



# A Preliminary Study on the Formation of the General Layouts on the Northern Neighborhood Community Based on GauGAN Diversity Output Generator

Yuzhe Pan, Jin Qian, and Yingdong Hu<sup>(✉)</sup>

School of Architecture and Design, Beijing Jiaotong University, Beijing, China  
yduhu@bjtu.edu.cn

**Abstract.** Recently, the mainstream gradually has become replacing neighborhood-style communities with high-density residences. The original pleasant scale and enclosed residential spaces have been broken, and the traditional neighborhood relations are going away. This research uses machine learning to train the model to generate a new plan, which is used in today's residential design. First, in order to obtain a better generation effect, this study extracts the transcendental information of the neighborhood community in north of China, using roads, buildings etc. as morphological representations; GauGAN, compared to the pix2pix and pix2pixHD, used by predecessors, can achieve a clearer and a more diversified output and also fit irregular contours more realistically. ANN model trained by 167 general layout samples of a neighborhood community in north of China from 1950s to 1970s can generate various general layouts in different shapes and scales. The experiment proves that GauGAN is more suitable for general layout generation than pix2pix (pix2pixHD); Distributed training can improve the clarity of the generation and allow later vectorization to be more convenient.

**Keywords:** Neighborhood community in north of China · General layout morphology generation · GauGAN · pix2pix · Machine learning

## 1 Introduction

Driven by the value of urban areas, from the 1980s to the 1990s, high-density housing has gradually replaced neighborhood community as the mainstream. The original pleasant scale and enclosed residential spaces have been broken, ignoring human scale and psychological considerations. The settlement environment has become larger in scale and richer in form, but the inclusiveness for diversified functions and behaviors has become less and less, and it is more difficult to change the fact that neighborhood relations are gradually indifferent. How to create a habitable public space and reshape the neighborhood social network has become the difficulty of residential planning today.

These authors contributed equally to this work and should be considered co-first authors. Project supported by Beijing Jiaotong University Training Program of Innovation and Entrepreneurship for Undergraduates.

© The Author(s) 2021

P. F. Yuan et al. (Eds.): CDRF 2020, *Proceedings of the 2020 DigitalFUTURES*, pp. 179–188, 2021.

[https://doi.org/10.1007/978-981-33-4400-6\\_17](https://doi.org/10.1007/978-981-33-4400-6_17)

Neighborhood community originated from the former Soviet Union were widely used in cities construction in China because it met the need of rapid construction new settlements in the Early Stage of the Republic. Most buildings are 3–5 floors, with streets outside and courtyards inside. The semi-public space organization in the courtyard and the mode of the dense road network in small blocks are the main reasons for its good neighborhood relations [1].

Nowadays, if you want to reproduce the neighborhood scene, you need to explore the characteristics of the public space organization and street scale of neighborhood community first. Existing neighborhood communities are the result of a long time trial and error process by predecessors, and it is undoubtedly more efficient than groping. However, the analysis of potential mechanism problems has always been a major problem facing the architectural discipline. There are certain limitations in the related research, no matter the methods and tools used. First, relying on the subjective judgment of researchers, deducing the constituent elements of independent cases, limited by differences of knowledge and background, cognition is also different, and researches based on a small number of cases cannot accurately extract common characteristics; secondly, generation laws are complex, it is difficult to qualitatively and quantitatively analyze and convert into measurable indicators, specific design methods and executable programs. This research hopes to extract objective factors and discover hidden laws [2].

Recently, advances in deep learning technology, represented by computer vision, have made it possible for computers to collect and analyze implicit feature laws. This study uses machine learning to conduct large-scale sample training, to cover the potential laws of neighborhood community and automatically generate output to provide a more rational and comprehensive working method and ideas for the protection and innovation of northern neighborhoods.

## 2 Research

### 2.1 AI Application in Architecture

The application of computers in the architecture field has gone through four stages: Modularity, Computational Design, Parametricism and AI. In the first three stages, computer was used as design aids due to its high computing performance, such as BIM and digital construction, which couldn't design independently without architects. The development of AI in the 21st century makes it possible to replace designers in specific fields like building form generators, digital construction and so on, especially in the field of building layout in which predecessors have used computer vision to explore.

### 2.2 Deep Learning Architectural Plan Generator Application

Through the retrieval of existing articles, the following three used related methods of computer vision achieving substantial results in the field of building plan generation. The first two articles are mainly focused on the generation of indoor floor plans, while the last article is aimed at general layout. The results all adopt a similar method: first convert building layout into bitmaps, and then use image to image algorithm to train the deep learning model (Table 1).

