

## Article

# Enhancing the Sustainability of AI Technology in Architectural Design: Improving the Matching Accuracy of Chinese-Style Buildings

Feiran Chen <sup>1</sup>, Mengran Mai <sup>1</sup>, Xinyi Huang <sup>2</sup> and Yinghan Li <sup>1,\*</sup><sup>1</sup> Department of Architecture, School of Fine Art, South-Central Minzu University, Wuhan 430074, China<sup>2</sup> Jiangsu Foreign Affairs Translation and Interpretation Center, Nanjing 210024, China

\* Correspondence: 2022100@mail.scuec.edu.cn

**Abstract:** This study discusses the application of AI technology in the design of traditional Chinese-style architecture, aiming to enhance AI's matching accuracy and sustainability. Currently, there are limitations in AI technology in generating details of traditional Chinese-style architecture, so this study proposes a method of fine-tuning AI pre-training models, by extracting samples of traditional architectural style elements, to enhance the trajectory and output accuracy of AI generation. The research method includes constructing AI pre-training models, using DreamBooth and ControlNet tools for personalized training and perspective control. Through experimental verification, this study found that pre-trained models can effectively enhance the accuracy and controllability of AI in the preliminary design of architecture. At the same time, the application of ControlNet technology has significantly improved the accuracy and realism of architectural rendering. The value of this study lies in proposing a new method that combines AI technology with the process of traditional Chinese architectural design, which can help architects better protect and inherit the culture of traditional Chinese architecture. Through this method, it can reduce the difficulty of learning traditional Chinese architectural design, optimize the design process, enhance design efficiency, and provide strong support for the sustainable development of traditional Chinese architecture.



**Citation:** Chen, F.; Mai, M.; Huang, X.; Li, Y. Enhancing the Sustainability of AI Technology in Architectural Design: Improving the Matching Accuracy of Chinese-Style Buildings. *Sustainability* **2024**, *16*, 8414. <https://doi.org/10.3390/su16198414>

Academic Editor: Mohamed Gamal Abdelmonem

Received: 15 July 2024

Revised: 14 September 2024

Accepted: 24 September 2024

Published: 27 September 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Keywords:** architectural representation; architectural design; artificial intelligence; DreamBooth; Stable Diffusion

## 1. Introduction

### 1.1. Existing Problems

Traditional architecture is one of the core elements of Chinese historical and cultural heritage, with significant historical and cultural value. However, the complex wooden structures of Chinese traditional buildings pose a huge challenge for the average architect, who often requires years of accumulated experience and practice to design buildings with appropriate proportions and detailed perfection in traditional Chinese style. In China's architectural design industry, the design of traditional style buildings is separate from modern architectural design. Although this separation allows dedicated traditional-style architects to undergo continuous training and rapid growth, it also makes it difficult for technological advancements in the field of modern architectural design to be directly applied to the design process of traditional Chinese-style buildings. In the preliminary design phase, Chinese architects would provide clients with at least three design proposals, mainly including 2D floor plans and 3D architectural renderings. The designs must often go through numerous revisions during multiple discussions with clients and government regulatory departments, costing architects time and pointless works. The main reason for this efficiency is the insufficient communication and discussion during the preliminary design phase. Architects often have to revise or even redo designs whenever new ideas are proposed by the client and officials in the later stages.

Therefore, finding a way to integrate traditional Chinese architectural elements into the modern architectural design process has become a significant challenge that contemporary Chinese architects must face.

Currently, as a creative auxiliary tool, AI generative technology has made significant progress in the field of modern architectural design. However, many traditional Chinese architectural images primarily exist in the form of paintings on the internet. Therefore, AI training based on these online image resources tends to present traditional Chinese architecture in a painted style, as shown in Figure 1, which illustrates this problem in both Midjourney and DALL-E 3. However, the paintings displayed in Figure 1 are significantly different in form and structure compared to real buildings. Thus, even if AI can generate images of traditional Chinese architecture by learning from painted images, these images have limited applicability in actual architectural design. This results in a problem: the application of AI technology in designing traditional Chinese architectural styles has not yet been practical. Additionally, research on the use of AI in traditional Chinese architecture primarily focuses on related AI technologies like BIM (Building Information Modeling), with relatively few studies involving recent Diffusion Models.



**Figure 1.** Huizhou architecture generated by other AI tools.

Considering the aforementioned factors, the current application of AI technology in the design process of traditional Chinese architecture faces numerous limitations and challenges, rendering it incapable of addressing the existing issues within the traditional architectural design process. Overall, the main challenges of applying AI technology in traditional Chinese style architectural design are first, ordinary large AI models lack a deep understanding of traditional Chinese architectural styles, and second, these AI models are weak in generating details of traditional Chinese-style architecture.

### *1.2. The Sustainability of This Technological Approach*

This study demonstrates the potential of AI in preserving and promoting traditional architectural culture. By accurately capturing and reproducing the unique characteristics of Chinese traditional architecture, AI empowers architects to seamlessly integrate traditional elements into contemporary designs. This not only simplifies the learning process, but also facilitates the widespread application of traditional architectural styles, allowing traditional culture to transcend time and space. Moreover, this integration of AI and traditional architecture promotes sustainable development by ensuring that traditional architectural forms can adapt to modern needs, fostering a harmonious coexistence of culture and technology.

It proposes a novel AI framework to enhance the design process of traditional architecture. By leveraging a dataset of traditional architectural elements and a denoising diffusion model, the AI can generate visually accurate and structurally sound design schemes. This approach not only streamlines the design process, but also promotes sustainable practices by preserving the unique characteristics of traditional architecture while incorporating modern design principles.

### 1.3. Materials and Methods

The relationship between AI-assisted design and construction design can be viewed as an iterative optimization process. AI-generated initial design schemes provide a foundation for subsequent construction design, while feedback from the construction design process can, in turn, optimize the AI's initial design. This cyclical process enables close integration of design and construction, achieving more efficient and accurate design goals. Additionally, artificial intelligence technology also plays a significant role in the construction phase. Taking the BRAILS (Building Recognition and Intelligence for Large-Scale Structures) system developed by the US National Science Foundation as an example, this system can assess the vulnerability of urban buildings to natural disasters such as earthquakes, hurricanes, or tsunamis, providing data support for seismic retrofitting and resilience enhancement of buildings.

The method proposed in this research primarily targets the early stages of the architectural design process (Figure 2). Throughout the entire design process, this method aims to address the existing issues in Section 1 related to AI, including the inadequate reflection of traditional architecture and others. By harnessing AI, it outputs visual architectural schemes containing traditional architectural elements, even while reducing the precision of the modeling in traditional architectural design. This approach will lessen the workload in traditional architectural design projects and alleviate the burden on designers.

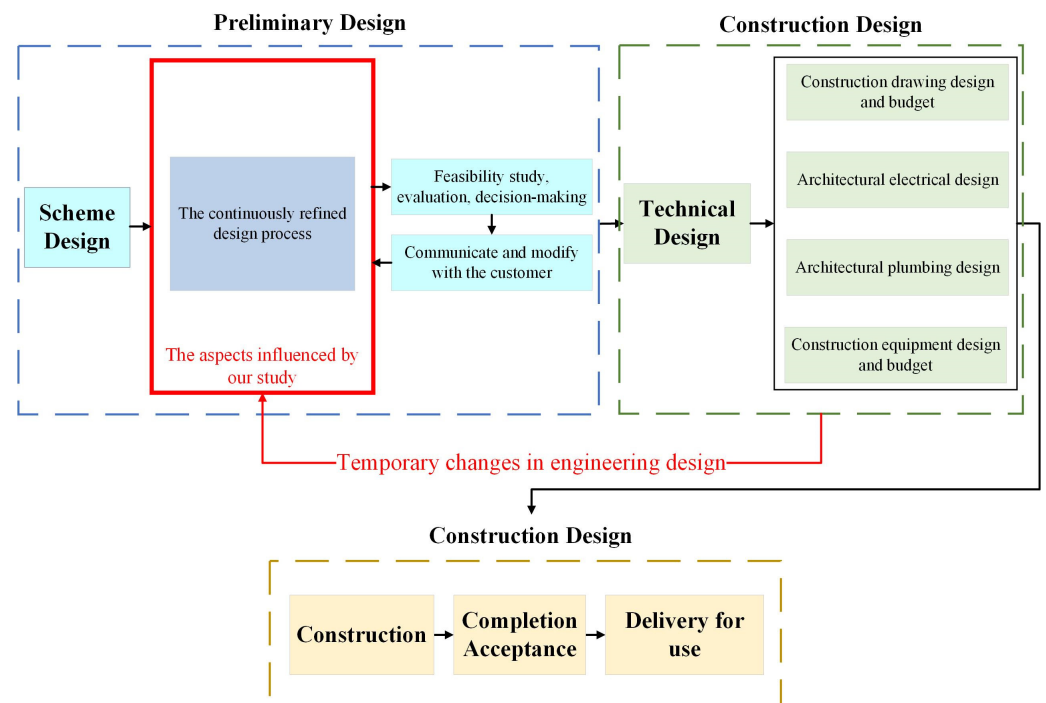


Figure 2. Design Scheme and communicate with and modify according to the wishes of the customer.

### 1.4. The Significance of the Research

During the rapid urbanization process in China over the past thirty years, many ancient buildings have been damaged, and numerous cities have been filled with international style buildings. This phenomenon has not only led to the loss of historical and cultural heritage,

but has also severely impacted the uniqueness and cultural depth of cities. Against this backdrop, the development of pseudo-classic buildings has accelerated. Pseudo-classic buildings reflect the essence of traditional Chinese culture, enhancing cultural identity and pride among people, while also playing a significant role in tourism by attracting a large number of visitors and bringing substantial economic benefits to local areas. The government has strongly supported the construction of pseudo-classic buildings at the policy level, encouraging social forces to participate in the protection and utilization of such buildings through various financial support and incentive measures. Moreover, pseudo-classic buildings provide a relaxing environment for people in modern urban life, helping them escape the pressures of daily life through their unique aesthetics and cultural atmosphere. In the context of China's rapid urbanization, the imitation of ancient architecture not only serves as a means of preserving and inheriting historical culture but also responds to the psychological needs of modern people.

By utilizing artificial intelligence (AI) to assist in the design of ancient-style buildings, the phenomenon of "thousand cities with the same appearance" can be effectively avoided. Particularly, through the integration of regional characteristics of Chinese ancient architecture, buildings with unique cultural and historical value can be created. AI can analyze large volumes of historical architectural data, extract unique architectural elements and styles from different regions, and ensure that each project reflects distinct local features. It also aids in selecting and optimizing the use of local building materials and traditional craftsmanship. By generating multiple design variants, AI provides more options, helping designers explore different design possibilities while ensuring the accurate incorporation of cultural elements into new designs.

## 2. Literature Review

### 2.1. Cultural Significance of Traditional Chinese Architecture

Traditional Chinese architecture is unique and has a long history, dating back to the Zhou Dynasty, which was 2500 years ago [1]. Compared to modern architecture, traditional Chinese architecture emphasizes a harmonious coexistence with the natural environment. Simultaneously, there exists a tense competitive relationship between traditional architecture and modern architecture [2]. Traditional Chinese architecture embodies a wealth of cultural values, and in the context of China's rapid social transformation, protecting traditional Chinese architecture is extremely important for the inheritance of Chinese cultural values [3]. Over the past two decades, China's rapid urbanization has given rise to a large number of buildings with low cultural value and high energy consumption, making "a thousand cities all look alike" in contemporary society [4,5]. The group of traditional Chinese buildings reflects a long history of evolution. Its form, style, and regional culture are closely related, making it an important window for studying Chinese culture [6]. Furthermore, many traditional Chinese architectural elements and decorations have formed typical Chinese cultural symbols and traditional philosophical thoughts [7].

Traditional Chinese architecture centers on the core values of symmetry and balance. By arranging major architectural elements symmetrically around a central axis, it creates a harmonious overall structure. This symmetrical layout reflects the aesthetic concept of balance and coordination pursued in ancient Chinese architecture. The design incorporates the philosophy of the Five Elements (Wu Xing), striving for balance and harmony through the selection and use of materials, colors, and forms that complement or counteract each other. Feng Shui plays a crucial role in traditional architecture, emphasizing the importance of choosing suitable geographic locations, orientations, and environmental layouts to achieve favorable Feng Shui and auspicious energy flow, thereby bringing prosperity and good fortune [8].

Additionally, carvings, patterns, and decorations in architectural elements often embody mythological stories and symbolic meanings, representing wishes for good fortune, wealth, and protection, thus infusing the buildings with unique cultural significance. The use of wooden structures and dougong (interlocking wooden brackets) is a distinctive feature of Chinese architecture. Through the ingenious mortise and tenon technique, these

structures are made solid and stable, showcasing the exquisite craftsmanship of ancient Chinese woodworking. Traditional buildings frequently use local materials such as wood, brick, earth, and glazed tiles, which play significant roles in both structure and decoration while highlighting the unique charm of regional culture.

