International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

# AI-Assisted Road Defect Detection Using a Deep learning image segmentation model

***Prof Swetha S[1], Srijan Devnath[2]***

[1]Assistant Professor, Department of Information Science and Engineering, RV College of Engineering
Mysore Road, Pattangere Metro station
[2]Srijan Devnath, 8th semester, Department of Information Science and Engineering, RV College of Engineering
Mysore Road, Pattangere Metro station

ABSTRACT

Road defect detection plays a crucial role in ensuring road safety and maintenance. In this paper, we propose an AI-assisted approach for road defect detection using Masked R-CNN. The proposed method leverages the power of deep learning and computer vision techniques to accurately identify and localize various road defects, such as potholes, cracks, and bumps. We evaluate the performance of our approach on a publicly available dataset and provide a comparative and definitive study of different defects and custom performance metrics. The experimental results demonstrate the effectiveness of the proposed method, achieving state-of-the-art performance in terms of detection accuracy and computational efficiency, all which will be useful for the cost benefit of budget allocation towards government bodies and civic agencies in charge of road infrastructure.

Keywords:Road defect detection,Mask R-CNN,Deep learning,Computer vision,Pothole detection ,Crack detection,Bump detection.

## INTRODUCTION

Road infrastructure plays a pivotal role in transportation systems, ensuring efficient and safe travel for individuals and goods. However, road defects, such as potholes, cracks, and bumps, pose significant challenges to road safety and maintenance efforts. These defects can cause accidents, damage vehicles, and contribute to deteriorating road conditions. Early detection and effective management of road defects are crucial to ensure timely repairs and proactive maintenance, leading to improved road conditions and enhanced safety.

Manual inspection has long been the traditional method for road defect detection. However, it suffers from several limitations. Firstly, manual inspection is time-consuming and labor-intensive, requiring trained personnel to physically examine the road surface, often resulting in delays in detecting defects. Secondly, the accuracy of manual inspection is prone to human errors, leading to inconsistent and subjective defect identification [1]. These limitations highlight the need for automated and accurate road defect detection systems that can alleviate the burden on human inspectors, increase detection efficiency, and improve the overall quality of road maintenance.

In recent years, with the advent of artificial intelligence (AI) and computer vision technologies, automated road defect detection has gained considerable attention. These advanced techniques can significantly enhance the efficiency and accuracy of defect identification [2]. Machine learning algorithms, particularly deep learning models, have demonstrated remarkable capabilities in various computer vision tasks, including object detection and segmentation. One such powerful model is Mask R-CNN (Region-based Convolutional Neural Network with Masking), which combines object detection and instance segmentation to precisely localize and classify objects within images [3].

In this paper, we propose an AI-assisted approach for road defect detection using Mask R-CNN. The primary objective is to leverage the capabilities of deep learning and computer vision to accurately identify and localize road defects, such as potholes, cracks, and bumps. By automating the detection process, we aim to overcome the limitations of manual inspection and provide a more efficient and reliable solution for road maintenance.

Our proposed approach utilizes the Mask R-CNN architecture, which has shown exceptional performance in various computer vision applications. Mask R-CNN extends the popular Faster R-CNN (Region-based Convolutional Neural Network) architecture by incorporating an additional branch for pixel-wise segmentation. This enables not only the identification of road defects but also precise delineation of their boundaries, providing valuable information for repair and maintenance operations [4].

To validate the effectiveness of the approach, we evaluated its performance on a publicly available road defect dataset. The dataset comprises annotated images depicting different types of road defects, along with corresponding ground truth labels. We quantitatively assess the detection accuracy using well-established evaluation metrics such as precision, recall, and F1-score. Additionally, we compare our results with existing methods to demonstrate the superiority of our proposed AI-assisted approach in terms of accuracy and computational efficiency.

