

Crack Detection using Pretrained Deep Learning under Varying UAV Imaging Configurations

Esmat, S. Z. S.,¹ Zainuddin, K.,^{1*} Rasib, A. W.,² Majid, Z.,² Ghazali, M. D.,¹ Marzukhi, F.,¹ Yusof, N. M.³ and Aziz, A. F. A.³

¹Faculty of Built Environment, Universiti Teknologi MARA, Perlis Branch, Perlis, Malaysia

²Faculty of Built Environment and Surveying (FABU), Universiti Teknologi Malaysia, Malaysia

³Innovation & Centre of Excellence, PLUS Berhad

*Corresponding Author

DOI: <https://doi.org/10.52939/ijg.v22i1.4711>

Abstract

Crack detection is critical for maintaining the structural integrity of infrastructure, particularly in large-scale applications. Recent advances in deep learning have demonstrated high potential for automating crack detection using unmanned aerial vehicle (UAV) imagery. However, most existing approaches are based on 2D image analysis, limiting spatial context and hindering accurate dimensional measurement. This study investigates the performance of a pretrained deep learning crack detection model embedded in Bentley Systems' iTwin Capture Modeler. UAV images were acquired at varying camera orientations and distances, and detection outputs from 2D images were projected onto a 3D mesh model generated via photogrammetry, enabling spatially referenced analysis. The UAV configurations included vertical, horizontal, and oblique (60°) flight angles. Among the tested scenarios, the slanting configuration at 40m yielded the highest F1 score of 92.28%. The findings highlight the critical influence of camera orientation and image resolution on detection accuracy and demonstrate the applicability of pretrained models in 3D digital twin workflows for structural inspection tasks.

Keywords: Crack detection, Deep Learning, iTwin Capture Modeler, Imaging Configuration, UAV Photogrammetry

1. Introduction

Crack detection plays a fundamental role in structural health monitoring, especially for critical infrastructure such as buildings and bridges. Cracks may develop due to material degradation, environmental exposure, or mechanical stress, and if not addressed, can threaten structural integrity and public safety. Traditional inspection methods, such as manual visual assessments, are labour-intensive, time-consuming, and prone to human error, rendering them inefficient for large-scale or inaccessible structures [1] and [2].

Numerous studies have demonstrated the growing effectiveness of deep learning approaches, particularly convolutional neural networks, in automated crack detection tasks. Pretrained models such as U-Net and ResNet have been applied in recent works to detect cracks on various material surfaces with promising levels of accuracy [3][4] and [5]. These models are often integrated with unmanned aerial vehicles (UAVs), allowing efficient processing of large volumes of high-resolution imagery. Despite their advantages, many of these

studies are based on controlled imaging conditions and do not fully consider real-world variations in camera configuration, viewing angles, or camera-to-object distances. As a result, although pretrained models perform well under standard conditions, their reliability when applied to variable UAV imaging scenarios remains insufficiently explored.

Even though deep learning-based crack detection has advanced considerably, most existing applications are still limited to two-dimensional (2D) image analysis. This restriction reduces the ability to localize cracks spatially or perform accurate dimensional measurements due to the absence of scale and depth information in images [6][7][8] and [9]. In addition to dimensional limitations, several imaging factors also affect model performance. These include image resolution, ground sampling distance, camera orientation, and flight altitude. For example, vertical images often fail to detect façade cracks due to limited visibility, while oblique or horizontal views may introduce geometric distortion and inconsistent image resolution [10][11] and [12].

To address the lack of spatial context in 2D crack detection, Bentley Systems introduced a pretrained deep learning model within its iTwin Capture Modeler software. The model works by detecting and tracing the cracks that appear in the oriented 2D images in the form of vector lines and then projects the vector on the generated three-dimensional (3D) model, which allows for spatial visualization and analysis [13]. This method facilitates accurate crack measurement, including length and width, and supports the integration of spatially referenced inspections [14] and [15]. However, the reliability of this pretrained model under different UAV imaging conditions has not been thoroughly investigated. The present study addresses this research gap by evaluating the model's performance across a range of camera orientations and distances within a 3D reconstruction framework.

2. Literature Review

2.1 UAV Photogrammetry for Structural Monitoring

UAV photogrammetry has been widely used in structural monitoring due to its ability to capture high-resolution images in areas that are difficult to access. It allows accurate observation and mapping of cracks by generating 3D models of structural surfaces [15] and [16].

One of the main advantages of UAV photogrammetry is its ability to support remote and virtual inspections. For example, [14] and [17] showed that integrating 3D models with automated crack detection algorithms can improve the efficiency and consistency of structural health monitoring. While some studies have focused on the quality of the 3D models produced, others have emphasized the benefits of automation in identifying surface defects. Together, these studies support the use of UAV photogrammetry as part of the digital inspection process. Rather than being used only for image capture, UAV photogrammetry is increasingly being positioned as a key component in a more comprehensive and objective inspection workflow. The integration of UAV imaging, 3D modelling, and crack analysis reflects a shift towards a more systematic approach to structural monitoring.

