

Exploring the Potential of Super-Resolution for Crack Analysis in UAV Facade Orthomosaics of Small Bridges

Dahye Jang¹, Hyunsoo Park¹, Euiik Jeon^{1,2}

¹AI team, Innopam Co., Ltd., Yongsan-gu, Seoul, South Korea - dhjang@innopam.com, hspark@innopam.com, euiik@innopam.com

²Department of Urban Big Data Convergence, University of Seoul, Dongdaemoon-gu, Seoul, South Korea- euiik0323@uos.ac.kr

Keywords: Super-Resolution, Crack Segmentation, Facade Orthomosaic, UAV(Uncrewed Aerial Vehicles), Small Bridge

Abstract

UAV-based bridge inspections offer significant advantages in efficiency and safety, yet they face a fundamental trade-off between achieving the low Ground Sample Distance (GSD) required for high-precision damage analysis and maintaining operational efficiency. Acquiring imagery fine enough to quantify fine cracks (e.g., < 0.3 mm width) necessitates close-range flights that increase flight time and data volume, thereby diminishing the core benefits of UAVs. This study proposes and validates a workflow that leverages Super-Resolution (SR) technology to enhance the accuracy of quantitative analysis from efficiently captured, low-resolution orthomosaics. To achieve this, we first conducted a comparative analysis of four representative SR models (FSRCNN, SRGAN, Real-ESRGAN, and SwinIR) to identify the optimal architecture for bridge crack restoration. Second, the selected model was applied to a real-world facade orthomosaic (GSD ≈ 0.3 mm) generated from UAV imagery, followed by a quantitative comparison of crack length and width measurement accuracy before and after SR application. The results showed that Real-ESRGAN delivered the best performance. Most notably, the application of SR dramatically reduced the average relative error in crack width measurement from a prohibitive 149.11% to a practical 10.03%, while also more than halving the error in length measurement from 4.80% to 1.93%. This study demonstrates that SR is not merely a visual enhancement technique but a practical solution that enables the acquisition of high-precision, quantitative data comparable to that of a detailed safety inspection, all from safe and efficient UAV operations.

1. Introduction

The maintenance and safety assessment of aging infrastructure, such as bridges, are critical for public safety. In recent years, inspections using Unmanned Aerial Vehicles (UAVs) have become a standard practice, offering significant advantages in efficiency and accessibility over traditional methods (KALIS, 2022). However, a fundamental challenge remains: the trade-off between operational efficiency and measurement accuracy. To accurately detect and quantify structurally significant fine cracks (e.g., widths < 0.3 mm), a very low Ground Sample Distance (GSD) is required. Achieving this necessitates flying the UAV dangerously close to the structure, which drastically increases the number of required images, flight time, and operational complexity, thereby undermining the primary benefits of using a UAV.

To address this limitation, Super-Resolution (SR) technology has emerged as a promising solution. SR aims to computationally reconstruct a high-resolution (HR) image from one or more low-resolution (LR) inputs, potentially bridging the gap between efficient data acquisition and high-precision analysis. While various deep learning-based SR models, including those based on CNNs, GANs, and Transformers, have been developed, their comparative performance on the specific task of enhancing bridge facade orthomosaics for quantitative crack analysis has not been thoroughly investigated. It remains unclear which architectural approach is most suitable for restoring the unique textures of concrete cracks and how the application of SR quantitatively impacts the accuracy of crack measurements in a real-world scenario.

This study aims to systematically evaluate the effectiveness of SR technology for high-precision bridge crack analysis from

UAV-captured orthomosaics. To achieve this, our research is structured into two main experiments. First, we conduct a comparative performance analysis of four representative SR models (FSRCNN, SRGAN, Real-ESRGAN, and SwinIR) to identify the optimal model for this task. Second, we apply the selected optimal model to a real-world, UAV-generated facade orthomosaic (GSD ≈ 0.3 mm) to create an SR-enhanced version (GSD ≈ 0.1 mm). Finally, we perform a rigorous quantitative analysis by comparing crack segmentation results, as well as length and width measurements, between the original and the SR-enhanced orthomosaics against ground-truth data. This study provides a quantitative validation of SR as a practical tool to overcome the GSD limitations inherent in UAV-based inspections.

2. Methodology

The overall workflow of this study, as illustrated in Figure 1, consists of two main experiments.

