Research on Architectural Generation Design of Specific Architect's Sketch Based on Image-To-Image Translation

Yuqian Li, Weiguo Xu[✉], and Xingchen Liu

School of Architecture, Tsinghua University, Beijing, China
{liyuqian17,liuxc20}@mails.tsinghua.edu.cn,
xwg@mail.tsinghua.edu.cn

Abstract. Sketch is a way for architects to communicate with others. Architects record their own ideas through rapid drawing. However, sketches are abstract, vague, and even ambiguous. To this end, architects need to spend a lot of time, through modeling and other means, to present the architectural plan that can be understood by people. However, this method is time-consuming and laborious. Due to the development of deep learning technology, especially convolutional neural networks (CNN) and generative adversarial networks (GAN), they have shown great advantages in the field of image recognition and generation. With the help of these technologies, ambiguous architectural sketches can be directly transformed into architectural scheme drawings, and architects' creative intentions can be continuously improved and developed, It will be very convenient and efficient. Therefore, based on the image-to-image translation, this paper realizes the mapping from architectural sketches to architectural scheme drawings with the help of CycleGAN. Through the analysis of the architectural generation design results of Frank Gehry's and Alberto Campo Baeza's architectural sketches, firstly, the feasibility of this method is verified. Secondly, it is found that this method can well complete the identification of sketch boundaries. In the generated scheme drawings, it can not only reflect the volume and lighting changes of the building, but also reflect the architect's creative intention and style to a large extent, The side reflects the cognitive ability of this method to architectural design.

Keywords: Architectural sketch · Architectural scheme drawings · Image-to-Image translation · CycleGAN

1 Introduction

With the development of computer technology, especially computer graphics, the use of 3D modeling, rendering and other means of design has become very convenient and efficient, but the concept sketch for architects in the process of architectural creation, always has a unique irreplaceable. A sketch has a definite goal and is a picture with a specific intention. Sketch is an action process in which architects record their own ideas through rapid drawing.

Le Corbusier said: I like to use painting to express design ideas. Painting can be faster and more realistic. In addition to Corbusier, famous architects such as Mies van der Rohe, Frank Gehry and Zaha Hadid all like to express their design ideas and intentions through sketches. To them, the sketch is like an unfinished architectural work of an architect, a way of communication between the architect and others, which is full of the complex mental process of the architect and shows the architectural ideal of the architect.

However, sketches are often abstract, fuzzy and even ambiguous. Take Zaha Hadid as an example. Her sketches are very abstract. Zaha once said that she hopes to use abstract expression to break the thinking of traditional architecture. For this kind of intentional design expression, people's interpretation of it is often not clear, even sometimes, for the architect himself can not imagine such a fuzzy sketch, its development into a building will be. So, generally, architects and their teams need to spend a lot of time, through modeling to present architectural scheme drawings that can be understood by people. If the output is not satisfactory, they also need much time to modify, and constantly repeat this process.

Because of the development of deep learning, especially convolutional neural networks (CNNs) and generative adversarial networks (GANs), they have shown great advantages in the field of image recognition and generation. If we develop these technologies with architectural design, ambiguous sketches can be directly transformed into scheme drawings, and architects' creative intentions can be continuously improved and developed, then the work will be very convenient and efficient.

It is worth mentioning that from modernism, post-modernism to deconstructionism, architects of various schools emerge in endlessly, and their design styles are quite different. Extensive research increases the difficulty of understanding sketches. It is very necessary to study from individual to general.

Therefore, with the help of deep learning, this paper will extract the architectural sketch features of a specific architect, and produce the corresponding architectural scheme drawings, so as to realize the translation between the sketch and the building image, which could help the process of architectural design and achieve the purpose of convenience and efficiency.

2   Related Work

In the early stage of AI, machine learning has become a cutting-edge technology and has been applied in many researches. The black box process of machine learning is very similar to the process of people's cognition of the world. Through the learning and training of a large number of data, we can find its inherent law and map this law to the generated results. At the beginning of the twenty-first century, deep learning has shown greater advantages in big data training by increasing the depth of hidden layer of neural network in machine learning, and has successfully made a breakthrough in speech, image and other processing. This also makes the field of image translation as a branch of deep learning research develop rapidly.

