



Integrated 3D Structural Element and Damage Identification: Dataset and Benchmarking

Katrina Montes¹(✉), Molan Zhang², Jiaming Liu¹, Lama Hajmoussa¹,
ZhiQiang Chen², and Ji Dang¹(✉)

¹ Saitama University, Saitama, Japan
dangji@mail.saitama-u.ac.jp

² University of Missouri-Kansas City, Kansas, USA

Abstract. Traditional bridge inspection is a manually performed visual process that is time-consuming, costly, and requires significant support from equipment and resources. Recent advances in artificial intelligence (AI) have accelerated the advances toward inspection automation enabled by robotic imaging and machine vision. Prolific studies in recent literature have reported that damage types and their locations can be identified in 2-dimensional (2D) images with complex scenes using deep learning (DL) based techniques (e.g., semantic object detection or segmentation). However, these efforts have not achieved a level of applicability as meaningful as those from traditional human-based bridge inspection. To enable practical applicability, structural elements and damage patterns must be identified (including detection, localization, and quantification) in a 3D space. To this end, one significant research challenge is the lack of 3D databases for learning both structural elements and damage patterns. This study first developed a unique 3D dataset based on a real bridge structure, and a low-cost LiDAR-enabled imaging device (Intel RealSense) was adopted during the data collection process. Semantic annotations were added for structural elements and damage based on the point clouds and the associated RGB imagery data. Furthermore, a DL-based method was developed to benchmark the usefulness and validity of the dataset. The proposed dataset and DL methods will be open-sourced and expected to facilitate the advances toward engineering inspection automation for bridge structures.

Keywords: structural damage · inspection automation · Machine Vision · Point Cloud · Images

1 Introduction

By 2023, approximately 43% out of about 700,000 bridges in Japan will exceed 50 years of service which will require periodic inspection and maintenance [1]. Aside from that, in 2019, around 42% out of 617,000 bridges across the United States have been put into service for at least 50 years and 7.5% are in poor condition [2]. The continued aging can render bridges more vulnerable when resisting extreme events, and more likely to impact human safety and economy. Conventional visual bridge inspection requires a lot of manpower, equipment, and time. Thus, it is important to have more efficient ways of inspection and maintenance for infrastructures such as bridges.

Researchers have proposed to use advanced to improve the bridge inspection process. Machine vision based deep learning (DL) methods are among recent advances, which attract much attention from both researchers and practitioners. To realize damage segmentation and classification, DL architectures based on convolutional neural networks (CNN) were proved to be promising. It is noted that most of these efforts adopt a data-centric approach by using, tuning, or modifying existing architectures. For instances, Mask R-CNN model was used to segment multiple damages like corrosion, cracks, spalling, etc. on bridges using a large dataset of bridge inspection reports [3]. YOLO v3 was applied to classify different bridge damages like cracks and corrosion [4]. For specific bridge components, such as rubber bearings, VGG-Unet model was proposed to segment the cracks on the rubber cover [5]. Aside from the detection of damages, damage quantification was also proposed by estimating the width of structural cracks using deep learning [6].

For bridge component recognition, different convolutional neural network models were proposed to segment bridge components against complex scene in images. Structural components such as columns, beams and slabs, other structural members, and nonstructural components were considered [7]. Due to the limited availability of real data, synthetic data was simulated, and the road bridges components and damages were extracted and annotated. After that, a CNN model was proposed to segment the components and bridge damages [8]. A continued study of [8] proposed a method for bridge component segmentation using the images collected by a UAV. The images were used to reconstruct a point cloud data then the components were categorized [9]. The bridge component using 3-dimensional data has an extra dimension that can be useful for further assessment that's why this study trained a CNN model that can segment the bridge component from a point cloud data collected by a laser scanner. The points in point cloud data were classified into three categories: deck, pier, and background [10].

In order to visualize the damage and its location on the bridge, 3D model reconstruction using Structure from Motion (SfM) and other techniques are useful tools. Inadomi and Chun proposed a method to convert the point cloud data of a bridge into 2D images to segment the components using DeepLab v3+ model, then the segmented components can be reflected to the original point cloud model [11]. Yamane and Chun furtherly improved the method by introducing deep learning methods to detect damages in 2D image and then projected them back the 3D model [12].