**Table 1.** Comparison of various studies in the research

Author	Algorithms	Result	Deficiency
Hao Zheng	pix2pixHD	Indoor floor plan and Mutual mapping of functional color block diagrams	Can't match irregular boundaries (lack of corresponding samples); Unable to output diversely
Stanislas Chaillou	pix2pix	Step training, designers can intervene in the intermediate link, controllability; Realize the output from the outline of the plane to the indoor layout	Can't match irregular boundaries (lack of corresponding samples; limitation of algorithm itself); Unable to output diversely
Yubo Li	Deep learning	Generic flat contour and road generate function color map	Unable to output diversely

In 2018, Hao Zheng trained 100 samples using pix2pixHD to realize the recognition and generation of architectural layouts. First, it trains with indoor layout as label and function color map as result; at the same time, it reserves training to achieve the mutual map of the indoor layout and the function color map; The visualization of the training process proves that pix2pixHD has similarities with human cognition. The author mentioned that results couldn't match the irregular boundary well that may be due to samples which have irregular boundary are few [5].

Stanislas Chaillou produced the indoor layout generator program ArchiGan in 2019. Using GANs and step training, three models can realize from site conditions to building outline to function color map to indoor layout. When inputting the site condition, it can output the indoor layout. In the third step, it conducts separate training the furniture layout of each type of room through different colors. Step training is to control the single steps of machine learning, to realize the human intervention and ensure generation quality. It trained four architectural styles to achieve style transfer to be selected according to pros and cons when designing [6].

The schgan produced by Yubo Liu in 2019 can automatically generate a functional layout of elementary school based on road and contour conditions. According to the user evaluation mechanism set by it, the plan generated by AI is higher than the original plan. Specific experimental method was not disclosed, but the results prove that the deep learning can achieve generation of general layouts [8].

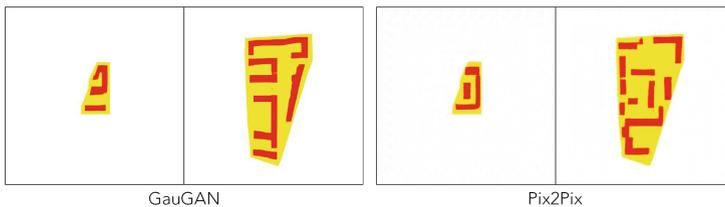
The above three articles have the following questions: (1) Due to the limits of algorithm and the number of samples, the generated results have a lot of noise, and the visual recognition is greatly affected; (2) The above articles exist that the direction of elements in the results can only be orthogonal, and cannot change according to site condition. This experiment proves that reasons include not only the above-mentioned oblique arrangement, but also the limitations of pix2pix itself; (3) The above researches can not realize diversity generated results with a single input, which is impossible to provide designers with multiple possibilities.

### 3 Methodology

#### 3.1 GauGAN

Compared with pix2pix [7] used by predecessors, GauGAN [7] performs better in high-resolution image synthesis, and natively supports diverse output. The reason lies in the SPADE generator and Multiscale Discriminator that constitute the GauGAN neural network structure. The first image to image algorithm is pix2pix, then improved pix2pixHD [8]. Compared to pix2pix, pix2pixHD removes the U-Net structure in the generator, and uses Resnet and local Enhancer to increase the network depth, and improves on the basis of pix2pix's patch discriminator to multiscale discriminator, which contributes to the better high resolution Images (above  $512 \times 512$ ) synthesis. And the difference between GauGAN and pix2pixHD is mainly the generator model architecture, which has the following advantages:

1. GauGAN's generator has a better ability to obtain information because it uses SPADE (Spatially-adaptive denormalization) as the norm layer. Compared to the instance/batch norm layer used in pix2pix and pix2pixhd, SPADE can better retain the information in the labelmap, and finally can make the generator output more in line with the information provided by the labelmap. This may also be the reason why the generated result of gauGAN can correspond to the inclined profile. Figure 1 shows the difference between the output results of gauGAN and pix2pix when facing oblique land use conditions. It can be seen that GauGAN adopts the SPADE norm layer, which makes the shape of the plane of the oblique layout of the building correspond well in the generated results. The outline of the hypotenuse of the land used in the labelmap, but pix2pix does handle with the outline of the hypotenuse in the labelmap, and still reflects the orthogonal form.