In the deep neural network, convolution neural network is the most effective network for image processing. CNN can extract the specific features in the image, that is, a group of computational elements can process the visual information hierarchically

through forward feedback. When using CNN for model recognition, the representation information learned by CNN will change with the change of network level, that is, with the increase of network level, the focus of image representation will change from the initial specific pixel to the image content. Therefore, the low-level CNN will capture the style representation of the image, while the high-level network will capture the content representation of the image. On the basis of this theory, scholars such as Gatys et al. [2, 3] proposed the style transfer and applied it to the creation of artists' style paintings.

Although the style transfer can achieve image translation successfully, because this method needs deep neural network in training, and the trained model can only be applied to specific image migration, the image translation algorithm based on GAN shows better performance.

In 2014, Goodflow et al. [4] and others proposed the GAN, which is composed of two models: generator and discriminator. Since the discriminator will constantly judge the similarity between the model generated by the generator and the original model, Isola [5] and others proposed Pix2Pix, an image translation framework based on GAN. After that, Pix2PixHD solves the problem of high-resolution image translation on the basis of Pix2Pix, and Vid2Vid solves the problem of high-resolution video image translation on the basis of Pix2PixHD. Since then, many image-to-image algorithms based on GAN have emerged.

Whether it is based on CNN or GAN, its goal is to learn the mapping between the input image and the output image, that is, image-to-image translation requires not only the generation of dual image in the target domain according to the image in the source domain, but also the consistency of the image in the translation process.

3   Methodology

3.1   Network Architecture

As mentioned above, GAN has obvious advantages in the field of image generation. Pix2Pix, Pix2PixHD and Vid2Vid are known as the trilogy in the field of image translation, which can achieve most of the work of image-to-image translation. Scholar Huang and Zheng [6] and Zheng et al. [7] have done a lot of research on the generation of room types through Pix2Pix. These results have inspired architects' design ideas to a certain extent.

However, in the process of architectural design, as a way to express the design intention, sketch is often not achieved overnight, it needs to be continuously developed and improved, it is a relatively circular process, rather than one-way. Therefore, this paper chooses the Cyclegan based on Pix2Pix. Cyclegan has a continuous cycle of generation process, in order to simulate the design process of sketch implementation scheme.

CycleGAN is a technology that uses unpaired image collections from two different domains to train an unsupervised image conversion model through the GAN architecture.

The CycleGAN (Fig. 1) 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 → Y, such that the distribution of images from G(X) is indistinguishable from the distribution Y using an adversarial loss. Because this mapping
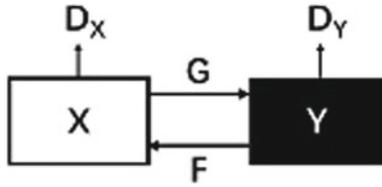
Fig. 1. The network architecture of CycleGAN

is highly under-constrained, CycleGAN couple it with an inverse mapping F: Y → X and introduce a cycle consistency loss to enforce F(G(X))≈X (and vice versa).

There are many creative applications of CycleGAN. It was first used in the transfer of photographs to an artist's painting such as converting a photograph into a Van Gogh painting. Jack Clark used this algorithm to convert ancient maps of Babylon, Jerusalem, and London into modern Google Maps and satellite views. Mario Klingemann used the code to translate portraits into dollface. Besides, CycleGAN is also used in medical fields, such as translating MRI to CT data.

3.2 Data Preparation

3.2.1 Sample Selection

The purpose of this paper is to extract sketch features and generate architectural scheme drawings through cyclegan algorithm. As mentioned above, the sketch itself has ambiguity, and each architect has his own design style. Collecting mixed architectural style data will increase the difficulty of computer cognition.

Frank Gehry is a master of deconstruction. His design is characterized by peculiar and irregular curves, and his sketches are relatively abstract. Alberto Campo Baeza is a Spanish modernist architect. His architecture is pure, elegant and poetic. He is good at using the combination of flowing space and light. His sketches are simple and readable. Therefore, this paper selects these two famous architects who are good at sketching as learning samples, and because of their different sketching styles, we can get a closer understanding of computer cognition through comparison.

3.2.2 Data Collection

Due to the limitation of the number of projects, architects can not collect a large number of data, and to achieve one-to-one correspondence also increases the difficulty of data collection. In this paper, through the collection of different perspectives of the same building, different scenes of same building, for the two architects, we collected 100, that is, 50 pairs, 80, that is, 40 pairs of data as the research samples, 80% as training data, 20% as test data.

