A Framework of Reconstructing Piping Systems on Class-imbalanced 3D Point Cloud Data from Construction Sites

Yilong Chen[1], Seongyong Kim[1], Yonghan Ahn[2], and Yong K. Cho[1]

[1]School of Civil and Environmental Engineering, Georgia Institute of Technology, U.S.A
ychen3339@gatech.edu, skim3310@gatech.edu, yong.cho@ce.gatech.edu

[2]Department of Architecture and Architectural Engineering, Hanyang University, South Korea
yhahn@hanyang.ac.kr

Abstract –

In construction environments, modifications to the dimensions, positioning, and trajectory of plumbing infrastructure within edifices are frequently necessitated by on-site conditions and pragmatic installation procedures. Recent advancements in Scan-to-BIM technology have streamlined pipe construction processes by monitoring development through a 3D model. However, existing 3D point cloud processing methods rely heavily on given local geometric information to distinguish pipes from adjacent components. Furthermore, point clouds originating from construction environments are mostly class-imbalanced data which could negatively impact the date-driven approach. This paper proposed a novel framework for segmenting and reconstructing piping systems utilizing raw 3D point cloud data acquired from construction sites, addressing the aforementioned challenges. The data firstly undergoes preprocesssing, including the elimination of redundant points, rotational adjustments, and sampling procedures. Subsequently, a point cloud semantic segmentation network is trained to predict the per-point class labels after adding local features and mitigating the class imbalance issues. Finally, Efficient RANSAC is employed to identify cylinder-shaped pipes based on the prediction outcomes. The proposed framework shows superior performance compared to existing semantic segmentation methods and exhibits considerable promise for piping system reconstruction.

Keywords –

Point cloud segmentation; pipe reconstruction; construction sites; deep learning

1 Introduction

Mechanical, electrical, and plumbing (MEP) system makes up a large amount of construction cost and asset value. Recent advances in laser scanning and Building Information Modeling (BIM) technologies have demonstrated the great potential of Scan-to-BIM application for MEP systems. However, compared with visible structural and architectural components like columns, walls, floors, ceilings, windows, and doors in an as-built interior environment, pipes and ducts are less frequently extracted, analyzed, and reconstructed [1]. The workflow for extracting MEP systems directly from raw point cloud data with architectural and structural components in close proximity remains to be discussed [2].

3D point cloud data applications in the construction field, such as geometry quality inspection and semantic model reconstruction, are limited by requiring the existence of a BIM model as a precondition. Construction sites are ideal places to capture the point cloud data before pipes are covered by other structural components to record the size, location and route [1]. However, only a partial scan of the pipe surface might be available due to the space limitation for scanning [3] .

Machine learning and deep learning are fast-growing techniques in 3D point cloud reconstruction. Current state-of-the-art deep learning model, like PointNet [4], uses no local structure information within neighboring points [5] and thus cannot correctly separate pipes from neighboring contexts, like ceilings and walls, in terms of semantic segmentation. There are other works focusing on classifying point cloud data, including pipes by traditional machine learning methods like Support Vector Machine (SVM) [6], AdaBoost [7], Conditional Random Field and Markov Random Field [2], which do not directly feed raw point cloud data into the neural network but instead calculate and use predefined geometric features to train the classifier.

This paper proposes a framework to segment and reconstruct a pipe model using raw 3D point cloud data from construction sites with architectural and structural components in close proximity. An overview is shown in Figure 1. The onsite raw scans are first preprocessed by removing redundant information, applying rotation to align data to the x-axis and y-axis directions and sampling points for training from evenly divided blocks. Then the training data sets containing six classes of structural components are passed to the proposed neural network. The backbone network is robust to semantic segmentation tasks on pipes by adding local features like intensity and normal, modifying the classifier such as adding weight for loss function and using a threshold for prediction, and using ROC AUC as alternative evaluation metrics. The output of the deep learning model is then processed with Efficient RANSAC algorithms to detect cylinder shapes with heavy noise in a constrained time. The obtained cylinder's parameters could then be used to create a potential BIM model. Finally, the proposed framework is verified by comparing reconstruction results using ground truth data collected from a laser scanner on a construction site with the prediction result made by the neural network. The key contributions of this work can be delineated as follows: (a) the study summarized research related to reconstruction of piping systems and extended them to be generally applicable as a framework, (b) the research proposes an procedure to modify and tune a deep learning model suitable for partially scanned objects, highly imbalanced, and over-noisy datasets, and (c) the experiment explored and assessed the proposed semantic segmentation and shape detection framework by empirically employing real-world construction site dataset.

