Proceedings Paper:

Zhao, X., Wang, T.H. orcid.org/0000-0002-4500-8415 and Peng, C. orcid.org/0000-0001-8199-0955 (2022) Automatic room type classification using machine learning for two-dimensional residential building plans. In: Pak, B., Wurzer, G. and Stouffs, R., (eds.) Proceedings of the 40th Conference on Education and Research in Computer Aided Architectural Design in Europe. eCAADe 2022, 13-16 Sep 2022, Ghent, Belgium. eCAADe (Education and research in Computer Aided Architectural Design in Europe) and KU Leuven Faculty of Architecture , pp. 593-600. ISBN 9789491207334

Automatic Room Type Classification using Machine Learning for Two-Dimensional Residential Building Plans

Xingjian Zhao[1], Tsung-Hsien Wang[2], Chengzhi Peng[3]
[1,2,3]University of Sheffield
[1,2,3]{xzhao66|tsung-hsien.wang|c.peng}@sheffield.ac.uk

Building plan semantic retrieval is of interest in every stage of construction and facility management processes. A conceptual design model with a space layout can be used for the early building evaluation, such as functional spatial validation, circulation and security checking, cost estimation, and preliminary energy consumption simulation. With the development of information technology, existing machine learning methods applied to semantic segmentation of building plan images have successfully identified building elements such as doors, windows, and walls. However, for the higher level of room type/function recognition, the prediction accuracy is low when building plans do not contain sufficient details such as furniture. In this paper, we present a workflow and a predictive model for residential room type classification. Given a building plan image, the building elements are first identified, followed by room feature extraction by connectivity and morphological characterization using a rule-based algorithm. The Multi-Layer Perceptron (MLP) is trained with the feature set and then predicts the room type of test samples. We collected 1,586 residential room samples from 165 building layout plans and categorized rooms into nine types. Finally, our current model can achieve a classification accuracy of 0.82.

Keywords: *Floor Plan Semantic Retrieval, Room Type Classification, Machine Learning*

INTRODUCTION

In architectural projects, the building layout plan diagrams play a crucial role in building construction and facility management (Lee et al. 2012). According to Eastman (2009), a space layout can be used for estimating spatial validation, circulation analysis, preliminary cost estimation and energy assessment. In the early design stage, the project information tends to be saved and delivered in the form of two-dimensional diagrams (Zhao, Deng and Lai 2021). While understanding building plans can be straightforward for architects (Zeng et al. 2019), the manual processes of building semantic information acquisition, processing and management limit productivity and are prone to errors and omissions (Bazjanac and Crawley 1997).

Automatic building plan interpretation has been investigated for a long time. Liu et al (2017) employed Convolutional Neural Networks (CNN) to detect junction points in the building plan drawings to obtain the location of walls and doors. Cho et al (2020) adopted style-transfer algorithms based on Conditional Generative Adversarial Networks (CGAN) to extract the wall, door, and window pixels; their network performed well on the hand-drawn diagrams. However, there are several challenges regarding the higher-level room function semantic recognition. Current approaches focus on image-based deep learning techniques, categorizing room regions on the pixel level. These methods use a very small convolution kernel to scan the input image, extract texture distribution or color difference features to classify each pixel (Wang et al. 2018). For building plan diagrams, the classification accuracy depends on the details contained in the building plan (such as furniture layout), which are usually not available in early building plans. A building plan image always has many blank pixels without any semantic content and is different from a natural photograph that has an essential feature in every pixel, therefore, the former requires specific information for its description to make it more meaningful (Goyal, Chattopadhyay and Bhatnagar 2021).

In this paper, we propose a new tool for semantic analysis of residential room types. Referring to the method developed by Zeng et al (2019), which can identify the wall and opening pixels in plan diagrams, we improved their method and further differentiated walls, windows, and doors. The closed boundaries surrounded by the pixels of these building elements are then marked as room regions. Through our algorithm, the morphological characteristics and spatial accessibility of room regions, as well as the information of building elements in adjacent rooms, are automatically computed. The feature of each room region is represented by a series of numerical values, and we then manually labeled the corresponding room type to produce the training dataset. A Multi-Layer Perceptron (MLP) model was trained with the labeled rooms. We evaluated the resultant model showing a room type classification accuracy of 0.82.