2.2 UAV-based Crack Detection with Deep Learning Models

Deep learning has become a widely used approach for detecting cracks in concrete structures, particularly because of its ability to perform detailed image segmentation. Several studies have used pretrained models such as U-Net, DeepLabv3+, YOLOv5, and TransUNet, which have proven effective in handling pixel-level tasks using UAV-captured images [7][16] and [18]. The deep learning

approach offers clear advantages in terms of accuracy and robustness compared with traditional image processing techniques. For example, pretrained convolutional neural networks (CNNs) consistently outperform methods such as edge detection and histogram-based segmentation in detecting concrete surface cracks [12][19] and [20]. Deep learning can also handle image noise and environmental variation more effectively than conventional techniques, helping to minimize false detections and improve overall detection accuracy [21]. Overall, previous studies have shown that deep learning provides a reliable and scalable solution for crack detection, particularly when applied to UAV imagery. Pretrained models with transfer learning not only reduce dependence on extensive annotated datasets but also enable automated analysis that is both faster and more consistent than manual inspection or classical image processing.

2.3 Impact of Camera Configuration on Crack Detection

The configuration of UAV cameras plays an important role in determining the quality and reliability of image-based crack detection. Factors such as flight altitude, viewing direction, and camera angles directly affect image resolution, surface visibility, and the ability to detect fine concrete surface defects [10]. Flight altitude is closely related to ground sampling distance (GSD), where the closer the camera is to the object, the better the GSD, allowing cracks to be detected efficiently. In addition, camera geometry also contributes to the effectiveness of crack detection, where the use of an oblique (slanting) camera angle generally provides better results by successfully detecting cracks that are often missed in vertical views [12] and [22]. Hence, camera configuration is not merely a technical consideration but a significant factor influencing the accuracy and reliability of crack detection models. Tailoring UAV image acquisition strategies to the geometry and orientation of the target structure can result in more consistent and accurate inspection outcomes.

2.4 Evaluation Metrics for the Crack Detection Model

The performance of crack detection models is commonly assessed using standard classification metrics, including precision, recall, and the F1 score. These metrics provide critical insights into how accurately and consistently a model can identify cracks under different operational conditions. Precision refers to the proportion of correctly identified cracks among all instances detected as cracks. It plays a key role in reducing false positives

and ensuring that irrelevant features, such as shadows or surface textures, are not mistakenly classified as cracks, particularly when working with pretrained segmentation models [18].

While precision focuses on the accuracy of detected cracks, recall addresses the model's ability to capture as many true cracks as possible. In other words, recall measures the proportion of actual cracks that are successfully detected by the model, including small or fine cracks. The F1 score represents the harmonic mean of precision and recall, offering a balanced assessment of a model's performance. Relying too heavily on a single metric may lead to misleading conclusions. Therefore, the F1 score is recommended as a more comprehensive indicator of detection robustness [15].

However, achieving reliable performance in real-world applications remains a challenge. Factors such as lighting variability, surface occlusion, and environmental noise can significantly affect model effectiveness during field inspections [22] and [23]. Therefore, a combination of precision, recall, and F1 score is necessary for evaluating the practical viability of crack detection systems when deployed in UAV-based inspections.

3. Methodology

By referring to Figure 1, this study began with site selection and planning, where a suitable building with visible cracks was identified for monitoring. Ground control points (GCPs) were established to enhance georeferencing accuracy during the photogrammetric process. Following this, images were acquired using UAVs based on different camera configurations. The collected images were subsequently processed using photogrammetry to generate 3D models. These models were then analysed using deep learning techniques for crack detection. Finally, an assessment was conducted to evaluate the reliability of the detection results based on the different UAV imaging setups. The following subsections explain the research methodology in detail.

3.1 Study Area

This study was conducted at Universiti Teknologi MARA (UiTM), Perlis Branch, Arau Campus, located in the northern region of Peninsular Malaysia. One of the buildings at this campus (Figure 2) suffers from cracks and requires frequent monitoring and was therefore selected as the experimental object.