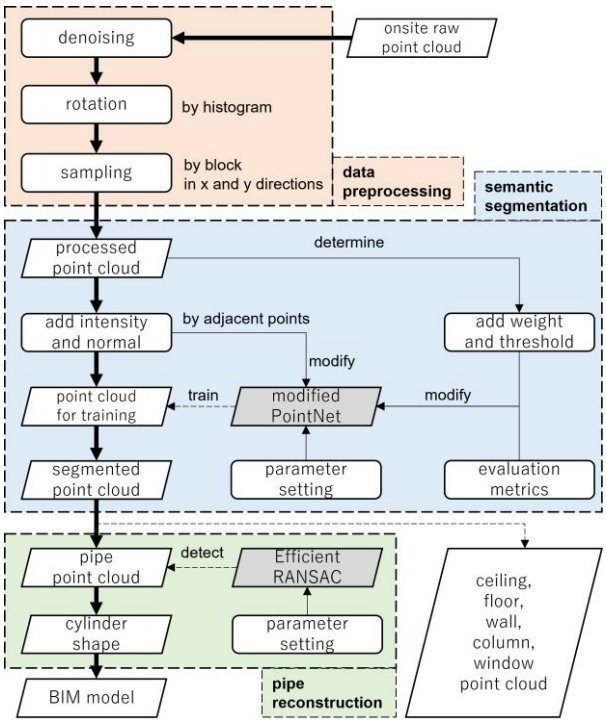

Figure 1. Overview of pipes segmentation and reconstruction process

2 Related Work

2.1 3D Models in Construction Field

There are mainly two kinds of 3D model reconstruction based on the information contained in the model: (1) geometry model reconstruction and (2) semantic model reconstruction [8]. The difference is that the latter includes the geometric information and the object's semantic information. Both two categories depend heavily on the implementation of computer vision algorithms such as Structure from Motion (SFM) [9] and random sample consensus (RANSAC) [10]. The semantic information for point clouds acquired from the construction field refers to building components such as ceilings, walls, and MEP systems. Instance-wise or point-wise object classification has been widely studied to reconstruct the as-built model with the purpose of comparing it with an pre-design 3D model. Given existing BIM models, 3D point clouds can be used for geometry quality inspection and construction process monitoring [11]. Even if not, as-built 3D model reconstruction can still be helpful for facility management, such as renovation [12], maintenance [13], and energy analysis [14]. As the 3D BIM model is beneficial for various applications in the whole life cycle of a building, there still lacks a general framework for segmentation and reconstruction of the as-built building, where onsite 3D point cloud data would be practical for constructing a workflow.

2.2 3D Semantic Segmentation

Semantic segmentation of 3D point clouds is to generate semantic information for every point. It is usually considered supervised learning methods, including traditional machine learning techniques and state-of-the-art deep learning approaches [5]. Supervised machine learning techniques are further divided into two groups: (1) feature based on every single point, such as Support Vector Machine [6], AdaBoost [7] and (2) feature based on contextual information in graphical models, such as Conditional Random Field and Markov Random Field [2]. Compared with traditional machine learning methods, deep learning approaches feed training data directly into a neural network to acquire high-dimensional features without specific knowledge for feature design and extraction. The deep learning approaches can then be divided into three categories

based on the format in which the point cloud is fed into the neural network: (1) Multiview [15], (2) voxel, and (3) raw point cloud. Since Multiview and voxel forms need to convert 3D point cloud data to 2D images or low-resolution forms, directly binding transformation in the neural network and performing on point cloud data have advantages in controlling computational overhead and maintaining original information of feature space. However, these deep learning approaches deal with only single-shape recognition problems or lack local structure information to correctly separate certain target classes from neighbor structural components in close proximity. A refinement for the architecture of the current neural network used in 3D point cloud semantic segmentation on construction sites needs to be further explored.

