Communication

Base Study of Bridge Inspection by Modeling Touch Information using Light Detection and Ranging

Tomotaka Fukuoka [1], Takahiro Minami [2] and Makoto Fujiu [1,*]

[1] Institute of Transdisciplinary Sciences for Innovation, Kanazawa University, Kanazawa 920-1192, Japan; tfukuoka@se.kanazawa-u.ac.jp

[2] Institute of Technology, Shimizu Corporation, Tokyo 135-0044, Japan; minami_t@shimz.co.jp

[*] Correspondence: fujiu@se.kanazawa-u.ac.jp

Abstract: In Japan, bridges are inspected via close visual examinations every five years. However, these inspections are labor intensive, and a shortage of engineers and budget constraints will restrict such inspections in the future. In recent years, efforts have been made to reduce the labor required for inspections by automating various aspects of the inspection process. In this study, we proposed and evaluated a method of applying super-resolution technology to obtain precise point cloud information to create distance information images to enable the use of tactile information (e.g., human touch) on the surface to be inspected. We measured the distance to the specimen using LiDAR, generated distance information images, performed super-resolution on the pseudo-created low-resolution images, and evaluated them in comparison with the existing magnification method. The evaluation results suggest that the adaptation of the super-resolution technique is effective in increasing the resolution of the boundary of the distance change.

Keywords: bridge maintenance; haptics; point cloud data; LiDAR; 3D modeling; super-resolution

1. Introduction

Japan contains nearly 730,000 bridges with a length of 2 m or more, most of which were constructed during a period of high economic growth. Therefore, the proportion of aged bridges is expected to increase at an accelerated pace in the future, and the strategic maintenance, management, and renewal of this simultaneously aging infrastructure will soon pose an imminent issue [1].

Currently, bridge maintenance and management involve preventive maintenance measures with periodic inspections to prevent minor damages. In 2014, road administrators were obliged to conduct 100% of their surveillance tasks by close visual inspection once every five years, according to the uniform standard set by the government [2]. However, certain bridges could not be thoroughly inspected owing to the shortage of skilled labor and financial resources of the local governments that manage these bridges. In the foreseeable future, continuous maintenance of such infrastructure will be difficult, and, since 2019, inspections have been conducted by close visual inspection or a method that provides equivalent results.

One of the major limitations pertains to the high cost of close visual inspections. Depending on the surrounding environment of the bridge, close visual inspection requires the construction of scaffolding and the use of expensive specialized vehicles, which increases the cost of such inspections. Additionally, the traffic regulations prevailing during inspection can cause economic losses. Therefore, the development of a more economical and versatile inspection method is required.

To develop alternative methods for close visual inspections, Minami et al. verified the feasibility of visual inspection of bridges using images captured by an ultrahigh-resolution camera [3]. An inspection method utilizing an ultrahigh-resolution camera can

acquire detailed images of the bridge sections from a long distance, even on bridges where close visual inspection is difficult. Moreover, this does not require a dedicated vehicle or traffic regulation, which can be expected to reduce inspection costs. In this related study, the results of crack diagnosis using only the obtained images were consistent with those derived from a close visual inspection. Research on crack detection by means of image processing has also been previously conducted [4,5]. In recent years, crack detection methods have been proposed based on deep learning, and their detection accuracy has improved each year [6–8].

Close visual inspection is also conducted with information other than vision as well. For instance, while examining the presence or absence of concrete flaking, the inspector performs a hammering test and assesses the presence or absence of concrete flaking using acoustic information. Thus, research is being conducted to improve the efficiency and accuracy of flaking detection, and certain approaches have been proposed to automate the recording of hammering tests [9] and mounting a hammering mechanism on a drone [10].

