The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLVIII-G-2025
ISPRS Geospatial Week 2025 “Photogrammetry & Remote Sensing for a Better Tomorrow…”, 6–11 April 2025, Dubai, UAE

A Simplified Gaussian Approach for Asphalt Crack Detection based on Deep Learning and RGB Images

Walid Darwish[1,2] *, Wael Ahmed[1,2]

[1] Public Works Department, Faculty of Engineering, Cairo University,
Giza, Egypt, 12613.(walid.darwish & wael.mohamed) @cu.edu.eg
[2] NAMAA for Engineering Consultations, Dokki , Giza, Egypt

Keywords: Surface Pavement, Cracks, Road Inspection, YOLO, Gaussian, Camera Calibration.

Abstract

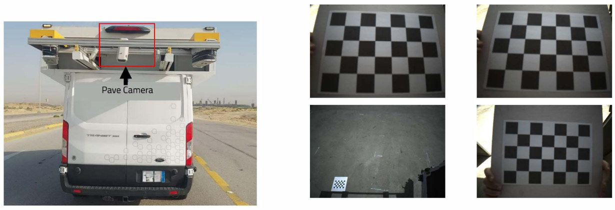
Monitoring pavement condition is a crucial aspect for pavement maintenance management systems (PMMS). There are several pavement characteristics that affect the pavement condition, Crack distress is a highly representative type of pavement distress and often serves as an early indicator of more extensive pavement issues. Cracks impact both the operational efficiency and safety of road pavements and significantly influence maintenance decisions. We propose a workflow to detect cracks using YOLOv9 deep learning algorithm combined with statistical analysis through principal component (PCA) and Gaussian distribution. The proposed workflow includes camera calibration to address the metric issues in vision-based crack detection methods, utilizing Zhang's calibration method to compute the camera's internal and external parameters. To validate the proposed framework, three different datasets were acquired. Laser Crack Measurement System (LCMS) was used as a ground truth data for further verification the proposed method. Experimental results demonstrate that the proposed method achieves millimeter-level accuracy (std = $\pm 1.0mm$) compared to LCMS. This indicates the method's potential applicability for asphalt road crack segmentation and crack width estimation.

1. Introduction

Roadway infrastructure plays a vital role in a country's economy (Qureshi et al., 2023). The condition of pavement is a fundamental infrastructure component that significantly impacts national development. Regular inspections of pavement surfaces are necessary due to traffic loads and climate conditions (Qureshi et al., 2023; Azam et al., 2023). Recently, pavement condition monitoring has been enhanced by using optical images and laser scanners (Al-Sabaeei et al., 2024).

Functional and structural failures are the main two types of pavement failure. While functional failure is caused by surface distresses, structural failure involves the collapse or breaking of pavement layers (Hasan et al., 2024). There are several pavement characteristics that affect the pavement condition index such as roughness, surface condition, surface skid resistance, and pavement strength. Surface condition and its distresses play a significant role in pavement assessment. Surface distresses are divided to cracks (longitudinal, transverse, alligator, block, and edge cracks), surface openings (patches, potholes), surface deformation (rutting, depression), and surface defects (ravelling, bleeding) (Lv et al., 2023; Qureshi et al., 2022).

Crack distress is a highly indicative type of pavement distress and often serves as an early warning of more extensive pavement issues. If not repaired promptly, cracks can evolve into more severe forms of pavement deterioration (Lv et al., 2023). Recently, deep learning techniques have been widely implemented for rating pavement condition directly from images and to determine cracks and assessing their severity. A deep learning framework was proposed for intelligent pavement condition rating based on the Pavement Surface Condition Index (PSCI) scale, with images labeled from 1 to 10 by experts for Irish roads (Qureshi et al., 2022). The dataset was cleaned to eliminate images exhibiting motion blur, insufficient lighting, and focus blur.



(a) MFV with pave camera installed at the back

(b) Sample of calibration images for Pave camera

Figure 1. System configuration and calibration data

Guo et al. (2022) introduced Boundary Aware Refinement Network for Crack Detection (BARNet) network to obtain a robust crack detection from images. It consists of three modules; Base predictor, Edge adaptation, and Refinement modules. This deep learning approach was evaluated on five different datasets. Ji et al. (2020) proposed an approach for crack detection using DeepLabv3+ in asphalt pavement at the pixel level. Five aspects of the crack were employed using the crack quantification algorithm. These are: crack length, mean width, maximum width, area, and ratio.