3.3 Training Process

Take the cycle generation from Gehry's sketch to scheme drawing as an example (Fig. 2).
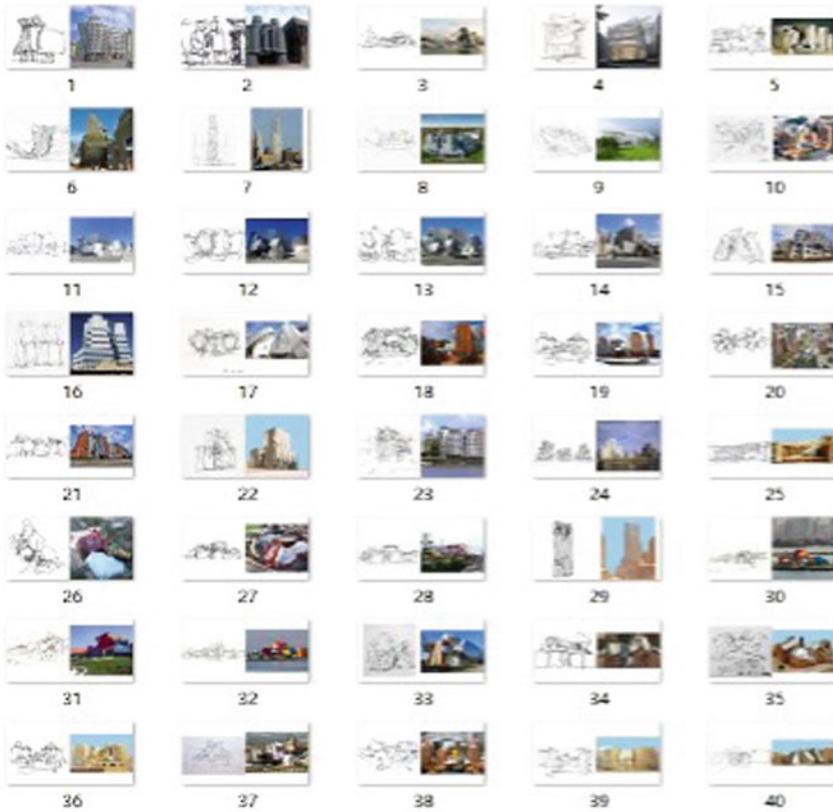
Fig. 2. The data of Gehry's work

Gehry's sketches and building images are taken as two groups of pictures and they are unpaired, while the sketch and building images are one-to-one correspondence.

The CycleGAN will develop an architecture of two GANs, and each GAN has a discriminator and a generator model, meaning there are four models in total in the architecture.

The first GAN will generate pictures of sketches given pictures of building images, and the second GAN will generate scheme drawings given pictures of sketch.

Each GAN has a conditional generator model that will synthesize an image given an input image. And each GAN has a discriminator model to predict how likely the generated image is to have come from the target image collection. The discriminator and generator models for a GAN are trained under normal adversarial loss like a standard GAN model.

So far, the models are sufficient for generating plausible images in the target domain but are not translations of the input image.

Each of the GANs are also updated using cycle consistency loss. This is designed to encourage the synthesized images in the target domain that are translations of the input image.

Cycle consistency loss compares an input picture to the Cycle GAN to the generated picture and calculates the difference between the two, e.g. using the L1 norm or summed absolute difference in pixel values.

There are two ways in which cycle consistency loss is calculated and used to update the generator models each training iteration.

The first GAN (GAN 1) will take an image of sketches, generate image of scheme drawings, which is provided as input to the second GAN (GAN 2), which in turn will generate an image of sketches. The cycle consistency loss calculates the difference between the image input to GAN 1 and the image output by GAN 2 and the generator models are updated accordingly to reduce the difference in the images.

This is a forward-cycle for cycle consistency loss. The same process is related in reverse for a backward cycle consistency loss from generator 2 to generator 1 and comparing the original pictures of the buildings to the generated picture of scheme drawings.