This study focuses on touch information, which contains information other than the visual information used by inspectors during inspection. The inspector conducts tactile inspection as required and may examine the damage using tactile sensation. For example, tactile inspection is conducted for diagnosing concrete in bridge piers, using the texture of the concrete surface for these assessments. Thus, this research aimed to model the tactile sensation during bridge inspection by constructing a 3D model of the bridge and modeling the sensations perceived during a tactile inspection. This paper aims to generate a high-resolution 3D model as the basis of this study. Moreover, a 3D model of the target object was constructed to objectively perform the "subjectively conducted" tactile inspection as well as suppress the variations in the inspection results. This study aims to objectify the tactile inspection conducted by an engineer at the inspection site and improve inspection efficiency by creating a detailed 3D model of the target bridge and acquiring the palpation results of the target bridge using the developed tactile sensation model that accurately reproduces the surface conditions of the bridge.

Human haptics can perceive diverse textures owing to the micrometer-level variations in roughness of a target surface [11,12]. Research on the effects of touch, such as attention training using tactile feedback, has also been conducted [13]. Thus, the point cloud density used for 3D model generation should preferably reflect the same order of roughness. However, the current laser surveying instruments record measurements from distances of several tens of meters or more and provide point cloud information spaced in units of millimeters (Figure 1). Also, the smallest unit of coordinate information is generally millimeters. There are surveying instruments that can identify roughness in the order of microns, but these are for surveying very small areas or samples that can be placed on top of the instrument and cannot be used to survey large structures (such as bridges).

Figure 1. Image of model of specimen reproduced with point cloud information (**top**) and partially enlarged image with a sparse point cloud (**bottom**).

To solve this problem, we focused on a method to generate pseudo-dense point cloud information using sparse point cloud information in millimeter units. One such method is densification using distance images. There have been studies on methods to generate dense point cloud information from sparse point cloud information [14], but we consider these methods to be still in the process of research. We thought that there was a possibility to obtain dense point cloud information by converting the point cloud information into a distance image and then converting the result of the distance image into the point cloud information. In this paper, we propose a method of generating pseudo-dense information from sparse information using deep learning. Additionally, this paper describes a method for generating pseudo-information for shooting point distance information, which is information used for calculating point cloud information. Herein, distance information was acquired using Light Detection and Ranging (LiDAR). In principle, LiDAR conducts laser irradiation to calculate the distance between the measuring instrument and the target based on the time required by light to return from the target after reflection.

Our proposed method utilized generative adversarial networks (GANs) [15], i.e., a generative model, to generate dense distance information from sparse distance information. A generative model outputs data similar to those used when training a model.

As shown in Figure 1, point cloud data acquired by LiDAR are sparse. When distance information is calculated from these, point cloud information is generated into an image, and a low-resolution image is generated. One of the existing GAN-based technologies is a super-resolution technology, which generates pseudo-high-resolution images based on low-resolution images. By generating a high-resolution distance information image from a low-resolution distance information image using super-resolution technology, it is considered possible to obtain pseudo-dense distance information (Figure 2). Based on the pseudo-distance information obtained by super-resolution, it is thought that dense point cloud information can be computed. In this paper, we propose and evaluate a method for generating dense distance images from sparse distance images by applying this super-resolution technique.

Figure 2. Example of super-resolution of the pair of images of a raw RGB image and its distance image.

2. Related Research

In recent years, many new inspection methods have been proposed to reduce economic costs and simplify the process of inspections [16–20]. Especially, deep learning-based methods have been studied well. For example, research is being conducted on the automatic diagnosis of buildings and the automatic generation of inspection results [21–23]. Several techniques have been proposed using 3D models for bridge inspection [24,25]. These existing methods utilize laser survey instruments either installed on the ground or mounted on a drone, conduct surveys from multiple measurement points, acquire point cloud information of a bridge, and create a 3D model of the entire bridge. The point cloud information acquired using these instruments can be used to create a 3D model with an accuracy that provides visual confirmation of concrete delamination and cracking. However, the density of the point cloud information may be inadequate for the quantitative evaluation of the texture.

The generated content cannot be easily controlled while generating image data with GAN because model training is conducted using only the authenticity judgment of the generated result as an evaluation index. In this regard, conditional GAN (cGAN) [26,27] has been proposed as a method for controlling images generated with GAN. In this method, the generated data are adjusted by adding information that controls the output result for the input data used in generating the pseudo-data. A type of cGAN called Pix2Pix [28] uses annotated images that are color-coded for each region as input signals for pseudo-image generation. Accordingly, the image generation process is controlled by the regions segmented by this annotation, thereby generating images with distinct features.

