Research Article

Development of a Cloud-Based Road Surface Quality Assessment System

Lim Ka Quan[1], Nor Azuana Ramli[1]

1. Centre for Mathematical Sciences, University Malaysia Pahang, Malaysia

This research aims to develop a cloud-based system utilizing the You Only Look Once version 8 (YOLOv8) model for assessing road surface quality. The system is designed to address critical road maintenance challenges and the need for high accuracy and fast response road surface quality monitoring. Data acquisition involved images from the Internet, dashcams, and smartphones, with subsequent processing through advanced image techniques. The YOLOv8 model demonstrated efficacy in detecting various road surface defects, achieving a precision of 0.457 and a recall of 0.486. While exhibiting potential in identifying patches and potholes, further refinement is required for crack detection. The model's processing speed, with 9.7 milliseconds per image, indicates its capability for near real-time analysis. Finally, the model is deployed on cloud infrastructure hosted by Digital Ocean to provide scalability and accessibility. The cloud-based system enables users to upload videos for automated defect detection and offers downloadable results, fostering collaborative initiatives in road surface monitoring. While the model shows promise, particularly in detecting patches and potholes, crack detection has room for improvement. Future work could focus on enhancing the model's performance for this challenging defect class.

Correspondence: papers@team.qeios.com — Qeios will forward to the authors

1. Introduction

The inconvenience and uncertainty of public transportation in Malaysia cause many citizens to rely on driving. However, the condition of the roads they traverse is a critical factor that impacts their safety. Maintaining high-quality road surfaces is essential for ensuring road-user safety, facilitating economic activities, and providing access to essential services. Recognizing this, the Malaysian government must prioritize addressing road surface quality to mitigate the risks associated with poor infrastructure.

 doi.org/10.32388/04NBI2.2

The Ministry of Transport Malaysia (MOT) reports that around 600,000 road accident cases have been recorded throughout 2023[1]. In that time, 6,443 lives were lost to road accidents, which translates to an average of 18 deaths every day last year[2]. A study by the Malaysian Institute of Road Safety Research (MIROS) found that road accidents were mainly caused by human behaviour, followed by the design, condition of road infrastructure, and vehicles' condition[3]. This study shows that poor road conditions significantly contribute to road accidents, although they are not the leading cause. There is still a potential to reduce the number of road accidents and injuries by improving road surfaces, thereby fostering sustainable cities and communities.

Understanding the nature of road surface defects is the first step toward addressing this issue. According to Rolt et al.[4], road surface defects pose significant challenges to the durability and safety of transportation infrastructure[4]. For instance, potholes, which are depressions in the road surface, form due to factors such as water infiltration, freezing and thawing cycles, and traffic wear and tear. These defects can cause vehicle damage and present safety hazards for drivers. Another common defect, fatigue cracking, often referred to as crocodile cracking, emerges from the continuous pressure exerted by heavy vehicles, combined with temperature changes and natural aging. These cracks resemble the skin of a crocodile and, if not addressed promptly, can intertwine and exacerbate structural vulnerabilities in the roadway.

To tackle these issues, modern technology offers promising solutions. Emara et al.[5] highlighted the importance of detecting and assessing road surface conditions to ensure the safety and efficiency of transportation networks. The rise of smartphones and their built-in sensors has led to innovative methods like mobile crowdsensing (MCS) for evaluating road surface quality. This approach leverages the widespread use of smartphones to collect data on road conditions, offering a cost-effective and scalable solution[5]. The future of road surface quality assessment lies in integrating advanced technologies. Lasers, cameras, sensors, smartphones, and cloud storage services are revolutionizing how we monitor road conditions.

Although the initial cost of these systems may be high, their benefits in terms of accuracy and coverage make them a worthwhile investment. A cloud-based road surface quality assessment system can be developed by harnessing the computing power and various sensors in smartphones and dashcams, addressing the growing need for efficient monitoring. The primary goal of this study is to design a cost-effective and highly accurate road surface quality assessment system that effectively monitors road

doi.org/10.32388/04NBI2.2

damage and ultimately enhances road user safety. By leveraging deep learning and cloud-based technology, such a system can provide real-time insights and timely interventions, ensuring safer and more reliable transportation infrastructure for all.

2. Methodology

Figure 1 illustrates the framework of the research study related to the cloud-based road surface quality assessment. Data for this study was collected from site visits and the internet, such as Google Images. These images were pre-processed and partitioned into test data and train data. A model was then developed based on the train data. This model was evaluated using the model's predicted value and test data through cross-validation. If the results were insignificant, the process returned to the model development stage for refinement. Once the model evaluation was satisfactory, the model was deployed in the Digital Ocean cloud service.

