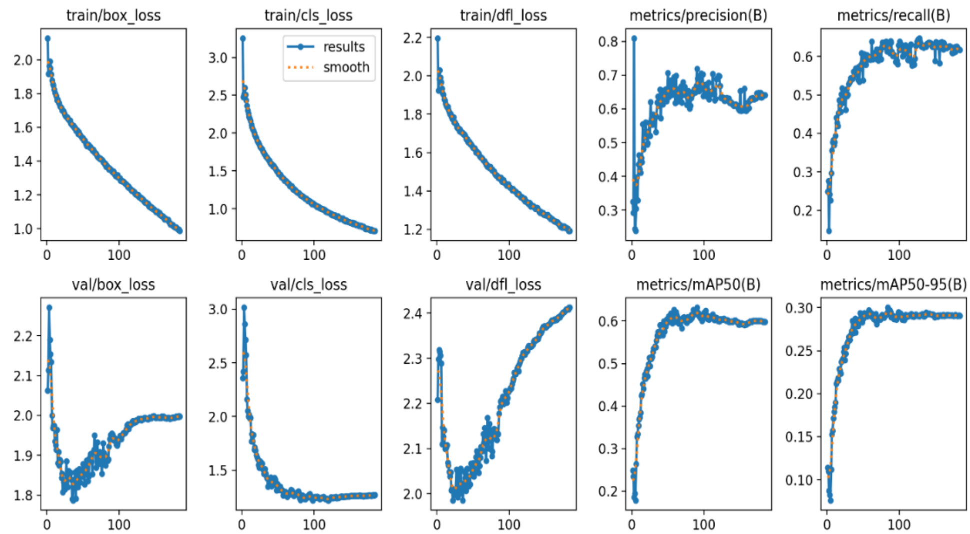
4. Results

In this section, the outcomes of training and validating the YOLOv9 model for detecting road damage are discussed. The study looks at how the model performs using measures, like accuracy, recall, average precision (mAP) and F1 score. The examination points out where the model shines in identifying road damages and pinpoints areas that need adjustments. It explores differences in performance among categories. Looks into potential problems, with overfitting offering a detailed assessment of the model's effectiveness and possible enhancements.

Analysis of Performance Metrics and Image-Based Results

The training and validation losses, for box, class and DFL (Distance from Line) show a decrease during the training phase indicating learning. However, while validation losses initially drop they later begin to rise suggesting overfitting. This implies that although the model performs well on training data, its performance on validation data worsens over time. Precision and recall metrics exhibit progress throughout the training process albeit with some fluctuations. This indicates that the model is getting better at making predictions but struggles to maintain performance across all categories. The mean Average Precision (mAP) metrics, mAP50 and mAP50 95 also demonstrate enhancements, indicating precision and recall across different Intersection over Union (IoU) thresholds. Nonetheless the mAP50 95 metric remains lower highlighting the difficulties, in achieving precision and recall simultaneously at thresholds. These findings are depicted in Figure 1.

Figure 1. Training and Validation Losses, Precision, Recall, and mAP Training and Validation Losses



The confusion matrix reveals class-specific performance variations, with certain classes such as longitudinal and transverse cracks exhibiting higher misclassification rates. This correlates with the lower precision and recall observed in the mean metrics. Conversely, classes like patches and potholes show higher true positive rates, which align with their higher precision and recall values. These discrepancies underscore the model's varying effectiveness across different types of road damage, as illustrated in Figure 2.

The confusion matrix shows that some classes, like transverse cracks have misclassification rates impacting the precision and recall metrics. On the hand classes such, as patches and potholes exhibit true positive rates leading to better precision and recall values. These differences highlight how the model performs differently for types of road damage as depicted in Figure 2.