As a super-resolution method using deep learning, SRCNN [29] was first proposed by Dong et al. Although this method was a simple model consisting of three convolution layers, it set a state-of-the-art record at the time of publication. Kim et al. proposed VDSR [30], which solves the problems of the training speed and the single resolution of generated images in SRCNN. Dong et al. also proposed FSRNN [31], which solves the computational cost and processing speed problems of SRCNN. As a GAN-based super-resolution method, Ledig et al. proposed SRGAN [32] using a ResNet-based generator, which can generate more precise images than previous methods. Wang et al. improved the generator, classifier, and loss function of SRGAN and proposed ESRGAN [33], which produces more precise images with smaller fluctuations. In this paper, we use Real-ESRGAN [34], which improves on ESRGAN's dataset and model structure to better the generalization performance.

3. Overview of Proposed Method

In this study, a high-resolution distance image is generated using a GAN-based super-resolution method (Figure 3). This image is generated from a low-resolution distance image generated from sparse point cloud information. The accuracy of the generated images is evaluated by distance images generated at different resolutions for the test object.
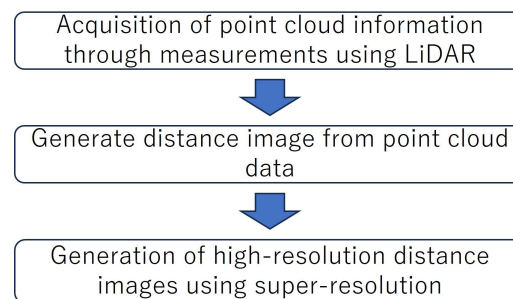

Figure 3. Architecture of the proposed method.

3.1. Deep Learning Model

In this study, Real-ESRGAN is used as the super-resolution method. As one of the well-known super-resolution models, it is a model that is easily adaptable to various types of images. The model is trained by fine-tuning the realesrgan-x4plus model, a generic model of super-resolution. For training the model, the optimization algorithm is adam, the learning rate is 0.0001, the beta parameter is 0.9 to 0.99, the batch size is 12, and the iterator is 20,000. As a comparison, distance images enlarged by bilinear interpolation, a commonly used conventional image enlargement method, are generated and evaluated.

3.2. Dataset

In this study, measurements were made at a distance of approximately 200 mm from the test specimen. An Intel L515 LiDAR camera was used for the measurements, and the measurement distance was an average of 50 acquisitions. As shown in Figure 4, the

specimens to be photographed had large differences in unevenness, with the closest distance measured at approximately 202 mm and the farthest distance at approximately 227 mm. The minimum distance specified in the product specifications is 25 cm. However, in this paper, measurements were conducted at the actual measurable limit distance in order to obtain point cloud data with the highest possible density. In the evaluation, the distance image was a monochrome image, and the pixel values of the distance image represented the acquired distance information in mm. Figure 5 shows the distance image. To train the model, distance images were acquired at a resolution of 1280 pixels in height and 720 pixels in width. Due to the effective range of LiDAR, the upper 160 pixels and the rightmost 30 pixels are where the background is reflected. These areas have a far distance and different features from the specimens and would usually be the noise used to train model. Thus, we deleted these areas from the image data in this paper. The upper 200 × 1250 pixels were divided into an evaluation area and the lower 360 × 1250 pixels into a training area (Figure 6). The image of the training area was further divided into 100 square images of 256 pixels with an overlap of 10 pixels vertically and 100 pixels horizontally. Low-resolution images were generated from the segmented images at 0.75×, 0.5×, and 0.33× magnifications, respectively, for a total of 400 images to be used as the training dataset. In addition, as data for evaluation, low-resolution images with 0.5× and 0.25× magnifications of the data divided into regions for evaluation were created.