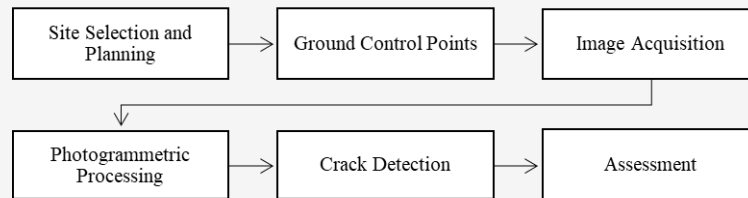


Figure 1: Crack detection using UAV study workflow

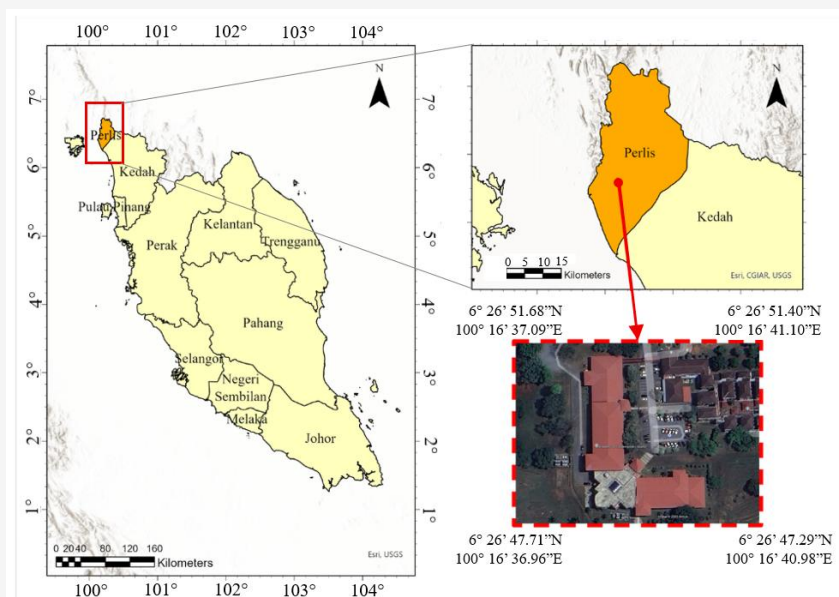


Figure 2: Universiti Teknologi MARA (UiTM), Perlis Branch, Arau Campus, Malaysia

The area of interest in this experiment comprised the flat rooftop and façade of the building, as these sections are more exposed to environmental conditions and are prone to structural cracks due to weathering and the age of the building. The structure stands at an estimated height of 15m, making it suitable for testing the effects of different altitude configurations. The selection of this location and building was also influenced by its suitability for inspection purposes, due to the presence of visible surface cracks, accessibility to UAV imaging, safety considerations, drone operation authorization, ease of conducting visual inspections, and the ability to verify field data.

3.2 Ground Control Point (GCP)

Prior to image acquisition, a reconnaissance survey was conducted to plan and manage the data collection workflow. As part of this process, Ground Control Points (GCPs) were established to ensure high positional accuracy and enable the generation of a georeferenced 3D model. In total, 11 GCPs were strategically placed around the study area, covering both the ground and concrete rooftop surfaces of the building. Of these, five were located on the ground surface and distributed across the area surrounding the building, while the remaining GCPs were placed on the concrete rooftop. The distribution of GCPs ensured coverage of both horizontal and vertical surfaces during reconstruction. Their placement was carefully planned to support accurate photogrammetric georeferencing and to minimize distortion during 3D reconstruction. The GCPs were measured using the Fast Static GPS surveying method, which provides high precision through extended observation periods and post-processing corrections. The WGS 84 (World Geodetic System 1984) coordinate system was selected to standardize the coordinates obtained from the drone for processing in the iTwin software. 11 Verification Points (VPs) were marked on the building façade to assess the accuracy of the 3D model. Table 1 summarizes the GNSS parameters used for GCP observation.

Table 1: GNSS Parameters used for GCPs observation

Aspect	Parameters
GNSS Method	Fast Static
Observation interval	1 second
PDOP Value	< 3
Observation Duration	30 minutes/points

3.3 Configuration for Crack Detection

In this study, three camera configurations were tested for data collection to evaluate their influence on crack detection using the pretrained model iTwin software. These configurations comprised vertical (nadir), slanting (oblique at 60°), and horizontal (side view) image capture. All images were acquired using a DJI Phantom 4 drone with 80% front and 75% side overlaps. The drone is equipped with a 1-inch CMOS sensor, a 24mm (equivalent) f/2.8 lens, and captures images at a photo resolution of 12 megapixels. The vertical configuration was set to capture images directly downward from the aerial perspective, as implemented in standard aerial mapping practices. The flying heights were set at 40m, 50m, and 60m above ground level. Considering that the building is 15m tall, the camera-to-object distances were approximately 25m, 35m, and 45m, respectively. The vertical aerial mapping configuration is commonly used to generate digital maps, surface models, and orthophotos, as the geometry of these products requires projection from above. However, this configuration is limited in reconstructing façades or side views of objects such as buildings.

Therefore, this study also experimented with a slanting configuration to capture both the concrete rooftop and the façade of the building for a complete 3D model reconstruction. The camera was set at a 60° angle and flown at the same altitude as in the vertical configuration. Meanwhile, the horizontal configuration refers to the side-view capture of the building façade, a technique that is commonly implemented for structural inspection and excludes the top view. For this configuration, images were captured at distances of 10m, 20m, and 30m from the building surface to evaluate the effect of proximity on detection performance. Table 2 summarizes the camera configurations used in this study, while Figure 3 shows the distribution of GCP and VP.

3.4 Data Processing

All acquired image datasets then underwent photogrammetric processing using the iTwin software. The first step involved image alignment, where the software performed photogrammetric triangulation to align the overlapping images and establish a spatial framework. Once the triangulation step was completed, the model was georeferenced before proceeding with 3D mesh reconstruction. The spatial framework was refined to process only selected areas of interest to reduce processing time and computer resources, specifically focusing on the flat concrete rooftop and its façade, where cracks were observed.