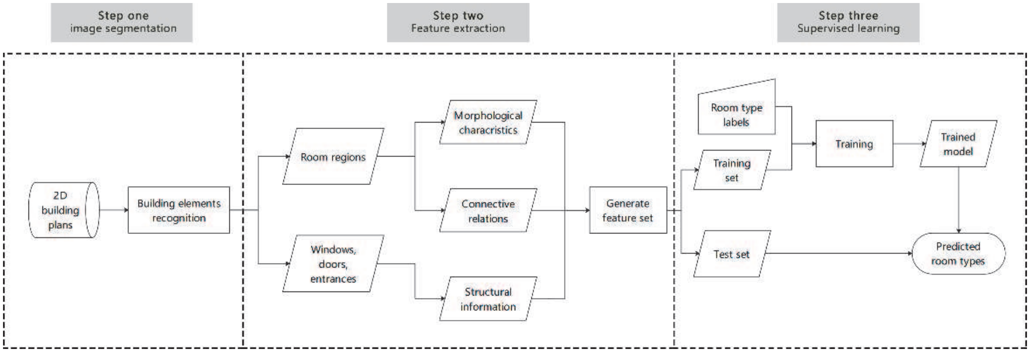
Our contribution is two-fold. First, our method can classify room functions at the level of objects. Compared to previous pixel-based classification methods, our method does not require specific details contained in the building plan. Besides, our method can construct connectivity information between rooms and adjacent relationships between rooms and building elements, which can be used for further automatic three-dimensional modeling or building plan-based energy performance analysis. For example, He et al (2021) proposed a predictive model for daylight performance of general floor plans. The automatic floor plan understanding can help to omit the process of manual floor plan encoding.

ROOM CLASSIFICATION

With the development of computer vision, image-based deep learning methods are popular to address image segmentation and semantic retrieval problems (O'Mahony et al. 2019). Huang and Zheng (2018) used Generative Adversarial Networks (GAN) to identify room functions and building elements in the architectural drawings. Their method performed well in recognizing rooms with orthogonal shapes and furniture symbols but led to errors when a room has irregular shapes or insufficient details. Additionally, they labeled some complete room regions into separate room types, e.g., living room and dining room, causing the network to produce significant noise close to boundaries. Lu et al (2021) and Ahmed et al (2012) detected the room type text in the image to identify room regions. The approach is obviously not applicable when the input floor plan image has no text information. Zeng et al (2019) designed a room-boundary guided multi-task network. They proposed a hierarchy mechanism to firstly identify pixels as room-boundary pixels or room-type pixels, and then classify them into individual elements. Their network can recognize wall and opening pixels even if the input building plan has curve boundaries. Although they developed a post-process that determined the predicted room type based on dominant pixels, the prediction accuracy of room type is still much lower than walls and openings.

Some researchers adopted data-driven methods for room function classification. Mewada et al (2020) collected the room area and aspect ratio as input to distinguish a region as a room or non-room type. Bloch and Sacks (2018) extracted five feature values, including area, number of doors, number of windows, number of room boundary lines, and floor level offset from residential Building Information

Figure 1
The integrated workflow of automatic room type classification



Modeling (BIM) models. They utilized the Artificial Neural Network (ANN) to learn these features and classified rooms into 15 types with a maximum accuracy of 0.82. Buyuklieva (2020) raised a hypothesis that there is a link between room geometry and usage in office buildings. Visual Graph Analysis (VGA) measurement in the space syntax theory was computed to categorize spaces. Each space was divided into many nodes and the amount of each node that can be connected to another inter-visible node was used to quantify the spatial visibility to predict space functions. Wang, Sacks and Yeung (2022) generated the graph representation of the apartment layout image. They extracted rooms as nodes and connections as edges, to distinguish room functions through Graph Neural Networks (GNN). However, existing data-driven approaches require different platforms and software to collect room feature information or need manual data preparation, which is time-consuming and error-prone.

In conclusion, using an image-based or a data-driven approach alone cannot fully address the challenge of semantic segmentation of residential room types. This paper proposes an integrated method to facilitate both building plan information acquisition and semantic analysis by combining the above-mentioned approaches.

METHODOLOGY

An integrated workflow

The purpose of this research is to classify room functions in two-dimensional residential building plan diagrams through automatic rule-based feature extraction and calculation.