The first experiment is the process of selecting the optimal Super-Resolution (SR) model. For this purpose, four SR models—FSRCNN, SRGAN, Real-ESRGAN, and SwinIR—are trained and their performance is comparatively evaluated using a self-constructed high-resolution dataset.

The second experiment aims to verify the practical applicability of the selected optimal model. Facade orthomosaics generated from low-resolution images acquired by a UAV, to which the optimal SR model selected in Experiment 1 is applied. Finally, crack segmentation, length, and width measurements are performed on the facade orthomosaics before and after SR application. The results are then compared to quantitatively analyze the effectiveness of the super-resolution technology.

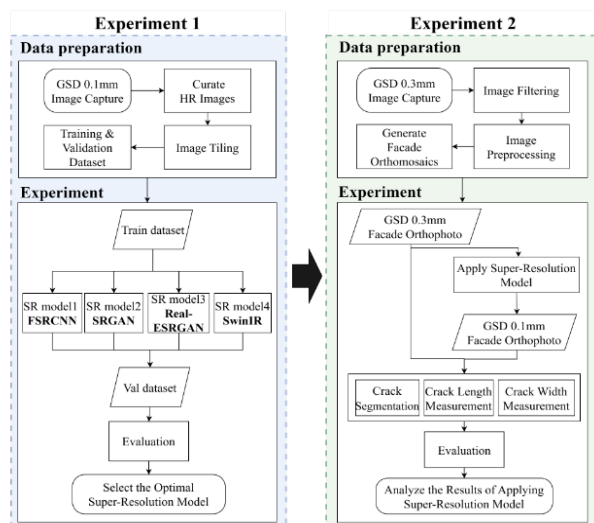


Figure 1. Workflow of the data preparation, processing and analysis used in this study

2.1 Data Preparation

2.1.1 Data Acquisition

This study constructed two types of image datasets with different purposes and specifications for the development and performance verification of Super-Resolution (SR) models. Both datasets were acquired from bridges classified as "Category 3 facilities" under South Korea's "Special Act on the Safety Control and Maintenance of Establishments."

First, the High-Resolution (HR) dataset for SR model training was acquired at Dongcheon Bridge (an RC slab bridge with a length of 38.4 m) located in Daegu, South Korea. To secure ultra-high-resolution images with a Ground Sample Distance (GSD) of 0.1 mm, a high-performance mirrorless camera, the Fujifilm GFX100 II (11648×8736 pixels), equipped with a GF 55mm lens was used. A total of 690 HR images of the abutments and piers were obtained while maintaining a shooting distance of approximately 1.4 m from the structure's surface.

Second, the Low-Resolution (LR) dataset for façade orthomosaic generation and verification of the SR model's applicability was acquired at Jujeong Bridge (an RC box girder bridge with a length of 24.0 m) in Jinju, Gyeongsangnam-do, South Korea. A DJI Mavic 3 Classic UAV (5280×3956 pixels) was utilized for image acquisition. The flight was conducted maintaining an offset distance of 1 m from the structure, aiming for a GSD of approximately 0.3 mm. However, image distortion occurred during the capture of the bridge's bottom slab due to the UAV gimbal's upward angle limitation (+35°). Furthermore, the non-uniformity of the shooting distance caused by manual flight resulted in the final acquired images having a GSD distribution ranging from approximately 0.25 mm to 0.4 mm.

2.1.2 Facade Orthomosaic Generation

In this study, Agisoft Metashape software was used to generate facade orthomosaics from the bridge images captured by the UAV. Metashape was selected for its robust image alignment capabilities, considering the characteristic of the bridge's underside being a GPS signal-shaded area.

The facade orthomosaic generation process is as follows. First, the relative positions of the images were estimated through Bundle Adjustment. At this stage, a model in a relative coordinate system without an absolute scale is created due to the absence of GPS information. To enable quantitative analysis of damages such as cracks, the absolute scale of the model was calibrated using a scale reference (a triangular ruler) attached to the structure.

Second, to enhance the quality of the 3D model, a Dense Point Cloud was first generated, followed by a refinement process of manually removing the background and noise. This was to prevent the generation of unnecessary data and to reduce the model's complexity, which can occur when generating a mesh directly from a depth map.