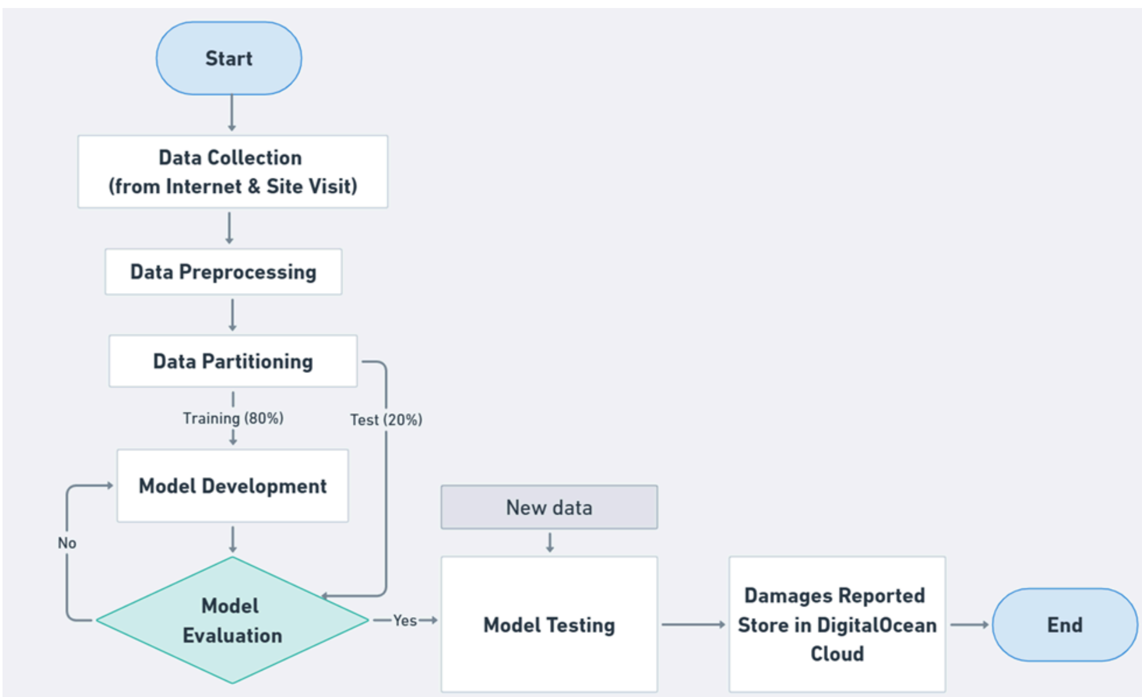

Figure 1. Research Framework

3. Result

In this research, Python was used in the Jupyter Notebook to run all the image preprocessing and analysis. Figure 2 shows the video recorded by using dashcam.

