Received: 7 November 2024 | Accepted: 4 January 2025

DOI: 10.1111/mice.13422

RESEARCH ARTICLE

COMPUTER-AIDED CIVIL AND INFRASTRUCTURE ENGINEERING WILEY

A structure-oriented loss function for automated semantic segmentation of bridge point clouds

Chao Lin[1] | **Shuhei Abe**[1] | **Shitao Zheng**[1,2] | **Xianfeng Li**[1,2] | **Pang-jo Chun**[1,2]

[1]Department of Civil Engineering, The University of Tokyo, Tokyo, Japan

[2]Institute of Engineering Innovation, The University of Tokyo, Tokyo, Japan

Correspondence
Pang-jo Chun, Department of Civil Engineering, The University of Tokyo, Tokyo, Japan.
Email: chun@g.ecc.u-tokyo.ac.jp

Funding information
Council for Science, Technology and Innovation; Cross-Ministerial Strategic Innovation Promotion Program (SIP), Grant/Award Number: JPJ012187; Public Works Research Institute; JSPS Grant-in-Aid, Grant/Award Numbers: 21H01417, 22H01561, 23H00198; China Scholarship Council, Grant/Award Number: 202306210048

Abstract

Focusing on learning-based semantic segmentation (SS) methods for bridge point cloud data (PCD), this study proposes a structure-oriented concept (SOC) with training focused on the spatial distribution patterns of bridge components, including both the horizontally absolute location of each component and its vertically relative position compared with other components. Then a structure-oriented loss (SOL) function, which embodies the core of SOC, is defined accordingly, and it is compared to five cutting-edge loss functions on a collected bridge PCD dataset. In contrast to the limitations of other loss functions, SOL significantly improves the overall evaluation metrics of overall accuracy (6.53%) and mean intersection over union (mean IoU: 8.67%). The IoU of the category "others" is improved by 8.44%, which is very important for automating the time-consuming denoising process. Furthermore, the demonstrated robustness of SOC and SOL reveal great potential to improve the performance of other SS models.

1 | INTRODUCTION

Bridges are essential elements of infrastructure in the modern transportation network playing an irreplaceable role in spanning physical obstacles (Kunlamai et al., 2024). Unfortunately, a large and still growing number of bridges are aging, leading to a globally concerning issue that cannot be overlooked (Chun et al., 2021). For example, approximately 42% of the bridges in the United States (ASCE, 2021) and 43% of the bridges in Japan (MLIT, 2024) are over 50 years old. Therefore, reliable inspection of bridges must be guaranteed periodically (Adeli & Saleh, 1999), and most commonly this process is still undertaken manually by trained staff, making it labor-intensive, time-consuming, and prone to error (Lamas et al., 2023). Due to the increasing number of aging bridges, there is an urgent need to develop more efficient and automated inspection methods (Adeli & Kim, 2009).

As-is model representations of constructed infrastructures are of great value for various applications in civil engineering, including inspection and maintenance tasks (Noichl et al., 2024). With the development of laser-scanning (H. S. Park et al., 2007; S. W. Park et al., 2015), point cloud data (PCD) has been increasingly utilized for digitizing existing structures not only because of its efficient and non-contact acquisition but also its potential to support necessary inspection and management tasks after analysis and processing (H. Liang et al., 2024). A point cloud is a discrete set of three-dimensional (3D) data points in space. Each point position has X, Y, and Z coordinates and optionally contains other attributes such as color, normal and intensity information, and

so forth. However, the raw PCD of a whole bridge is infrastructure-level, where each point does not have any semantic information so component-level information cannot be directly revealed from it. Therefore, it is necessary to decompose full-scale bridge PCD at the component level to obtain geometric, topological, and semantic information of bridge components.

Semantic segmentation (SS), which has shown outstanding performance in image processing (Chen et al., 2024; Chu et al., 2024), also serves as a powerful tool for decomposing PCD. SS becomes an inevitable step in establishing comprehensive and accurate 3D models from PCD since the manageable component with associated geometry, feature, and condition information is the elementary unit in 3D models. PCD after SS allows the creation of 3D models, which contain abundant information (e.g., the building information model [BIM]; Hattori et al., 2024; Romero-Jarén & Arranz, 2021), the digital twin (Lu & Brilakis, 2019), and finite element models (G. Yu & Adeli, 1993), offering a geometric basis for inspection tasks such as damage detection (Huang et al., 2024; Qin et al., 2024; Sun et al., 2024), health monitoring (Riveiro et al., 2016), and vibration control (Ghaedi et al., 2017; Gutierrez Soto & Adeli, 2019; Hongjin Kim & Adeli, 2005). The accuracy of PCD SS is vital for determining the reliability of the 3D models and the performance of subsequent applications that rely on 3D data processing. Therefore, improving the accuracy of SS models stands out as a primary focal point within civil engineering.

Existing SS approaches for PCD can be classified into algorithm-based and learning-based methods. When applied to bridges, shortcomings of algorithm-based methods were exposed, such as the involvement of intensive manual work (Lamas et al., 2023), sensitivity to noise and point density (Riveiro et al., 2016), inadaptability to components with complex shapes (Lu et al., 2019), and so forth. In contrast, learning-based methods have attracted plenty of interest and attention because they are capable of being generalized to different bridge types. Representative learning-based methods include PointNet (Qi, Su, et al., 2017) and its extensions such as PointNet++ (Qi, Yi, et al., 2017), superpoint graph (SPG; Landrieu & Simonovsky, 2018), and dynamic graph convolutional neural network (DGCNN; Y. Wang et al., 2019). Many studies have deployed these SS models on bridge PCD and achieved segmentation results. In these studies, many deep learning techniques are utilized in the models, such as CNN, the attention mechanism, the encoder–decoder structure, loss functions, and so forth.

In the field of bridge PCD SS, however, previous studies either borrowed loss functions designed for other targets or sometimes made slight modifications to them for bridges. As a result, current loss functions constrain the performance of SS models. Until now, there has been no loss function that is developed based on the characteristics of bridges. In this study, we propose a novel structure-oriented concept (SOC) and a structure-oriented loss (SOL) function according to it where the geometrical pattern of bridges is utilized to determine the samples worth focusing on. The comparison results show that SOL outperforms the existing cutting-edge loss functions in terms of all quantitative metrics.

The principal contributions of this study are: (1) An SOC that quantitatively represents the spatial distribution of bridge components; (2) an SOL function that significantly improves the performance of bridge PCD SS models; and (3) the verification of the substantial potential inherent in loss functions.

The rest of this paper is structured as follows. Section 2 reviews the current state of learning-based research toward the SS of bridge PCD and summarizes efforts focused on developing loss functions. Section 3 first introduces the methodology of this paper. The data preparation process and implementation details are then presented. Section 4 elaborates the SOC and subsequently defines the SOL function. The best values of the hyperparameters in SOL are obtained according to experiment results. Section 5 shows a comparison between the proposed SOL and other existing loss functions. The predominance of SOL is demonstrated by both numerical and visualized results. Section 6 concludes the paper and discusses the limitations and potential future research directions.

2 | LITERATURE REVIEW

Extensive studies have managed to segment bridge PCD by leveraging cutting-edge deep learning models where measures were taken from multiple perspectives to accommodate the characteristics of bridge PCD. While there have been limited attempts at developing more advanced loss functions among these studies, numerous existing loss functions can be considered as baselines or potential candidates for this task. In the subsequent subsections, notable learning-based applications of bridge PCD SS and popular loss functions in deep learning are reviewed, followed by a discussion on the corresponding research gaps.

2.1 | Learning-based bridge PCD SS

PointNet (Qi, Su, et al., 2017) is a pioneering deep learning framework for PCD processing that can extract geometric features from PCD while remaining invariant to permutations of the input point sets. Hyunjun Kim et al. (2020) leveraged PointNet on subspaces of full-scale bridge PCD

to detect piers and decks. Hyunsoo Kim and Kim (2020) compared the performance of PointNet, PointCNN, and DGCNN in classifying bridge components. Each bridge PCD was segmented into two parts as training and testing samples. The results showed that DGCNN was proven to outperform PointNet and PointCNN. These two studies automated the SS of bridge PCD while discarding the manual removal of noisy points, demonstrating the capability of deep learning-based methods to detect and classify bridge components from PCD. The issue revealed from these two studies presents a paradox: bridges, as relatively large infrastructures (N. Wang & Adeli, 2015), require partition of their PCD before deep learning to preserve the high spatial resolution. However, global features are lost during this partition procedure, thereby influencing the SS results. This issue was taken into consideration in later studies.

Lee et al. (2021) proposed a graph-based hierarchical DGCNN (HGCNN) model for the SS of railway bridges with electric poles. HGCNN possesses a broader range of local features and improves the ability to detect the boundary between two different components. According to the results, HGCNN outperformed PointNet and DGCNN comprehensively, with a notable improvement of about 3% in the intersection over union (IoU) of the electric pole when using HGCNN. Y. C. Lin and Habib (2022) used SPG to semantically segment bridge components and road infrastructure such as man-made terrain, buildings, vegetation, and so forth. The good performance of SPG verified its feasibility in processing large-scale PCD with efficiency and high accuracy. Yang et al. (2022) introduced a weighted cross entropy (WCE) loss function (Han et al., 2021) to SPG and thus presented a novel method called weighted SPG (WSPG). The unequal distribution between different bridge components was mitigated, and WSPG achieved the best performance, compared to PointNet, DGCNN, and the original SPG. Additionally, five loss functions were compared in this study, and none of them achieved optimal results in all component categories. Subsequently, Yang et al. (2023) upgraded WSPG by using a graph-structured deep metric learning method and replacing the previous loss function with a graph-structured contrastive loss function (Landrieu & Boussaha, 2019), which improved the boundary segmentation quality and boosted the general performance of the proposed model. Xiao et al. (2024) proposed RandLA-BridgeNet by making slight adjustments to RandLA-Net (Hu et al., 2020), which showed impressive performance, compared to other state-of-the-art models. When feeding the same dataset collected by Lu et al. (2019) to RandLA-BridgeNet with background points (Xiao et al., 2024) and to PointNet, DGCNN, SPG, and WSPG after denoising (Yang et al., 2022), RandLA-BridgeNet performed only slightly lower in most evaluation metrics than the other models.