Third, a 3D mesh model was constructed based on the refined point cloud, and manual editing was performed to simultaneously improve processing efficiency and model precision. Finally, facade orthomosaics for each bridge member were generated as the final output from the completed 3D model by applying a Planar Projection method.

The size and GSD information of the generated facade orthomosaics are presented in Table 1, and the facade orthomosaics are shown in Figure 2. The facade orthomosaics were generated and classified by bridge component, and the front and back of the piers were specified as 'Front' and 'Back', respectively. Due to the non-uniform shooting distance between the UAV camera and the structure during manual flight under the bridge, the GSD of the generated facade orthomosaics ranged from 0.25 mm to 0.4 mm.

Bridge	Image name	Pixels (W x H)	GSD (mm)
Jujeong	A1	31473 x 9612	0.25
Jujeong	A2	23505 x 9278	0.34
Jujeong	S1	19356 x 9731	0.42
Jujeong	S2	18059 x 9592	0.45
Jujeong	S3	23186 x 12545	0.35
Jujeong	S4	23164 x 11683	0.35
Jujeong	P1_Front	27867 x 9391	0.3
Jujeong	P1_Back	22861 x 7533	0.35
Jujeong	P2_Front	24203 x 8867	0.33
Jujeong	P2_Back	19830 x 7203	0.4
Jujeong	P3_Front	19963 x 7218	0.4
Jujeong	P3_Back	24157 x 9674	0.33

Table 1. Pixel dimensions (W x H) and Ground Sampling Distance for the Orthomosaics of Jujeong bridge



Figure 2. Facade Orthomosaic of a Pier from the Jujeong Bridge

2.1.3 Super-Resolution Training Dataset Construction

The dataset for training the SR models was constructed based on single high-resolution images taken with a digital camera at Dongcheon Bridge. Out of a total of 690 original high-resolution images, 345 images containing crack objects were selected for use.

A dataset composed of paired LR and HR images is essential for SR model training. In this study, the tile size of the LR images was fixed at 256×256 pixels, and the tile size of the HR images was set according to the target scale factor. As the target scale factor in this study was set to 3x, the HR images were tiled into 768×768 pixels. Subsequently, only the tiles that clearly contained crack objects were further selected.

The LR tiles corresponding to the selected HR tiles were generated by downsampling them to a size of 256×256 pixels using the Bicubic interpolation method provided by the OpenCV library. Through this process, a final training dataset consisting of LR-HR image pairs was constructed. The entire dataset was randomly split into a training set and a validation set at a 9:1 ratio for model training and performance verification. The composition of the constructed dataset is summarized in Table 2.

Tile size	Image	# Train	# Validation	# Sum
768x768	HR	5539	615	6154
256x256	LR	5539	615	6154

Table 2. Overview of the SR Dataset Based on Tile Size

2.2 Analysis Methodology

2.2.1 Super-Resolution models

In this study, to select the model most optimized for the super-resolution of facade orthomosaics, deep learning models based on three representative architectures were comparatively analyzed. The selected models are the CNN (Convolutional Neural Network)-based FSRCNN, the GAN (Generative Adversarial Network)-based SRGAN and Real-ESRGAN, and the Transformer-based SwinIR. Each model possesses distinct advantages in computational efficiency, realistic texture restoration, and learning long-range dependencies, respectively.

The performance of the models was evaluated through a comprehensive integration of both quantitative metrics and qualitative assessment. For the quantitative evaluation, Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index Measure (SSIM) were utilized. Qualitative assessment involved a comparative analysis of the visual quality of the generated output. Notably, for GAN-based models, which are capable of exhibiting superior perceptual quality despite yielding lower PSNR scores, both quantitative metrics and visual naturalness were concurrently considered during the final model selection process.

2.2.2 Crack segmentation and Quantitative Analysis

In this study, the Mask2Former model, which uses a Swin Transformer as its backbone network, was adopted for crack segmentation (Cheng et al., 2022). A key feature of Mask2Former is the application of a masked attention-based transformer structure in its decoder. According to previous research, combining it with a Swin Transformer backbone has been reported to achieve particularly high segmentation accuracy (Liu et al., 2021). For model training, instead of constructing

a new dataset, pre-trained model weights from the study by Jeon (2023) on concrete tunnel crack images were utilized.

