UNB | Off-site Construction Research Centre
EST. 1785 UNIVERSITY OF NEW BRUNSWICK

CSCE CCGC 1887 Canadian Society for Civil Engineering

Transforming Construction with Reality Capture Technologies: The Digital Reality of Tomorrow

August 23-25, 2022, Fredericton, New Brunswick, Canada

PERFORMANCE ASSESSMENT OF DEEP NEURAL NETWORKS FOR CLASSIFICATION OF IFC OBJECTS FROM POINT CLOUD WITH LIMITED LABELED DATA

Seydgar, Majid[1,*], Motamedi, Ali[1], and Poirier, Érik A[1]
[1] Department of Construction Engineering, École de Technologie Supérieure, Université du Québec, Canada
[*] majid.seydgar.1@ens.etsmtl.ca

Abstract: Point cloud (PC) processing using deep neural networks (DNNs) has been attracting an increasing amount of attention due to its high performance in tasks, such as Industry Foundation Classes (IFC)-based object classification. DNNs typically rely on large labeled-sample sizes for training. On the other hand, 3D IFC object annotation is a time-consuming task, and the existing datasets are only small-sized labeled samples. In this study, we perform a set of experiments to assess the training progress and generalization capacity of two state-of-the-art DNNs for IFC object classification. The results show that limited training samples can lead to inefficient learning by DNNs, even if a large portion of the annotated samples are allocated to the training task. The conducted experiments indicate that overfitting on the training set as well as low speed convergence during the training are some of the main issues for classification of 3D IFC objects using DNNs. To tackle these issues, future studies could take advantage of unsupervised and semi-supervised learning methods to use a greater volume of unlabeled samples for training and decrease the reliance of DNNs on annotated data.

Keywords: Deep Neural Networks (DNN); Industry Foundation Classes (IFC); Building Information Modeling (BIM); Object classification; Scan-to-BIM

1 INTRODUCTION

As research into the development and application of building information modeling (BIM) in the field of Architecture, Engineering, Construction and Facilities Management (AEC/FM) continues to grow, the numerous advantages BIM provides as a centralized information exchange platform is becoming increasingly prevalent (Motamedi, Hammad, and Asen 2014; Shahinmoghadam, Natephra, and Motamedi 2021). Despite the great benefits of BIM, one of the challenges it faces is its reliance on manual modeling for the creation of BIModels. Various approaches have investigated the application of artificial intelligence (AI) in the creation of BIM using generative design approaches (Karan and Asadi 2019; Karan, Safa, and Suh 2020) and in performing efficient operation and maintenance (Cheng et al. 2020; Kwon et al. 2021). In this study, the focus is on the reconstruction of a BIM using scanned data, an approach called scan-to-BIM. The scanned data enables the representation of the “as-is” conditions of the building. The creation of BIM from scanned data has been the subject of many research projects over the past two decades. The primary methods mainly relied on labor-intensive and arduous manual modeling (Tang et al. 2010). However, with the introduction of deep neural networks (DNNs) and machine learning (ML)-based methods, researchers have been trying to enhance the Scan-to-BIM pipeline. In particular, IFC-based object detection using scanned data is one of Scan-to-BIM’s sub-tasks that has been addressed by several recent DNN studies

(Koo, Jung, and Yu 2021; Emunds et al. 2021; 2022). In general, object detection from scanned data is a challenging task due to various noise effects (Charles R Qi et al. 2017) and occultations (Kim and Kim 2021). In recent years, DNNs have performed efficiently in the detection (Charles R Qi et al. 2017; Charles Ruizhongtai Qi et al. 2017) and classification of various objects from scan data using well-known datasets (Dai et al. 2017; Avetisyan et al. 2019; Chang et al. 2017). However, IFC-based objects in indoor environments (e.g., stair, pipe, railing) are under-represented in most of the existing datasets. To address this issue, the IFC object detection methods typically rely on manual data annotation and then use the annotated data for the DNNs training. Since 3D data annotation is a time-consuming task, the annotated samples are usually limited, which can lead to inefficient DNNs training for IFC object detection. In this study, we assess the learning progress of the state-of-the-art DNNs for 3D IFC object detection in two scenarios: (a) training with a larger set of training samples; and (b) training with limited training samples using the publicly available IFCNet dataset (Emunds et al. 2021). These experiments will enable us to evaluate the effects of limited training samples in the training and inference of DNNs. Moreover, this study discusses the possible research avenues to tackle the performance degradation of DNNs, caused by the lack of labeled data. The paper starts by describing related studies that tackled the problem of limited training data for DNN training. The methodology is then described, followed by the experimental configuration and applied methods for the comparison task. Finally, the results are presented and the effects of limited training samples based on the results are discussed.