2.3 Plumbing Pipe Reconstruction

There has been extensive interest in research on the reconstruction of pipes by 3D point cloud data in computer vision and construction management. Ahmed et al. [16] use Hough-transform and domain constraints to develop an effective approach to locating and reconstructing 3D pipes. Dimitrov et al. [17] introduce a new region-growing method with a single parameter, which accounts for the desired level of abstraction for context-free segmentation of building point clouds, including MEP systems. Due to a large number of noises and different sizes of pipelines that cannot be directly used by traditional methods such as region-growing, RANSAC and Hough transformation, Liu et al. [18] proposed a simpler method to find circles in 2D space rather than finding cylinders in 3D space. Since most of the above research targets are pipes in scenes where there are dominantly planar or non-planar surfaces, Perez-Perez et al. [2] presented a Markow Random Field learning method to assign semantic labels to point cloud segment where MEP systems are close to building components. Studies also use neural networks to reduce the time-consuming and labor-intensive process of pipe recognition and segmentation. Kim et al. [19] showed an automatic pipe and elbow recognition method, including CNN, to filter falsely classified points against occlusion between pipes and elbows. Cheng et al. [20] utilized PointCNN [21] to learn the point cloud feature and classify points into 6 different classes and no-part for later clustering and aggregation. Kim et al. [22] tested the performance of MCVNN [15] and PointNet [4] on retrieving piping component types from an already segmented point cloud. In most of the research, the MEP system data has already had a high resolution before classification or shape detection, while the point cloud data from the construction site is usually a partial scan due to the obstructions. There is also little discussion about evaluating reconstruction results using the original ground truth data with reconstruction results by the prediction result.

2.4 Imbalanced Classification

While the convolutional neural network is gaining significance in perception fields, one big challenge is that, some classes in the data set have a remarkably higher portion than other classes; it's also referred to as class imbalance. The class imbalance would affect convergence during training and a model's generalization while testing. [23] Strategies for addressing the class imbalance problem are mainly divided into two categories [24]: (1) data-level sampling methods and (2) classifier-level weight methods. The main idea of oversampling is replicating samples from minority classes, while undersampling refers to removing examples from majority classes. Both methods would address the problem to some extent, but there are also shortcomings: while oversampling may cause an overfitting problem, undersampling would lose specific valuable data. Cost-sensitive learning [25] assigns different weights for wrong predictions from different classes. One way to implement it in a neural network is to pass the weighted loss to the backpropagation process. Another classifier-level weight method is tantamount to adjusting the threshold for the last layer of the neural network. Usually, the output is divided by its prior probability, which equals the ratio of the corresponding instances in the same class. These strategies are mainly proposed for 2D image classification tasks, and the effectiveness of 3D point cloud semantic segmentation on data from construction sites remains to be tested.

3 Datasets

For this study, the datasets were acquired by a LiDAR scanner at the student center construction site at Georgia Institute of Technology. They are manually labeled into 10 classes: ceiling, column, floor, pipe, wall, window, noise, and others. The numbers of points for the first 6 classes (Figure 2) in total 4 scans were shown in Table 1.

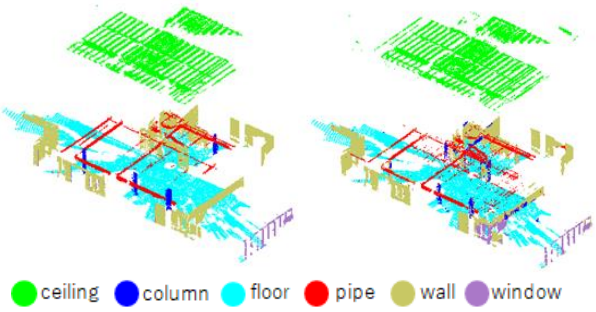

Figure 2. Example of ground truth (left) and segmentation result (right). The ceiling is moved upwards for better visualization for other classes.

4 Data Preprocessing

4.1 Denoising

Raw point cloud data from construction sites contain redundant information that should be removed before preprocessing. Except for six structural components, including ceilings, columns, floors, pipes, walls and windows, other movable onsite components such as workers, tools, materials, and noise are removed to keep the data specialized, clean, and distinct. We use statistical outlier removal to remove noise that deviate from their neighbors compared with the average of the point cloud. There are primarily two benefits to denoise raw point cloud data. The first one is that movable data like workers, tools, and materials occupy only a small percentage of the total data; however, the amount can vary largely across different sites, aggravating the imbalanced-classification problem and influencing the accuracy of other desired structural components. The second reason is that those movable data do not have regular geometric shapes; thus, the normal for small local faces can differ even on the same object. In Table 2, we calculate and add the normal of small faces formed by local points to the backbone network. Removing these irregular components would improve the model's performance by making the cylinder shapes more distinguishable from the plane context, since the normal of cylinders and planes are more consistent on a local surface than movable opponents.

