FloorPlan-LLaMa: Aligning Architects' Feedback and Domain Knowledge in Architectural Floor Plan Generation

Jun Yin[1,*], Pengyu Zeng[1,*], Haoyuan Sun[1], Yuqin Dai[1], Han Zheng[2], Miao Zhang[1], Yachao Zhang[1], Shuai Lu[1,†]

[1]Tsinghua University
[2]Guangzhou University

* Equal contribution. † Corresponding author.

Contact: shuai.lu@tsinghua.edu.cn, yinj24@mails.tsinghua.edu.cn

Abstract

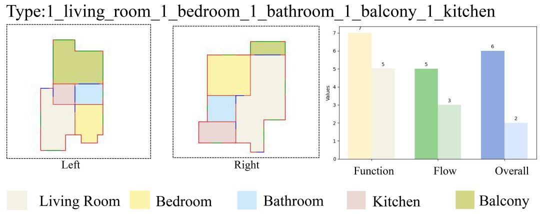
Floor plans serve as a graphical language through which architects sketch and communicate their design ideas. Actually, in the Architecture, Engineering, and Construction (AEC) design stages, generating floor plans is a complex task requiring domain expertise and alignment with user requirements. However, existing evaluation methods for floor plan generation rely mainly on statistical metrics like FID, GED, and PSNR, which often fail to evaluate using domain knowledge. As a result, even high-performing models on these metrics struggle to generate viable floor plans in practice.

To address this, (1) we propose **ArchiMetricsNet**, the first floor plan dataset that includes functionality, flow, and overall evaluation scores, along with detailed textual analyses. We train **FloorPlan-MPS** (Multi-dimensional Preference Score) on it. (2) We develop **FP-LLaMa**, a floor plan generation model based on an autoregressive framework. To integrate architects' professional expertise and preferences, FloorPlan-MPS serves as the reward model during the RLHF (Reinforcement Learning from Human Feedback) process, thereby aligning FP-LLaMa with the needs of community. (3) Comparative experiments demonstrate that our method outperforms baseline models in both text-conditional and class-conditional tasks. Validation by professional architects confirms that our approach yields more rational plans and aligns better with their preferences. Project page is available at: website.

1 Introduction

In architectural design, drafting floor plans is a complex and iterative process (Bakhoum and Wakita, 2023; Zeng et al., 2025a; Qin et al., 2024; Yin et al., 2025). Architects continuously refine floor plan components to balance creative vision with practical requirements (Nauata et al., 2020; Zeng et al.; Zong et al., 2025; Zeng et al., 2025b).

Type:1_living_room_1_bedroom_1_bathroom_1_balcony_1_kitchen



Right:The circulation is very reasonable, but the layout is not well-organized and the overall shape is irregular. The shape of the living room is extremely unreasonable, the overall space utilization is low, but the functional configuration is complete and reasonable.
Left:The circulation is poor, the layout disorganized, and the overall shape irregular. The balcony shape and size are unreasonable, and bathroom/kitchen lighting surpassing the bedroom is also problematic. However, the functional configuration is reasonable.

Figure 1: Architects assess building floor plans through various professional lenses, which can be broadly categorized into functionality, flow, and overall layout.

The design process involves considering numerous factors, such as room number and arrangement, functional distribution, and orientation, while addressing diverse user needs. Traditional methods for evaluating floor plan generation often use metrics such as FID (Heinonen et al., 2008), PSNR (Hore and Ziou, 2010), and GED (Heckman and Rubinstein, 2001). While these metrics assess image quality, distribution similarity, or structural correspondence, they do not evaluate room functionality and may overlook inherent design flaws in the architectural floor plans within the input dataset. Actually, it is important to note that when an architect evaluates a floor plan, their primary focus extends beyond basic image quality. Essentially, they are assessing the layout, design ideas, and other conceptual aspects represented by the image. We have summarized three major limitations in existing research on architectural floor plans.

Major Limitations

1. The current floor plan datasets do not include architects' feedback;
2. Existing metrics cannot effectively evaluate floor plan quality;
3. Current methods for generating floor plans do not integrate the RLHF process.

Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pages 6640–6662
July 27 - August 1, 2025

Recently the GPT (Generative Pre-trained Transformer) framework and the RLHF mechanism have played a significant role in contributing to the advancements of large language models in terms of general capability and user-centered alignment (Ouyang et al., 2022; Chan, 2023; Chaudhari et al., 2024). Our insights derive from viewing architectural floor plans as a professional language of architects, enabling us to leverage methodologies from NLP (Natural Language Processing) tasks. Our contributions are summarized as follows:

1. **ArchiMetricsNet Dataset:** We introduce ArchiMetricsNet, the first floor plan dataset that incorporates three professional evaluation dimensions: functionality, circulation, and overall assessment, accompanied by detailed textual analyses. Figure 1 provides examples. Using this dataset, we train FloorPlan-MPS (Multi-dimensional Preference Score), a model designed for the multi-dimensional evaluation of floor plans. And we refer to the scores generated by this model as ARS (Architectural Reasonableness Scores).

2. **FP-LLaMa Model:** We develop FP-LLaMa, a floor plan generation model based on an autoregressive framework. To integrate architects' professional expertise and preferences, FloorPlan-MPS has been utilized as the reward model during the RLHF process, aligning FP-LLaMa's outputs with the requirements of the community.

3. **Comprehensive validation:** Comparative experiments show that our method surpasses baseline models, including HouseDiffusion, HouseGan++, and Tell2Design, in both text-conditional and class-conditional tasks. Validation by professional architects confirms that our method enhances the rationality of generated plans and aligns with their preferences.

2 Related Works

2.1 Floor Plan Generation

Recent advancements in machine learning have significantly improved floor plan generation. Graph2Plan (Hu et al., 2020) uses graph neural networks to convert layout graphs into functional plans by modeling spatial relationships. House-GAN (Nauata et al., 2020) generates house layouts from functional graphs, while House-GAN++ (Nauata et al., 2020) enhances this by incorporating graph-constrained and conditional GANs for iterative optimization. Diffusion models, such as HouseDiffusion (Shabani et al., 2023) and FloorPlanDiffusion (Ploennigs and Berger, 2023), use denoising processes to create diverse, realistic plans, but rely on traditional metrics like FID and PSNR, which overlook architectural functionality. In terms

Dataset	OpenSource	Size	Desc.	Scores
Rplan	✓	80,788	×	×
CubiCasa5K	✓	5,000	×	×
Tell2Design	✓	80,788	✓	×
SESYD	✓	1,000	×	×
RuralHomeData	✓	800	×	×
ArchiMetricsNet	✓	120,000	✓	✓

Table 1: Summary of dataset characteristics.

of the datasets, Rplan (Wu et al., 2019), with over 80,000 residential plans, provides a robust foundation in floor plan generation tasks; Tell2Design (Leng et al., 2023), introduces language-guided generation with paired descriptions but lacks feasibility evaluations and analyses; CubiCasa, RuralHomeData and SESYD (Kalervo et al., 2019; Lu et al., 2021; Delalandre et al., 2010) offer diverse architectural styles for generalization tests. However, these datasets lack professional evaluations, risking biases from inherently flawed data, and may lead to harmful model outputs. In this paper, we introduce ArchiMetricsNet, the first floor plan dataset incorporating three professional evaluation dimensions: functionality, flow, and overall assessment.

2.2 Reinforcement Learning from Human Feedback

Reinforcement learning from human feedback (RLHF) has enabled language models to produce more truthful, helpful, and harmless outputs. For example, InstructGPT (Ouyang et al., 2022) successfully applied RLHF to GPT-3, achieving significant performance gains. In text-to-image generation, studies have explored learning human feedback to improve visual models. HPS (Wu et al., 2023b) trained a model using the HPD dataset with Stable Diffusion outputs, aligning results with user preferences. HPS v2 (Wu et al., 2023a) expanded this to include multiple generative models and larger datasets but still focused on overall preferences, neglecting diversity. PickScore (Kirstain et al., 2023) collected real-user prompts and annotations, creating a larger dataset, while ImageReward (Xu et al., 2023) added detailed scoring dimensions like alignment and fidelity. MPS (Zhang et al., 2024b) introduced a multi-dimensional preference scoring framework, evaluating aesthetics, semantic alignment, detail quality, and overall satisfaction. Moreover, recent studies (Sun et al., 2025a,b,c) have also explored the paradigm of eliminating the need for explicit reward models by directly fine-tuning models using human preferences. However, these methods focus on natural images and are unable to evaluate or optimize architectural floor plans in the AEC domain. In this paper, we train FloorPlan-MPS, a model designed for the multi-dimensional evaluation of floor plans.

2.3 Autoregressive Generative Model

With the rapid advancement of machine learning technologies(Zhang et al., 2025, 2024a; Wang et al., 2025; YIN et al.), autoregressive models generate text through "next token prediction", a paradigm that has demonstrated exceptional scalability and human-like conversational capabilities in language tasks. Similarly, autoregressive models have been applied to image generation, with notable examples including VQVAE (Razavi et al., 2019), DALL-E (Ramesh et al., 2021), and Parti (Yu et al., 2021). These models use tokenizers to convert images into discrete tokens and predict them sequentially. Extensions like MaskGIT (Chang et al., 2022) and MAGVIT (Yu et al., 2023) adopt masked image modeling, while VAR (Tian et al., 2024) employs hierarchical attributes. Despite surpassing diffusion models in performance, it remains uncertain whether original language model architectures can achieve similar success in floor plan generation. In this paper, we develop FP-LLaMa, a floor plan generation model based on an autoregressive framework and utilize FloorPlan-MPS to integrate architects' professional expertise and preferences.

3 Dataset

Our aim is to build a dataset of architectural floor plans with expert evaluations and professional text analysis, using a systematic processing and evaluation workflow.

3.1 Data Statistics

We collected a total of 24,000 ratings and textual analysis, including 14,000 from Rplan; combining with 2,000 generated from House-GAN, 2,000 generated from HouseGAN++, 2,000 generated from HouseDiffusion, 2,000 generated from Tell2Design, and 2,000 generated from FP-LLaMa.

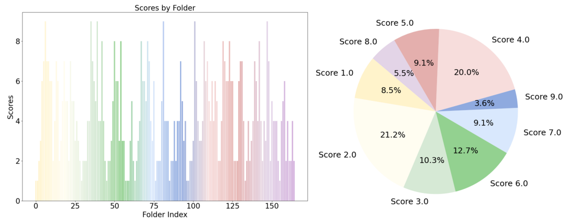


Figure 2: The Overall Scores of different floor plan types.

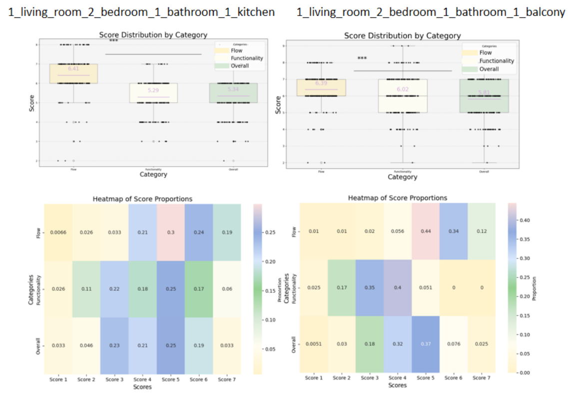


Figure 3: Scores of two typical layouts

Subsequently, we further performed data augmentation and synthesis as shown in Figure 10, resulting in a total of 120,000 pairs of ratings.

3.2 Prompt Collection and Annotation

Initially, we divided them into 165 categories based on the number of rooms and annotated these categories in Figure 9 and Figure 12 (see Appendix A for more details on the ArchiMetricsNet annotation pipeline). Analysis in the Figure 2 shows that 63% of these categories scored below 5, indicating a significant number of unreasonable room types in the original R-plan dataset and the output results of unaligned models. Furthermore, Figure 3 shows that even if the number and types of rooms are reasonable, the layout may still contain unreasonable floor plans. Although some generation models allow users to specify room types and numbers through text or graph structures, bypassing reliance on such data, for contour-conditional generation methods, these data can introduce harmful biases.