In our previous effort, we proposed to use DeepLab v3+ to segment the corrosion from the RGB images of a steel bridge [13]. Then, the segmented corrosion damages were visualized into the 3D bridge model. The feature points of the damages were projected into a 3D bridge model reconstructed using structure from motion to know the location of the damages. In addition to that, the 3D model was saved and can be viewed remotely through mixed reality platform. However, the corrosion was segmented using the 2D images then there is a need to conduct SfM to reconstruct the 3D bridge model which takes time depending on the hardware used.

It seems that the component detection and the damage segmentations were conducted in separated works and for different purpose. However, it is important for the stage of AI and 3D damage detection and diagnose to let AI understand the relationship between damage and its location in the bridge to enhance its ability to further understand the

reason of damage and diagnose the cause and propose the countermeasure for them. The bridge structural component and existing damages have a relationship which can help in the evaluation and prediction of the cause or progression of the damage. In addition to that, the usage of 3-dimensional data contains additional information such as component surfaces and the damage continuation patterns which are highly significant. However, one of the main challenges is how to incorporate and know the location of the damage to the bridge component. In addition to that, segmenting the bridge component and existing damages in a three-dimensional environment have lack of available real-world 3-dimensional bridge data.

Therefore, this study aims to incorporate the segmented bridge damages to the specific bridge component for further determination of the damage cause and diagnosis. Using a low-cost lidar device, RGBD data was collected during the onsite bridge inspection and annotated using two open-sourced software's. A CNN model was proposed to do the point cloud bridge component segmentation and RGB damage segmentation. The raw and annotated point cloud dataset of the bridge component and RGB damages is open accessed.

2 Proposed 3D Damage and Component Segmentation Method

This study created a unique 3-dimensional dataset based on a real bridge structure using a low-cost LiDAR-enabled imaging device (Intel RealSense) and proposed a CNN model for segmentation of bridge components in 3-dimension and damage segmentation in RGB images as shown in Fig. 1. Semantic annotation for the structural elements using the point-cloud data through an open-sourced software were conducted. The damages were also annotated using the RGB image data. Furthermore, a deep learning method was established as a benchmark model to validate the dataset. The proposed dataset and DL methods are open-sourced and expected to facilitate the advances toward engineering inspection automation for bridge structures.

As a future study as shown in Fig. 2, after the creation of point cloud dataset and training of the deep learning model for component and damage segmentation. The trained weights will be used to deploy in smartphones applications, UAV onboard processing, and HoloLens for the segmentation of damages and component in the actual bridge inspection. The data collected will be transfer through cloud which will be used for the damage diagnosis and bridge 3D model reconstruction for the whole bridge considering the original data and the data with segmented damages and components. The damage diagnosis report will consist of the type of component, damage type, and damage location, cause of damage, and further evaluation. The 3D model can be saved and repeatedly done through periodic inspection to compare the deterioration of the bridge through time.

3 Proposed 3D Damage and Component Segmentation Method

The bridge as shown in Fig. 3, is a concrete pedestrian bridge located inside Saitama University campus in Japan. A low-cost Lidar device as shown in Fig. 4 was used to collect the bridge point cloud, RGB image, and depth data as shown in Fig. 5. The dataset for bridge components segmentation focuses on beam, column, transverse girder, and

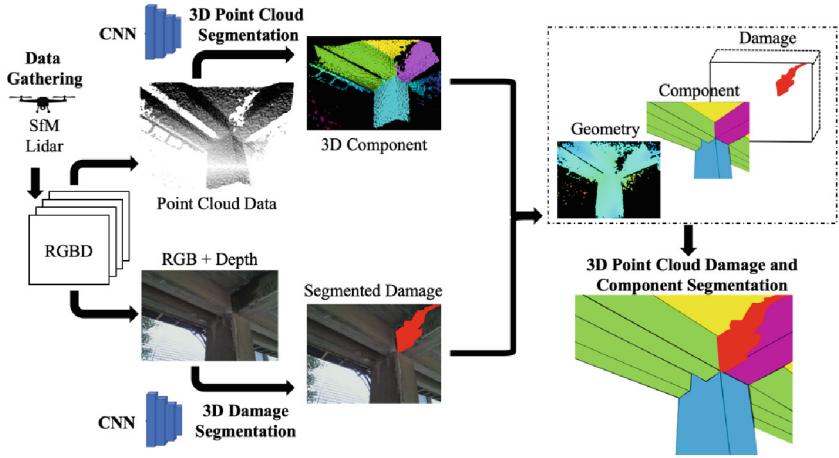


Fig. 1. Proposed Methodology.