4   Results

See (Fig. 3).

a. The output   b. The output of Alberto's work

Fig. 3. The results of the test training

4.1 Data Preparation

Whether it is Gary's work (Fig. 4) or Alberto's work (Fig. 5), in the generated scheme drawing, the sketch boundary recognition is well completed. In the scheme drawing, the architecture, sky and other environmental factors are clearly expressed.
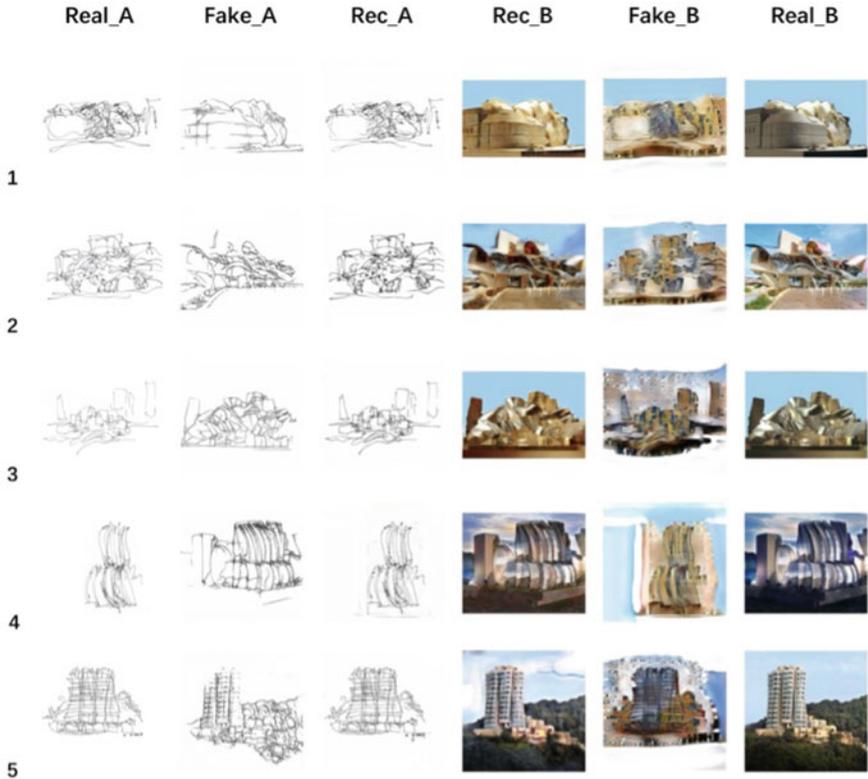

