VAIDE: Virtual AI Designer for Web3D Exhibition Layout Creation

Bixiao Zhao
Tongji University
Shanghai, China
2231517@tongji.edu.cn

Nanxi Chen*
Shanghai Institute of Microsystems and Information Technology, CAS
Shanghai, China
University of Chinese Academy of Sciences
Beijing, China
nanxi.chen@mail.sim.ac.cn

Jinyuan Jia*
School of Game, Jilin Animation Institute
Changchun, China
csjyjia@aliyun.com

Enyang Feng
School of Game, Jilin Animation Institude
Changchun, China
fengenyang2021sh@163.com

Xiaohen Li
Tongji University
Shanghai, China
2333071@tongji.edu.cn

2024 IEEE International Conference on Web Services (ICWS) | 979-8-3503-6855-0/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICWS62655.2024.00158

***Abstract*—The design of a virtual exhibition layout is a meticulous process. It usually starts with floor plan design, adhering to guidelines such as visitor traffic flow, exhibit prominence, and aesthetic considerations. Based on floor plans, 3D models are constructed, followed by the creation of final scenes. However, there is a notable lack of an efficient method to streamline this complex design process, which often requires profound professional expertise. To address this gap, we propose a novel approach leveraging generative AI and a floor plan recognition algorithm to reconstruct 3D exhibition layouts, thereby enhancing designers' efficiency. We have achieved the design of exhibition hall floor plans using Stable Diffusion and completed the automatic conversion from floor plans to 3D models, as well as the web-based operation and display platform. We undertake a comparative analysis between our solution and a CGAN-based method, evaluating the results through both subjective and objective measures. Furthermore, we present a comprehensive flowchart outlining the process of generating a museum exhibition hall on a web platform. Additionally, we demonstrate the design impact of our methodology in creating a virtual exhibition hall utilizing Web3D technology.**

***Index Terms*—Web3D, Stable Diffusion, AIGC, Exhibition layout design, Generative AI**

I. Introduction

With the expansion of the cultural industry, there has been a noticeable uptick in the number of exhibitions, accompanied by a rapid evolution in the themes of exhibition design. This dynamic landscape necessitates a streamlined approach to exhibition hall design to mitigate costs, simplify the design process, and foster more effective communication with laypersons, thereby elevating the overall efficacy of exhibition designs. Currently, the design and communication process for exhibition hall layouts faces several challenges.

Elevated Communication Expenses: Traditional renderings fall short in fully capturing the intricate details and overall design impact of exhibition hall layouts. Given the collaborative nature of exhibition design involving exhibitors, organizers, and designers, this shortfall can significantly escalate communication costs. Adjustments to renderings are notably challenging, exacerbating the expense and time required to address unforeseen changes, such as the need for last-minute exhibit substitutions.

Increased 3D Modeling Expenditures: Although 3D modeling offers a more detailed representation of exhibition hall designs, the associated costs for creation and alteration are substantial. This makes 3D modeling less feasible in the preliminary stages of design, where requirements are prone to frequent changes. Moreover, sharing the source files of 3D models poses a risk to the safeguarding of designers' intellectual property. Additionally, panoramic videos of 3D models may omit crucial details, further compounding the issue.

The domain of Artificial Intelligence Generative Content (AIGC) [1] is witnessing rapid progress in indoor design, offering more precise and efficient design plans than traditional sketches and thus reducing the initial costs of design. In the realm of exhibition hall design, there has been a push to employ AIGC techniques for crafting floor plans [4]. Nevertheless, these efforts typically employ a broad generative approach, lacking the nuanced control needed over specific exhibition areas. Detailed control is pivotal for designing exhibition spaces that enhance the viewing experience and dictate the flow of visitor traffic.

Research into fine-grained control has been conducted within the home design domain [5], utilizing technologies like Stable Diffusion [6] or GAN [7] models for floor plan creation and AI for the conversion of these plans into vector graphics. These methodologies, however, are primarily focused

Our work is supported by Ant Group and General Program of National Natural Science Foundation of China (NSFC, No. 6207071897).
* Corresponding author

2836-3868/24/$31.00 ©2024 IEEE
DOI 10.1109/ICWS62655.2024.00158

on residential layouts, which are designed with functionality in mind, using a standard template to segment the space into bedrooms, kitchens, and other clearly defined functional zones [5]. This approach is not directly applicable to exhibition hall layouts, which demand a greater degree of flexibility to accommodate the diverse needs of stakeholders. Moreover, the lack of a uniform standard in output complicates the automatic transformation of two-dimensional floor plans into three-dimensional models, highlighting the need for tailored solutions in the exhibition design sector.