4 Method

4.1 FP-LLaMa

In our method, the autoregressive generation mechanism is employed, dexterously eliminating the need for noise prediction in reverse diffusion process during RLHF phase. Specifically, we transform floor plans into discrete token sequences with the *FloorPlan Tokenizer* and employ the

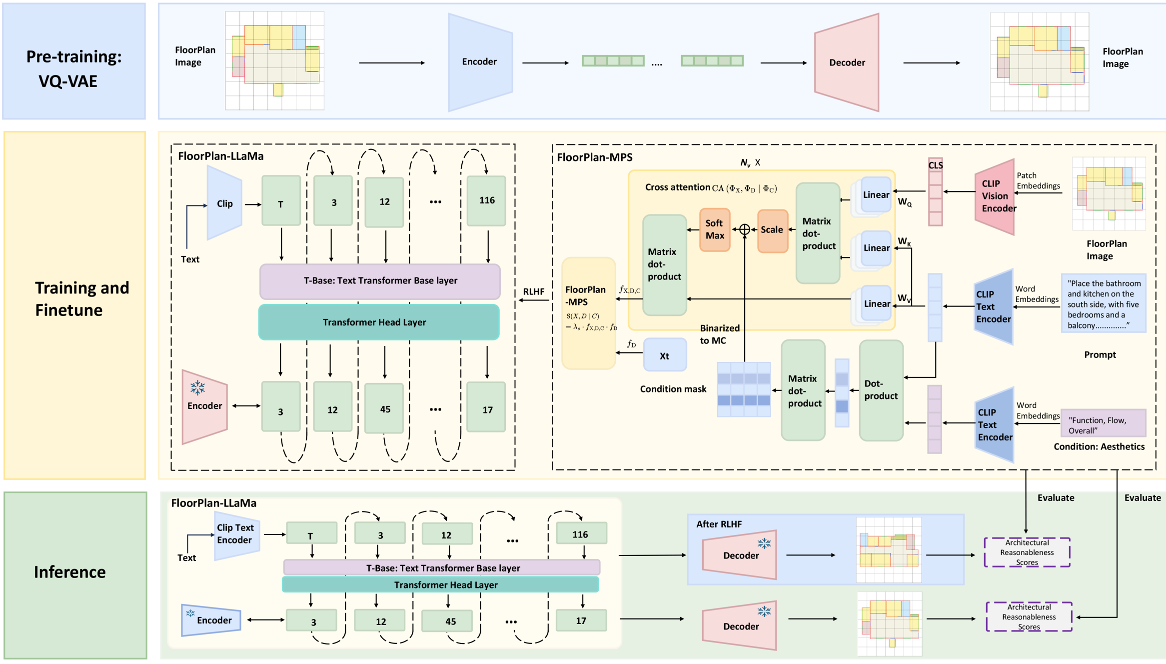


Figure 4: The framework of FP-LLaMa involves several steps. First, we train a VQ-VAE model, followed by training the FP-LLaMa. Subsequently, we apply RLHF using FloorPlan-MPS to align the model with architects' feedback.

Next-Token Prediction to generate floor plans sequentially. Furthermore, in order to enhance efficiency, the *Class-Conditional Mechanism* and the *Classifier-Free Guidance* are incorporated.

FloorPlan Tokenizer. Architecture utilized in our approach is similar to that of VQGAN (Esser et al., 2021), an encoder-quantizer-decoder framework, converting continuous floor plan data into discrete tokens. Firstly, a floor plan $\mathrm{X} \in \mathbb{R}^{\mathrm{H}\times\mathrm{W}\times 3}$ is mapped to the feature matrix $\mathrm{F} \in \mathbb{Q}^{\mathrm{h}\times\mathrm{w}}$ by the encoder, where $\mathrm{h} = \mathrm{H}/p$, $\mathrm{w} = \mathrm{w}/p$, p is the downsample ratio. Furthermore, each element $\mathrm{F}_{(i,j)}$ in the feature matrix is then mapped to its nearest code index $\mathrm{Q}_{(i,j)}$ of the nearest one $\mathrm{F}'_{(i,j)}$ in the codebook. Finally, remapping the index $\mathrm{Q}_{(i,j)}$ back to $\mathrm{F}_{(i,j)}$, and then the decoder reconstructs the floor plan X' from the feature space. For FloorPlan Tokenizer training, three types of loss functions are considered: the reconstruction loss, the perceptual loss from LPIPS (Zhang et al., 2018) and the adversarial loss from the discriminator of PatchGAN (Isola et al., 2017) that is concurrently trained:

$$\mathcal{L}_{\mathrm{FP}} = \underbrace{\|\mathrm{X}-\mathrm{X}'\|_2^2}_{\text{Reconstruction Loss}} + \underbrace{\mathcal{L}_{\mathrm{P}}(\mathrm{X},\mathrm{X}')}_{\text{Perceptual Loss}} + \underbrace{\lambda_{\mathrm{ad}}\mathcal{L}_{\mathrm{ad}}(\mathrm{X}')}_{\text{Adversarial Loss}}.$$

Next-Token Prediction. When it comes to inference, FP-LLaMa generates tokens sequentially with an autoregressive framework. For a floor plan token sequence $\mathrm{Q} = \{\mathrm{Q}_1, \mathrm{Q}_2, \ldots, \mathrm{Q}_{\mathrm{h}\times\mathrm{w}}\}$, the next-token is given by:

$$p(\mathrm{Q}_{\mathrm{next}}) = \prod_{t=1}^{\mathrm{h}\times\mathrm{w}} p(\mathrm{Q}_t|\mathrm{Q}_{<t}, \mathrm{C}),$$

where C represents conditions.

Class-Conditional Mechanism. Floor plan generation is categorized into 165 classes by us, as illustrated in Figure 9, which serve as conditional parameters for generation. Practically, the class embedding is utilized as the prefilling token embedding, from whom starting the generation process.

Classifier-Free Guidance. Although CFG is commonly used in diffusion models (Ho and Salimans, 2022), its application in AR-based image generation is gaining increasing attention. Recent work such as LlamaGen (Sun et al., 2024) has explored this direction. Inspired by these efforts, we integrate CFG into ourl FP-LLaMa. During training, the conditional input is randomly removed and replaced with a null unconditional embedding. Then, in the inference phase, each token's logit ℓ_g is calculated by conditional logit ℓ_c and unconditional logit ℓ_u:

$$\ell_g = \ell_u + s(\ell_c - \ell_u),$$

where s is the scale of classifier-free guidance.

4.2 FloorPlan Multi-dimensional Preference Score

To evaluate architectural floor plans across the three dimensions of function, flow, and overall, a multimodal feature integration framework has been designed, which is illustrated in Figure 4.

4.2.1 Model Structure of FloorPlan-MPS

Firstly, we extract respective features of architectural floor plan X, descriptive text D, and scoring preference condition C with three encoders:

$$\Phi_{\mathrm{X}} = E_{\mathrm{v}}(\mathrm{X}) \quad \Phi_{\mathrm{D}} = E_{\mathrm{t}}(\mathrm{D}) \quad \Phi_{\mathrm{C}} = E_{\mathrm{c}}(\mathrm{C}),$$

where $E_{\mathrm{v}}(\mathrm{X})$ is vision encoder of the CLIP, mapping patch embeddings to visual feature space Φ_{X}; $E_{\mathrm{t}}(\mathrm{D})$ and $E_{\mathrm{c}}(\mathrm{C})$ are text encoders of the CLIP, mapping description embeddings and condition embeddings respectively to description feature space Φ_{D} and condition feature space Φ_{C}.

Furthermore, in order to integrate visual modality Φ_{X} and textual modality Φ_{D} effectively, the Cross Attention (CA) module is introduced by us:

$$\mathrm{CA}(\Phi_{\mathrm{X}}, \Phi_{\mathrm{D}}) = \sigma\left(\frac{\Phi_{\mathrm{X}} W_q (\Phi_{\mathrm{D}} W_k)^T}{\sqrt{n_d}}\right) \Phi_{\mathrm{D}} W_v, \tag{1}$$

where $\sigma(\cdot)$ is the SoftMax function; W_q, W_k, W_v $\in \mathbb{R}^{n_d \times n_d}$ are learnable parameters.

Moreover, let's consider integrating the preference condition C. Specifically, certain words in the description should be given different emphasis depending on varying conditions. To achieve such a target, a condition mask $\hat{\mathrm{M}}$ is utilized to filter out the relevant tokens. *Firstly*, we use the following form to compute the relevance score matrix between the description D and the condition C:

$$\mathrm{RS}_{D,C} = \Phi_{\mathrm{C}} \Phi_{\mathrm{D}}^T W_R + b_R,$$

where W_R and b_R are learnable parameters of the relevance score. *Secondly*, we perform the average pooling and then repeat n_v (first dimension of the vision encoder) times on the first dimension of $\mathrm{RS}_{D,C}$, obtaining the mask M preliminarily. *Finally*, we binarize the mask M into $\hat{\mathrm{M}}$ with the rule: elements below the relevance threshold are assigned a value of negative infinity, whereas those exceeding the threshold are set to zero:

$$\hat{\mathrm{M}}_{(i,j)} = \begin{cases} -\infty, & \mathrm{M}_{(i,j)} < \delta; \\ 0, & \mathrm{M}_{(i,j)} \geq \delta, \end{cases}$$

where δ is the relevance threshold; $i \in [1, n_v]$, $j \in [1, n_d]$ are indexes of M and $\hat{\mathrm{M}}$.

Then, enhancing the CA module shown in Equation (1) with the binarized conditional mask $\hat{\mathrm{M}}$ for fusing the preference condition C:

$$\mathrm{CA}(\Phi_{\mathrm{X}}, \Phi_{\mathrm{D}} | \Phi_{\mathrm{C}}) = \sigma\left(\frac{\Phi_{\mathrm{X}} W_q (\Phi_{\mathrm{D}} W_k)^T}{\sqrt{n_d}} + \hat{\mathrm{M}}\right) \Phi_{\mathrm{D}} W_v, \tag{2}$$

Finally, we extract the first dimension (also known as class token) of $\mathrm{CA}(\Phi_{\mathrm{X}}, \Phi_{\mathrm{D}} | \Phi_{\mathrm{C}})$ in prediction period, denoting as the $f_{\mathrm{X,D,C}}$. However, in practice, we have further observed that none of the descriptive words are associated with the condition when the description is too brief, in which the cross attention module is incapable of extracting information from the description. To address such a dilemma, we supplement the final dimension of Φ_{D}, donating as f_{D}, with the $f_{\mathrm{X,D,C}}$. Hence, the FloorPlan-MPS is calculated by:

$$\mathrm{S}(X, D|C) = \lambda_s \cdot f_{\mathrm{X,D,C}} \cdot f_{\mathrm{D}} \tag{3}$$

where λ_s is the learned scale; X, D, and C donate floor plan, description and preference condition, respectively.

