Smart City Construction Based on Deep Learning and Building Information Modeling

Jiaheng Lv[#1,a], Yujia Liu[#2,b], Tao Fu[2], Xiran Bai[3], Sirui Wang[1], Jiangjun Li[2,c,*]

[1]School of Electrical and Information Engineering, Beijing University of Civil Engineering and Architecture, Beijing, China
[2]School of Urban Economics and Management, Beijing University of Civil Engineering and Architecture, Beijing, China
[3]School of Environmental and Energy Engineering, Beijing University of Civil Engineering and Architecture, Beijing, China
[#]Contributed equally

[a]1992177512@qq.com (J.L),[b]2071681108@qq.com (Y.L)
* Correspondence: [c]ljjfuture@126.com (J.Li)

Abstract. In recent years, BIM research has significantly influenced the economic planning of smart cities globally. With advancements in big data, the Internet of Things, and other emerging technologies, BIM technology, has evolved, and CIM based on BIM has become crucial in urban construction. However, managing large and complex traditional model data is costly and inefficient. To address this, the study proposes using a deep learning neural network to automatically generate BIM models from 3D point cloud data. By integrating deep learning techniques with BIM, the potential of CIM can be realized, enhancing the efficiency and accuracy of urban planning. The methodology involves collecting and analyzing datasets, training and simulating the framework using RandLA-Net for deep learning, and ultimately confirming the feasibility and efficiency of combining AI and BIM through testing.

Keywords: Building Information Modeling · Point Cloud · RandLA-Net · City Information Modeling

1 Introduction

With the rise of smart city development and global urbanization. Urban economic planning is receiving attention as a facilitation.

In recent years, the rapid development of new technologies, such as the Internet, big data, and the Internet of Things (IoT), has presented both opportunities and challenges for urban economic planning. Artificial Intelligence (AI) and Building Information Modeling (BIM) [1], among other emerging technologies, are becoming increasingly important tools for urban economic planning. However, traditional BIM still has some limitations in data processing, such as the amount of data required being too large to handle. So, it requires optimization and processing.

K. Subramaniyam et al. (eds.), *Proceedings of the 3rd International Academic Conference on Blockchain, Information Technology and Smart Finance (ICBIS 2024)*, Atlantis Highlights in Computer Sciences 21,
https://doi.org/10.2991/978-94-6463-419-8_18

The swift advancement AI offers increased data processing capabilities for smart city development through deep learning. The application of AI technology can enable the deep mining of various data resources in the city to generate more accurate BIM models. Additionally, it can predict and analyze the development trend of the urban economy. In addition, deep learning technology can effectively solve complex problems in urban planning and improve the science and effectiveness of urban planning. For traditional BIM models, the large amount of data leads to an increase in manpower and material resources, and the cost required increases significantly. In terms of software implementation, it has been shown through Shakil Ahmed's research on BIM barriers[2] that traditional training and learning curves are expensive and the software is costly to purchase and the datasets and image sets required are also very large. So, to address this aspect, we collect building datasets and acquire point cloud data through deep learning by efficient semantic segmentation of neural network RandLA-Net to automatically generate BIM models. After testing, the accuracy of our model is greatly improved, and the overall complexity as well as the efficiency is also improved. The following is our specific analysis and implementation of the combination of RandLA-Net deep learning and BIM.

2 Related Work

2.1 Generated BIM from CAD

Nowadays for the development of cities, the construction of City Information Modeling (CIM)[3] is an extremely important part. BIM, as the root of CIM, is even more important as it lays the foundation for constructing a good CIM. However, the data involved in this process is extensive and challenging to implement. In their study, Bin Yang[4] presented a semi-automatic layer classification method to generate a semantically rich structural BIM model as well as to generate the BIM model from the two-dimensional (2D) Computer Aided Design (CAD) drawings. The method extracts semantic information related to individual elements by extracting them from a floor plan and inferring their spatial location. This information is then converted into element parameters. This method semi-automates various processes and simplifies the huge workload of manual modeling. However, it can only generate one structural BIM model for a single floor at a time. The user must manually set up the floors and import the drawings for each floor plan, which requires a technician to implement. Another team, Lu[5] also developed a semi-automatic method to assist in building BIM by extracting structural components from CAD drawings to create a foundation of IFC-based structural BIM objects. These methods have limitations. BIM constructions need to be cut from CAD drawings and cannot be fully automated to generate BIM models. Fig. 1 is CAD to generate BIM flowchart.

