FloorplanSBS: Synthesizing Vector Floorplans by Patch-Based Floorplan Segmentation - Supplementary Material

Wenming Wu
Hefei University of Technology
Hefei, China
wwming@hfut.edu.cn

Tianlei Sheng
Hefei University of Technology
Hefei, China
2023110572@mail.hfut.edu.cn

Gaofeng Zhang
Hefei University of Technology
Hefei, China
g.zhang@hfut.edu.cn

Liping Zheng*
Hefei University of Technology
Hefei, China
zhenglp@hfut.edu.cn

1 Application

Our model excels at generating floorplans under various input constraints, supporting both boundary-only inputs and combinations of layout graphs and boundaries. Users can simply provide a house boundary and optionally include a layout graph specifying room types and sizes, to effortlessly generate structured floorplans. For users without a clear layout in mind, our model offers an intuitive and flexible solution: by adding a few simple segmentation lines, users can indicate how the space should be partitioned. Meanwhile, layout graphs can be used to define room categories, approximate locations, and connectivity, allowing users to customize the design based on their preferences. This flexibility empowers even non-experts to explore and experiment with diverse layout configurations.

To further improve usability, we developed an interactive application, as shown in Figure 1. The interface allows users to easily add horizontal or vertical segmentation lines by clicking on the design canvas. In addition, it features an automatic segmentation mode powered by our model, enabling one-click floorplan generation based on the provided boundary. This dual functionality—supporting both manual customization and automated generation—offers a versatile tool for users of all levels, from beginners to professionals, to explore, refine, and perfect their floorplan designs.

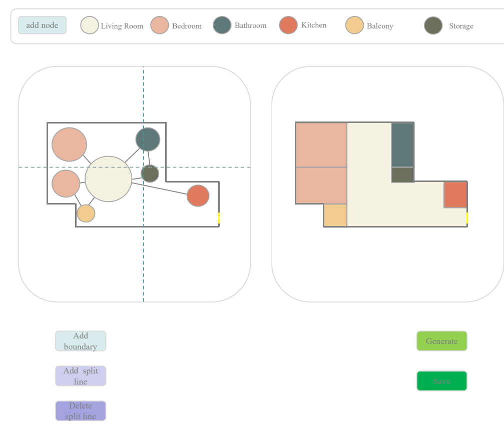

Figure 1: Our interface for user design of vector floorplans.

2 Evaluation metric

Unlike other floorplan design methods that focus on building generative models based on diversity and compatibility and evaluating their performance, our approach emphasizes accurately predicting floorplan segmentation. To ensure the predicted floorplans align closely with the ground truth, we primarily use class-wise mean Accuracy (mAcc) to compute the average classification accuracy for each category. Additionally, we calculate the mean Intersection over Union (mIoU) on rendered images for pixel-level evaluation. Fréchet Inception Distance (FID) and Kernel Inception Distance (KID) are employed to measure the differences between the generated floorplan images and real floorplan images.

To further quantify the distributional differences between the predicted results and ground truth, we introduce Maximum Mean Discrepancy (MMD) [4] as an additional evaluation metric. Specifically, we use a pretrained ResNet-18 model to extract high-dimensional feature representations from both the predicted and ground truth images, excluding the final fully connected layer. For the feature vectors X and Y, the MMD value is computed using the following formula :

$$\begin{aligned}\mathrm{MMD}^2(X,Y) = &\frac{1}{m^2}\sum_{i,j}k(x_i,x_j)+\frac{1}{n^2}\sum_{i,j}k(y_i,y_j)\\ &-\frac{2}{mn}\sum_{i,j}k(x_i,y_j),\end{aligned} \tag{1}$$

where k is a Gaussian kernel function defined as:

$$k(x,y)=\exp\left(-\frac{\|x-y\|^2}{2\sigma^2}\right). \tag{2}$$

These metrics evaluate the quality of the generated floorplans both at the pixel level and overall. However, they are less sensitive to issues that significantly affect the completeness and usability of floorplans in downstream tasks. In intermediate results generated by *Graph2Plan* [1] and *RPLAN* [2], the bounding boxes of rooms often fail to align perfectly. To address this, we use the Room Intersection (RI) [3] metric to compare our results and demonstrate our superior performance. The RI metric matches predicted rooms (p) with ground truth rooms (g) based on their Intersection over Union (IoU). A pair of rooms is considered matched if and only if

*Corresponding author.

MM '25, Dublin, Ireland

ACM ISBN 979-8-4007-2035-2/2025/10
https://doi.org/10.1145/3746027.3754556

their $IoU(p, g)$ is greater than or equal to 0.5 and the two rooms have the same type. Matched pairs are considered True Positives (TP), while predicted rooms without a matching ground truth are counted as False Positives (FP), and ground truth rooms without a matching predicted room are counted as False Negatives (FN). The RI value is computed using the following formula:

$$\mathrm{RI} = \frac{\sum_{(p,g)\in T_P} \mathrm{IoU}(p, g)}{|T_P| + \frac{1}{2}|F_P| + \frac{1}{2}|F_N|} \tag{3}$$

References

[1] Ruizhen Hu, Zeyu Huang, Yuhan Tang, Oliver Van Kaick, Hao Zhang, and Hui Huang. 2020. Graph2plan: Learning floorplan generation from layout graphs. *ACM Transactions on Graphics (TOG)* 39, 4 (2020), 118–1.

[2] Wenming Wu, Xiao-Ming Fu, Rui Tang, Yuhan Wang, Yu-Hao Qi, and Ligang Liu. 2019. Data-driven interior plan generation for residential buildings. *ACM Transactions on Graphics (TOG)* 38, 6 (2019), 1–12.

[3] Bingchen Yang, Haiyong Jiang, Hao Pan, and Jun Xiao. 2023. Vectorfloorseg: Two-stream graph attention network for vectorized roughcast floorplan segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 1358–1367.

[4] Guandao Yang, Xun Huang, Zekun Hao, Ming-Yu Liu, Serge Belongie, and Bharath Hariharan. 2019. Pointflow: 3d point cloud generation with continuous normalizing flows. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 4541–4550.