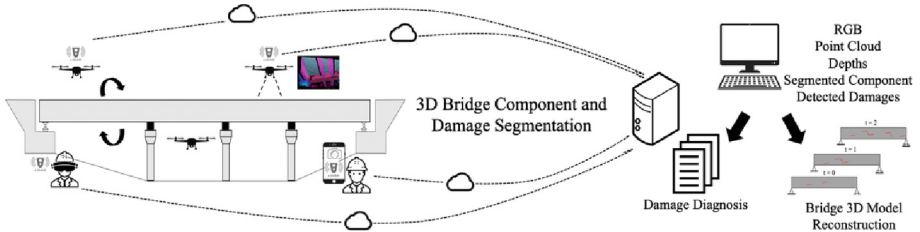


Fig. 2. Proposed bridge inspection automation and report generation.

main girder, while the damages include corrosion, spalling, cracks, and leaking water. Using the Intel® RealSense™ Lidar camera, both RGB image and depth image can be captured at the same time. The resolution of RGB image is 640 by 480 pixels, while the depth image size is 320 by 240 pixels. Some challenges in gathering the data include the lighting condition, distance of the Lidar from the surface, and the quality of the point cloud data. The point cloud annotation and RGB annotation are all in json file.



Fig. 3. Bridge site and structure.

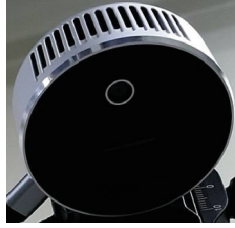


Fig. 4. Intel® RealSense™ Lidar Camera L515.

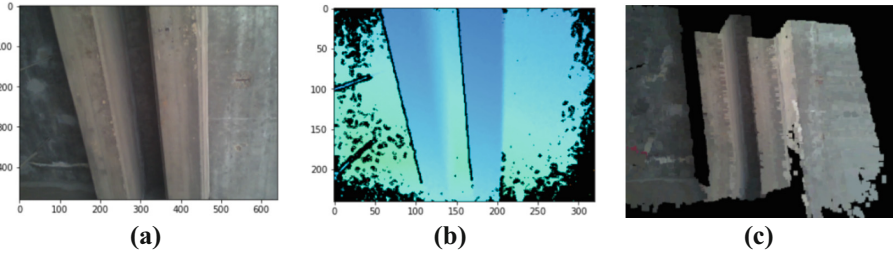


Fig. 5. Captured 3D data: (a) RGB Image, (b) Depth Image, and (c) Point Cloud Data

The open-source VGG image annotator [14] was used to annotate the damages using the RGB images. A sample annotation is shown in Fig. 6, in which corrosion damage is labeled. In our dataset, the damage categories include cracking, corrosion, spalling, and moisture marks (the latter creates an adverse condition enabling potential damage).



Fig. 6. Damage annotation in color images: (a) color image and (b) annotated image with damage.

The bridge component point cloud annotation was conducted using an open-sourced software named “supervisely” [15]. The structural component class names are “main girder”, “transverse girder”, “deck” and “column”. The point cloud annotation result can be visualized as shown in Fig. 7, in where the components were marked by different colors: main girder as violet, transvers girder as green, column as blue, and deck as yellow. The final annotation output is in “json” format prior to training.

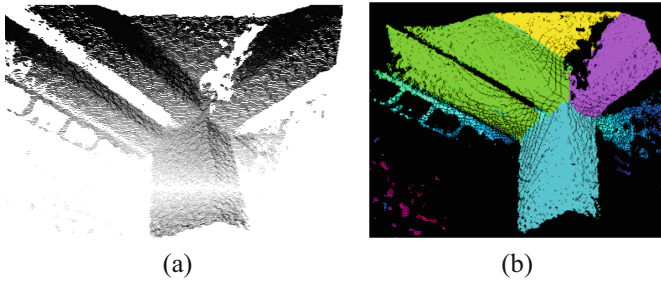


Fig. 7. Point Cloud Bridge Component Annotation (a) RGB (b) Annotation

4 Enhanced 3D GNN for Bridge Damage and Component Segmentation

Representing data in 3D is becoming increasingly important in computer vision. In recent years, more and more point cloud has been used to show 3D data. For the color point cloud database, we have considered its practicability in multiple aspects. Here, the author proposes a semantic segmentation network based on 3D data points to verify the practical significance and usability of our database.

In computer vision, the task of semantic segmentation is to segment images or point clouds and distinguish different objects. When semantic segmentation is used, it divides an image or point cloud into semantically meaningful parts and then semantically labels for each part as one of the predefined classes. Identifying objects within different point clouds or image data is useful in many applications. However, more difficulties are encountered when performing semantic segmentation on point cloud data than in semantic segmentation of 2D images.

