Semantic Segmentation of Real and Synthetic Point Cloud Data for Digital Twinning of Bridges

M. Saeed Mafipour[1], Cagatay Alici[1], S. Saadat Shakeel[1] and Affan Kalkavan[1]
[1]Chair of Computational Modeling and Simulation, Technical University of Munich, Arcisstraße 21, 80333 Munich, Germany
E-mail(s): m.saeed.mafipour@tum.de, cagatay.alici@tum.de, saadat.shakeel@tum.de, affan.kalkavan@tum.de

Abstract: Digital twins (DT) can decrease the maintenance costs of bridges through a 3D geometric-semantic model representing the current status of the structure. The 3D model of a DT is generally created from point cloud data (PCD) captured by laser scanning or photogrammetry. Semantic segmentation is an essential step in processing PCD and 3D bridge modeling. Deep learning models are efficient tools to automate semantic segmentation. These models, however, need a large dataset for training which is challenging to achieve for bridges. This paper proposes an approach to generate synthetic PCD of bridges for training deep learning models. The parametric model of bridges is created, and their synthetic PCD is simulated. This dataset is then added to the real dataset of bridges for training a deep learning model. The paper results show that the synthetic PCD can be used for data augmentation and improving the performance of deep learning models.

Keywords: Digital Twins, Point Cloud Data, Semantic Segmentation, Synthetic, Deep Learning

1 Introduction

Bridges, as critical structures, require maintenance during their service life. Recent ASCE report card [1] shows that the number of deficient bridges is increasing as the rate of deterioration is higher than rehabilitation and replacement. Despite the feasibility of the conventional methods for evaluation and condition assessment of bridges, they are loosely and only partially supported by digital methods. Digitization of existing bridges can support the inspection and maintenance process of the structure and decrease the applying costs to a large extent.

The concept of digital twin (DT), originated from the industry [2], is concerned with providing the digital counterpart of a system or object. In the domain of building information modeling (BIM), a digital twin represents the digital replica of an existing structure [3]. This 3D model can incorporate all the collected information from the construction site. This information can include the current geometry of the structure, cracks and their location on the body of the structure. In comparison with a building information model, a DT is largely concerned with the as-performed and as-is phases of the structure [4]. Also, it illustrates the interaction of humans with the structure and can be updated in specific intervals. In bridges, these intervals might be longer as the physical features of the asset vary gradually [5]. A digital twin is generally created from the point cloud data (PCD) resulting from terrestrial and aerial capturing methods. Laser scanning and photogrammetry are two common methods that can capture existing assets with high measurement accuracy [6], [7]. Compared to a visual inspection, these methods are faster and less labor-intensive. To create the DT model of a bridge, its PCD needs to be processed. Semantic segmentation is an essential step in processing PCD. Through semantic segmentation, the entire PCD of the bridge is divided to clusters representing the point cloud of elements. In the current practice, this step is conducted manually that in turn, increases the costs of digital twinning to a large extent. To benefit from the advantages of a DT, semantic segmentation is required to be automated or at least semi-automated.

2 Related research

Recently, there have been efforts to automate the point cloud segmentation of bridges. The proposed methods can be divided to heuristic algorithms and deep learning models. In what follows, some of these methods are reviewed.

Lu, Brilakis, and Middleton [8] detected elements in the point cloud of concrete bridges by a heuristic algorithm following the relative distance of points in the point clusters of elements. Hu, Zhao, Huang, *et al.* [9] used a multi-view convolutional neural network (CNN) to extract features from photogrammetry and linked it with a multi-layer perceptron (MLP) to segment the point cloud of bridges. Lee, Park, and Ryu [10] added contextual features by kd-tree and K-nearest neighbors (KNN) to PointNet [11] and deep graph-convolutional neural network (DGCNN) [12] and improved the performance of the models. Yan and Hajjar [13] segmented the point cloud of steel bridges by a heuristic algorithm based on the connection rules that generally exist in the bridges. Truong-Hong and Lindenbergh [14] proposed a heuristic algorithm to segment elements in the point cloud of bridges by a voxel growing algorithm. Xia, Yang, and Chen [15] calculated a local feature descriptor for classifying points in the PCD of bridges.