There were also studies that established SS models from scratch using existing frameworks only for reference. Xia et al. (2022) proposed a straightforward scheme to detect structural components from bridge PCD. Geometric features were extracted from normals of neighboring points and then fed into a classification neural network. Though manual postprocessing was required, this scheme improved the mean IoU (mIoU) from 45.92% to 94.72% as compared to PointNet, which is much more complicated. Jing et al. (2022) designed an SS model using the popular encoder–decoder structure for masonry arch bridges called BridgeNet. Some methods and hyperparameters were borrowed from PointNet++ (Qi, Yi, et al., 2017), RandLA-Net (Hu et al., 2020), and FG-Net (Liu et al., 2020), such as maxpooling, global feature extraction, random sampling, downsampling, and drop ratios. As a result, BridgeNet outperformed PointNet++ and RandLA-Net by a significant margin. Jing et al. (2024) subsequently enhanced BridgeNet to BridgeNetv2 by incorporating the Transformer architecture. While maintaining lightweight and memory-efficient, BridgeNetv2 surpassed PointNet++, RandLANet, and BridgeNet across all metrics.

Learning-based methods have proven to be reliable and effective in automating PCD SS for different types of bridges. In particular, studies (Xiao et al., 2024; Yang et al., 2022, 2023) have demonstrated significant improvements in model performance by replacing the loss functions used. However, it is worth noting that many studies still retain the inherent loss function of the deep learning model they employ.

2.2 | Loss functions

Cross entropy (CE) loss has been extensively used to describe the loss of classification results (Hui & Belkin, 2020). Since the SS of point clouds is point-wise classification in essence, CE is intuitively adopted in most SS models. Based on CE, the concept of weighted CE (WCE) was proposed to allocate different weights to certain classes so that specific concentrations can be rendered to targeted objects. For example, data imbalance among different categories always exists in PCD and the classification results of minority categories are unsatisfied as a consequence (Zhu et al., 2019), but the utilization of WCE can compensate categories with fewer points by raising their weights. Additionally, WCE loss functions were raised from different patterns according to the research objectives (Han et al., 2021; Milioto et al., 2019; Phan & Yamamoto, 2020). Furthermore, focal loss (FL; T. Y. Lin et al., 2017) promoted CE from another perspective by setting a factor to focus more on misclassified examples. Essentially, FL accords with the

concept of WCE, but higher weights are allocated to hard examples instead of certain categories.

The features employed during learning can also be used to improve the discriminative power of loss functions, such as geometric features, embedding features, and so forth. Wen et al. (2016) developed the center loss function, which utilizes the sum of Euclidean distances between the deep features and their center to represent the intra-class variation. According to experiments, the center loss function effectively minimized the intra-class variations while keeping the features of different classes separable. Inspired by the central idea of the center loss, Z. Liang et al. (2020) proposed a structure-aware loss (SAL) function for PCD instance segmentation where the geometric center was also taken into consideration to make good use of the rich geometric information in PCD. Moreover, inter-class discrimination is also included in the function.

There have been loss functions that were specifically developed for image SS tasks. Dice loss (Milletari et al., 2016) measures the overlap between predicted and target segmentation masks in image SS. Jaccard loss, also known as IoU loss (J. Yu et al., 2016), evaluates object detection algorithms by measuring the overlap between predicted and ground truth bounding boxes. Despite the success of these loss functions in image SS, their performance on point cloud SS has been found to be limited according to Yang et al. (2022).

To summarize, loss functions are still underdeveloped in the realm of deep learning for PCD. There is a need for further research on designing loss functions that are specifically tailored to address the characteristics of PCD, especially in the context of bridge PCD. Bridge PCD has distinct characteristics, such as its large scale (Adeli & Saleh, 1997), data imbalance among categories, regular distribution pattern of components, and so forth. On the one hand, some characteristics posed barriers or challenges when deploying learning-based methods, and researchers have addressed them from various perspectives, as mentioned earlier. On the other hand, potentially useful characteristics, such as the geometric relationship of bridge components, have not attracted enough attention as they did not impede the deep learning process. Therefore, approaches to utilize the regular distribution pattern of bridge components to orient the learning process of SS models for bridge PCD have yet to be researched.

2.3 | Loss functions for comparison

To demonstrate the predominance of SOL, five representative cutting-edge loss functions are selected to compare with the SOL function when deployed on the same dataset, namely, negative log likelihood (NLL) loss, CE loss, WCE loss (Han et al., 2021), FL (T. Y. Lin et al., 2017), and SAL (Z. Liang et al., 2020). Specifically, these loss functions are defined as follows.

2.3.1 | NLL

NLL loss is widely utilized when dealing with large datasets because of its computational advantages. It is defined as

$$L = -\frac{1}{N}\sum_{i=1}^{N} x_i \tag{1}$$

where N represents the point number, and x_i denotes the probability of the true class after a LogSoftmax layer for the ith point.

2.3.2 | CE

The suitability for classification tasks of the CE loss made it popular in SS models, it is defined as

$$L = -y\log(p) \tag{2}$$

where y and p are label vectors of ground truth and predicted results, respectively.

2.3.3 | WCE

Assigning different weights to training classes is currently a popular method to alleviate data imbalance. Various weight allocation strategies have been proposed for diverse purposes. Han et al. (2021) presented a novel WCE and it exhibited high performance in the SS of bridge point clouds (Yang et al., 2022). The WCE loss function is defined as

$$L = -\sum_{c=1}^{M} \frac{\prod_{i=1,\, i\neq c}^{M}\sqrt{N_i}}{\sum_{i=1}^{M}\sqrt{N_i}} y_c \log(p_c) \tag{3}$$

where M denotes the number of classes, N_c represents the point number of the cth class during training, while y_c and p_c are label vectors of ground truth and predicted results of the cth class, respectively.

2.3.4 | FL

FL, as indicated by its name, contains a focus on hard examples. Class imbalance can also be dealt with by FL.

FL for binary classification is defined as

$$L = \begin{cases} -\alpha(1-p)^{\gamma}\log(p), & y=1 \\ -(1-\alpha)\,p^{\gamma}\log(1-p), & otherwise \end{cases} \tag{4}$$

where $p \in [0, 1]$ is the model's estimated probability for the class with label $y = 1$, $y \in \{\pm 1\}$ specifies the ground truth class, γ is the tunable focusing parameter, and α is the weighting factor. T. Y. Lin et al. (2017) indicated that extending the FL to the multi-class case is straightforward and $\gamma = 2, \alpha = 0.25$ works best. Therefore, the same values for these two hyperparameters were adopted in this study.

2.3.5 | SAL

Z. Liang et al. (2020) took the structure information of point clouds into consideration and thus proposed a SAL function. It is defined as

$$L = \frac{1}{M}\sum_{i=1}^{M}\sum_{k=1}^{N_i}\text{sigmoid}\left(\|p'_{i,k}\|\right)\left[\|s'_{i,k}\|-\alpha\right]_+^2 + \frac{1}{M(M-1)}\sum_{i=1}^{M}\sum_{j=1,j\neq i}^{M}\left[\beta-\|\mu_{s,i}-\mu_{s,j}\|\right]_+^2 \tag{5}$$

where M is the number of classes in the scene; N_i is the point number of the ith class; $\|\cdot\|$ denotes the Euclidean distance; $[x]_+$ means $\max(0, x)$; $p'_{i,k}$ and $s'_{i,k}$ are the 3D coordinates and the embedding after mean centering, respectively; $\mu_{s,i}$ means the embedding center of the ith class; α and β are two thresholds that are set to 0.7 and 1.5, respectively, based on the experiment results.

3 | RESEARCH WORKFLOW

3.1 | Methodology

There is great potential in incorporating the structural schema of bridges into the SS process, and loss functions can be a good platform for doing this because of their improvement to deep learning models (Chern et al., 2023; Deng et al., 2024; Q. Wang et al., 2020). Therefore, both horizontal and vertical distributions of different bridge components are considered in the proposed SOL when calculating the loss during deep learning. To illustrate the effectiveness of SOC and SOL in leveraging the distribution patterns of bridge components for better SS, the subsequent two sections will introduce and validate these methods. In Section 4, the theory of SOC is narrated in detail, revealing the robustness and generalization capacity of the concept. Then the SOL function that is based on SOC is specifically defined, and the values of hyperparameters within it are calibrated. In Section 5, SOL is compared with existing state-of-the-art loss functions through both numerical metrics and visual representations. This comparative evaluation serves to underscore the comprehensive and insightful superiority of SOL in the realm of bridge PCD SS.

In the following subsections, the dataset to test the performance of loss functions and the quantitative metrics for performance evaluation are introduced.

TABLE 1 Technical specifications of the terrestrial laser scanner.