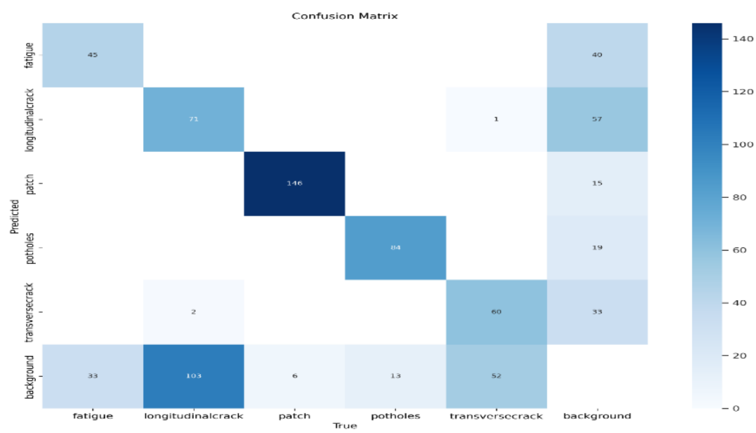


Figure 2. Confusion Matrix

The graph, in Figure 3 shows how performance changes for classes as confidence levels go up. For example, the class for patches consistently maintains precision at all confidence levels. On the hand classes like cracks see a noticeable decrease in precision at higher confidence levels. This difference as shown in Figure 3 leads to a precision value of 0.612 which's decent but suggests some variations in precision, at different confidence thresholds.

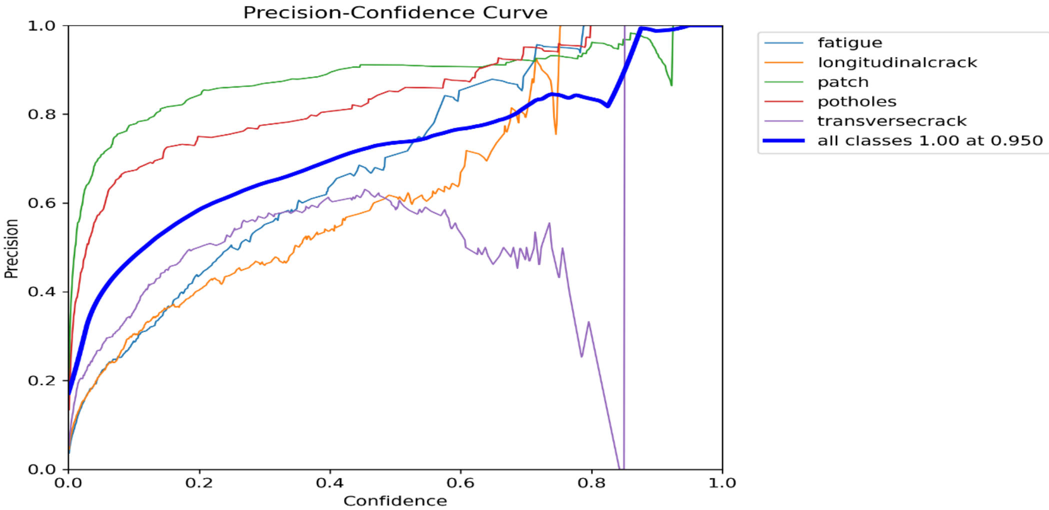


Figure 3. Precision-Confidence Curve

The graph, in Figure 4 highlights the differences in recall based on confidence levels. When confidence levels rise classes like patches maintain a recall rate. Classes like longitudinal cracks experience a significant drop in recall. This suggests that while some classes consistently achieve recall rates others face challenges, in maintaining performance as confidence levels increase.