4.2.2 FloorPlan-MPS Model Training

During the training phase, inputs of our objective mainly include: the FloorPlan-MPS $\mathrm{S}(X, D|C)$, two floor plans X_1 and X_2, description D, preference condition C and the preference score vector P that annotated by human. Motivated by PickScore (Kirstain et al., 2023) and MPS (Zhang et al., 2024b), we binarize preference score vector P with the following rule: P is assigned the value of (1, 0) if X_1 is preferred, (0, 1) if X_2 is preferred, and (0.5, 0.5) in the event of a tie. Moreover, we set optimization objective of the scoring model as the KL-divergence between the human annotation P and the softmax-normalized predictions $\hat{P}$:

$$\begin{aligned} \hat{P}_{i,c} &= \frac{\exp \mathrm{S}(X_i, D|C)}{\sum_{j=1}^{2} \exp \mathrm{S}(X_j, D|C)} \\ \mathcal{L}_{\text{pred}} &= \sum_{C} \sum_{i=1}^{2} P_{i,c} (\log P_{i,c} - \log \hat{P}_{i,c}) \end{aligned} \tag{4}$$

4.3 FloorPlan Autoregressive Optimization

Building upon the aforementioned work, we further consider fine-tuning floor plan autoregressive models using FloorPlan-MPS with the RLHF methodology. The popular PPO algorithm (Schulman et al., 2017), grounding in policy gradient optimization, is utilized to achieve such target. In

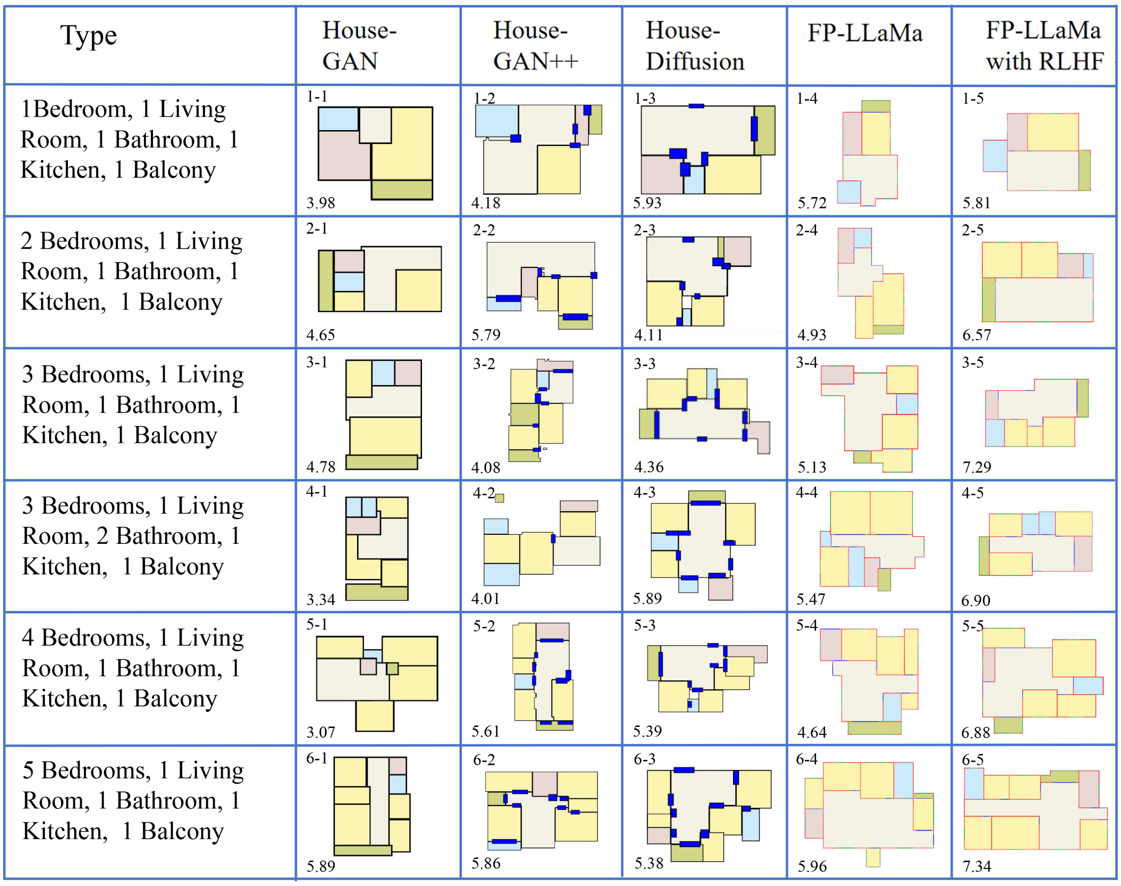


Figure 5: Qualitative results of floor plans generated by different methods on Class-Conditional task.

Algorithm 1 FloorPlan Autoregressive Optimization (FPAO)

Dataset: Description set $\mathcal{D} = \{d_1, d_2, ..., d_n\}$
Input: FloorPlan autoregressive model with pre-trained parameters θ, FloorPlan-MPS model as reward model r, reward-to-loss map function ϕ.
for $d_i \in \mathcal{D}$ **do**
  Generation: $\{x_1, \ldots, x_k\} \leftarrow$ FloorPlan ARM π_{θ_i}
  for $x_j \in \{x_1, \ldots, x_k\}$ **do**
    $\mathcal{L}_j \leftarrow \phi(r(x_j, d_i))$ // FloorPlan-MPS as loss
  end for
  $\mathcal{L}_{reward} \leftarrow \sum_{j=1}^{k} \mathcal{L}_j$ // Sum the $\mathcal{L}_j$ as FPAO loss
  $\theta_{i+1} \leftarrow \theta_i$ // Update π_{θ_i} with PPO pipeline
end for

essence, with human preferences encapsulated in FloorPlan-MPS acting as a reward function, it further guides model's policy through gradient optimization based on assigned rewards of the k generations, thereby yielding an enhanced generation policy. The pseudocode for this process is summarized in Algorithm 1.

5 Experiments

5.1 Implementation Details

All experiments are conducted on a high-performance server equipped with four NVIDIA A100 PCIe GPUs (80GB memory each), a 32-core Intel(R) Xeon(R) Gold 6326 CPU @ 2.90GHz, and 256GB of system memory. The system runs Ubuntu with CUDA 11.8 and PyTorch 2.1.x, using mixed-precision training and gradient accumulation across GPUs to optimize memory usage. We initialize the text and vision encoders, E_D and E_F, with pre-trained parameters from the CLIP model, while the remaining parameters are initialized randomly. The model is trained on our dataset for 30,000 steps, with a batch size of 128, a learning rate of 3×10^{-6}, and a 500-step warm-up.

5.2 Datasets and Evaluation Metrics

We conduct experiments on the proposed ArchiMetricsNet dataset, employing four key evaluation metrics: FID (Frechet Inception Distance), SSIM (Structural Similarity Index Measure), PSNR (Peak Signal-to-Noise Ratio) and our proposed ARS (Architectural Reasonableness Scores) which is the result of the FloorPlan-MPS model. We test two different generation tasks: class-conditional generation and text-conditional generation. The following baseline methods are selected: House-GAN (Nauata et al., 2020), HouseGAN++ (Nau-

Figure 6: Qualitative results of floor plans generated by different methods on Text-Conditional task. The detailed prompt input is listed in the Appendix G.

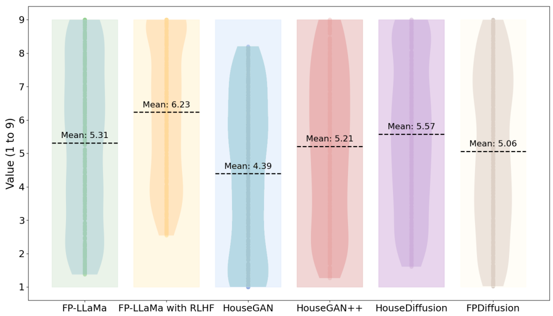


Figure 7: Average human evaluation score on the Class-Conditional task.

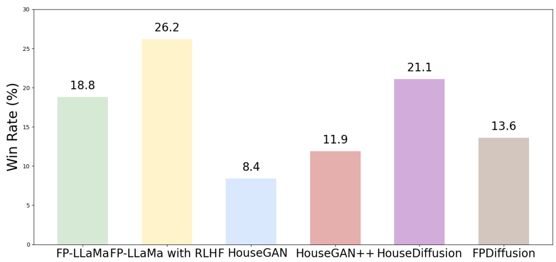


Figure 8: Human evaluation win rate on the Class-Conditional task.

ata et al., 2021), HouseDiffusion (Shabani et al., 2023), FloorPlanDiffusion (Zeng et al., 2024), Tell2Design (Leng et al., 2023), Obj-GAN (Li et al., 2019), Qwen2.5-7B-Instruct (Yang et al., 2024), ChatGLM4V (GLM et al., 2024).

5.3 Qualitative Experiments

Class-Conditional Generation: Existing state-of-the-art models such as HouseGAN, HouseGAN++, and HouseDiffusion generate layouts based on graphs, while FloorPlanDiffusion generates layouts using input room blocks. These input conditions, encompassing room counts and types, can be regarded as a specialized form of class-conditional floor plan generation. To ensure a fair comparison, we sample 500 images for each method and select examples representing the median architectural reasonableness scores for display. The color configuration is based on the scheme in FloorPlanDiffusion. Figure 5 shows that our method is capable of generating more reasonable floor plans, where the top-left corner of each image displays its ID, and the bottom-left corner shows its overall score. To assist non-architectural professionals in understanding, we have included a summary of common layout issues provided by architects (see Appendix B) and detailed expert analysis for each figure in Appendix H).

Text-Conditional Generation: Inspired by the Tell2Design, we apply large language models to the "Next-Room Sequence Prediction" task. Figure 11 presents an example of an input prompt and an output sequence, introducing a benchmark for evaluating large language models. Utilizing Llama-3.2-11B-Vision-Instruct (Touvron et al., 2023) as our backbone, we compare its post-training performance with open-source models such as ChatGLM4V (GLM et al., 2024) and Qwen2.5-7B-Instruct (Yang et al., 2024).

As shown in Figure 6, existing methods generate high-quality, high-resolution images from the traditional image generation perspective, showcasing its advancements for downstream tasks. However, from an architect's viewpoint, traditional methods often yield suboptimal designs, with issues like impractical functionality, inefficient flow, or overly small rooms. In contrast, FP-LLaMa with RLHF generates more practical layouts, featuring proper room proportions, areas, and adjacency relationships.

5.4 Quantitative Experiments

Method	FID↓	PSNR↑	SSIM↑
FP-LLaMa	**15.92**	**62.23**	**0.907**
FP-LLaMa with RLHF	<u>16.72</u>	60.03	0.892
HouseGAN	26.36	55.28	0.766
HouseGAN++	23.72	56.01	0.807
HouseDiffusion	17.71	<u>61.44</u>	0.868
FPDiffusion	18.98	60.18	<u>0.895</u>

Table 2: Results of Class-Conditional generation among different methods. The best results are **bolded**, and the second-best are <u>underlined</u>.

Method	Functionality↑	Flow↑	Overall↑
FP-LLaMa	<u>6.27</u>	5.91	6.31
FP-LLaMa with RLHF	**7.18**	**6.53**	**6.62**
HouseGAN	4.82	4.38	5.07
HouseGAN++	5.83	5.04	5.89
HouseDiffusion	5.88	<u>6.13</u>	<u>6.34</u>
FPDiffusion	6.14	6.06	5.23

Table 3: ARS of Class-Conditional generation among different methods.

Method	FID↓	PSNR↑	SSIM↑
FP-LLaMa	**15.43**	**65.87**	**0.864**
FP-LLaMa with RLHF	<u>16.73</u>	63.26	<u>0.842</u>
Tell2Design	17.72	60.33	0.825
Qwen2.5-7B-Instruct	16.94	<u>64.15</u>	0.832
ChatGLM4V	16.83	63.14	0.837
Obj-GAN	26.43	54.23	0.653

Table 4: Results of Text-Conditional generation of different methods.

Method	Functionality↑	Flow↑	Overall↑
FP-LLaMa	<u>7.08</u>	<u>7.02</u>	7.08
FP-LLaMa with RLHF	**7.30**	**7.15**	**7.25**
Tell2Design	6.82	6.73	6.95
Qwen2.5-7B-Instruct	7.02	6.90	<u>7.12</u>
ChatGLM4V	6.72	6.85	6.78
Obj-GAN	4.15	3.42	3.78

Table 5: ARS of Text-Conditional Generation of Different Methods.

For a fair comparison, we utilize the FloorPlan-Diffusion color scheme when computing the metrics for floor plans generated by different methods. Table 2 and Table 4 present the performance of different methods based on traditional metrics, and Table 3 and Table 5 evaluate their performance using Architectural Reasonableness Scores. FP-LLaMa achieves the best performance on traditional metrics. It can be observed that the performance of FP-LLaMa on traditional metrics shows a modest decline after applying RLHF. Nevertheless, as shown in Table 3, the average architectural reasonableness scores indicate that FP-LLaMa with RLHF outperforms the second-best method by 8.54% and maintains high consistency across all dimensions of architect preferences. In contrast, the scores of FP-LLaMa without RLHF are nearly on par with those of traditional methods. This suggests that further improvements in clarity and similarity to the ground truth do not necessarily lead to increased design reasonableness. The data analysis in the dataset (Section 3) partially explain this phenomenon: a portion of the original dataset includes inherently unreasonable layouts, and such biased data could act as a limiting factor for the model in generating reasonable results. We also provide further ablation study in Appendix C.