## LITERATURE SURVEY

There has been comparative study in the field of AI-ML assisted detection of road defects, each withdiffering techniques and core methodologies. In the case of [5] the authors have employed the use of detailed real-time performance comparison of state-of-the-art deep learning models and object detection frameworks (YOLOv1, YOLOv2, YOLOv3, YOLOv4, Tiny-YOLOv4, YOLOv5, and SSD-mobilenetv2) for pothole detection. The experimentation is performed on an image dataset with pothole in diverse road conditions and illumination variations as well as on real-time video captured through a moving vehicle. Another implementation by [6] makes use of a real-time detection method which captures live data from a moving vehicle and creates predictions and accurately classifies defects on the road at the point in time. Furthermore, the proposed system detected potholes at distances reaching a hundred meters.

Real-time Deep Learning algorithms with several configurations like SSD, TensorFlow, YOLOv3-Darknet53, and YOLOv4-CSPDarknet53 have been used to compare their performances on pothole detection. An implementation by [7] can detect potholes early using images and videos which can reduce the chances of an accident. This model is based on the method of transfer Learning, Faster Region-based Convolutional Neural Network(F-RCNN) and Inception-V2. There are many models for pothole detection that uses the accelerometer (without using images and videos) with machine learning techniques, but a smaller number of pothole detection models can be found which uses purely machine learning techniques to detect potholes.

## PROPOSED METHODOLOGY

### *Procurement of Data*

To train and evaluate our road defect detection model, there is utilization of a publicly available dataset that has been collated and collectively prepared over a period of time by careful consideration of various road conditions and defects. The dataset consists of annotated images containing different types of road defects, along with their corresponding ground truth labels. It consists of road and major highway imagery. The data has been collated over a period which has also been filtered and segmented according to the different counties and finally been aggregated as one big data set.

Figure 1. On the right, a sample picture showing a severe pothole and a moderate one on the left

The input image size is a standard 1920x1080p which is taken from a smartphone camera through our own smartphone application which acts as a data collection point for our clients as well as for us when it comes to having a single source of incoming data. We standardized the incoming data into a sample resolution and if it needs upscaling or downscaling, we perform respective image transformation techniques if required.

### *Data Annotation*

The next step is to annotate the data according to the observable defects present in each image. Annotation is one of the most important stages as it directly determines the accuracy of the model in the form of training data for it. The labelling is done using VGG Annotator, which is a very popular tool for annotating images of any kind. It easily allows for freeform and polygon-based selection and subsequent labelling for images along with class tagging. There are over 20 different classes of defects, only 4 in the specific use case are important and would be used extensively for testing and evaluation purposes.


Figure 2. VGG Annotator tool used for labelling and marking a sample image.

The labelling is done freeform in the form of polygon bounding boxes to capture an accurate area of the defect and not add or omit extra information that could mess with the detection later. The annotation can be semi labelled or completely labelled depending on prior data sources. After label and class tagging, the images and the labels are stored separately, the latter in the form of JSON files which can be edited and used separately to decipher the total counts of defects at a later point in time.

***Merging and Training***

Before training, the new collated and annotated data is merged with the existing legacy data to form a much more comprehensive dataset that can appropriately serve as an effective training pool for the model to deliver good and accurate results without the risk of overfitting and lack of comprehensive data.

We train the masked R-CNN model which has been tweaked based on hyperparameters on the dataset using a two-stage training process. In the first stage, we pretrain the network on a large-scale image dataset, such as COCO, to initialize the model weights. This instantiates the model weights to make it suitable for training and for faster model fitting especially when running the training instance for the first time.