Accuracy	Operating range	Acquisition rate	Field of view	
			Horizontal	Vertical
±20 mm at 10 m	0.5–100 m	10^5 points per second	360°	295°

3.2 | Dataset

To establish a bridge PCD dataset for validation, a terrestrial laser scanner (Matterport Pro3) was used to collect the PCD of 12 reinforced concrete (RC) bridges in Japan. The technical specifications of the scanner are shown in Table 1. The number of scans for each bridge was determined by its size, with a minimum of ten scan positions per bridge being taken to ensure data quality.

The collected PCD was manually annotated, classifying the points into five classes, namely, abutment, girder, deck, parapet, and others. The original and annotated PCD of the 12 bridges is shown in Figure 1, and Table 2 presents the number of points in each class for all bridges.

As for the aforementioned 12 bridges in Japan, Bridges 1 to 9 are located in a rural region of Fukushima Prefecture, and Bridges 10 to 12 are in urban areas near Tokyo. Bridges 4, 8, and 9 contain steel girders beneath their RC girders, whereas the remaining only possess RC girders. The collected PCD of Bridges 5, 10, 11, and 12 are incomplete. Bridge 5 was only scanned from the sides and top due to the inaccessibility of the underside, and the role of Bridges 10 to 12 is to compensate for the point scarcity in the class parapet so only parts that are close to parapets were scanned, including the girder and deck.

3.3 | Implementation details

Considering the scale of bridge PCD in the dataset, the original PCD of each bridge was empirically downsampled to 5% for computational efficiency using the random downsampling technique. The proportion of each class in the whole dataset after downsampling is shown in Figure 2.

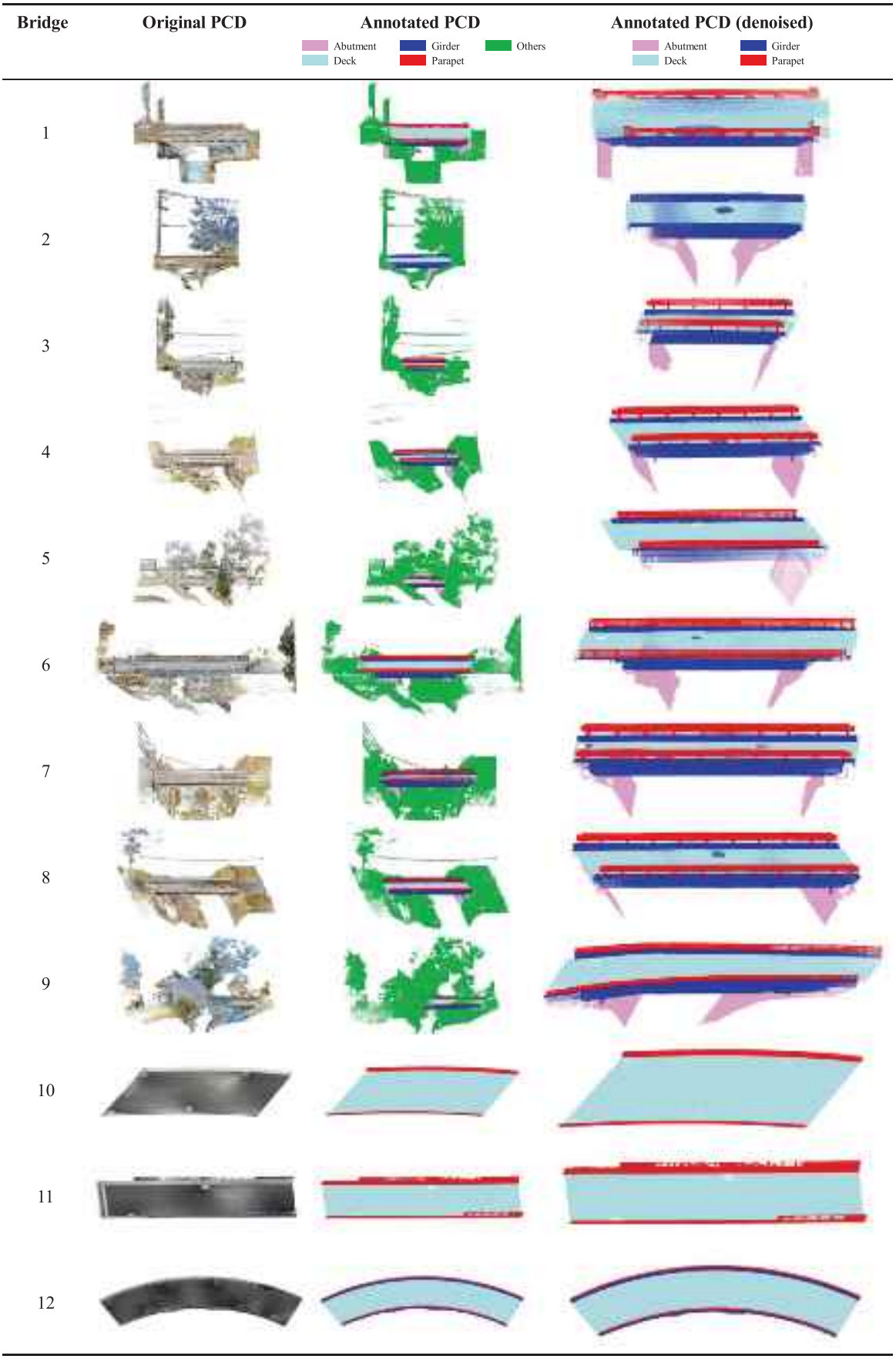

FIGURE 1 Dataset used to test loss functions. PCD, point cloud data.

TABLE 2 Metadata of the bridge point cloud data dataset.

Bridge	Total number of points	Number of points in each class				
		Abutment	Girder	Deck	Parapet	Others
1	28,956,502	1,615,265	11,756,082	1,959,218	300,831	13,325,106
2	27,155,818	5,814,749	7,091,466	2,116,838	/	12,132,765
3	39,118,997	5,241,694	9,889,022	2,743,044	743,967	20,501,270
4	29,741,828	1,329,542	6,471,526	3,759,269	934,157	17,247,334
5	20,674,051	113,713	1,200,867	4,597,179	906,135	13,856,157
6	38,702,332	7,658,837	13,036,570	6,199,693	889,295	10,917,937
7	48,206,907	3,236,297	16,608,943	4,191,342	1,855,722	22,314,603
8	30,687,112	528,687	8,319,788	5,481,559	1,378,384	14,978,694
9	80,792,348	14,653,328	32,099,530	8,331,194	1,958,971	23,749,325
10	6,102,567	/	/	4,973,275	1,128,553	739
11	4,595,953	/	/	3,125,038	1,470,915	/
12	10,824,090	/	1,257,849	8,306,717	1,257,506	2018

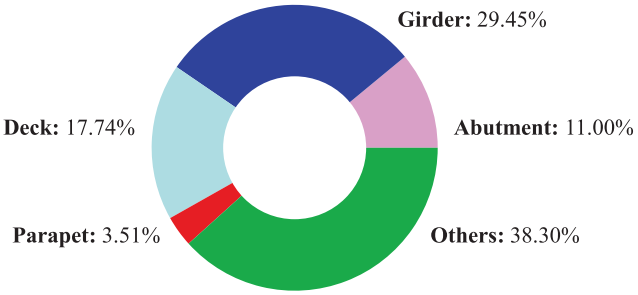

FIGURE 2 Proportion of each class after downsampling.

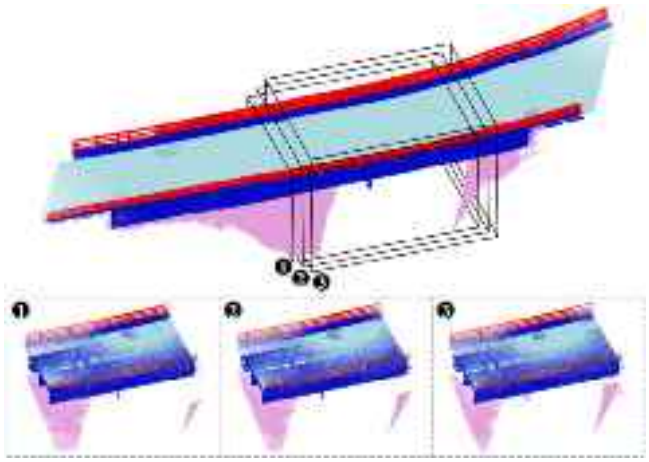

FIGURE 3 Partitioned bridge point cloud data (PCD) with overlap.

During the deep learning process, Bridges 3, 4, and 7 were selected for testing and the others for training. PointNet++ (Qi, Yi, et al., 2017) is selected to serve as the backbone network based on its prevalence. Moreover, every bridge PCD was partitioned with a subspace length of 10 m and a 95% overlap to appropriately fit the receptive field of PointNet++ as inspired by Hyunjun Kim et al. (2020), which is shown in Figure 3. As a result, 528 bridge segments were obtained as training data.

To quantitatively evaluate the effectiveness and performance of different loss functions on bridge PCD SS, four common evaluation metrics were used, namely, overall accuracy (OA), mean class accuracy, IoU, and mIoU.

The deep learning framework PyTorch was utilized and the multi-scale grouping strategy was applied in PointNet++ during training as recommended by the authors (Qi, Yi, et al., 2017). Default parameters were maintained in other parts of PointNet++ to fairly compare the performance of different loss functions. More specifically, the Adam optimizer was employed, whose initial learning rate and decay rate were set to 0.001 and 0.0001, respectively.