Web3D technology [8], leveraging the power of the internet to display three-dimensional graphics, now possesses enhanced rendering capabilities that meet the requirements for depicting exhibition halls without sharing the source files of 3D models. Web3D offers a dynamic and comprehensive visualization to all stakeholders involved in exhibition design, effectively circumventing the financial inefficiencies linked with partial renderings. Furthermore, Web3D enables real-time updates and effortless sharing of design visualizations, significantly lowering communication overhead. However, the initial hurdle of constructing 3D models on the web, due to its complexity and associated costs, presents a challenge. Traditional Web3D uses a GL Transmission Format (glTF) 3D model format, which is more complex and contains useless information, such as the vertex matrix of walls, but just needs two corner positions.

This paper proposes VAIDE, which stands for a Virtual AI Designer for Web3D Exhibition layout creation. VAIDE revolutionized the process of transforming the basic outline of an exhibition hall into a comprehensive Web3D model. This transformation involves a meticulous process that incorporates three core components: a Stable Diffusion-based model for creating detailed floor plans, a post-processing and recognition model for interpreting these plans, and a web-based platform for constructing the 3D models. The recognition model plays a crucial role by identifying the vertex positions of walls and exhibition areas and translating this information into JSON format for seamless integration with the Web3D construction process. Our goal is to streamline the creation of 3D layouts, thereby overcoming the traditional cost barriers associated with Web3D model development. The principal achievements of our research are encapsulated in the following contributions:

1) **Enhanced Control via ControlNet:** We have refined the Stable Diffusion model using ControlNet, enabling more precise and controllable outcomes for exhibition hall floor plans. Additionally, we've developed a layout model that generates a mask for the floor plan, further aiding designers in tailoring the Stable Diffusion outputs to their specific needs.
2) **Accurate Floorplan Recognition for 3D Reconstruction:** Our tailored exhibition hall floorplan recognition model facilitates the conversion of 2D designs into 3D models with a high accuracy rate. This high level of precision ensures that the final 3D models are accurate to the original designs.
3) **A Unified Web Platform for Comprehensive Design and Visualization:** By integrating the Stable Diffusion design system with our floorplan recognition model into a singular web platform, we've created a versatile tool that allows designers and stakeholders to engage directly with the exhibition hall planning process. This platform not only enables the efficient design and adjustment of floor plans but also supports the 3D construction of these designs. Furthermore, it provides functionalities for placing exhibits within the exhibition space, offering an initial glimpse into the spatial dynamics and visual appeal of the proposed setup.

Our contributions mark significant advancements in the field of exhibition hall design, offering a set of tools that enhance the efficiency, accuracy, and creativity of the design process. By facilitating a more interactive and user-friendly approach to 3D model creation, we aim to set a new standard for exhibition planning and execution. Source code at: https://github.com/zbx99/VAIDE

II. Related Work

AIGC is growing rapidly in areas like natural language, images, and 3D due to large models. In the field of image generation, diffusion models [2], DALL-E-2 [3] and Midjourney generate images from natural language descriptions based on user prompts, both achieving excellent effects. Moreover, Stable Diffusion [6] not only possesses powerful image generation capabilities but also has cultivated a large open-source community, helping users build personalized AI.

Additionally, many large model fine-tuning methods, such as Hypernetworks [9], DreamBooth [10], and LoRA [11], have enabled the public to obtain their own personalized AI design assistants. Particularly, LoRA [11] allows for cost-effective fine-tuning of large models and produces satisfactory results.

In comparison to GANs [13], Stable Diffusion (SD) [6], under the control of related small models such as ControlNet [12], can provide a more stable and controllable image generation process, which is suitable for exhibit floorplan design that requires precise control over design elements and details. Furthermore, compared to GANs, which require more detailed adjustments and are prone to mode collapse, fine-tuning of Stable Diffusion based on LoRA [11] provides a simpler and more effective adjustment method, offering a feasible self-training approach for designers.