2 RELATED WORK

Many studies have introduced different DNN-based methods to overcome the limitations of traditional methods in the tasks of interpreting 3D data, such as object recognition and classification (Feng et al. 2019; Wang et al. 2019). Typically, DNNs use hierarchical feature-extractor architectures to robustly extract semantic and abstract features from the PCs (Charles Ruizhongtai Qi et al. 2017; Li et al. 2018). The robust feature-learning pipeline of DNNs provides remarkable results in IFC object detection (Emunds et al. 2022; Koo, Jung, and Yu 2021; Emunds et al. 2021). However, DNNs require a large set of labeled samples to update and tune the weights of their layers (i.e., parameters). This is because DNNs utilize a significant number of parameters that are explicitly optimized via training samples. This said, datasets with parametric objects and attributes (e.g., IFC-based objects) are very limited since object annotation from scanned data is time-consuming and arduous. Recently, a limited number of IFC-based datasets, such as IFCNet (Emunds et al. 2021) and BIMGEOM (Collins et al. 2021) have become publicly available to promote Scan-to-BIM using DNNs. However, the size of these datasets is usually insufficient to robustly train the DNNs. In the literature on DNN training, many studies have addressed the problem of performance degradation of DNNs under the conditions of limited training samples (Zhao, Chua, and Lee 2020; Meng et al. 2021). Unsupervised learning (Ma and Leite 2022), weakly supervised learning (Beyer and Dai 2022), and semi-supervised learning (Zhao, Chua, and Lee 2020) are among the well-known approaches to alleviate the problem of limited training data. However, these methods are rarely investigated as a potential solution to overcoming the issues related to 3D IFC object detection from PC.

3 ASSESSMENT METHOD

This section briefly describes the employed dataset and DNNs for the 3D IFC object classification. Also, the assessment criteria and experimental configurations to perform the comparison are explained.

3.1 DATASET

To perform 3D object classification for IFC objects, the publicly available IFCNet dataset (Emunds et al. 2021) was used. IFCNet contains roughly 8,000 samples in twenty classes, such as, air terminal, beam, and duct. The dataset was divided into two subsets, namely train and test. However, the samples in two subsets were combined to perform random sample selection for the training scenarios in this study.

3.2 EXPERIMENTAL DNNs AND ASSESSMENT CRITERIA

To perform 3D IFC object detection, two well-known DNNs, namely DGCNN (Wang et al. 2019) and MeshNet (Feng et al. 2019) were used. DGCNN is a dynamic graph convolution-based network that has demonstrated high performance in the task of 3D IFC object classification (Emunds et al. 2021; 2022). This method uses an edge-preserving strategy, which enables it to capture the geometrical information of the PC more effectively. DGCNN also shows superior performance compared to other state-of-the-art rivals such as PointNet ++ (Charles Ruizhongtai Qi et al. 2017) on the ShapeNet dataset (Wu et al. 2015). Therefore, DGCNN was employed in this study under the condition of limited labeled data. The general pipeline of DGCNN utilization for the 3D IFC object classification is illustrated in Figure 1, where the Blender application ("Blender Foundation" 2002) was used to perform uniform point sampling.