Majidifard et al. (2020) deployed YOLO deep learning model to classify distresses and developed a U-net based model to quantify their severity using Google Street View images of pavements. Each pavement image underwent rating using a comprehensive pavement condition assessment system by employing both linear regression and weighted regressors. In this system, only two distresses are used to compute the rating: cracks and potholes. Evidence showed that when evaluating segmentation algorithms using deep learning models, the performance is robust when test and training images come from the same datasets collected by the same device. However, performance degrades

* Corresponding author

This contribution has been peer-reviewed.
https://doi.org/10.5194/isprs-archives-XLVIII-G-2025-345-2025 | © Author(s) 2025. CC BY 4.0 License.

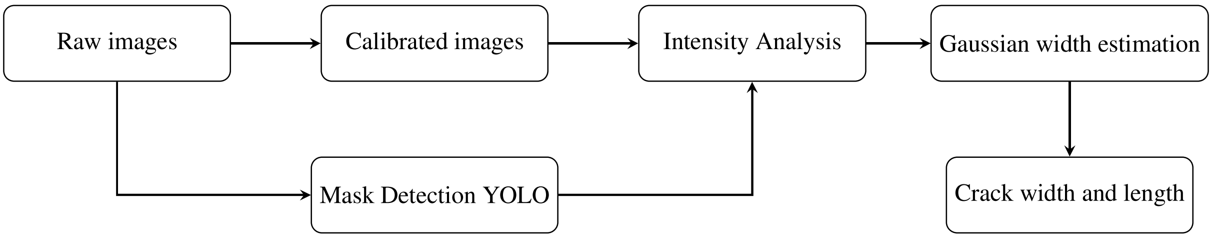


Figure 2. Proposed workflow

when datasets acquired by different devices are used (Qureshi et al., 2023).

Visual and geometry-based methods for crack detection utilize image processing techniques that focus on the photometric and spatial characteristics of cracks compared to their surrounding areas. These methods leverage variations in light intensity and geometric features to accurately identify and analyze the presence of cracks (Quan et al., 2019). Kapela et al. (2015) proposed an approach based on RGB images and Histograms of Oriented Gradients (HOG) (Dalal and Triggs, 2005) to detect road cracks, which is prone to noise. The RGB images were captured from a camera mounted on top of a car and faced-down orthogonal to the pavement. The preprocessing steps included doubling the gray-scale image, applying a blur filter, and using a contact stitching algorithm. The images were then processed through the Histogram of Oriented Gradients (HOG) for crack detection and to prepare the final results.

Liu et al. (2018) proposed a workflow for the detection of approximate area that contains surface defects utilizing the Gradient Local Binary Pattern (GLBP) technique. The developed approach employs weighted binary output values across eight neighborhood directions to represent local variations in grayscale. Quan et al. (2019) introduced a workflow for surface cracks detection by the improved Otsu threshold (Otsu et al., 1975). Median filter was applied to gray scale image then followed by Otsu threshold segmentation for feature extraction and detection. Mubashshira et al. (2020) proposed an unsupervised approach for detecting road surface cracks by applying color histogram analysis. Segmentation was involved in the results of using k-means clustering for crack detection in 2D image.

In the proposed workflow, cracks were detected from optical images calibrated for actual measurements of crack width and length. Deep learning YOLO was utilized for a rough estimation of boxes containing cracks, followed by spatial and color analysis using histograms for improved detection. Contributions to the proposed approach for automatic crack detection can be summarized as follows:

- Camera calibration approach for accurate measurement from photos
- Introducing a hybrid approach consists of Deep learning that apply YOLOV9 followed by local statistical analysis of calibrated optical image for crack width analysis.
- Compare the measurement of pavement cracks from calibrated image to results from laser crack detection system.

2. Crack detection workflow

The proposed workflow (as shown in Figure 2) consists of four main steps: Camera calibration, Mask Detection, Intensity Analysis, Crack detection and extraction. The following subsection will discuss each step in details.

2.1 Camera Calibration

Figure 1a illustrates the proposed cracking detection system that consist of one camera installed on the top of Multi Functional Vehicle (MFV). In order to overcome the metric problem of the vision based crack detection methods, camera internal and external calibration parameters were computed. Assuming that the camera height is constant for all taken images and the intrinsic parameters are known, relationship between pixel and metric domains could be determined. Considering the pinhole camera model, Zhang's calibration method to estimate camera parameter (Zhang, 2000) was applied. Sample of the collected calibration data is shown in Figure 1b. Table 1 shows the calibration parameters of the pave camera that can be used to construct orthogonal images with the called (Bird's Eye images). For this research purpose, Bird's Eye Image will be called Calibrated images. The calibration process includes lens distortion removal and orthogonal projection of the captured images.