A. Research of Exhibition Hall Floorplans Design

At present, there is a growing interest in floorplan design for exhibition halls. Wang et al. [14] have proposed artificial-intelligence-assisted floorplan design methods for exhibition halls. Also, Min et al. [4] proposed an exhibition hall floorplan design model based on a conditional GAN to assist designers in completing exhibition hall designs.

B. Research of Home Floorplans Design

Our research about home floorplans primarily focuses on two aspects: one is the segmentation of floorplans based on boundary constraints. Floorplan design based on boundary constraints achieves a reasonable structural design and room division within the boundaries. Early floorplan generation algorithms [15] construct the objective function based on architectural quality indicators and resident preferences, starting from initial randomly generated floorplans that are iteratively optimized using evolutionary algorithms. Wu et al. [16] simulates the human design process, first locating rooms, then positioning walls, and simultaneously adjusting the input architectural boundaries to generate high-demand floor plan designs. This method uses a living room priority strategy to predict room locations, thereby improving the rationality of the generated floorplans.

Nauata et al. [21] propose the House-GAN++ model, which uses a bubble diagram to present rooms and their relationship and convert it into floorplans. He et al. [17] proposes a new generation model, iPLAN, which decomposes the design process into three steps: obtaining room types, positioning rooms, and finally determining room partitions. This process allows the model to accept input from designers at any stage to generate floorplans that meet user preferences. Sun et al. [18] designed a wall-oriented method called WallPlan, which innovatively uses two modules to predict the floorplan's walls and room functions separately. Then, these two modules alternately generate partial floorplans until no new walls can be generated. WallPlan can produce high-quality floor plans without post-processing.

The other aspect in this area primarily focuses on converting floorplans into vector graphics for subsequent processing.

Floorplan vectorization methods are broadly grouped into two: segmentation-based for element recognition and detection-based for junction points. Chen et al. [19] use a detection model and CNNs to identify wall corners, but it fails for slanted walls and false detections can ruin the entire reconstruction. Zeng et al. [20] enhance the deep neural network and loss function to improve floorplan segmentation precision, but they cannot restore the vectorized floorplan structure.

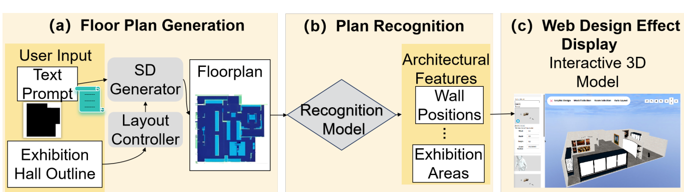

Fig. 1. The main process of VAIDE

III. WEB3D EXHIBITION LAYOUT CREATION

The flowchart of VAIDE (Virtual AI Designer for Web3D Exhibition Layout Creation) serves as a blueprint for our innovative approach to exhibition layout design, seamlessly integrating AI technology with web-based visualization tools. The flowchart, delineated in Figure 1, organizes the VAIDE process into three distinct yet interconnected stages:

Floor Plan Generation: This initial phase leverages the Stable Diffusion (SD) AI model [6] to generate detailed floor plans for exhibition halls. Here, designers can input specific parameters or criteria to influence the layout generation, ensuring the final design aligns with their vision and the exhibition's requirements. The use of ControlNet [12] for fine-tuning enables even more precise control over the generated layouts, allowing for the customization of spaces to suit various themes or exhibit sizes.

Plan Recognition: Following the creation of the floor plan, the recognition model comes into play. This advanced AI model analyzes the generated floor plan to identify key elements such as wall positions, exhibition areas, and other relevant architectural features. The accuracy of this model is crucial, as it translates the two-dimensional design into data that can be used to construct a three-dimensional model. This model ensures a high level of fidelity between the original design and its 3D representation.

Web Design Effect Display: The culminating phase of the VAIDE process involves the visualization of the exhibition layout in a 3D Web3D model. The data processed by the recognition model is utilized to build an interactive 3D model that can be viewed and navigated.

A. Floor Plan Generation

In VAIDE, we employ a combination of SD image Generator and a Controllor, serving as the foundational architecture of our model. This architecture is crafted to include five key components: a text encoder, layout separation, ControlNet, an image creator, and an image decoder. Each component plays a crucial role in translating descriptive prompts of exhibition halls into detailed, controllable floorplan designs.