Table 2: Camera configurations

Configurations	Camera Angle (°)	Flying Height/ Camera-to-Object Distance (m)	Ground Sampling Distance (mm)
Vertical	0	25	16.408
		35	21.281
		45	25.121
Slanting	60	25	37.280
		35	41.685
		45	46.090
Horizontal	90	10	4.524
		15	8.599
		20	13.780



(a)



(b)

Figure 3: The distribution of Ground Control Points (GCP): (a) The distribution of Ground Control Points (labelled as CP) on the ground, (b) The distribution of Verification Points (VP) on the façade wall

The rest of the building, which was covered with metal panels, was excluded from the area of interest to avoid misdetection. In addition, the ‘Geometric Precision’ for the reconstruction setting was set to Ultra to ensure that the resulting 3D surfaces were sharply defined and free from blurring or smoothing artifacts. This was critical for enhancing the visibility of fine surface details such as cracks. Following the generation of the high-resolution 3D mesh model, crack detection was conducted using the ‘Extract Features’ tool in the iTwin Capture Modeler. This tool applies a pretrained deep learning model directly to the 2D UAV images, detecting cracks and representing them as cyan vector lines. The resulting crack vectors were then projected onto the generated 3D mesh model, allowing spatial visualization and measurement of the detected features, as shown in Figure 4. No fine-tuning or retraining of the model was performed, as the objective of this study was to assess the reliability of the pretrained system in its default configuration under real-world UAV imaging

scenarios.

3.5 Analysis

The ground truth (GT) for detection model assessment was established using manual visual inspection as well as inspection conducted using digital orthophotos. For the rooftop surface, crack data were manually recorded on-site through visual inspection. Each crack instance was logged in detail to provide a reliable reference for comparison with the output of the detection model. In addition to manual visual inspection, orthophotos of the rooftop were generated using Agisoft Metashape to verify whether any cracks were missed during the initial inspection. For the façade surfaces, GT data were obtained entirely through orthophoto analysis, as physical access to elevated vertical surfaces was limited. Agisoft Metashape was chosen for orthophoto generation due to its flexible control over processing parameters, allowing the production of high-resolution imagery suitable for visual crack

identification. This dual approach ensures comprehensive and verifiable crack annotations for evaluating the detection performance of the model.

The detection results were assessed using three standard metrics, which are precision, recall, and F1 score. These metrics are widely used in evaluating segmentation tasks, especially when dealing with imbalanced datasets such as crack detection, where true cracks (positive cases) are relatively sparse compared with background pixels. Precision quantifies the proportion of detected cracks that are actual cracks found in the images (Equation 1). Recall reflects the model's ability to detect actual cracks out of all existing cracks, including those that may have been missed or falsely detected (Equation 2). The F1 score (Equation 3) combines both precision and recall into a single metric and serves as a balanced indicator of the model's performance. These metrics are formally defined as:

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}}$$

Equation 1

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}}$$

Equation 2

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

Equation 3

In the context of crack detection, precision, recall, and F1 score are commonly used to evaluate model performance. A high recall indicates that the model successfully detects most existing cracks, whereas a high precision suggests that the detected cracks are mostly accurate. The F1 score is particularly valuable when there is a need to balance between identifying all true cracks and minimizing false alarms. Accuracy, on the other hand, is not typically used as an evaluation metric in crack detection tasks. This is because it requires a reliable count of true negatives, which represent correctly identified non-crack areas. In practice, it is difficult to define clear boundaries for non-crack regions, especially in complex datasets such as rooftop and façade imagery. These areas may contain diverse backgrounds and textures that are not easily categorized, making them unreliable for determining true negatives. Therefore, accuracy may become a misleading measure. Precision, recall, and F1 score are more appropriate because they focus on the detection of cracks, which are the positive class of interest, particularly in imbalanced datasets where crack features are sparse compared to the background.

4. Results and Discussion

This study evaluated the performance of a pretrained deep learning crack detection model across nine UAV imaging configurations, comprising vertical, slanting, and horizontal views at three different distances. The detection performance was assessed using precision, recall, and F1 score, as shown in Table 3 and Figure 5, to capture both the correctness and completeness of the model's predictions under varying conditions.