Fig. 4. xxx

4.2 Certain Cognitive Ability for Different Perspectives of the Same Building

In order to expand the sample size, this study collected different perspectives of the same building. It can be seen in the Fig. 6 that due to the continuity of design elements from different perspectives, the generated scheme drawings have similar color attributes. This may be because the computer recognizes the similarity between images to a certain extent.
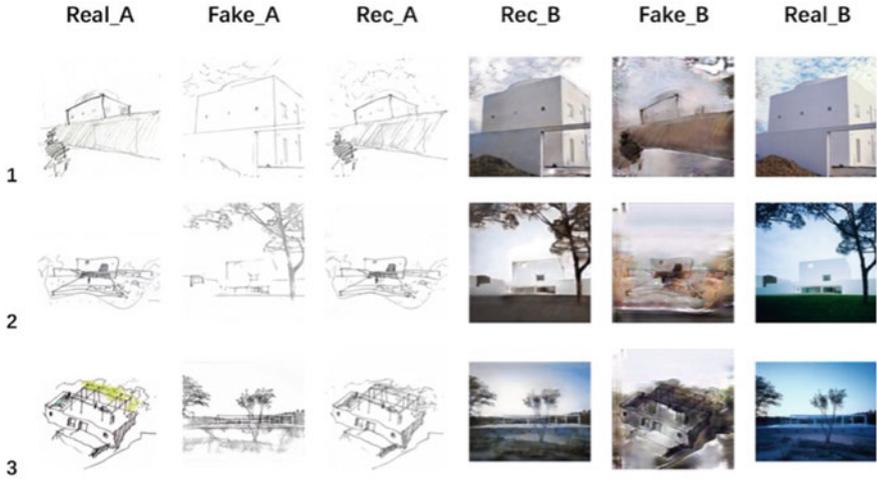
Fig. 5. xxx
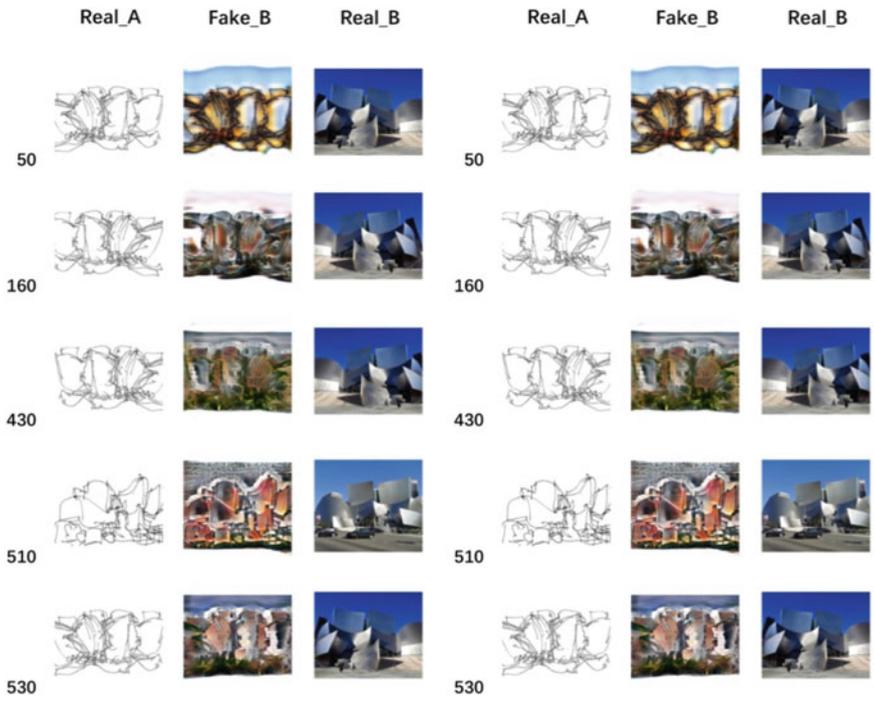
Fig. 6. xxx

4.3   The Generated Scheme Drawings Can Reflect the architect's Creative Style

In Alberto's design, he is good at using natures and hopes to integrate architecture and environment. It can be seen from the Fig. 7 that the generated architectural scheme drawing has a very similar architectural scene, and the natural elements, light and environment are fully expressed.
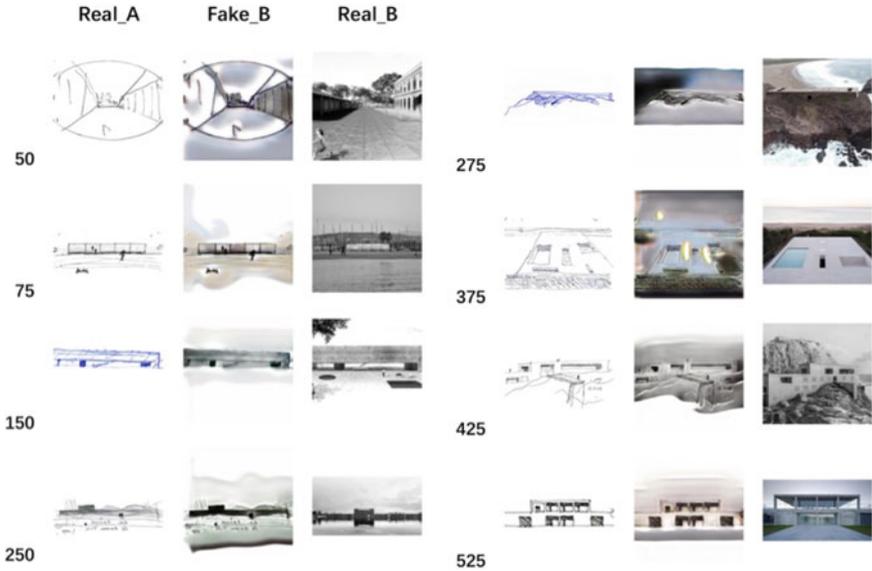

Fig. 7. xxx

4.4   The Generated Scheme Drawings Can Reflect Good Shadow Changes

In the Fig. 8, the architectural light and shadow expressed by Alberto in real buildings are well interpreted in the generated renderings. It creates the same interior atmosphere as the architect's intention.