The initial step in our process involves the text encoder, which is responsible for converting descriptive prompts about the exhibition hall into token embeddings. By utilizing CLIP [25], a pre-trained model, in a manner akin to its application in Stable Diffusion, we ensure that our model comprehends the descriptive prompts with high accuracy and contextual relevance. SD relies on ControlNet to exert precise control over the outputs. By introducing a mask that segments the exhibition hall into various functional and thematic areas, ControlNet enables the targeted generation of walls and other structural elements, significantly enhancing the customization capabilities offered to designers.

We propose a Layout Controller to parse the encoded information and delineate the exhibition hall into distinct regions or zones. This segmentation forms the basis for the subsequent detailed design steps, establishing a framework within which specific elements of the exhibition hall can be developed. The process of transforming the layout of an exhibition hall into a detailed floorplan involves an intricate model that divides the hall into distinct regions for exhibits and booths. This model operates through two primary stages: region localization and mask generation, each leveraging different technologies to achieve precise and customizable design outcomes.

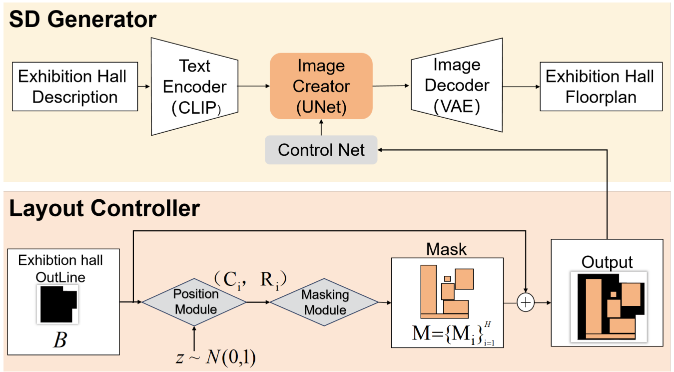

Fig. 2. The architecture of floor plan generation

The whole design procedure can be presented as this formula.

$$M = \{M_i\}_{i=1}^{H} = \{GM(B, C_i, R_i)\}_{i=1}^{H} \\ = \{GM(B, P_i(B, z))\}_{i=1}^{H} \quad (1)$$

The mask output of the exhibition hall is $M = \{M_i\}_{i=1}^{H}$, which M_i represents the mask of i-th exhibitionable area. B presents the boundary of the whole exhibition hall. Based on B and $z \sim N(0,1)$, we generate the i-th center C_i and i-th mask regions R_i, which presents the top-left and bottom-right corners of the bounding box of the exhibitionable area iteratively. The GM presents the mask generation model, and the P_i presents the i-th mask position generation model.

For the task of region localization within the exhibition hall, we employ ResNet18 as the base model. ResNet18, known for its efficiency and effectiveness in image recognition tasks, serves to identify specific areas designated for exhibits and booth placements. Through a regression technique, the model computes the coordinates of bounding box vertices around the booth areas. This step is crucial for delineating the spatial constraints and organization within the hall, ensuring that each exhibit or booth is allocated its rightful place according to the design parameters.

Following the localization of regions, the next step involves generating a mask for the exhibit areas. We utilize OpenCV, a comprehensive library for computer vision tasks, to create these masks. The generated mask is then superimposed on the original floorplan image, resulting in a composite that is fed into ControlNet. This composite image allows ControlNet to apply localized control over the drawing output, ensuring that the design accurately reflects the intended layout and segmentation of the exhibition space.

With the regions defined and control parameters set, the Image Creator takes over. Utilizing the groundwork laid by the previous stages, it generates high-level latent features of the exhibition hall floorplan. Then, the Image Decoder decodes the latent features into the virtual output of the exhibition hall floorplan, readying the floorplan for further processing, such as recognition and 3D modeling.

We extend a virtual museum floor plan dataset [4], which comprises 66 distinct floor plans, and further use LoRA (Low-Rank Adaptation) to fine-tune the Stable Diffusion model. LoRA requires relatively small amounts of data to achieve significant adaptability to the target exhibition scenario. In order to train the layout separation model, we labeled the dataset of exhibitionable area masks and got the center and region of each mask. So, the whole dataset contains four parts, including input.png, mask.png, output.png, and CR.json.