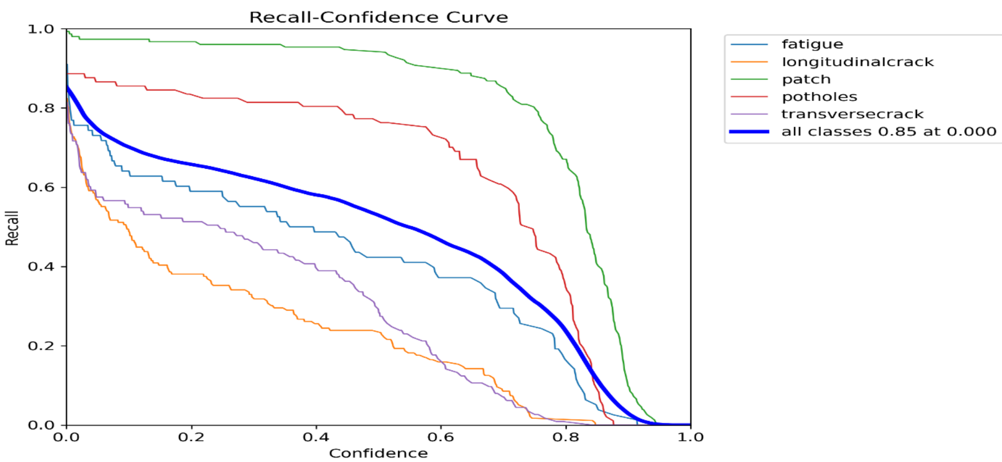


Figure 4. Recall-Cofidence Curve

The graph, in Figure 5 depicts the balance between precision and recall for categories, such as patches, longitudinal cracks and transverse cracks. While patches demonstrate precision and recall the values are lower for transverse cracks. This leads to a recall of 0.577 and an average F1 score of 0.594. These measurements underscore the difficulties, in maintaining performance across all categories showcasing both the strengths and weaknesses of the model in identifying different types of road damage.

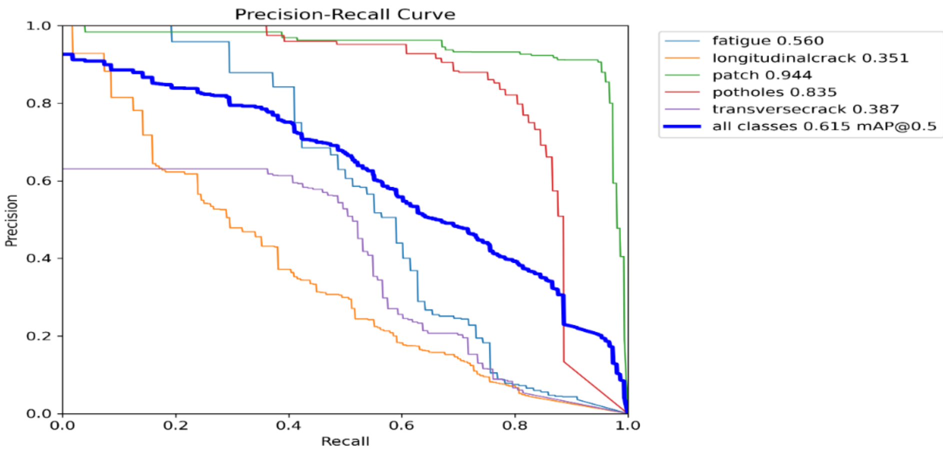


Figure 5. Precision-Recall Curve

Finally, the graph showing the F1 confidence curve, in Figure 6 represents how well precision and recall are balanced. It provides a measure that combines both aspects. The F1 score remains consistently high for categories like patches at confidence levels indicating an effective performance. However, there is a decrease in the F1 score for categories such as cracks at higher confidence levels underscoring the challenge of maintaining consistent performance across all types of road damage. The average F1 score of 0.63 at a confidence level of 0.293 highlights the importance of tuning to achieve performance, across all categories.