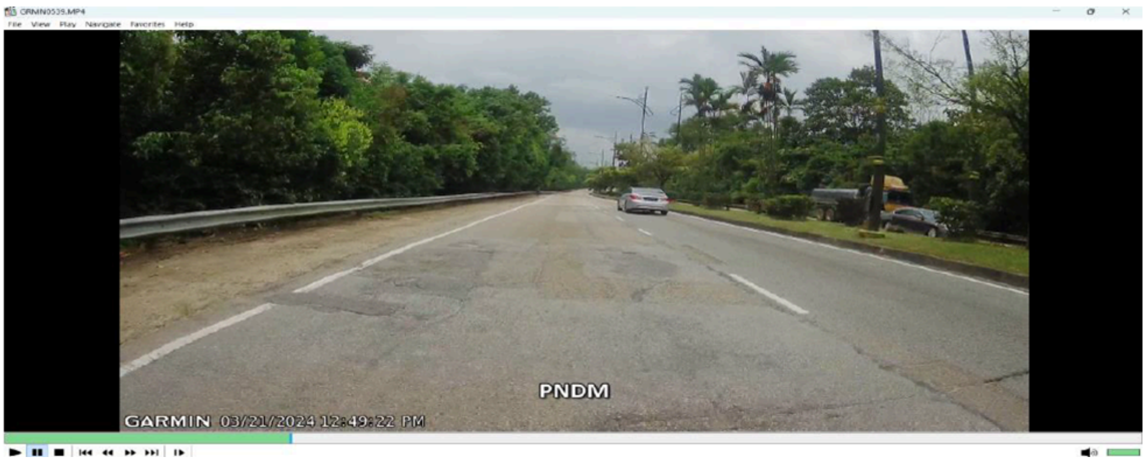

Figure 2. Video Recorded using Dashcam

Before model development, video and image preprocessing were performed. First, frames were extracted from the video and resized to 640*640 pixels. Roboflow was implemented to perform image preprocessing, such as annotation, resizing, augmentation, and data partitioning. After completing all the preprocessing, this dataset was exported to a Jupyter Notebook environment for model training. Model used for this study is YOLOv8. YOLOv8 was chosen because of its state-of-the-art (SOTA) object detection model, which balances speed, accuracy, and flexibility[6]. Building on its predecessors in the YOLO series, YOLOv8 introduces architectural advancements that improve precision and recall, enabling highly accurate detections even in complex environments. It maintains the hallmark real-time inference speed of YOLO models, making it ideal for time-sensitive applications like autonomous vehicles and video surveillance. The model's versatility is another key strength, as it supports object detection, instance segmentation, and image classification within a unified framework. With multiple model sizes available (e.g., nano, small, medium, large), YOLOv8 is scalable for deployment on devices ranging from resource-constrained edge devices to powerful GPUs.

For this study, the number of epochs for model development was 200, and the batch size was 16. By default, the model training will trigger early stopping if there is no improvement in the last 50 epochs to avoid overfitting; thus, this model stopped at epoch 153. The best result was obtained at epoch 103.

 doi.org/10.32388/04NBI2.2

Overall, for 263 images and 1180 instances, the model achieved a precision of 0.457, a recall of 0.486, a mAP50 of 0.441, and a mAP50-95 of 0.237. Specifically, for the “crack” class, which had 711 instances, the model demonstrated lower performance with a precision of 0.2, recall of 0.254, mAP50 of 0.136, and mAP50-95 of 0.0373. In contrast, the “patch” class, with 113 instances, showed the highest performance, achieving a precision of 0.651, recall of 0.743, mAP50 of 0.703, and mAP50-95 of 0.504. The “fatigue” class, with 242 instances, had moderate performance, with a precision of 0.415, recall of 0.471, mAP50 of 0.409, and mAP50-95 of 0.154. Lastly, the “pothole” class, comprising 114 instances, performed reasonably well with a precision of 0.564, recall of 0.476, mAP50 of 0.515, and mAP50-95 of 0.251. These metrics indicate that while the model performs reasonably well in detecting “patch” and “pothole,” it struggles more with identifying “crack,” highlighting areas for potential improvement.

The model’s speed is also broken down as follows: it takes 0.7 milliseconds (ms) for preprocessing, 7.5 ms for inference, 0.0 ms for loss computation, and 1.5 ms for postprocessing per image. This breakdown underscores that the model is optimized for real-time or near-real-time analysis, with the majority of the processing time dedicated to inference, the task of detecting objects in the image. Preprocessing and postprocessing times are relatively minimal, suggesting that the model efficiently prepares the data for analysis and processes the results. The 0.0 ms loss time confirms that loss computation, typically used during training to optimize the model, is not a factor during the evaluation phase. This efficient processing is a key feature for road anomaly detection in dashcam footage applications, instilling confidence in the model’s applicability.

A precision versus recall (PR) curve is called out to illustrate the trade-off between precision and recall for different thresholds. A perfect PR curve would achieve a precision of 1 and a recall of 1, which are at the top right of the plot. For instance, the “pothole” class achieves high precision but at the cost of lower recall. This means the model is very accurate when predicting a pothole, but it might miss some actual one. mAP is a metric that summarizes the PR curve by averaging the precision across different recall thresholds. The value “0.632 mAP@0.5” indicates the mean Average Precision (mAP) at a recall level of 0.5, as shown in Figure 3.