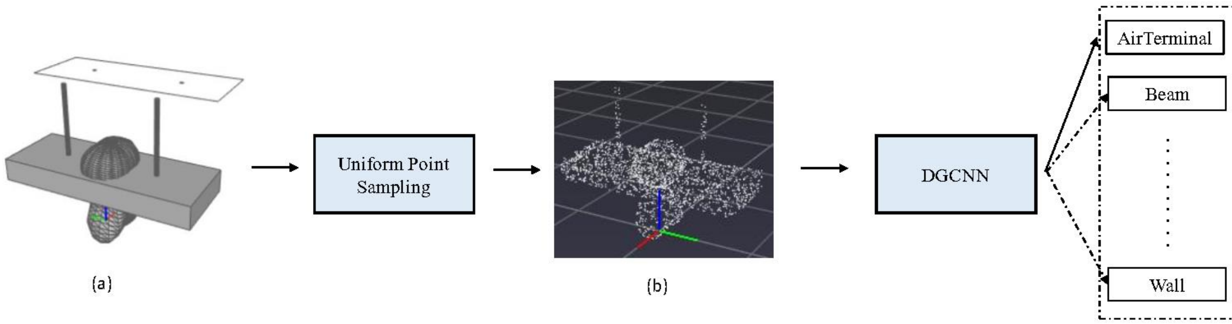


Figure 1: Pipeline of 3D IFC object classification with DGCNN; a) IFC object b) generated points using uniform sampling.

The second DNN-based method, MeshNet, takes advantage of the geometric properties of the mesh data (Feng et al. 2019) first by converting the raw PC to a set of convex polygons (i.e., faces) along with their adjacency information (i.e., neighbors), then by using the extracted information to perform 3D object classification. MeshNet demonstrates comparable performance for 3D IFC object classification, but its performance under limited labeled data has not yet been investigated for 3D IFC object classification. Figure 2 shows the pipeline of MeshNet for the task of mesh generation and IFC object classification. Specifically, MeshLab (Cignoni et al. 2008) was used to convert the raw point sets to mesh data.

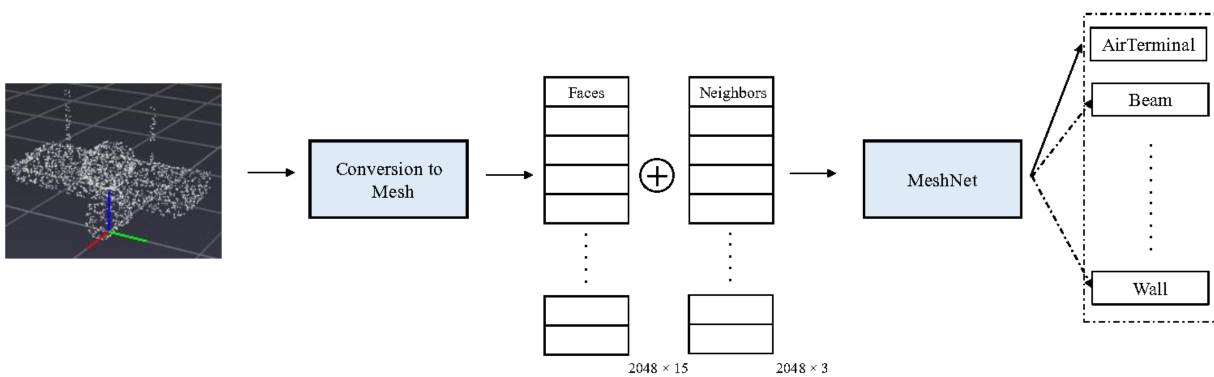


Figure 2: Pipeline of the 3D IFC object classification using MeshNet.

To perform the training of the chosen methods, the methods' hyperparameters were set using the Tune platform (Liaw et al. 2018), and the methods were trained for β epochs on the training set. To assess the performance of the experimental methods, five assessment criteria, namely accuracy (AS), F_1, precision (PS), recall (RS), and balanced accuracy (BAS) were used. In particular, AS was calculated to determine the percentage of predictions that matched with the test set and F_1 to reflect the harmonic mean accuracy score of the PS and RS. F_1 was defined using Eq. 1

$$F_1 = 2 \times \frac{PS \times RS}{PS + RS} , \quad \text{(Eq. 1)}$$

where

$$PS = \frac{S_P^T}{S_P^T + S_P^F}\ ; RS = \frac{S_P^T}{S_N^T + S_N^F}\ . \quad \text{(Eq. 2)}$$

S_P^T, S_P^F, and S_N^F are the number of true positives, the number of false positives, and the number of false negatives predictions, respectively. Lastly, BAS, which was used as an accuracy metric for an imbalanced dataset, was defined in Eq. 3 as

