



Using CycleGAN to Achieve the Sketch Recognition Process of Sketch-Based Modeling

Yuqian Li and Weiguo Xu(✉)

School of Architecture, Tsinghua University, Beijing, China

liyuan17@mails.tsinghua.edu.cn, xwg@mail.tsinghua.edu.cn

Abstract. Architects usually design ideation and conception by hand-sketching. Sketching is a direct expression of the architect's creativity. But 2D sketches are often vague, intentional and even ambiguous. In the research of sketch-based modeling, it is the most difficult part to make the computer to recognize the sketches. Because of the development of artificial intelligence, especially deep learning technology, Convolutional Neural Networks (CNNs) have shown obvious advantages in the field of extracting features and matching, and Generative Adversarial Neural Networks (GANs) have made great breakthroughs in the field of architectural generation which make the image-to-image translation become more and more popular. As the building images are gradually developed from the original sketches, in this research, we try to develop a system from the sketches to the images of buildings using CycleGAN algorithm. The experiment demonstrates that this method could achieve the mapping process from the sketches to images, and the results show that the sketches' features could be recognised in the process. By the learning and training process of the sketches' reconstruction, the features of the images are also mapped to the sketches, which strengthen the architectural relationship in the sketch, so that the original sketch can gradually approach the building images, and then it is possible to achieve the sketch-based modeling technology.

Keywords: Sketch-based modeling · Sketch recognition · Image-to-image translation · CycleGAN

1 Introduction

The concept design is the initial in the architectural design, and it is also the most important part in the whole process. Once the concept is determined, the design direction is also determined. And architects usually design ideation and conception by hand-sketching which is a direct expression of the architect's creativity. But with the computer aided architecture design system, you will spend a lot of time to convert the sketch to a 3D modeling. However, if the sketch could directly generate the computer architectural concept model which could be edited and developed by the architect, it will be efficient to the design process.

At present, the sketch-based modeling is a relatively popular research direction. Compared with the traditional 3D software modeling method, the sketch in the sketched-based modeling has replaced the “Window, Icon, Menu, Pointer” (WIMP) interactive method in the traditional 3D software. The sketch expresses the designer’s intention and then completes the modeling task. Since sketching is one of the architect’s professional competence, this modeling method is very friendly to the architect, and because of its easy operation, the whole modeling process can be completed by one person alone.

However, for a sketch-based modeling system, it is very difficult to understand the design intent expressed by the sketch. That is, the realization of feature mapping from 2D sketches to 3D modeling is one of the difficulties in the system. Due to the differences in hand-sketching expressions, the ambiguity of the sketch itself increases the difficulty of understanding the sketch. So, additional knowledge and corresponding methods need to be added in the modeling process to reduce the difficulty of understanding the sketch as much as possible. People tend to use simple sketches to express initial ideas and concept and want to use as few strokes as possible to convey information. Therefore, if researches want to realize the feature map from 2D sketches to 3D modeling, the first step is to achieve of sketch recognition.

Because of the development of artificial intelligence, especially machine learning technology, Convolutional Neural Networks (CNNs) have shown obvious advantages in the field of extracting features and matching, and Generative Adversarial Neural Networks (GANs) have made great breakthroughs in the field of architectural generation which make the image-to-image translation become more and more popular.

As the building images are gradually developed from the original sketches, in this research, we try to develop a sketch-to-image translation system which could map the images’ features to the sketch and in the process of the sketch reconstruction, the architectural relationships of the sketches have been strengthened, and then achieve the sketch recognition process in the sketch-based modeling.

2 Related Works

Sketch-based modeling is a research about computer graphics, and there are many related research results. The earliest Sketch-based modeling study was based on contour sketch modeling. Igarashi et al. (1999) proposed a method of judging 3D geometric shapes by recognizing the contour curve of the sketch. Xu et al. (2014) developed a sketch-based True2Form modeling system, which uses selective regularization algorithms from 3D shape information such as curvature, symmetry, parallelism and other shape attributes. Bui et al. (2015) developed a method to generate 3D appearance shadow illustrations by recognizing the outline and shadow of the sketch. Xu et al. (2013) proposed the Sketch2scene framework, which can automatically infer multiple scene objects from a hand-sketching to generate a good 3D model scene. Huang et al. (2017), developed a deep convolutional neural network, in which the features of the 2D sketch are calculated as the parameters of the model, and these parameters in turn produce multiple sketches similar to the input, then the user can select an output shape, or further modify the sketch to explore other shapes.