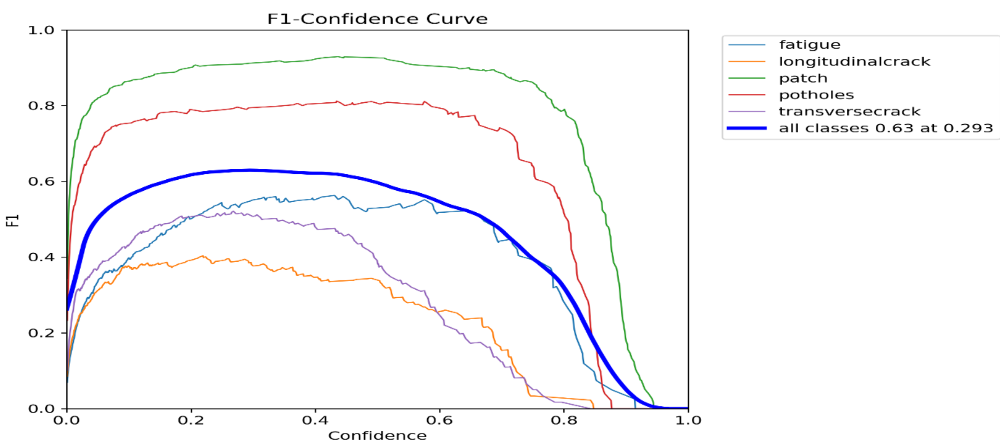


Figure 6. F1-Confidence Curve

Overall Performance Across All Classes

The models overall performance, across all categories appears satisfactory with a precision of 0.643 recall of 0.626 F1 score of 0.634 and mAP of 0.615. These metrics suggest that the model performs adequately but still has room for improvement to enhance accuracy and consistency across all categories.

When it comes to fatigue cracks the model demonstrates precision at 0.541 indicating some positives and a recall of 0.551 implying some missed instances. The balanced F1 score of 0.546 reflects the trade off between precision and recall in this category. The mAP value of 0.560 indicates a capability in identifying and categorizing fatigue cracks.

On the hand when dealing with cracks the model encounters more challenges with lower precision at 0.465 suggesting a higher false positive rate and a recall of 0.336 indicating many missed instances. The F1 score of 0.391 underscores the difficulty in detecting cracks while the relatively low mAP value of 0.351 points to challenges in precise localization and classification.

In contrast to these findings the model excels in recognizing patches with a precision of 0.867 and a high recall of .961 suggesting accurate detection, for most instances involving patches which results in an outstanding F1 score .911.

The impressive mAP score of 0.944 suggests performance, in pinpointing and categorizing patches. When it comes to potholes the model exhibits a precision of 0.772 and a recall rate of 0.814 indicating its ability to accurately spot and detect most potholes. A solid F1 score of 0.792 signifies a balance between precision and recall while the high mAP value of 0.835 showcases its efficiency in localizing and categorizing potholes.

On the hand the model's performance with cracks is average with a precision score of 0.569 suggesting some false detections and a lower recall rate of 0.468 indicating missed instances. The F1 score of 0.514 reflects this performance while the low mAP value of 0.387 points towards challenges in precise localization and classification.

In this breakdown, by category as shown in Table 1 we gain understanding of the models' strengths and areas for enhancement in various types of road damage. It emphasizes the model's proficiency, in detecting patches and potholes while pinpointing aspects that could use improvement with regards to transverse cracks.

Table 1. Performance Across All Classes

Class	Images	Instances	Precision (P)	Recall (R)	mAP50	mAP50-95
All	227	616	0.643	0.626	0.615	0.382
Fatigue	45	78	0.541	0.551	0.560	0.195
Longitudinal Cracks	67	176	0.465	0.336	0.351	0.137
Patch	102	152	0.867	0.961	0.944	0.621
Potholes	56	97	0.772	0.814	0.835	0.402
Transverse Cracks	56	113	0.569	0.468	0.387	0.143

5. Discussion

The YOLOv9 model showed performance, in identifying and categorizing types of road damage like patches and potholes. These categories displayed levels of accuracy and consistency as seen in the precision confidence and recall confidence curves across confidence thresholds. The models ability to detect patches and potholes effectively can be credited to their characteristics and larger sizes making them easily recognizable. The strong mean F1 score and mAP for these categories indicate the models capacity to maintain a balance between accuracy and completeness ensuring identification and classification in real world scenarios.