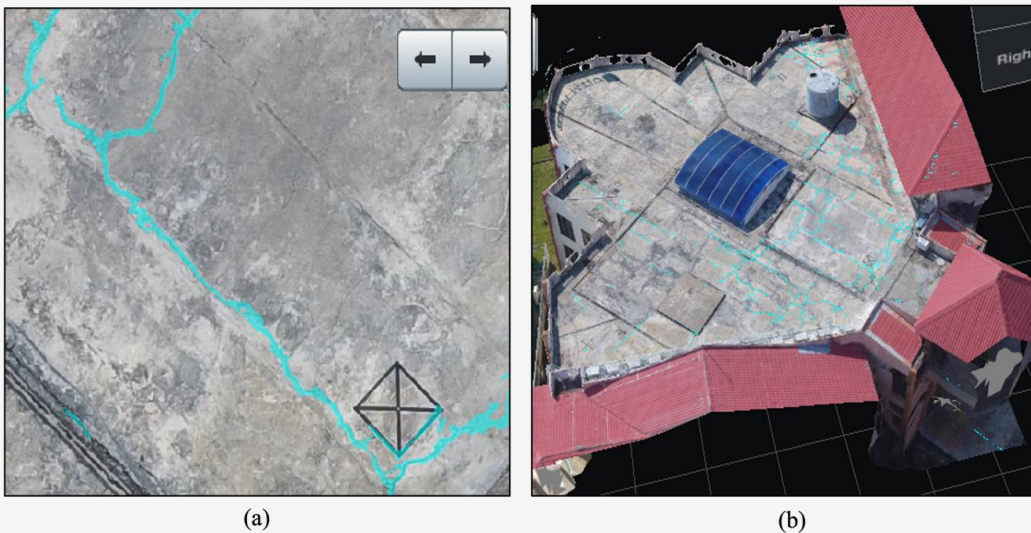
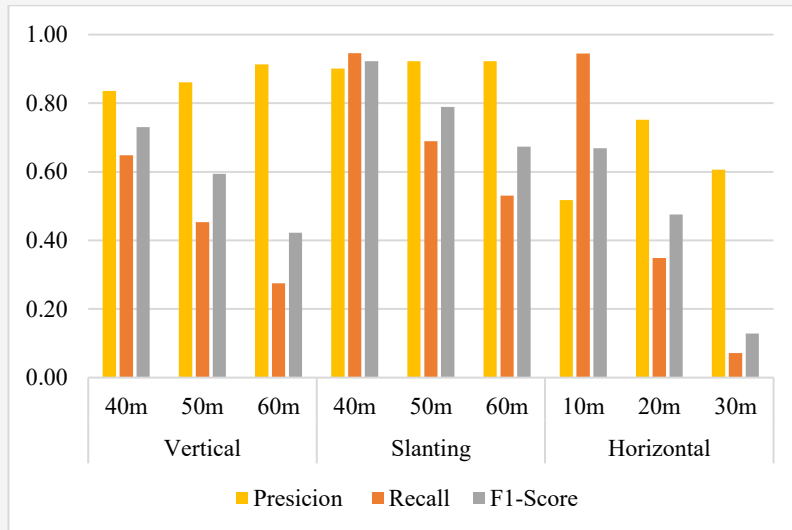


Figure 4: The detected cracks are represented as cyan vectors:

- (a) The detection was initially performed on 2D images, and
 (b) The same detected vector highlighted in red box projected onto the respective location on the 3D model

Table 3: Performance of pretrained crack detection model for different camera configurations

Configuration	Height/ Distance (m)	Precision %	Recall %	F1 score %
Vertical	40	83.57	64.84	73.02
	50	86.11	45.35	59.41
	60	91.29	27.48	42.24
Slanting	40	90.10	94.56	92.28
	50	92.24	68.93	78.90
	60	92.27	53.05	67.37
Horizontal	10	51.77	94.48	66.89
	20	75.14	34.83	47.60
	30	60.62	7.21	12.88

**Figure 5:** Precision, recall, and F1 score comparison across camera configurations

The vertical configuration simulated conventional photogrammetric mapping, in which the camera was oriented directly downward. At 40m, the model yielded a moderate F1 score of 73.02%, with a precision of 83.57% and a recall of 64.84%. However, as the flight altitude increased to 50m and 60m, the recall values dropped sharply to 45.35% and 27.48%, respectively, despite a continuous increase in precision that reached 91.29% at 60m. This trend indicates that the model becomes more conservative at higher altitudes. While it successfully filtered out false detections, it also failed to identify many actual cracks. The main factor contributing to this performance drop is the reduced image resolution at higher altitudes, which results in the loss of fine texture and edge features required by the pretrained model to recognize subtle crack patterns. As the flight altitude increased from 40m to 60m, the corresponding image resolution decreased from 16 mm/pixel to 25 mm/pixel, as shown in Figure 6. Although the cracks remain visually discernible to the human eye, the reduced pixel-level detail impacts the model's ability to detect them, especially when

those cracks are narrow or faint. However, the lower resolution at higher altitudes also reduces the model's sensitivity to surface noise and minor artifacts, which in turn leads to fewer false positives and hence higher precision values.

This effect is further illustrated in Figure 7, which shows the spatial distribution of detected cracks across the building surface at different altitudes. At 40m, the detection coverage is dense and widespread, especially in regions highlighted by red boxes, many of which correspond to false positives. As the altitude increases to 50m and 60m, the number of such detections decreases noticeably. At 60m, the detection becomes sparse or absent in several regions, indicating the model's reduced sensitivity due to lower resolution. However, this also results in fewer false crack detections, contributing to the higher precision observed at greater heights. This trade-off confirms the model's conservative behaviour at higher altitudes, as it successfully avoids misclassifying non-crack textures as cracks but simultaneously risks failing to detect fine or incipient cracks.