2.2 Mask Detection

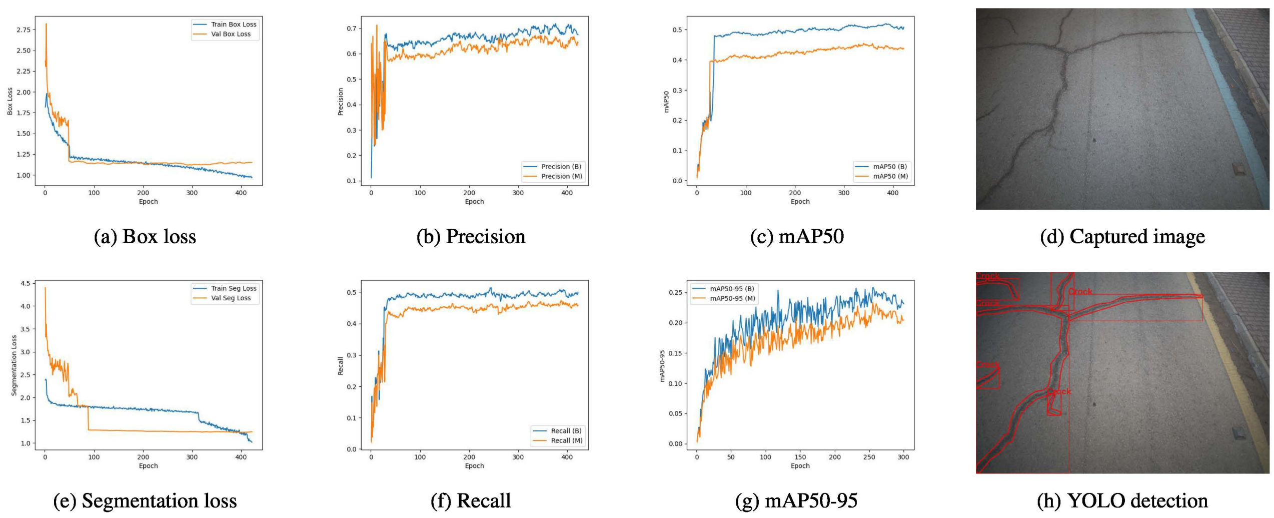
The images are analysed for coarse detection of boxes that have cracks using YOLOv9 deep learning architecture. The network has three main cores, which are: Backbone, which is responsible for feature extraction and representation; Auxiliary core, which enhances the feature extraction and performs multi-scale feature extraction; Neck, which is responsible for feature aggregation and integrates different resolution features; and finally Head core, which gives the final result if needed (Imran et al., 2024).

YOLOv9 was trained on a combination of three different datasets as listed below. The whole dataset was split to 5200 images for training and 780 images were retained for validation purpose.

- **In-House Dataset:** The dataset was collected by our vehicles and labeled by our Pavement Management System (PMS) Team. This dataset includes high-resolution images of road surfaces, annotated with various types of pavement distresses and other relevant features.

- **Edmcrack600:** The dataset consists of 600 images with detailed annotations for crack detection, providing a diverse range of examples to improve the model's generalization capabilities (Mei et al., 2020).

- **DeepCrack:** This dataset further enhances the diversity and richness of the training data, contributing to improved model performance on crack detection tasks (Liu et al., 2019)