As for the computational environment, an Intel Core i9-14900KF CPU with 128 GB of RAM clocked at 3.20 GHz, and an NVIDIA GeForce RTX 4090 GPU was deployed in a PC.

4 | SOC AND SOL FUNCTION

4.1 | SOC

As illustrated in Figure 4, in the horizontal direction, abutments are located on both sides along the length of the bridge, and parapets lie on both sides along the width; and vertically, abutments, girders, decks, and parapets are arranged from the bottom to top. In this paper, we propose the SOC that represents the geometrical pattern quantitatively and defines corresponding judging conditions so that the fixed spatial relationship between bridge

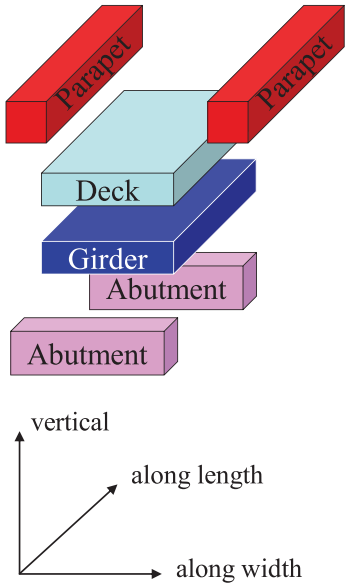

FIGURE 4 Spatial distribution pattern of bridge components.

components can be utilized to help evaluate the accuracy of deep learning models. To reduce the computational burden, only the spatial center of all points that belong to one certain component in a given space is used to represent its location rather than calculating the location of each point individually.

4.1.1 | Horizontal distribution

To clearly observe the horizontal distribution of bridge components, their horizontal locations must be represented. Rectangles were selected as the minimum surrounding borders considering the shape of the horizontal projection of most components and calculation simplicity. Before training, each minimum rectangle that enclosed all points in a single component instance was calculated and the coordinates of all vertices were recorded as shown in Figure 5. Girders and decks are chunky components, so a single rectangle was enough for each of them; however, abutments and parapets are paired components so a pair of rectangles was needed. When each point was classified into one category by the deep learning model during training, the recorded coordinates could be used to judge whether the point was in the right region.

4.1.2 | Vertical distribution

To observe the vertical relationship between components precisely, their vertical centers were calculated for each bridge PCD as illustrated in Table 3.

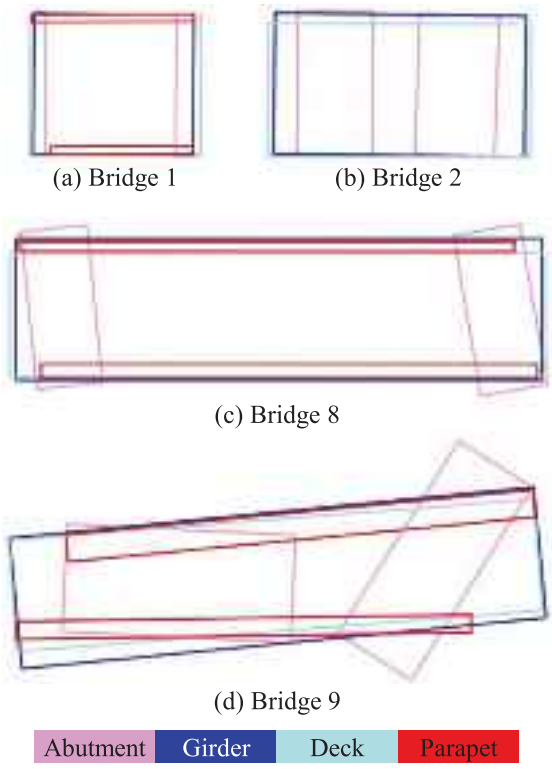

FIGURE 5 Minimum rectangles of components in bridges.

TABLE 3 Vertical centers of each bridge component.

Bridge	Vertical center (m)			
	Abutment	Girder	Deck	Parapet
1	1.64	2.15	2.30	2.46
2	−0.52	0.84	1.25	/
3	1.67	3.65	4.20	5.20
4	−2.44	−0.26	0.11	0.96
5	−2.20	0.18	0.03	0.68
6	−0.18	1.42	2.07	2.82
7	1.49	3.49	4.05	4.90
8	−1.49	0.47	0.82	1.67
9	1.63	3.70	4.13	4.78
10	/	/	−0.10	0.43
11	/	/	−0.28	0.59
12	/	0.08	−0.11	0.64

As can be seen from Table 3, all complete bridge PCD satisfies the vertical distribution pattern, while some incomplete bridge PCD does not, namely, Bridge 5 and Bridge 12. Specifically, the vertical centers of girders are above those of decks for the two bridges, which is contrary to the assumed pattern. Figure 6 shows details of the collected PCD of Bridge 5 and Bridge 12. The collected girder PCD is primarily located above decks, while the portion below decks is dramatically missing, leading to a reversed

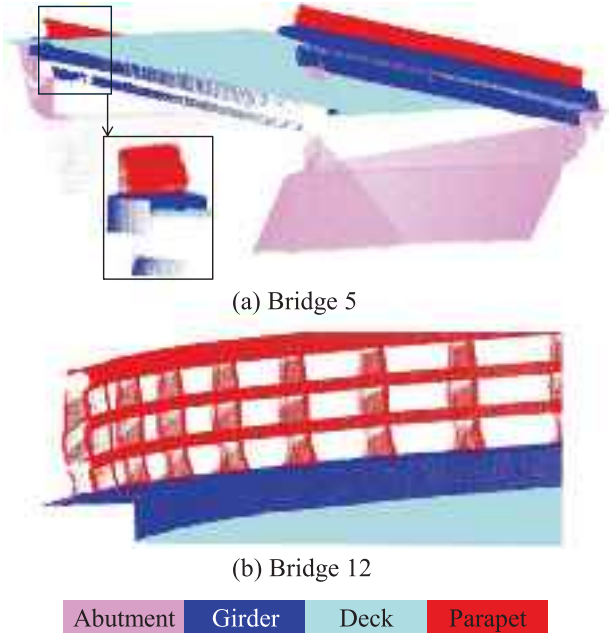

FIGURE 6 Incomplete girder in bridge PCD.

vertical relationship between girders and decks. However, it is worth noticing that the relationships between other components in Bridge 5 and Bridge 12 still accord with the distribution pattern.

Table 3 presents the vertical centers of each bridge component from the perspective of global bridge PCD, highlighting that defects in the PCD of a certain component can impact the relationships between its vertical center and that of other components. SS models generally employ a slicing window to extract local features and the window size varies significantly across different models. Within each slicing window, it is highly probable that the PCD of certain components is largely truncated while that of other components remains relatively complete, introducing significant uncertainties to the relationships between vertical centers. Additionally, the slicing window is randomly positioned on the entire bridge PCD during training, further rendering the vertical distribution within it more unpredictable.

To make use of the vertical relationships of bridge components while eliminating the aforementioned uncertainties, a judging condition can be defined to select components that satisfy the vertical distribution pattern in a given space. Then more focus is allocated to selected components whose vertical relationships with others are reversely predicted by the model during training. Additionally, wrongly predicted relationships between vertically non-adjacent elements in abutments, girders, decks, and parapets lead to larger errors. As the vertical distance between elements in the list increases, the magnitude of errors tends to amplify.

4.2 | SOL function

4.2.1 | Definition

The suitability and robustness of CE loss made it prevalent in SS tasks. Therefore, CE was selected as the basis to deploy SOC, leading to the formulation of a new loss function named SOL. Essentially, SOL shares the same core concept as WCE, which leverages the geometric relationship of bridge components to allocate the weights.

The loss of each slicing window is calculated individually, and the total loss can then be acquired by summing up all individual losses. The SOL function for every slicing window is defined as

$$L = -\sum_{c=1}^{M} w_c y_c \log(p_c) \tag{6}$$

where M denotes the number of component classes in each slicing window, w_c represents the weight given to the cth class, while y_c and p_c are label vectors of ground truth and predicted results, respectively.

In detail, w_c is defined as

$$w_c = 1 + \beta_c + \sum_{i=1, i\neq c}^{M} \alpha_{c,i} \tag{7}$$

where β_c and $\alpha_{c,i}$ are named as the horizontal and vertical weighting factors, respectively, according to their functions, and they can be calculated by:

$$\beta_c = \begin{cases} \beta, & condition\ C\ is\ True \\ 0, & otherwise \end{cases} \tag{8}$$

$$\alpha_{c,i}|_{|c-i|=k} = \begin{cases} k \cdot \alpha, & condition\,(c,i)\ is\ True \\ 0, & otherwise \end{cases} \tag{9}$$

In Equation (5), β_c is employed to punish the cth class if it is out of the right region horizontally, which makes "*condition C*" true; and in Equation (9), $\alpha_{c,i}$ is utilized to penalize the cth and ith class when "*condition*(c,i)" is true, indicating a reversed prediction of the vertical relationship between class c and i. The calculation process of the conditions for horizontal and vertical distribution is depicted in Figure 7, and the pseudo-code is presented in Table 4. Additionally, the vertical relationship between non-adjacent elements in the sequence of abutment, girder, deck, and parapet is more prominent. The further apart a pair of elements are in the sequence, the stronger the vertical relationship becomes. Hence, Equation (9) imposes a heavier penalty when a stronger vertical relationship is inverted.