doi.org/10.32388/04NBI2.2

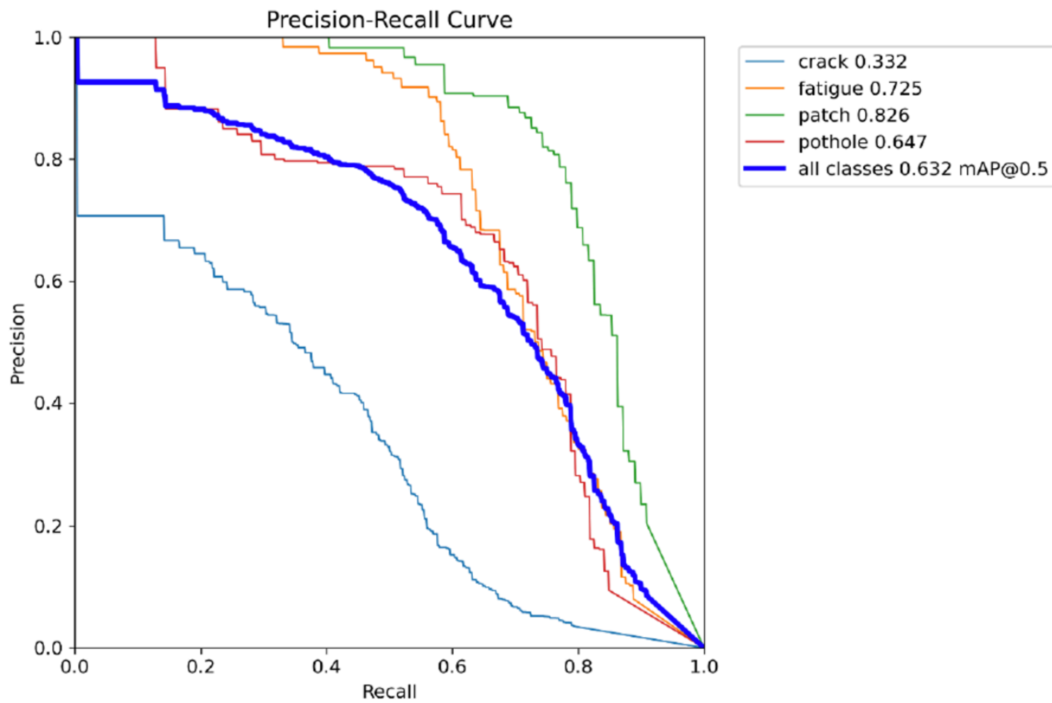

Figure 3. Precision versus Recall Curve

Figure 4 illustrates the training and validation performance of an object detection model. The training loss curves for box loss, classification loss, and distribution focal loss show a steady decrease over epochs, indicating effective learning during the training phase. Similarly, the precision and recall metrics improve consistently before stabilizing, reflecting the model's ability to correctly identify positive cases while reducing false positives and capturing more true positives over time.

For the validation metrics, the losses decrease initially but exhibit some fluctuations, particularly the box loss, which may suggest potential overfitting or challenges in generalizing to unseen data in later epochs. However, the mAP metrics (mean Average Precision at 50% IoU and across IoUs from 50% to 95%) improve and plateau, highlighting strong detection accuracy and consistency. Despite the validation loss oscillations, the overall performance trends indicate the model is achieving decent results. To further enhance the model, it may be beneficial to investigate the causes of validation loss fluctuations and apply regularization techniques such as dropout or data augmentation. Additionally, using early stopping could help mitigate overfitting. Finally, testing the model on a separate dataset would confirm its ability to generalize effectively.