3 Overview of the paper

Heuristic algorithms are capable of labeling the input PCD of bridges without training. These algorithms, however, are mostly limited to presumptions and conditions that might not be satisfied in all bridges. On the other side, deep learning models are highly generic and can learn the features of points. These models, however, need a large dataset for training. Despite the recent advances in technologies such as laser scanning and photogrammetry, the capturing process of existing bridges is still challenging and time-consuming. Also, deep models generally need more than a few samples of PCD for training. The main objective of this paper is to provide a workflow for simulation of bridges and generating their PCD. For this purpose, the 3D models of bridges are created in a BIM-authoring system. These models are then used as input in Helios++ [16] to generate their PCD. RGB channels are also added by interpolating the points of the real PCD. First, the model is trained only by using real PCD. Next, to evaluate the augmentation impact of synthetic data, the model is trained with both real and synthetic PCD. Finally, the statistical metrics obtained from the models with and without augmentation are compared.

4 Synthetic point cloud data

Synthetic PCD of bridges can be created and used to augment the dataset for training the deep learning models. Generating synthetic PCD requires modeling and simulating the existing bridges along with the capturing process by scanning tools. In what follows the required steps to create synthetic PCD are described.

4.1 Parametric modeling

Parametric modeling is a computer-aided design (CAD) approach for creating dynamic models. As a result of parametric modeling, the model of bridges achieves the capability of reshaping as the value of parameters change. To generate the synthetic PCD of bridges, the parametric model of bridges is created on Revit [17] using a plugin named SOFiSTiK Bridge Modeler [18]. To provide more realistic bridge models, the dimension of structural elements are considered following the structural drawings of actual bridges. Each part of the bridge, including the surrounding background, is designed to resemble the real PCD of bridges. The dimensions of the components, number of spans, and the shape of elements are varied to increase the diversity of the parametric models. Trees with dense vegetation are also modeled around the structure, as shown in Figure 1.

4.2 Simulation

To create the PCD of the modeled bridges, the laser scanning process of actual bridges is simulated by Helios++ [16]. In this software, tripods are placed around the structure to ensure the full coverage of the models, as shown in Figure 1. The scanning process of the bridge models is conducted by rotating optics with 120Hz scanning frequency, 5000Hz pulse frequency, 180° scanning angle, and 10 head rotate per sec/deg.

To automate the annotation process of the synthetic PCD, the 3D model of each structural element, including piers, railings, deck, and background, is scanned separately. This process results in the exact labeling of synthetic PCD; however, some regions that cannot be generally captured in real bridges are scanned as well. To simulate occlusion that generally exist in the PCD of real bridges, these regions are cropped out and deleted. To add RGB channels to the synthetic PCD, the elements of real PCD are selected and aligned with the corresponding elements of the synthetic PCD. Next, the RGB channels are interpolated from the real data to the synthetic data, as shown in Figure 1.

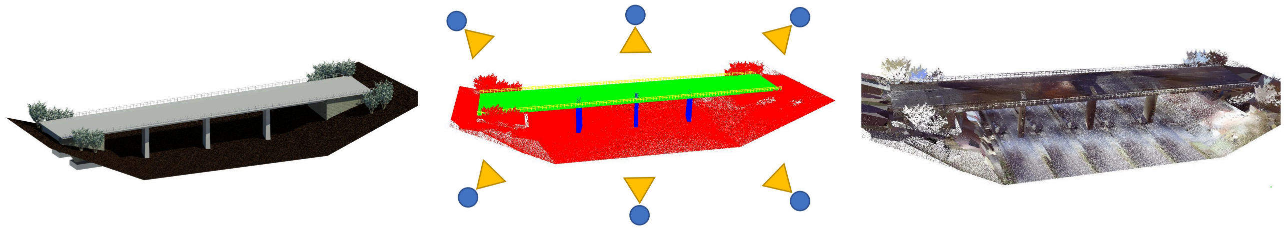

Figure 1: Creating synthetic PCD: parametric modeling (left); laser scanning of the bridge model (middle); resulting synthetic PCD (right).

5 Semantic segmentation

semantic segmentation of PCD can be defined as the labeling process of the data at a point level. Through semantic segmentation, the entire PCD is summarized to segments that show the element of interest in bridges. In this section, details of the deep learning model for processing points and the dataset are mentioned.

5.1 Deep learning model

As a deep learning model capable of processing points, RandLA-Net [19] is utilized for semantic segmentation. This model can process and learn features of large-scale point clouds. It applies random sampling in subsequent layers of the network to reduce the number of points. Simultaneously, to prevent the loss of key features through sampling, a local feature aggregation module is proposed to progressively increase the receptive field for each point. This module contains a local spatial encoding (LocSE) and an attentive pooling block, as shown in Figure 2. LocSE computes the neighboring points of each point by K-nearest neighbors (KNN) and sends the relative Euclidean distance of the point to its neighbors to a multi-layer perceptron (MLP). Next, the resulting features are concatenated with the input features of the point. In the attentive pooling unit, these features are aggregated through a weighted sum. The weights (scores) of the operation are obtained from a shared MLP to emphasize the more important features. Finally, every two feature aggregation modules are stacked to expand the receptive field of the point.