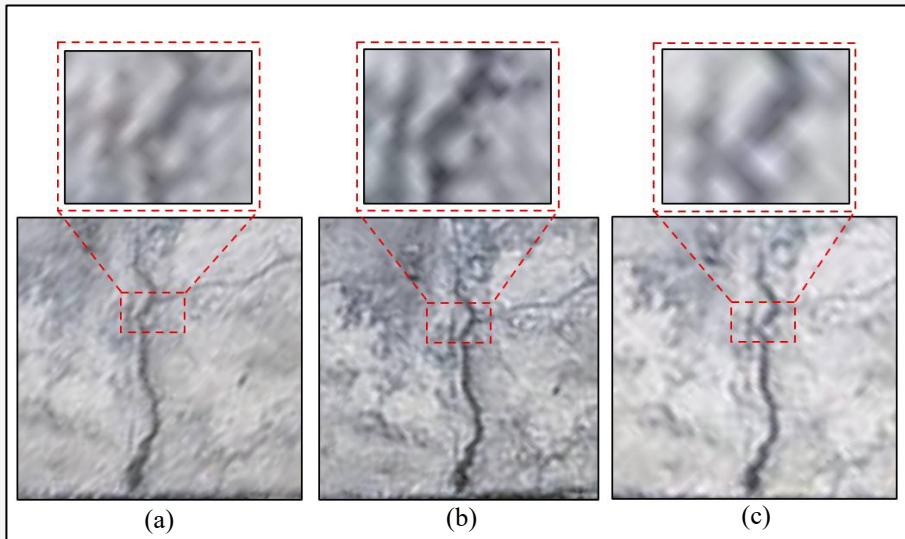


Figure 6: Close-up images of the same crack at (a) 40m, (b) 50m, and (c) 60m. Higher altitudes reduce image resolution, making cracks less visually distinct and harder for the model to detect

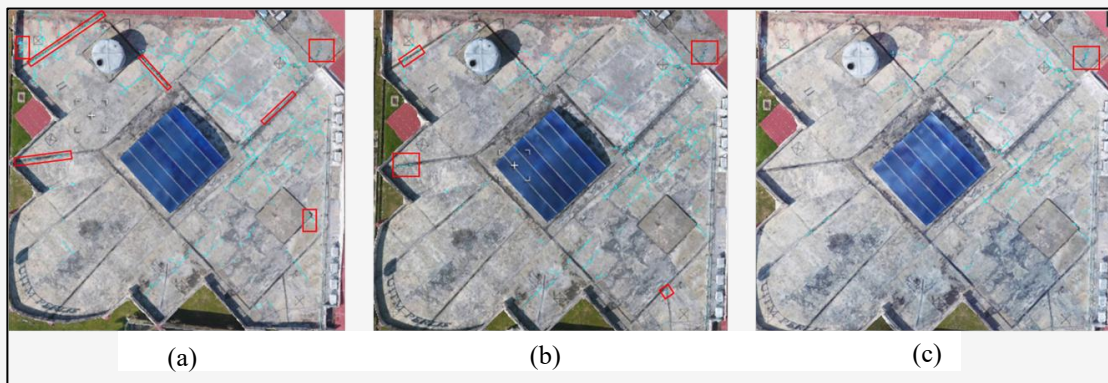


Figure 7: Crack detection at (a) 40m, (b) 50m, and (c) 60m. High resolution at 40m causes over-detection (red boxes), while at higher altitudes the model's sensitivity reduces false positives but may miss finer cracks

The horizontal configuration, involving side-view imaging at 10m, 20m, and 30m from the building façade, revealed a contrasting detection pattern compared to the vertical setup. At 10m, the model yielded a recall of 94.48%, indicating high sensitivity to actual cracks. However, precision was relatively low at 51.77%, resulting in an F1 score of 66.89%. As the distance increased, both precision and recall declined further, leading to a significant drop in overall performance. Unlike the vertical configuration, where the camera captured mostly flat concrete surfaces with minimal texture, the horizontal setup introduced a texture-rich environment. Vertical imaging typically involves rooftops or pavement areas with uniform tone and fewer visual distractions, which allows the model to focus more effectively on crack features. In contrast, the side-view images of the building façade included various architectural elements such as window frames, joint lines, and decorative grooves.

Although horizontal imaging at close range provided high image resolution, the abundance of linear and edge-like patterns within the scene increased the chances of false positives. As shown in Figure 8, the model incorrectly detected window edges and repetitive façade structures as cracks. The presence of strong contrast between windowpanes, shadows, and wall boundaries further contributed to the confusion. These visual complexities reduced the model's precision and highlighted the challenge of applying pretrained detection models in scenes with diverse background textures. Among all configurations tested, the slanting orientation at 40m yielded the highest overall performance, with an F1 score of 92.28%, a precision of 90.10%, and a recall of 94.56%. Even at increased altitudes of 50m and 60m, the model maintained a consistently high level of precision, exceeding 92%, along with a respectable recall.