In traditional Chinese architecture, gardens and buildings often merge to form a harmonious coexistence with the natural environment. Elements such as artificial hills, ponds, pavilions, and plants in the gardens serve as important complements to the architecture, creating a poetic and pleasant living environment. Traditional Chinese architecture, with its unique design concepts, architectural forms, and cultural connotations, displays a rich historical and cultural heritage, reflecting the wisdom and traditional values of the Chinese people [9,10].

## 2.2. Application of AI in Architectural Design

Starting from Disco Diffusion in February 2022 [11], with the widespread application of the Diffusion Model [12], AI technology has brought revolutionary changes to architectural design, providing architects with greater creative space. Based on this technology, AI-generated architectural images can not only resonate with the original design, but also provide new perspectives for architectural design [13]. By constructing a dataset specific to indoor decoration styles and leveraging the robust Diffusion Model as a foundation [14], it enables the efficient generation of interior decoration designs that cater to specific styles and functions. Insights drawn from studies on certain generative AI optimization techniques underscore the significance of enhancing the accuracy of AI-generated designs [15]. AI design tools with enhanced precision can better assist architects in improving design quality and efficiency [16]. Li, Chengyuan et al. discussed the trend in applying generative AI technology in architectural design, particularly the rapid development of deep generative models [17]. Furthermore, a series of AI platforms based on the diffusion model, led by MidJourney, have been extensively incorporated into the architectural design process. This type of AI technology is also referred to as one of the representatives of the Fourth Industrial Revolution [18,19].

With respect to the effects of other studies that utilize the Diffusion Model to optimize interior design [14], it can be concluded that similar methods have received ample empirical validation in interior design, and Midjourney, based on the principle of the Diffusion Model, has been proven able to generate architectural and site images close to their original forms in the context of Islamic architectural heritage [20].

And text-to-image [21] technology can directly generate architectural renderings through textual descriptions of architectural details and image style parameters; while the Image-To-Image technology is based on providing drawings with an architectural perspective to AI, combined with textual descriptions, to produce the final architectural rendering. Chen, J et al. [16] indicate that the use and selection of guiding words not only influence the direction and style of the generated architectural design, but also have a decisive impact on the final design outcome.

Gan, R et al. proposed a novel text-to-image model, iDesigner, for the field of interior design, and the quality of the generated images is enhanced by strategically applying prompt engineering and large language models (LLMs) [22]. Moreover, by using technologies such as image processing, 3D CAD-2024 design, and additive manufacturing, the images recommended by AI can be transformed into tangible, manufacturable projects [23]. Similarly, AI is used in the preliminary sketch design of high-rise buildings, generating sketches based on user's geometric-shape preferences and building-integration preferences, while preserving creativity and diversity [24]. This preliminary design sketch can also be colored and blended by a novel Y-shaped Generative Adversarial Network (GAN) applied to the given architectural sketch. This new type of Y-shaped GAN, through embedding an attention mechanism model, colors the sketch while maintaining a state of triangular balance [25,26].

AI technology not only includes design optimization and architectural performance prediction, but also has rendering capabilities comparable to those of professional software.

The application of AI in design tasks can not only assist designers, but also improve the entire design process [27]. Even when design requirements are not yet clear, AI is capable of overcoming challenges within a defined space [28]. Meanwhile, not only in the realm of 3D design, but also during the 2D design phase, AI can play a role. There are currently several platforms available online, such as Finch, Mnml AI, and Maket.

The construction details of traditional architecture will indeed have an impact on AI, and the study of traditional architecture will also affect how AI interprets other buildings [29]. Based on the outstanding rendering and creative capabilities of AI, the architect Kaveh Najafian used the AI program Midjourney, with iconic architecture as the theme, to explore the possibilities of designing that architecture in other architectural styles. They challenge the inherent impression of traditional architectural singular forms, showing the dynamic, changeable, and innovative potential of architecture [30]. Architect Rolando Cedeño de la Cruz also utilized AI to reimagine the ziggurat of Mesopotamia, blending elements of ancient and modern, such that the new ziggurat design retains the mystery and solemnity of ancient architecture while infusing the briskness and brightness of modern art, displaying a unique aesthetic of the combination of the past and present [31].

In terms of interior design, Chen, J. et al. also optimize interior design effects using architectural model fine-tuning [32]. They proposed a new improved aesthetic diffusion model for generating aesthetically pleasing interior designs in batches. This model combines diffusion models with a semantically diffusion-guided AI architectural design workflow, enhancing the practicality of diffusion models in the field of interior design. Designers can obtain corresponding results by inputting design requirements in text form. This method provides a novel design approach for interior designers.

In the field of historical architecture, scholars have explored the performance of AI in the context of architectural heritage [33]. The AI-embedded teaching model has a positive impact on student learning, especially in terms of “innovative ability” and “work efficiency” [34,35]. Sukkar et al. [20] discuss the limitations of AI generation tools such as Midjourney and how they influence the definition of Islamic architecture in AI. A small number of studies have explored the current state and future outlook of AI in architectural design [36]. Additionally, the value of AI in assisting designers to solve more complex problems was explored [32]. At the same time, the relationship between some famous architects, such as Antoni Gaudí, and designs generated by AI has also been explored [37]. The results of semi-structured interviews with 16 experienced architectural designers also demonstrated that these models contribute to enhancing creativity, visualization, and imagination, especially in the early stages of design. However, there are also difficulties and potential threats, necessitating further improvements to the tools to adapt to the field of architectural design [38].

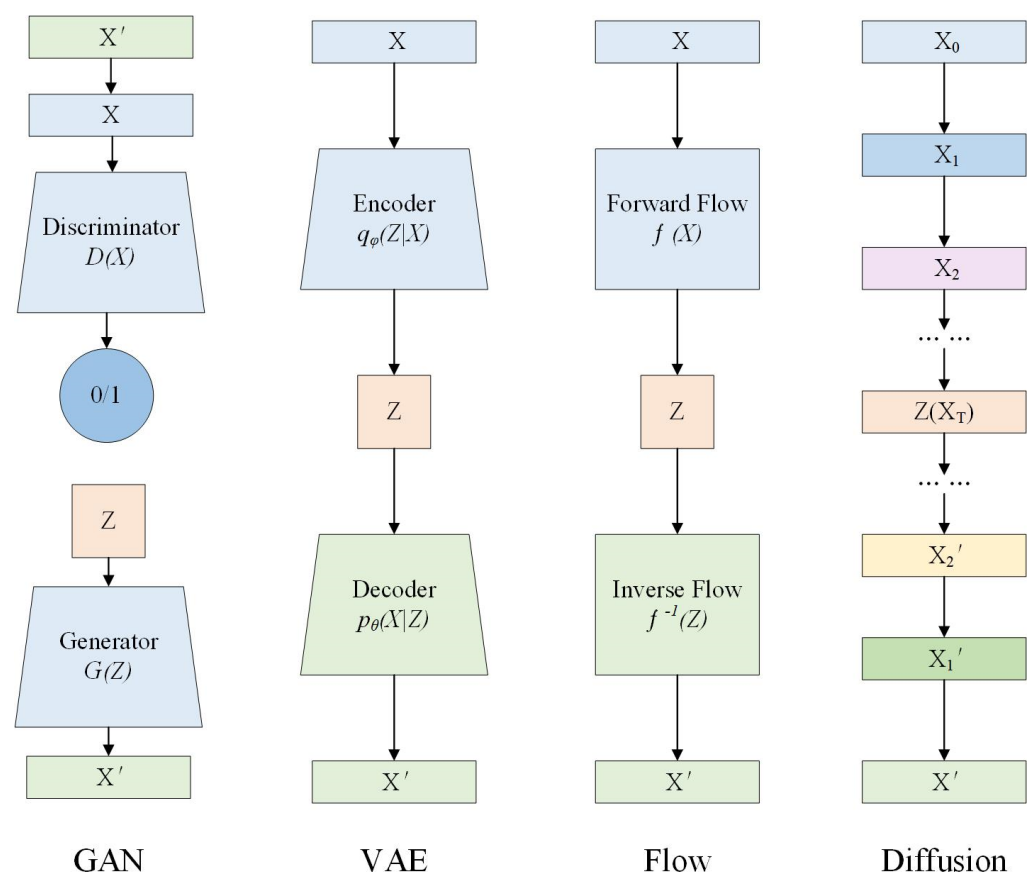
As indicated, AI can assume a crucial role in the field of architectural design education [39]. Establishing an understanding of AI among designers is instrumental in driving the development of education. It supports the fostering of personalized academic learning paths, the cultivation of collaborative learning platforms with educators, and the realization of simulations and scenario construction [40,41]. For the realm of architectural design, such cognitive establishment is of paramount importance [42].

The impact, challenges, and future prospects brought about by AI image generators have become a focal point of current discussions [43]. AI still falls short in understanding key constraints in architectural design, such as plot area, height restrictions, and design regulations [44]. Although AI offers exciting new avenues for architectural creativity, its practical application still requires the formulation of creative solutions to address the aforementioned complex issues [17]. The currently commercially available text-to-image systems have issues with creativity and appropriateness when generating images, which may limit their effectiveness in engineering design [45].

### 2.3. The Principles of Generative AI

Diffusion models, also widely known as diffusion probabilistic models, are a family of generated models that are Markov chains trained with variational inference. The logic of this technology is iteratively refining noise into a vector representation of high-resolution images under the guidance of text prompts. Internally, the prompts are tokenized and encoded into vector representations using the CLIP (Contrastive Language–Image Pre-training) neural network. Guided by the text representation, Stable Diffusion progressively refines the vector representation of the image by using the U-Net (Convolutional Networks for Biomedical Image Segmentation) neural network to predict and eliminate noise, thereby enhancing image quality and adherence to the prompts. Finally, the image representation is upscaled into a high-resolution image [46].

Initially the inspiration came from nonequilibrium thermodynamics techniques, which introduced a novel high-resolution image synthesis approach, termed denoising diffusion probabilistic models [12]. It allows the models to generate images of a specific subject, given a few reference examples. This approach can greatly reduce computational resources while retaining image details. This method can be used for generating various types of images, including text-to-image synthesis, unconditional image generation, class-conditional image synthesis and super resolution. This page also provides some experimental results, demonstrating that this method can greatly reduce computational resources, while maintaining high quality (Figure 3).



**Figure 3.** Generative strategies of diverse generative models.

In summary, while all four models aim to generate new data, they achieve this goal through different mechanisms [47]. Generative Adversarial Networks (GANs) exhibit certain limitations in text-to-image generation tasks. Although conditional GANs can achieve some degree of association between text and images by introducing text embeddings, they still suffer from unstable generation results and difficulties in precisely controlling

the correspondence between text and images. Particularly when generating images with specific architectural styles or themes, GANs often struggle to capture subtle text differences, leading to inconsistencies between generated images and textual descriptions. VAEs rely on a proxy loss, and flow models require specialized architectures to build reversible transformations.

In contrast, Diffusion Models demonstrate stronger potential in text-to-image generation tasks. By gradually recovering images from noise, Diffusion Models can generate high-quality and diverse images. More importantly, the text embeddings in Diffusion Models are tightly coupled with the image generation process, enabling the model to produce images that are highly consistent with text descriptions. Especially when handling fine-grained text prompts commonly found in architectural design, Diffusion Models can generate images with rich details and stylistic variations. Additionally, Diffusion Models can effectively avoid mode collapse and maintain the ability to generate images from different categories through their intrinsic regularization mechanism. By introducing category and theme information during training, Diffusion Models can generate architectural images with specific styles or themes, while preserving an overall understanding of the category [48].

The training pseudo-codes for the Diffusion Model and the GAN model are shown in Algorithms 1 and 2. Their differences can primarily be discussed from the perspectives of generation method, loss function, and training objective:

---

**Algorithm 1.** Pseudo-code for Training a Diffusion Model

---

**Input:**

Pre-trained Stable Diffusion model

Subject images ‘subject\_images’

Class images ‘class\_images’

**Output:**

Fine-tuned model

1: Load pre-trained Stable Diffusion model;

2: Load images of subject and class;

3: Generate text embeddings for subject and class using text encoder;

4: **for** ‘epoch = 1, 2, . . . , num\_epochs’ **do**

5:     **for** ‘subject\_image’, ‘class\_image’ in zip (subject\_data, class\_data) **do**

6:         Encode subject and class images into latent space;

7:         Add noise to the latents;

8:         Perform denoising step on the noisy latents using subject and class embeddings;

9:         Calculate loss for subject and class;

10:         Add regularization term to the class loss;

11:         Backpropagate and optimize;

12:     **end for**

13:     Print epoch loss;

14: **end for**

15: Save the fine-tuned model;

---

**Generation Method:** the GAN generates fake samples through the generator and uses the discriminator to distinguish between real and fake samples, whereas the Diffusion Model generates images through a denoising process in the latent space.