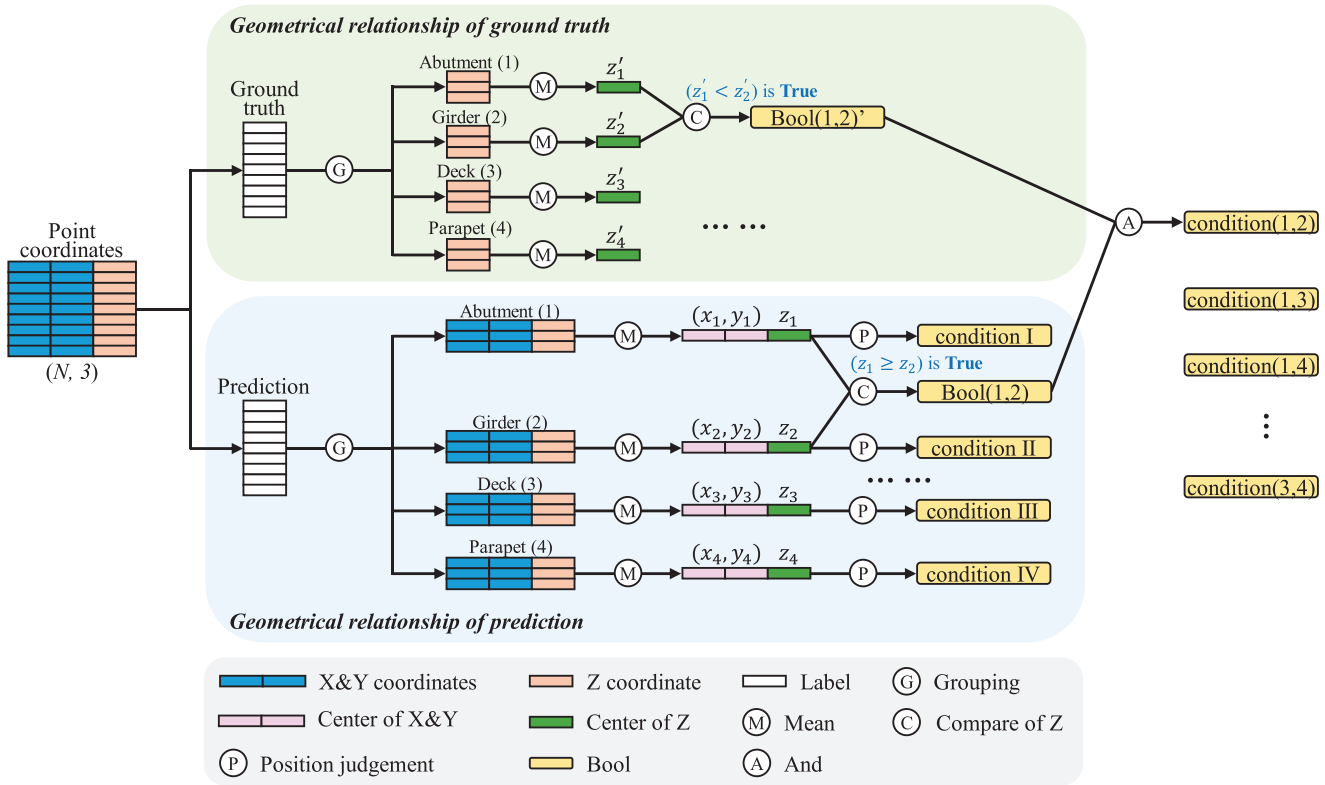

FIGURE 7 Calculation process of the conditions for horizontal and vertical distribution.

TABLE 4 Pseudo-code for conditions calculation.

Algorithm 1 Calculation process of horizontal and vertical conditions

Require: The numbers of component classes and points M and N; the horizontal and vertical weighting coefficients β and α; the 3D coordinates of points $C = \{(x^k, y^k, z^k)\}_{k=1}^{N}$; the label vectors of prediction and ground truth L_{pre} and L_{gt}; the weights for each class during training $W = \{w^k = 1\}_{k=1}^{M}$; the horizontal location judging function $H(coordinate, class)$; the cross entropy loss function $G(prediction, ground\_truth, weight)$

1: Group points based on L_{pre} and L_{gt} and calculate centers from C
$Z_{gt} = \{z_{gt}^k\}_{k=1}^{M}$ $C_{pre} = \{(x_{pre}^k, y_{pre}^k, z_{pre}^k)\}_{k=1}^{M}$
2: **for** iteration $i = 1$ to M **do**
3: **if** not $H(coordinate = (x_{pre}^i, y_{pre}^i), label = i)$ **then**
4: $w^i = w^i + \beta$
5: **end if**
6: **for** iteration $j = (i+1)$ to M **do**
7: **if** $z_{gt}^j > z_{gt}^i$ and $z_{pre}^j < z_{pre}^i$ **then**
8: $w^i = w^i + (j-i) \times \alpha$, $w^j = w^j + (j-i) \times \alpha$
9: **end if**
10: **end for**
11: **end for**
12: $loss = G(prediction = L_{pre}, ground\_truth = L_{gt}, weight = W)$
13: **return** *loss*

4.2.2 | Calibration

In most studies, "others" is generally set as a class segmented by SS models instead of being deleted manually before training to improve the efficiency and automation of the SS process. Therefore, the performance of segmenting the raw dataset is the most prioritized consideration for an SS model. The raw dataset is fed into the SS model to experiment on the performance by varying α and β, which are the hyperparameters in the proposed SOL function. The experiment results are shown in Table 5. It can be observed that $\alpha = 200$, $\beta = 300$ worked best

TABLE 5 Results (%) of different values for the hyperparameter α and β in the structure-oriented loss (SOL) function.

α	β	Overall accuracy (OA)	Mean class accuracy (mAcc)	Mean IoU (mIoU)	Intersection over union (IoU) for each class				
					Abutment	Girder	Deck	Parapet	Others
150	200	81.42	83.57	66.34	40.07	82.42	67.91	74.00	67.32
150	300	79.01	82.98	65.44	38.35	73.44	61.91	**83.54**	69.96
150	400	79.65	81.02	63.62	36.15	82.27	67.77	69.55	62.35
200	200	82.27	84.85	68.45	49.98	83.94	64.02	78.38	65.92
200	300	**87.67**	86.65	**74.43**	50.40	85.10	**75.73**	82.84	**78.06**
200	400	82.64	87.32	70.20	52.77	78.53	66.03	82.73	70.93
200	500	81.16	85.19	64.30	49.36	84.89	55.15	64.05	68.06
200	600	78.93	80.12	62.68	42.87	75.16	64.77	62.98	67.61
250	300	82.66	**87.48**	66.04	55.38	**87.61**	56.28	61.70	69.23
250	400	79.34	85.01	63.78	48.69	79.58	59.79	67.81	63.03
250	500	81.59	84.64	66.39	45.51	79.28	69.68	67.10	70.37
300	200	82.50	**87.48**	65.39	**60.16**	85.16	53.18	59.59	68.88
300	300	83.51	86.41	68.53	53.30	84.80	62.75	72.58	69.24
300	400	82.59	84.28	68.99	49.28	77.86	72.19	73.96	71.65

TABLE 6 Results (%) of different loss functions on the dataset.

Loss function	OA	mAcc	mIoU	IoU for each class				
				Abutment	Girder	Deck	Parapet	Others
Negative log likelihood (NLL)	76.89	82.92	60.52	33.82	79.25	56.09	72.47	60.96
Cross entropy (CE)	81.14	86.62	65.76	44.96	82.40	57.19	76.08	68.15
Weighted CE (WCE)	80.68	83.28	63.82	37.04	78.79	60.14	73.51	69.62
Focal loss (FL)	76.30	83.09	59.90	32.54	74.97	63.41	68.11	60.47
Structure-aware loss (SAL)	19.95	22.68	8.11	1.28	14.16	1.43	3.32	20.39
SOL	**87.67**	**86.65**	**74.43**	**50.40**	**85.10**	**75.73**	**82.84**	**78.06**

in the experiments, so α and β were set to 200 and 300, respectively, for the SOL function.

5 | COMPARISON

5.1 | Performance comparison

The raw dataset was fed into the backbone in which the loss function is replaced by the five selected loss functions, and their results are compared with the SOL function in Table 6. Comparing the results in Tables 5 and 6, SOL outperforms other loss functions when α and β are set to multiple values, demonstrating its robustness and effectiveness. Moreover, the best SOL exceeds the five selected loss functions in every category and the improvements on two of the three overall evaluation metrics are significant (OA by 6.53% and mIoU by 8.67%). In particular, the IoU of "others" was improved by 8.44%, which helped to automate the time-consuming denoising process and contributed greatly to the comprehensive performance of the model since "others" accounts for the largest proportion according to Figure 2. It is worth noticing that SAL does not exhibit SS ability when collaborating with PointNet++.

5.2 | Discussion

The generally lower metrics of our dataset demonstrate that this dataset is more difficult to segment than other datasets (Hyunsoo Kim & Kim, 2020; Lee et al., 2021; Xiao et al., 2024; Yang et al., 2022) using deep learning models. Bridges 1 to 9 are located in the countryside, so two main challenges exist for SS: (1) Too much noise was inevitably included during scanning as elaborated in Figure 1 and Table 2, such as vegetation, electric poles, houses, and so forth. (2) The bridges are built over rivers so the tripod of the scanner has to be put in rivers when scanning the bottom part of the bridges. The water flow caused the scanner to constantly shake and thus influenced the PCD quality.