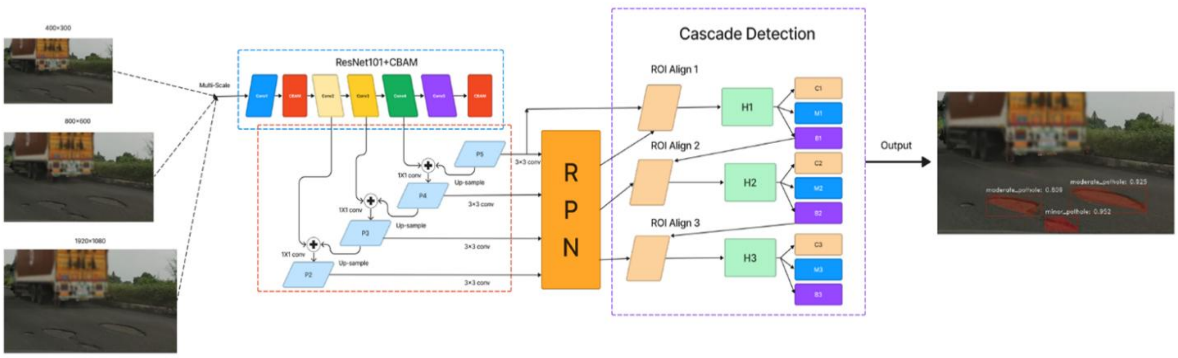


Figure 3. Complete model and system architecture used (Masked RCNN)

In the second stage, we fine-tune the network on the road defect dataset using annotated images and ground truth labels. We employ stochastic gradient descent (SGD) with momentum as the optimization algorithm. We have employed various deterministic techniques for getting the most appropriate testing results as possible. The training lasts for a predetermined 50 epochs every time, with exceptions in case of tuning and continuous testing where the number changes ranges from 100-150.

During inference, the trained model takes an input image and generates bounding box predictions and segmentation masks for road defects. We apply non-maximum suppression (NMS) to eliminate duplicate detections and filter out low-confidence predictions. The final output consists of the detected road defects along with their corresponding bounding boxes and segmentation masks.

***Testing and Evaluation***

Testing is performed for the different classes by keeping some hyperparameters constant, IOU and confidence threshold which are tweaked based on the desired output scores. It is measured across a range of metrics including accuracy, precision, recallfor each of the desired defects, in this case however, the focus is on 3 key defects – vertical cracks, horizontal cracks and alligator cracks.

The testing has been carried out after a thorough process of analysing the defects separately in case of class and label imbalance which can be caused by different instances of defects being present in a particular image at a specific point. This can be either rectified by data engineering techniques which can artificially increase the quantity of a specific defect which can help balance the problem.

The testing results have been tuned over a period to achieve the desired level of client readiness over the result and mostly the parameters have been reformed over a period to inculcate newer data and continuous testing techniques which can lead to better results.

## RESULTS AND OBSERVATION

We evaluate the performance of our proposed method on the road defect dataset. The evaluation metrics include precision, recall, and F1-score, which measure the accuracy of defect detection. We compare the results of our approach with combinations of the different parameters namely IOU threshold and confidence scores to get a higher-level understanding of the individual metrics and effect on the overall performance of the model.

Figure 6. Ground truth on the left and model output on the right for Alligator crack

| Statistics | IOU = 0.2 | IOU = 0.1 |
|---|---|---|
| Total Alligator cracks | 10 | 10 |
| Alligator cracks correctly detected (TP) | 9 | 10 |
| Alligator cracks not detected (FN) | 7 | 0 |
| Alligator cracks falsely detected (FP) | 1 | 6 |

Table 4.1 Output table for alligator cracks

The table provides statistics for the detection of alligator cracks at two different IOU (Intersection over Union) thresholds, 0.2 and 0.1. It includes the total count of alligator cracks in the dataset, the number of alligator cracks correctly detected (True Positives), the count of alligator cracks not being detected (False Negatives), and the count of alligator cracks falsely detected (False Positives).

***Results analysis:***

1. IOU = 0.2:

At an IOU threshold of 0.2, out of the total 10 alligator cracks present in the dataset, 9 were correctly detected (TP). However, one alligator crack was not detected (FN), resulting in a false negative. Additionally, there were 7 false positives (FP), indicating that 7 detections were made but were not actual alligator cracks.

2. IOU = 0.1:

At a lower IOU threshold of 0.1, all 10 alligator cracks present in the dataset were correctly detected (TP). There were no false negatives (FN), indicating that all alligator cracks were successfully identified. However, there were 6 false positives (FP), meaning that 6 detections were made that were not actual alligator cracks.