It can also be seen from the Fig. 9 that in the design of Gehry, the light and shadow between the building volumes are also expressed, which interprets the relationship between the building volumes.

5   Conclusion and Discussion

In this study, based on image to image translation, with the help of GycleGAN, the mapping from architectural sketches to building images is realized. Through the analysis of the architectural generation design results of Gehry's and Alberto's architectural sketches, the feasibility of this method is verified. Secondly, it is found that this method can well complete the identification of sketch boundaries. In the generated scheme
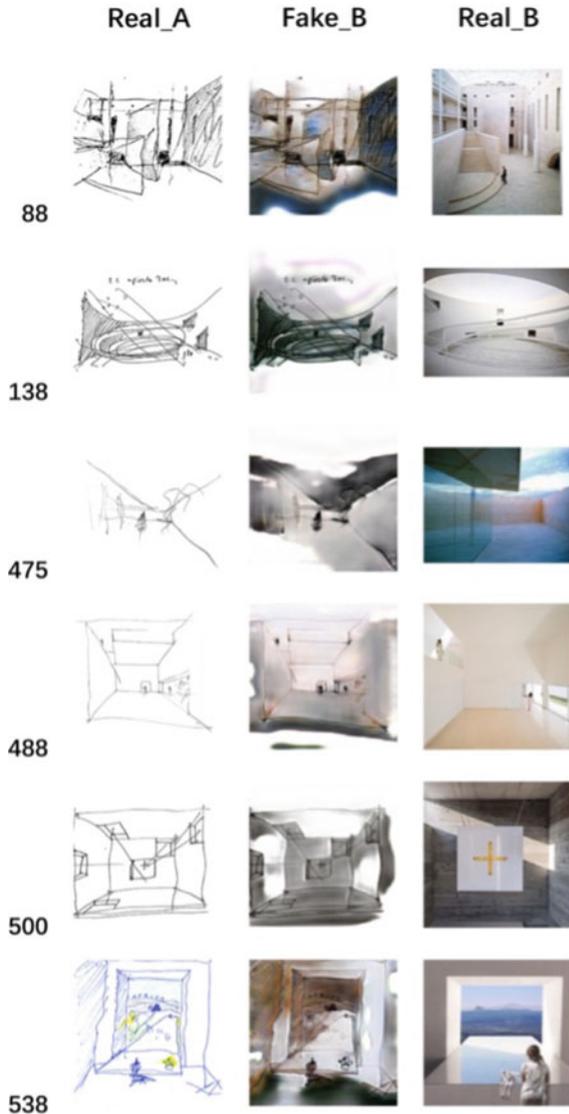
Real_A   Fake_B   Real_B

88

138

475

488

500

538

Fig. 8. The results of the test training

drawings, it can not only reflect the volume and lighting changes of the building, but also reflect the architect's creative intention and style to a large extent, the side reflects the cognitive ability of this method to architectural design.

Through the horizontal comparison of the sketch generation results of the two architects, we can see that the more abstract the sketch is, the more difficult it is to identify, the clearer and simpler the sketch is, the better the effect of the generated scheme drawing will be.

Fig. 9. The results of the test training

Of course, this also shows that the cognitive ability of computer needs to be further strengthened. This is also one of the important tasks of the next research.

References

1. Simonyan K, Zisserman A (2014). Very deep convolutional networks for large-scale image recognition. arXiv
2. Gatys LA, Ecker AS, Bethge M (2015) Texture synthesis using convolutional neural networks. MIT Press
3. Gatys LA, Ecker AS, Bethge M (2015) A neural algorithm of artistic style. J Vis
4. Goodfellow IJ, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S et al (2014) Generative adversarial networks. In: Advances in neural information processing systems, vol 3, pp 2672–2680
5. Isola P, Zhu JY, Zhou T, Efros AA (2016) Image-to-image translation with conditional adversarial networks. In: IEEE conference on computer vision & pattern recognition. IEEE
6. Huang W, Zheng H (2018) 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 Oct 2018, pp 156–165. ISBN 978-0-692-17729-7
7. Zheng H, An K, Wei J, Ren Y (2020) 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 Aug 2020, pp 599–608
8. Zhu JY, Park T, Isola P, Efros AA (2017) Unpaired image-to-image translation using cycle-consistent adversarial networks