6 Human Evaluation

We invite eight Ph.D. students majoring in architecture and five professional architects that have not been involved in the previous dataset annotation, to evaluate the architectural layout rationality of 150 sets of images generated by different methods, using a scoring scale from 1 to 10. Furthermore, two expert evaluators (one male and one female) with extensive expertise in architectural residential design are appointed to supervise the evaluation process. Figure 7 shows that our method receives the highest average score from architects across all images, while Figure 8 demonstrates that the images generated by our method under equivalent and fair input conditions are the most preferred by evaluators. These findings align well with the aforementioned qualitative and quantitative analyses.

We report the variance and standard deviation of the scores for each method, as shown in Table 6. The relatively high values reflect the inherent subjectivity of human assessments—different designers may apply their own evaluation criteria—and the diverse nature of floor plan generation, where a single input may correspond to multiple valid outputs. Nevertheless, our method exhibits the lowest variance and standard deviation among all methods, indicating that the generated samples received more consistent evaluations.

Model	Variance↓	Standard Deviation↓
FP-LLaMa	4.852	2.202
FP-LLaMa+RLHF	**3.593**	**1.895**
HouseGAN	4.538	2.130
HouseGAN++	4.426	2.103
HouseDiffusion	4.263	2.064

Table 6: Variance and standard deviation in human evaluation scores.

7 Conclusion

In this paper, we seek to address the limitations of existing architectural floor plan generation and evaluation methods by introducing ArchiMetricsNet, a dataset with human feedback; and FP-LLaMa, a fine-tuned autoregressive model to align with professional architects. This work advances the integration of human expertise in architectural floor plan generation technologies and demonstrates the potential of large autoregressive generation models in the AEC domain.

8 Acknowledgements

This work is supported by the National Natural Science Foundation of China (Grant No. 52078294); Department of Education of Guangdong Province (2023ZDZX4078); Shenzhen Science and Technology Innovation Committee (WDZC20231129201240001).

9 Limitations

While our approach achieves significant improvements, there are still areas for further refinement:

1. **Architectural Codes:** In complex architectural scenarios, such as hospital layouts, compliance with intricate architectural codes remains a challenge. Future work could integrate Retrieval-Augmented Generation (RAG) and multi-agent systems to enhance code adherence.

2. **Evaluation Metrics:** ArchiMetricsNet introduces multi-dimensional evaluations to better align with architectural expertise. However, these metrics primarily reflect human preferences and do not fully capture aspects like structural feasibility or energy efficiency, which could be explored in future extensions.

3. **Dataset Scope and Diversity:** The ArchiMetricsNet dataset provides 120,000 professionally evaluated samples, primarily focusing on residential layouts. While this ensures a high-quality benchmark, expanding to commercial, industrial, and public architecture will further enhance the model's versatility and applicability across diverse design contexts.

10 Potential Risks

While it brings significant improvements, there are a few potential risks to consider:

1. **Alignment with Human Preferences:** Our model is trained with RLHF, which enhances alignment with human evaluations. However, as preferences can be subjective and context-dependent, there is a risk that the generated layouts might reflect biases inherent in the training data. Future iterations could benefit from broader and more diverse expert feedback to ensure balanced design outcomes.

2. **Interpretability of AI-Generated Designs:** While our approach improves layout rationality, AI-generated designs still require architect refinement. Providing clearer design rationales could enhance human-AI collaboration.

3. **Use in Early-Stage Design Assistance:** Designed for early-stage assistance, our method should complement rather than replace human expertise. Future work could develop interactive frameworks to strengthen human-AI collaboration.

11 Ethics Statement

The floor plans used in FP-LLaMa are sourced from publicly available datasets (RPLAN and Tell2Design) and are intended solely for research purposes. All data has been anonymized to remove personal or confidential information. Annotators involved in labeling were fairly compensated based on annotation quality and standard architectural wage rates.

References

Nagy R Bakhoum and Osamu A Wakita. 2023. The Professional Practice of Architectural Working Drawings. John Wiley & Sons.

Anastasia Chan. 2023. Gpt-3 and instructgpt: technological dystopianism, utopianism, and “contextual” perspectives in ai ethics and industry. AI and Ethics, 3(1):53–64.

Huiwen Chang, Han Zhang, Lu Jiang, Ce Liu, and William T Freeman. 2022. Maskgit: Masked generative image transformer. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 11315–11325.

Shreyas Chaudhari, Pranjal Aggarwal, Vishvak Murahari, Tanmay Rajpurohit, Ashwin Kalyan, Karthik Narasimhan, Ameet Deshpande, and Bruno Castro da Silva. 2024. Rlhf deciphered: A critical analysis of reinforcement learning from human feedback for llms. arXiv preprint arXiv:2404.08555.

Mathieu Delalandre, Ernest Valveny, Tony Pridmore, and Dimosthenis Karatzas. 2010. Generation of synthetic documents for performance evaluation of symbol recognition & spotting systems. International Journal on Document Analysis and Recognition (IJDAR), 13(3):187–207.

Patrick Esser, Robin Rombach, and Bjorn Ommer. 2021. Taming transformers for high-resolution image synthesis. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pages 12873–12883.

Team GLM, Aohan Zeng, Bin Xu, Bowen Wang, Chenhui Zhang, Da Yin, Dan Zhang, Diego Rojas, Guanyu Feng, Hanlin Zhao, et al. 2024. Chatglm: A family of large language models from glm-130b to glm-4 all tools. arXiv preprint arXiv:2406.12793.

James J Heckman and Yona Rubinstein. 2001. The importance of noncognitive skills: Lessons from the ged testing program. American economic review, 91(2):145–149.

Markus Heinonen, Ari Rantanen, Taneli Mielikäinen, Juha Kokkonen, Jari Kiuru, Raimo A Ketola, and Juho Rousu. 2008. Fid: a software for ab initio structural identification of product ions from tandem mass spectrometric data. Rapid Communications in Mass Spectrometry: An International Journal Devoted to the Rapid Dissemination of Up-to-the-Minute Research in Mass Spectrometry, 22(19):3043–3052.

Jonathan Ho and Tim Salimans. 2022. Classifier-free diffusion guidance. arXiv preprint arXiv:2207.12598.

Alain Hore and Djemel Ziou. 2010. Image quality metrics: Psnr vs. ssim. In 2010 20th international conference on pattern recognition, pages 2366–2369. IEEE.

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):118–1.

Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros. 2017. Image-to-image translation with conditional adversarial networks. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 1125–1134.

Ahti Kalervo, Juha Ylioinas, Markus Häikiö, Antti Karhu, and Juho Kannala. 2019. Cubicasa5k: A dataset and an improved multi-task model for floorplan image analysis. In Image Analysis: 21st Scandinavian Conference, SCIA 2019, Norrköping, Sweden, June 11–13, 2019, Proceedings 21, pages 28–40. Springer.

Yuval Kirstain, Adam Polyak, Uriel Singer, Shahbuland Matiana, Joe Penna, and Omer Levy. 2023. Pick-a-pic: An open dataset of user preferences for text-to-image generation. Advances in Neural Information Processing Systems, 36:36652–36663.

Sicong Leng, Yang Zhou, Mohammed Haroon Dupty, Wee Sun Lee, Sam Conrad Joyce, and Wei Lu. 2023. Tell2design: A dataset for language-guided floor plan generation. arXiv preprint arXiv:2311.15941.

Wenbo Li, Pengchuan Zhang, Lei Zhang, Qiuyuan Huang, Xiaodong He, Siwei Lyu, and Jianfeng Gao. 2019. Object-driven text-to-image synthesis via adversarial training. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 12174–12182.

Zhengda Lu, Teng Wang, Jianwei Guo, Weiliang Meng, Jun Xiao, Wei Zhang, and Xiaopeng Zhang. 2021. Data-driven floor plan understanding in rural residential buildings via deep recognition. Information Sciences, 567:58–74.

Nelson Nauata, Kai-Hung Chang, Chin-Yi Cheng, Greg Mori, and Yasutaka Furukawa. 2020. House-gan: Relational generative adversarial networks for graph-constrained house layout generation. In Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part I 16, pages 162–177. Springer.

Nelson Nauata, Sepidehsadat Hosseini, Kai-Hung Chang, Hang Chu, Chin-Yi Cheng, and Yasutaka Furukawa. 2021. House-gan++: Generative adversarial layout refinement network towards intelligent computational agent for professional architects. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 13632–13641.

Long Ouyang, Jeffrey Wu, Xu Jiang, Diogo Almeida, Carroll Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, et al. 2022. Training language models to follow instructions with human feedback. Advances in neural information processing systems, 35:27730–27744.

Joern Ploennigs and Markus Berger. 2023. Diffusion models for computational design at the example of floor plans. arXiv preprint arXiv:2307.02511.

Sizhong Qin, Chengyu He, Qiaoyun Chen, Sen Yang, Wenjie Liao, Yi Gu, and Xinzheng Lu. 2024. Chathousediffusion: Prompt-guided generation and editing of floor plans. arXiv preprint arXiv:2410.11908.

Aditya Ramesh, Mikhail Pavlov, Gabriel Goh, Scott Gray, Chelsea Voss, Alec Radford, Mark Chen, and Ilya Sutskever. 2021. Zero-shot text-to-image generation. In International conference on machine learning, pages 8821–8831. Pmlr.

Ali Razavi, Aaron Van den Oord, and Oriol Vinyals. 2019. Generating diverse high-fidelity images with vq-vae-2. Advances in neural information processing systems, 32.

John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. 2017. Proximal policy optimization algorithms. arXiv preprint arXiv:1707.06347.

Mohammad Amin Shabani, Sepidehsadat Hosseini, and Yasutaka Furukawa. 2023. Housediffusion: Vector floorplan generation via a diffusion model with discrete and continuous denoising. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 5466–5475.

Haoyuan Sun, Bo Xia, Yongzhe Chang, and Xueqian Wang. 2025a. Generalizing alignment paradigm of text-to-image generation with preferences through f-divergence minimization. In Proceedings of the AAAI Conference on Artificial Intelligence, volume 39, pages 27644–27652.

Haoyuan Sun, Bo Xia, Yifei Zhao, Yongzhe Chang, and Xueqian Wang. 2025b. Identical human preference alignment paradigm for text-to-image models. In ICASSP 2025-2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pages 1–5. IEEE.

Haoyuan Sun, Bo Xia, Yifei Zhao, Yongzhe Chang, and Xueqian Wang. 2025c. Positive enhanced preference alignment for text-to-image models. In ICASSP 2025-2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pages 1–5. IEEE.

Peize Sun, Yi Jiang, Shoufa Chen, Shilong Zhang, Bingyue Peng, Ping Luo, and Zehuan Yuan. 2024. Autoregressive model beats diffusion: Llama for scalable image generation. arXiv preprint arXiv:2406.06525.

Keyu Tian, Yi Jiang, Zehuan Yuan, BINGYUE PENG, and Liwei Wang. 2024. Visual autoregressive modeling: Scalable image generation via next-scale prediction. In The Thirty-eighth Annual Conference on Neural Information Processing Systems.

Hugo Touvron, Louis Martin, Kevin Stone, Peter Albert, Amjad Almahairi, Yasmine Babaei, Nikolay Bashlykov, Soumya Batra, Prajjwal Bhargava, Shruti Bhosale, et al. 2023. Llama 2: Open foundation and fine-tuned chat models. arXiv preprint arXiv:2307.09288.

Jianhui Wang, Yangfan He, Kun Li, Sida Li, Lan Zhao, Jun Yin, Miao Zhang, Tianyu Shi, and Xueqian Wang. 2025. Mdanet: A multi-stage domain adaptation framework for generalizable low-light image enhancement. Neurocomputing, page 129572.

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):1–12.

Xiaoshi Wu, Yiming Hao, Keqiang Sun, Yixiong Chen, Feng Zhu, Rui Zhao, and Hongsheng Li. 2023a. Human preference score v2: A solid benchmark for evaluating human preferences of text-to-image synthesis. arXiv preprint arXiv:2306.09341.

Xiaoshi Wu, Keqiang Sun, Feng Zhu, Rui Zhao, and Hongsheng Li. 2023b. Better aligning text-to-image models with human preference. arXiv preprint arXiv:2303.14420, 1(3).

Jiazheng Xu, Xiao Liu, Yuchen Wu, Yuxuan Tong, Qinkai Li, Ming Ding, Jie Tang, and Yuxiao Dong. 2023. Imagereward: learning and evaluating human preferences for text-to-image generation. In Proceedings of the 37th International Conference on Neural Information Processing Systems, pages 15903–15935.