The above-mentioned studies put forward a variety of recognition methods in the sketch-based modeling, which provide methodological reference to our study. However,

because of the researchers’ computer professional background, the results are universal and impractical. To develop the sketch-based modeling is undoubtedly the most suitable candidate for architects. This group is well aware of the logic of architectural design, can understand the design intent of architectural sketches, and also has strong 3D space capabilities.

Of course, architects and scholars have tried to use the machine learning and its algorithm results to study building generation tasks. For example, Matias Del Campo tried to use style transfer algorithms to generate the building skin (2019) and plan the urban city (2019). Weixin Huang from Tsinghua University and the University of Pennsylvania Hao Zheng from the University of Pennsylvania also have done some studies about the generation of indoor units through the pix2pix algorithm (2018). These results have inspired the architect’s design.

In this study, we try to make a sketch-to-image translation in order to achieve the sketch-based modeling, which is also a study about architectural generation.

3 Methodology

3.1 Network Architecture

As mentioned above, architects have tried several different algorithms to achieve the image-to-image translation, such as style transfer algorithm, pix2pix algorithm and so on. The style transfer algorithm is actually developed from the texture generation area, which combined with the deep object recognition area, so the core of the algorithm is still a texture style; the pix2pix algorithm is an optimized version of the cGAN, and its requirement about the data is very demanding, which require paired data. However, in many tasks, paired training data will not be available. Such as the data in this study—the sketch and the image of the building, it is a set of unpaired data, which is equivalent to two modes of the same scene. For this kind of data set, the algorithm of CycleGAN could improve the problem of pix2pix algorithm’s stringent data pair requirements (Fig. 1).

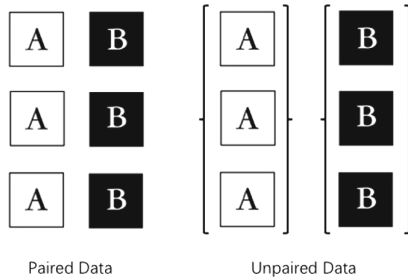


Fig. 1. The different between paired data and unpaired data

The CycleGAN presents an approach for learning to translate an image from a source domain X to a target domain Y in the absence of paired examples. The goal is to learn a mapping $G: X \rightarrow Y$, such that the distribution of images from $G(X)$ is indistinguishable

from the distribution Y using an adversarial loss. Because this mapping is highly under-constrained, CycleGAN couple it with an inverse mapping $F: Y \rightarrow X$ and introduce a cycle consistency loss to enforce $F(G(X)) \approx X$ (and vice versa) (Fig. 2).

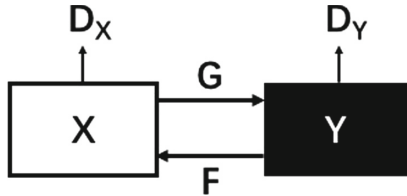


Fig. 2. The network architecture of CycleGAN

3.2 Data Preparation

Principles of Data Collection

Before the data collection, we made some principles: First, the sketch and the image of the building must be one building. It means that a one-to-one correspondence between the sketch and the image of the building in the data, although this is not required in the CycleGAN, we believed that such a data set may improve the effectiveness of model training. Second, all the designs are well-known and the sketches are made by the famous architects themselves. Third, the data collected should be extensive. Due to the subjective nature of the architect's drawing of sketches, and the design techniques of architectural schemes are diverse. By collecting a wider range of data samples, the scope of the data set is more comprehensive.

Data Collection

Since it is difficult to collect the architect's sketches and corresponding images, the data that can be collected is limited. After screening and processing the collected data, a total of 200 data were selected, namely 100 sketch data and 100 image data.

Data Processing

First, normalize the collected data, and each picture is $256 * 256$. After that, 160 data, that is, 80 pairs of samples are used as training data, and 40 data, that is, 20 pairs of samples are used as test data. Among them, the sketch data set is placed in the trainA folder as the source data domain X , which corresponds to the target data domain Y ; the image data set is placed in the trainB folder as the source data domain Y , which corresponds to the target data domain X .

3.3 Training Process

The CycleGAN is a ring structure, with two generators $G (X \rightarrow Y)$ and $F (Y \rightarrow X)$, two discriminators D_X and D_Y : in the generator part, because the image in this study

is $256 * 256$, so 9 residual blocks are used; in the discriminator part, through five-layer convolution, the number of channels is reduced to 1, and finally the average pooling size is also reduced to $1 * 1$.