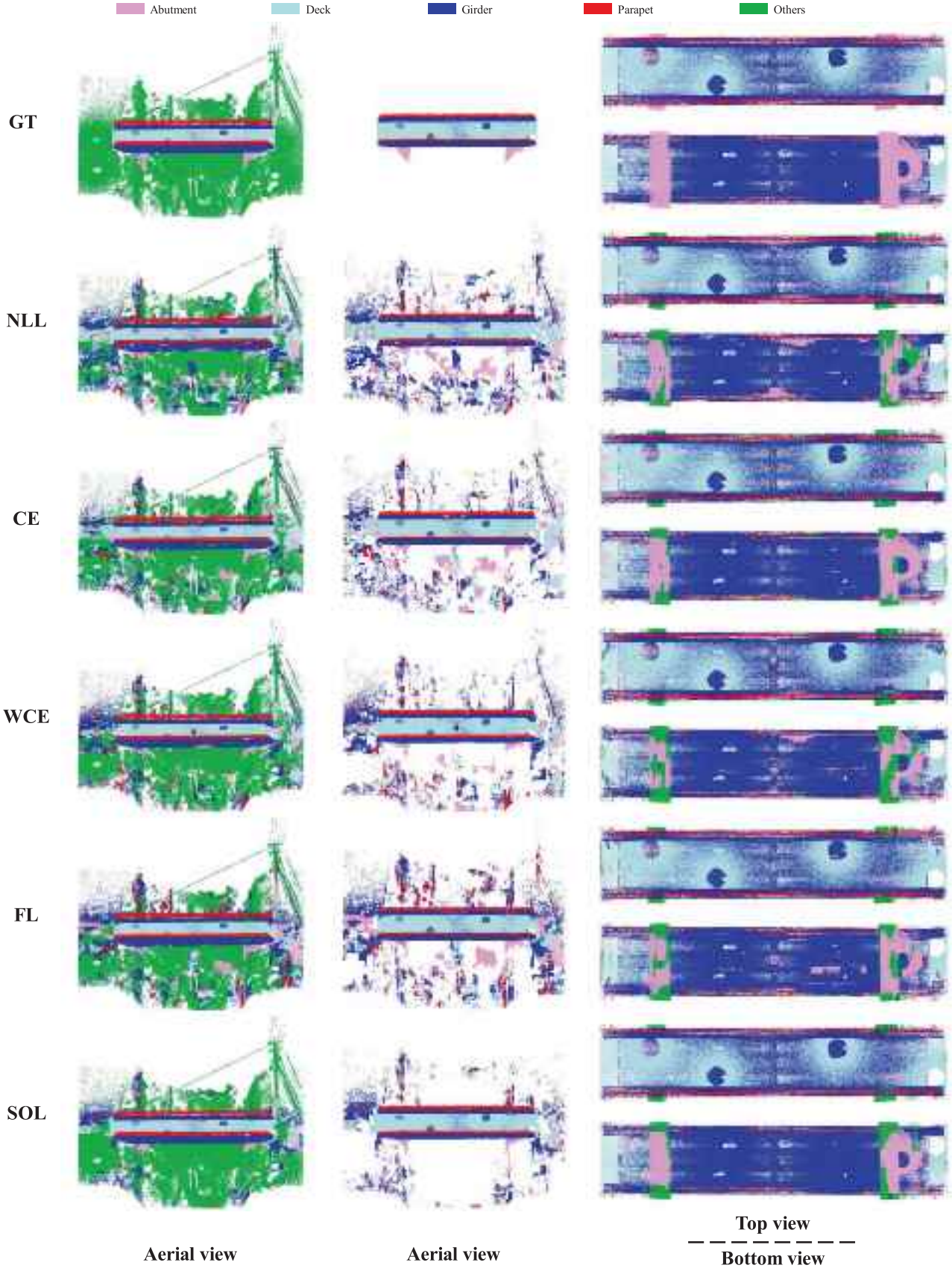

FIGURE 8 Visualization of prediction results on Bridge 7. Left: complete PCD; middle: full-scale PCD without points that are predicted as "others"; right: PCD in the position of bridge components. CE, cross entropy; FL, focal loss; GT, ground truth; NLL, negative log likelihood; SAL, structure-aware loss; SOL, structure-oriented loss; WCE, weighted cross entropy.

TABLE 7 Training time per epoch of different loss functions.

Loss function	Training time (s)
NLL	1019.6
CE	1014.6
WCE	1032.6
FL	1006.4
SAL	1132.1
SOL	1071.5

The predicted results of Bridge 7 by NLL, CE, WCE, FL, and SOL are visualized in Figure 8. The significant differences between SOL and other loss functions demonstrate the predominance of SOL. In detail, SOL has better performance in classifying "others," especially for the area under the bridge and the electric pole. SOL almost categorizes the whole electric pole as "others," while other loss functions do not. As for the bridge components, SOL identifies a more complete deck and has the best performance for girders, while other loss functions wrongly predicted more abutments as girders. However, failure still persists at the interfaces of neighboring bridge components despite the exceptional overall performance of SOL.

The predominance of SOL demonstrates its strength. SOL is calculated in each individual slicing window, which guarantees enough training samples. Furthermore, judging conditions from both horizontal and vertical perspectives help SOL to examine each part meticulously, rendering as many features as possible. Due to the weight calculation process, SOL needs more computation than vanilla CE. The training time per epoch of each loss function is calculated as shown in Table 7. The training time of SOL falls in the middle among all loss functions, which is acceptable considering its better performance.

In particular, the IoU of the abutment was the lowest among all classes for every loss function. This is primarily because of the extensive weeds on abutments whose color information influenced the model. Figure 9 shows the confusion matrix of SOL, which had the best performance among all loss functions. As depicted in the top right corner of the confusion matrix, 29% of the points in abutments were misclassified into "others."

Moreover, the failure or restriction of other loss functions in SS is also worth noticing. SAL failed to segment bridge PCD when used with the PointNet++ framework, and that was mainly because SAL was originally designed for one specific model (Graham et al., 2018). The dimension of embedding in SAL was set to four to fit the specific model but that of PointNet++ is 128 (Qi, Yi, et al., 2017). Therefore, the overly solidified configuration, such as

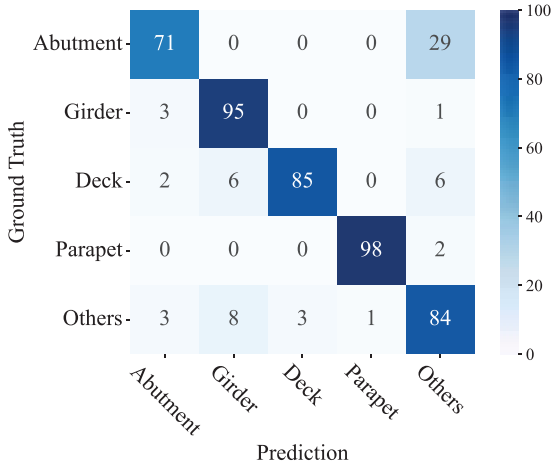

FIGURE 9 Confusion matrix of structure-oriented loss.

hyperparameters, of loss functions can sometimes hinder them from being generalized to other deep learning frameworks. In Table 6, CE comprehensively outperforms WCE, which is contrary to the results obtained by Yang et al. (2022). Wrongly allocated weights can damage the SS ability of vanilla CE. Great care should be taken when calculating and allocating weights.

The robustness of SOL has been verified, which guarantees its compatibility with various deep learning models and applicability to diverse bridge PCD datasets. Theoretically, the proposed SOL can cooperate with any deep learning model based on distinct practical scenarios and improve its outcomes. When generalizing SOL to other PCD datasets whose bridges have different layouts or more component categories (e.g., piers, pier caps, diaphragms, lamp posts, cables, etc.), the subsequent principles should be followed. (1) Slightly calibrate the two hyperparameters in SOL from $\alpha = 200$ and $\beta = 300$ to achieve better performance. (2) Horizontally, directly add minimum rectangles of extra components to the calculation process as depicted in Figure 7. (3) Vertically, divide categories that are located at similar vertical heights into the same groups and only compare components between groups rather than within the same groups during training. For example, set piers, pier caps, and abutments in a group, girders and diaphragms in a group, parapets, lamp posts, cables in a group, and so forth.

6 | CONCLUSION

Loss functions in learning-based SS models for PCD have been persistently underrated, especially in the realm of bridge PCD. In this paper, an SOC was proposed that quantitatively represented the spatial relationship between bridge components and designs' corresponding methods

to utilize them to orient the training process of deep learning models. Then an SOL was defined according to SOC where CE served as the baseline. A PCD dataset containing 12 bridges was established, and PointNet++ was selected as the backbone framework. According to the comparison between SOL and five existing cutting-edge loss functions, SOL outperformed these functions comprehensively, and the better results for "others" were particularly satisfactory. Additionally, SOL improved the capacity of PointNet++ without significantly increasing the training time. Therefore, the effectiveness and efficiency of SOC and SOL were demonstrated, showing their great potential to improve the performance of other SS models. Moreover, this paper proved that appropriate loss functions are no less important than other sections in the learning framework when considering the contribution to the comprehensive performance of SS models. More efforts should be put into loss functions in the future to fulfill this research gap so that the capabilities of deep learning models can be improved comprehensively.