**Loss Function:** the GAN updates the model using both the generator’s loss and the discriminator’s loss, while the Diffusion Model updates the model using the loss of the subject and class images, as well as the loss from the denoising diffusion process.

**Training Objective:** the goal of the GAN is for the generator to produce fake samples that are indistinguishable from real samples, while the goal of the Diffusion Model is to generate accurate and expected images through the denoising process.

When applied to architecture, combining the latest Diffusion Model technology with previous image-generation techniques can bring a completely new workflow to the entire architectural design process. Nowadays, there is research [49] proposing a workflow



that interconnects different deep learning models, such as GANs (Generative Adversarial Networks) [50] and NLMs (Normalized Least Mean Squares), to decompose and handle multiple levels and features of architectural design, to address the current limitations of language model-based image generation.

---

**Algorithm 2.** GAN Training Process

---

**Input:** Real samples  $S_{\text{real}}$ , Generator  $G$ , Discriminator  $D$ , Number of epochs  $\text{num\_epochs}$

**Output:** Trained Generator  $G$ , Trained Discriminator  $D$

1. **Initialize** Generator  $G$  and Discriminator  $D$
2. **Define** loss function and optimizers for  $G$  and  $D$
3. **for** epoch = 1, 2, . . . ,  $\text{num\_epochs}$  **do**
4.   **for each** batch of real samples  $S_{\text{real}}$  **do**
5.     Generate random noise  $z = \text{torch.randn}(\text{batch\_size}, \text{noise\_dim})$
6.     Generate fake samples  $S_{\text{fake}} = G(z)$
7.     **Train Discriminator D:**
8.       Calculate loss for real samples:  
 $\text{loss\_D\_real} = \text{loss\_function}(D(S_{\text{real}}), \text{real\_labels})$
9.       Calculate loss for fake samples:  
 $\text{loss\_D\_fake} = \text{loss\_function}(D(S_{\text{fake}}.\text{detach}()), \text{fake\_labels})$
10.       Total Discriminator loss:  
 $\text{loss\_D} = (\text{loss\_D\_real} + \text{loss\_D\_fake})/2$
11.       Update Discriminator  $D$  parameters:  
 $\text{optimizer\_D.zero\_grad}()$   
 $\text{loss\_D.backward}()$   
 $\text{optimizer\_D.step}()$
12.     **Train Generator G:**
13.       Generate fake samples  $S_{\text{fake}} = G(z)$
14.       Calculate Generator loss:  
 $\text{loss\_G} = \text{loss\_function}(D(S_{\text{fake}}), \text{real\_labels})$
15.       Update Generator  $G$  parameters:  
 $\text{optimizer\_G.zero\_grad}()$   
 $\text{loss\_G.backward}()$   
 $\text{optimizer\_G.step}()$
16.     **end for**
17.     **Output training loss**  $\text{loss\_D}, \text{loss\_G}$
18. **end for**

**Return:** Trained Generator  $G$ , Trained Discriminator  $D$

---

In this experiment, the authors employed two methods to validate their workflow. The first method involved using individual NLMs (Normalized Least Mean Squares) to test a design scenario, comparing the impact of different text prompts and visual references on the results. The second method utilized a hybrid workflow, combining NLMs (Normalized Least Mean Squares) with other GANs (Generative Adversarial Networks), such as StyleGAN [51] and CycleGANs (Cycle-Consistent Generative Adversarial Networks) [52], using a specific architectural dataset for feature decoupling and domain transformation. The authors presented the results of both methods and analyzed the subjectivity and creativity of designers in different design tasks and processes, and explored how this workflow supports creative thinking and design innovation.

Regarding the model training section, research on fine-tuning diffusion models abroad stems from the method proposed by Imagen for developing custom text-to-image diffusion models. Foreign research on finetuning Diffusion Models originates from the method of customizing text-to-image Diffusion Models proposed by Imagen research and development.

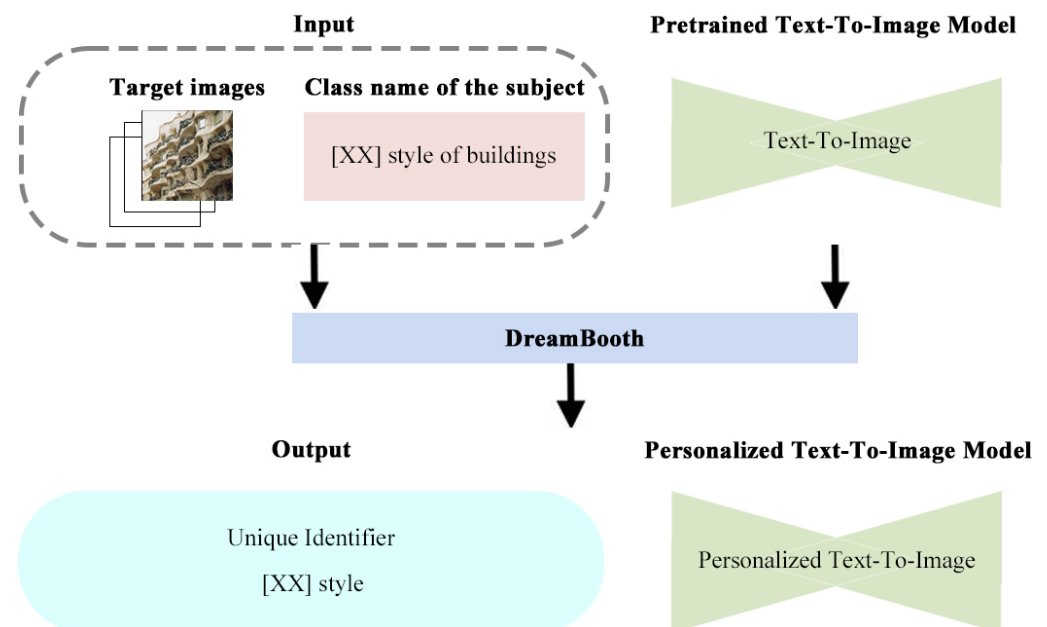
Besides training Diffusion Models, researchers at Stanford University have studied methods to assist image generation in order to generate results that conform to the constructed models, to aid architectural design and computer vision [53]. They give a new framework that leverages lightweight adapters to enable precise controls over pre-trained models. It enables the input of additional conditions, such as edge maps, segmentation maps, and key points, and can be trained on a small dataset in an end-to-end manner [54]. This method is used to control large Diffusion Models to facilitate the input of additional conditions. By leveraging various conditioning techniques, such as boundary map control, pose control, segmentation map control, and normal map control, it can enhance the flexibility and diversity of image generation. Compared to general text-to-image or image-to-image [55], the differences between the two are as follows.

ControlNet's current pre-trained models are mature and applicable to more scenarios, being able to support more kinds of condition detectors (nine major categories), while the text-to-image and image-to-image adapter is designed and implemented in an engineering approach that is simpler, more flexible, and easier to integrate and extend.

With ControlNet, text-to-image and image-to-image can be guided by more than one condition model. For example, it can use both depth and a segmentation map as input conditions at the same time, or use a sketch as guidance within an inpainting mask region.

### 3. The AI Training Method

A thorough understanding of the training principles (Figure 4) is crucial for mastering the method to enhance architectural rendering effects. Inspiration for this research method partly comes from existing research by Han, Q et al. on AI learning of traditional Chinese architecture [56]. Building upon this inspiration, the present study employs CLIP (Contrastive Language–Image Pre-Training)-based contrastive learning in DreamBooth to validate its learning capability for traditional Chinese architecture.



**Figure 4.** Schematic diagram of DreamBooth training.

The principle of DreamBooth is to form a new language–visual dictionary for Diffusion Models, linking new vocabulary to specific themes that users want to generate. It is to embed images of a given theme into the output domain of the model, so that unique identifiers can be used to synthesize them, to reduce the issue of similar AI models in Figure 5 being unable to correspond to specific architectural styles.



Figure 5. Model response effect before training.

In order to distinguish the model effects before and after architectural fine-tuning, the architectural training set of this session needs to meet the following two conditions (Figure 5):

- The original model has a poor response to this design style;
- The images used for training should have a clear and distinct design style;

### 3.1. Integration of Architectural Training Datas

The Huizhou architecture dataset utilized in AI training should encompass a rich array of distinctive features that are emblematic of this traditional Chinese architectural style. This includes the iconic horse-head walls, intricate Huizhou-style carvings, and the characteristic white walls and black tiles that define the aesthetic of Huizhou buildings. Additionally, the dataset should emphasize the unique group layout of these structures, showcasing how individual buildings are arranged in harmony to form cohesive communities (Figure 6).

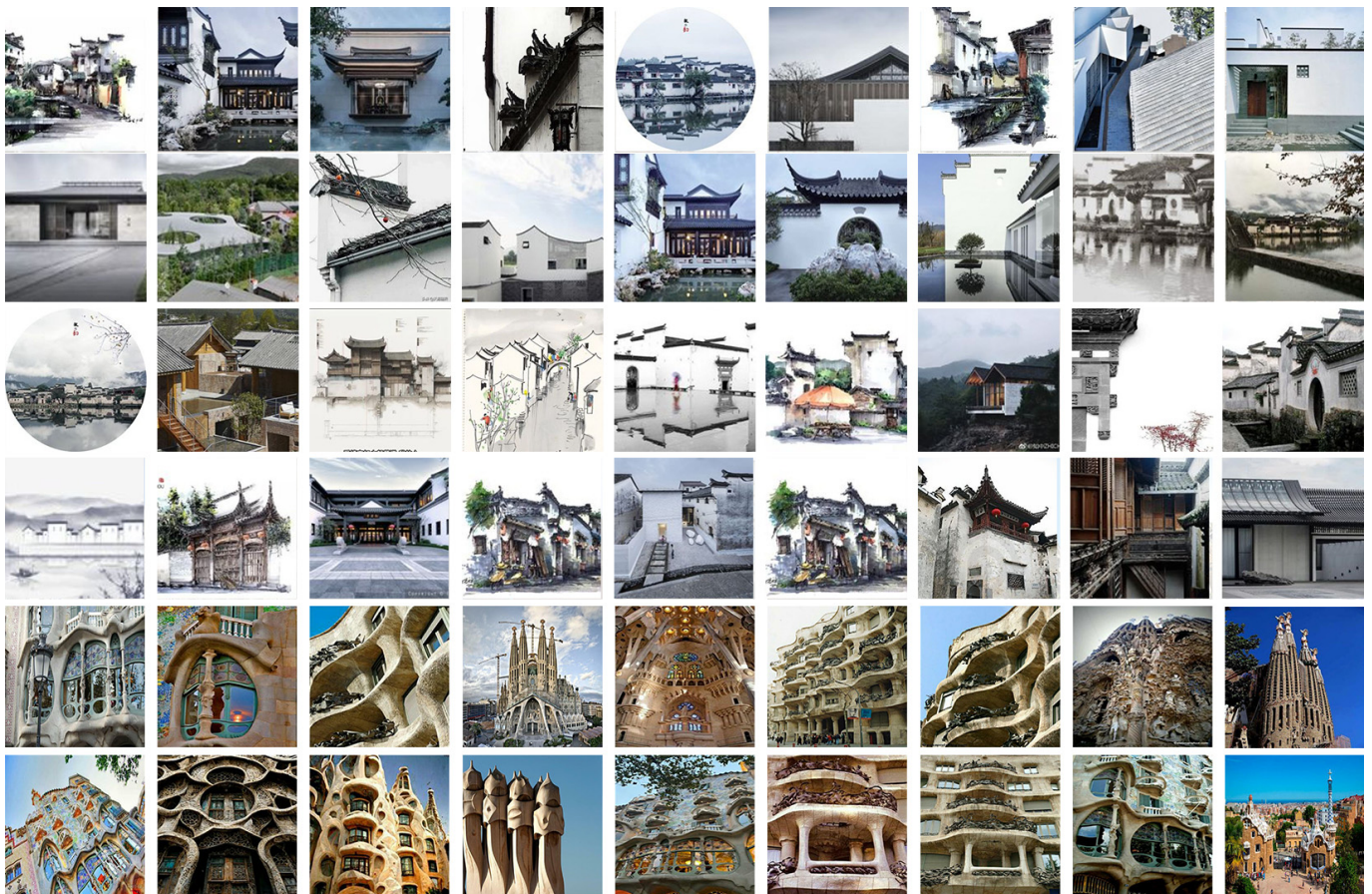


Figure 6. Partial display of two data sets.