The training process is that X represents the image in the sketch domain, and Y represents the image in the building image domain. The image of the sketch domain is generated by the generator G to the image of the building image domain, and then reconstructed back to the original image input in the sketch domain by the generator F; the image of the building image domain is generated by the generator F to generate the image of the sketch domain, and then generated by the device F reconstructs back to the original image input in the building image map domain. It is worth noting that CycleGAN

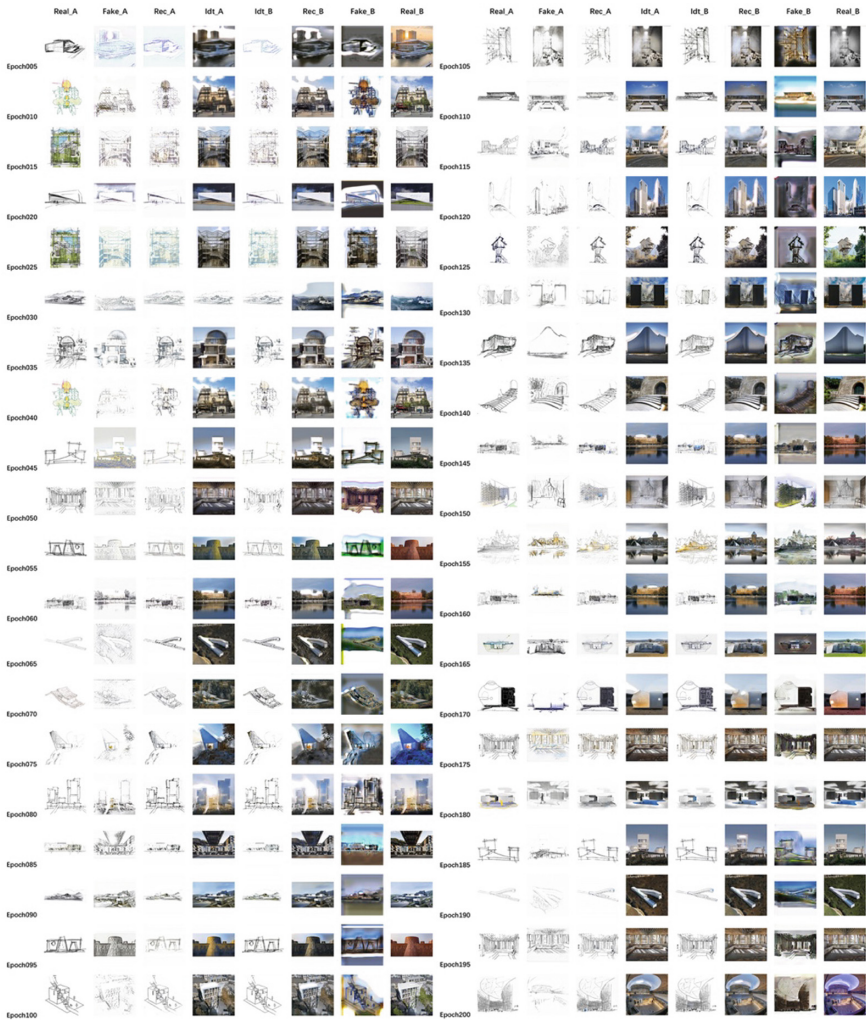


Fig. 3. The part process of the training

adds an identity mapping part, that is, generator G uses sketches to generate building images, but if the input itself is a building image, then it should generate an image belonging to the building image. In addition, for the stability of training, historically generated fake samples are used to update the discriminator instead of the currently generated fake samples (Fig. 3).

4 Results

From the Fig. 4, we can see that the training from the sketch to the building image has completed the sketch recognition and through the training of the reconstruction, the features of the building images are mapped to the sketches, which strengthens the architectural relationship in the sketch, which could make the original sketch to approach the building images step by step.