The future work could focus on data fusion where PCD and images are jointly analyzed. Recently, extensive studies on crack and corrosion detection using image analysis have been published but they have primarily relied on 2D data and have not fully utilized the information inherent to the entire 3D structure of each bridge (Chu & Chun, 2024; Chun & Kikuta, 2024; Chun et al., 2021). The detection performance could be further improved with the help of advanced techniques (Pereira et al., 2020), such as the neural dynamic classification algorithm (Rafiei & Adeli, 2017), the dynamic ensemble learning algorithm (Alam et al., 2020), and the self-supervised learning approach (Rafiei et al., 2024a, 2024b). Critical factors in PCD, such as the spatial distribution of damage and the relationships between components, provide new value to bridge inspection and maintenance when PCD is analyzed in combination with images. In any case, this study contributes to PCD SS, which enables detailed component identification with 3D positional information, and by facilitating the creation and integration of BIM, the study is expected to contribute to more effective bridge maintenance and management.

ACKNOWLEDGMENTS

This work was partially supported by the Council for Science, Technology and Innovation (CSTI), Cross-Ministerial Strategic Innovation Promotion Program (SIP), the 3rd period of SIP "Smart Infrastructure Management System" Grant Number JPJ012187 (Funding agency: Public Works Research Institute) and JSPS Grant-in-Aid for Scientific Research Grant Numbers 21H01417, 22H01561, and 23H00198, and China Scholarship Council Grant Number 202306210048.

ORCID

Chao Lin https://orcid.org/0000-0002-1517-8392
Shitao Zheng https://orcid.org/0009-0008-1357-6080
Xianfeng Li https://orcid.org/0000-0003-2200-1453
Pang-jo Chun https://orcid.org/0000-0002-9755-8435

REFERENCES

Adeli, H., & Kim, H. (2009). *Wavelet-based vibration control of smart buildings and bridges*. CRC Press. https://doi.org/10.1201/9780367813451

Adeli, H., & Saleh, A. (1997). Optimal control of adaptive/smart bridge structures. *Journal of Structural Engineering*, *123*(2), 218–226. https://doi.org/10.1061/(ASCE)0733-9445(1997)123:2(218)

Adeli, H., & Saleh, A. (1999). *Control, optimization, and smart structures: High-performance bridges and buildings of the future*. John Wiley. https://www.wiley.com/en-us/Control%2C+Optimization%2C+and+Smart+Structures%3A+High-Performance+Bridges+and+Buildings+of+the+Future-p-9780471350941

Alam, K. M., Siddique, N., & Adeli, H. (2020). A dynamic ensemble learning algorithm for neural networks. *Neural Computing and Applications*, *32*(10), 8675–8690. https://doi.org/10.1007/s00521-019-04359-7

American Society of Civil Engineers (ASCE). (2021). *ASCE's 2021 infrastructure report card*. https://infrastructurereportcard.org/cat-item/bridges-infrastructure/

Chen, D., Li, Y., Tao, J., Li, Y., Zhang, S., Shan, X., Wang, T., Qiao, Z., Zhao, R., Fan, X., & Zhou, Z. (2024). Deep learning-based segmentation model for permeable concrete meso-structures. *Computer-Aided Civil and Infrastructure Engineering*, *39*, 3626–3645. https://doi.org/10.1111/mice.13300

Chern, W. C., Nguyen, T. V., Asari, V. K., & Kim, H. (2023). Impact of loss functions on semantic segmentation in far-field monitoring. *Computer-Aided Civil and Infrastructure Engineering*, *38*(3), 372–390. https://doi.org/10.1111/mice.12832

Chu, H., & Chun, P. J. (2024). Fine-grained crack segmentation for high-resolution images via a multiscale cascaded network. *Computer-Aided Civil and Infrastructure Engineering*, *39*(4), 575–594. https://doi.org/10.1111/mice.13111

Chu, H., Long, L., Guo, J., Yuan, H., & Deng, L. (2024). Implicit function-based continuous representation for meticulous segmentation of cracks from high-resolution images. *Computer-Aided Civil and Infrastructure Engineering*, *39*, 539–558. https://doi.org/10.1111/mice.13052

Chun, P. J., & Kikuta, T. (2024). Self-training with Bayesian neural networks and spatial priors for unsupervised domain adaptation in crack segmentation. *Computer-Aided Civil and Infrastructure Engineering*, *39*(17), 2642–2661. https://doi.org/10.1111/mice.13315

Chun, P. J., Yamane, T., & Maemura, Y. (2021). A deep learning-based image captioning method to automatically generate comprehensive explanations of bridge damage. *Computer-Aided Civil and Infrastructure Engineering*, *37*(11), 1387–1401. https://doi.org/10.1111/mice.12793

Deng, L., Yuan, H., Long, L., Chun, P. J., Chen, W., & Chu, H. (2024). Cascade refinement extraction network with active boundary loss for segmentation of concrete cracks from high-resolution images. *Automation in Construction*, *162*, 105410. https://doi.org/10.1016/j.autcon.2024.105410

Ghaedi, K., Ibrahim, Z., Adeli, H., & Javanmardi, A. (2017). Invited review: Recent developments in vibration control of building and bridge structures. *Journal of Vibroengineering*, *19*(5), 3564–3580. https://doi.org/10.21595/jve.2017.18900

Graham, B., Engelcke, M., & Van Der Maaten, L. (2018). 3D semantic segmentation with submanifold sparse convolutional networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT (pp. 9224–9232). https://doi.org/10.1109/cvpr.2018.00961

Gutierrez Soto, M., & Adeli, H. (2019). Semi-active vibration control of smart isolated highway bridge structures using replicator dynamics. *Engineering Structures*, *186*, 536–552. https://doi.org/10.1016/j.engstruct.2019.02.031

Han, X., Dong, Z., & Yang, B. (2021). A point-based deep learning network for semantic segmentation of MLS point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing*, *175*, 199–214. https://doi.org/10.1016/j.isprsjprs.2021.03.001

Hattori, K., Oki, K., Sugita, A., Sugiyama, T., & Chun, P. J. (2024). Deep learning-based corrosion inspection of long-span bridges with BIM integration. *Heliyon*, *10*(15), e35308. https://doi.org/10.1016/j.heliyon.2024.e35308

Hu, Q., Yang, B., Xie, L., Rosa, S., Guo, Y., Wang, Z., Trigoni, N., & Markham, A. (2020). RandLA-Net: Efficient semantic segmentation of large-scale point clouds. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Seattle, WA (pp. 11108–11117). https://doi.org/10.1109/cvpr42600.2020.01112

Huang, Y., Liu, Y., Liu, F., & Liu, W. (2024). A lightweight feature attention fusion network for pavement crack segmentation. *Computer-Aided Civil and Infrastructure Engineering*, *39*, 2811–2825. https://doi.org/10.1111/mice.13225

Hui, L., & Belkin, M. (2020). *Evaluation of neural architectures trained with square loss vs cross-entropy in classification tasks*. arXiv preprint arXiv:2006.07322. https://doi.org/10.48550/arXiv.2006.07322

Jing, Y., Sheil, B., & Acikgoz, S. (2022). Segmentation of large-scale masonry arch bridge point clouds with a synthetic simulator and the BridgeNet neural network. *Automation in Construction*, *142*, 104459. https://doi.org/10.1016/j.autcon.2022.104459

Jing, Y., Sheil, B., & Acikgoz, S. (2024). A lightweight Transformer-based neural network for large-scale masonry arch bridge point cloud segmentation. *Computer-Aided Civil and Infrastructure Engineering*, *39*(16), 2427–2438. https://doi.org/10.1111/mice.13201

Kim, H., & Adeli, H. (2005). Wavelet-hybrid feedback linear mean squared algorithm for robust control of cable-stayed bridges. *Journal of Bridge Engineering*, *10*(2), 116–123. https://doi.org/10.1061/(ASCE)1084-0702(2005)10:2(116)

Kim, H., & Kim, C. (2020). Deep-learning-based classification of point clouds for bridge inspection. *Remote Sensing*, *12*(22), 3757. https://doi.org/10.3390/rs12223757

Kim, H., Yoon, J., & Sim, S. H. (2020). Automated bridge component recognition from point clouds using deep learning. *Structural Control and Health Monitoring*, *27*(9), e2591. https://doi.org/10.1002/stc.2591

Kunlamai, T., Yamane, T., Suganuma, M., Chun, P. J., & Okatani, T. (2024). Improving visual question answering for bridge inspection by pre-training with external data of image-text pairs. *Computer-Aided Civil and Infrastructure Engineering*, *39*(3), 345–361. https://doi.org/10.1111/mice.13086

Lamas, D., Justo, A., Soilán, M., Cabaleiro, M., & Riveiro, B. (2023). Instance and semantic segmentation of point clouds of large metallic truss bridges. *Automation in Construction*, *151*, 104865. https://doi.org/10.1016/j.autcon.2023.104865

Landrieu, L., & Boussaha, M. (2019). Point cloud oversegmentation with graph-structured deep metric learning. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Long Beach, CA (pp. 7440–7449). https://doi.org/10.1109/cvpr.2019.00762

Landrieu, L., & Simonovsky, M. (2018). Large-scale point cloud semantic segmentation with superpoint graphs. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT (pp. 4558–4567). https://doi.org/10.1109/cvpr.2018.00479

Lee, J. S., Park, J., & Ryu, Y. M. (2021). Semantic segmentation of bridge components based on hierarchical point cloud model. *Automation in Construction*, *130*, 103847. https://doi.org/10.1016/j.autcon.2021.103847

Liang, H., Yeoh, J. K. W., & Chua, D. K. H. (2024). Material augmented semantic segmentation of point clouds for building elements. *Computer-Aided Civil and Infrastructure Engineering*, *39*(15), 2312–2329. https://doi.org/10.1111/mice.13198