4.2 Point Transformation

In the PointNet framework, an affine transformation matrix is used for the input and feature transformations in the model, and it is claimed that combining both the regularization and transformations can achieve its best performance [4]. In our implementation of raw point cloud data from construction sites, we found that higher accuracy can be achieved by applying a rotation on the point cloud data to align most walls to the x-axis and y-axis before feeding it into the neural network. Raw point cloud data from construction sites have more complicated shapes and contain more outlier points than the indoor environment, so alignment to x-axis and y-axis would help each block have a more reasonable distribution of different categories of points in the afterward sampling process. In actual practice, we divided the angle from -45° to 45° by small intervals. For each small angle, we calculated the histogram of the distribution of the x-axis and y-axis for the point cloud data. Then we operate max pooling on the number of points in each bin by step size 3 and sum up the top 20 results in the x-axis and y-axis. By maximizing the density of bins in this way, we could acquire an angle by which most of the walls are rotated in x and y directions. For the four datasets we used for training and evaluating, the angles are, respectively -33.44°, -42.48°, 3.86° and -15.29°.

4.3 Point Sampling

We apply the same sampling method used in [4] for indoor environments. We divide the area of construction sites in x and y directions with a block size of 2 meters. In every single block, we sample 1024 points randomly through the z direction and use stride size 1 meter in x and y directions to overlap part of the sampling space. The training result determines these numbers that a sparser sampling would not likely provide enough neighbor points for later normal calculation, thus harming the accuracy of prediction, while a denser sampling would significantly increase the number of points as well as the time used for training and computing. Note that in this study, we do not apply any sampling methods such as oversampling or undersampling. Further research involves using other sampling strategies to improve prediction accuracies of minor classes such as window and column.

5 Semantic Segmentation

5.1 Backbone Network

PointNet is a neural network framework that can consume point cloud data directly to perform 3D recognition tasks, such as object classification, part segmentation, and semantic segmentation, efficiently and effectively without transforming the data to 3D voxel grids or 2D images. The detailed design of the network structure uses the max pooling layer, global feature combined with the local feature, and transformation matrices to deal with the three main properties of point cloud data, unordered, interaction among points, and invariance under transformations.

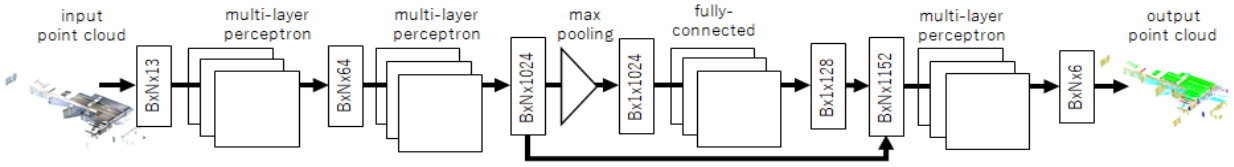

Figure 3. The architecture of the modified PointNet

In our implementation, we use a similar neural network structure to the PointNet paper for indoor semantic segmentation, except for increasing input data dimension from 9 to 13. The architecture of our backbone network is shown in **Error! Reference source not found.**

5.2 Intensity and Normal

In the original PointNet framework, the 9 features of the point cloud used in the neural network are, respectively local x, y, z coordinates, global x, y, z coordinates, and normalized R, G, B values of color. Although local coordinates combined with global coordinates provide certain knowledge for point segmentation, in practice, this information is still insufficient for extracting pipe systems with structural components like ceilings in proximity for semantic segmentation.

Therefore, we consider adding other features such as intensity and normal. Intensity is a feature that we could get directly from raw point cloud data from construction sites. Since it records the return strength of the beam, it has been used in recent research to help recognize the material information of construction sites [26]. However, since additional factors such as scan angle, range, and moisture content can also affect the intensity, it does not always lead to a consistent result. Indeed, in our experiment, we expected intensity could help us recognize columns and windows apart from walls, but the result was not promising. We investigated the specific intensity field of the data and learned that the intensity of columns is relatively stable, but there are also such values in the same range for walls. Considering the close spatial relationship, it makes sense that the intensity of these structural components would be similar, especially for raw point cloud data from construction sites.