An Yang, Baosong Yang, Beichen Zhang, Binyuan Hui, Bo Zheng, Bowen Yu, Chengyuan Li, Dayiheng Liu, Fei Huang, Haoran Wei, Huan Lin, Jian Yang, Jianhong Tu, Jianwei Zhang, Jianxin Yang, Jiaxi Yang, Jingren Zhou, Junyang Lin, Kai Dang, Keming Lu, Keqin Bao, Kexin Yang, Le Yu, Mei Li, Mingfeng Xue, Pei Zhang, Qin Zhu, Rui Men, Runji Lin, Tianhao Li, Tingyu Xia, Xingzhang Ren, Xuancheng Ren, Yang Fan, Yang Su, Yichang Zhang, Yu Wan, Yuqiong Liu, Zeyu Cui, Zhenru Zhang, and Zihan Qiu. 2024. Qwen2.5 technical report. arXiv preprint arXiv:2412.15115.

Jun Yin, Wen Gao, Jizhizi Li, Pengjian Xu, Chenglin Wu, Borong Lin, and Shuai Lu. 2025. Archidiff: Interactive design of 3d architectural forms generated from a single image. Computers in Industry, 168:104275.

JUN YIN, PENGJIAN XU, WEN GAO, PENGYU ZENG, and SHUAI LU. Drag2build: Interactive point-based manipulation of 3d architectural point clouds generated from a single im-age.

Jiahui Yu, Xin Li, Jing Yu Koh, Han Zhang, Ruoming Pang, James Qin, Alexander Ku, Yuanzhong Xu, Jason Baldridge, and Yonghui Wu. 2021. Vector-quantized image modeling with improved vqgan. arXiv preprint arXiv:2110.04627.

Lijun Yu, Yong Cheng, Kihyuk Sohn, José Lezama, Han Zhang, Huiwen Chang, Alexander G Hauptmann, Ming-Hsuan Yang, Yuan Hao, Irfan Essa, et al. 2023. Magvit: Masked generative video transformer. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 10459–10469.

Pengyu Zeng, Wen Gao, Jizhizi Li, Jun Yin, Jiling Chen, and Shuai Lu. 2025a. Automated residential layout generation and editing using natural language and images. Automation in Construction, 174:106133.

Pengyu Zeng, Wen Gao, Jun Yin, Pengjian Xu, and Shuai Lu. 2024. Residential floor plans: Multi-conditional automatic generation using diffusion models. Automation in Construction, 162:105374.

Pengyu Zeng, Maowei Jiang, Zihang Wang, Jizhizi Li, Jun Yin, and Shuai Lu. Card: Cross-modal agent framework for generative and editable residential design. In NeurIPS 2024 Workshop on Open-World Agents.

Pengyu Zeng, Jun Yin, Yan Gao, Jizhizi Li, Zhanxiang Jin, and Shuai Lu. 2025b. Comprehensive and dedicated metrics for evaluating ai-generated residential floor plans. Buildings, 15(10):1674.

Miao Zhang, Yiqing Shen, Jun Yin, Shuai Lu, and Xueqian Wang. 2024a. Adagent: Anomaly detection agent with multimodal large models in adverse environments. IEEE Access.

Miao Zhang, Jun Yin, Pengyu Zeng, Yiqing Shen, Shuai Lu, and Xueqian Wang. 2025. Tscnet: A text-driven semantic-level controllable framework for customized low-light image enhancement. Neurocomputing, page 129509.

Richard Zhang, Phillip Isola, Alexei A Efros, Eli Shechtman, and Oliver Wang. 2018. The unreasonable effectiveness of deep features as a perceptual metric. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 586–595.

Sixian Zhang, Bohan Wang, Junqiang Wu, Yan Li, Tingting Gao, Di Zhang, and Zhongyuan Wang. 2024b. Learning multi-dimensional human preference for text-to-image generation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 8018–8027.

Ziyang Zong, Guanying Chen, Zhaohuan Zhan, Fengcheng Yu, and Guang Tan. 2025. Housetune: Two-stage floorplan generation with llm assistance. Preprint, arXiv:2411.12279.

A Details on ArchiMetricsNet Annotation Pipeline

We hired professional annotators to evaluate building image pairs. The evaluation was based on two sub-dimensions (functionality and flow design) and an overall dimension (overall score), defined as follows: Functionality, scored from 1 to 10, requires annotators to assess the rationality of the functional layout, including whether the number of rooms meets requirements, whether room usage distribution is reasonable, and whether the relationships between functional areas are scientifically arranged to meet daily or commercial needs.

flow design, also scored from 1 to 10, involves reviewing the quality of flow design. Annotators focus on whether pedestrian flow is efficient, unnecessary crossings or congestions are minimized, and privacy and convenience are considered, such as positioning the master bedroom far from public areas or situating the kitchen near the dining area.

The overall score, ranging from 1 to 10, combines the functionality and flow design scores with the annotators' subjective preferences. It reflects the overall quality of the design, considering factors like innovation, project-specific requirements, and the balance between aesthetics and practicality.

For text descriptions, annotators provide concise summaries of each building plan's strengths and weaknesses. Strengths highlight notable advantages in functionality, flow design, or innovation. Weaknesses identify issues such as unreasonable functional configurations or complex flow design.

Each building plan was scored on the three dimensions on a scale of 1 to 10, with scores normalized to [0, 1]. This annotation task was completed by 32 experts in architectural design. Before formal annotation, all members underwent pre-annotation training, and those failing to meet consistency standards were disqualified. In the end, 29 experts participated, with 25 responsible for annotation and 4 for quality checking. Each image was independently rated by 5 annotators, and the final score was the average of the 5 ratings. When annotators indicate that they are unable to provide an appropriate score based on their own experience and understanding, they are encouraged to skip the current item, allowing another annotator to complete it. Thirty percent of the annotations were reviewed by quality checkers, and data with significant discrepancies were re-annotated.

B Cases of Unreasonable Room Layouts

Based on the results of analysis and discussions with architects, we have identified and summarized six typical cases of unreasonable room layouts in the figure 13, accompanied by simplified explana-

Figure 9: The 165 categories of floor plans.

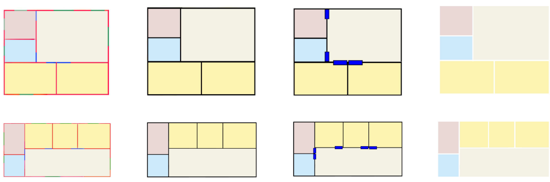


Figure 10: Different floor plan representation configurations in the dataset.

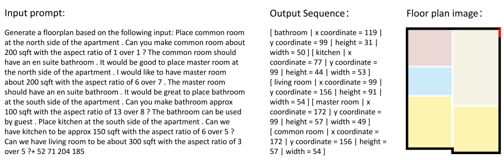


Figure 11: Example input and output in the Text-Conditional planar generation task.

tions. Generally, the first three issues primarily concern functionality, while the latter three pertain to flow design:

The Room layout outline is unreasonable. As shown in the first row of the figure 13, if the room's layout outline contains excessive indentations or protrusions, it may deviate significantly from practical or aesthetic standards, leading to inefficient space utilization and unnecessary economic costs. No one would want their home to have excessively irregular walls, making even the act of hanging a painting awkward.

The Room shape proportion is unreasonable. As illustrated in the second row of the figure 13, unbalanced proportions can result in rooms appearing overly narrow or excessively spacious. For spaces like bedrooms, such imbalances can directly affect the psychological comfort of inhabitants. In architectural design practices, the recommended length-to-width ratio for rooms is generally between 0.5 and 1.

The Room's area is inappropriate. The area of a room may exceed or fall short of what is required for its intended function. As shown in the third row, if a bedroom cannot even accommodate a bed, how can it fulfill the role of a bedroom? Such issues are particularly common in layouts with multiple bedrooms or bathrooms, as evidenced by the examples in the third row.

The Number of rooms is unreasonable. An excessive or insufficient number of rooms can lead to inefficient space utilization or wasted resources. As shown in the fourth row, why would a house without a living room have such a large kitchen and three bathrooms? Less severe cases might include a lack of essential rooms, such as no bathroom or an absent kitchen.

The Room adjacency is unreasonable. Poor adjacency relationships between rooms can disrupt accessibility and logical workflows. For instance, as depicted in the fifth row, having two bathrooms adjacent to one another within a house is unreasonable. Typically, when architects design two bathrooms, one is positioned near the living room, while the other is placed near a bedroom for logical accessibility.

The Room hierarchy is unreasonable. Inefficient nesting relationships can result in confusing or impractical spatial hierarchies. As illustrated in the sixth row, some layouts feature only a sin-

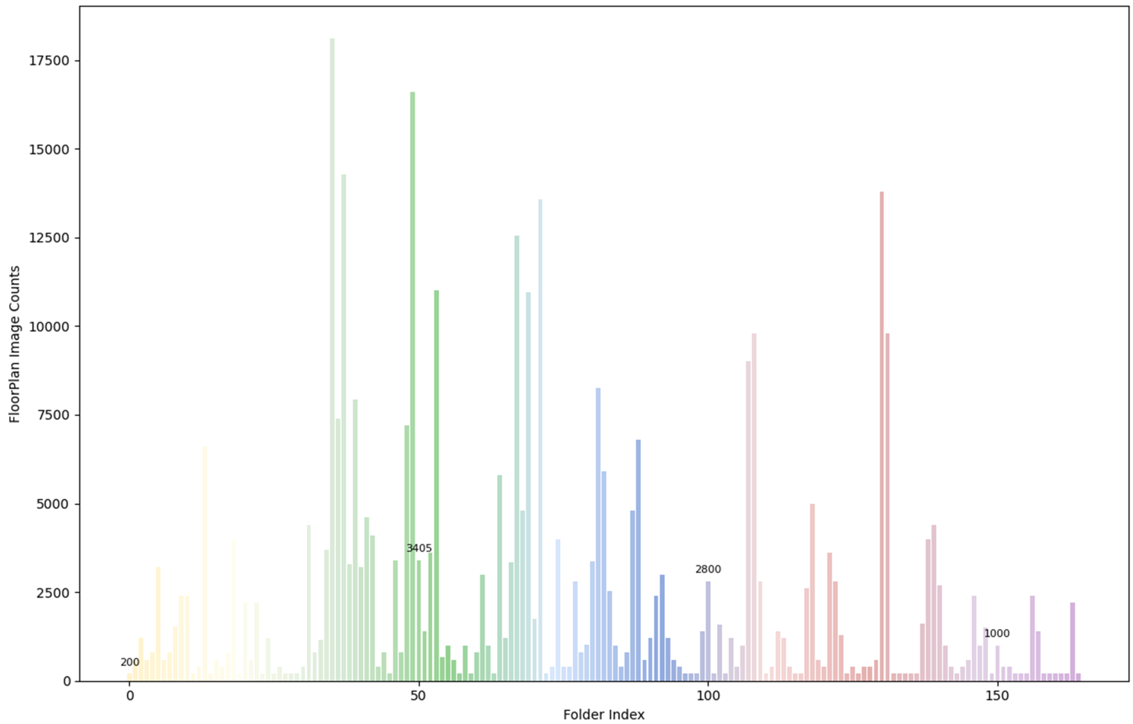


Figure 12: The number of floor plans in each category.

gle flow path, requiring one to sequentially pass through the living room, kitchen, and bathroom to reach the bedroom. For example, if a house contains two bathrooms, it is reasonable to place one in the master bedroom and the other in the living room. However, if a house has only one bathroom, and it is located within the master bedroom, it becomes problematic as it compromises the privacy of the master bedroom when accessed by other occupants.

C Ablation Study

ID	Module		Metrics		
	CA	Mask	Functionality	Flow	Overall
1	×	×	6.05	6.12	6.08
2	✓	×	6.43	6.18	6.32
3	✓	✓	**6.72**	**6.37**	**6.59**

Table 7: **Ablation study for different components in FloorPlan-MPS.** CA refers to the cross-attention module, and Mask represents the preference condition mask $\hat{\mathrm{M}}$

We conducted ablation studies to verify the effectiveness of each component in FloorPlan-MPS. As shown in the table, the cross-attention module and the mask $\hat{\mathrm{M}}$ contribute to improving the model's performance in floor plan preference.