Figure1 illustrates the process of our method. Given a building plan image, we start by recognizing the basic elements (walls, windows, and doors) and detect the room boundaries. Next, we compute the morphological characteristics of closed regions. The building elements are associated with the adjacent room region as structural information, in which the doors are used to construct the connectivity matrix between room regions and to calculate the spatial accessibility of the room regions. We constructed the room feature set and manually labeled them into nine room types. Finally, the MLP is trained with our training set, and the test set is used to evaluate the model performance.

The Dataset

In this paper, we adopted the R3D dataset proposed by Zeng et al (2019), containing plans with both regular and non-regular shapes. These versatile polygonal shapes provide potential to investigate the correlation between the morphology of the room and its associated function. We selected 165 floor plan images in R3D and obtained 1,649 room samples. A total of 14 room types were manually labeled for these samples, namely *balcony*, *bathroom*, *bedroom*, *closet*, *corridor*, *courtyard*, *garage*, *hallway*, *kitchen*, *laundry*, *living room*, *maid room*, *study*, and *utility*. As shown in Figure 2, different room types vary hugely in the numbers of samples, in which the lower, median and upper quartiles are 12, 55, and 244.75. We removed the courtyard, garage, laundry, maid room and study whose sample numbers are below or close to the lower quartiles. The final dataset contains 1590 room samples.

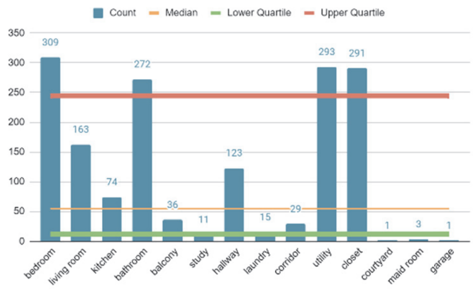


Figure 2
The number of different room types

Building element recognition

The deep multi-task network proposed by Zeng et al (2019) successfully recognize walls and openings (including doors and windows) in the floor plan image. We improved their model to further distinguish windows and doors. Specifically, we defined the door adjacent to the outside as the entrance, which is marked with red pixels (Figure 3 (b)). The room regions are extracted as shown in Figure 3 (c).

Feature extraction

Morphological characteristics. Björk (1992) thinks that space can be defined in two complementary ways; one is based on physical separation; the other one is through homogeneous activity. Our implementation extracted the room boundary as a polygon and adopted the Douglas-Peucker (DP) algorithm for simplification. DP is a well-known line simplification method that reduces curve complexity and preserves geometry features by deleting non-characteristic points and extracting character points (Ari et al. 2022). Figure 4 presents a schematic explanation of the DP algorithm. Assume that a trajectory S is composed of a collection of seven points (S_1, S_2, ⋯, S_7):

1. Connect the start point S_1 to its farthest point S4 and get the line segment $L_{S1 \to S4}$.

Figure 3
(a) the input floor plan image (b) recognized building elements including windows (pink), doors (blue) and entrances (red) (c) detected room regions

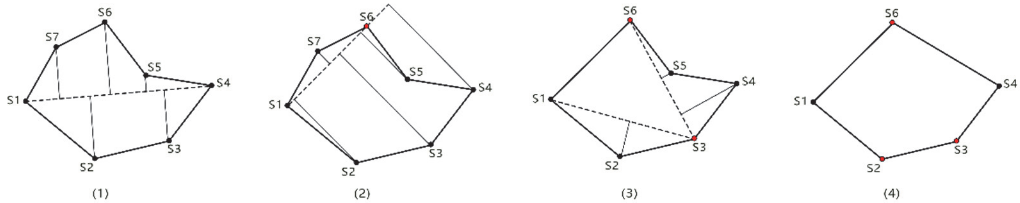


Figure 4
Process of the Douglas-Peucker algorithm

2. Calculate the Euclidean distance of all other points on the S to the $L_{S1 \rightarrow S4}$, record the maximum distance as Dmax and find the corresponded point S_6.
3. Compare the D_{max} with the predefined tolerance D. If the Dmax is smaller than D, then the line $L_{S1 \rightarrow S4}$ is treated as the approximation of the original segment.
4. Otherwise, the trajectory S is divided into two parts by the point S_6, and repeat the process 1 ~ 3 for these two parts.

After the simplification, we calculated the number of boundary lines, the area ratio, and the distance between room center and floor plan center to describe the morphological feature of room regions. In particular, we mapped this distance to 0 to 1 to eliminate the possible errors caused by different image sizes.