B. Plan Recognition

For the virtual museum floor plan recognition model, its main function is to eliminate design imperfections such as image burrs and edge blurring in different areas generated by Stable Diffusion. By post-processing optimization and relatively precise pixel positioning, the goal is to achieve 3D reconstruction of virtual museum floor plans.

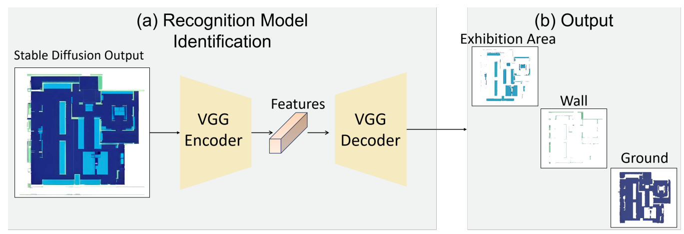

Fig. 3. Recognition model input and outputs

To achieve the objectives mentioned above, we designed a model as depicted in Figure 3 to separately extract walls and exhibition areas from the design diagrams. Using OpenCV for post-processing of the wall and exhibition area images, we achieved the final wall reconstruction and the delineation of the exhibition areas.

We initially employed a recognition model to identify and extract different regions from the exhibition hall floorplan, categorizing the overall floorplan into Wall Output and Exhibition Area Output. For each extracted region, we use OpenCV to refine image artifacts. For the wall output, we applied a centerline extraction algorithm to get the wall center lines, accurately pinpointing the start and end positions of the walls, which were then saved in JSON data. The specific processing flow and effects are shown in Figure 4.

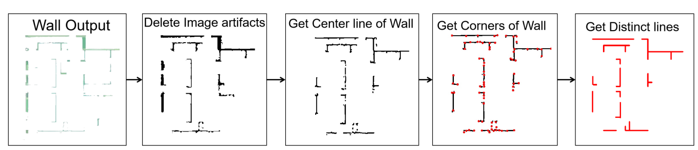

Fig. 4. Flow chart of wall output post-processing

For the Exhibition Area Output, we use a minimum bounding box algorithm to get the bounding box of each exhibition area. By comparing the differences between the bounding circles and rectangles and the original recognized areas, we determined the approximate shapes of each exhibition area. This approach achieved a balance between the accuracy of exhibition area recognition and the regularity of the defined areas. Ultimately, we saved the results of the exhibition area

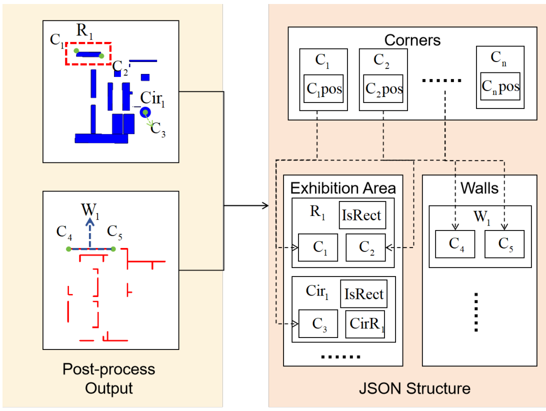

Fig. 5. JSON Data of exhibition hall layout output

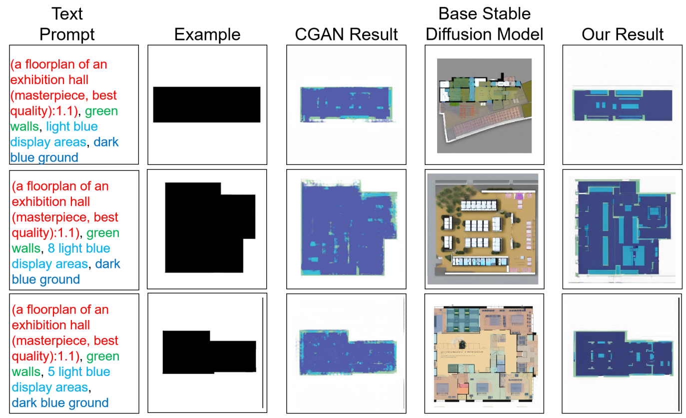

Fig. 7. Comparison of CGAN model, Stable Diffusion and our generation output

recognition in JSON format for rendering by Web3D. The final JSON data format we used for reconstructing the 3D exhibition hall is listed in Figure 5.