The model's segmentation output is in the form of a binary mask but tends to be predicted thicker than the actual crack. Therefore, to extract the precise center path of the crack, the segmentation area was first converted into a 1-pixel-thick skeleton using sci-kit-image's skeletonization algorithm. From this skeleton, Open CV's findContours was used to extract continuous contours, laying the groundwork for calculating the crack's length.

The crack width was measured through a multi-directional search centered on each point of the skeleton. From a center point $c=(x_0, y_0)$, pixels were searched in multiple directions at 10-degree intervals to find the boundary where the crack region (value 1) ends. A threshold-refined segmentation mask was used to improve the accuracy of boundary detection. As specified in the equation below, the width in a given direction (d_i) is calculated by summing the distances to the boundary in the search direction (r_f) and the opposite direction (r_b). Among the several width values calculated at a single point, the minimum value was adopted as the final width for that point. This value was then multiplied by the GSD to convert it to millimeters (mm).

$$r_f = \max \{r \mid 1 \leq r \leq r_{\max}, I(x_0 + r \cos \theta_i, y_0 + r \sin \theta_i) = 1\}$$

$$r_b = \max \{r \mid 1 \leq r \leq r_{\max}, I(x_0 - r \cos \theta_i, y_0 - r \sin \theta_i) = 1\}$$

$$d_i = r_f + r_b$$

3. Results and Discussions

3.1 Super-resolution models comparison

To objectively compare the performance among models, all experimental conditions except for the model architecture were set identically. Specifically, all models used the same input and output data. The training process was conducted for 20 epochs with a batch size of 2. The Adam optimizer was used with a learning rate of 1×10^{-4} . Mean Squared Error (MSELoss) was commonly applied as the loss function. No data augmentation techniques were used, and training was performed in a DataParallel environment using 2 GPUs. This approach was intended to clearly analyze the impact of structural differences between models on super-resolution performance.

In the quantitative evaluation, the changes in Loss, Peak Signal-to-Noise Ratio (PSNR), and Structural Similarity Index Measure (SSIM) for the training data of each model are visualized per epoch in Figure 3. Analysis of the graphs in Figure 3 shows that in 'Loss per Epoch', the Real-ESRGAN model maintained the lowest loss values throughout the training process and exhibited the most stable learning trend. In Figure 3, the Real-ESRGAN model also sustained the highest PSNR values across all epochs, demonstrating its superior pixel-level image restoration capability.

The final PSNR and SSIM performance of each model on the validation dataset is presented in Table 3. The results show that the Real-ESRGAN model achieved a PSNR of 34.20 dB and an SSIM of 0.82, outperforming the other compared models in quantitative performance metrics. Considering the stability and performance improvement trend during the training process, as well as the final performance metrics on the validation dataset, the Real-ESRGAN model demonstrated the most outstanding super-resolution performance among the SR models compared in this study. This proves that Real-ESRGAN possesses excel-

lent restoration capabilities in both pixel-level accuracy (PSNR) and visual quality (SSIM).