On the hand the model encountered challenges when detecting transverse cracks leading to lower precision and recall rates. The precision confidence and recall confidence curves for these categories exhibited decreases at confidence levels suggesting a struggle for the model in maintaining accuracy with increased confidence thresholds. This difficulty implies that longitudinal and transverse cracks, which are more subtle and visually similar to road features present a task for the model. The mAP values for these categories underscore the complexities involved in achieving localization and classification indicating a requirement, for additional optimization of the model and enhancement of data resources.

The examination of training and validation losses uncovered an issue of overfitting, where the validation loss began to rise after decreasing while the training loss kept decreasing. This pattern suggests that the model works well on the training data but struggles to generalize to validation data. Overfitting is a problem, in learning models, where the model gets too specialized in the training data, capturing noise and specific patterns that don't transfer well to fresh data.

To tackle overfitting various strategies can be utilized. Integrating regularization methods like dropout and weight decay can help alleviate overfitting. Dropout randomly eliminates units from the network during training pushing the model to learn durable features. Weight decay penalizes weights, encouraging models that generalize better. Enhancing data augmentation techniques can boost variability in the training data aiding in generalization by the model. Advanced augmentation methods such as mixup and cutmix can generate intricate training samples enhancing the model's adaptability to diverse variations. Implementing stopping during training can prevent overfitting by stopping training when the validation loss begins to rise (Vishwakarma & Vennelakanti 2020; Li et al., 2024).

To enhance the model's performance in identifying classes, like transverse cracks several avenues can be explored.

Enhancing the variety and volume of training data particularly focusing on represented categories, like transverse cracks can enhance the models learning process. Targeted collection of data and the creation of data sets are strategies to achieve this objective. It is crucial to ensure high quality annotations to enable the model to learn from the data. The use of annotation tools and consistent labeling techniques can elevate the quality of the dataset. Exploring deep learning architectures that incorporate attention mechanisms can help direct the models focus towards image components resulting in improved accuracy in detecting subtle and intricate damage patterns. Furthermore, employing scale feature extraction methods can boost performance in identifying damages of various sizes.

Implementing transfer learning from models pre trained on tasks can significantly enhance the model's performance levels. Fine tuning a trained model for road damage detection tasks allows for leveraging previously learned features, from extensive datasets thereby accelerating training processes and increasing accuracy rates. Additionally adopting methods that combine predictions from models can enhance both robustness and accuracy. Utilizing a comprising models trained with configurations and subsets of data enables capturing a broader spectrum of patterns thereby reducing chances of misclassifications.

The YOLOv9 models practical application, in automated road inspection systems shows promise. Mobile apps utilizing the model can instantly. Categorize road damage during inspections allowing for reporting and documentation. Drones equipped with high quality cameras and the model can efficiently survey road networks in remote or difficult to access areas. Fixed cameras strategically placed can monitor road conditions continuously using the model to aid in maintenance and repairs. Maintenance vehicles with automated features and the model can decrease the need for inspections streamlining maintenance processes. Customizing the model with data can enhance its performance in regional settings boosting effectiveness across different landscapes. These examples demonstrate how the YOLOv9 model enhances efficiency, cost effectiveness and responsiveness in road maintenance practices resulting in more well-kept road infrastructure (Wang et al., 2023; Yu & Zhou 2023).

Future studies should address real time implementation challenges like efficiency and hardware demands. Ensuring performance of the model on devices such as mobile phones and embedded systems is essential, for practical usage.

Furthermore, delving into how the model can be combined with data sources such, as traffic and weather updates could offer an insight into road conditions and improve decision making in maintenance planning. By overcoming these challenges and exploring methods we can greatly enhance the reliability and efficiency of automated systems for detecting road damage especially when it comes to complex types like longitudinal and transverse cracks. It is crucial for studies to prioritize these approaches in order to boost the model's performance overall and ensure its adaptability, for a range of real-world scenarios (Yu & Zhou 2023; Wang & Yoon 2021).