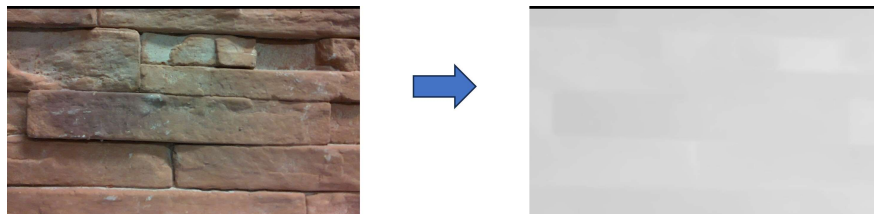**Figure 4.** Image of the test specimen.

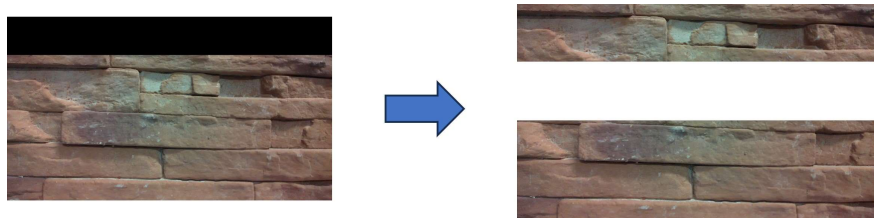**Figure 5.** The results of converting RGB images into distance images.

Figure 6. The results of splitting the image dataset into a test image (upper) and a training image (lower).

3.3. Evaluation Method

This study evaluates the distance image data obtained by super-resolving the test data and the correct distance image data by calculating the root mean squared error

(RMSE) value per pixel. The lower the RMSE calculation result, the smaller the error in the distance information. In our study, the fine-tuned Read-ESRGAN model received test images 200 × 1250 pixels in size and output same-sized distance images.

The result of super-resolving the input image, which was 0.5 times the size of the original evaluation image, two times using Real-ESRGAN was $Real - ESRGAN_{x2}$, and the result of super-resolving the input image, which was 0.25 times the size of the original evaluation image, four times was $Real - ESRGAN_{x4}$. As the baseline, the results of the bilinear interpolation method were denoted as $Baseline_{x2}$ and $Baseline_{x4}$, respectively, for the distance image generated in the same way using the enlarged result. Table 1 shows the results of calculating the RMSE values using the values for each pixel of each image. The images are also shown in Figure 7, in which the pixels which matched the correct distance to the generated distance image are painted in green, those which matched only the results generated by Real-ESRGAN are painted in blue, those which matched only the baseline results are painted in yellow, and those which did not match the results of both methods are painted in gray. Figure 7 also shows an image in which the tones of the distance image are changed with emphasis.

Table 1. Comparison of the delamination detection results.

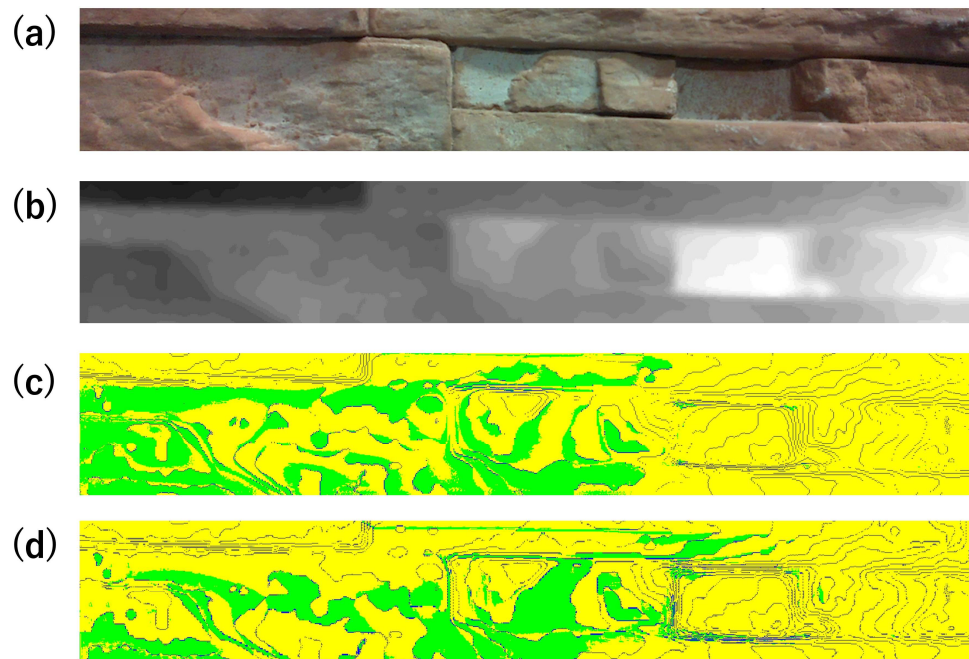|	$Baseline_{x2}$	$Baseline_{x4}$	$Real - ESRGAN_{x2}$	$Real - ESRGAN_{x4}$
RMSE	0.215	0.263	2.085	2.651