**Fig. 1.** Comparison between GauGAN and Pix2Pix

2. Multi-modal synthesis is natively implemented in GauGAN. Similar to ordinary GANs, the input of SPADE Generator is a normal distributed random vector (latent), and labelmap is the input of each SPADE Resnet in the Generator, so GauGAN can make generation diverse as general GANs. However, the generator input of pix2pix and pix2pixHD is directly a semantic map without latent, so the result diversity cannot be directly achieved. The results of generating diversity are conducive to designers to filter, reflecting the concept of human-computer interaction.
3. The calculation efficiency of GauGAN is higher than pix2pixHD. The resolution changing part of the pix2pix and pix2pixHD generator is executed by the deconv/conv

layer, while this part of SPADE Generator is executed by the interpolate sample layer result in less parameters and higher efficiency than pix2pixHD.

### 3.2 Step Training

In order to improve the clarity of the generated results, this experiment compares the overall and step training. The overall training is to train the model to directly map the contour graph to the layout morphology, while step training is to map the contour graph to the building layout graph first, and then map the building layout graph to the road network graph. When testing, matrix multiplication is applied on the obtained building layout and road network to obtain the final result.

Comparing the generated results obtained by testing two methods (Fig. 1), it can be distinguished that the generated results of step training are clearer and less noisy, so the step training method is finally adopted in this experiment (Fig. 2).

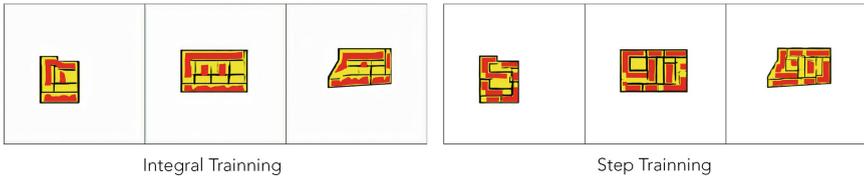


Fig. 2. Comparison between Integral training and Step training

## 4 Machine Learning for the General Layout Shapes of the Northern Neighborhoods in China

See Fig. 3.

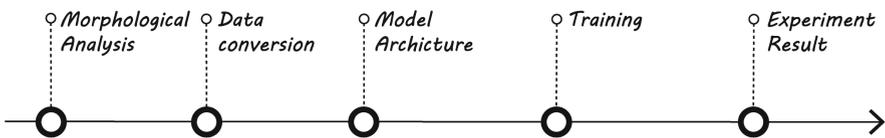


Fig. 3. Process

### 4.1 Morphological Analysis

First step is to analyze the composition of the general layout of neighborhood community and extract the elements. The public space formed by buildings and roads is the main place to communicate. The buildings determine the division, and the road as a special public space also affects communication. So, this study extracts roads, buildings, and courtyard spaces as morphological representations, shields redundant information that is not relevant, using machine learning to extract common features of public space organization and street scale in the samples.

## 4.2 Data Conversion

This step is to convert the general layout into bitmap data that the program can process. This study explored 167 neighborhood communities in Beijing and Tianjin, compiling and plotting as data sets. In order to adapt step training, each sample includes four types: outline map, non-road set, road set and complete sample set (Fig. 4). After balancing the calculation amount of neural network and the resolution, the  $512 * 512$  pixel bitmap is used uniformly, and the maximum range is  $330 \text{ m} * 330 \text{ m}$ . All samples are entered with the same scale. The scale of the 167 sample ranges from  $110 \text{ m} * 110 \text{ m}$  to  $330 * 250 \text{ m}$ , all within the maximum range. Use three colors to represent the three extracted elements, and use HSV to control the variables: building:  $H = 0, S = 100$ ; courtyard space:  $H = 58, S = 100$ ; road:  $H = 0, S = 0$ . Some samples have the characteristics of oblique field contours so that whether GauGAN can match irregular boundaries can be confirmed.

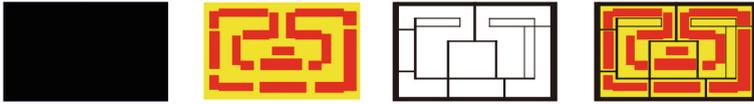


Fig. 4. Sample examples (from left to right: input, output non-road, output road, full output)

## 4.3 Model Architecture

This step uses the GauGAN algorithm to complete the model architecture and uses step-by-step training (outline-> buildings and buildings-> roadmap)

### 1. GauGAN neural network architecture

GauGAN is an image translation algorithm published by Nvidia Lab in 2019 that can achieve multi-modal synthesis. This experiment is implemented according to the paper, including VAE, which is used to achieve style guide multi-modal synthesis. The implementation of GauGAN is shown in Fig. 5.