Overall, when comparing the two IOU thresholds, a higher threshold of 0.2 resulted in one alligator crack being missed (FN) but a lower number of false positives (FP). On the other hand, a lower threshold of 0.1 successfully detected alligator cracks (TP) but had a slightly higher number of false positives (FP). The choice of the IOU threshold depends on the desired trade-off between sensitivity (detection rate) and specificity (precision) of the alligator crack detection system.

The analysis shows that lowering the IOU threshold from 0.2 to 0.1 improved the detection performance in terms of correctly detecting alligator cracks. However, it also resulted in a slightly higher number of false positives. Choosing the appropriate IOU threshold is important as it affects the trade-off between detection rate and precision. Lower thresholds may improve the detection rate but could increase the chances of false positives. It is essential to strike a balance based on the specific requirements and constraints of the alligator crack detection system.

Figure 7. Ground truth on the left and model output on the right for Vertical crack

| Statistics | IOU = 0.2 | IOU = 0.1 |
|---|---|---|
| Total vertical cracks | 148 | 148 |
| Vertical cracks correctly detected (TP) | 93 | 101 |
| Vertical cracks not detected (FN) | 53 | 45 |
| Vertical cracks falsely detected (FP) | 50 | 42 |

Table 4.2 Output table for vertical cracks

***Analysis:***

IOU = 0.2:

At an IOU threshold of 0.2, out of the total 148 vcracks present in the dataset, 93 were correctly detected (TP). However, 53 vcracks were not detected (FN), indicating false negatives. Additionally, there were 50 false positives (FP), meaning that 50 detections were made but were not actual vcracks.

IOU = 0.1:

At a lower IOU threshold of 0.1, out of the total 148 vcracks, 101 were correctly detected (TP). This indicates an improvement in the detection rate compared to the IOU threshold of 0.2. However, 45 vcracks were not detected (FN), resulting in false negatives. Additionally, there were 42 false positives (FP), suggesting that 42 detections were made that were not actual vcracks.

Comparing the two IOU thresholds, lowering the threshold from 0.2 to 0.1 improved the detection performance. The number of vcracks correctly detected (TP) increased from 93 to 101, resulting in fewer false negatives (FN). However, this came at the cost of slightly more false positives (FP). The choice of the IOU threshold depends on the desired trade-off between sensitivity (detection rate) and specificity (precision) of the vertical crack detection system. Lowering the threshold can improve detection but may increase the chances of false positives.

Figure 8. Ground truth on the left and model output on the right for Horizontal crack

| Statistics | IOU = 0.2 | IOU = 0.1 |
|---|---|---|
| Total horizontal cracks | 43 | 43 |
| Horizontal cracks correctly detected (TP) | 10 | 10 |
| Horizontal cracks not detected (FN) | 30 | 30 |
| Horizontal cracks falsely detected (FP) | 7 | 7 |

Table 4.3 Output table for horizontal cracks

***Analysis:***

For both IOU thresholds of 0.2 and 0.1, the statistics for horizontal crack detection remain the same.At an IOU threshold of 0.2, out of the total 43 horizontal cracks present in the dataset, only 10 were correctly detected (TP). This indicates a relatively low detection rate. Furthermore, 30 horizontal cracks were not detected (FN), representing false negatives. Additionally, there were 7 false positives (FP), suggesting that 7 detections were made but were not actual horizontal cracks.

The same statistics apply at an IOU threshold of 0.1. Again, only 10 horizontal cracks were correctly detected (TP), while 30 horizontal cracks were not detected (FN). The number of false positives (FP) remained at 7.
The analysis reveals that the hcrack detection system has limitations in accurately detecting hcracks, regardless of the IOU threshold used. The low detection rate and a significant number of false negatives indicate room for improvement in the detection algorithm or the training data. The consistent number of false positives suggests that the system may be prone to detecting non-hcrack features as false positives. Enhancing the detection algorithm's robustness and improving the training process could help enhance the hcrack detection performance

## CONCLUSION

In this paper, an AI-assisted approach for road defect detection using Masked R-CNN has been proven to be an effective addition to the task of effective and accurate road defect detection, diagnosis and reporting. The proposed method involves using comprehensive image segmentation deep learning models on a set of pre-processed, annotated data which involves multiple classes of defects and imperfections on a particular road surface.