As shown in Figure 6, during the dataset construction process, to address the poor response of the original model to traditional Huizhou architecture and Antoni Gaudí architectural styles, this research collected over 50 and 150 images, respectively, related to these two architectural styles, from the Pinterest website. Each image was meticulously annotated with its corresponding architectural style and spatial function in a text file. Thus, the “Traditional Huizhou Architecture and Antoni Gaudí Dataset” was established.

### 3.2. Analysis of the Model-Training Convergence Process

The classification method for the dataset is also explained in Figure 7. Figure 8 diagrams are intended to illustrate the key steps and concepts in the training process, enabling architects to have a clearer understanding of the principles and methods of model training.

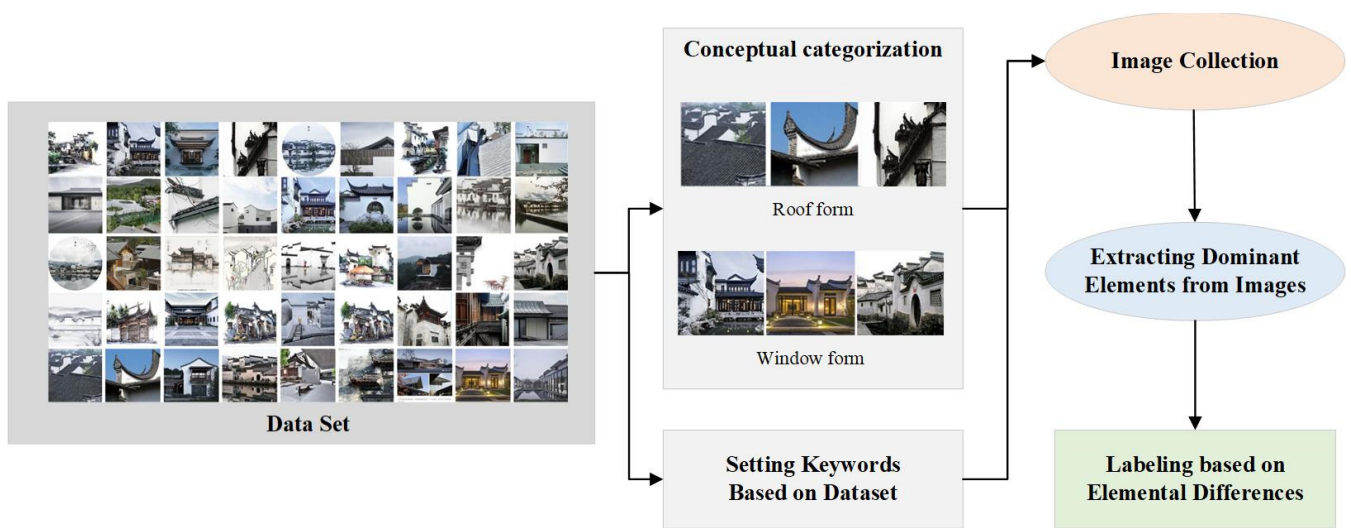


Figure 7. Classification methods for datasets.

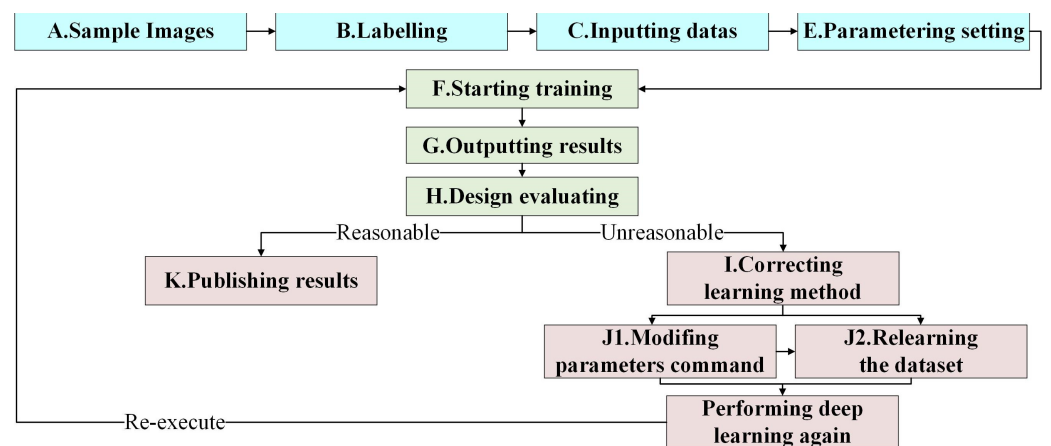


Figure 8. Inputting data, parameter setting, design evaluation, and parameter command.

This study utilizes the open-source project “sd\_dreambooth\_extension” available on GitHub, for training on the WEBUI interface (Figure 9).

By training according to the training values specified in the parameters mentioned above and recording the loss and learning rate (lr) at each stage, the resulting graph curve graph is shown in Figure 10. As seen from the graph, with the increase in epochs, the loss values also begin to converge.

Figure 10 indicates that as the number of training iterations increases, the model’s understanding of the training set deepens, and the feature learning becomes more refined,

resulting in a decrease in loss values and convergence towards a stable state. This result further demonstrates the improvement of the model in the learning process during this study. Additionally, in the later stages of training, the text learning rate and the Unit learning rate gradually decrease. This is to prevent potential overfitting of the model to the training set, due to excessively high learning rates.

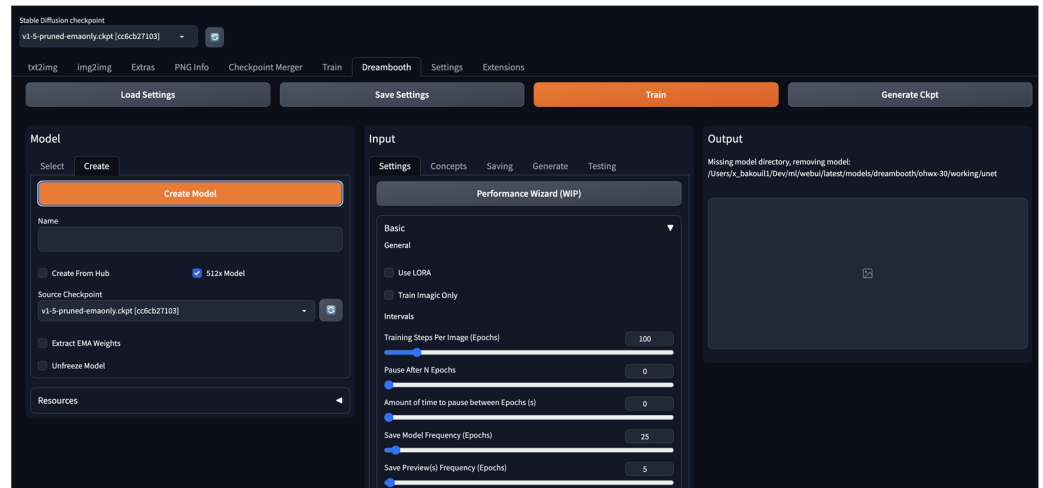


Figure 9. Software used in the paper: Stable Diffusion Webui Version 1.5.2.

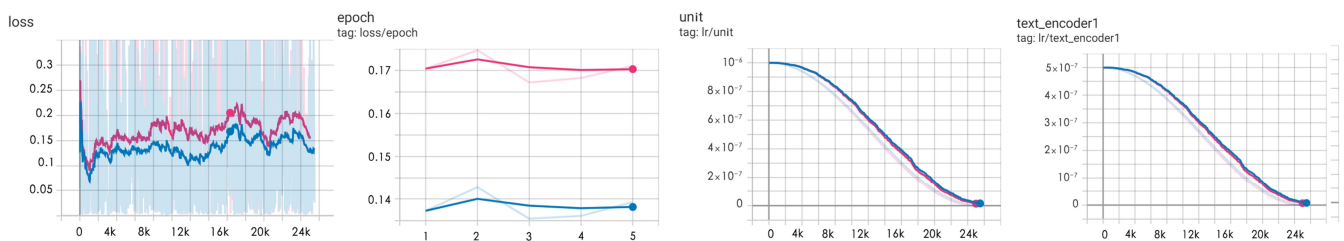


Figure 10. Curves of the variation in loss values and learning rates during training.

In Figure 11, it is apparent that as the number of training steps increases and the number of epochs gradually rises, the model’s fit to the training sets of Antoni Gaudí and Huizhou architecture continuously and steadily improves. This trend indicates that, over the long-term training process, the model in this study gradually enhances its ability to recognize and understand these two architectural styles.



Figure 11. The training performance corresponding to each epoch.

For each epoch stage in this study, a model was set up, and consistent parameters were used to generate the models at each stage for comparison. This approach allows for a clearer observation of the process by which the model gradually learns the features of the training set. Particularly when directly comparing the effects of the first Epoch with those of the fifth Epoch in Figure 11, the difference in the degree of fit becomes especially pronounced. These two epochs represent the results before and after training, respectively, clearly demonstrating the model's progressive improvement in its ability to recognize the types of architecture in the training set during the learning process.

### 3.3. The Selection of Experimental Subjects in DreamBooth

The training platform and AI model (non-Chinese origin) need validation for traditional Chinese architecture. To verify, select an internationally renowned, highly distinctive architectural style. A well-known style aids AI understanding, while a distinctive style helps assess research method effectiveness. Rapid transformation of a generic building into a specific style will demonstrate efficacy. The validation process allows optimization of the technical roadmap through parameter adjustments, establishing a robust foundation for future applications in traditional Chinese architecture.

From a logical perspective, it would be most reasonable to initially select a traditional Chinese architectural style. However, the various styles of traditional Chinese architecture do not reflect preferences of individual architects, but rather regional architectural characteristics [57]. It was not until the 1920s that China saw the emergence of a professional architect in the modern sense, when a group of Chinese students who had studied architecture at the University of Pennsylvania in the United States returned to China [58]. Prior to this, traditional Chinese architecture was predominantly designed by craftsmen, whose names were mostly unknown.

Therefore, it is challenging to select a group of internationally renowned and highly unique styles within traditional Chinese architecture to initially validate the effectiveness of our research method. Choosing Antoni Gaudí as our first example in this study, is due to the following reasons:

- (1) Although Gaudí's career spans different stylistic periods, his works exhibit a unique "architectural marvel", with a strong personal style. In world architecture history, Gaudí's works are unparalleled. Thus, we use Gaudí's works for AI model training and experimentation to intuitively assess our methodology's effectiveness.
- (2) Gaudí's era marked the transition from classical to modern architecture, blending both elements in his works. Similarly, this study aims to present Chinese traditional architectural styles using AI technology, inevitably incorporating modern features due to functional modernization. This fusion parallels Gaudí's work.

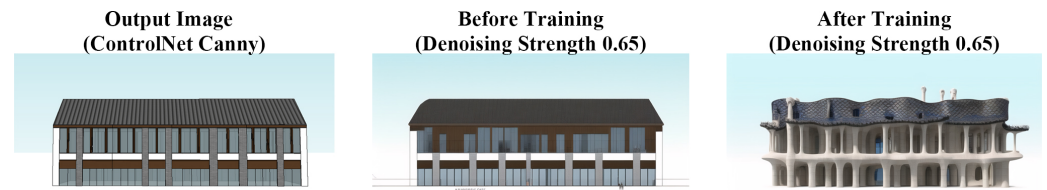
## 4. Results

It is necessary to conduct preliminary experiments on the research methods proposed in this paper to verify that the method can reflect the cultural characteristics of the region in the preliminary design stage of architecture, and to see whether it can effectively improve the accuracy and controllability of AI-rendered images in the current environment.

The CFG scale is set to 7.0 to enhance the persuasiveness of the results. The default number of Steps for image Diffusion is 20.

### 4.1. Visual Evaluation Metrics

Firstly, it is essential to evaluate the potential of DreamBooth in image style restoration, because if its capability in style restoration is inadequate, discussing its practical assistance to architects becomes meaningless. Therefore, this paper validates the actual effect of DreamBooth through a comparative study of the architectural art styles of Antoni Gaudí. Details regarding the response of the original model to this design style can be seen in Figure 12.



**Figure 12.** Comparison of elevation renderings.

When this study assesses the results after using the basic untreated pre-trained model, it can be observed that although the output architectural style diagram responds to the embedded text identifiers such as “building facade diagram” in the instructions, the response to the “by Antoni Gaudí” text identifier is not ideal (Figures 5 and 10). In contrast, after evaluating the results by the fine-tuned pre-trained model, it can observe that the effect diagrams are closer to the curved, arched and arc features that are unique to Antoni Gaudí’s architectural works (Figure 12). This phenomenon indicates that the model is a better fit and match for the “by Antoni Gaudí” text identifier. It can generate novel variant architectural effect images according to the dataset used for training, while retaining the features of the dataset (Figure 6).