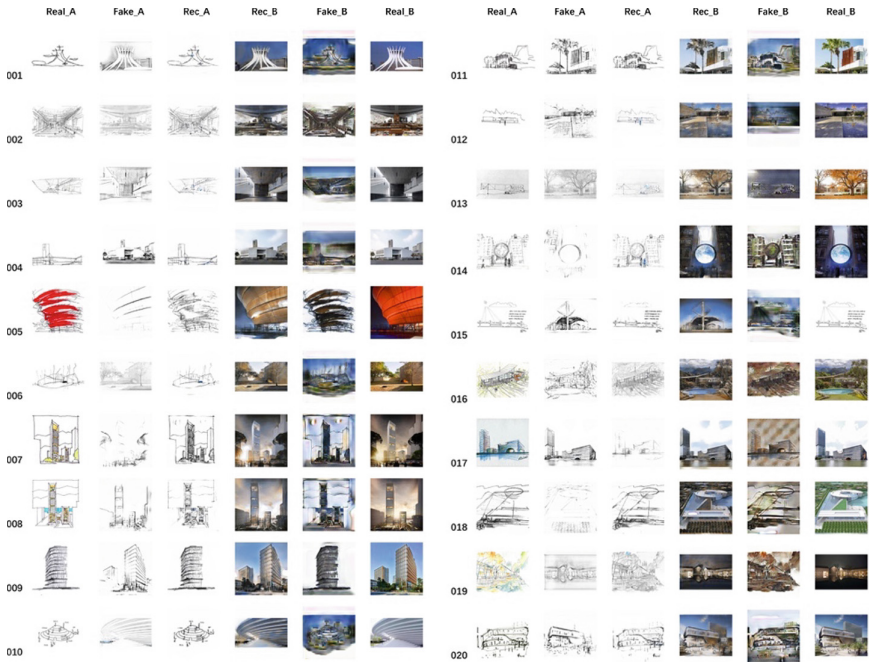


Fig. 4. The results of the test training

4.1 Recognition of Sketch and Generation of Corresponding Building Image

First, it can be seen from the Fig. 5 that in the generation of the sketch to the building image, the boundary of the sketch has been recognized. The training process has identified the building's exterior images and interior images, because the sky of the generated exterior images has been rendered to blue and in the generated interior images, the original color state of the building images has been retained.

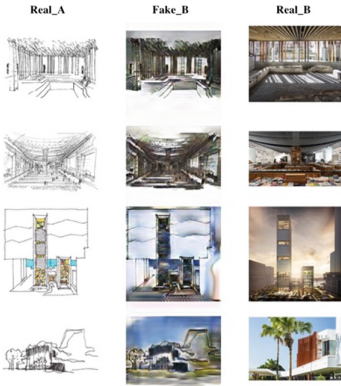


Fig. 5. Recognitions of sketches

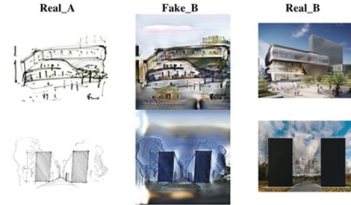


Fig. 6. The building volume relationship



Fig. 7. The environmental relationship

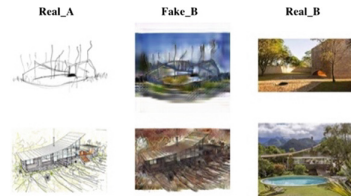


Fig. 8. The horizontal comparisons

Second, in the Fig. 6, the building volume relationship of the building image is well recognized and mapped in the sketch. In more detail, the virtual-real relationship of the three building volumes has also been well studied.

Third, in the Fig. 7, the environmental relationship of the building, such as shadow changes, light transmission and reflection of windows has been well reflected in the generated image.

Also, through the horizontal comparison of the different sketches and the corresponding images pairs of the generated building images in the Fig. 8, it is found that there will be differences in the generation results with different drawing levels. The simpler the sketch is, the worse the building image it generates, and the more complex the sketch, the better the result.

4.2 Sketch Reconstruction

As there is an image reconstruction part in the CycleGAN, it has been reflected in the output. By training the features of the building images, a new sketch based on the original sketch is reconstructed. It can be seen from the Fig. 9 that the reconstructed sketch maps certain features of the building images and strengthens the architectural relationship in the sketch.