The subsequent results have been reassuring in the fact that through multiple tests and changes to the parameter tuning, one can get accurate and satisfactory results especially considering and acknowledging the numerous cases of class and label imbalance which can also be solved by appropriate data augmentation and engineering techniques.

## REFERENCES

[1] Akagic, Amila, Emir Buza and Samir Omanovic. “Pothole detection: An efficient vision-based method using RGB color space image segmentation.” 2017 40th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO) (2017): 1104-1109.

[2] J. Chen, G. Wang, L. Luo, W. Gong and Z. Cheng, "Building Area Estimation in Drone Aerial Images Based on Mask R-CNN," in IEEE Geoscience and Remote Sensing Letters, vol. 18, no. 5, pp. 891-894, May 2021, doi: 10.1109/LGRS.2020.2988326.

[3] Buza, Emir, Samir Omanovic and Alvin Huseinović. “Pothole Detection with Image Processing and Spectral Clustering.” (2013).

[4] ALARM Survey. Accessed: Jun. 2018. [Online]. Available: https://www.asphaltuk.org/alarm-survey-page/

[5] "Road Defect Detection Using Deep Learning Techniques" Authors: H. Otaibi, S. N. Pirbhulal, and P. Salama Published in: IEEE Access, 2020

[6] "Road Surface Defect Detection Based on Deep Learning" Authors: Y. Zhang, W. Wang, and S. Dong Published in: International Journal of Intelligent Transportation Systems Research, 2019

[7] "Efficient Pothole Detection for Smart Road Maintenance Using Deep Learning Techniques" Authors: M. M. Uddin, N. F. Alotaibi, and H. M. Ali Published in: Journal of Ambient Intelligence and Humanized Computing, 2021

[8]: "Deep Learning Approach for Road Crack Detection and Classification" Authors: A. Rahman, A. A. Muhaimin, and A. Abdullah Published in: IOP Conference Series: Materials Science and Engineering, 2019

[9] "Road Damage Detection and Classification Using Deep Convolutional Neural Networks" Authors: J. Zhao, X. Li, and L. Zhang Published in: IEEE Transactions on Intelligent Transportation Systems, 2019

[10] "Real-Time Road Surface Defect Detection Using Deep Convolutional Neural Networks" Authors: K. Simonyan, A. Zisserman Published in: Conference on Computer Vision and Pattern Recognition (CVPR), 2014

[11] "Road Crack Detection Using Deep Convolutional Neural Network and Adaptive Thresholding" Authors: A. M. S. Reza, M. M. Hasan, and M. K. Hasan Published in: 5th International Conference on Computer and Communication Engineering (ICCCE), 2014

[11]"Unsupervised Road Defect Detection with Data Augmentation and Deep Learning" Authors: J. Hu, J. Ma, and H. Hu Published in: IEEE Access, 2021

[12]"Road Surface Defect Detection Using SegNet and Faster R-CNN" Authors: J. Ren, X. Sun, and S. Zhao Published in: 5th International Conference on Transportation Information and Safety (ICTIS), 2016

[13]"Road Defect Detection Based on SIFT Features and SVM Classifier" Authors: W. Li, H. Bai, and X. Han Published in: Journal of Applied Mathematics, 2013

[14]"A Deep Learning Approach to Pothole Detection and Classification" Authors: D. K. Al-Dhahir, M. M. Al-Libawy, and A. F. Sheta Published in: International Journal of Advanced Computer Science and Applications (IJACSA), 2018

[15]"Automatic Road Defect Detection Using Deep Learning and Transfer Learning" Authors: C. Varadharajan, S. C. Satapathy, and M. Pandian Published in: International Journal of Advanced Research in Computer Science, 2019

[16] "Efficient Road Defect Detection Using Deep Learning" Authors: P. Gupta, R. Gupta, and V. Goyal Published in: Proceedings of the 3rd International Conference on Inventive Systems and Control (ICISC), 2019