(a) Box loss (b) Precision (c) mAP50 (d) Captured image

(e) Segmentation loss (f) Recall (g) mAP50-95 (h) YOLO detection

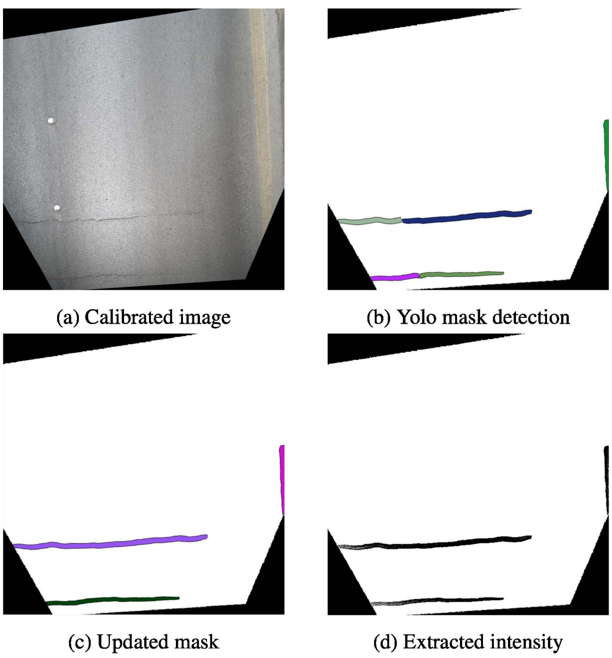
Figure 3. Performance analysis of yolov9e-seg results

Table 1. Pave camera calibration parameters focal lengths and principal point in pixels, distortion parameters are dimensionless, and Height of the camera (H) is in meters

Parameter	H	F_x	F_y	C_x	C_y	k_1	k_2	k_3	p_1	p_2
Value	2.237	1787.75	1786.58	1640.34	1092.44	-0.0186	-0.0372	0.0157	-0.0014	-0.0011

The training process for the YOLOv9e-seg model was developed by Ultralytics. It starts with initializing the model using pre-trained weights from the COCO dataset that provides a robust starting point for fine-tuning. The optimization was carried out using the Adam optimizer. A composite loss function, comprising classification loss, localization loss, and segmentation loss, were employed to train the model effectively. The model's performance was evaluated using metrics such as mean Average Precision (mAP) for object detection and segmentation accuracy. The following Figure 3 shows the performance results for mask and box predictions.

To enhance the accuracy of segmentation masks during inference, we employed a patch-based approach. The inference process involves dividing each input image into sixteen patches, with each patch covering 25% of the original image and a stride of 0.125 of the original image in both the X and Y axes. This strategy effectively enlarges the objects within smaller patches, allowing the model to produce more accurate segmentation masks. Once the individual patches are processed, they are combined to reconstruct the original image, resulting in more precise and accurate object masks. Figure 3h shows the results of segmentation on the captured RGB image.



(a) Calibrated image (b) Yolo mask detection

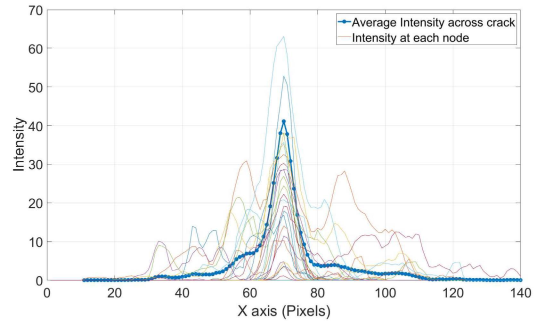
(c) Updated mask (d) Extracted intensity

Figure 4. Processing of optical images

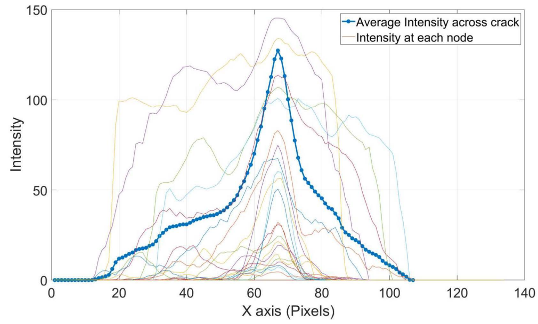
2.3 Intensity Analysis

As a results of the proposed inference method, YOLO can introduce errors by detecting the same crack in multiple segments. To address this issue, an analysis of the detected masks is performed using a covariance matrix for alignment, and nearest neighbor distance. The procedure involves determining the alignment of a single mask using the covariance matrix. If two different masks are found to have nearest points within a threshold difference in alignment and distance, these masks are merged.

The global application of photo enhancement for crack detection can sometimes yield negative effects due to variations in light and shadow. To address this issue, a process is developed to begin with creating a boundary mask for each crack and the optical photo, which is then converted to grayscale. Recognizing that cracks typically exhibit lower intensity compared to other pavement features, a median thresholding method is employed to filter out pixel values unrelated to cracks. Unlike other methods that convert this step into a binary analysis, this presented approach maintains intensity levels and excludes points above the median, effectively reducing non-crack points and enhancing crack detection accuracy.