C. Web Design Effect Display

At last, we implemented an online exhibition hall with an automated layout design and placement system based on Web3D technology. In this system, users first draw or upload the outline of the exhibition hall, and then the initial design of the exhibition hall's internal layout is realized based on the Stable Diffusion model. After users choose their relatively preferred design diagram, they enter the floorplan recognition system. The floor plan recognition system completes the identification of exhibit spacing and walls in the exhibition hall and their 3D reconstruction. The overall working effect of the system is shown in Figure 6.

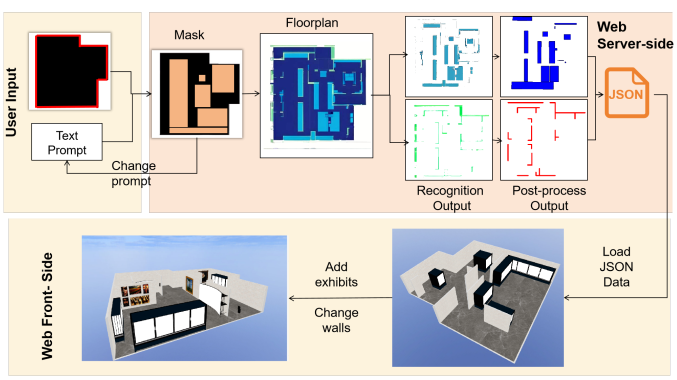

Fig. 6. The VAIDE service process

IV. Experiments

A. Exhibition Hall Floorplan Design

VAIDE involves training and fine-tuning of exhibition hall floorplans based on the Stable Diffusion v1.5 base model. Through iterative training, an initial design of the exhibition area within the exhibition hall has been preliminarily realized. A comparison of the training results from ours with the Conditional GAN (CGAN) method proposed in [4] is conducted. We also compared the Stable Diffusion v1.5 base model with the same prompt. The comparison results, as illustrated in Figure 7 and Table I, indicate that the exhibition hall floorplans generated by the Stable Diffusion model are relatively richer in layout compared to the CGAN model.

Given the absence of a publicly available CGAN model for direct comparison, we used the proposed virtual dataset to generate floor plans for both of the solutions. By using the same dataset to generate outcomes with Stable Diffusion, we were able to conduct a thorough analysis of the strengths and weaknesses of Stable Diffusion relative to the CGAN model. The specifics of this comparative study are detailed in a dedicated section of our paper, with comprehensive results tabulated for easy reference.

From Figure 7, we can find out that compared with our output, the exhibition areas of the floorplan generated by CGAN have blurry image edges. For designers, unclear edges need more human intervention, and it is inconvenient to construct 3D models or generate traditional rendering output.

The base diffusion model generated floorplans with many meaningless parts. Additionally, in the last result of the base Stable Diffusion model, the floorplan is much more similar to the home floorplan rather than the exhibition hall. It is likely that the baseline StableDiffusion model's output is polluted by a large number of household floor plans on the Internet, which leads to incorrect design output. Our fine-tuned model increased the weight of the exhibition hall floor plan. Thus, the outputs generated by our model are closer to the actual design results of the exhibition halls. Compared to CGAN, our model preserves relatively clear boundaries between areas, which significantly facilitates further processing by subsequent recognition models and designers.

Based on the proposed virtual museum dataset, we proceeded with the LoRA fine-tuning of the Stable Diffusion model, utilizing a user-friendly graphical interface (GUI) provided by the Stable Diffusion platform. The efficiency of this fine-tuning process was demonstrated through our experimental results, which showed that after merely 20 epochs of training, the model was capable of producing satisfactory design outcomes. The success of these experiments and the quality of the resulting designs are further documented, with specific examples and comparative analyses presented to underscore the effectiveness of our approach.