### 2. Step Training

The training process of the model is not directly from contour to final result, but a step-by-step training (ArchiGAN: a Generative Stack for Apartment Building Design, 2019) to obtain a clearer bitmap result, which is divided into the following two parts: (1) Take the out-of-plane contour as the input dataset and train the building layout as the output dataset; (2) Then use the building layout as the input dataset and the road network map as the output dataset for training.

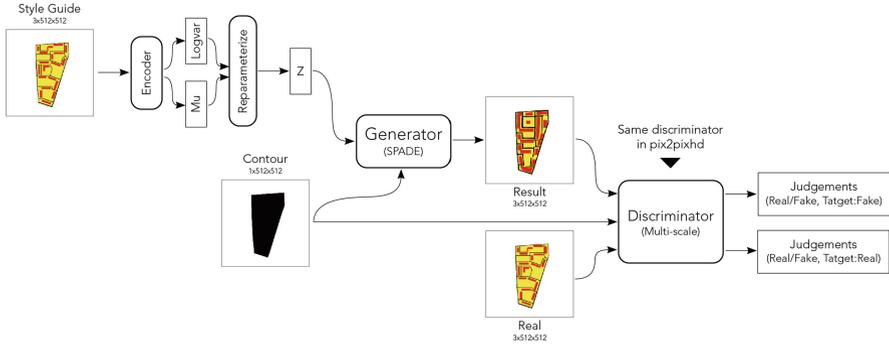


Fig. 5. Detail of GauGAN architecture

When testing, input the outline of the plane into the model trained in the first step, and the output building layout map is input into the second model to obtain the road network map, and then matrix multiplication of the building layout map and the road network map to obtain the final result. The content is shown in Fig. 6

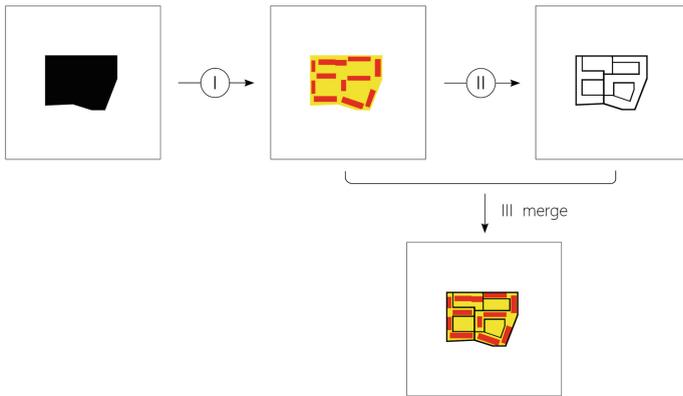
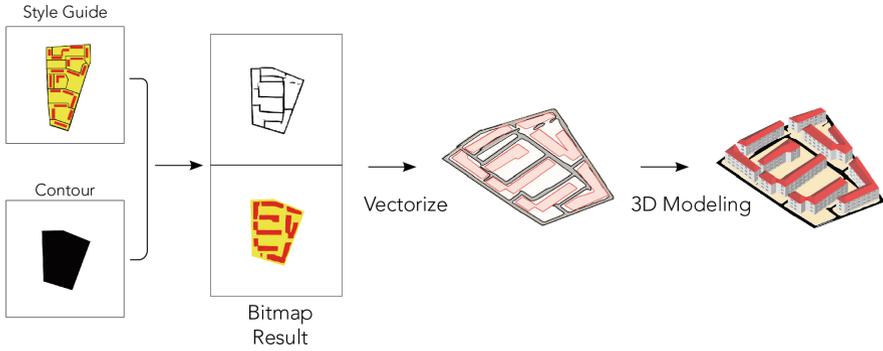