Figure 7. Images of the evaluation. (**a**) Input color image; (**b**) input distance image; (**c**) result of ×2; and (**d**) result of ×4.

The results in Table 1 show that the distance image super-resolved using Real-ESRGAN has a larger error than the magnified image using the conventional method. For both the super-resolution and conventional methods, the RMSE value increases with the magnification factor, and the accuracy of the generated distance image decreases. The results in Figure 7 show scattered gray pixels, indicating that the baseline method makes errors at boundaries with varying distances. On the other hand, blue pixels are more prevalent at the boundaries of the distance change, and the distance image generated by Real-ESRGAN is more accurate than the baseline at the boundary pixels where the distance changes. This tendency becomes stronger as the magnification factor increases. The Real-

ESRGAN results also show results that are correct for a coherent range of the same distance, failing to generate certain distances. This may be due to a bias in the number of training data per distance in the model's training data. These results suggest that the super-resolution of distance images using Real-ESRGAN can obtain more accurate distances at the boundaries than conventional methods of enlargement. Even if the model is trained with a small number of training data (about 100 original images), it is possible to obtain a specific distance, suggesting the possibility of highly accurate image generation by expanding the training data.

4. Conclusions

In recent years, the demand for alternative methods of close visual inspections of bridges has increased, and extensive research has been conducted on damage detection methods using image processing. However, only a few studies focused on the tactile inspections used by inspectors. Thus, this study aimed to model the touch information of an inspector and construct a 3D model of the target object to automate and improve the efficiency of tactile inspections. Accordingly, we applied a GAN-based super-resolution method to improve the accuracy of the coordinate point density and the distance information obtained using LiDAR for developing a high-definition 3D model. We performed the evaluation of the accuracy of a super-resolution image generated from a low-resolution input image. The evaluation results show that the super-resolution images are superior in estimating the boundary areas as well as in the performance and limitations of a small-scale learning database.

In future work, we will train the super-resolution model with various kinds of surface condition images to evaluate its generalization performance.

Author Contributions: Conceptualization, T.F. and T.M.; methodology, T.F.; writing—original draft, T.F.; investigation, T.F., T.M. and M.F.; project administration, M.F. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to the fact they are not open data.

