HOKKAIDO UNIVERSITY

Title	Crack detection using CNN with video images of railway tunnel central passage
Author(s)	Kadota, Yuichiro; Takuwa, Koichiro; Kitahara, Takeshi
Relation	Proceedings of the 10th International Symposium on Reliability Engineering and Risk Management (ISRERM 2026), June 28 - July 1, 2026, Sapporo, Japan
Citation	Proceedings of the 10th International Symposium on Reliability Engineering and Risk Management (ISRERM 2026)
Issue Date	2026-07-01
Doc URL	https://hdl.handle.net/2115/100421
Type	conference paper
File Information	C000010_1.pdf

Hokkaido University Collection of Scholarly and Academic Papers : HUSCAP

10th International Symposium on Reliability Engineering and Risk Management (ISRERM 2026),
28 June-01 July 2026, Sapporo, Japan
Kitahara, T., Miyamori, Y. & Beer, M. (editors)

Crack detection using CNN with video images of railway tunnel central passage

Yuichiro Kadota[1], Koichiro Takuwa[2], and Takeshi Kitahara[3]*, [4]

[1] *Graduate School of Engineering, Kanto Gakuin University, Yokohama, Japan (Presenting author).*
E-mail: y-kadota@jrcc.co.jp
[2] *Investigation Department, JR Central Consultants Company, Tokyo, Japan*
E-mail: k-takuwa@jrcc.co.jp
[3] *Department of Civil Engineering, Kanto Gakuin University, Yokohama, Japan*
*E-mail: kitahara@kanto-gakuin.ac.jp, *Corresponding Author*
[4] *Graduate School of Global Environmental Studies, Kyoto University, Kyoto, Japan*
kitahara.takeshi.8t@kyoto-u.ac.jp

Improving the efficiency of crack inspections is a challenge in the maintenance and management of concrete railway structures. Conventional close-up visual inspections have been problematic due to errors and oversights arising from subjective judgment, as well as the effort required for recording them. Therefore, in this study, we attempted to detect cracks by video recording on the side walls of the central passage of a railway tunnel. Given the on-site conditions, we built a video recording system using a tablet and lighting equipment and considered shooting parameters such as resolution, frame rate, and illuminance. In the future, we will convert videos into still images and identify cracks, aiming to improve the efficiency of inspections and advance the system through automatic diagnosis

Keywords: Crack detection, Video-based inspection, Railway tunnel, Orthophoto generation, Image processing.

1 Introduction

In the maintenance and management of concrete railway structures, it is essential to periodically identify structural deterioration such as cracking. At present, crack inspections are mainly conducted through close visual inspection by inspectors. In this process, crack widths are measured using crack gauges, and the results are manually recorded in inspection reports (Railway Technical Research Institute 2007).

In railway tunnels, as shown in Figure 1, a central passage is constructed along the tunnel axis. As part of routine tunnel maintenance, vertical ground displacement beneath the track bed is regularly monitored through leveling surveys. When significant displacement is detected by comparing leveling results obtained at different inspection periods, a detailed investigation is conducted. This investigation typically involves a close visual inspection of cracks on the side walls of the railway tunnel's central passage, as illustrated in Figure 2.

However, close visual inspection largely relies on inspectors' subjective judgement, which can lead to human error, including overlooking cracks and difficulties in accurately quantifying crack progression between inspection cycles. In addition, on-site inspection work and subsequent organization of inspection records are time-consuming. Given the expected reduction in the number of experienced engineers, there is a growing need to conduct inspections more efficiently.

In recent years, image processing technologies and artificial intelligence (AI) techniques have been increasingly applied to the detection of surface deterioration, such as crack (Ukai 2000). Nevertheless, applying existing image analysis methods to railway tunnels presents several challenges. Railway tunnel environments are characterized by limited illumination, resulting in low-contrast images that are susceptible to noise. Consequently, conventional image analysis techniques may fail to achieve sufficient crack detection accuracy under such conditions.

In this study, we aim to establish image acquisition methods and shooting conditions suitable for the side walls of railway tunnel central passages, which are concrete structures subjected to low-light environments. As a preliminary investigation, trial experiments were conducted on a reinforced concrete (RC) structure at Kanto Gakuin University to evaluate the applicability of the proposed approach under conditions simulating those of actual railway tunnels.