Figure 8: Detection output for the horizontal configuration at 10m. The model incorrectly identified window frames and façade elements as cracks due to high image resolution and the presence of texture-rich surfaces

This performance suggests that slanting imaging provides an ideal balance between surface visibility and feature clarity. The oblique angle improves the visual differentiation of surface discontinuities, such as cracks, by introducing shadows and variations in reflectance. This enhanced contrast helps the model distinguish crack features more effectively. Additionally, the slanted perspective increases the apparent size of vertical surfaces within the image frame, which improves the effective image resolution along the plane of interest compared to a vertically oriented image. The consistently high F1 scores across multiple altitudes indicate that the pretrained model generalises well when presented with imagery that contains rich visual features, which aligns with the expected behaviour of deep learning-based detectors.

Compared to the vertical configuration, the slanting setup preserves more surface detail at higher altitudes, allowing the model to maintain strong recall performance. In contrast with horizontal imaging, it avoids excessive background complexity and minimizes the risk of false positives. These advantages contribute to the superior and consistent performance of the slanting configuration, making it the most reliable option among those tested for pretrained crack detection.

5. Conclusions

This study investigated the applicability and reliability of a pretrained crack detection model embedded in Bentley Systems' iTwin Capture Modeler software. The original model was developed

using a combination of UAV and handheld imagery with a resolution of 1 cm/pixel. In contrast, this research used only UAV-acquired images, captured under varying camera configurations, to assess how these variations influence model performance when applied to a single structure. The results show that while the model was able to detect cracks under all tested configurations, clear differences were observed in detection quality. Slanting and vertical views produced more complete and accurate crack detection outcomes compared to horizontal views. This indicates that camera orientation and viewing angle significantly affect the effectiveness of pretrained detection models, even when used on the same target structure.

By applying a consistent pretrained model across images captured with different UAV settings, this study demonstrates that image acquisition strategies play a crucial role in shaping detection outcomes. The findings contribute to a deeper understanding of how pretrained models behave when deployed under conditions different from their original training specifications, especially in practical UAV inspection workflows. These insights may inform decision-making in future UAV-based structural assessments that rely on pretrained models without retraining.

Despite these promising results, the study is subject to several limitations. The pretrained model was applied in its default form without any fine-tuning, which may limit its ability to handle surface textures or lighting conditions not represented in the original training data. Ground truth data for vertical

and slanting configurations were generated through manual annotation and visual inspection on rooftop areas, which may introduce human bias or subjectivity. For the horizontal configuration, ground truth was derived from orthophotos, which also carry limitations related to resolution accuracy and visual interpretation. Additionally, the image dataset was limited to a single building type under specific environmental conditions, which may restrict the generalizability of the findings to other structural contexts.

Acknowledgment

The authors would like to express their sincere appreciation to Universiti Teknologi MARA for providing the research facilities under research grant 100-TNCPI/GOV 16/6/2 (018/2024). The authors also extend their gratitude to the Faculty of Built Environment and Surveying (FABU), Universiti Teknologi Malaysia (UTM), and PLUS Malaysia Berhad for their invaluable technical assistance and financial support under grants VOT R.J130000.7652.4C773 and Q.J130000.3052.04M83. Additionally, the authors acknowledge the Land Surveyors Board Malaysia (LJT) for sponsoring the presentation of this research at the 10th International Conference on Geomatics and Geospatial Technology.