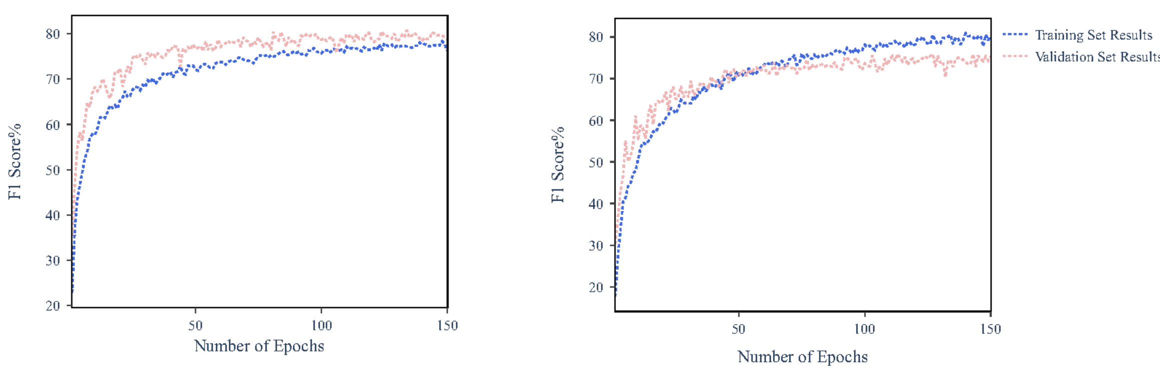
$$BAS = \frac{S_P^T}{S_P^T + S_N^F} + \frac{S_N^T}{S_N^T + S_P^F}\ , \quad \text{(Eq. 3)}$$

where S_N^T is the number of true negatives in the confusion matrix between the predicted labels and true labels for the IFC objects. These assessment criteria were used to evaluate the performance of the competing methods in the experiments.

4 EXPERIMENTAL RESULTS

In this section, the results of the two sets of experiments are discussed. Scenario (a) and (b) were performed to assess the sensitivity of the methods to the number of available training samples. For this, we randomly selected 70% of the samples for the training and reserved the rest for the test assessment in scenario (a). For scenario (b), 30% of the samples were randomly selected for the training and the rest were reserved for the test assessment. The β was set to 150 to train and compare the methods equally and fairly. It should be noted that the training task was performed using two NVIDIA RTX™ 3090 Ti graphic processor units (GPUs) and the training time for DGCNN and MeshNet were 235 and 45 minutes, respectively.

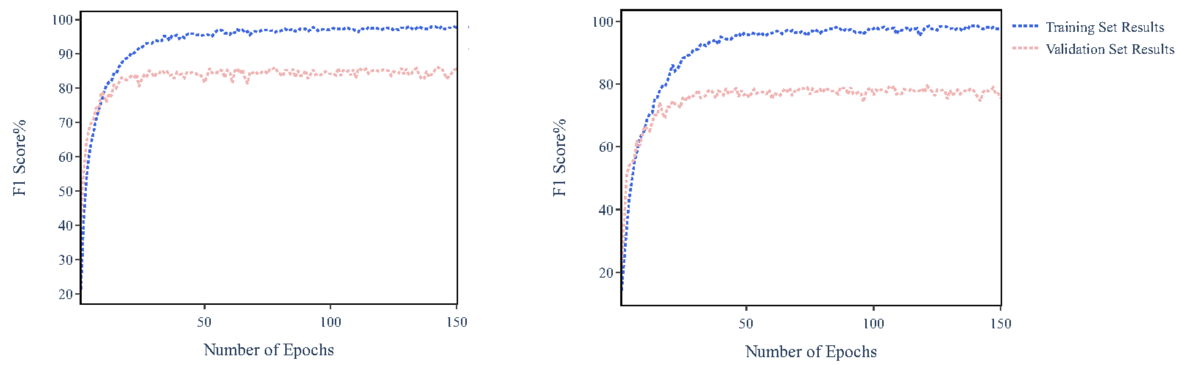
In the first set of experiments, our aim was to assess the effects of limited training samples in the speed and quality convergence of the methods to the optimum learning spot. To this end, we used 30% of the training samples in each scenario as a validation set and obtained the F_1 accuracy score after each epoch training. This enabled us to assess not only the generalization performance of the competing methods on an unseen sample set, but also to compare the learning curves of the training and validation set. The comparison between the learning progress of the two sets of samples, where one was seen by the methods (i.e., training set) and the other, indirectly seen (i.e., validation set) enabled us to detect the over-fitting issue during the learning procedure (Srivastava et al. 2014). Figure 3 illustrates the results of the learning progress and generalization capacity for the DGCNN in both the training and validation sets.