(a) Intensity value across the crack and Average intensity value across the crack of an estimated crack width of 5.8 mm



(b) Intensity value across the crack and Average intensity value across the crack of an estimated crack width of 8.4 mm

Figure 5. Results for crack width estimation

2.4 Crack Detection and Extraction

Each crack’s intensity mask points are processed individually using a histogram of pixel values. Depending on its orientation, the crack mask is rotated to align with vertical (Y) axis along the image. At each specific interval, the histogram was computed and then an aggregated histogram for all section is used to estimate crack width. The crack width is estimated to be the standard deviation of the Gaussian shape. Figure 5 shows the Gaussian shape of the accumulated intensity.

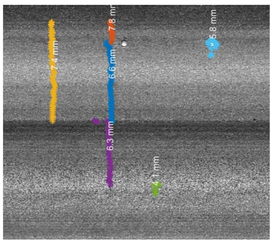
3. RESULTS and DISCUSSION

To test the proposed approach, images of three different roads were acquired using a Multi-Function Vehicle (MFV). A Laser Cracking Measurement System (LCMS) was mounted on the MFV to detect cracks with sub-millimeter accuracy. Since a crack may be detected in several segments of varying widths, the average crack width was computed from the LCMS data and then compared to the crack width obtained using the proposed approach.

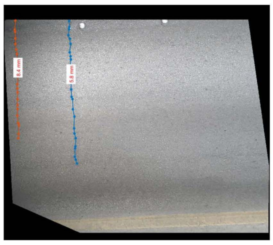
The proposed system uses one camera inclined to the pavement, which has a coverage area nearly 5×5 meters, while the LCMS data was collected for an area of 4×10 meters. To mitigate this problem, we aligned the calibrated data to be in the direction of the lane. It worth to note that, the detection of the crack masks is applied to the original images without calibration and then the masks are mapped to the calibrated image. Figure 6 shows the detected crack width compared to the LCMS output results, it can be notice that the proposed method can achieve comparable accuracy to the LCMS results. Table 2 shows a quantitative results for more than 70 images which contains cracks. The proposed system achieved an accuracy of $\pm 1.0mm$.

Table 2. Statistics for 82 detected cracks (dimensions are in mm)

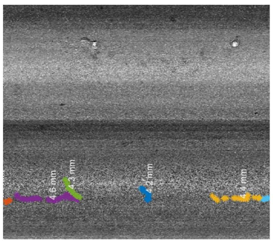
Parameter	LCMS Results	Our Results
Maximum width	13.7	14.8
Minimum width	4.2	3.6
Mean absolute error	–	1.2
Standard deviation	–	1.0



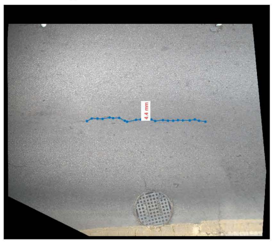
(a) LCMS data for road section 1



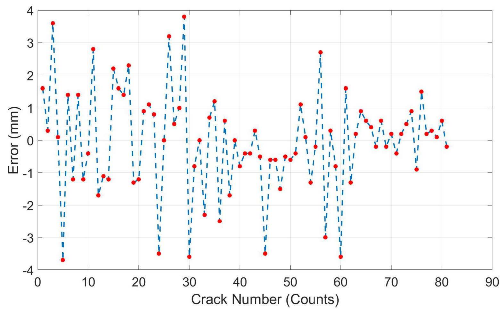
(b) Proposed method (road 1)



(c) LCMS data for road section 2



(d) Proposed method (road 2)



(e) Error of estimated crack width for total 82 cracks

Figure 6. Processed data for crack detection, on the left is LCMS output, while on the right is Our proposed method

4. Conclusion

In this research, a novel method for crack detection is proposed. Unlike existing methods, the proposed approach estimates crack width in metric units (e.g., millimeters) through a calibration process using the Pave camera. The Gaussian model is used to estimate the bell shape resulting from the inverted intensity value around the crack line, which is identified by a mask detected using YOLOv9. This approach was applied to three different real datasets, each containing data from a Laser Cracking Measurement System (LCMS). The results demonstrate that the proposed method achieves millimeter-level accuracy compared to LCMS results, indicating its potential applicability for asphalt road crack segmentation and crack width estimation.

5. Acknowledgements

Authors would like to extend their sincere gratitude to PMMS team at NAMAA for Engineering Consultations for their invaluable support and contributions to this research.