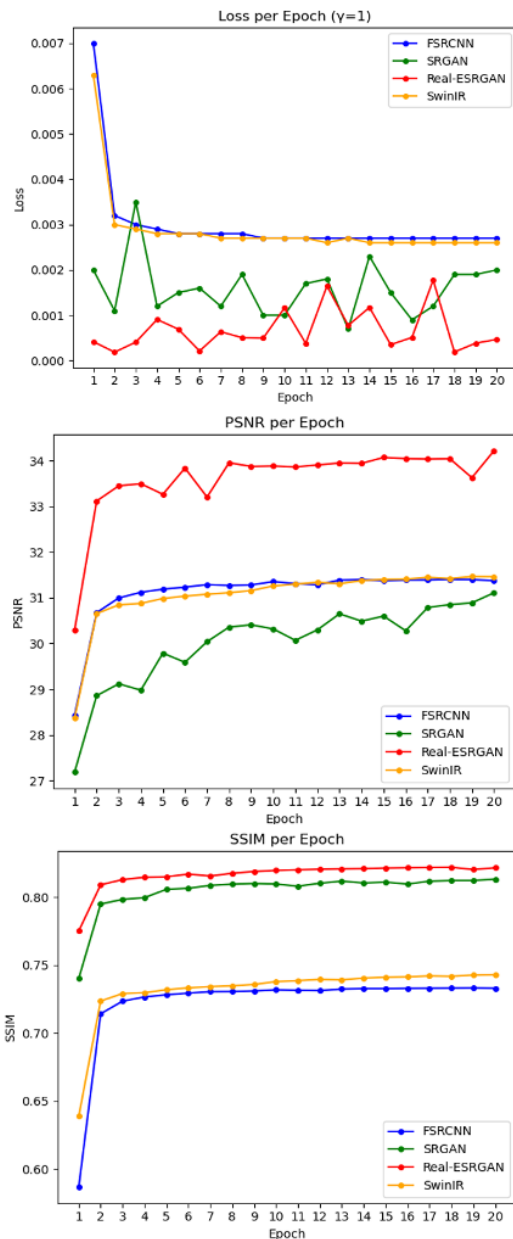


Figure 3. Training performance of super-resolution models per epoch

Model	FSRCNN	SRGAN	Real-ESRGAN	SwinIR
PSNR(dB)	31.46	31.0	34.2	31.6
SSIM	0.73	0.81	0.82	0.75

Table 3. Comparison of SR models using PSNR and SSIM metrics

3.2 Effect of Super-resolution

This section comparatively analyzes the crack width and length measurement results between the original facade orthomosaics and the super-resolved results using the Real-ESRGAN model.

3.2.1 Crack Length Comparison

The accuracy of crack length measurements before and after the application of super-resolution was evaluated by comparing them against the Ground Truth (GT), and the results are presented in Table 4. The GT lengths of the 15 crack samples used in the analysis were calculated based on manually created 2D crack vectors. For the facade orthomosaics before and after SR application, crack vectors were automatically extracted and their lengths were calculated using a crack segmentation model.

The analysis revealed that the average relative error for the SR-applied facade orthomosaic (GSD ≈ 0.1 mm) was 1.93%, which is less than half of the 4.80% error for the non-SR facade orthomosaic (GSD 0.25-0.4 mm). The maximum relative error also showed twice the stability at 5%, compared to 10% for the original image. For instance, crack #6 exhibited a high error of 10% without SR, but this was significantly improved to 2% after SR application. These results suggest that the SR model reduces the variance in crack measurements and provides more consistent results.

Furthermore, the errors in crack lengths calculated from the SR facade orthomosaic showed a balanced characteristic, with a mix of overestimation (negative error) and underestimation (positive error), indicating no significant prediction bias in a particular direction. In contrast, the original facade orthomosaic showed a clear tendency to underestimate the length of most cracks. This is attributed to the fact that the resolution of the original imager was not high enough to adequately capture the detailed structure of the cracks.

Overall, super-resolution (SR) technology effectively restores fine structural details, such as the subtle curvatures of cracks, thereby significantly enhancing the accuracy and stability of crack measurements (Figure 4). Accordingly, the application of SR models to facade orthomosaics can be regarded as a practical and reliable approach to overcome the inherent limitations of conventional low-resolution image analysis.

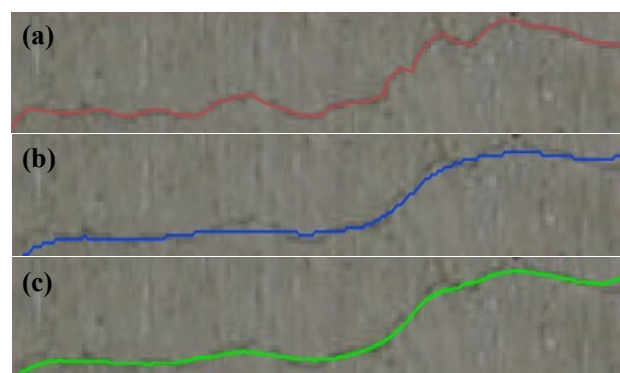


Figure 4. Comparison of Crack length, (a) GT by Crack Segmentation: Length = 64.69 mm, (b) Crack Segmentation in LR Facade Orthomosaic: Length = 62.24 mm, (c) SR Façade Orthomosaic: Length = 63.46 mm