5.2 Dataset

The bridge PCD of Cambridge [8] is used for training and testing the deep learning model. This *Real* dataset contains 10 samples of reinforced concrete bridges captured by laser scanning. Following section 4, 10 samples of synthetic PCD is also created for bridges. Next, the synthetic data is added to the dataset of real bridges to provide an *Augmented* dataset containing 20 bridges in total. Figure 3 depicts two typical samples of the real and synthetic PCD, respectively.

5.3 Hyperparameters

The model is trained for semantic segmentation of four classes, namely: piers, deck, railings, and background. The hyperparameters of RandLA-Net are tuned through a trial and error process, and the parameters with the best validation accuracy are selected. The model is trained for 512 epochs with a batch size of two. The input channel of the network is considered 6, representing *X, Y, Z* and *R, G, B*. The number of 16 neighbors is selected for the KNN search block of the model. Also, 6 neurons are used in the attentive pooling block of the network.

6 Results

To evaluate the impact of the synthetic data on the results, the model is trained on two datasets and tested on the same samples of the real bridges. The first dataset is the PCD of the real bridges containing 10 samples, and the next one is the PCD of the real bridges combined with 10 samples of the synthetic PCD. The validation is performed on only the real samples of the bridges using leave one out cross validation (LOOCV) method. In every iteration, a real PCD is left out for testing, and the training process is conducted on the other remaining samples. The test results of the models are reported by the mean intersection over union (MIoU) and mean accuracy (MAcc) over classes. These statistical indices are calculated from the confusion matrix of the predicted labels in validation or testing phases. Table 1 shows the results of the model on each test sample of the real PCD. As can be seen, the results have improved for some of the test samples. Averaging the results over all the test samples illustrates that augmentation has increased MIoU by 5.2% and MAcc by 3.1%.

Considering the prediction results over each test sample, a noticeable change is observed in the accuracy of Bridge 02 and Bridge 08 where MIoU has increased more than 20% after synthetic data augmentation. One reason for such improvement could be due to the slight horizontal curvature of these bridges. Adding synthetic samples with horizontal curvature has resulted in a more generic dataset for the model to learn the features of curve bridges. Figure 4 visually depicts the prediction

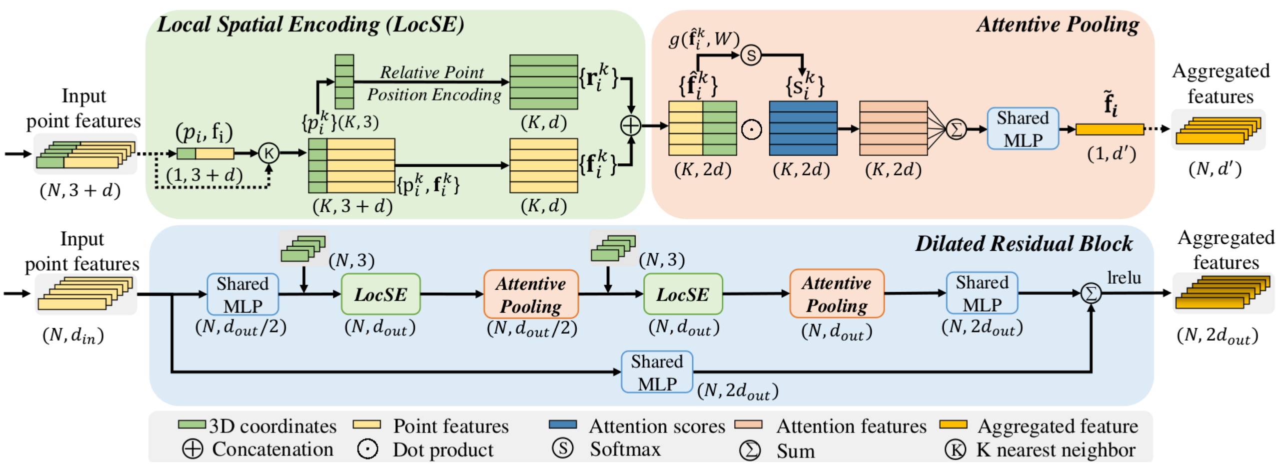

Figure 2: Architecture of RandLA-Net [19].