Liang, Z., Yang, M., Li, H., & Wang, C. (2020). 3D instance embedding learning with a structure-aware loss function for point cloud segmentation. *IEEE Robotics and Automation Letters*, *5*(3), 4915–4922. https://doi.org/10.1109/lra.2020.3004802

Lin, T. Y., Goyal, P., Girshick, R., He, K., & Dollár, P. (2017). Focal loss for dense object detection. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Venice, Italy (pp. 2980–2988). https://doi.org/10.1109/iccv.2017.324

Lin, Y. C., & Habib, A. (2022). Semantic segmentation of bridge components and road infrastructure from mobile LiDAR data. *ISPRS Open Journal of Photogrammetry and Remote Sensing*, *6*, 100023. https://doi.org/10.1016/j.ophoto.2022.100023

Liu, K., Gao, Z., Lin, F., & Chen, B. M. (2020). FG-Net: Fast large-scale LiDAR point clouds understanding network leveraging correlated feature mining and geometric-aware modelling. *2021 IEEE International Conference on Robotics and Automation (ICRA)*, Xi'an, China (pp. 12896–12902). https://doi.org/10.1109/icra48506.2021.9561496

Lu, R., & Brilakis, I. (2019). Digital twinning of existing reinforced concrete bridges from labelled point clusters. *Automation in Construction*, *105*, 102837. https://doi.org/10.1016/j.autcon.2019.102837

Lu, R., Brilakis, I., & Middleton, C. R. (2019). Detection of structural components in point clouds of existing RC bridges. *Computer-Aided Civil and Infrastructure Engineering*, *34*(3), 191–212. https://doi.org/10.1111/mice.12407

Milioto, A., Vizzo, I., Behley, J., & Stachniss, C. (2019). Rangenet++: Fast and accurate lidar semantic segmentation. *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Macau, China (pp. 4213–4220). https://doi.org/10.1109/iros40897.2019.8967762

Milletari, F., Navab, N., & Ahmadi, S. A. (2016). V-Net: Fully convolutional neural networks for volumetric medical image segmentation. In B. Chen (Ed.), *2016 fourth international conference on 3D vision (3DV)* (pp. 565–571). IEEE. https://doi.org/10.1109/3dv.2016.79

Ministry of Land, Infrastructure, Transport and Tourism (MLIT). (2024). *Road maintenance in Japan: Problems and solutions*. https://www.mlit.go.jp/road/road_e/s3_maintenance.html

Noichl, F., Collins, F. C., Braun, A., & Borrmann, A. (2024). Enhancing point cloudsemantic segmentation in the data-scarce domain of industrial plants throughsynthetic data. *Computer-Aided Civil and Infrastructure Engineering*, *39*, 1530–1549. https://doi.org/10.1111/mice.13153

Park, H. S., Lee, H. M., Adeli, H., & Lee, I. (2007). A new approach for health monitoring of structures: Terrestrial laser scanning. *Computer-Aided Civil and Infrastructure Engineering*, *22*(1), 19–30. https://doi.org/10.1111/j.1467-8667.2006.00466.x

Park, S. W., Park, H. S., Kim, J. H., & Adeli, H. (2015). 3D displacement measurement model for health monitoring of structures using a motion capture system. *Measurement*, *59*, 352–362. https://doi.org/10.1016/j.measurement.2014.09.063

Pereira, D. R., Piteri, M. A., Souza, A. N., Papa, J., & Adeli, H. (2020). FEMa: A finite element machine for fast learning. *Neural Computing and Applications*, *32*(10), 6393–6404. https://doi.org/10.1007/s00521-019-04146-4

Phan, T. H., & Yamamoto, K. (2020). *Resolving class imbalance in object detection with weighted cross entropy losses*. arXiv preprint arXiv:2006.01413. https://doi.org/10.48550/arXiv.2006.01413

Qi, C. R., Su, H., Mo, K., & Guibas, L. J. (2017). Pointnet: Deep learning on point sets for 3d classification and segmentation. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, HI (pp. 652–660). https://doi.org/10.1109/cvpr.2017.16

Qi, C. R., Yi, L., Su, H., & Guibas, L. J. (2017). PointNet++: Deep hierarchical feature learning on point sets in a metric space. *Advances in Neural Information Processing Systems*, *30*, Long Beach, CA.

Qin, S., Qi, T., Deng, T., & Huang, X. (2024). Image segmentation using Vision Transformer for tunnel defect assessment. *Computer-Aided Civil and Infrastructure Engineering*, *39*, 3243–3268. https://doi.org/10.1111/mice.13181

Rafiei, M. H., & Adeli, H. (2017). A new neural dynamic classification algorithm. *IEEE Transactions on Neural Networks and Learning Systems*, *28*(12), 3074–3083. https://doi.org/10.1109/tnnls.2017.2682102

Rafiei, M. H., Gauthier, L., Adeli, H., & Takabi, D. (2024a). Self-supervised learning for electroencephalography. *IEEE Transactions on Neural Networks and Learning Systems*, *35*(2), 1457–1471. https://doi.org/10.1109/tnnls.2022.3190448

Rafiei, M. H., Gauthier, L. V., Adeli, H., & Takabi, D. (2024b). Self-supervised learning for near-wild cognitive workload estimation. *Journal of Medical Systems*, *48*(1), 107. https://doi.org/10.1007/s10916-024-02122-7

Riveiro, B., DeJong, M. J., & Conde, B. (2016). Automated processing of large point clouds for structural health monitoring of masonry arch bridges. *Automation in Construction*, *72*, 258–268. https://doi.org/10.1016/j.autcon.2016.02.009

Romero-Jarén, R., & Arranz, J. J. (2021). Automatic segmentation and classification of BIM elements from point clouds. *Automation in Construction*, *124*, 103576. https://doi.org/10.1016/j.autcon.2021.103576

Sun, L., Yang, Y., Zhou, G., Chen, A., Zhang, Y., Cai, W., & Li, L. (2024). An integration-competition network for bridge crack segmentation under complex scenes. *Computer-Aided Civil and Infrastructure Engineering*, *39*, 617–634. https://doi.org/10.1111/mice.13113

Wang, N., & Adeli, H. (2015). Self-constructing wavelet neural network algorithm for nonlinear control of large structures. *Engineering Applications of Artificial Intelligence*, *41*, 249–258. https://doi.org/10.1016/j.engappai.2015.01.018

Wang, Q., Ma, Y., Zhao, K., & Tian, Y. (2020). A comprehensive survey of loss functions in machine learning. *Annals of Data Science*, *9*, 187–212. https://doi.org/10.1007/s40745-020-00253-5

Wang, Y., Sun, Y., Liu, Z., Sarma, S. E., Bronstein, M. M., & Solomon, J. M. (2019). Dynamic graph CNN for learning on point clouds. *ACM Transactions on Graphics (TOG)*, *38*(5), 1–12. https://doi.org/10.1145/3326362

Wen, Y., Zhang, K., Li, Z., & Qiao, Y. (2016). A discriminative feature learning approach for deep face recognition. In B. Leibe, J. Matas, N. Sebe, & M. Welling (Eds.), *Lecture notes in computer science: Vol. 9911. Computer vision—ECCV 2016* (pp. 499–515). Springer. https://doi.org/10.1007/978-3-319-46478-7_31

Xia, T., Yang, J., & Chen, L. (2022). Automated semantic segmentation of bridge point cloud based on local descriptor and machine learning. *Automation in Construction*, *133*, 103992. https://doi.org/10.1016/j.autcon.2021.103992

Xiao, J. L., Fan, J. S., Liu, Y. F., Li, B. L., & Nie, J. G. (2024). Region of interest (ROI) extraction and crack detection for UAV-based bridge inspection using point cloud segmentation and 3D-to-2D projection. *Automation in Construction*, *158*, 105226. https://doi.org/10.1016/j.autcon.2023.105226

Yang, X., del Rey Castillo, E., Zou, Y., & Wotherspoon, L. (2023). Semantic segmentation of bridge point clouds with a synthetic data augmentation strategy and graph-structured deep metric learning. *Automation in Construction*, *150*, 104838. https://doi.org/10.1016/j.autcon.2023.104838

Yang, X., del Rey Castillo, E., Zou, Y., Wotherspoon, L., & Tan, Y. (2022). Automated semantic segmentation of bridge components from large-scale point clouds using a weighted superpoint graph. *Automation in Construction*, *142*, 104519. https://doi.org/10.1016/j.autcon.2022.104519

Yu, G., & Adeli, H. (1993). Object-oriented finite element analysis using EER model. *Journal of Structural Engineering*, *119*(9), 2763–2781. https://doi.org/10.1061/(asce)0733-9445(1993)119:9(2763)

Yu, J., Jiang, Y., Wang, Z., Cao, Z., & Huang, T. (2016). Unitbox: An advanced object detection network. In A. Hanjalic, C. Snoek, & M. Worring (Eds.), *Proceedings of the 24th ACM international conference on multimedia* (pp. 516–520). ACM. https://doi.org/10.1145/2964284.2967274

Zhu, B., Jiang, Z., Zhou, X., Li, Z., & Yu, G. (2019). Class-balanced grouping and sampling for point cloud 3d object detection. arXiv preprint arXiv:1908.09492. https://doi.org/10.48550/arXiv.1908.09492

How to cite this article: Lin, C., Abe, S., Zheng, S., Li, X., & Chun, P.-j. (2025). A structure-oriented loss function for automated semantic segmentation of bridge point clouds. *Computer-Aided Civil and Infrastructure Engineering*, *40*, 801–816. https://doi.org/10.1111/mice.13422