However, traditional AI rendering workflows often result in architectural renderings that deviate significantly from the expected images in terms of local details. Therefore, enhancing the controllability of renderings becomes a crucial concern.

#### 4.2. Precision and Controllability Enhancement Comparison

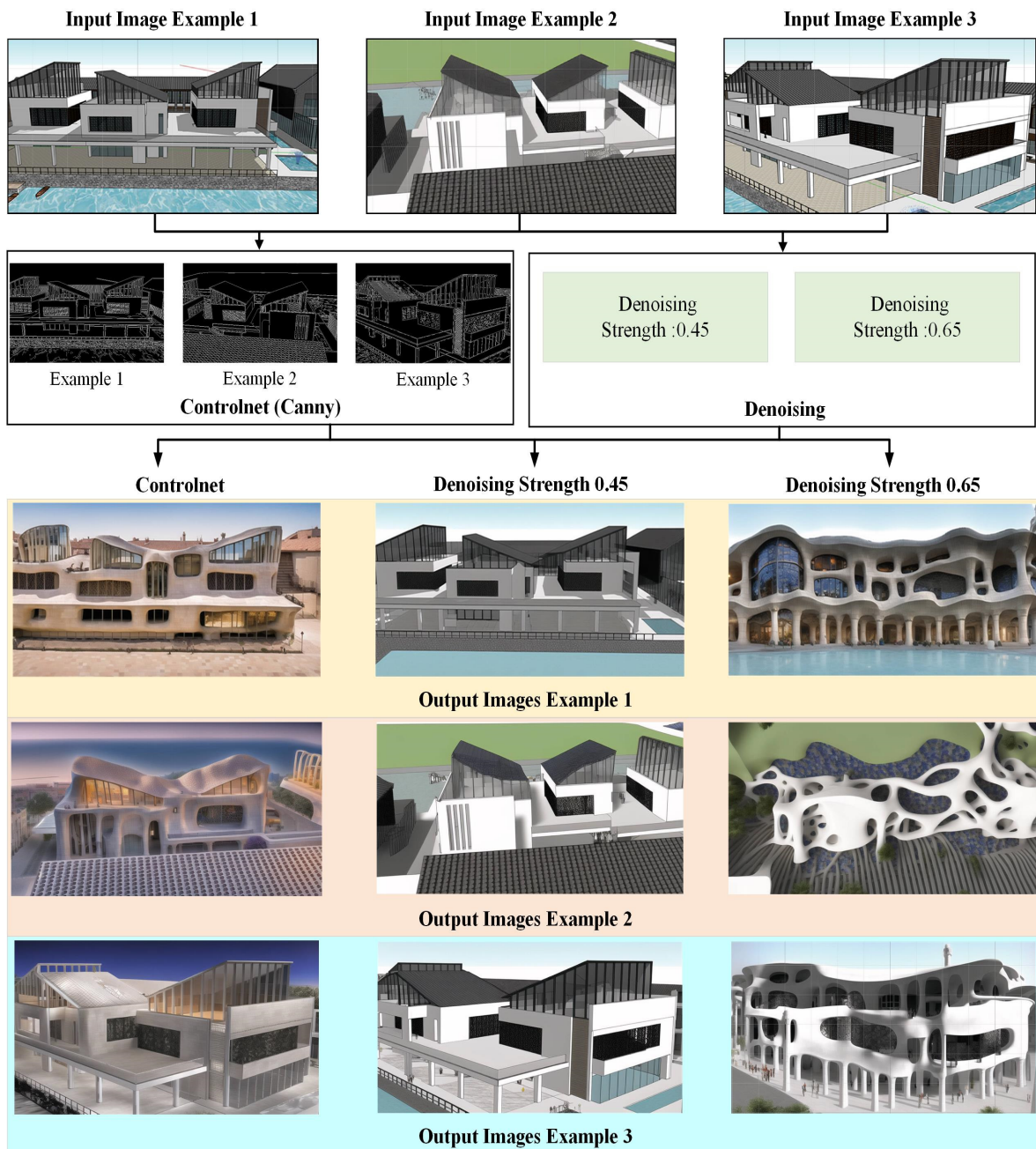
The previous practical workflows may lead to diffusion and fusion of elements. This may not achieve the desired effects for architectural renderings that require high accuracy in the later stages of design. For this reason, this paper introduces the new technology ControlNet, and in practical applications, its effectiveness is validated by evaluating its accuracy in mapping features of the original images (Figure 13).

To highlight the importance of ControlNet in the preliminary design of buildings, Figure 12 also presents a comparative analysis on results of using and not using ControlNet.

Firstly, when denoising the original images solely using Image-To-Image image-to-image without employing ControlNet, a significant difference is observed between the rendered results and the original architectural images when the denoising strength is adjusted to 0.65 (Figure 13). Specifically, there are distortions in architectural details and structural deformations, leading to less practical rendered results. Conversely, when the denoising strength is reduced to 0.45, the generated image shows little difference from the original image, but lacks noticeable rendering effects (Figure 13).

However, the use of ControlNet enhances the accuracy of generated images in depicting scene structure, reducing “jagged” deformations.

Results in Figure 14 result show that MLSD tends to read the outlines of external straight edges in this set of images, while curvy structures are often ignored in the mapping. Therefore, in one of the comparative groups with relatively simple building outlines, MLSD can perform better, but in the other two architectural renderings, the expression of building structures is relatively chaotic on MLSD. The Depth neural network structure will obtain a logical image identifying the foreground and background. It based on white and black colors after mapping the original image information, so the architectural renderings of buildings have a relatively obvious depth of field and vividly restore the shadows of buildings. The generated depth maps clearly show that it focuses on the spatial distance relationships between objects within the image. However, it grasps the details of buildings poorly. In comparison, the Canny neural network structure model can well identify the edge contours of various objects in the three contrast groups in the images.



**Figure 13.** Comparison of before and after effects using ControlNet.

#### 4.3. Evaluation of Traditional Architectural Design

Based on the previous methods, this section will issue clear commands to AI, forming a visualized design process for Huizhou architecture with reduced modeling workload, providing real-time information feedback for architects. Through the objective evaluation of visualization schemes by multimodal AI, the value of each scheme can be judged, thereby enhancing the sustainability of AI technology in Chinese-style buildings.

The initial modeling images of the four groups of input images only include basic modeling of the site and the arrangement of doors and windows. The modeling part is relatively rough, as it does not follow the specific form of Huizhou architecture. The reason behind this is to validate whether AI can provide visually complete designs even with lower modeling accuracy.



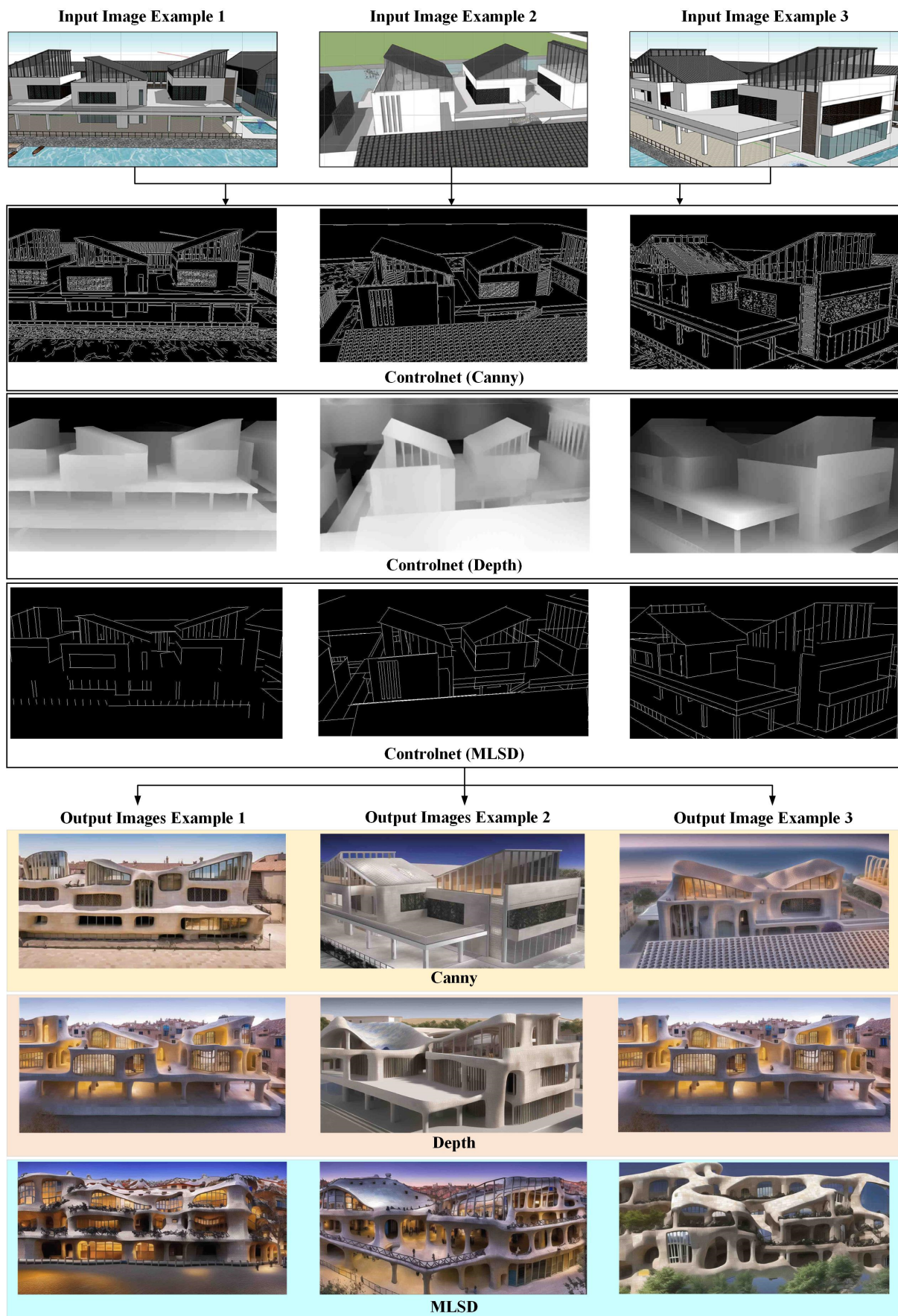


Figure 14. Comparison of effects of different ControlNet preprocessors.

Since displaying all results would make each image too small to be easily identifiable, this paper randomly selected three generated results from each group as examples, to showcase the overall effects (Figure 15).

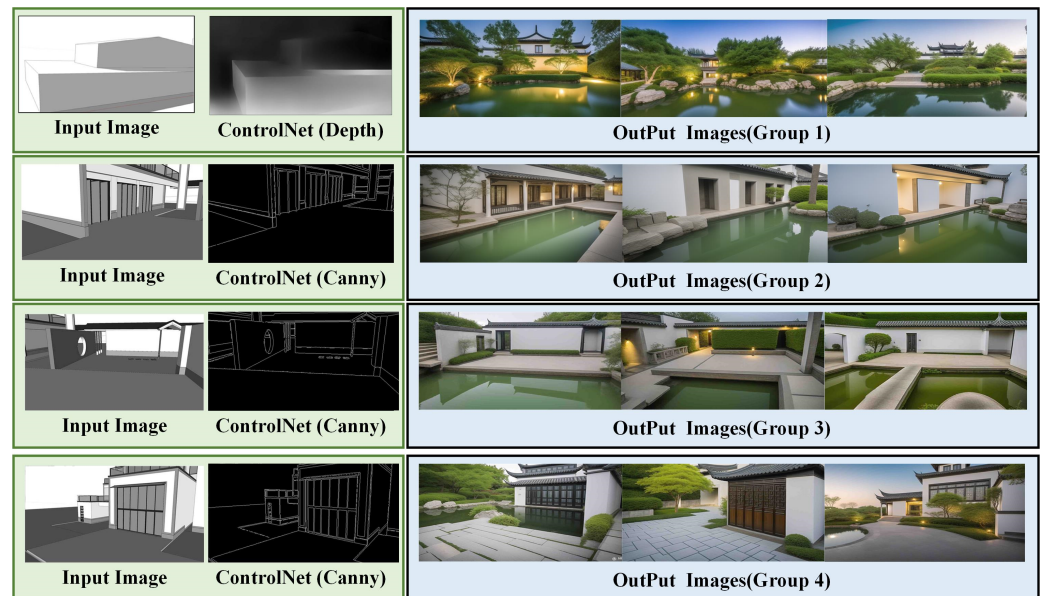


Figure 15. Partially generated visualization of Huizhou architectures.