A big challenge is the sparseness of points between point clouds, which makes it possible to see through objects. This makes it difficult to see the structure in the point cloud and distinguish which object a point belongs to. In order to deal with this challenge, this paper chooses the 3D GNN network [16] as the basic model, which is an end-to-end 3D graph neural network that can learn representations directly from 3D point clouds. On this basis, we added Transform Network to the network and proposed a chain training method. Here, Fig. 8 shows the structure of the Enhanced 3D GNN network:

In this network, we use the chain method of training, because some disaster types will only appear on specific components, such as spalling only appearing on columns. We hypothesize that mentioning component type as one of the inputs to damage type detection part can provide such meaningful information that can improve model performance. The component map needs to be obtained from the first step of the network. We all know that in the real world, buildings or bridges have their specific geometric structure, which plays a vital role in determining the component type, and the best way to express the geometric structure and spatial position is 3D point cloud information.

Therefore, we first take the colored point cloud image as the input of the SubNet1, using the location information to build a directed graph. Here we treat each pixel as a node and connect it to KNN in 3D space through directed edges. After constructing the graph, we use the color information in the color point cloud and transformed CNN to

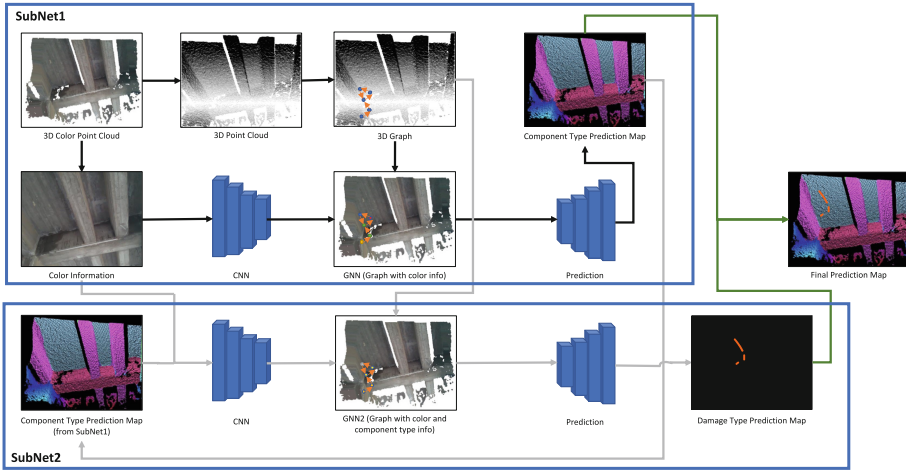


Fig. 8. Enhanced 3D GNN network structure

calculate the feature of each pixel as the initial information of the corresponding node, so as to ensure that the graph neural network can utilize texture and geometric information at the same time. The output of the network in the first step will be the component prediction map corresponding to the input.

The second step is to re-assign initial information and weights to each node based on the results of the previous sub-network and the constructed graph network. Here we combine the color information and element types of nodes and perform initial convolution through CNN and use this as the initial node information of the second network for training. The output of the network in this step will be the damage type prediction map of the input image.

In summary, the Enhanced 3D GNN network we proposed contains two sub-networks, which share directed graphs with each other, but have different initial information, and can output different results for different applications.

For the network model we proposed, we used it as the benchmark model of the database for experiments. First, we used 70% of the database as training data and 30% as validation data. After 100 epoch training, we obtained Tables 1 and 2 as our validation results. Through the analysis of experimental results, it is not difficult to see that the model has relatively excellent performance for small sample databases and can successfully ignore noise and accurately identify most components. Of course, since there are too few training samples of the damage type in this database, like there are only two cracks, the prediction result of the damage type is not ideal even if the transform layer model is used. But even so, it is not difficult to see that with enough data, such as spalling data, the model can still successfully identify the type of disaster at that location.

Table 1. Validation results for component and damage type detection.

Component Type	IoU %	Accuracy %	Damage Type	IoU %	Accuracy %
Deck	69.498	89.781	Crack	0	99.531
Transverse Girder	63.347	78.908	Corrosion	0	98.802
Main Girder	37.440	86.484	Leaking Water	78.455	91.567
Column	78.655	97.900	Spalling	57.483	85.749
Total Mean	62.235	88.265	Total Mean	33.984	93.914

Table 2. Validation confusion matrix for component and damage type prediction.