Fig. 6. Step Training

#### 4.4 Vectorization and 3D Procedural Modeling

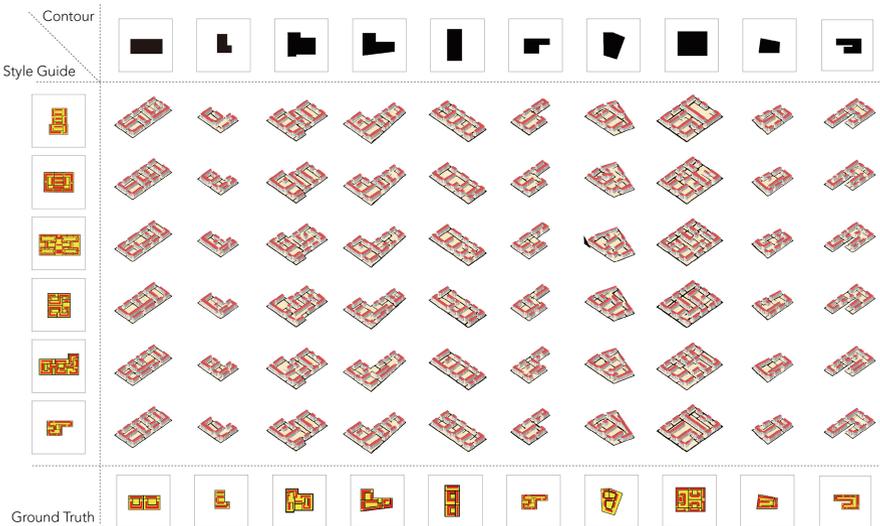
Models trained by the step training generates road bitmap and building layout bitmap results. The two bitmap results are vectorized and merged, and then 3d procedural modeling and visualization are performed. The black polygons represent the road net, and it can also be seen the windows and red roof, which are similar to the real world neighborhood community building style, so that the generated results are more intuitive and easier to be accepted by architects (Fig. 7).



**Fig. 7.** Vectorization and 3d procedural modeling

### 4.5 Experiment Result

Using the trained model for the result generation test, compared with the real plan (Fig. 8), it can be seen that for the input of the same site contour map, the model can generate various different results, reflecting the diversity of the GauGAN algorithm. Secondly, from test results of the third sample in the figure, it can be seen that GauGAN makes good use of the information provided by the contour map to obtain generated results that conform to the input, so that the generated building shape fits the oblique outline fitted the sample status happening. However, the results also show that the style guide has less effect on the general layout generation results.



**Fig. 8.** Result

## 5 Conclusion

### 5.1 GauGAN Is More in Line with Architectural Design Needs Than Pix2pix (Pix2pixHD)

GauGAN is better than pix2pix in terms of diversity and contouring which is undoubtedly more in line with the needs of architectural design. Through experiments, it compared the GauGAN algorithm with the predecessor's pix2pix algorithm, and finally realized the multiple different flats including roads and buildings for a single plane outline; due to the use of GauGAN algorithm and the non-orthogonal samples included in the sample set, the results generated in this experiment are already achievable compared to the results of the predecessors Fit and change irregular contours. The above results prove that the method of this research experiment can be used to diversify and automatically generate the overall functional layout of the building, but whether Pix2pixHD can achieve contour fit is not yet accurately concluded.

### 5.2 The Use of Step Training Can Improve the Clarity of Generated Results and Allow the Later Vectorization to Be More Convenient

Experiments have found that the use of step training helps to improve image quality, resulting in clearer, less noisy, and more recognizable results; second, this method allows the results of building and roads to be completely separated, making it easier to separate buildings and buildings. The bitmap result of the road is vectorized to avoid the mixing of elements, which can not accurately vectorize different elements, so that the vectorized result becomes more accurate.

## References

1. Xuan, Y.: The Effect of Neighborhood Mode in Modern Residential District Planning. *Chinese & Oversea Architecture*, pp. 94–95 (2011)
2. Jia, R., Sun, Y.: Cultural characteristics and existence value of buildings in Beijing million village. *J. Beijing Inst. Civ. Eng. Arch.*, 76–80 (2012)
3. Huang, W., Zheng, H.: *Architectural Drawings Recognition and Generation through Machine Learning*. Tsinghua University, University of Pennsylvania (2018). <https://www.researchgate.net/publication/328280126>
4. Chaillou, S.: *AI Architecture Towards a New Approach* (2019). [https://www.academia.edu/39599650/AI\\_Architecture\\_Towards\\_a\\_New\\_Approach](https://www.academia.edu/39599650/AI_Architecture_Towards_a_New_Approach)
5. Isola, P., Zhu, J.-Y., Zhou, T., Efros, A.A.: Image-to-image translation with conditional adversarial networks. Berkeley AI Research (BAIR) Laboratory, UC Berkeley. <https://arxiv.org/abs/1611.07004> (2016)
6. Wang, T.-C., Liu, M.-Y., Zhu, J.-Y., Tao, A., Kautz, J., Catanzaro, B.: High-resolution image synthesis and semantic manipulation with conditional GANs. <https://arxiv.org/abs/1711.11585v1> (2017)
7. Park, T., Liu, M.-Y., Wang, T.-C., Zhu, J.-Y.: Semantic image synthesis with spatially-adaptive normalization. <https://arxiv.org/abs/1903.07291> (2019)
8. Liu, Y., Deng, Q., Liang, L.: SchGAN: Primary school campus layout generation, intelligent assistant for architectural design. (2020) <https://blog.csdn.net/shadowcz007/article/details/104035601>

**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.