D Experimental Settings

We empirically tuned the relevance threshold for mask generation via grid search on a held-out validation subset and selected 0.3 as the final value.

Regarding the reconstruction quality of the VQ-VAE decoder, we leveraged well-established parameter and model configurations that have previously demonstrated strong performance on large-scale natural image datasets such as ImageNet (Sun et al., 2024). For our dataset, we trained the model using the same settings and evaluated reconstruction fidelity on a held-out validation set of 1,000 floor plans. Specifically, we used relative FID (rFID), SSIM, and PSNR as evaluation metrics. With a codebook size of 16,384 and a code vector dimension of 8, the model achieved the following results:

Metric	Value
FID (↓)	8.83
SSIM (↑)	0.786
PSNR (↑)	24.72

Table 8: Reconstruction quality of the VQ-VAE decoder on the validation set.

These results indicate that our VQ-VAE encoder provides sufficient compression while preserving both perceptual and structural quality, thereby establishing a robust foundation for the downstream autoregressive modeling.

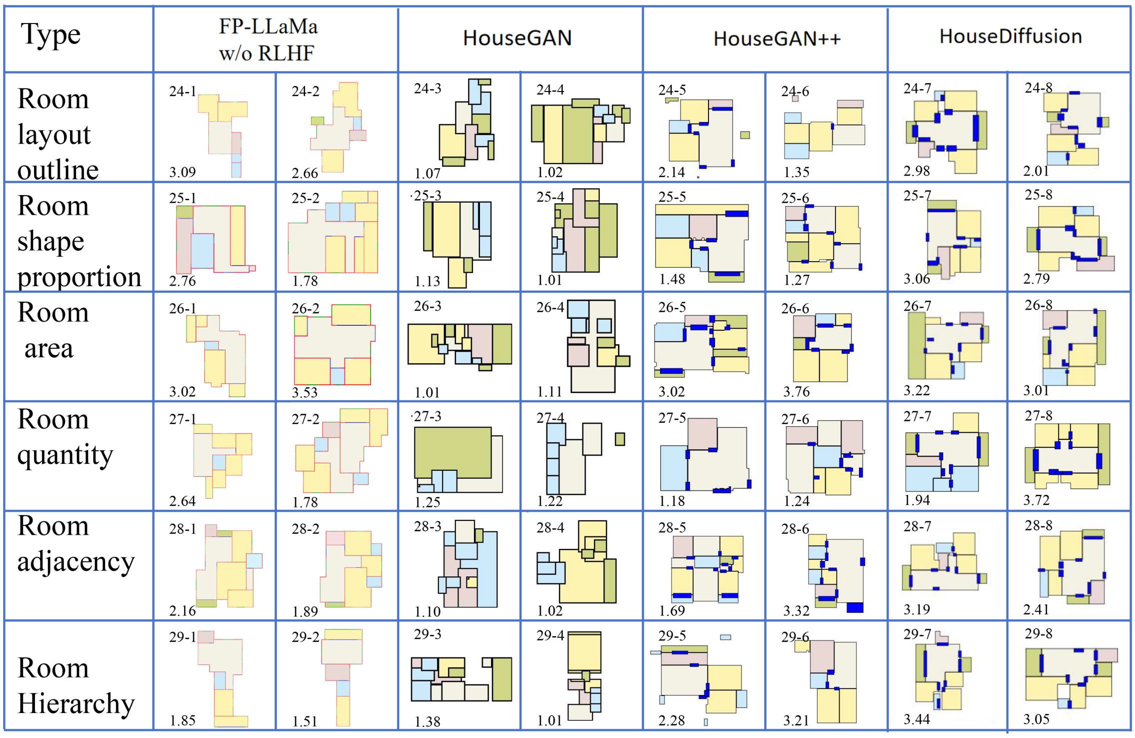


Figure 13: Six typical cases of unreasonable room layouts

E Future Work

FP-LLaMa has potential to integrate seamlessly with workflows like RAG(Retrieval Augmented Generation) and multi-agent systems, enabling adherence to building codes and real-world team collaboration. Figure 14 below presents a possible framework based on FP-LLaMa. By fine-tuning large language models with the ArchiMetricsNet dataset, we can develop FloorPlanChat, which combines the ability to comprehend architectural plans with the rich prior LLM. Identifying shortcomings in generated designs helps refine instructions to produce optimized plans.

F Discussion

Why don't we standardize the model's output solely through predefined rules? Residential design is inherently a complex and iterative process, characterized by the absence of a universally optimal solution. Moreover, the requirements for residential spaces vary significantly across regions and individuals, rendering the reliance on rigid rules insufficient. The imposition of overly stringent constraints may also limit the diversity and creativity of generated outputs, reducing the model's adaptability across diverse contexts.

In line with this, we deliberately avoid rigid hierarchical structures during dataset preprocessing. Instead of organizing layouts via function-driven graph hierarchies—as adopted in models like HouseGAN and HouseGAN++—we partition the dataset into 165 categories based on the number of rooms. While hierarchical structures may enhance input controllability by encoding spatial and functional relationships, they often demand higher user expertise and introduce additional constraints into the generation process. From a different perspective, we treat floor plans as a special type of natural image. This perspective allows us to apply general-purpose generative models and paves the way for unifying layout generation with broader image generation tasks in the future.

Overall, these design choices reflect our broader goal: to embrace the diversity and ambiguity of real-world design scenarios, enabling more flexible and human-centered generative tools.

G The Detailed Prompt for the Text-Conditional FloorPlan Generation.

In this section, we provide a detailed list of the text descriptions for each floor plan in the text-conditional floor plan generation task. We use the numbers in the upper-left corner of the images for

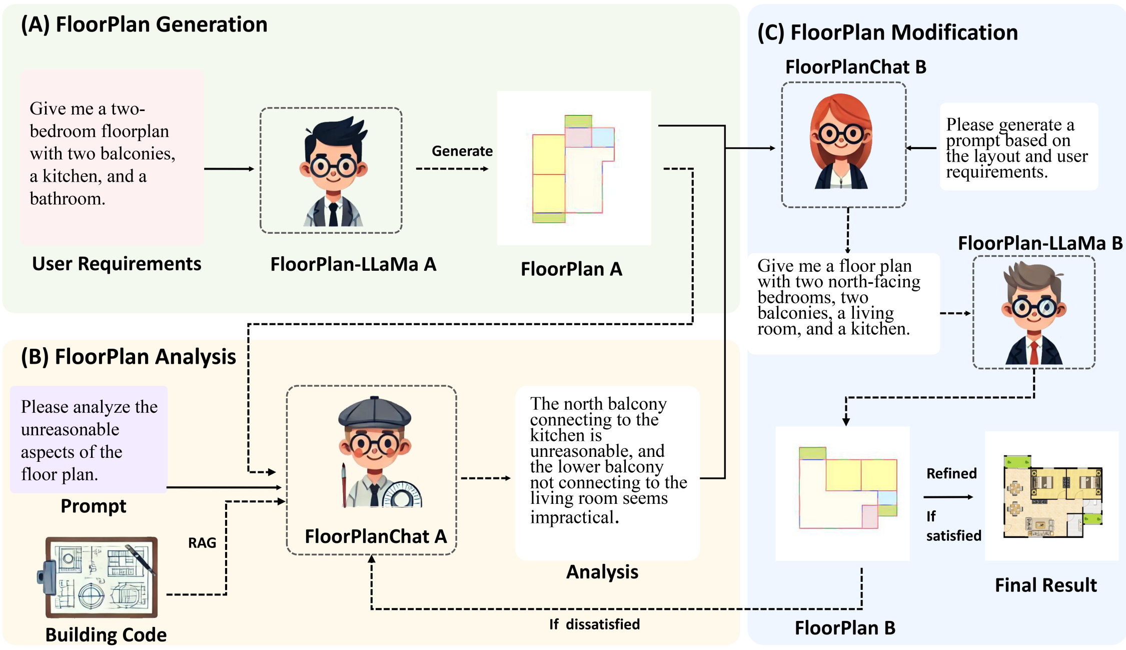


Figure 14: By coordinating multiple FP-LLaMas and FloorPlanChats, the process simulates the interaction between a human architect and a user, iteratively modifying the design until it meets satisfaction.

correspondence. **Row 16:** Generate a floorplan based on the following input: Make living room around 350 sqft with the aspect ratio of 15 over 7 . Place master room at the north side of the apartment . I would like to have master room about 300 sqft with the aspect ratio of 12 over 11 . The master room should have an en suite bathroom . Place balcony at the north side of the apartment . Make balcony about 100 sqft with the aspect ratio of 13 over 3 . The balcony is private . Place kitchen at the south side of the apartment . Make kitchen approx 100 sqft with the aspect ratio of 14 over 11 . The bathroom should be at the south east corner of the apartment . Can we have bathroom to be around 100 sqft with the aspect ratio of 16 over 15 ? The bathroom can be used by guest .+ 35 72 221 184 - 90 233 174 149 - 34 103 66 70

Row 17: Generate a floorplan based on the following input: It would be good to place balcony at the north side of the apartment . I would like to have balcony approx 50 sqft with the aspect ratio of 3 over 1 . Place master room at the north east corner of the apartment . The master room should be around 200 sqft with the aspect ratio of 10 over 13 . The master room should have an en suite bathroom . I would like to place common room at the north west corner of the apartment . Can we have common room to be about 150 sqft with the aspect ratio of 7 over 11 ? The common room should have an en suite bathroom . I would like to place kitchen at the south west corner of the apartment . Make kitchen about 50 sqft with the aspect ratio of 13 over 7 . Make living room approx 450 sqft with the aspect ratio of 10 over 11 . The bathroom should be at the south east corner of the apartment . Can we have bathroom to be around 50 sqft with the aspect ratio of 5 over 3 ? The bathroom can be used by guest .+ 72 46 184 210 - 164 217 177 204 - 164 60 178 46 - 162 190 201 151 - 162 88 197 54 - 96 86 136 45

Row 18: Generate a floorplan based on the following input: I would like to place kitchen at the south west corner of the apartment . Can we have kitchen to be about 100 sqft with the aspect ratio of 4 over 11 ? It would be great to have living room around 600 sqft with the aspect ratio of 13 over 16 . The common room should be at the south side of the apartment . Can we have common room to be about 150 sqft with the aspect ratio of 13 over 15 ? The common room should have an en suite bathroom . The bathroom should be at the north side of the apartment . Make bathroom about 100 sqft with the aspect ratio of 11 over 15 . The bathroom can be used by guest . It would be good to place master room at the north side of the apartment . The master room should be about 200 sqft with the aspect ratio of 9 over 14 . The master room should have an en suite bathroom .+ 36 65 220 191 - 208

192 221 178 - 148 108 190 66 - 147 147 218 76 - 126 84 146 64 - 35 175 68 142

Row 19: Generate a floorplan based on the following input: The balcony should be at the north side of the apartment . The balcony should be approx 100 sqft with the aspect ratio of 11 over 4 . It would be great to place master room at the north side of the apartment . The master room should be around 250 sqft with the aspect ratio of 10 over 11 . The master room should have an en suite bathroom . Place bathroom 1 at the north side of the apartment . It would be great to have bathroom 1 approx 100 sqft with the aspect ratio of 15 over 8 . The bathroom 1 can be used by guest . It would be great to place common room at the south west corner of the apartment . Can you make common room approx 200 sqft with the aspect ratio of 10 over 13 ? The common room should have an en suite bathroom . Can we have living room to be about 900 sqft with the aspect ratio of 5 over 7 ? It would be good to place bathroom 2 at the south side of the apartment . Can we have bathroom 2 to be around 100 sqft with the aspect ratio of 5 over 6 ? The bathroom 2 can be used by guest .+ 43 49 213 207 - 110 116 179 48