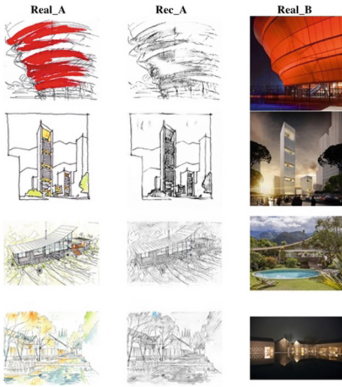


Fig. 9. The reconstructed sketches

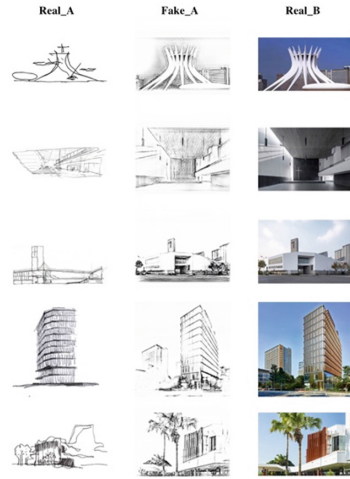


Fig. 10. The generations from building images to sketches

4.3 Building Images to Sketches

It can be seen from the Fig. 10 that the generation from building images to sketches is also successful, even better than the result of the sketch-generated-building-image. For the sketch, its features are relatively unified and more obvious, that is, a sketch with a single color. This result reflects that if the features of the building images are uniform, the final results of the sketch-generated-image could be better.

5 Conclusion and Discussion

This study is a sketch-to-image translation based on CycleGAN. Through the training of 160 data and the testing of 40 data, the study has completed the mapping process from sketch to building images. The results show that the CycleGAN can achieve the sketch recognition and reconstruction. Training is to map the features of the building image to the sketch, which strengthens the architecture relationship in the sketch, so that the original sketch can approach the building image gradually. And the sketch's reconstruction is also very consistent to the architect's cycled workflow and developed logic in the architectural design process.

Of course, the study still has some limit. First, the number of the data is not enough. Secondly, the data in this study is complex and extensive. If we add a single style or a comparison between the sketches of a certain architect and the building images, we could be able to compare the ability of data with different levels of complexity in the direction of generation from sketches to building images.

Acknowledgement. This research is supported by National Natural Science. Fund of China (NO. 51538006).

References

- Xu, B., Chang, W., Sheffer, A., Bousseau, A., McCrae, J., Singh, K.: True2form: 3D curve networks from 2D sketches via selective regularization. *ACM Trans. Graph.* (2014)
- Bui, M.T., Kim, J., Lee, Y.: 3D-look shading from contours and hatching strokes. *Comput. Graph.* **51**, 167–176 (2015)
- Del Campo, M., Manninger, S., Sanche, M., Wang, L.: The church of AI - an examination of architecture in a posthuman design ecology. In: Haeusler, M., Schnabel, M.A., Fukuda, T. (eds.) *Intelligent & Informed - Proceedings of the 24th CAADRIA Conference*, Victoria University of Wellington, Wellington, New Zealand, 15–18 April 2019, vol. 2, pp. 767–772 (2019)
- Del Campo, M., Carlson, A., Manninger, S.: Machine hallucinations: a comprehensive interrogation of neural networks as architecture design. In: *Proceedings of IASS Annual Symposia* (2019)
- Huang, H., Kalogerakis, E., Yumer, E., Mech, R.: Shape synthesis from sketches via procedural models and convolutional networks. *IEEE Trans. Visual. Comput. Graphics* **23**, 2003–2013 (2017)
- Huang, W., Zheng, H.: Architectural drawings recognition and generation through machine learning. In: *Proceedings of the 38th Annual Conference of the Association for Computer Aided Design in Architecture (ACADIA)*, Mexico City, Mexico, 18–20 October 2018, pp. 156–165 (2018). ISBN 978-0-692-17729-7
- Igarashi, T., Matsuoka, S., Tanaka, H.: Teddy: a sketching interface for 3d freeform design. In: *ACM SIGGRAPH*, Los Angeles, August, vol. 6, no. 1, pp. 409–416 (1999)
- Xu, K., Chen, K., Fu, H., Sun, W.L., Hu, S.M.: Sketch2scene: sketch-based co-retrieval and co-placement of 3D models. *ACM Trans. Graph.* **32**(4CD), 1–15 (2013)
- Zheng, H., An, K., Wei, J., Ren, Y.: Apartment floor plans generation via generative adversarial networks. In: *Anthropocene, Design in the Age of Humans-Proceedings of the 25th CAADRIA Conference*, vol. 2, Chulalongkorn University, Bangkok, Thailand, 5–6 August 2020, pp. 599–608 (2020)
- Zhu, J.Y., Park, T., Isola, P., Efros, A.A.: Unpaired image-to-image translation using cycle-consistent adversarial networks (2017)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