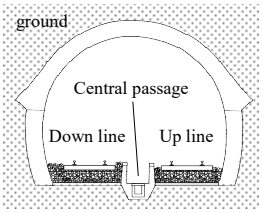


Figure 1. Cross section of a railway tunnel.

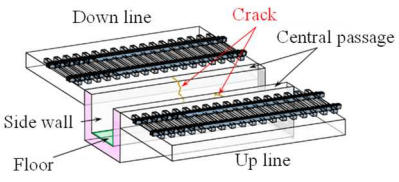


Figure 2. Central passage side wall and crack locations.

2 Challenges in Image-Based Crack Detection

When applying existing image analysis techniques to crack inspection on the side walls of railway tunnel central passages, several challenges arise.
First, railway tunnels are characterized by limited illumination, making it difficult to obtain sufficient lighting conditions. As a result, images captured in such environments tend to be dark and low in contrast, which significantly hinders accurate crack detection using conventional image analysis methods.
Second, traditional image-based crack detection approaches generally rely on a large number of still images captured from multiple viewing angles. In practice, this requires inspectors to take several hundred photographs per field of view in order to ensure adequate coverage of the inspection area. However, the central passage of a railway tunnel is a confined space, and the available inspection time is often limited. Consequently, conducting extensive still-image acquisition on site places a substantial burden on inspectors.
To address these challenges, this study explores an alternative approach in which video footage of the inspection area is captured as the operator moves along the tunnel. The recorded videos are subsequently converted into still images for crack detection. By adopting this method, it is expected that cracks can be detected effectively even under low-light conditions and with relatively low image quality, while significantly reducing the workload associated with on-site inspection activities (Kim et al. 2018).

3 Investigation of Image Analysis Methods

This section examines image analysis methods for crack detection on the side walls of the railway tunnel's central passages. To establish a practical inspection approach for low-light tunnel environments, suitable video acquisition methods, shooting conditions, and image processing procedures were investigated. Particular attention was paid to selecting video-recording equipment, camera settings, and image generation techniques to produce orthophotos suitable for image-based crack detection.

3.1 *Image acquisition method*

The video recording system developed for image acquisition is shown in Figures 3 and 4. In order to ensure practical applicability for on-site inspections, devices with simple data-handling capabilities were selected. Specifically, a tablet and an action camera were used for video recording (Wang 2023).
During video acquisition, the recording devices were mounted and fixed on a trolley to maintain a constant field of view. Video footage was captured while illuminating the target area with a handheld lighting device. The camera angle was set to be perpendicular to the side wall of the railway tunnel's central passage as the baseline configuration in order to minimize geometric distortion and ensure consistent image quality.

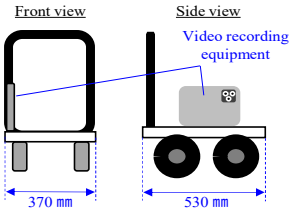


Figure 3. Video recording system.



Figure 4. Video recording system (overall view).

3.2 *Selection of video recording equipment*

The video recording equipment used in this study is summarized in Table 1. Four types of equipment that could be mounted and securely fixed on a trolley were selected for evaluation. The equipment was evaluated against three parameters: image resolution, frame rate, and illuminance.

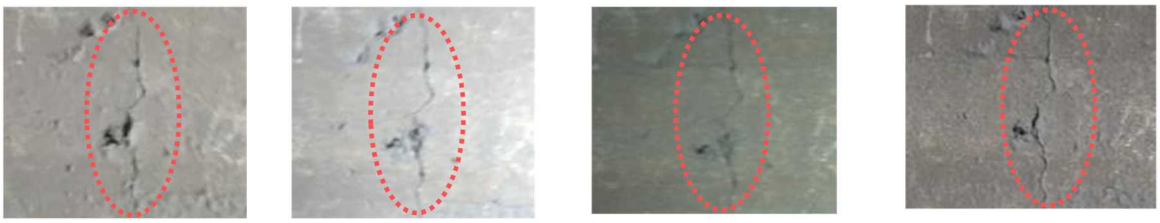
Table 1. List of video recording equipment.