This study evaluates the degree of heritage and the usability value of the fine-tuned model for Huizhou architecture design through two assessment metrics:

- Assessment of the restoration extent of the Huizhou architecture style and its landscape.
- Evaluation of whether the generated Huizhou architecture designs meet the rationality requirements for spatial structure as stipulated in the original project brief.

This paper uses the CogAgent multimodal model to evaluate the results, which are as follows (Figures 16 and 17):

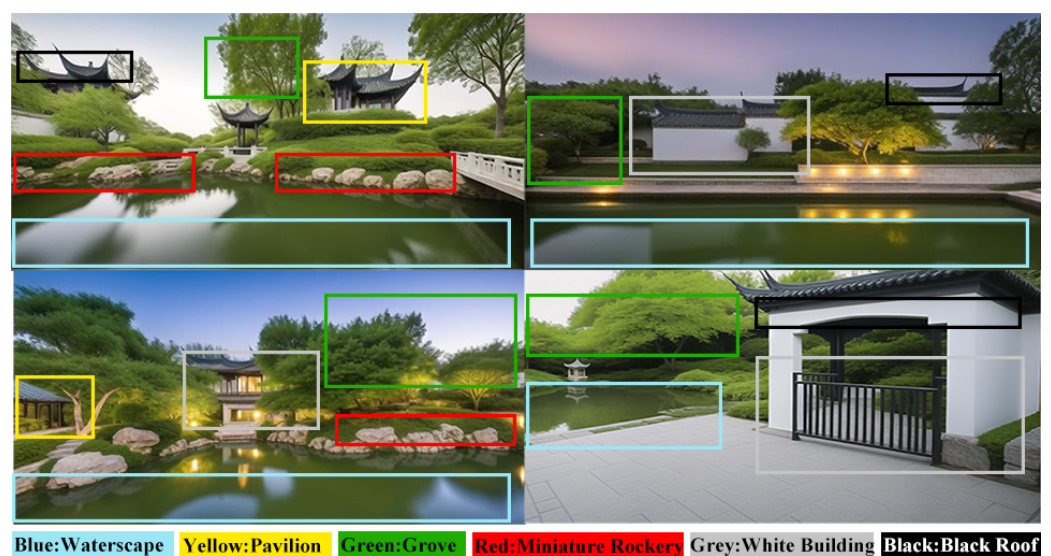
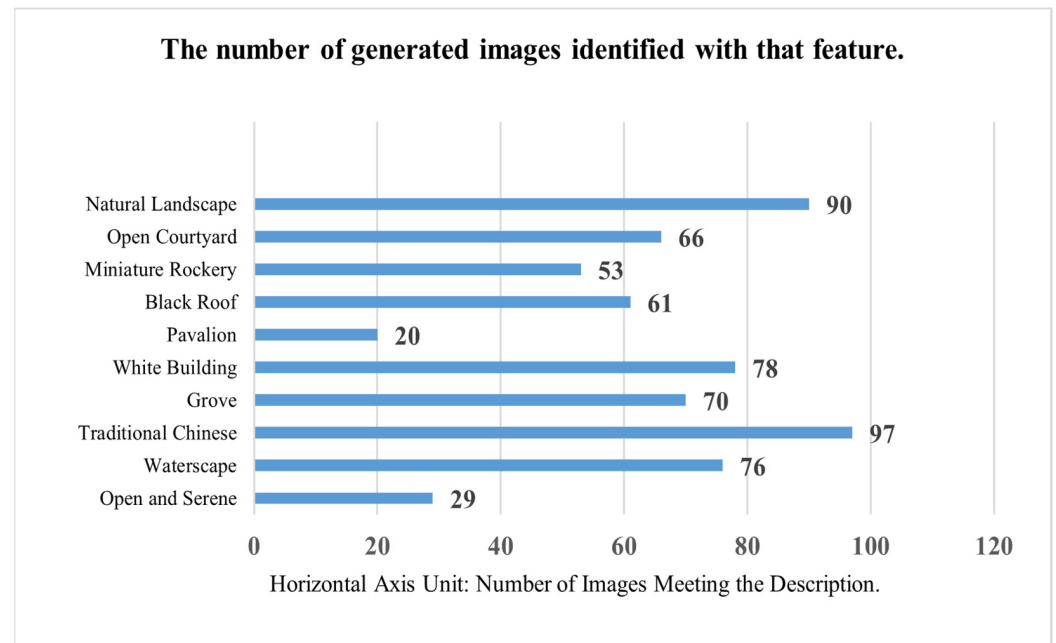


Figure 16. The recognition marking of image features.

The recognition results of the generated Huizhou architecture designs are shown in Figure 17, where the vertical axis represents words that frequently appear in the descriptions

identified by the CogAgent model, and the horizontal axis represents the number of images that meet these descriptions.



**Figure 17.** Recognition frequency of architectural features.

From the quantity statistics chart of Huizhou-architecture design results generated by the CogAgent model (Figure 16), it can clearly be seen that the model has the ability to generate core features of Huizhou architecture after fine-tuning. The model can generate images that reflect key features of Huizhou architecture with a ratio of over 70%, including Natural Landscape, Grove, White Building, etc. The low recognition numbers of certain keywords (such as Black Roof) are primarily due to the smaller proportion of these elements within the overall architecture, leading to fewer recognitions.

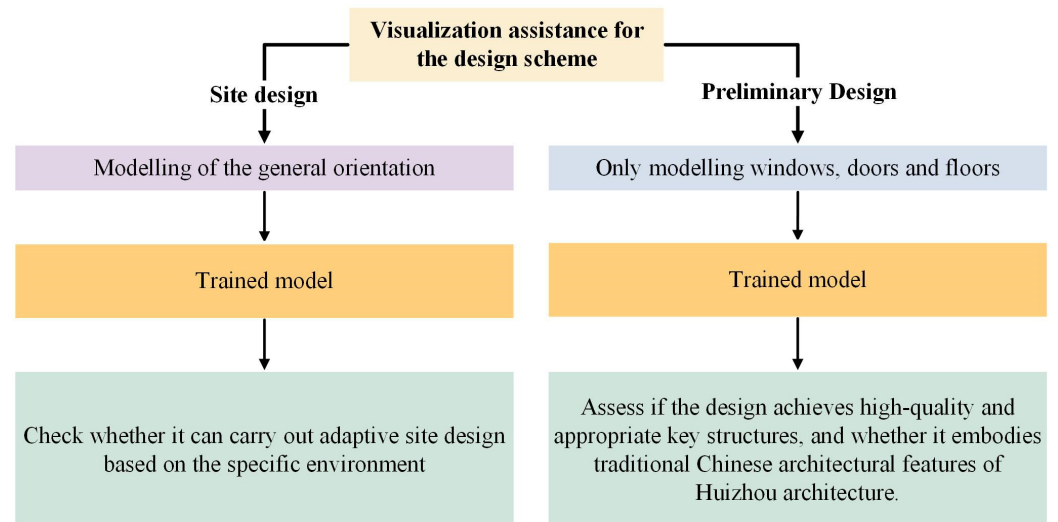
The analysis shows the model can capture Huizhou architecture's core principles: nature integration and Yin–Yang balance. AI successfully integrates these unique features into various aspects, including building shape, structure, materials, and decoration. Rendered images also display these traits in door and window design and cantilevered roof eaves (Figure 16).

## 5. Discussion

### 5.1. Methodological Practices in the Application of Traditional Chinese Architecture

As stated in Section 4, this paper has successfully demonstrated the effectiveness of the proposed method through empirical research. However, supporting the method theoretically is not sufficient. It is crucial to empirically demonstrate its practical applicability within the architect's workflow, as shown in Figures 1 and 2. The method needs to transition from a theoretical level to a concrete form of practical application (Figure 18).

Therefore, to comprehensively investigate its application effects and value, this study focused on a practical application surrounding the ancient-town design project currently being advanced in Huangshan City, Anhui Province, which is part of an old-city revival plan. This verification aims to demonstrate the effectiveness of integrating AI technology with traditional architectural design, enabling the application of traditional architecture within modern architectural processes. The convergence of this technology and culture represents a crucial avenue for achieving long-term sustainable development, aligning with the dual requirements of innovation and continuity in sustainable design, thereby promoting the sustainable inheritance of ancient architectural culture.



**Figure 18.** The verification process from results-based methodology to discussed.

### 5.2. Challenges in Traditional Chinese Architectural Design

In contemporary China, the design of ancient cities faces numerous challenges in the practice of cultural heritage preservation. The complex forms have restricted the development of traditional Chinese architecture. Imagine the modeling time required for these forms, which would be several times that of general modern architecture.

In architectural design, Huizhou architecture embodies several unique design features. These features include the following aspects (Figure 19):



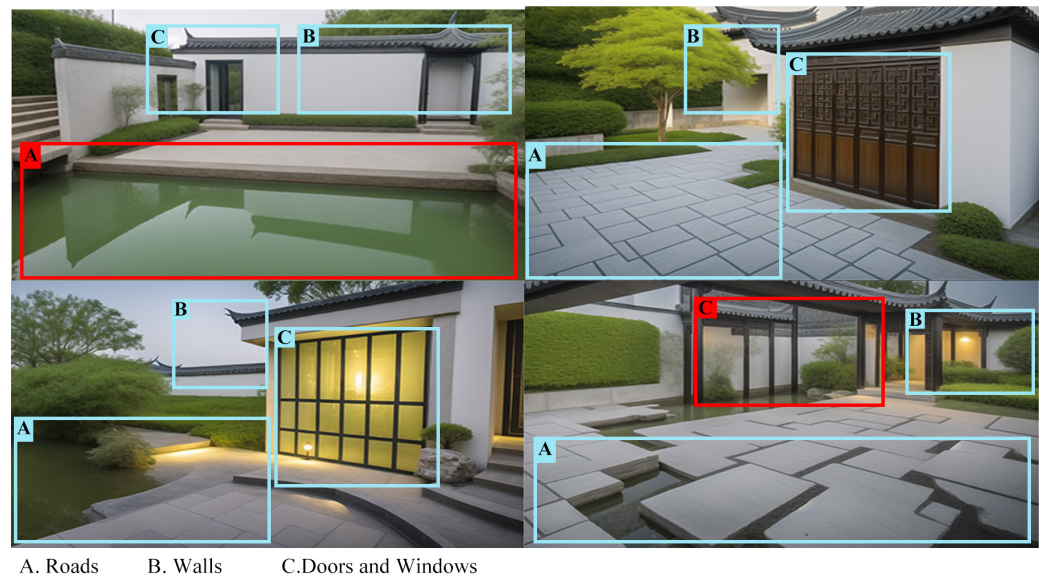
**Figure 19.** The traditional Huizhou-style architecture which reflecting the typical Huizhou architectural style.

Huizhou architecture often integrates with the surrounding landscape, emphasizing the layout based on Feng Shui principles. The village structures are compact, with narrow streets and alleys, presenting a maze-like layout. The buildings predominantly adopt the courtyard compound form. Structurally, Huizhou architecture relies primarily on wooden frames, with distinctive horse-head walls. Additionally, Huizhou architecture boasts a rich array of decorative arts, such as brick carving, wood carving, and stone carving, which can be found throughout. Roof forms include “hard top” (gabled roof), “suspended top” (hipped roof), and “resting top” (gable-and-hip combination), with upturned eaves and covered with small blue–green tiles. Interior design places emphasis on axial symmetry, featuring spacious and bright halls. Furniture often showcases Ming and Qing Dynasty styles.

Huizhou architecture is closely intertwined with clan culture. Architectural structures like ancestral halls and memorial archways embody the honor and status of clans, while academies and schools also hold significant positions. Furthermore, Huizhou architecture prioritizes environmental sustainability. Through design elements such as courtyards, yards, and windows, it maximizes the utilization of natural light and air circulation, enhancing residential comfort. Moreover, local materials such as wood, blue bricks, and stones are extensively used, promoting a harmonious coexistence with the natural environment.

### 5.3. Evaluation of Spatial Rationality

This evaluation primarily focuses on Groups 2, 3, and 4, which, unlike Group 1, have explicit requirements for the architectural structure set in the Input Images (Figure 20). For the parts involving doors and windows, there are clear specifications for the position, size, and even the architectural style.



**Figure 20.** Scoring Demonstration.

In the process of evaluating structural rationality, this research adopted three main architectural structure-location accuracy indicators: the position of doors and windows, the layout of walls, and the design of pavements. These indicators correspond to important elements in architectural design. Precise assessment and scoring of these three critical factors are carried out.

In Group 2, 3, and 4 (Figure 21), each group contains ten Huizhou architecture schemes. These schemes have been scored, and the scoring trend for each scheme is illustrated in the following three tables.