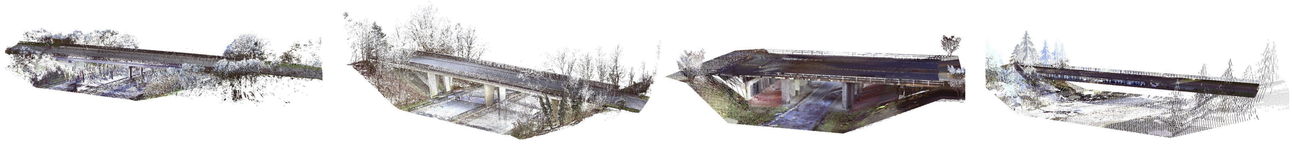

Figure 3: Two typical point clouds of each dataset: real samples [8] (left); synthetic samples (right).

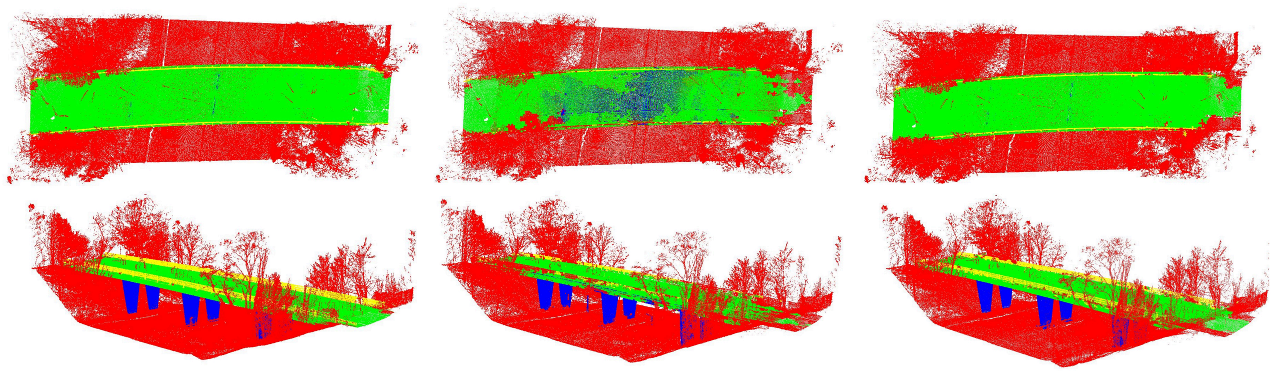

Figure 4: Prediction results of the model: ground truth (left), training on real PCD (middle), training on augmented data (right). Class labels: railings (yellow), background (red), deck (green), piers (blue).

Table 1: The results of semantic segmentation of the real and augmented PCD (%)

Test Sample	MAcc Real	MAcc Aug	ΔMAcc	MIoU Real	MIoU Aug	ΔMIoU
Bridge 01	86.7	84.2	-2.9	81.2	79.6	-2.0
Bridge 02	82.2	90.4	10.0	71.1	88.0	23.8
Bridge 03	97.0	95.7	-1.3	92.7	91.5	-1.3
Bridge 04	93.2	95.9	2.9	89.9	92.2	2.6
Bridge 05	96.4	98.0	1.7	92.6	95.5	3.1
Bridge 06	95.6	98.2	2.7	87.3	89.0	2.0
Bridge 07	93.7	92.2	-1.6	76.0	72.5	-4.6
Bridge 08	78.1	93.8	20.1	70.6	90.5	28.2
Bridge 09	94.1	94.0	-0.1	89.8	89.2	-0.7
Bridge 10	85.0	84.8	-0.2	76.2	76.6	0.5
Average	90.2	92.7	3.1	82.7	86.5	5.2

results of the model on Bridge 08. As can be seen, data augmentation has improved the prediction accuracy of the model especially in regions where the piers and railings are connected to the deck.

7 Conclusion

Semantic segmentation is an essential step in the 3D modeling of bridges and creating digital twins. Despite the recent advances in technologies such as laser scanning and photogrammetry, the scanning process of real bridges is still labor-intensive and costly. Deep learning models have recently shown successful performance in automating the labeling process of PCD. These models, however, need a large dataset for training which is hard to achieve for bridges. This paper proposed an approach to generate synthetic PCD of bridges for data augmentation. Parametric models of real bridges were created, and their PCD was simulated. This synthetic dataset was combined with the dataset of real bridges and used for training a deep learning model. Training on the real and augmented datasets demonstrated that data augmentation by synthetic PCD improves the accuracy of the models. On average, the MIoU of the deep model in predicting the labels of real PCD increased by 5.2%. This improvement shows the feasibility of using synthetic PCD for augmentation. However, the factors affecting the results are still unclear. To provide more accurate point clouds for data augmentation and training deep learning models, these factors need to be investigated in the future.