6. Conclusion

In conclusion, significant progress has been made in determining whether the YOLOv9 model can effectively identify and categorize road damages. The model demonstrates performance, in spotting patches and potholes boasting precision and recall values of over 0.8 for these categories. However, it encounters challenges when it comes to identifying transverse cracks with precision and recall values dropping below 0.6. These findings suggest that while the model holds promise, for automated road maintenance systems, further research and refinement are necessary to enhance its accuracy and dependability across a spectrum of road damage types.

References

1. Arya, D., Maeda, H., Ghosh, S. K., Toshniwal, D., Mraz, A., Kashiyama, T., & Sekimoto, Y. (2020). Transfer learning-based road damage detection for multiple countries. arXiv preprint arXiv:2008.13101.
2. Arya, D., Maeda, H., Ghosh, S. K., Toshniwal, D., Mraz, A., Kashiyama, T., & Sekimoto, Y. (2021). Deep learning-based road damage detection and classification for multiple countries. Automation in Construction, 132, 103935.
3. Naddaf-Sh, S., Naddaf-Sh, M. M., Kashani, A. R., & Zargarzadeh, H. (2020, December). An efficient and scalable deep learning approach for road damage detection. In 2020 IEEE International Conference on Big Data (Big Data) (pp. 5602-5608). IEEE.
4. Doshi, K., & Yilmaz, Y. (2020, December). Road damage detection using deep ensemble learning. In 2020 IEEE International Conference on Big Data (Big Data) (pp. 5540-5544). IEEE.
5. Kulambayev, B., Beissenova, G., Katayev, N., Abduraimova, B., Zhaidakbayeva, L., Sarbassova, A., ... & Shyrakbaye, A. (2022). A deep learning-based approach for road surface damage detection. Applied Sciences, 12(8), 3360.
6. Arya, D., et al. (2020). Deep learning-based road damage detection and classification for multiple countries. In 2020 IEEE International Conference on Big Data (Big Data).
7. Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 779-788).
8. Redmon, J., & Farhadi, A. (2017). YOLO9000: better, faster, stronger. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 7263-7271).
9. Redmon, J., & Farhadi, A. (2018). Yolov3: An incremental improvement. arXiv preprint arXiv:1804.02767.
10. Bochkovskiy, A., Wang, C. Y., & Liao, H. Y. M. (2020). Yolov4: Optimal speed and accuracy of object detection. arXiv preprint arXiv:2004.10934.
11. Jocher, G. (2020). YOLOv5. Ultralytics. Retrieved from https://ultralytics.com.
12. Li, C., Li, L., Jiang, H., Weng, K., Geng, Y., Li, L., ... & Wei, X. (2022). YOLOv6: A single-stage object detection framework for industrial applications. arXiv preprint arXiv:2209.02976.
13. Wang, C. Y., Bochkovskiy, A., & Liao, H. Y. M. (2023). YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 7464-7475).
14. Jocher, G. (2023). YOLOv8. Ultralytics. Retrieved from https://ultralytics.com.
15. Al Rabbani Alif, M., & Hussain, M. (2024). YOLOv1 to YOLOv10: A comprehensive review of YOLO variants and their application in the agricultural domain. arXiv e-prints, arXiv-2406.
16. Hussain, M. (2023). YOLO-v1 to YOLO-v8, the rise of YOLO and its complementary nature toward digital manufacturing and industrial defect detection. Machines, 11(7), 677.
17. Gopalakrishnan, K. (2018). Deep learning in data-driven pavement image analysis and automated distress detection: A review. Data, 3(3), 28.
18. Maeda, H., Sekimoto, Y., Seto, T., Kashiyama, T., & Omata, H. (2018). Road damage detection using deep neural networks with images captured through a smartphone. arXiv preprint arXiv:1801.09454.
19. Mukherjee, R., Iqbal, H., Marzban, S., Badar, A., Brouns, T., Gowda, S., ... & Zonooz, B. (2021). AI driven road maintenance inspection. arXiv preprint arXiv:2106.02567.
20. Redmon, J., & Farhadi, A. (2018). "YOLOv3: An Incremental Improvement." arXiv preprint arXiv:1804.02767.
21. Shim, S., Kim, J., Lee, S. W., & Cho, G. C. (2021). Road surface damage detection based on hierarchical architecture using a lightweight auto-encoder network. Automation in Construction, 130, 103833.