Using some generated Huizhou architecture proposals as examples, this section explains the scoring criteria for architectural structure rationality. The assessment is based on a comparison with the structures in the original Input Images (Figures 20 and 21). Doors and windows, walls, and pavements are evaluated factors, and scores are assigned based on their accuracy and rationality. If any obviously irrational structures are identified during the comparison process, such as parts highlighted by red rectangles, no scores will be awarded for those sections.

Based on detailed analysis of the data in Figure 20, it is apparent that within all evaluated images, the structural accuracy of the walls is undoubtedly the most prominent. This observation is not coincidental, but is proven by consistency shown across 30 sets of generated results. In fact, out of these 30 sets, only one did not receive a score, which speaks volumes about the fine-tuned model's performance in designing complexity and achieving precision.

On the other hand, although the structural accuracy of the doors and windows may not be on par with that of the walls, its performance is still quite commendable. Out of 30 generated results, scores were awarded to 24 sets. Despite being slightly inferior to the performance of the wall section in numerical terms, doors and windows are undoubtedly also an important consideration in the overall assessment.

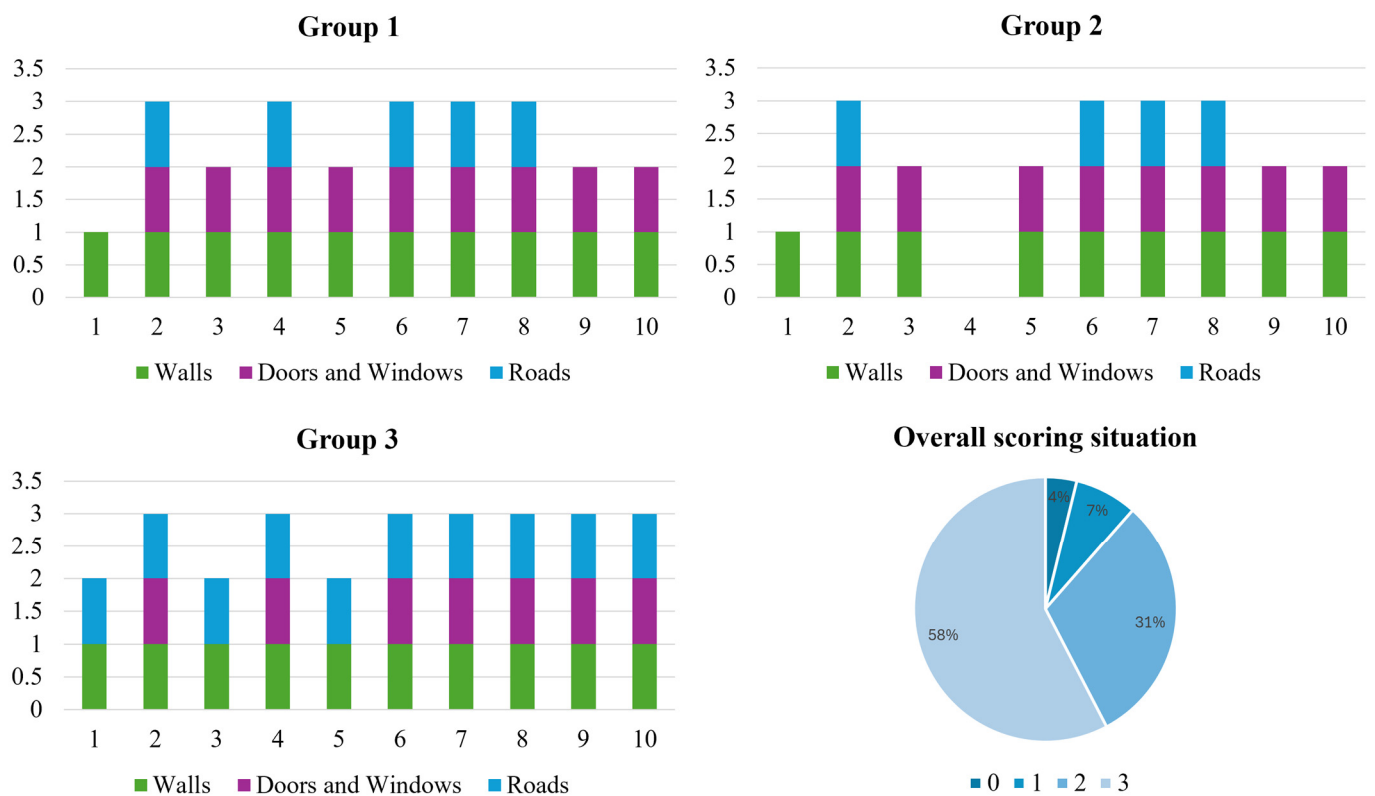


Figure 21. Specific scoring for each group.

#### 5.4. Improving Design Efficiency

As shown in Figure 22, even with significantly reduced modeling accuracy and the exclusion of Huizhou architectural factors, the visualization rendering process in the design phase can still achieve the same visual effect as Huizhou architecture in reality. It shows that even with a limited budget, architects can realize the refinement and functional features of Huizhou architecture style within a small-scale construction, thereby effectively revitalizing China’s ancient techniques, particularly exhibiting the innovation of Huizhou architecture style and presenting a broad and creative architectural vision.

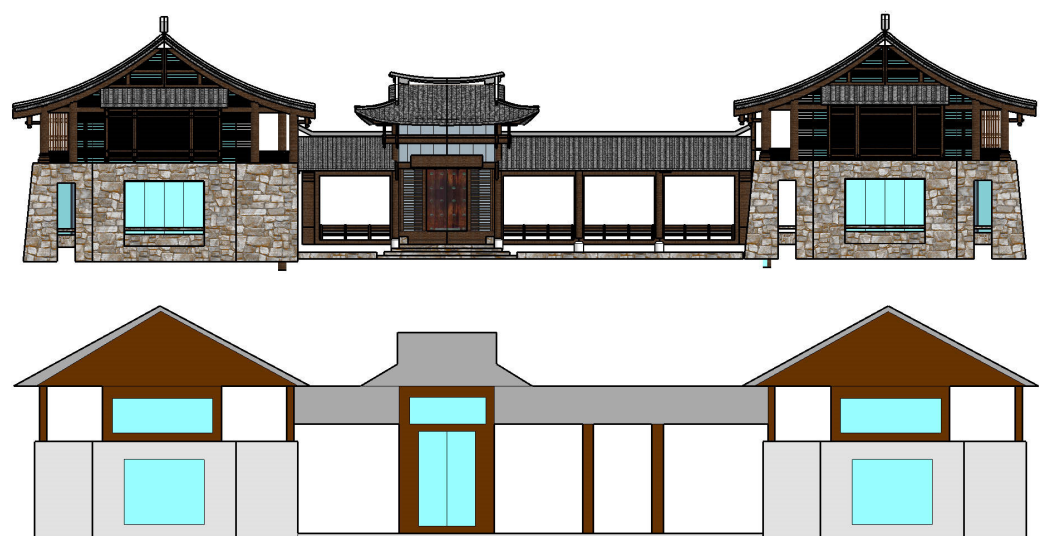


Figure 22. Comparison of modeling workload.

## 6. Conclusions

### 6.1. Application and Value

This research confirms the method's practicality in traditional architecture, providing a theoretical foundation for Huizhou architecture in China. By comparing DreamBooth renderings to Antoni Gaudí's architectural style, we validate DreamBooth's ability to achieve stylized generalization in architecture. ControlNet technology enhances rendered image controllability, reducing distortions and improving precision and authenticity. Real project tasks verify its applicability in traditional design.

This approach integrates AI into traditional Chinese architecture design, simplifying the preliminary process. It introduces new possibilities and addresses the decline of traditional architecture in modern building environments. This helps traditional architecture adapt to contemporary design, positively impacting its long-term development.

Compared to traditional methods, this research significantly simplifies modeling time for traditional architecture. It demonstrates AI's potential in preserving intangible traditional architectural heritage. By deeply learning traditional architectural styles and elements, AI can extract their essence and generate creative design expressions. Furthermore, AI can merge traditional architectural features with modern design philosophies, resulting in green buildings that are both time-specific and conducive to the continuity and innovative advancement of traditional cultural heritage.

### 6.2. Future Research

In contemporary society, the application of AI technology in architectural design has sparked widespread discussion and controversy. Proponents argue that AI can significantly enhance design efficiency by rapidly generating and optimizing design plans, saving time and effort. Additionally, AI contributes to the preservation and restoration of historical buildings, utilizing big data analysis to uncover patterns in traditional design, promoting innovative designs, and providing personalized design solutions. AI can also handle large numbers of data and complex computational tasks, making the design process more precise and scientific, thereby improving overall design quality.

However, opponents worry that AI may overlook the cultural connotations of traditional Chinese architecture, leading to a break in cultural heritage. They believe that AI lacks human creativity and artistic sense, and cannot fully replace human designers. Moreover, over-reliance on AI may result in the degradation of designers' skills, and ethical issues concerning data privacy and intellectual property need to be addressed. The focal point of the debate is whether the design quality generated by AI can meet or surpass the level of human designers, and whether AI can truly understand and adapt to the architectural styles and needs of different cultures.

Although AI holds great potential in traditional architectural design, how to balance technological innovation with cultural heritage requires careful consideration. While technological development is inevitable, its application must respect and protect traditional culture, ensuring the harmonious development of technology and a humanistic spirit. Meanwhile, designers should continuously enhance their skills and collaborate with AI technology, leveraging each other's strengths to advance and innovate in architectural design. Only on this basis can AI technology truly bring revolutionary changes to architectural design while avoiding the destruction of traditional culture.

Studying and applying the latest deep learning techniques involves using conditional generative adversarial networks (cGANs) [59], Variational AutoEncoder (VAE) [60,61] and U-Net (Convolutional Networks for Biomedical Image Segmentation) [62,63] to enhance the model's expression ability and generation effect.

**Author Contributions:** Conceptualization, Y.L. and F.C.; methodology, F.C.; software, F.C.; validation, F.C., Y.L. and M.M.; formal analysis, F.C.; investigation, F.C.; resources, Y.L.; data curation, M.M.; writing—original draft preparation, F.C. and X.H.; writing—review and editing, Y.L. and X.H.; visualization, Y.L.; supervision, Y.L.; project administration, Y.L.; funding acquisition, Y.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** Supported by “the Fundamental Research Funds for the Central Universities”, South-Central MinZu University (Grant Number: CSZL23010).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The original data presented in the study are openly available in FigShare at <https://figshare.com/s/553b517891ef0ef70533> (accessed on 14 July 2024).