Conflicts of Interest: Author Takahiro Minami was employed by the company Shimizu Corporation. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Ministry of Land, Infrastructure, Transport and Tourism. White Paper. 2019; Volume 2020. Available online: http://www.mlit.go.jp/hakusyo/mlit/h30/hakusho/r01/pdf/np202000.pdf (accessed on 2 October 2020).
2. Ministry of Land, Infrastructure, Transport and Tourism. Road Bridge Periodic Inspection Procedures, Road Bureau. 2022. Available online: https://www.mlit.go.jp/road/sisaku/yobohozen/tenken/yobo4_1.pdf (accessed on 30 July 2023).
3. Minami, T.; Fujiu, M.; Takayama, J.; Suda, S.; Okumura, S.; Watanabe, K. A basic study on diagnostic imaging technology for bridge inspection using super high resolution camera. *Sociotechnica* **2018**, *15*, 54–64.
4. Yamaguchi, T.; Nakamura, S.; Saegusa, R.; Hashimoto, S. Image-based crack detection for real concrete surfaces. *IEEJ Trans. Elec. Electron.* **2008**, *3*, 128–135. https://doi.org/10.1002/tee.20244.
5. Nguyen, H.-N.; Kam, T.-Y.; Cheng, P.-Y. An automatic approach for accurate edge detection of concrete crack utilizing 2D geometric features of crack. *J. Signl Process. Syst.* **2014**, *77*, 221–240. https://doi.org/10.1007/s11265-013-0813-8.
6. Chun, P.-J.; Igo, A. Crack detection from image using Random Forest. *J. Jpn Soc. Civ. Eng. F3*. **2015**, *71*, I-1–I-8.
7. Yokoyama, S.; Matsumoto, T. Development of an automatic detector of cracks in concrete using machine learning. *Procedia Eng.* **2017**, *171*, 1250–1255. https://doi.org/10.1016/j.proeng.2017.01.418.
8. Cha, Y.-J.; Choi, W.; Büyüköztürk, O. Deep learning-based crack damage detection using convolutional neural networks. *Comput. Aid. Civ. Infrastruct. Eng.* **2017**, *32*, 361–378. https://doi.org/10.1111/mice.12263.

9. Watanabe, A.; Even, J.; Morales, L.Y.; Ishi, C.T. Robot-assisted hammer sounding inspection of infrastructures. *J. Robot. Soc. Jpn.* **2015**, *33*, 548–554. https://doi.org/10.7210/jrsj.33.548.

10. Miura, T.; Nitta, M.; Wada, H.; Nakamura, H. The Development of Inner Defect Detecting Method by Using UAV Having Hammering Mechanism and Application to Actual Bridges. *J. Struct. Eng. A* **2019**, *65*, 607–614.

11. Watanabe, S.; Ozaki, K.; Yamazaki, Y.; Yamamoto, S. Recognition and language estimation of fine particles through tactile sensing with fingers. *J. Jpn Soc. Precis. Eng. Contrib. Pap.* **2005**, *71*, 1421–1425. https://doi.org/10.2493/jspe.71.1421.

12. Chiba, T.; Kuroda, S.; Yamaguchi, M. Evaluating method for tactile sensations of leathers using multi-physical properties. *J. Life Support Eng.* **2018**, *30*, 44–50. https://doi.org/10.5136/lifesupport.30.44.

13. Wang, D.; Li, T.; Afzal, N.; Zhang, J.; Zhang, Y. Haptics-mediated approaches for enhancing sustained attention: Framework and challenges. *Sci. China Inf. Sci.* **2019**, *62*, 1869–1919.

14. Tianchang, S.; Jun, G.; Kangxue, Y.; Ming-Yu, L.; Sanja, F. Deep Marching Tetrahedra: A Hybrid Representation for High-Resolution 3D Shape Synthesis. *Adv. Neural Inf. Process. Syst.* **2021**, *34*, 6087–6101.

15. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial nets. *Adv. Neural Inf. Process. Syst.* **2014**, *27*, 2672–2680.

16. Menendez, E.; Victores, J.G.; Montero, R.; Martínez, S.; Balaguer, C. Tunnel structural inspection and assessment using an autonomous robotic system. *Autom. Constr.* **2018**, *87*, 117–126.

17. Pham, N.H.; La, H.M.; Ha, Q.P.; Dang, S.N.; Vo, A.H.; Dinh, Q.H. Visual and 3D mapping for steel bridge inspection using a climbing robot. In Proceedings of the ISARC 2016—33rd International Symposium on Automation and Robotics in Construction, Auburn, AL, USA, 18–21 July 2016; pp. 141–149.

18. Xie, R.; Yao, J.; Liu, K.; Lu, X.; Liu, Y.; Xia, M.; Zeng, Q. Automatic multi-image stitching for concrete bridge inspection by combining point and line features. *Autom. Constr.* **2018**, *90*, 265–280.