Measurement (mm)		GSD $\approx$ 0.1mm (After SR)			GSD $\approx$ 0.3mm (Before SR)		
No.	Length	Calculation	Absolute error	Relative error	Calculation	Absolute error	Relative error
1	88.2816	87.5711	0.7105	1%	85.9825	2.2991	3%
2	49.6905	47.9821	1.7084	3%	46.2209	3.4696	7%
3	129.7856	123.0473	6.7383	5%	119.0611	10.7245	8%
4	62.049	60.6508	1.3982	2%	58.5354	3.5136	6%
5	123.8035	123.5571	0.2464	0%	121.3312	2.4723	2%
6	145.1841	142.2324	2.9517	2%	130.1144	15.0697	10%
7	126.1787	124.9406	1.2381	1%	120.831	5.3477	4%
8	89.361	87.666	1.695	2%	86.8565	2.5045	3%
9	44.3138	42.8118	1.502	3%	41.3132	3.0006	7%
10	83.0751	83.7291	-0.654	1%	80.6431	2.432	3%
11	153.1216	147.9271	5.1945	3%	145.6694	7.4522	5%
12	106.8712	107.2806	-0.4094	0%	105.5292	1.342	1%
13	62.173	60.2012	1.9718	3%	57.3425	4.8305	8%
14	76.6865	77.4377	-0.7512	1%	75.968	0.7185	1%
15	80.1966	78.7508	1.4458	2%	76.6302	3.5664	4%

Table 4. Comparison of Crack length measurements

Measurement (mm)		GSD $\approx$ 0.1mm (After SR)			GSD $\approx$ 0.3mm (Before SR)		
No.	Width	Calculation	Absolute error	Relative error	Calculation	Absolute error	Relative error
1	0.1	0.1	0	0%	0.33	0.23	230%
2		0.12	-0.02	20%	0.33	0.23	230%
3		0.12	-0.02	20%	0.36	0.26	260%
4		0.1	0	0%	0.77	0.67	670%
5		0.2	-0.1	100%	0.39	0.29	290%
6	0.15	0.12	-0.02	20%	0.42	0.32	320%
7		0.12	0.03	20%	0.36	0.21	140%
8		0.12	0.03	20%	0.6	0.45	300%
9	0.2	0.2	0	0%	0.66	0.46	230%
10		0.2	0	0%	0.6	0.4	200%
11		0.2	0	0%	0.6	0.4	200%
12		0.2	0	0%	0.36	0.16	80%
13	0.3	0.31	-0.01	3%	0.36	0.06	20%
14		0.29	0.01	3%	0.7	0.4	133%
15		0.29	0.01	3%	0.44	0.14	47%
16		0.31	-0.01	3%	0.8	0.5	167%
17	0.4	0.29	0.01	3%	0.66	0.36	120%
18		0.42	-0.02	5%	0.44	0.04	10%
19		0.42	-0.02	5%	0.44	0.04	10%
20	0.45	0.42	0.03	7%	0.44	0.01	2.2%
21		0.42	0.03	7%	0.66	0.21	47%
22	0.5	0.52	-0.02	4%	0.74	0.24	48%
23		0.48	0.02	4%	0.66	0.16	32%
24		0.48	0.02	4%	0.8	0.3	60%
25	0.55	0.52	0.03	5%	0.6	0.05	9%
26	0.6	0.62	-0.02	3%	0.6	0.1	14%

Table 5. Comparison of Crack width measurements

3.2.2 Crack Width Comparison

The impact of applying super-resolution to facade orthomosaics on the accuracy of crack width measurement was quantitatively evaluated, and the results are summarized in Table 5. The GT was based on field measurements taken with a crack scale, while the crack widths from the images before and after SR application were calculated using an automated algorithm and compared with the GT. In the width calculation process, a robust crack outline was extracted by taking the intersection of the Meijering filter and Adaptive Mean binarization. Then, the minimum diameter was calculated through a radial search and converted to a physical unit (mm).

The analysis showed that the average relative error of the crack widths measured from the SR-applied image ($GSD \approx 0.1$ mm) was 10.03%, indicating a high degree of agreement with the GT. In contrast, the average relative error for the original non-SR image ($GSD \approx 0.3$ mm) reached 150.53%, indicating very low reliability of the measurements.

This extreme error stems from the fundamental limitation that the GSD of the original image (≈ 0.3 mm) exceeds the width of the micro-cracks being measured (0.1–0.2 mm). Given that the pixel size is larger than the crack itself, the algorithm fails to resolve the true boundaries, leading to a significant overestimation of the width. It was only when super-resolution (SR) technology enhanced the GSD to approximately 0.1 mm, thereby reducing the pixel size to below that of the crack width, that reliable and accurate measurements became feasible. Consequently, the application of SR reduced the average relative error to approximately one-fifteenth of the original, clearly demonstrating its transformative impact on measurement precision.