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

- Steinhardt, N.S. Chinese architectural history in the twenty-first century. *J. Soc. Archit. Hist.* **2014**, *73*, 38–60. [CrossRef]
- Chen, F. Traditional architectural forms in market oriented Chinese cities: Place for localities or symbol of culture? *Habitat Int.* **2011**, *35*, 410–418. [CrossRef]
- Liu, Q.; Liao, Z.; Wu, Y.; Mulugeta Degefu, D.; Zhang, Y. Cultural sustainability and vitality of Chinese vernacular architecture: A pedigree for the spatial art of traditional villages in Jiangnan region. *Sustainability* **2019**, *11*, 6898. [CrossRef]
- Sicheng, L. Why Study Chinese Architecture? *J. Soc. Archit. Hist.* **2014**, *73*, 8–11. [CrossRef]
- Sun, L. The evolution of Liang Sicheng’s construction of Chinese architectural traditions in his drawings (1920s–1930s). *Front. Archit. Res.* **2023**, *12*, 319–336. [CrossRef]
- Wang, Y.; Woods, P.C.; Koo, A.C. Exploring the Evolution and Inheritance of Traditional Chinese Architectural Forms in Jiehua. *Int. J. Technol.* **2023**, *14*, 1196. [CrossRef]
- Keswick, M. *The Chinese Garden: History, Art, and Architecture*; Harvard University Press: Cambridge, CA, USA, 2003.
- Kim, Y.J.; Park, S. Tectonic Traditions in Ancient Chinese Architecture, and Their Development. *J. Asian Archit. Build. Eng.* **2017**, *16*, 31–38. [CrossRef]
- Zhang, D. Cultural Symbols in Chinese Architecture. *Archit. Des. Rev.* **2019**, *1*, 17.
- Wei, X.; Si, Z. Fully exploring traditional Chinese culture and promoting organic development of green city. *Procedia Eng.* **2017**, *180*, 1531–1540. [CrossRef]
- Li, L. The Impact of Artificial Intelligence Painting on Contemporary Art From Disco Diffusion’s Painting Creation Experiment. In Proceedings of the 2022 International Conference on Frontiers of Artificial Intelligence and Machine Learning (FAIML), Hangzhou, China, 19–21 June 2022; pp. 52–56.
- Ho, J.; Jain, A.; Abbeel, P. Denoising diffusion probabilistic models. *Adv. Neural Inf. Process. Syst.* **2020**, *33*, 6840–6851.
- Zhang, Z.; Fort, J.M.; Giménez Mateu, L. Decoding emotional responses to AI-generated architectural imagery. *Front. Psychol.* **2024**, *15*, 1348083. [CrossRef] [PubMed]
- Chen, J.; Shao, Z.; Hu, B. Generating interior design from text: A new diffusion model-based method for efficient creative design. *Buildings* **2023**, *13*, 1861. [CrossRef]
- Liao, W.; Lu, X.; Fei, Y.; Gu, Y.; Huang, Y. Generative AI design for building structures. *Autom. Constr.* **2024**, *157*, 105187. [CrossRef]
- Chen, J.; Wang, D.; Shao, Z.; Zhang, X.; Ruan, M.; Li, H.; Li, J. Using artificial intelligence to generate master-quality architectural designs from text descriptions. *Buildings* **2023**, *13*, 2285. [CrossRef]
- Li, C.; Zhang, T.; Du, X.; Zhang, Y.; Xie, H. Generative AI for Architectural Design: A Literature Review. *arXiv* **2024**, arXiv:2404.01335.
- Jaruga-Rozdolska, A. Artificial intelligence as part of future practices in the architect’s work: MidJourney generative tool as part of a process of creating an architectural form. *Architectus* **2022**, *3*, 95–104. [CrossRef]
- Hegazy, M.; Mohamed Saleh, A. Evolution of AI role in architectural design: Between parametric exploration and machine hallucination. *MSA Eng. J.* **2023**, *2*, 262–288. [CrossRef]
- Sukkar, A.W.; Fareed, M.W.; Yahia, M.W.; Abdalla, S.B.; Ibrahim, I.; Senjab, K.A.K. Analytical Evaluation of Midjourney Architectural Virtual Lab: Defining Major Current Limits in AI-Generated Representations of Islamic Architectural Heritage. *Buildings* **2024**, *14*, 786. [CrossRef]
- Saharia, C.; Chan, W.; Saxena, S.; Li, L.; Whang, J.; Denton, E.L.; Ghasemipour, K.; Gontijo Lopes, R.; Karagol Ayan, B.; Salimans, T. Photorealistic text-to-image diffusion models with deep language understanding. *Adv. Neural Inf. Process. Syst.* **2022**, *35*, 36479–36494.
- Gan, R.; Wu, X.; Lu, J.; Tian, Y.; Zhang, D.; Wu, Z.; Sun, R.; Liu, C.; Zhang, J.; Zhang, P. iDesigner: A High-Resolution and Complex-Prompt Following Text-to-Image Diffusion Model for Interior Design. *arXiv* **2023**, arXiv:2312.04326.



23. Taiwo, R.; Bello, I.T.; Abdulai, S.F.; Yussif, A.-M.; Salami, B.A.; Saka, A.; Zayed, T. Generative AI in the Construction Industry: A State-of-the-art Analysis. *arXiv* **2024**, arXiv:2402.09939.
24. Qian, W.; Yang, F.; Mei, H.; Li, H. Artificial intelligence-designer for high-rise building sketches with user preferences. *Eng. Struct.* **2023**, *275*, 115171. [[CrossRef](#)]
25. Zhao, L.; Song, D.; Chen, W.; Kang, Q. Coloring and fusing architectural sketches by combining a Y-shaped generative adversarial network and a denoising diffusion implicit model. *Comput.-Aided Civ. Infrastruct. Eng.* **2024**, *39*, 1003–1018. [[CrossRef](#)]
26. Chen, R.; Zhao, J.; Yao, X.; He, Y.; Li, Y.; Lian, Z.; Han, Z.; Yi, X.; Li, H. Enhancing Urban Landscape Design: A GAN-Based Approach for Rapid Color Rendering of Park Sketches. *Land* **2024**, *13*, 254. [[CrossRef](#)]
27. Petráková, L.; Šimkovič, V. Architectural alchemy: Leveraging Artificial Intelligence for inspired design—a comprehensive study of creativity, control, and collaboration. *Archit. Pap. Fac. Archit. Des. STU* **2023**, *28*, 3–14. [[CrossRef](#)]
28. Pena, M.L.C.; Carballal, A.; Rodríguez-Fernández, N.; Santos, I.; Romero, J. Artificial intelligence applied to conceptual design. A review of its use in architecture. *Autom. Constr.* **2021**, *124*, 103550. [[CrossRef](#)]
29. Ploennigs, J.; Berger, M. AI art in architecture. *AI Civ. Eng.* **2023**, *2*, 8. [[CrossRef](#)]
30. Najafian, K. *Maximalist AI Explorations Reimagine the Versailles Palace with Mesmerizing Gold Facades*; Mango, Z., Ed.; Designboom: Milan, Italy; New York, NY, USA; Beijing, China; Tokyo, Japan, 2022. Available online: <https://www.designboom.com/architecture/maximalist-ai-explorations-versailles-palace-gold-facades-kaveh-najafian-09-15-2022> (accessed on 14 July 2024).
31. Khan, R. Midjourney Reinvents Ancient Ziggurat Pyramid as Modern Cultural Landmarks. 2023. Available online: <https://www.designboom.com/architecture/midjourney-ancient-ziggurat-pyramid-temple-modern-arts-venue-rolando-cedeno-de-la-cruz-04-27-2023/> (accessed on 14 July 2024).
32. Chen, J.; Shao, Z.; Zheng, X.; Zhang, K.; Yin, J. Integrating aesthetics and efficiency: AI-driven diffusion models for visually pleasing interior design generation. *Sci. Rep.* **2024**, *14*, 3496. [[CrossRef](#)]
33. Sukkar, A.W.; Fareed, M.W.; Yahia, M.W.; Mushtaha, E.; De Giosa, S.L. Artificial Intelligence Islamic Architecture (AIIA): What Is Islamic Architecture in the Age of Artificial Intelligence? *Buildings* **2024**, *14*, 781. [[CrossRef](#)]
34. Jin, S.; Tu, H.; Li, J.; Fang, Y.; Qu, Z.; Xu, F.; Liu, K.; Lin, Y. Enhancing Architectural Education through Artificial Intelligence: A Case Study of an AI-Assisted Architectural Programming and Design Course. *Buildings* **2024**, *14*, 1613. [[CrossRef](#)]
35. Cudzik, J.; Nyka, L.; Szczepański, J. Artificial intelligence in architectural education-green campus development research. *Glob. J. Eng. Educ.* **2024**, *26*, 20–25.
36. Mustoe, J.E. *Artificial Intelligence and Its Application in Architectural Design*. Ph.D. Thesis, University of Strathclyde Glasgow, Glasgow, UK, 1990.
37. Zhang, Z.; Fort, J.M.; Mateu, L.G. Exploring the Potential of Artificial Intelligence as a Tool for Architectural Design: A Perception Study Using Gaudí's Works. *Buildings* **2023**, *13*, 1863. [[CrossRef](#)]
38. Albaghajati, Z.M.; Bettaieb, D.M.; Malek, R.B. Exploring text-to-image application in architectural design: Insights and implications. *Archit. Struct. Constr.* **2023**, *3*, 475–497. [[CrossRef](#)]
39. Yoshimura, Y.; Cai, B.; Wang, Z.; Ratti, C. Deep learning architect: Classification for architectural design through the eye of artificial intelligence. *Comput. Urban Plan. Manag. Smart Cities* **2019**, *16*, 249–265.
40. Indonesia, I.A.; Keprofesian, B. *Pedoman Hubungan Kerja Antara Arsitek Dengan Pengguna Jasa*; Badan Sistem Informasi Arsitektur, Ikatan Arsitek Indonesia: Kepulauan Riau, Indonesia, 2007.
41. Fernandes, D.; Garg, S.; Nikkel, M.; Guven, G. A GPT-Powered Assistant for Real-Time Interaction with Building Information Models. *Buildings* **2024**, *14*, 2499. [[CrossRef](#)]
42. Fareed, M.W.; Bou Nassif, A.; Nofal, E. Exploring the Potentials of Artificial Intelligence Image Generators for Educating the History of Architecture. *Heritage* **2024**, *7*, 1727–1753. [[CrossRef](#)]
43. Beyan, E.V.P.; Rossy, A.G.C. A review of AI image generator: Influences, challenges, and future prospects for architectural field. *J. Artif. Intell. Archit.* **2023**, *2*, 53–65.
44. Zhang, C.; Wang, W.; Pangaro, P.; Martelaro, N.; Byrne, D. Generative Image AI Using Design Sketches as input: Opportunities and Challenges. In Proceedings of the 15th Conference on Creativity and Cognition, Virtual, 19–21 June 2023; pp. 254–261.
45. Brisco, R.; Hay, L.; Dhami, S. Exploring the role of text-to-image AI in concept generation. *Proc. Des. Soc.* **2023**, *3*, 1835–1844. [[CrossRef](#)]
46. Lee, Y. The parametric design genealogy of Zaha Hadid. *J. Asian Archit. Build. Eng.* **2015**, *14*, 403–410. [[CrossRef](#)]
47. Dhariwal, P.; Nichol, A. Diffusion models beat gans on image synthesis. *Adv. Neural Inf. Process. Syst.* **2021**, *34*, 8780–8794.
48. Corvi, R.; Cozzolino, D.; Poggi, G.; Nagano, K.; Verdoliva, L. Intriguing properties of synthetic images: From generative adversarial networks to diffusion models. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Vancouver, BC, Canada, 17–24 June 2023; pp. 973–982.
49. Bolojan, D.; Yousif, S.; Vermisso, E. Latent Design Spaces: Interconnected Deep Learning Models for Expanding the Architectural Search Space. In *Architecture and Design for Industry 4.0: Theory and Practice*; Barberio, M., Colella, M., Figliola, A., Battisti, A., Eds.; Springer International Publishing: Cham, Switzerland, 2024; pp. 201–223.
50. Goodfellow, I.J.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.C.; Bengio, Y. Generative Adversarial Nets. In Proceedings of the Neural Information Processing Systems, Montreal, QC, Canada, 8–13 December 2014.

51. Karras, T.; Laine, S.; Aittala, M.; Hellsten, J.; Lehtinen, J.; Aila, T. Analyzing and Improving the Image Quality of StyleGAN. In Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 14–19 June 2019; pp. 8107–8116.
52. Zhu, J.-Y.; Park, T.; Isola, P.; Efros, A.A. Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks. In Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017; pp. 2242–2251.
53. Zhang, L.; Rao, A.; Agrawala, M. Adding conditional control to text-to-image diffusion models. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Paris, France, 4–6 October 2023; pp. 3836–3847.
54. Dhesikan, R.; Rajmohan, V. Sketching the Future (STF): Applying Conditional Control Techniques to Text-to-Video Models. *arXiv* **2023**, arXiv:2305.05845.
55. Isola, P.; Zhu, J.-Y.; Zhou, T.; Efros, A.A. Image-to-Image Translation with Conditional Adversarial Networks. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2016; pp. 5967–5976.
56. Han, Q.; Yin, C.; Deng, Y.; Liu, P. Towards classification of architectural styles of Chinese traditional settlements using deep learning: A dataset, a new framework, and its interpretability. *Remote Sens.* **2022**, *14*, 5250. [[CrossRef](#)]
57. Zhu, J. *Architecture of Modern China: A Historical Critique*; Routledge: London, UK, 2013.
58. Xue, C.Q. *Building a Revolution: Chinese Architecture Since 1980*; Hong Kong University Press: Hong Kong, China, 2005; Volume 1.
59. Loey, M.; Manogaran, G.; Khalifa, N.E.M. A deep transfer learning model with classical data augmentation and CGAN to detect COVID-19 from chest CT radiography digital images. *Neural Comput. Appl.* **2020**, 1–13. [[CrossRef](#)] [[PubMed](#)]
60. Peng, J.; Liu, D.; Xu, S.; Li, H. Generating diverse structure for image inpainting with hierarchical VQ-VAE. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, Nashville, TN, USA, 20–25 June 2021; pp. 10775–10784.
61. Rosca, M.; Lakshminarayanan, B.; Warde-Farley, D.; Mohamed, S. Variational approaches for auto-encoding generative adversarial networks. *arXiv* **2017**, arXiv:1706.04987.
62. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Proceedings of the Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015: 18th International Conference, Munich, Germany, 5–9 October 2015; Proceedings, Part III 18; 2015; pp. 234–241.
63. Alom, M.Z.; Hasan, M.; Yakopcic, C.; Taha, T.M.; Asari, V.K. Recurrent residual convolutional neural network based on u-net (r2u-net) for medical image segmentation. *arXiv* **2018**, arXiv:1802.06955.

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.