(a) (b)

Figure 3: Training progress analysis of the DGCNN; a) using large training set b) using small training set.

According to the results presented in Figure 3, the DGCNN is unable to converge to the optimum learning spot in both scenarios, which shows its lack of capacity in extracting abstract and semantically rich features from the point set of the IFC objects regardless of the size of the training set. Specifically, the F_1 scores of the training sets in both scenarios did not exceed 80%. Moreover, shrinking the size of the training set leads to an overfitting issue and decreases the generalization performance in scenario (b). The training progress in scenario (a) shows more consistent convergence of both training and validation sets. Nevertheless, the learning curves show limited progress after epoch 50, which could be due to the gradient saturation and optimization inefficiency of the DGCNN for the task of IFC object classification.

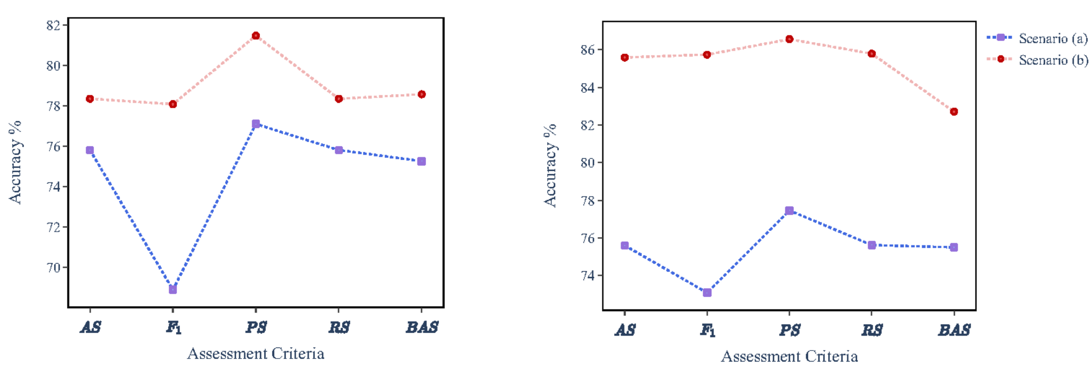


(a) (b)

Figure 4: Training progress analysis of the MeshNet in scenarios (a) and (b).

The same assessment was then performed for MeshNet, and this is illustrated in Figure 4. According to the results, the overfitting problem is inevitable for the MeshNet in both scenarios, which can show the extreme reliance of the method on the training data and the lack of capacity to perform regularization to decrease the gap between the performance on the training and validation sets. However, by comparing the F_1 score on the validation sets in scenarios (a) and (b), we can infer that the overfitting problem intensified after the size of the training samples was decreased. Overall, we can observe the high sensitivity of MeshNet to the training samples in the learning progress as well as to the high degree of uncontrolled overfitting issue.

After the evaluation of the applied methods during the training stage, the performance on the test set using the assessment criteria, described in the methodology section, was assessed in two scenarios. Figure 5 shows the assessment results of the applied methods on the test set. According to these results, decreasing the size of the training set leads to a greater performance drop (up to 12%) in the MeshNet method. This validates the effect of overfitting, observed during the training progress. DGCNN also shows performance degradation, specifically in the F_1 score, but there was only a 2% performance drop, which could indicate that DGCNN is more effective in training progress with limited label samples compared to MeshNet.