On the other hand, normal is beneficial for distinguishing pipes from ceilings. We use a K-d Tree with dimension three to search neighbor points for each point and then calculate the normal vectors of the plane constructed by those points. We set the radius to 1 meter and the maximum number of neighbors to be searched as 30. Since pipes have cylindrical shapes and ceilings have plane shapes, the direction and the distribution of normal vectors of pipes and ceilings should be distinct even in a small close proximity. The result shows that by adding normal and combining with other proposed methods, we can largely improve the accuracy of pipes without affecting the accuracy of other major classes such as ceiling, wall, and floor.

5.3 Weight and Threshold

The class imbalance has been a big problem in convolutional neural networks. For point cloud data from construction sites, this problem is more severe than indoor environments. The ratios of each class to the total number of points of our data are respectively 35.84% for ceiling, 3.13% for column, 26.93% for floor, 9.30% for pipe, 23.28% for wall and 1.52% for window. Major classes like ceiling, floor and wall make up over 86% of the whole dataset, while the smallest minority class only contains 1.52% of the total points, which causes the ratio of imbalance reaches to over 23 [23]. What makes the problem worse is that those categories occupying a larger percentage of the total data have relatively simple shapes compared with the remaining categories. This would make the model redundant to change once it finds a straightforward relationship between certain features and corresponding labels, say the z-axis and floor, and reach high accuracy. To mitigate the negative influence of class imbalance, we try to apply weight while calculating loss and threshold while making a prediction. Before adding the loss together, we assign a different weight w_i for wrong predictions from different classes by multiplying $(1/r_i)/\sum(1/r_i)$ for each class, where r_i is the prior probability of each class, equalling the ratio of each class, calculated by the total number of points n_i for each class (**Error! Reference source not found.**). Note here that we roughly use each class's ratio in the whole dataset instead of each training batch; we defer this exploration to future work. The other strategy we try is to adjust the decision threshold of the output layer in the test phase by multiplying the equivalent weight we use for calculating loss. We found that weight for training combined with a threshold for testing would improve the accuracy for minor classes like pipe, column, and window, although it might make accuracy for other major classes lower. Since recall might not fully reflect the correctness of prediction, we further investigate the IoU of each class before and after applying weight and threshold. The result is shown in **Error! Reference source not found.**. For pipes, from the 4th configuration to the 5th configuration, the accuracy increases while IoU decrease means that both true positive points and false positive points increase. This result is favorable for later shape fitting since little more false positive outlier points for pipes should not affect cylinder fitting too much because RANSAC algorithm uses normals to detect shape.

Table 1. Number of points (million) and weight (threshold) for six classes.

	ceiling	column	floor	pipe	wall	window
1	2.33	0.07	2.74	0.82	2.20	0.04
2	2.63	0.45	2.17	0.53	2.71	0.19
3	3.39	0.07	1.30	0.51	1.65	0.17
4	3.02	0.40	2.34	1.09	0.83	0.09
n_i	11.37	0.99	8.54	2.95	7.38	0.48
r_i	0.36	0.03	0.27	0.09	0.23	0.02
w_i	0.02	0.27	0.03	0.09	0.04	0.55

5.4 Evaluation Metrics

Since overall accuracy favors over-represented majority classes, leading to a highly misleading assessment of both training and evaluating [23], we use the area under the receiver operating characteristic curve, ROC AUC [27], to determine when to stop while training. ROC AUC is essentially a plot of the false positive rate to the true positive rate for different prediction thresholds. While the area under the curve depicted by plots of false positive rate and true positive rate pairs gets larger and near to 1, the model's ability to distinguish two different classes becomes better. Usually, ROC AUC is used for binary classification, but it can also be extended to multiple classes. Here we utilize the One-vs-rest algorithm to compute the AUC of each class against the rest classes [28] to let the model understand the whole scene and achieve better performance than binary classification. We save the model with the highest ROC AUC score for testing during training and use this model for further evaluation.