References

- [1] Hamishebahar, Y., Guan, H., So, S. and Jo, J., (2022). A Comprehensive Review of Deep Learning-Based Crack Detection Approaches. *Applied Sciences*, Vol. 12(3). <https://doi.org/10.3390/app12031374>.
- [2] Peruñ, G., (2024). Advances in Non-Destructive Testing Methods. *Materials*, Vol. 17(5). <https://doi.org/10.3390/ma17030554>.
- [3] Golding, V. P., Gharineiat, Z., Munawar, H. S. and Ullah, F., (2022). Crack Detection in Concrete Structures using Deep Learning. *Sustainability*, Vol. 14(13). <https://doi.org/10.3390/su14138117>.
- [4] Chakurkar, P. S., Vora, D., Patil, S., Mishra, S. and Kotecha, K., (2023). Data-Driven Approach for AI-Based Crack Detection: Techniques, Challenges, and Future Scope. *Frontiers in Sustainable Cities*. Vol. 5. <https://doi.org/10.3389/frsc.2023.1253627>.
- [5] Yuan Q., Shi, Y. and Li, M., (2024). Review of Computer Vision-Based Crack Detection Methods in Civil Infrastructure: Progress and Challenges. *Remote Sensing*, Vol. 16(16). <https://doi.org/10.3390/rs16162910>.
- [6] Syetiawan, A., Maharani, M., Nugroho, A., Setyo, R., Apriyanti, D., Augustian, M., and Prihatin, S. (2025). Precise 3d Modelling Using Combination Depth Maps Processing and Ground Control Point. *International Journal of Geoinformatics*, Vol. 21(4), 71–81. <https://doi.org/10.52939/ijg.v21i4.4079>.
- [7] Katsigiannis, S., Seyedzadeh, S., Agapiou, A. and Ramzan, N., (2023). Deep Learning for Crack Detection on Masonry Façades Using Limited Data and Transfer Learning. *Journal of Building Engineering*, Vol. 76. <https://doi.org/10.1016/J.Job.2023.107105>.
- [8] Lu, X., Li, Q., Li, J. and Zhang, L., (2025). Deep Learning-Based Method for Detection and Feature Quantification of Microscopic Cracks on the Surface of Concrete Dams. *Measurement*, Vol. 240. <https://doi.org/10.1016/j.measurement.2024.115587>.
- [9] Deng, P., Yao, J., Li, C., Wang, S., Li, X., Ojha, V., He, X. and Matsumoto, T., (2025). Unified Few-Shot Crack Segmentation and its Precise 3D Automatic Measurement in Concrete Structures. *arXiv Preprint*. [Online]. Available: <https://doi.org/10.48550/arXiv.2501.09203>. [Accessed April 10, 2025].
- [10] Ahmed, S., El-Shazly, A., Abed, F. and Ahmed, W., (2022). The Influence of Flight Direction and Camera Orientation on the Quality Products of UAV-Based SfM-Photogrammetry. *Applied Sciences*, Vol. 12(20). <https://doi.org/10.3390/app122010492>.
- [11] Chen, X., Liu, C., Chen, L., Zhu, X., Zhang, Y. and Wang, C., (2024). A Pavement Crack Detection and Evaluation Framework for a UAV Inspection System Based on Deep Learning. *Applied Sciences*, Vol. 14(3). <https://doi.org/10.3390/app14031157>.
- [12] Li, Y., Ma, J., Zhao, Z. and Shi, G., (2022). A Novel Approach for UAV Image Crack Detection. *Sensors*, Vol. 22(9). <https://doi.org/10.3390/s22093305>.
- [13] Bentley Systems. *iTwin Capture Detectors Download Page*. Bentley Knowledge Base. [Online]. Available: https://bentleysystems.service-now.com/community?id=kb_article&sysparm_article=KB0107801 [Accessed Mar. 10, 2025].
- [14] Yiğit, A. Y. and Uysal, M., (2024). Automatic Crack Detection and Structural Inspection of Cultural Heritage Buildings using UAV Photogrammetry and Digital Twin Technology. *Journal of Building Engineering*, Vol. 94. <https://doi.org/10.1016/j.job.2024.109952>.

- [15] Zhang, Z., He, Y., Hu, D., Jin, Q., Zhou, M., Liu, Z., Chen, H., Wang, H. and Xiang, X., (2025). Algorithm for Pixel-Level Concrete Pavement Crack Segmentation Based on an Improved U-Net Model. *Scientific Reports*, Vol. 15. <https://doi.org/10.1038/s41598-025-91352-x>.
- [16] Song, F., Liu, B. and Yuan, G., (2024). Pixel-Level Crack Identification for Bridge Concrete Structures using Unmanned Aerial Vehicle Photography and Deep Learning. *Structural Control and Health Monitoring*, Vol. 2024(1). <https://doi.org/10.1155/2024/1299095>.
- [17] Wang, J., Ueda, T., Wang, P., Li, Z. and Li, Y., (2025). Building Damage Inspection Method using UAV-Based Data Acquisition and Deep Learning-Based Crack Detection. *Journal of Civil Structural Health Monitoring*, Vol. 15(1), 151–171, <https://doi.org/10.1007/s13349-024-00836-3>.
- [18] Jin, T., Zhang, W., Chen, C., Chen, B., Zhuang, Y. and Zhang, H., (2023). Deep-Learning- and Unmanned Aerial Vehicle-Based Structural Crack Detection in Concrete. *Buildings*, Vol. 13(12), <https://doi.org/10.3390/buildings13123114>.
- [19] Chen, K., Reichard, G., Xu, X. and Akanmu, A., (2021). Automated Crack Segmentation in Close-Range Building Façade Inspection Images Using Deep Learning Techniques. *Journal of Building Engineering*, Vol. 43. <https://doi.org/10.1016/j.jobe.2021.102913>.
- [20] Özgenel, Ç. F. and Sorguç, A. G., (2018). Performance Comparison of Pretrained Convolutional Neural Networks on Crack Detection in Buildings. *ISARC Proceedings*, Vol. 35, 94–101. <https://doi.org/10.22260/ISARC2018/0094>.
- [21] Dorafshan, S., Thomas, R. J. and Maguire, M., (2018). Comparison of Deep Convolutional Neural Networks and Edge Detectors for Image-Based Crack Detection in Concrete. *Construction and Building Materials*, Vol. 186, 1031–1045. <https://doi.org/10.1016/j.conbuildmat.2018.08.011>.
- [22] Jang, K., Park, S., Jung, H., Yoo, H. and An, Y. K., (2024). Deep Learning-Based 3D Digital Damage Map of Vertical-Type Tunnels Using Unmanned Fusion Data Scanning. *Automation in Construction*, Vol. 162. <https://doi.org/10.1016/j.autcon.2024.105397>.
- [23] Mei, Q. and Gül, M., (2020). A Cost Effective Solution for Pavement Crack Inspection Using Cameras and Deep Neural Networks. *Construction and Building Materials*, Vol. 256. <https://doi.org/10.1016/j.conbuildmat.2020.119397>.