Item	Pattern (1)	Pattern (2)	Pattern (3)	Pattern (4)
Video recording equipment	Tablet (Apple iPad Pro) 	Smartphone (Apple iPhone 15 Pro Max) 	Digital camera (Sony α7R IV) 	Action camera (GoPro HERO9) 
Image stabilization	Optical Image Stabilization (OIS)	Optical Image Stabilization (OIS)	Optical Image Stabilization (OIS)	Electronic Image Stabilization (EIS)
Number of pixels	Approx. 0.92 MP 720p (1280×720)	Approx. 8.3 MP 4K (3840×2160)	Approx. 8.3 MP 4K (3840×2160)	Approx. 8.3 MP 4K (3840×2160)

Regarding image resolution, previous investigations have confirmed that the typical crack width detectable is approximately 1.0 mm. Based on this requirement, an image resolution of 720p (approximately 0.92 megapixels), corresponding to a spatial resolution of 0.78 mm per pixel or finer, was adopted as the minimum specification. In order to obtain higher image quality and potentially improve crack detection accuracy, equipment patterns (2) to (4) were configured to record video at 4K resolution (approximately 8.3 megapixels).
The frame rate was set to 30 frames per second (fps), assuming a traveling speed equivalent to a person walking while pushing the trolley during on-site inspections. This frame rate was considered sufficient to ensure continuous image acquisition without excessive motion blur.As for illuminance, a handheld lighting device commonly used in field inspections was employed. Under these conditions, the illuminance on the target surface was approximately 2300 lx, which was considered representative of practical inspection environments in railway tunnels.

3.3 *Consideration of video recording equipment*

Figure 5 shows images of the same crack extracted from videos recorded using the four types of video recording equipment listed in Table 1. Since video recording was conducted while moving along the central passage on a trolley, the videos captured using equipment patterns (1) to (3) were affected by trolley-induced vibration, resulting in insufficient image sharpness and difficulty in clearly capturing cracks.
In contrast, the action camera used in pattern (4) compensated for image blur caused by trolley vibration via Electronic Image Stabilization (EIS). As a result, cracks were captured more clearly compared with the other equipment. This indicates that EIS plays a significant role in suppressing motion-induced blur during video acquisition in tunnel environments.
Based on these results, the action camera was selected as the most suitable video-recording device for this study. In subsequent experiments, video data acquired with the action camera were used to generate a three-dimensional (3D) model and for image-based crack detection.



Pattern (1): Tablet. Pattern (2): Smartphone. Pattern (3): Digital camera. Pattern (4): Action camera.

Figure 5. Crack images captured using different video recording equipment.

3.4 *Generation of orthophotos*

To detect cracks, it is necessary to convert the acquired video data into still images. Therefore, RealityScan (Capturing Reality Ltd.) software was used to generate three-dimensional (3D) models and orthophotos of the target surfaces of reinforced concrete (RC) structures. RealityScan is photogrammetry-based software that enables the creation of 3D models from video data captured from multiple viewpoints.
However, as shown in Figure 6, when a 3D model was generated using video recorded from only a single frontal direction, the resulting model exhibited curvature. This is considered to be due to the lack of depth information when images are acquired from a single viewing direction.
To address this issue, in addition to the video recorded from the frontal direction, supplementary videos were captured from viewing angles of 45° to the left and 45° to the right, as illustrated in Figure 7, in order to provide sufficient depth information. Using videos acquired from these three directions, a 3D model was generated. The resulting 3D model is shown in Figure 8. By using images captured from three different directions, a high-accuracy 3D model and orthophotography of the target structure were successfully generated without geometric distortion.

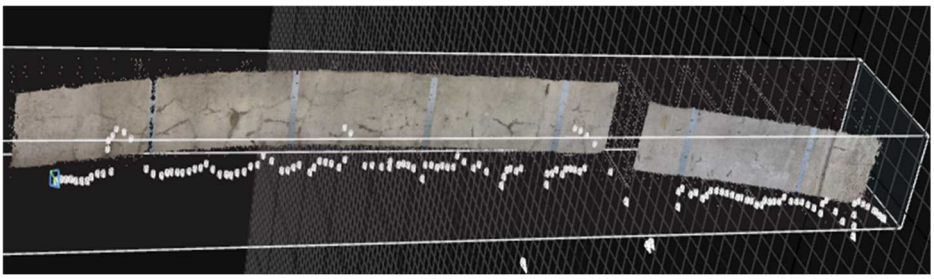


Figure 6. 3Dmodel generated from video captured from a single direction.