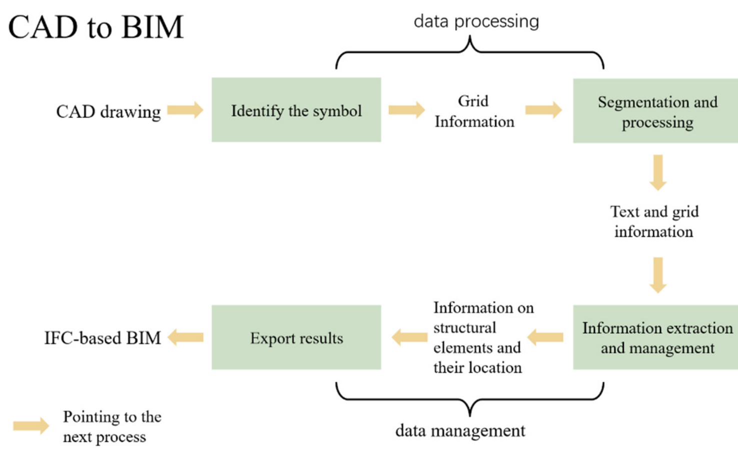

Fig. 1. CAD Generated BIM Flowchart

2.2 BIM generation based on Pointnet++

This study by Charles R. Qi from Stanford University analyses the efficiency of Ponitnet++ [6] in processing data. The study used the farthest point sampling method FPS, which has a time complexity reached O(n2), making Pointnet++ is unsuitable for large-scale semantic segmentation of point clouds[7]. Instead, it is only suitable for localized feature learning. As a result, it is not recommended for large-scale BIM model generation. So, this is where a method is needed to handle large-scale semantic segmentation. Fig. 2 illustrates the basic principle.

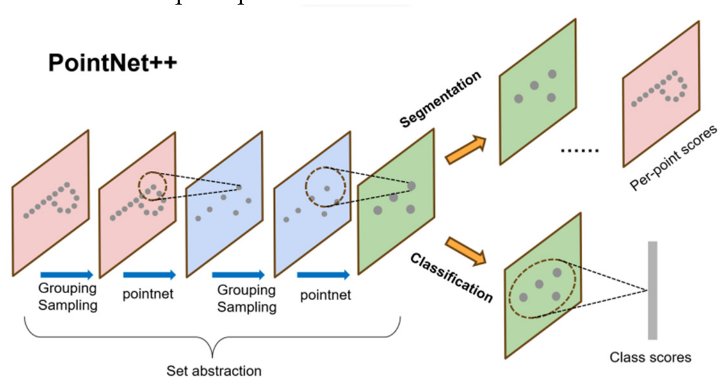

Fig. 2. PointNet++

3 Methodology

3.1 Overview of the methodology

To enhance semantic segmentation, this paper proposed RandLA-Net, a deep learning network that aids in processing the point cloud data to generate BIM models more effectively. The main behand RandLA-Net is to perform local analysis on the point cloud data and learn from randomly selected local samples. The system can categorize each point in the point cloud into various classifications, including walls, floors, doors, and windows, to achieve semantic comprehension and segmentation of the point cloud data. This process enables the generation of precise BIM models.

RandLA-Net disassembles the input point cloud data into localized regions, each containing a set of ordered points. The local regions are sorted using the inverse Euclidean distance, prioritizing the most important ones. This guarantees that local regions of different scales are fully learned during the training process.

After selecting the local regions, RandLA-Net uses local equalization sampling to obtain point pairs with similar distributions. This method can effectively eliminate the category imbalance problem in the point cloud data, improving the network's ability to learn in the classification tasks.