To further compare the generation effects of the CGAN model, the fine-tuned Stable Diffusion model, we compared them with the standards below:

1) **Valid Plan Rate**: We evaluate how many of the generated images show a valid floor plan that is not cut off and contains at least some recognizable symbology in each 100 images.
2) **Overfit Check**: We check if the Stable Diffusion models do not overfit on the training samples by prompting for human faces, expecting that we do not get floor plans. So, we check the number of non-floorplan designs by prompting for human faces in each 100 images.
3) **Exhibition Hall Layout Organization**: We first calculate the ratio of unique rectangular exhibition areas to the total exhibition space in a generated image. Then, we find the average ratio for n images, denoted by P_o. Here, P_i presents the i-th image. O_j and $\hat{O}_j$ present distinct and rectangularly exhibition show areas, and the whole exhibition areas

$$P_o = \frac{\sum_{i=1}^{n} P_i}{n} \quad P_i = \frac{\sum_j O_j}{\sum_j \hat{O}_j} \tag{2}$$

4) **Exhibition Hall Layout Filling Degree**: First, we get the ratio between the Exhibition area and ground to get the filling degree of a generated image. Second, we get the average ratio of 100 images. The evaluation method of generation output is similar to that of the Exhibition Hall Layout Organization.
5) **Fréchet Inception Distance**[22]: A quantitative evaluation metric widely used in the assessment of image generation models such as GANs and Variational Autoencoders (VAEs [24]). FID is employed to compute the distance between the Inception feature vectors of real images and generated images. This metric can effectively reflect the quality and diversity of image generation. A lower FID value indicates a smaller discrepancy between the distribution of generated images and the distribution of real images, suggesting that the generated images are closer to the real images.

TABLE I
QUANTITATIVE EVALUATION OF CGAN, OURS

Standards	Real	CGAN model	Ours
Valid Plan Rate	100%	52%	69%
Overfit Check			1.51%
Exhibition Hall Layout Organization	70.00%	54.95%	63.36%
Exhibition Hall Layout Filling Degree	19.83%	10.03%	5.96%
FID		122.31	117.19

Table I, shows that we get more valid exhibition hall floorplan than CGAN output, though we still have invalid output, such as exhibition hall floorplan with no exhibition area. Furthermore, by comparing the performance of our model with that of the CGAN model in terms of Exhibition Hall Layout Organization, we find that our generated results are more distinct floor plans for exhibitions. In addition, we evaluate the FID (Fréchet Inception Distance) between the images generated by CGAN and the actual images, as well as the FID between the images generated by our fine-tuned model and the actual images. Our results show that our model has a lower FID, with a distribution closer to the real data, leading to clearer and more effective generated floorplans.

B. Recognition Model Accuracy

We do not find a model to recognize exhibition hall design output or datasets about this, so we compared our model to a home floorplan recognition model [20], and labeled the dataset in [4] to get the recognized accuracy. For quantitative evaluation, we adopted two widely-used metrics [23]: the overall pixel accuracy and the per class pixel accuracy:

$$overall_{a}ccu = \frac{\sum_i N_i}{\sum_i \hat{N}_i} \quad class_{a}ccu(i) = \frac{N_i}{\hat{N}_i} \tag{3}$$

Where $\hat{N}_i$ and N_i are the total number of the ground-truth pixels and the correctly predicted pixels for the i-th floor plan element, respectively. Table II shows that in the exhibition hall floorplan recognition task, our model has higher accuracy in locating the exhibition hall exhibition area and wall. This is mainly because Zeng's [20] recognition model design is focused on home room type and their relationship, which leads to recognition mistakes in the exhibition hall floorplan.

TABLE II
COMPARISON BETWEEN HOME FLOORPLAN RECOGNITION MODEL AND OUR MODEL ON EXHIBITION HALL FLOORPLAN DATASET

Model	overall_accu	class_accu		
		Wall	Ground	Exhibition Area
Zeng et al. [20]	79.26%	91.02%	87.09%	94.12%
Ours	95.06%	98.56%	97.75%	98.75%

V. CONCLUSION

The experiments conducted in this study demonstrate that the system can effectively assist designers in their work. The system can optimize the design process of professionals, provide a more convenient initial design process for virtual museums, and offer a platform for designers to showcase the effects of their preliminary virtual museum design floor plans. However, there is room for further improvement in the system. Firstly, the training effect of Stable Diffusion and the quality of the generated floor plans are not stable and are significantly influenced by the dataset and prompt words. Decomposing the design steps, such as layouting the floor plan based on the virtual museum tour route provided by the designer and offering a process that allows designers to adjust the design effects of the floor plans, is a more reliable approach.