References

[1] ASCE, “Asce’s 2021 infrastructure report card”, American Society of Civil Engineers, Report, 2021.

[2] E. Negri, L. Fumagalli, and M. Macchi, “A review of the roles of digital twin in cps-based production systems”, *Procedia Manufacturing*, vol. 11, pp. 939–948, 2017, ISSN: 2351-9789.

[3] F. Collins, M. Mafipour, F. Noichl, Y. Pan, and M. Vega, “Towards applicable scan-to-bim and scan-to-floorplan: An end-to-end experiment”, in *Proc. of the 32th Forum Bauinformatik*, 2021.

[4] Y. Pan, A. Borrmann, H.-G. Mayer, *et al.*, “Built environment digital twinning”, Technical University of Munich, Report, 2019.

[5] M. S. Mafipour, S. Vilgertshofer, and A. Borrmann, “Deriving digital twin models of existing bridges from point cloud data using parametric models and metaheuristic algorithms”, in *Proc. of the EG-ICE Conference 2021*, 2021.

[6] Z. Zhu, S. German, and I. Brilakis, “Detection of large-scale concrete columns for automated bridge inspection”, *Automation in construction*, vol. 19, no. 8, pp. 1047–1055, 2010, ISSN: 0926-5805. DOI: http://dx.doi.org/10.1016/j.autcon.2010.07.016.

[7] Z. Zhu, S. German, and I. Brilakis, “Visual retrieval of concrete crack properties for automated post-earthquake structural safety evaluation”, *Automation in Construction*, vol. 20, no. 7, pp. 874–883, 2011, ISSN: 0926-5805. DOI: http://dx.doi.org/10.1016/j.autcon.2011.03.004.

[8] R. Lu, I. Brilakis, and C. R. Middleton, “Detection of structural components in point clouds of existing rc bridges”, *Computer-Aided Civil and Infrastructure Engineering*, vol. 34, no. 3, pp. 191–212, 2019, ISSN: 1093-9687. DOI: http://dx.doi.org/10.1111/mice.12407.

[9] F. Hu, J. Zhao, Y. Huang, and H. Li, “Structure-aware 3d reconstruction for cable-stayed bridges: A learning-based method”, *Computer-Aided Civil and Infrastructure Engineering*, vol. 36, no. 1, pp. 89–108, 2021, ISSN: 1093-9687.

[10] J. S. Lee, J. Park, and Y.-M. Ryu, “Semantic segmentation of bridge components based on hierarchical point cloud model”, *Automation in Construction*, vol. 130, p. 103 847, 2021, ISSN: 0926-5805.

[11] C. R. Qi, H. Su, K. Mo, and L. J. Guibas, “Pointnet: Deep learning on point sets for 3d classification and segmentation”, in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 652–660.

[12] Y. Wang, Y. Sun, Z. Liu, S. E. Sarma, M. M. Bronstein, and J. M. Solomon, “Dynamic graph cnn for learning on point clouds”, *ACM Transactions on Graphics (TOG)*, vol. 38, pp. 1–12, 2019.

[13] Y. Yan and J. F. Hajjar, “Automated extraction of structural elements in steel girder bridges from laser point clouds”, *Automation in Construction*, vol. 125, p. 103 582, 2021, ISSN: 0926-5805.

[14] L. Truong-Hong and R. Lindenbergh, “Automatically extracting surfaces of reinforced concrete bridges from terrestrial laser scanning point clouds”, *Automation in Construction*, vol. 135, p. 104 127, 2022, ISSN: 0926-5805.

[15] T. Xia, J. Yang, and L. Chen, “Automated semantic segmentation of bridge point cloud based on local descriptor and machine learning”, *Automation in Construction*, vol. 133, p. 103 992, 2022, ISSN: 0926-5805.

[16] L. Winiwarter, A. M. E. Pena, H. Weiser, *et al.*, “Virtual laser scanning with helios++: A novel take on ray tracing-based simulation of topographic full-waveform 3d laser scanning”, *Remote Sensing of Environment*, vol. 269, p. 112 772, 2022.

[17] Autodesk, *Revit*, version 2021. [Online]. Available: https://www.autodesk.com/products/revit/overview.

[18] SOFiSTiK, *Sofistik bridge + infrastucture modeler*, version 2021. [Online]. Available: https://www.sofistik.com/products/bim-cad/bridge-infrastructure-modeler.

[19] Q. Hu, B. Yang, L. Xie, *et al.*, “Randla-net: Efficient semantic segmentation of large-scale point clouds”, in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 11 108–11 117.