By analyzing and comparing the various Point Cloud sampling methods, it was concluded that Random Sampling is the most suitable for large-scale point sets. A Local Feature Aggregation (LFA) is designed to prevent feature loss caused by random sampling.

This is achieved by gradually increasing the perceptual field of each point. RandLA-Net uses a Multi-Layer Perceptron (MLP)[8] to encode the features in the local region and uses a Fully Connected Neural Network (FCN) to classify the encoded features. For MLP, the transfer from the input layer to the hidden layer can be represented by the following equation.

$$z_j = \sum_{i=1}^{n} \omega_{ij} \cdot x_i + b_j \tag{1}$$

Where,ω_{ij} is the weight connecting between the ith neuron of the input layer and the jth neuron of the hidden layer,x_i is the output of the ith neuron of the input layer,b_j is the bias term of the jth neuron of the hidden layer.

The input can then be nonlinearly transformed by an activation function (activation function) to obtain the output a_j of the jth neuron of the hidden layer. where $f(\cdot)$ is the activation function.

$$a_j = f(z_j) \tag{2}$$

In this way, for each neuron in the hidden layer, its inputs and outputs can be computed. These outputs will be used as inputs to the hidden or output layer, which will continue to be passed and computed.

In this process, the network uses global contextual information and correlations between local regions to enhance the accuracy of semantic segmentation of point cloud data. Fig. 3 illustrates the fundamental structure of the network.

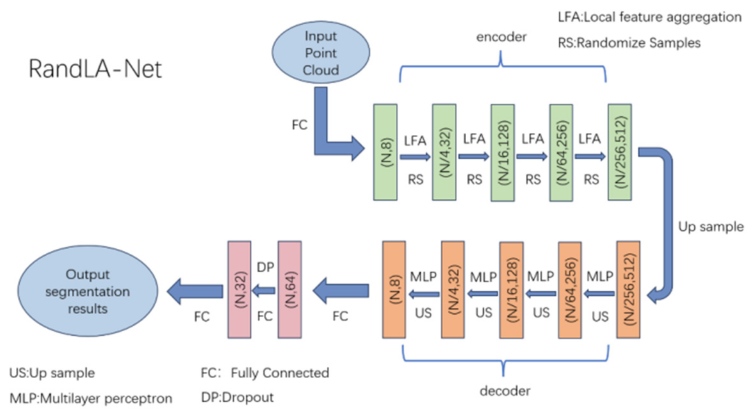

Fig. 3. RandLA-Net structure

3.2 Result

In this case, the dataset S3DIS[9] is used and a random sampling method to reduce the time complexity to O(1), resulting in a significant increase in the data processing performance.

Table 1. Experimental results based on the S3DIS dataset

Methods	Time (second)	Parameters (millions)	Maximum inference points(millions)	mIoU (%)	OA (%)
RandLA-Net	**183**	**1.30**	**1.02**	**79.2**	**94.1**
PointNet++[6]	9231	1.12	0.94	75.7	93.7
PointNet[10]	194	0.93	0.52	72.1	91.2
KPConv[11]	708	15.2	0.55	74.4	92.8

Table 1 The table shows our experimental procedure. By comparing the data such as mIoU value and accuracy of different models, we conclude that Rand-LA-Net is more suitable to deal with large-scale point cloud models. Rand-LA-Net is more accurate and more suitable to deal with large-scale point cloud models.

Table 2. Experimental results based on the SemanticKITTI dataset

Methods	Time (second)	Parameters (millions)	Maximum inference points(millions)	mIoU (%)	OA (%)
RandLA-Net	**171**	**1.27**	**1.31**	**53.4**	**95.1**
PointNet++[6]	9825	5.3	1.16	20.1	93.5
PointNet[10]	209	2.4	0.72	15.2	92.8
KPConv[11]	954	11.2	0.94	30.6	93.2

To test the effect on different data sources and different scenarios. We trained again on the SemanticKITTI dataset[12]. The corresponding results are shown in **Table 2**.