19. Ivanovic, A.; Markovic, L.; Car, M.; Duvnjak, I.; Orsag, M. Towards Autonomous Bridge Inspection: Sensor Mounting Using Aerial Manipulators. *Appl. Sci.* **2021**, *11*, 8279.

20. Caballero, R.; Parra, J.; Trujillo, M.Á.; Pérez-Grau, F.J.; Viguria, A.; Ollero, A. Aerial Robotic Solution for Detailed Inspection of Viaducts. *Appl. Sci.* **2021**, *11*, 8404.

21. Rui, Z.; Ruqiang, Y.; Zhenghua, C.; Kezhi, M.; Peng, W.; Robert, X.G. Deep learning and its applications to machine health monitoring. *Mech. Syst. Signal Process.* **2019**, *115*, 213–237.

22. Soleimani-Babakamali, M.H.; Esteghamati, M.Z. Estimating seismic demand models of a building inventory from nonlinear static analysis using deep learning methods. *Eng. Struct.* **2022**, *266*, 114576.

23. Zambon, I.; Vidović, A.; Strauss, A.; Matos, J. Condition prediction of existing concrete bridges as a combination of visual inspection and analytical models of deterioration. *Appl. Sci.* **2019**, *9*, 148.

24. Bolourian, N.; Soltani, M.M.; Albahria, A.H.; Hammad, A. High level framework for bridge inspection using LiDAR-equipped UAV. In Proceedings of the International Symposium on Automation and Robotics in Construction 2017—ISARC, Taipei, Taiwan, 28 June–1 July 2017; Volume 34.

25. Jung, S.; Song, S.; Kim, S.; Park, J.; Her, J.; Roh, K.; Myung, H. Toward Autonomous Bridge Inspection: A framework and experimental results. In Proceedings of the 2019 16th International Conference on Ubiquitous Robots (UR), Jeju, Republic of Korea, 24–27 June 2019; pp. 208–211.

26. Mirza, M.; Osindero, S. Conditional Generative Adversarial Nets. *arXiv* **2014**, arxiv:1411.1784,

27. Reed, S.; Akata, Z.; Yan, X.; Logeswaran, L.; Schiele, B.; Lee, H. Generative Adversarial Text to Image Synthesis. In Proceedings of the International Conference on Machine Learning, New York, NY, USA, 20–22 June 2016.

28. Isola, P.; Zhu, J.; Zhou, T.; Efros, A.A. Image-to-Image Translation With Conditional Adversarial Networks. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017. https://doi.org/10.1109/CVPR.2017.632.

29. Chao, D.; Loy, C.C.; He, K.; Tang, X. Learning a Deep Convolutional Network for Image Super-Resolution. In *Computer Vision—ECCV 2014*, Springer International Publishing: Cham, Switzerland, 2014; pp 184–199.

30. Kim, J.; Lee, J.K.; Lee, K.M. Accurate Image Super-Resolution Using Very Deep Convolutional Networks. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 26 June–1 July 2016; pp. 1646–1654.

31. Chao, D.; Loy, C.C.; Tang, X. Accelerating the Super-Resolution Convolutional Neural Network. In *Computer Vision—ECCV 2016*; Springer International Publishing: Cham, Switzerland, 2016; pp. 391–407. .

32. Ledig, C.; Theis, L.; Huszar, F.; Caballero, J.; Cunningham, A.; Acosta, A.; Aitken, A.; Tejani, A.; Totz, J.; Wang, Z.; et al. Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network. In Proceedings of the Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017.

33. Wang, X.; Yu, K.; Wu, S.; Gu, J.; Liu, Y.; Dong, C.; Qiao, Y.; Loy, C.C. ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks. *arXiv* **2018**. https://doi.org/10.48550/arXiv.1809.00219.

34. Wang, X.; Xie, L.; Dong, C.; Shan, Y. Real-ESRGAN: Training Real-World Blind Super-Resolution with Pure Synthetic Data. In Proceedings of the International Conference on Computer Vision Workshops (ICCVW), Montreal, BC, Canada, 11–17 October 2021.