22. Zhang, Z., Xie, Y., Xing, F., McGough, M., & Yang, L. (2018). MDNet: A Semantically and Visually Interpretable Medical Image Diagnosis Network. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 779-788).
23. Kopsida, M., Brilakis, I., & Vela, P. A. (2015, October). A review of automated construction progress monitoring and inspection methods. In Proc. of the 32nd CIB W78 Conference 2015 (pp. 421-431).
24. Youwai, S., Chaiyaphat, A., & Chaipetch, P. (2024). YOLO9tr: A Lightweight Model for Pavement Damage Detection Utilizing a Generalized Efficient Layer Aggregation Network and Attention Mechanism. arXiv preprint arXiv:2406.11254.
25. Ultralytics. (2024). YOLOv9: Architecture and Techniques. Ultralytics Documentation. Retrieved from Ultralytics YOLO Docs.
26. Wu, C., Ye, M., Zhang, J., & Ma, Y. (2023). YOLO-LWNet: A lightweight road damage object detection network for mobile terminal devices. Sensors, 23(6), 3268.
27. Hao, X., Liu, L., Yang, R., Yin, L., Zhang, L., & Li, X. (2023). A review of data augmentation methods of remote sensing image target recognition. Remote Sensing, 15(3), 827.
28. Perez, L., & Wang, J. (2017). The effectiveness of data augmentation in image classification using deep learning. arXiv preprint arXiv:1712.04621.
29. Workspace Pavement. (2021, November). Pavement Distresses Dataset [Data set]. Roboflow. https://universe.roboflow.com/workspace-pavement/pavement-distresses.
30. Sami, A. A., Sakib, S., Deb, K., & Sarker, I. H. (2023). Improved YOLOv5-based real-time road pavement damage detection in road infrastructure management. Algorithms, 16(9), 452.
31. Vishwakarma, R., & Vennelakanti, R. (2020, December). CNN model & tuning for global road damage detection. In 2020 IEEE international conference on big data (Big Data) (pp. 5609-5615). IEEE.
32. Li, Y., Yin, C., Lei, Y., Zhang, J., & Yan, Y. (2024). RDD-YOLO: Road Damage Detection Algorithm Based on Improved You Only Look Once Version 8. Applied Sciences, 14(8), 3360.
33. Wang, X., Gao, H., Jia, Z., & Li, Z. (2023). BL-YOLOv8: An improved road defect detection model based on YOLOv8. Sensors, 23(20), 8361.
34. Yu, G., & Zhou, X. (2023). An improved YOLOv5 crack detection method combined with a bottleneck transformer. Mathematics, 11(10), 2377.
35. Liu, L., Zuo, H., & Qiu, X. (2021, April). Research on defect pattern recognition of light guide plate based on deep learning semantic segmentation. In Journal of Physics: Conference Series (Vol. 1865, No. 2, p. 022033). IOP Publishing.
36. Wan, F., Sun, C., He, H., Lei, G., Xu, L., & Xiao, T. (2022). YOLO-LRDD: A lightweight method for road damage detection based on improved YOLOv5s. EURASIP Journal on Advances in Signal Processing, 2022(1), 98.
37. Wang, L., & Yoon, K. J. (2021). Knowledge distillation and student-teacher learning for visual intelligence: A review and new outlooks. IEEE Transactions on Pattern Analysis and Machine Intelligence, 44(6), 3048-3068.