Spatial accessibility. Space syntax utilises analytical, quantitative and descriptive measures for analysing space relationships in buildings and cities (Hillier and Hanson, 1984). In space syntax, Visibility Graph Analysis (VGA) provides several ways to represent the continuous spatial layout of a building. A series of computed values can present the accessibility of spaces in the plan system (Turner, 2001). Most measurement of the VGA is based on the step depth. As shown in Figure 5, where the nodes represent rooms and the edges represent door connections, the step depth is the least number of edges between two rooms. Through analysing the adjacency relationship between doors and room regions, we can get the connectivity matrix of the floor plan to calculate the step depth of each room region to others. For example, the connectivity matrix of R_1 in Figure 5 can be represented as: $C_{R1} = \{0\ 1\ 0\ 1\ 0\ 0\ 0\ 0\ 0\}$. We further computed the mean depth and the entrance depth to express the spatial accessibility of each room. The former means the cumulative sum of the step depth of one room divided by the total number of rooms. The resulting mean depth value indicates the accessibility of a room within a plan layout. The latter is the step depth of each room to the entrance, revealing room accessibility from the outside.

Structural information. We calculated the number of hosted windows and doors for each room region. Since the size of windows and doors may vary from plan to plan, we also computed the pixel ratio of each window and door.

Machine Learning Approach

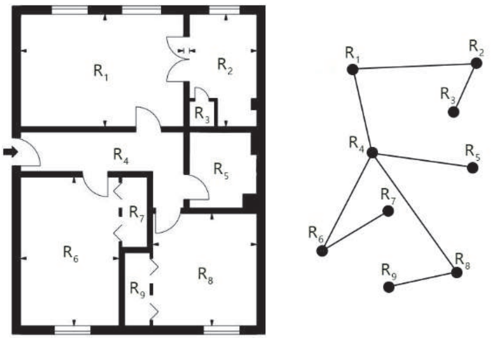


Figure 5
The floor plan image and its graph representation

In supervised machine learning, we use a set of geometrical, spatial, and structural features to represent each room object. The purpose is to discover underlying relationships between samples and assigned labels to enable the classifier for the new instance prediction (Kotsiantis, Zaharakis and Pintelas, 2007). We collected nine numerical features for each room sample, namely *area ratio*, *number of boundary lines*, *the distance between room centre and plan centre*, *number of windows*, *the ratio of window pixels*, *number of doors*, *the ratio of door pixels*, *mean depth* and *entrance depth*. The MLP model in the Scikit-learn library was employed to implement the machine learning procedure. Scikit-learn provides many machine learning algorithms for both classification and regression problems, in which the MLP is a well-known deep learning algorithm that simulates the biological structure of human brain (Pedregosa et al. 2011). To make the results

	Overall	Bedroom	Living	Kitchen	Bath	Balcony	Hallway	Corridor	Utility	Closet
Zeng et al	0.66	0.79	0.93	/	0.78	0.49	0.68	/	/	0.54
Ours	**0.82**	**0.86**	**0.98**	**0.22**	**0.79**	**0.83**	**0.93**	**0.50**	**0.62**	**1**

Table 1
Comparison with Deep multi-task network

	Bedroom	Living	Kitchen	Bathroom	Balcony	Hallway	Corridor	Utility	Closet
Bedroom	0.86	0.04	0.05	0.05	0	0	0	0	0
Living	0.02	0.98	0	0	0	0	0	0	0
Kitchen	0.46	0	0.22	0.25	0	0	0	0.01	0.05
Bathroom	0.02	0	0.01	0.79	0.02	0	0	0.02	0.14
Balcony	0	0	0	0	0.83	0	0	0	0.17
Hallway	0	0	0	0	0	0.93	0.07	0	0
Corridor	0	0	0	0	0	0.50	0.50	0	0
Utility	0	0	0	0.01	0	0	0	0.62	0.37
Closet	0	0	0	0	0	0	0	0	1

Table 2
Confusion matrix of prediction results

comparable, our training and test sets were consistent with those of Zeng et al (2019), with 1,649 room samples divided into 1,317 for training and 332 for testing. Since the MLP is sensitive to feature scaling, all feature values were normalized before training.