(a) (b)

Figure 5: Assessment of the experimental methods' performance on the test set in two scenarios; a) DGCNN b) MeshNet

5 CONCLUSIONS

In this study, we assessed the performance of two state-of-the-art DNNs for the task of 3D IFC object classification from point cloud, based on training and inference tasks in two scenarios: (a) large training and small test-sample sets; and (b) small training and large test-sample sets. Based on our observations, the performance of DNN-based methods for IFC 3D object classification significantly depends on the annotated data. Moreover, during the training, issues, such as overfitting and slow training convergence, were observed in the experiments for both scenarios.

The findings of this study highlight the importance of the regularization schemes in the DNNs configurations to improve the robustness of the generalization of methods and to decrease the gap between the training and validation results. Based on the observations of the result of the learning progress of DGCNN, more effective schemes to analyze the raw point cloud of IFC objects and to extract semantic features need to be investigated to enhance the learning progress during the training task. Moreover, since the existing dataset for the task of IFC object recognition is extremely limited, future studies can take advantage of unsupervised and semi-supervised methods for 3D object recognition and classification to alleviate the problem of performance degradation in DNNs with limited labeled data.

5.1 References

Avetisyan, Armen, Manuel Dahnert, Angela Dai, Manolis Savva, Angel X. Chang, and Matthias Nießner. 2019. “Scan2cad: Learning Cad Model Alignment in Rgb-d Scans.” In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2614–23.

Beyer, Tim, and Angela Dai. 2022. “Weakly-Supervised End-to-End CAD Retrieval to Scan Objects.” *ArXiv Preprint ArXiv:2203.12873*.

“Blender Foundation.” 2002. 2002. https://www.blender.org/get-involved/developers/.

Chang, Angel, Angela Dai, Thomas Funkhouser, Maciej Halber, Matthias Niessner, Manolis Savva, Shuran Song, Andy Zeng, and Yinda Zhang. 2017. “Matterport3d: Learning from Rgb-d Data in Indoor Environments.” *ArXiv Preprint ArXiv:1709.06158*.

Cheng, Jack CP, Weiwei Chen, Keyu Chen, and Qian Wang. 2020. “Data-Driven Predictive Maintenance Planning Framework for MEP Components Based on BIM and IoT Using Machine Learning Algorithms.” *Automation in Construction* 112: 103087.

Cignoni, Paolo, Marco Callieri, Massimiliano Corsini, Matteo Dellepiane, Fabio Ganovelli, Guido Ranzuglia, and others. 2008. “Meshlab: An Open-Source Mesh Processing Tool.” In *Eurographics Italian Chapter Conference*, 2008:129–36. Salerno, Italy.

Collins, Fiona C., Alexander Braun, Martin Ringsquandl, Daniel M. Hall, and André Borrmann. 2021. “Assessing IFC Classes with Means of Geometric Deep Learning on Different Graph Encodings.” In *Proc. of the 2021 European Conference on Computing in Construction*.

Dai, Angela, Angel X Chang, Manolis Savva, Maciej Halber, Thomas Funkhouser, and Matthias Nießner. 2017. “Scannet: Richly-Annotated 3d Reconstructions of Indoor Scenes.” In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 5828–39.

Emunds, Christoph, Nicolas Pauen, Veronika Richter, Jérôme Frisch, and Christoph van Treeck. 2021. “IFCNet: A Benchmark Dataset for IFC Entity Classification.” In *EG-ICE 2021 Workshop on Intelligent Computing in Engineering*, 166. Universitätsverlag der TU Berlin.

Emunds, Christoph, Nicolas Pauen, Veronika Richter, Jérôme Frisch, and Christoph van Treeck. 2022. “SpaRSE-BIM: Classification of IFC-Based Geometry via Sparse Convolutional Neural Networks.” *Advanced Engineering Informatics* 53: 101641.

Feng, Yutong, Yifan Feng, Haoxuan You, Xibin Zhao, and Yue Gao. 2019. “Meshnet: Mesh Neural Network for 3d Shape Representation.” In *Proceedings of the AAAI Conference on Artificial Intelligence*, 33:8279–86.

Karan, Ebrahim, and Somayeh Asadi. 2019. “Intelligent Designer: A Computational Approach to Automating Design of Windows in Buildings.” *Automation in Construction* 102: 160–69.