This indicates that Super-Resolution (SR) technology extends beyond mere visual enhancement, serving as a robust and practical tool for the quantitative assessment of micro-cracks below regulatory thresholds (e.g., < 0.3 mm) defined by facility safety standards, within the context of UAV-based remote inspection. Consequently, SR is expected to overcome the intrinsic limitations of conventional visual inspection and low-resolution image analysis by delivering objective and high-fidelity data essential for evidence-based infrastructure maintenance and decision-making.

4. Conclusion

This study successfully demonstrated that Super-Resolution (SR) technology can effectively overcome the inherent GSD limitations of UAV-based bridge inspections, enabling high-precision quantitative crack analysis from efficiently captured, lower-resolution imagery. Our comparative analysis identified Real-ESRGAN as the optimal model among CNN, GAN, and Transformer-based architectures, excelling in both quantitative metrics (PSNR/SSIM) and the perceptual quality required for restoring fine crack details.

When applied to a real-world facade orthomosaic, the SR-enhanced imagery ($GSD \approx 0.1$ mm) yielded significant improvements in measurement accuracy. Specifically, the average relative error in crack length measurement was more than halved, decreasing from 4.80% to 1.93%. More dramatically, the SR application corrected the severe overestimation of fine crack widths inherent in the original low-resolution data, reducing the average relative error from a prohibitive 149.11% to a practical 10.03%.

These findings are significant as they validate a practical workflow for bridging the gap between efficient, safe-distance UAV data acquisition and the high-precision data required for structural integrity assessments. This approach allows for the reliable quantification of critical damage metrics (e.g., crack widths under 0.3 mm) without resorting to risky, close-range flights, thereby enhancing the safety, cost-effectiveness, and reliability of digital bridge inspections. While this study provides a strong proof-of-concept on selected samples, future work should focus on scaling this methodology to evaluate full-facade segmentation and on testing the model's robustness under diverse environmental conditions.

Acknowledgement

This research was supported by the Korea Agency for Infrastructure Technology Advancement(KAIA) grant funded by the Ministry of Land, Infrastructure and Transport (Grant RS-2022-00142566).

References

- Dong, C., Loy, C.C., He, K., Tang, X., 2016. FSRCNN: Fast Super-Resolution Convolutional Neural Network. arXiv:1608.00367. doi: 10.1109/CVPR.2016.090
- Feng, C.-Q., Li, B.-L., Liu, Y.-F., Zhang, F., Yue, Y., Fan, J.-S., 2023: Crack assessment using multi-sensor fusion simultaneous localization and mapping (SLAM) and image super-resolution for bridge inspection. *Automation in Construction*, 155, 105047. ISSN 0926-5805, E-ISSN 1872-7891. doi.org/10.1016/j.autcon.2023.105047.
- Kim, J., Shim, S.B., Ahn, J.B., Cho, G.C., 2024: Performance evaluation of concrete crack detection using deep learning-based super-resolution image reconstruction. *Journal of Korean Tunnelling and Underground Space Association*, 26(6), 713-728. ISSN 2233-8292, E-ISSN 2287-4747.
- Korea Authority of Land and Infrastructure Safety (KALIS), 2022: A Study on Pilot Application for Facility Safety Inspection Using Drones. KALIS, OTKCRK221052.
- Ledig, C., Theis, L., Huszar, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A.P., Tejani, A., Totz, J., Shi, W., Wiggers, M., 2017. Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network. arXiv:1609.04802. doi: 10.1109/CVPR.2017.425
- Li, X., Zhang, J., Lu, Y., 2021. SwinIR: Image Restoration Using Swin Transformer. arXiv:2108.10257. doi: 10.1109/CVPR.2021.00728
- Yuan, Y., Gu, S., Fu, X., 2021. Real-ESRGAN: Training Real-World Blind Super-Resolution with Pure Synthetic Data. arXiv:2107.10833. doi: 10.1109/ICCV.2021.00529
- Jeon, E. I., Lee, I., Kim, D. (2023). Crack detection in concrete using deep learning for underground facility safety inspection. *Journal of Korean Tunnelling and Underground Space Association*, 25(6), 555-567. https://doi.org/10.9711/KTAJ.2023.25.6.555