5.5 Results and Analysis

We set the number of sampling points in each block as 1024, block size as 2 meters, stride size as 1 meter, batch size as 24, learning rate as 0.001 and stop training in 20 epochs. We use cross-validation across all four datasets, training on three datasets while testing on the remaining dataset four times before averaging the results, and our modified backbone network (5th configuration in **Error! Reference source not found.**) results in 89.83% average overall training accuracy, 99.08% training ROC AUC score with 84.54% average overall validation accuracy, 91.10% validation ROC AUC score on the semantic segmentation task. **Error! Reference source not found.** shows an example of a prediction result.

From Table 2, we can safely conclude that we have found a better way to solve issues caused by partially imbalanced noisy data sets. Firstly, comparing the result for the first 5 configurations, by utilizing rotation and adding intensity, normal, and weight, the average accuracy and IoU for pipes would increase from 70.32% and 50.88% to 86.26% and 66.42%, while the average accuracy and IoU for all 6 classes would increase from 61.75% and 51.03% to 77.92% and 61.78%. This means data preprocessing, adding local features, and dealing with class imbalance problems are useful for recognizing plumbing pipes from other structural components. Secondly, comparing the result for last 3 configurations, we found that minority classes like column and window would benefit from weight for loss calculation and threshold for the testing phase. For pipes, the accuracy would increase while the IoU would decrease a little, which means some points of the ceiling are recognized as pipes. However, this fact should not affect the cylinder fitting process too much because RANSAC uses normal to detect the shape, and the normal for cylinder and plane are different. Finally, we discovered that using weight combined with the threshold proposed in the last configuration is not as effective as using them separately. This might be because we have already compensated weight to the loss function through backpropagation in the training process. The repeating threshold in the test phase would conversely hurt the prediction results.

6 Pipe Reconstruction

6.1 Cylinder Shape Detection

RANSAC is a non-deterministic algorithm to estimate the parameters of a mathematical model among noisy data. [10] It has been extensively used in reverse engineering, like point cloud reconstruction, with its favorable properties such as generality, simplicity, and robustness. Efficient RANSAC is a variation of the original RANSAC algorithm that uses a lazy cost evaluation function with a structured sampling strategy for candidate shapes like planes, spheres, cylinders, and other additional primitives. [29] This method is suitable and efficient for automatically acquiring raw point cloud data under relatively adverse conditions, such as heavy noise in constrained processing time.

Table 2. Result of semantic segmentation.

configuration	ceiling	column	floor	**pipe**	wall	window	average
	acc/IoU	acc/IoU	acc/IoU	acc/IoU	acc/IoU	acc/IoU	acc/IoU
Base PointNet [4]	77.9/70.2	31.3/22.0	98.3/95.6	70.3/50.9	78.7/60.5	15.2/7.0	62.0/51.3
Ours (w/ rotation)	85.3/78.0	33.7/23.9	96.7/93.6	68.9/51.6	77.5/61.2	32.9/18.5	65.8/54.5
Ours (w/ rotation & intensity)	90.0/78.8	24.4/17.4	97.9/95.0	69.2/55.6	**80.2**/64.9	43.0/25.3	67.5/56.2
Ours (w/ rotation & intensity & normal)	**94.6**/**85.2**	23.7/12.1	98.7/97.6	80.3/**68.3**	79.8/65.0	54.7/25.1	72.0/58.9
Ours (w/o threshold)	86.0/80.4	**53.3**/**26.8**	98.1/97.1	86.3/66.4	76.3/**66.4**	67.5/33.5	**77.9**/**61.8**
Ours (w/o weight)	88.0/81.5	33.9/17.6	**99.2**/**98.1**	83.6/64.8	72.2/64.0	74.7/**34.8**	75.3/60.1
Ours (w/ full condition)	82.3/78.3	48.9/15.4	95.3/94.5	**89.3**/61.3	45.5/38.4	**81.0**/22.5	73.7/51.7

For the shape to estimate, a cylinder is generated from two points with normals and then verified by the two free parameters: max distance to primitive ε and max normal deviation α.

6.2 Experiment Results

Since there is no common rule for setting distance threshold ε and normal deviation threshold α, we found their optimal values as well as other parameters from the experiments. We set the distance threshold ε as 0.15 and 10 for normal deviation threshold α. For the minimum support points for each primitive, we set the value as 200. We use a uniform sampling method like we used in the training process. We also use cross-validation across all four datasets and make predictions for each area. **Error! Reference source not found.** shows the detection result on the ground truth point cloud and prediction result. The error rate for cylinders' center coordinates, radius r, and length h are listed in **Error! Reference source not found.**.