Karan, Ebrahim, Mahdi Safa, and Min Jae Suh. 2020. “Use of Artificial Intelligence in a Regulated Design Environment–A Beam Design Example.” In *International Conference on Computing in Civil and Building Engineering*, 16–25. Springer.

Kim, Hyunsoo, and Changwan Kim. 2021. “3D As-Built Modeling from Incomplete Point Clouds Using Connectivity Relations.” *Automation in Construction* 130: 103855.

Koo, Bonsang, Raekyu Jung, and Youngsu Yu. 2021. “Automatic Classification of Wall and Door BIM Element Subtypes Using 3D Geometric Deep Neural Networks.” *Advanced Engineering Informatics* 47: 101200.

Kwon, Nahyun, Yonghan Ahn, Bo-Sik Son, and Hyosoo Moon. 2021. “Developing a Machine Learning-Based Building Repair Time Estimation Model Considering Weight Assigning Methods.” *Journal of Building Engineering* 43: 102627.

Li, Yangyan, Rui Bu, Mingchao Sun, Wei Wu, Xinhan Di, and Baoquan Chen. 2018. “Pointcnn: Convolution on x-Transformed Points.” *Advances in Neural Information Processing Systems* 31.

Liaw, Richard, Eric Liang, Robert Nishihara, Philipp Moritz, Joseph E. Gonzalez, and Ion Stoica. 2018. “Tune: A Research Platform for Distributed Model Selection and Training.” *ArXiv Preprint ArXiv:1807.05118*.

Ma, Jong Won, and Fernanda Leite. 2022. “Performance Boosting of Conventional Deep Learning-Based Semantic Segmentation Leveraging Unsupervised Clustering.” *Automation in Construction* 136: 104167.

Meng, Qinghao, Wenguan Wang, Tianfei Zhou, Jianbing Shen, Yunde Jia, and Luc Van Gool. 2021. “Towards A Weakly Supervised Framework for 3D Point Cloud Object Detection and Annotation.” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1–1. https://doi.org/10.1109/TPAMI.2021.3063611.

Motamedi, Ali, Amin Hammad, and Yoosef Asen. 2014. “Knowledge-Assisted BIM-Based Visual Analytics for Failure Root Cause Detection in Facilities Management.” *Automation in Construction* 43: 73–83.

Qi, Charles R, Hao Su, Kaichun Mo, and Leonidas J Guibas. 2017. “Pointnet: Deep Learning on Point Sets for 3d Classification and Segmentation.” In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 652–60.

Qi, Charles Ruizhongtai, Li Yi, Hao Su, and Leonidas J Guibas. 2017. “Pointnet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space.” *Advances in Neural Information Processing Systems* 30.

Shahinmoghadam, Mehrzad, Worawan Natephra, and Ali Motamedi. 2021. “BIM- and IoT-Based Virtual Reality Tool for Real-Time Thermal Comfort Assessment in Building Enclosures.” *Building and Environment* 199: 107905. https://doi.org/10.1016/j.buildenv.2021.107905.

Srivastava, Nitish, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. 2014. “Dropout: A Simple Way to Prevent Neural Networks from Overfitting.” *The Journal of Machine Learning Research* 15 (1): 1929–58.

Tang, Pingbo, Daniel Huber, Burcu Akinci, Robert Lipman, and Alan Lytle. 2010. “Automatic Reconstruction of As-Built Building Information Models from Laser-Scanned Point Clouds: A Review of Related Techniques.” *Automation in Construction* 19 (7): 829–43.

Wang, Yue, Yongbin Sun, Ziwei Liu, Sanjay E Sarma, Michael M Bronstein, and Justin M Solomon. 2019. “Dynamic Graph Cnn for Learning on Point Clouds.” *Acm Transactions On Graphics (Tog)* 38 (5): 1–12.

Wu, Zhirong, Shuran Song, Aditya Khosla, Fisher Yu, Linguang Zhang, Xiaoou Tang, and Jianxiong Xiao. 2015. “3d Shapenets: A Deep Representation for Volumetric Shapes.” In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 1912–20.

Zhao, Na, Tat-Seng Chua, and Gim Hee Lee. 2020. “Sess: Self-Ensembling Semi-Supervised 3d Object Detection.” In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 11079–87.