doi.org/10.32388/04NBI2.2

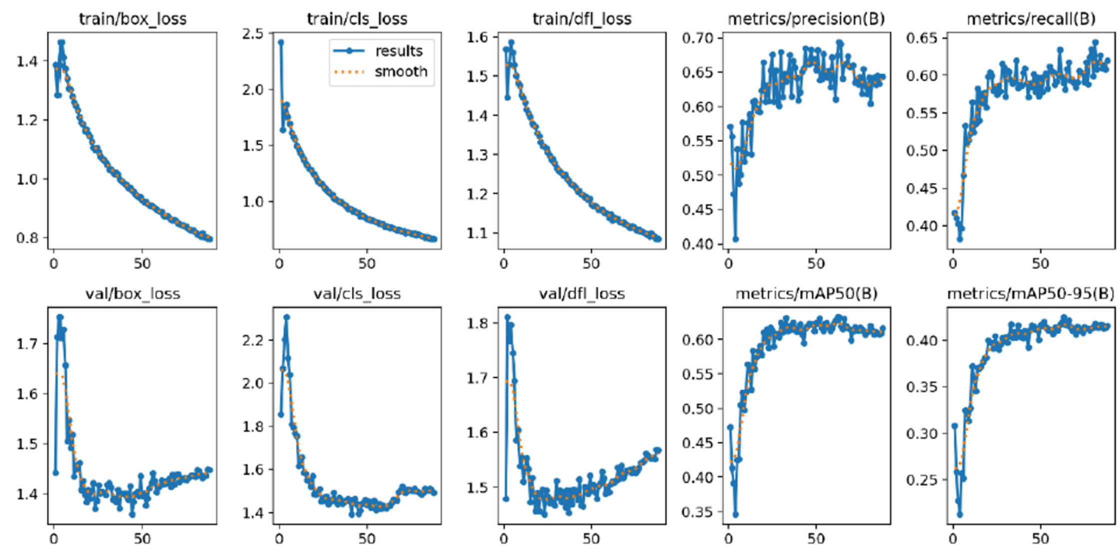

Figure 4. Overall Results for YOLOv8 Model Training

The model has been deployed on cloud infrastructure hosted by Digital Ocean to ensure scalability and accessibility. Users need to enter their username and password to access the system. Once logged in, they can upload videos to the cloud. The deployed model will then begin detecting potholes, cracks, fatigue, and patches in the videos. After detection, users can download the results, which will include the latitude and longitude of the road defects stored in a CSV file, as well as the frames of the detections stored in a ZIP file. To export the results, users can click on the export button for specific video files.

4. Conclusion

The study introduces a cloud-based method for assessing road surface quality, offering the potential to significantly reduce the time required for road defect detection, improve accuracy, and lower associated costs. By utilizing GIS technology, defect locations are accurately recorded using latitude and longitude coordinates, minimizing errors in the process. The research presents a deep learning model capable of accurately detecting and classifying road surface defects such as cracks, patches, potholes, and fatigue. Model performance was assessed based on accuracy, accuracy per epoch, and loss per epoch, with the YOLOv8 framework employed to identify and classify road surface defects. Video data was processed by extracting frames and converting them to JPG format for input. The sample data, labelled with the correct defect types using Roboflow, was utilized for model training to ensure accuracy and reliability.

doi.org/10.32388/04NBI2.2

The dataset for this research comprised 4,000 images collected from various sources, including Google, and on-site images from the Johor area. Preprocessing of the images involved enhancement, resizing, and annotation through Roboflow, which was pivotal in standardizing the images, ensuring accurate annotations, and addressing class imbalance issues. Once the pre-trained model was developed, it was applied to detect and classify road surface defects in real-world scenarios. The model's performance was evaluated using the F1-score, recall-precision graph, and confusion matrix. The findings revealed that YOLOv8 demonstrated high accuracy in detecting and classifying defects for the "patch," "fatigue," and "potholes" classes. However, its performance was less accurate for the "crack" class, displaying lower accuracy in detection and classification.

This research stands to benefit local government departments and enhance public safety by enabling timely and reliable identification of road surface defects, subsequently reducing the time required for maintenance and repairs. This improvement could significantly enhance road safety, minimizing the likelihood of accidents due to undetected defects. Furthermore, by improving overall road conditions, the research enhances the safety and comfort of drivers and passengers, potentially reducing vehicle damage and associated repair costs for the public.

References

1. ^*Ministry of Transport Malaysia. (2022). MALAYSIA ROAD SAFETY PLAN 2022-2030 MINISTRY OF TRANSPORT MALAYSIA. https://www.mot.gov.my/en/Pages/Land/Safety%20and%20Security/MRSP%202022-2030%20(1022).pdf.*
2. ^*Tarrence Tan, Ragananthini Vethasalam and Gerard Gimino. (2024). Road fatality statistics to be released daily, says Loke. The Star. Retrieved July 25, 2024, from https://www.thestar.com.my/news/nation/2024/03/06/road-fatality-statistics-to-be-released-daily-says-loke.*
3. ^*Gerard Gimino, Rahimy Rahim, Ragananthini Vethasalam. (2023). 915,874 road accidents recorded throughout 2021 and 2022, says Transport Ministry. The Star. Retrieved July 25, 2024, from https://www.thestar.com.my/news/nation/2023/06/13/915874-road-accidents-recorded-in-2021-and-2022-says-transport-ministry.*
4. [a, b]*Rolt J, Otto A, Endale AA. (2020). "Common surface defects in asphalt road pavements and unexpected causes". Proceedings of the Institution of Civil Engineers - Transport. 173 (1): 13–20. doi:10.1680/jtran.17.00088.*

5. [a, b] *Emara K, El-Kady A, Shaaban E, ElEliemy M. (2021). "Mobile Crowdsensing framework for road surface quality detection". International Journal of Intelligent Computing and Information Sciences. 21 (3): 95–106. doi:10.21608/ijicis.2021.91569.1119.*
6. [^] *Zoumana Keita. (2022, September). YOLO Object Detection Explained: A Beginner's Guide. Www.datacamp.com. https://www.datacamp.com/blog/yolo-object-detection-explained.*

Declarations

Funding: No specific funding was received for this work.

Potential competing interests: No potential competing interests to declare.