Prediction Label \	Deck	Transverse Girder	Main Girder	Column
Deck	2604725	449267	51876	3215
Transverse Girder	488386	3999856	337310	64976
Main Girder	117533	894973	871795	69558
Column	2006	59874	29155	838634
Prediction Label \	Crack	Corrosion	Leaking Water	Spalling
Crack	2	9	316	10
Corrosion	3	8	767	89
Leaking Water	0	2	51009	5005
Spalling	0	0	5185	9821

5 Conclusion

This study created an open-accessed 3D annotated dataset of a real bridge component and damages which can be used for future segmentation training. A low cost lidar was used to gather the RGB, depth, and converted point cloud data of the bridge components and damages. The dataset was published and can be access publicly for future 3D segmentation training.

The proposed benchmark model for 3D semantic segmentation achieved a relatively good results given the limitation of the training data. The mean IoU for component type

is 62% and 34% for the damages. Increasing the database might improve the accuracy and IoU.

The segmentation of bridge component and identification of the existing damages on the specific component surfaces can greatly help to further diagnose the cause of damage and the relationship of the nearby damages. Future studies include the bridge 3D model reconstruction and damage evaluation.

References

1. MLIT (Ministry of Land, Infrastructure, Transport and Tourism in Japan) White Paper, Section 2, Paragraph 2 (2014)
2. 2021 Report Card for America's Infrastructure, ASCE (2021). <https://infrastructurereportcard.org/at-item/bridges-infrastructure>
3. Taher, S., Dang, J.: Autonomous Multiple Damage Detection and Segmentation in Structures Using Mask R-CNN. Experimental Vibration Analysis for Civil Engineering Structures (2020). https://doi.org/10.1007/978-3-030-93236-7_45
4. Dang, J., Mizumoto, T., Chun, P., Liu, J., Fujishima, T.: Multi-Type Bridge Damage Detection Method Based on YOLO. Japan Society of Civil Engineering (2021). https://doi.org/10.11532/jsceiii.2.J2_447
5. Shi, J., Dang, J., Cui, M., Zuo, R., Shimizu, K.: Improvement of damage segmentation based on pixel-level data balance using VGG-Unet. Appl. Sci. **11**, 518.1–17 (2021). <https://doi.org/10.3390/app11020518>
6. Ni, F., Zhang, J., Chen, Z.: Zernike-moment measurement of thin-crack width in images enabled by dual-scale deep learning. Comput. Aided Civil Infrastruct. Eng. **34** (2018). <https://doi.org/10.1111/mice.12421>
7. Narazaki, Y., Hoskere, V., Hoang, T.A., Spencer Jr., B.F.: Vision-based automated bridge component recognition integrated with high-level scene understanding. In: The 13th International Workshop on Advanced Smart Materials and Smart Structures Technology (2017)
8. Narazaki, Y., Hoskere, V., Spencer Jr., B.F.: Synthetic environments for facilitating field implementation of vision-based structural management techniques. J. Artif. Intell. Data Sci. **2**(J2), 20–28 (2021). https://doi.org/10.11532/jsceiii.2.J2_20
9. Kim, H., Narazaki, Y., Spencer Jr., B.F.: Automated bridge component recognition using close-range images from unmanned aerial vehicles. En. Struct. **274** (2023)
10. Kim, H., Yoon, J., Sim, S.: Automated bridge component recognition from point clouds using deep learning. Struct. Control Health Monitor. **27**(9) (2022)
11. Inadomi, S., Chun, P.: Segmentation of bridge point clouds using point cloud imaging and deep learning. J. Artif. Intell. Data Sci. **2**(J2), 418–427 (2021)
12. Yamane, T., Chun, P., Honda, R.: Detecting and localizing damage based on image recognition and structure from motion, and reflecting it in a 3D bridge model. Struct. Infrastruct. Eng. (2022). <https://doi.org/10.1080/15732479.2022.2131845>
13. Montes, K., Dang, J., Liu, J., Chun, P.: Bridge status realization and management enhanced by UAV, SfM, and deep learning. European Workshop in Structural Health Monitoring (2023). <https://doi.org/10.1107/978-3-031-07258-1/55>
14. VGG Image Annotator. <https://www.robots.ox.ac.uk/~vgg/software/via/> last accessed in December (2022)
15. Supervisely. <https://supervise.ly>. Accessed Dec (2022)
16. Qi, X., Liao, R., Jia, J., Fidler, S., Urtasun, R.: 3d Graph Neural Networks for RGBD Semantic Segmentation. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 5199–5208 (2017)