To address this issue, in addition to the video recorded from the frontal direction, supplementary videos were captured from viewing angles of 45° to the left and 45° to the right, as illustrated in Figure 7, in order to provide sufficient depth information. Using videos acquired from these three directions, a 3D model was generated. The resulting 3D model is shown in Figure 8. By using images captured from three different directions, a high-accuracy 3D model and orthophotography of the target structure were successfully generated without geometric distortion.

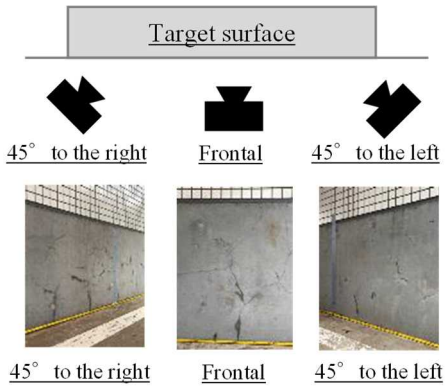
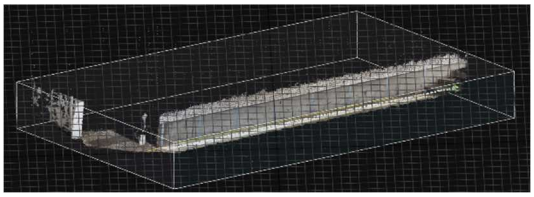
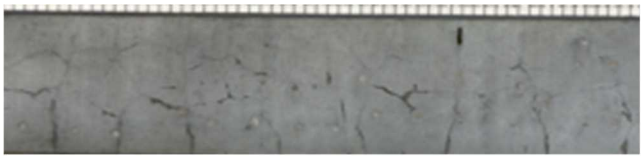


Figure 7. Video recording directions.



(a) 3D model (generated from videos captured from three directions).



(b) Orthophoto (generated from three-direction video data).

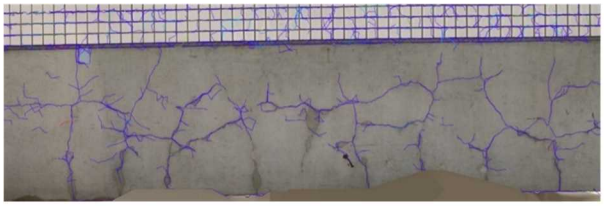
Figure 8. 3Dmodel generated from videos captured from three directions.

3.5 *Effect of shooting speed on crack detection results*

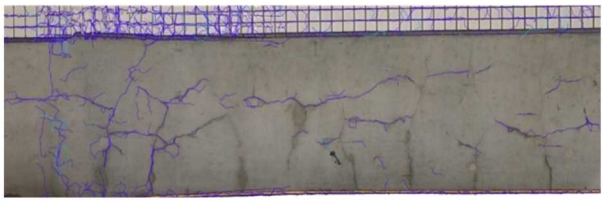
To examine the effect of shooting speed on crack detection results, videos were recorded using an action camera (GoPro) at two different shooting speeds: 1.0 km/h and 2.0 km/h. After video acquisition, orthophotos were generated from the recorded videos at each shooting speed, and crack detection was subsequently performed.

Crack detection was carried out using "Hibi-ken" (Crack Detection) (Chugai Technos Corporation; Infrastructure Inspection Support Technology Performance Catalog of the Ministry of Land, Infrastructure, Transport and Tourism, Technology ID: TN010010-V0325). "Hibi-ken" (Crack Detection) was adopted in this study because it can detect cracks with widths as small as 0.1 mm and allows efficient detection and editing of crack data. The crack detection results obtained from image analysis are shown in Figure 9. In the figure, blue lines indicate locations where cracks with widths of 0.2 mm or greater were detected.

Figure 10 compares the crack detection results obtained at different shooting speeds. Compared with the orthophotos generated from videos recorded at a shooting speed of 2.0 km/h, those generated at 1.0 km/h were clearer. Consequently, differences in crack detection accuracy were observed between the two shooting conditions. This suggests that the clearer orthophotos produced at the lower shooting speed contributed to improved crack detection performance.