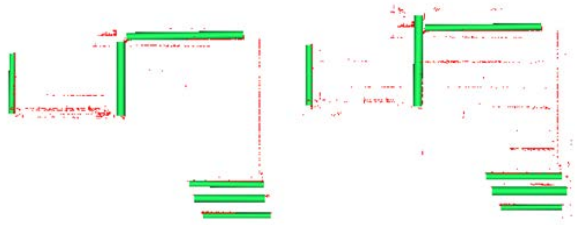

Figure 4. Example of reconstruction result of pipes by ground truth (left) and prediction result (right).

The outcome of cylinder shape detection on the prediction result is very satisfactory. The error rates for x, y, z coordinates of cylinders' centers are mostly under 1%, while the error rates for the radius r and length h of detected cylinders are under 10% and 5%, respectively. Although visibly, some leftover points could not be detected as cylinders under the current algorithm and parameter settings, either due to the unified distance and minimum support points we chose or the fact that some point clouds from other classes are mislabelled as pipes, these results do indicate the potential for directly using semantic segmentation prediction result from our framework to reconstruct plumbing pipes if the shape detection method itself reaches to a sufficient accuracy.

Table 3. Error rate for RANSAC estimation of cylinder parameters.

		x	y	z	r	h
error rate	mean (%)	0.40	0.57	0.54	9.60	4.70
	stdev	0.00	0.01	0.01	0.06	0.05

7 Conclusion

The proposed framework shows a novel approach that employs class-imbalanced point cloud data from construction sites to segement and reconstruct piping systems. The proposed techniques and enhanced backbone network exhibit robust performance in semantic segmentation of partially scanned, highly imbalanced, and noisy point cloud data. Moreover, the comparison with the ground truth demonstrates the feasibility of shape detection.

Despite the experiment results indicating robust segmentation performance, further improvements will be investigated by leveraing alternateve sampling methods and computing weight and threshold values per each batch training step. Also, ablation studies such as parameter selection will be considered to optimize the piping reconstruction performance across varying plumbing pipe demensions.

Acknowledgments

This research was supported by the National Science Foundation under Grant no. (2222723) and the Ministry of Science and ICT through the National Research Foundation of Korea (NRF-2022H1D3A2A01093584).

References

[1] F. Bosché, A. Guillemet, Y. Turkan, C. Haas and R. Haas. Tracking the built status of MEP works: Assessing the value of a Scan-vs-BIM system. *Journal of Computing in Civil Engineering,* 28(4), pages 1-28, 2014.

[2] Y. Perez-Perez, M. Golparvar-Fard and K. El-Rayes. Segmentation of point clouds via joint semantic and geometric features for 3D modeling of the built environment. *Automation in Construction,* 125, 2021.

[3] H. Son, C. Kim and Changwan Ki. Fully Automated As-Built 3D Pipeline Extraction Method from Laser-Scanned Data Based on Curvature Computation. *Journal of Computing in Civil Engineering,* 29(4), 2015.

[4] 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,* pages 652-660, 2017.

[5] Y. Xie, J. Tian and X. Zhu. Linking Points With Labels in 3D: A Review of Point Cloud Semantic Segmentation. *IEEE Geoscience and Remote Sensing Magazine,* 8, pages 38-59, 2020.

[6] Q. Wang and M.-k. Kim. Applications of 3D point cloud data in the construction industry: A fifiteen-

year reveiwo from 2004 to 2018. *Advanced Engineering Informatics,* 39, pages 306-319, 2019.

[7] M. Westoby, J. Brasington, N. Glasser, M. Hambrey and J. Reynolds. 'Structure-from-Motion' photogrammetry: A low-cost, effective tool for geoscience applications. *Geomorphology,* 179, pages 300-314, 2012.

[8] M. A. Fischler and R. C. Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM,* 4(6), pages 381–395, 1981.

[9] F. Bosché, M. Ahmed, Y. Turkan, C. T. Haas and R. Haas. The value of integrating Scan-to-BIM and Scan-vs-BIM techniques for construction monitoring using laser scanning and BIM: The case of cylindrical MEP components. *Automation in Construction,* 49, pages 201-213, 2015.