RESULTS AND DISCUSSION

Compared with the deep multi-task network proposed by Zeng et al (2019), our classification accuracy of all room types was improved (Table1). Specifically, the accuracy of the closet improved significantly from 0.54 to 1.0, and the accuracy of the hallway improved from 0.68 to 0.93. These two room types often do not have a detailed layout in the building plan, but the entrance depth and area ratio can distinguish them well. In addition, we added three new room types to classify, where the utility has a good prediction accuracy of 0.62, the corridor and kitchen have a poor identification accuracy of 0.50 and 0.22.

The confusion matrix (Table2) shows that the MLP recognized most of the bedrooms, living rooms, bathrooms, hallways, and all closets. These five room types tend to be easier distinguished. For example, the living room usually has larger areas, while the closet usually is very small; hallways are always adjacent to the entrance. Most kitchens are considered as bedrooms and bathrooms, suggesting that the MLP cannot identify kitchens and more decisive features are necessary. Hallways and corridors have similar functions, where the former connects with the main entrance, the latter connects with most rooms. In the R3D dataset, some of the hallways and corridors share one room region, which makes them confusing. Figure 6 shows the input image and output results. In this example, our model predicts the room type for 10 room regions, with the wrong prediction colored with red. W_1~W_7, D_1~D_9, *and* E_1 represent the recognized windows, doors, and entrance.

To summarize, for the problem of room semantic classification in two-dimensional residential building plan diagrams, the proposed method shows a better performance. For rooms that have similar features in the current dataset, more representative features need to be collected to differentiate them.

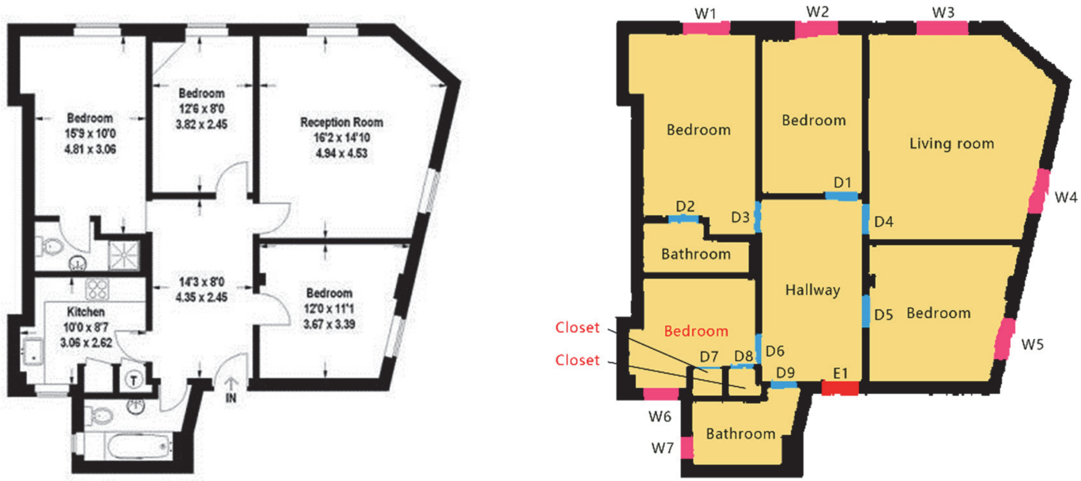


Figure 6
The input floor plan image, detected building elements and predicted room types

CONCLUSION

This paper presents an integrated method to classify room functions based on residential building plan diagrams. We developed a workflow integrating automatic feature extraction and rule-based reasoning. The results showed that our method could effectively recognize six room types, including bedroom, living room, closet, hallway, balcony and bathroom, given limited information about the wall, window and door. Meanwhile, our algorithm calculates the morphological characteristics, connective relations and adjacent building elements of room regions of the input building plan. For future improvement, the symbol detection algorithm can be used to identify the distribution of furniture in the floor plan and then establish a hierarchical room type prediction mechanism. For instance, furniture symbols such as cooking equipment or shower can effectively facilitate kitchen or toilet room type recognition. The hierarchical machine learning can then incrementally determine uncategorised rooms adjacent to detected ones.

REFERENCES

Ahmed, S., Liwicki, M., Weber, M. and Dengel, A., 2012, March. Automatic room detection and room labeling from architectural floor plans. In *2012 10th IAPR international workshop on document analysis systems* (pp. 339-343). IEEE.