REFERENCES

[1] Cao Y, Li S, Liu Y, et al. A comprehensive survey of ai-generated content (aigc): A history of generative ai from gan to chatgpt[J]. arXiv preprint arXiv:2303.04226, 2023.

[2] Ho J, Jain A, Abbeel P. Denoising diffusion probabilistic models[J]. Advances in neural information processing systems, 2020, 33: 6840-6851.

[3] Ramesh A, Pavlov M, Goh G, et al. Zero-shot text-to-image generation[C]//International conference on machine learning. Pmlr, 2021: 8821-8831.

[4] Min X, Zheng L, Chen Y. The Floor Plan Design Method of Exhibition Halls in CGAN-Assisted Museum Architecture[J]. Buildings, 2023, 13(3): 756.

[5] El-Basioni, Basma M. Mohammad, S. M. A. El-Kader, and Mahmoud Abdelmonim. "Smart home design using wireless sensor network and biometric technologies." information technology 1 (2013): 2.

[6] Rombach R, Blattmann A, Lorenz D, et al. High-resolution image synthesis with latent diffusion models[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2022: 10684-10695.

[7] Creswell, Antonia, et al. "Generative adversarial networks: An overview." IEEE signal processing magazine 35.1 (2018): 53-65.

[8] Chittaro, Luca, and Roberto Ranon. "Web3D technologies in learning, education and training: Motivations, issues, opportunities." Computers and Education 49.1 (2007): 3-18.

[9] Ha D, Dai A M, Le Q V. HyperNetworks[C]//International Conference on Learning Representations. 2016.

[10] Ruiz N, Li Y, Jampani V, et al. Dreambooth: Fine tuning text-to-image diffusion models for subject-driven generation[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2023: 22500-22510.

[11] Hu E J, Shen Y, Wallis P, et al. Lora: Low-rank adaptation of large language models[J]. arXiv preprint arXiv:2106.09685, 2021.

[12] Zhang L, Rao A, Agrawala M. Adding conditional control to text-to-image diffusion models[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2023: 3836-3847.

[13] Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets Advances in neural information processing systems[J]. arXiv preprint arXiv:1406.2661, 2014.

[14] Wang B. Digital design of smart museum based on artificial intelligence[J]. Mobile Information Systems, 2021, 2021: 1-13.

[15] Bahrehmand A.,Batard T., Marques R.,Evans A.,Blat J.:Optimizing layout using spatial quality metrics and user preferences. Graphical Models 93 (2017) ,25–38

[16] Wu W.,Fu X., Tang R., Wang Y ., Qi Y ., Liu L.: Data-driven interior plan generation for residential buildings. ACM Transactions on Graphics (TOG) 38 (2019), 1–12.

[17] He F., Huang Y ., Wang H.: iplan: Interactive and procedural layout planning. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (2022), pp. 7793–7802.

[18] Sun J., Wu W., Liu L., Min W., Zhang G., Zheng L.:Wallplan: Synthesizing floorplans by learning to generate wall graphs. ACM Transactions on Graphics (TOG) 41, 4 (2022), 1– 14.

[19] Chen Liu, Jiajun Wu, Pushmeet Kohli, and Y asutaka Furukawa. Raster-to-vector: Revisiting floorplan transformation. In Proceedings of the IEEE International Conference on Computer Vision, pages 2195–2203, 2017

[20] Zeng Z, Li X, Yu Y K, et al. Deep floor plan recognition using a multi-task network with room-boundary-guided

[21] Nauata N, Hosseini S, Chang K H, et al. House-gan++: Generative adversarial layout refinement network towards intelligent computational agent for professional architects[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021: 13632-13641.

[22] Heusel M, Ramsauer H, Unterthiner T, et al. Gans trained by a two time-scale update rule converge to a local nash equilibrium[J]. Advances in neural information processing systems, 2017, 30.

[23] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2015: 3431-3440.

[24] Kingma D P, Welling M. Auto-encoding variational bayes[J]. arXiv preprint arXiv:1312.6114, 2013.

[25] Radford A, Kim J W, Hallacy C, et al. Learning transferable visual models from natural language supervision[C]//International conference on machine learning. PMLR, 2021: 8748-8763.