[10] C. C. Aydin. Designing building façades for the urban rebuilt environment with integration of digital close-range photogrammetry and geographical information systems. *Automation in Construction,* 43, pages 38-48, 2014.

[11] E. Valero, F. Bosché and A. Forster. Automatic segmentation of 3D point clouds of rubble masonry walls, and its application to building surveying, repair and maintenance. *Automation in Construction,* 96, pages 29-39, 2018.

[12] C. Wang, Y. K. Cho and C. Kim. Automatic BIM component extraction from point clouds of existing buildings for sustainability applications. *Automation in Construction,* 56, pages 1-13, 2015.

[13] J. Zhang, X. Lin and X. Ning. SVM-Based Classification of Segmented Airborne LiDAR Point Clouds in Urban Areas. *Remote Sensing,* 5(8), pages 3749-3775, 2013.

[14] Z. Wang, L. Zhang, T. Fang, P. T. Mathiopoulos, X. Tong, H. Qu, Z. Xiao, F. Li and D. Chen. A Multiscale and Hierarchical Feature Extraction Method for Terrestrial Laser Scanning Point Cloud Classification. *IEEE Transactions on Geoscience and Remote Sensing,* 53(5), pages 2409-2425, 2015.

[15] H. Su, S. Maji, E. Kalogerakis and E. G. Learned-Miller. Multiview convolutional neural networks for 3d shape recognition. In *Proceedings of the IEEE international conference on computer vision,* pages 945-953, 2015.

[16] M. Ahmed, C. T. Haas and R. Haas. Autonomous Modeling of Pipes within Point Clouds. *In 2013 Proceedings of the 30th ISARC,* pages 1093-1100, 2013.

[17] A. Dimitrov and M. Golparvar-Fard. Segmentation of building point cloud models including detailed architectural/structural features and MEP systems. *Automation in Construction,* 51, pages 32-45, 2015.

[18] Y.-J. Liu, J.-B. Zhang, J.-C. Hou, J.-C. Ren and W.-Q. Tang. Cylinder Detection in Large-Scale Point Cloud of Pipeline Plant. *IEEE TRANSACTIONS ON VISUALIZATION AND COMPUTER GRAPHICS,* 19, pages 1700-1707, 2013.

[19] Y. Kim, C. H. P. Nguyen and Y. Choi. Automatic pipe and elbow recognition from three-dimensional point cloud model of industrial plant piping system using convolutional neural network-based primitive classification. *Automation in Construction,* 116, 2020.

[20] L. Cheng, Z. Wei, M. Sun, S. Xin, A. Sharf, Y. Li, B. Chen and C. Tu. DeepPipes: Learning 3D pipelines reconstruction from point clouds. *Graphical Models,* 111, 2020.

[21] Y. Li, R. Bu, M. Sun and B. Chen. PointCNN: Convolution On X-Transformed Points. *Proceedings of the 32nd International Conference on Neural Information Processing Systems,* pages 828–838, 2018.

[22] H. Kim, C. Yeo, I. D. Lee and D. Mun. Deep-learning-based retrieval of piping component catalogs for plant 3D CAD model reconstruction. *Computers in Industry,* 123, 2020.

[23] M. Buda, A. Maki and M. A. Mazurowski. A systematic study of the class imbalance problem in convolutional neural networks. *Neural Networks,* 106, pages 249-259, 2018.

[24] H. He and E. A. Garcia. Learning from Imbalanced Data. *IEEE Transactions on Knowledge and Data Engineering,* 21, pages 1263-1284, 2009.

[25] C. Elkan. The foundations of cost-sensitive learning. *International joint conference on artificial intelligence,* 17(1), pages 973-978, 2001.

[26] J. Park and Y. K. Cho. Point Cloud Information Modeling: Deep Learning–Based Automated Information Modeling Framework for Point Cloud Data. *Journal of Construction Engineering and Management,* 148, 2022.

[27] A. P. Bradley. The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognition,* 30, pages 1145-1159, 1997.

[28] F. Provost and P. Domingos. Well-trained PETs: Improving probability estimation trees. *CeDER Working Paper #IS-00-04, Stern School of Business, New York University,* 2000.

[29] R. Schnabel, R . Wahl and R. Klein. Efficient RANSAC for Point-Cloud Shape Detection. *Computer Graphics Forum,* 26(2), pages 214-226, 2007.