References

Al-Sabaeei, A. M., Souliman, M. I., Jagadeesh, A., 2024. Smartphone applications for pavement condition monitoring: A review.

Azam, A., Alshehri, A. H., Alharthai, M., El-Banna, M. M., Yosri, A. M., Beshr, A. A. A., 2023. Applications of Terrestrial Laser Scanner in Detecting Pavement Surface Defects. *Processes*, 11, 1370. https://www.mdpi.com/2227-9717/11/5/1370.

Dalal, N., Triggs, B., 2005. Histograms of oriented gradients for human detection. *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)*, 1, Ieee, 886–893.

Guo, J.-M., Markoni, H., Lee, J.-D., 2022. BARNet: Boundary Aware Refinement Network for Crack Detection. *IEEE Transactions on Intelligent Transportation Systems*, 23(7), 7343-7358.

Hasan, A. S., Roksana, K., Kabir, S. F., Abadin, M. J., Uddin, M. N., 2024. Investigation of Typical Distresses of Flexible Pavements in Dhaka City and Possible Remedies. *International Journal of Pavement Research and Technology*.

Imran, A., Hulikal, M. S., Gardi, H. A., 2024. Real Time American Sign Language Detection Using Yolo-v9. *arXiv preprint arXiv:2407.17950*.

Ji, A., Xue, X., Wang, Y., Luo, X., Xue, W., 2020. An integrated approach to automatic pixel-level crack detection and quantification of asphalt pavement. *Automation in Construction*, 114, 103176. https://linkinghub.elsevier.com/retrieve/pii/S0926580519308350.

Kapela, R., Sniatala, P., Turkot, A., Rybarczyk, A., Pozarycki, A., Rydzewski, P., Wyczalek, M., Bloch, A., 2015. Asphalt surfaced pavement cracks detection based on histograms of oriented gradients. *2015 22nd International Conference Mixed Design of Integrated Circuits Systems (MIXDES)*, IEEE, 579–584.

Liu, X., Xue, F., Teng, L., 2018. Surface defect detection based on gradient lbp. *2018 IEEE 3rd International Conference on Image, Vision and Computing (ICIVC)*, IEEE, 133–137.

Liu, Y., Yao, J., Lu, X., Xie, R., Li, L., 2019. DeepCrack: A Deep Hierarchical Feature Learning Architecture for Crack Segmentation. *Neurocomputing*, 338, 139–153.

Lv, Z., Cheng, C., Lv, H., 2023. Automatic identification of pavement cracks in public roads using an optimized deep convolutional neural network model. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 381. https://royalsocietypublishing.org/doi/10.1098/rsta.2022.0169.

Majidifard, H., Adu-Gyamfi, Y., Buttlar, W. G., 2020. Deep machine learning approach to develop a new asphalt pavement condition index. *Construction and Building Materials*, 247, 118513. https://linkinghub.elsevier.com/retrieve/pii/S0950061820305183.

Mei, Q., Gül, M., Azim, M. R., 2020. Densely connected deep neural network considering connectivity of pixels for automatic crack detection. *Automation in Construction*, 110, 103018.

Mubashshira, S., Azam, M. M., Ahsan, S. M. M., 2020. An unsupervised approach for road surface crack detection. *2020 IEEE Region 10 Symposium (TENSYMP)*, IEEE, 1596–1599.

Otsu, N. et al., 1975. A threshold selection method from gray-level histograms. *Automatica*, 11(285-296), 23–27.

Quan, Y., Sun, J., Zhang, Y., Zhang, H., 2019. The method of the road surface crack detection by the improved otsu threshold. *2019 IEEE International Conference on Mechatronics and Automation (ICMA)*, IEEE, 1615–1620.

Qureshi, W. S., Hassan, S. I., McKeever, S., Power, D., Mulry, B., Feighan, K., O'Sullivan, D., 2022. An exploration of recent intelligent image analysis techniques for visual pavement surface condition assessment.

Qureshi, W. S., Power, D., Ullah, I., Mulry, B., Feighan, K., McKeever, S., O'Sullivan, D., 2023. Deep learning framework for intelligent pavement condition rating: A direct classification approach for regional and local roads. *Automation in Construction*, 153.

Zhang, Z., 2000. A flexible new technique for camera calibration. *IEEE Transactions on pattern analysis and machine intelligence*, 22(11), 1330–1334.