Row 20: Generate a floorplan based on the following input: It would be good to place kitchen at the south side of the apartment . Can you make kitchen about 50 sqft with the aspect ratio of 13 over 10 ? I would like to place bathroom at the south west corner of the apartment . Make bathroom around 50 sqft with the aspect ratio of 3 over 7 . The bathroom can be used by guest . Make living room around 500 sqft with the aspect ratio of 2 over 5 . It would be good to place common room 1 at the south side of the apartment . Make common room 1 about 100 sqft with the aspect ratio of 3 over 4 . The common room 1 should have an en suite bathroom . Can we have common room 2 to be at south side . The common room 2 should be around 100 sqft with the aspect ratio of 16 over 13 . The common room 2 should have an en suite bathroom . I would like to place balcony at the north side of the apartment . Can we have balcony to be about 50 sqft with the aspect ratio of 15 over 7 ? Can we have master room to be at north side . Can we have master room to be approx 150 sqft with the aspect ratio of 11 over 16 ? The master room should have an en suite bathroom .+ 33 59 223 197 - 214 76 232 58 - 210 125 250 86 - 32 217 76 174 - 89 137 168 58

Row 21: Generate a floorplan based on the following input: I would like to place common room at the south side of the apartment . Make common room around 150 sqft with the aspect ratio of 11 over 14 . The common room should have an en suite bathroom . The balcony should be at the south side of the apartment . Make balcony approx 50 sqft with the aspect ratio of 5 over 2 . The bathroom should be at the north side of the apartment . Can we have bathroom to be around 50 sqft with the aspect ratio of 9 over 7 ? The bathroom can be used by guest . Can you make living room about 500 sqft with the aspect ratio of 5 over 12 ? It would be great to place master room at the north side of the apartment . Can you make master room about 200 sqft with the aspect ratio of 3 over 2 ? The master room should have an en suite bathroom . Place kitchen at the north side of the apartment . The kitchen should be around 50 sqft with the aspect ratio of 11 over 6 .+ 44 59 212 197 - 66 120 128 58 - 43 128 90 82 - 44 82 68 59

Row 22: Generate a floorplan based on the following input: Place common room 1 at the south side of the apartment . The common room 1 should be around 200 sqft with the aspect ratio of 4 over 5 . The common room 1 should have an en suite bathroom . I would like to place master room at the south side of the apartment . Can we have master room to be approx 250 sqft with the aspect ratio of 15 over 16 ? The master room should have an en suite bathroom . I would like to place common room 2 at the north side of the apartment . It would be good to have common room 2 about 100 sqft with the aspect ratio of 9 over 8 . The common room 2 should have an en suite bathroom . Place kitchen at the north side of the apartment . Make kitchen around 150 sqft with the aspect ratio of 8 over 5 . The bathroom should be at the north east corner of the apartment . Make bathroom approx 100 sqft with the aspect ratio of 4 over 5 . The bathroom is in an en suite bathroom . It would be good to have living room approx 350 sqft with the aspect ratio of 9 over 4 .+ 46 52 210 204 - 45 247 125 167

Row 23: Generate a floorplan based on the following input: It would be great to place balcony at the south side of the apartment . It would be great to have balcony approx 100 sqft with the aspect ratio of 1 over 1 . Can you make living room about 700 sqft with the aspect ratio of 9 over 11 ? Place master room at the south east corner of the apartment . The master room should be about 300 sqft with the aspect ratio of 11 over 12 . The master room should have an en suite bathroom . Place entrance at the north side of the apartment . Make entrance approx 100 sqft with the aspect ratio of 1 over 1 . Can we have common room 1 to be at south side . Can you make common room 1 around 100 sqft with the aspect ratio of 16 over 15 ? The common room 1 should have an en suite bathroom . Place common room 2 at the north east corner of the apartment . The common room 2 should be about 150 sqft with the aspect ratio of 11 over 8 . The common room 2 should have an en suite bathroom . Place bathroom at the north east corner of the apartment . Make bathroom around 50 sqft with the aspect ratio of 13 over 8 . The bathroom can be used by guest . I would like to place kitchen at the north side of the apartment . Make kitchen about 50 sqft with the aspect ratio of 8 over 9 .+ 27 51 229 205 - 190 214 221 183 - 188 191 223 156 - 188 88 226 50 - 169 230 204 194 - 27 186 162 51 - 86 241 135 192 - 26 90 66 50 - 26 204 60 170

H Expert Evaluation of Each Qualitative Image.

To help non-architectural professionals understand the rationale and flaws of each design presented in our paper, we invited three architects with doctoral degrees in architecture to jointly conduct the following evaluation.

Row 1: 1-1 The bedroom on the right side of the plane has an unreasonable L-shape, and the living room area is too small. 1-2 The length-to-width ratio of the kitchen room is greater than 2:1. 1-3 The area and proportion of each room are reasonable, the functional zoning and flow lines are reasonable, and the overall contour of the plane is regular and reasonable. 1-4 The length-to-width ratio of the kitchen room is greater than 2:1, and the shape of the living room is slightly irregular. 1-5 The area and proportion of each room are reasonable, the functional zoning and flow lines are reasonable, and the overall contour of the plane is regular and reasonable?

Row 2: 2-1 The area and proportion of each room are reasonable, the functional zoning and flow lines are reasonable, and the overall contour of the plane is regular and reasonable. 2-2 The length-to-width ratio of the bathroom is greater than 2:1, and the overall contour of the plane is irregular. 2-3 The bathroom is too small, and it cannot be directly accessed from the living room. 2-4 The overall shape is unreasonable, and the length-to-width ratio of the bedroom is greater than 2:1. 2-5 The area and proportion of each room are reasonable, the functional zoning and flow lines are reasonable, and the overall contour of the plane is regular and reasonable.

Row 3: 3-1 The bedroom room at the bottom of the plane is too large. 3-2 The bedroom at the bottom left corner of the plane cannot be directly accessed from the living room, the bathroom is too small, and the length-to-width ratio of the kitchen room is greater than 2:1. 3-3 The kitchen has an unreasonable L-shape, and the length-to-width ratio of the bathroom room is greater than 2:1. 3-4 The bedroom room at the bottom right corner has an unreasonable L-shape and the balcony area is too small. 3-5 The area and proportion of each room are reasonable, the functional zoning and flow lines are reasonable, and the overall contour of the plane is relatively reasonable.

Row 4: 4-1 The bedroom at the bottom left corner has an unreasonable L-shape, and the bathroom cannot be directly accessed from the living room. 4-2 The bathroom and bedroom on the left side of the floor plan cannot be directly accessed from the living room. The kitchen in the top-right corner of the floor plan cannot be directly accessed from the living room, and the kitchen's length-to-width ratio is greater than 2:1. 4-3 The area and proportion of each room are reasonable, the functional zoning and flow lines are reasonable, and the overall contour of the plane is relatively regular and reasonable. 4-4 The bathroom on the left side of the plane is too small, and the length-to-width ratio of the bathroom room at the bottom is greater than 2:1. 4-5 The area and proportion of each room are reasonable, the functional zoning and flow lines are reasonable, and the overall contour of the plane is relatively reasonable.

Row 5: 5-1 The kitchen, bathroom, and balcony are too small, and the bathroom has an unreasonable L-shape, and the balcony should have at least one side exposed to the outside environment. 5-2 The area and proportion of each room are reasonable, the functional zoning and flow lines are reasonable, and the overall contour of the plane is relatively reasonable. 5-3 The kitchen room at the top right corner has an unreasonable L-shape, and the bathroom cannot be directly accessed through the living room. 5-4 The length-to-width ratio of the bedroom at the top right corner is greater than 2:1, and the bedroom at the bottom right corner is too small. 5-5 The area and proportion of each room are reasonable, the functional zoning and flow lines are reasonable, and the overall contour of the plane is relatively reasonable.

Row 6: 6-1 The area and proportion of each room are reasonable, the functional zoning and flow lines are reasonable, and the overall contour of the plane is relatively reasonable. 6-2 The balcony area is too small, and the length-to-width ratio of the bathroom room is greater than 2:1, and the room in the middle of the right side has a length-to-width ratio greater than 2:1. 6-3 The overall contour of the plane is irregular, and the bathroom cannot be directly accessed from the living room. 6-4 The bedroom at the bottom of the plane is too small, and the overall contour of the plane is relatively unreasonable. 6-5 The area and proportion of each room are reasonable, the functional zoning and flow lines are reasonable, and the overall contour of the plane is relatively reasonable.

Row 7: 7-1 The kitchen and balcony areas are too small, and the bedroom at the bottom has an irregular shape. 7-2 The overall contour of the plane is unreasonable. 7-3 The area and proportion of each room are reasonable, the functional zoning and flow lines are reasonable, and the overall contour of the plane is relatively reasonable. 7-4 The overall contour is relatively unreasonable. 7-5 The area and proportion of each room are reasonable, the functional zoning and flow lines are reasonable, and the overall contour of the plane is regular and reasonable?

Row 8: 8-1 The length-to-width ratio of the bathroom room is greater than 2:1. 8-2 The length-to-width ratio of the kitchen room is greater than 2:1, and the bedroom at the top right corner has an irregular shape. 8-3 The area and proportion of each room are reasonable, the functional zoning and flow lines are reasonable, and the overall contour of the plane is regular and reasonable. 8-4 The length-to-width ratio of both the kitchen and bathroom rooms is greater than 2:1, and the overall contour of the plane is slightly unreasonable. 8-5 The area and proportion of each room are reasonable, the functional zoning and flow lines are reasonable, and the overall contour of the plane is regular and reasonable?

Row 9: 9-1 The balcony area is too small, and the kitchen and bedroom at the top cannot be directly accessed from the entrance hall. 9-2 The area and proportion of each room are reasonable, the functional zoning and flow lines are reasonable, and the overall contour of the plane is regular and reasonable. 9-3 The bedroom at the top right corner has an unreasonable L-shape. 9-4 The overall contour of the plane is slightly unreasonable. 9-5 The area and proportion of each room are reasonable, the functional zoning and flow lines are reasonable, and the overall contour of the plane is regular and reasonable?

Row 10: 10-1 The length-to-width ratio of the bedroom at the top right corner is greater than 2:1, and the bathroom on the left cannot be directly accessed from the living room. 10-2 The length-to-width ratio of the kitchen room is greater than 2:1. 10-3 The length-to-width ratio of the kitchen room is greater than 2:1, and the overall contour is irregular. 10-4 The irregular overall contour is not economical. 10-5 The area and proportion of each room are reasonable, the functional zoning and flow lines are reasonable, and the overall contour of the plane is regular and reasonable.

Row 11: 11-1 The kitchen area is too small, and the bedroom at the bottom left corner has an irregular shape. 11-2 The area and proportion of each room are reasonable, the functional zoning and flow lines are reasonable, and the overall contour of the plane is regular and reasonable. 11-3 The overall contour of the plane is slightly irregular. 11-4 The length-to-width ratio of the kitchen room is greater than 2:1. 11-5 The area and proportion of each room are reasonable, the functional zoning and flow lines are reasonable, and the overall contour of the plane is regular and reasonable.

Row 12: 12-1 The kitchen area is too large, but the overall contour of the plane is relatively reasonable. Besides, The balcony cannot be directly accessed from the living room. 12-2 The bedroom cannot be directly accessed from the living room. 12-3 The floor plan shape is irregular and not economical. 12-4 The length-to-width ratio of the bathroom room is greater than 2:1. 12-5 The area and proportion of each room are reasonable, the functional zoning and flow lines are reasonable, and the overall contour of the plane is regular and reasonable.

Row 13: 13-1 The kitchen area is too small, and the bedroom at the top left corner has an irregular shape. 13-2 The kitchen's length-to-width ratio is greater than 2:1. 13-3 The area and proportion of each room are reasonable, the functional zoning and flow lines are reasonable, and the overall contour of the plane is relatively reasonable. 13-4 The length-to-width ratio of the bathroom room is greater than 2:1, and the length-to-width ratio of the bedroom at the top left corner is also greater than 2:1. 13-5 The area and proportion of each room are reasonable, the functional zoning and flow lines are reasonable, and the overall contour of the plane is regular and reasonable.

Row 14: 14-1 The length-to-width ratio of the bathroom at the top left corner is greater than 2:1, and the bedroom in the middle at the bottom is too small. 14-2 The bathroom in the middle of the plane is a dark room, and the length-to-width ratio of the bedroom at the top is greater than 2:1. 14-3 The overall contour of the plane is irregular. 14-4 The bedroom on the left side of the plane is too small, and the overall contour of the plane is irregular. 14-5 The area and proportion of each room are reasonable, the functional zoning and flow lines are reasonable, and the overall contour of the plane is regular and reasonable.