4 Conclusions

In summary, the RandLa-Net model performs well in processing large point cloud data and generating BIM.

Firstly, it efficiently handles point cloud data through random sampling and local aggregation, avoiding computational difficulties and resource wastage. Secondly, it captures spatial relationships and semantic information through multi-resolution feature extraction to ensure reliable structural restoration. As a result, the generated BIM is more accurate and complete, facilitating building design, construction and maintenance. Finally, the model is highly versatile and scalable, performing well in a variety of scenarios. The enhanced ability to process large point cloud data improves construction efficiency and quality.

References

1. K. Barlish and K. Sullivan, "How to measure the benefits of BIM — A case study approach," *Autom. Constr.*, vol. 24, pp. 149–159, Jul. 2012, doi: 10.1016/j.autcon.2012.02.008.
2. S. Ahmed, "Barriers to Implementation of Building Information Modeling (BIM) to the Construction Industry: A Review," *J. Civ. Eng. Constr.*, vol. 7, no. 2, p. 107, May 2018, doi: 10.32732/jcec.2018.7.2.107.
3. M. L. Roumyeh and V. L. Badenko, "Integrating BIM and GIS to Move Towards CIM," in *BIM-моделирование в задачах строительства и архитектуры: материалы IV Международной научно-практической конференции. Под общ. ред. А. А. Семенова*, СПбГАСУ, Apr. 2021, pp. 14–26. doi: 10.23968/BIMAC.2021.002.
4. Yang B., Liu B., Zhu D., Zhang B., Wang Z., and Lei K., "Semiautomatic Structural BIM-Model Generation Methodology Using CAD Construction Drawings," *J. Comput. Civ. Eng.*, vol. 34, no. 3, p. 04020006, May 2020, doi: 10.1061/(ASCE)CP.1943-5487.0000885.
5. Q. Lu and S. Lee, "A Semi-Automatic Approach to Detect Structural Components from CAD Drawings for Constructing As-Is BIM Objects," in *Computing in Civil Engineering 2017*, Seattle, Washington: American Society of Civil Engineers, Jun. 2017, pp. 84–91. doi: 10.1061/9780784480823.011.
6. C. R. Qi, L. Yi, H. Su, and L. J. Guibas, "PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space".
7. Z. Zhang, Y. Dai, and J. Sun, "Deep learning-based point cloud registration: an overview," *Virtual Real. Intell. Hardw.*, vol. 2, no. 3, pp. 222–246, Jun. 2020, doi: 10.1016/j.vrih.2020.05.002.
8. A. Pinkus, "Approximation theory of the MLP model in neural networks," *Acta Numer.*, vol. 8, pp. 143–195, Jan. 1999, doi: 10.1017/S0962492900002919.
9. I. Armeni, S. Sax, A. R. Zamir, and S. Savarese, "Joint 2D-3D-Semantic Data for Indoor Scene Understanding." arXiv, Apr. 05, 2017. Accessed: Jan. 27, 2024. [Online]. Available: http://arxiv.org/abs/1702.01105
10. C. R. Qi, H. Su, K. Mo, and L. J. Guibas, "PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation." arXiv, Apr. 10, 2017. Accessed: Jan. 19, 2024. [Online]. Available: http://arxiv.org/abs/1612.00593
11. H. Thomas, C. R. Qi, J.-E. Deschaud, B. Marcotegui, F. Goulette, and L. J. Guibas, "KPConv: Flexible and Deformable Convolution for Point Clouds." arXiv, Aug. 19, 2019. Accessed: Jan. 19, 2024. [Online]. Available: http://arxiv.org/abs/1904.08889

12. J. Behley *et al.*, "SemanticKITTI: A Dataset for Semantic Scene Understanding of LiDAR Sequences," in *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, Seoul, Korea (South): IEEE, Oct. 2019, pp. 9296–9306. doi: 10.1109/ICCV.2019.00939.