Ari, B., Sobahi, N., Alçin, Ö.F., Sengur, A. and Acharya, U.R., 2022. Accurate detection of autism using Douglas-Peucker algorithm, sparse coding based feature mapping and convolutional neural network techniques with EEG signals. *Computers in Biology and Medicine, 143*, p.105311.

Bazjanac, V. and Crawley, D.B., 1997. The implementation of industry foundation classes in simulation tools for the building industry.

Björk, B.C., 1992. A conceptual model of spaces, space boundaries and enclosing structures. *Automation in construction, 1*(3), pp.193-214.

Bloch, T. and Sacks, R., 2018. Comparing machine learning and rule-based inferencing for semantic enrichment of BIM models. *Automation in Construction, 91*, pp.256-272.

Buyuklieva, B., 2020. Machine Learning for Exploring Spatial Affordance Patterns. *arXiv preprint arXiv:2005.08106*.

Cho, D., Kim, J., Shin, E., Choi, J. and Lee, J.K., 2020. Recognizing Architectural Objects in Floor-plan Drawings Using Deep-learning Style-transfer Algorithms.

Eastman, C., 2009. Automated assessment of early concept designs. *Architectural Design, 79*(2), pp.52-57.

Goyal, S., Chattopadhyay, C. and Bhatnagar, G., 2021. Knowledge-driven description synthesis for floor plan interpretation. *International Journal on Document Analysis and Recognition (IJDAR), 24*(1), pp.19-32.

Hillier and Hanson, J., 1984. *The social logic of space*. Cambridge university press.

Huang, W. and Zheng, H., 2018. Architectural drawings recognition and generation through machine learning.

Kotsiantis, S.B., Zaharakis, I. and Pintelas, P., 2007. Supervised machine learning: A review of classification techniques. *Emerging artificial intelligence applications in computer engineering, 160*(1), pp.3-24.

Lee, J.K., Lee, J., Jeong, Y.S., Sheward, H., Sanguinetti, P., Abdelmohsen, S. and Eastman, C.M., 2012. Development of space database for automated building design review systems. *Automation in Construction, 24*, pp.203-212.

Liu, C., Wu, J., Kohli, P. and Furukawa, Y., 2017. Raster-to-vector: Revisiting floorplan transformation. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 2195-2203).

Lu, Z., Wang, T., Guo, J., Meng, W., Xiao, J., Zhang, W. and Zhang, X., 2021. Data-driven floor plan understanding in rural residential buildings via deep recognition. *Information Sciences, 567*, pp.58-74.

Mewada, H.K., Patel, A.V., Chaudhari, J., Mahant, K. and Vala, A., 2020. Automatic room information retrieval and classification from floor plan using linear regression model. *International Journal on Document Analysis and Recognition (IJDAR), 23*(4), pp.253-266.

O'Mahony, N., Campbell, S., Carvalho, A., Harapanahalli, S., Hernandez, G.V., Krpalkova, L., Riordan, D. and Walsh, J., 2019, April. Deep learning vs. traditional computer vision. In *Science and Information Conference* (pp. 128-144). Springer, Cham.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V. and Vanderplas, J., 2011. Scikit-learn: Machine learning in Python. the Journal of machine Learning research, 12, pp.2825-2830.

He, Q., Li, Z., Gao, W., Chen, H., Wu, X., Cheng, X. and Lin, B., 2021. Predictive models for daylight performance of general floorplans based on CNN and GAN: A proof-of-concept study. Building and Environment, 206, p.108346.

Turner, A., 2001, May. Depthmap: a program to perform visibility graph analysis. In *Proceedings of the 3rd International Symposium on Space Syntax* (Vol. 31, pp. 31-12).

Wang, P., Chen, P., Yuan, Y., Liu, D., Huang, Z., Hou, X. and Cottrell, G., 2018, March. Understanding convolution for semantic segmentation. In *2018 IEEE winter conference on applications of computer vision (WACV)* (pp. 1451-1460). Ieee.

Wang, Z., Sacks, R. and Yeung, T., 2022. Exploring graph neural networks for semantic enrichment: Room type classification. *Automation in Construction, 134*, p.104039.

Zeng, Z., Li, X., Yu, Y.K. and Fu, C.W., 2019. Deep floor plan recognition using a multi-task network with room-boundary-guided attention. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 9096-9104).

Zhao, Y., Deng, X. and Lai, H., 2021. Reconstructing BIM from 2D structural drawings for existing buildings. *Automation in Construction, 128*, p.103750.