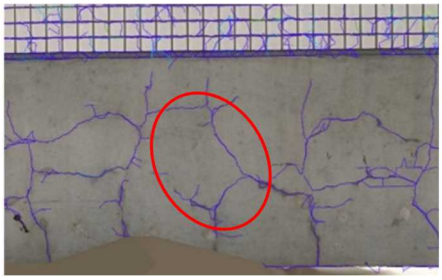


(a) Crack detection results (1.0 km/h).

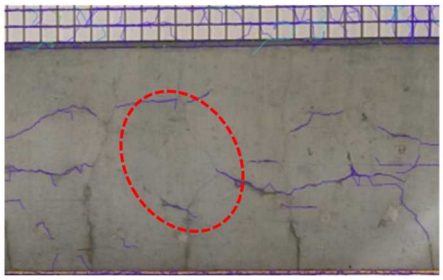


(b) Crack detection results (2.0 km/h).

Figure 9. Crack detection results.



(a) Crack detection results (1.0 km/h).



(b) Crack detection results (2.0 km/h).

Figure 10. Crack detection results at different shooting speeds.

4 Optimal Shooting Conditions

Based on the results of the above investigations, the optimal shooting conditions are summarized in Table 2. To suppress image blur and distortion in the generated three-dimensional (3D) models, an action camera equipped with Electronic Image Stabilization (EIS) was used for video recording device. The shooting speed was set to 1.0 km/h to ensure clear video acquisition suitable for orthophoto generation and stable crack detection. In addition, videos were recorded from three directions: a frontal direction, 45° to the left, and 45° to the right. By adopting this shooting configuration, high-quality orthophotos and accurate 3D models could be generated, enabling reliable crack detection on the side walls of the railway tunnel's central passage.

Table 2. Summary of optimal shooting conditions.

Item	Shooting conditions
Video recording equipment	Action camera (GoPro HERO9) 
Video acquisition method	Trolley-mounted lighting system
Number of pixels	Approx. 8.3 MP 4K（3840×2160）
Illuminance	Approx. 2300 lx (2200 lm)
Camera angles	Frontal / 45° left / 45° right (three directions)

5 Conclusion

In this study, a fundamental investigation was conducted to establish suitable image acquisition methods and shooting conditions for crack detection on the side walls of railway tunnel central passages, which are characterized by confined spaces and low-light environments. As a preliminary study, experiments were conducted on a reinforced concrete (RC) structure to simulate actual railway tunnel conditions.
The main findings of this study are summarized as follows.
First, video-based crack detection was investigated as an alternative to conventional close visual inspection, and practical shooting methods and conditions for on-site environments were organized.

Second, taking on-site applicability and ease of use into account, an action camera was selected as the video recording device. During video acquisition, the camera was mounted on a trolley, and video was recorded while illuminating the target area with a handheld lighting source.

Third, by using videos captured from three directions—frontal, 45° to the left, and 45° to the right—a three-dimensional (3D) model and orthophotos of the crack detection target surface could be generated without geometric distortion. This approach enabled the creation of high-accuracy 3D models and orthophotos suitable for reliable crack detection.

In railway tunnel environments with limited illumination, it is necessary to improve crack-detection accuracy even when using low-resolution orthophotos derived from videos recorded in darkness. In future work, annotations will be performed on crack images from railway tunnels to create training datasets, and the accuracy of machine-learning-based image analysis will be further improved.

Acknowledgments

The authors gratefully acknowledge Chugai Technos Corporation for providing the crack analysis software "Hibi-ken" (Crack Detection) used in this study. In particular, the authors would like to express their sincere appreciation for the valuable technical guidance on the software's functions and for the insightful advice on image acquisition throughout this research.

References

Kim, C. N., Kawamura, K., Tarighat, A. and Nakamura, H., A crack detection method for concrete infrastructures based on image processing technique and genetic algorithm, Proceedings of the Japan Concrete Institute, 40(1), 1785–1790, 2018.
Railway Technical Research Institute, Standards for Maintenance and Management of Railway Structures and Commentary, Railway Research-Culture Promotion Foundation, Japan, 2007.
Ukai, M., Development of image processing technique for detection of tunnel wall deformation using continuously scanned image, *Quarterly Report of RTRI*, 41(3), 120–126, 2000.
Wang, L., Automatic detection of concrete cracks from images using AdamSqueezeNet deep learning model, *Fracture and Structural Integrity*, 65, 289–299, 2023.