Row 15: 15-1 The bathroom at the top left corner has an irregular L-shape, and the bedroom at the bottom left corner also has an irregular L-shape. 15-2 The area and proportion of each room are reasonable, the functional zoning and flow lines are reasonable, and the overall contour of the plane is regular and reasonable. 15-3 The bathroom at the top right corner cannot be directly accessed from the living room. 15-4 The bedroom in the bottom-left corner of the floor plan has a length-to-width ratio greater than 2:1. 15-5 The area and proportion of each room are reasonable, the functional zoning and flow lines are reasonable, and the overall contour of the plane is regular and reasonable.

Row 16: 16-1 The area and proportion of each room are reasonable, the functional zoning and flow lines are reasonable, but the overall contour of the plane is slightly irregular. 16-2 The kitchen and balcony areas are too small, and the bedroom at the bottom has an irregular shape. 16-3 The area and proportion of each room are reasonable, the functional zoning and flow lines are reasonable, and the overall contour of the plane is relatively reasonable. 16-4 The area and proportion of each room are reasonable, the functional zoning and flow lines are reasonable, and the overall contour of the plane is relatively reasonable. 16-5 The area and proportion of each room are reasonable, the functional zoning and flow lines are reasonable, and the overall contour of the plane is relatively reasonable. 16-6 The area and proportion of each room are reasonable, the functional zoning and flow lines are reasonable, and the overall contour of the plane is relatively reasonable. 16-7 The area and proportion of each room are reasonable, the functional zoning and flow lines are reasonable, and the overall contour of the plane is regular and reasonable.

Row 17: 17-1 The area and proportion of each room are reasonable, the functional zoning and flow lines are reasonable, and the overall contour of the plane is relatively reasonable. 17-2 The bathroom area is too large, and the bedroom in the middle cannot be directly accessed from the living room. 17-3 The functional zoning is rational and economical, and the circulation is well - designed and cost - effective. 17-4 The area and proportion of each room are reasonable, the functional zoning and flow lines are reasonable, and the overall contour of the plane is relatively reasonable. 17-5 The area and proportion of each room are reasonable, the functional zoning and flow lines are reasonable, and the overall contour of the plane is relatively reasonable. 17-6 The area and proportion of each room are reasonable, the functional zoning and flow lines are reasonable, and the overall contour of the plane is relatively reasonable. 17-7 The area and proportion of each room are reasonable, the functional zoning and flow lines are reasonable, and the overall contour of the plane is regular and reasonable.

Row 18: 18-1 The bedroom at the bottom right corner has an unreasonable L-shape, and the length-to-width ratio of the kitchen room is greater than 2:1. 18-2 The bedroom is too large in area., and the bathroom cannot be directly accessed from the living room. 18-3 The overall contour of the plane is relatively unreasonable. 18-4 The length-to-width ratio of the kitchen room is greater than 2:1. 18-5 The length-to-width ratio of the kitchen room is greater than 2:1. 18-6 The length-to-width ratio of the kitchen room is greater than 2:1, and the overall contour of the plane is relatively unreasonable. 18-7 The area and proportion of each room are reasonable, the functional zoning and flow lines are reasonable, and the overall contour of the plane is regular and reasonable.

Row 19: 19-1 The area and proportion of each room are reasonable, the functional zoning and flow lines are reasonable, and the overall contour of the plane is regular and reasonable. 19-2 The bathroom at the top left corner has an unreasonable L-shape, and the bathroom in the middle is a dark room. 19-3 The bathroom on the left side is too small. 19-4 The area and proportion of each room are reasonable, the functional zoning and flow lines are reasonable, but the overall contour of the plane is relatively unreasonable. 19-5 The bathroom on the left side is too small. 19-6 The area and proportion of each room are reasonable, the functional zoning and flow lines are reasonable, but the overall contour of the plane is relatively unreasonable. 19-7 The area and proportion of each room are reasonable, the functional zoning and flow lines are reasonable, and the overall contour of the plane is regular and reasonable.

Row 20: 20-1 The area and proportion of each room are reasonable, the functional zoning and flow lines are reasonable, and the overall contour of the plane is relatively reasonable. 20-2 The bathroom has an unreasonable L-shape, and the kitchen is a dark room. 20-3 The gaps between the bedrooms are too large, making the layout uneconomical. 20-4 The bathroom cannot be directly accessed from the living room. 20-5 The gaps between the bedrooms are too large, making the layout uneconomical. 20-6 The area and proportion of each room are reasonable, the functional zoning and flow lines are reasonable, but the overall contour of the plane is relatively unreasonable. 20-7 The area and proportion of each room are reasonable, the functional zoning and flow lines are reasonable, and the overall contour of the plane is regular and reasonable.

Row 21: 21-1 The area and proportion of each room are reasonable, the functional zoning and flow lines are reasonable, and the overall contour of the plane is relatively reasonable. 21-2 The bedroom at the top right corner is too small, and the length-to-width ratio of the kitchen room is greater than 2:1. 21-3 The area and proportion of each room are reasonable, the functional zoning and flow lines are reasonable, and the overall contour of the plane is relatively reasonable. 21-4 The area and proportion of each room are reasonable, the functional zoning and flow lines are reasonable, and the overall contour of the plane is relatively reasonable. 21-5 The area and proportion of each room are reasonable, the functional zoning and flow lines are reasonable, and the overall contour of the plane is relatively reasonable. 21-6 The area and proportion of each room are reasonable, the functional zoning and flow lines are reasonable, and the overall contour of the plane is relatively reasonable. 21-7 The area and proportion of each room are reasonable, the functional zoning and flow lines are reasonable, and the overall contour of the plane is regular and reasonable.

Row 22: 22-1 The bathroom cannot be directly accessed from the living room, but the overall contour of the plane is relatively reasonable. 22-2 The bathroom area is too large, and the kitchen room has an irregular L-shape. 22-3 The kitchen cannot be directly accessed from the living room, but the overall contour of the plane is relatively reasonable. 22-4 The bathroom cannot be directly accessed from the living room, but the overall contour of the plane is relatively reasonable. 22-5 The bathroom cannot be directly accessed from the living room, but the overall contour of the plane is relatively reasonable. 22-6 The bathroom cannot be directly accessed from the living room, but the overall contour of the plane is relatively reasonable. 22-7 The area and proportion of each room are reasonable, the functional zoning and flow lines are reasonable, and the overall contour of the plane is regular and reasonable.

Row 23: 23-1 The bedroom at the top right corner has an irregular shape, and the overall contour of the plane is irregular. 23-2 The bathroom and the bedroom in the middle of the plane are dark rooms, and the kitchen area is too large. 23-3 The overall contour of the plane is irregular. 23-4 The overall contour of the plane is irregular. 23-5 The area and proportion of each room are reasonable, the functional zoning and flow lines are reasonable, and the overall contour of the plane is relatively reasonable. 23-6 The length-to-width ratio of the bathroom room is greater than 2:1, and the overall contour of the plane is relatively reasonable. 23-7 The area and proportion of each room are reasonable, the functional zoning and flow lines are reasonable, and the overall contour of the plane is regular and reasonable.

Row 24: 24-1 The two bedrooms at the top are both unreasonable L-shapes, and the bathroom at the bottom right corner is not connected to the living room. 24-2 The bedroom at the top right corner has an irregular shape, and the overall contour of the plane is irregular. 24-3 The bathroom at the bottom right corner cannot be directly accessed from the living room, and the overall contour of the plane is irregular. 24-4 The bathroom area is too small, and the length-to-width ratio of the bedroom room is greater than 2:1, and the balcony is too large. 24-5 The length-to-width ratio of the kitchen room is greater than 2:1, and the balcony is not connected to the bedroom or living room. 24-6 The two bathrooms on the left side cannot be directly accessed from the living room, and the length-to-width ratio of the kitchen room is greater than 2:1. 24-7 The bedroom and kitchen room have unreasonable L-shapes, and the bathroom area is too small. 24-8 The kitchen room has an unreasonable L-shape, and the overall contour of the plane is irregular.

Row 25: 25-1 The length-to-width ratio of both the bedroom and kitchen rooms is greater than 2:1, and the kitchen area at the bottom right corner is too small. 25-2 The length-to-width ratio of the bedroom on the right side is greater than 2:1, and the bedroom at the top right corner cannot be directly accessed from the living room. 25-3 The length-to-width ratio of the bathroom at the top right corner is greater than 2:1, and the overall contour of the plane is relatively irregular. 25-4 There are too many balconies, and the living room area is too small. 25-5 The length-to-width ratio of the bedroom at the top is greater than 2:1, and the bathroom on the left side cannot be directly accessed from the living room. 25-6 The length-to-width ratio of the bedroom at the bottom left corner is greater than 2:1, and there are too many living rooms. 25-7 The length-to-width ratio of the kitchen room is greater than 2:1, and the bathroom area is too small. 25-8 The bedroom at the bottom left corner has an unreasonable L-shape, and the bathroom cannot be directly accessed from the living room.

Row 26: 26-1 The overall contour of the plane is irregular, and the bedroom at the top left corner is too small. 26-2 The bedroom at the top left corner is too small. 26-3 There are too many balconies, and the kitchen area is too large, and the balcony is located in the middle of the room, and the bathroom cannot be accessed directly from the living room. 26-4 The bathroom on the right side is a dark room, and there are too many bathrooms. 26-5 The bedroom at the top right corner is too small, and there are too many balconies. 26-6 The bedroom at the top right corner is too small. 26-7 The bedroom at the top left corner is too small, and the balcony on the left side is too large. 26-8 The bedroom at the bottom left corner is too small, and the balcony on the right side is excessively narrow and elongated.

Row 27: 27-1 There are too many bedrooms, and the overall contour of the plane is irregular. 27-2 There are too many living rooms, and the overall contour of the plane is irregular, and the bedroom in the upper right corner of the plan has an irregular shape. 27-3 The balcony area is too large, and there are too many bathrooms, and the plan lacks the bedroom. 27-4 The plane lacks bedrooms, making the function unreasonable, and there are too many bathrooms. 27-5 The plane lacks bedrooms, making the function unreasonable, and the kitchen and bathroom are too large in area. 27-6 There are too many living rooms, and the bathroom in the middle of the plane is a dark room. 27-7 There are too many bathrooms, and the length-to-width ratio of the kitchen room is greater than 2:1. 27-8 The two bedrooms at the bottom both have unreasonable L-shapes, and the balcony on the right side is excessively narrow and elongated.

Row 28: 28-1 The bedroom at the bottom right corner cannot be directly accessed from the living room, and the bathroom on the right side cannot be directly accessed from the living room. 28-2 Both bathrooms on the right side cannot be directly accessed from the living room, and the overall contour of the plane is relatively irregular. 28-3 The bedroom area is too small, and there are too many bathrooms. 28-4 The plane lacks a living room, and the balcony should not be arranged inside the bedroom. 28-5 The bathroom is a dark room, and there are too many living rooms. 28-6 The bathroom at the bottom left corner cannot be directly accessed from the living room, and the kitchen area is larger than the bedroom. 28-7 The bathroom cannot be directly accessed from the living room, and the overall contour of the plane is relatively irregular. 28-8 The length-to-width ratio of the bathroom on the left side is greater than 2:1, and the bathroom on the left side cannot be directly accessed from the living room.

Row 29: 29-1 The bedroom at the bottom right corner cannot be directly accessed from the living room, and the length-to-width ratio of the bedroom at the bottom right corner is greater than 2:1. 29-2 The length-to-width ratio of the bedroom at the top is greater than 2:1, and the bedrooms and kitchen at the bottom cannot be directly accessed from the living room. 29-3 The bedroom has an unreasonable L-shape, and there are too many living rooms. 29-4 The bedroom area is too large, and the bedroom at the top cannot be directly accessed from the living room. 29-5 In the plan, there are too many bathrooms, and they are too small in area. The bathrooms in the upper right corner and at the bottom cannot be accessed directly from the living room. 29-6 The bedroom at the top left corner is too small, and the bedroom at the top left corner cannot be directly accessed from the living room. 29-7 The bathroom cannot be directly accessed from the living room, and the kitchen room has an unreasonable L-shape. 29-8 The bedroom at the bottom left corner has an unreasonable L-shape, and the bathroom cannot be directly accessed from the living room.