Urban Planning and Design Layout Generation Based on Artificial Intelligence

Ting Wan and Yuhang Ma
School of Garden, Northeast Forestry University, Harbin, Heilongjiang, China

Correspondence should be addressed to Ting Wan; wanting1970@nefu.edu.cn

Received 16 May 2022; Revised 7 June 2022; Accepted 9 June 2022; Published 28 June 2022

Academic Editor: Lianhui Li

Copyright © 2022 Ting Wan and Yuhang Ma. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Today’s cities are becoming more and more complex, the spatial layout is gradually becoming more and more complex, and all aspects of urban construction that need to be considered are increasing. The traditional urban planning and design methods have encountered new challenges. Based on the unique perspective of urban “mesoscale,” this study attempts to apply artificial intelligence technology in the early stage of urban planning and design, predict the positioning of design land based on the surrounding environment, so as to break the limitations of manual decision-making, explore the spatial layout problem from the perspective of machine, find the correlation between land data, and generate results with certain reference value to assist decision-making. By delimiting the research area and collecting and processing data, we trained and generated the artificial neural network model and selected three different areas for model test. The test results verify the feasibility and effectiveness of the method process.

1. Introduction

Human beings have created cities for survival and development. The most important and primary role of cities is to serve as gathering places for people living in them. Human activities connect people with urban space [1–4]. After the rapid development of urban construction, many urban plots have changed towards simplification and homogenization, the urban vitality has been weakened, and various problems have emerged in urban development. In fact, reasonable “mixing” is the form in which the city should exist and the ideal state for the sustainable development of the city. Even if the initial planning is relatively simple, the city still evolves in the direction of mixing in people’s actual use until it reaches a more balanced state [5, 6].

In recent years, with the increase of population and the enrichment of human activities, people also put forward new demands for urban construction [7]. The mixed development of cities is not only the current situation, but also a sustainable trend. In this context, the factors to be considered in urban construction will be more complex, comprehensive, and diversified [8–12]. The traditional experience-based urban planning and design methods have certain limitations, and the scheme ideas put forward by designers based on practical experience are sometimes not comprehensive and in-depth [13]. Today, in a digital era, computer technology is widely used in various fields [14–19], including architecture and urban areas. Concepts are parametric architectural design [20, 21], computational urban design [22, 23], urban big data analysis [24, 25], and smart city emerge [26] one after another.

Although the design field is more subjective than other science and engineering disciplines, it is not without rules to follow. When designers refer to and study excellent cases and apply them to their own schemes, they actually summarize the common laws and patterns of “good design.” Designers can find the law from the scheme. Similarly, if the data is input into the computer, the computer can also mine the law between the data. This kind of computer technology is called machine learning [27, 28], which is a subset of artificial intelligence technology. It aims to “learn” the correlation
between the characteristics of input data and label data through a large amount of data driving, so as to make a more accurate prediction. In urban research, machine learning technology can mine the potential laws between urban spatial layout data and try to generate urban spatial layout scheme from the perspective of computer to assist the existing manual planning and design.

Data input is an important part of machine learning model. The first law of geography reveals the close relationship between adjacent things. The surrounding environment is obviously one of the key factors to be considered in the design of urban spatial layout. The isolated scheme design separated from the environment will be very abrupt. Therefore, if the surrounding environment data of the design land is used as the input, the machine learning model can learn the relationship between the design land and the surrounding environment. The input data will cover multiple dimensions related to urban spatial layout, so as to give full play to the advantages of computer, put the city in a more complete quantitative environment for analysis, and comprehensively consider multiple attribute data of the plot to generate results. To sum up, in urban planning and design, how to use artificial intelligence technology to realize the generation of hybrid urban space layout scheme and create richer urban space to better meet the needs of residents is of great significance for the exploration of new urban research roads in the digital background.

2. Urban Planning and Design
   Layout Generation

With the help of artificial intelligence technology, we describe the construction process of the urban spatial layout generation method explicitly. The necessary steps of method construction are presented universally to form a standardized process [3, 5]. The premise of method construction is to clarify the basic research unit form of urban spatial layout. This study selected the artificial neural network as the main model algorithm. The relevant data of urban spatial structure is input into the model in a numerical matrix. This study will deal with these data in a grid and take the data grid obtained by dividing the research area into the most basic research unit.

2.1. Overall Framework. This study explores how to construct a method for generating urban spatial layout based on artificial intelligence technology. The method in this study is as follows: the build will consist of specific data preparation and algorithm development. In practical application, the design land may occupy multiple divided grid units. The algorithm development will be subdivided into two parts: single-grid algorithm development and multigrid algorithm development. In the case of an unknown grid, the multigrid algorithm expands the scope of application. The method construction is generally divided into three parts: data preparation, single-grid essential algorithm development, and multigrid expansion algorithm development.

The data preparation part aims to provide data input for the algorithm development part. After data acquisition, data preprocessing, and dataset generation, three-dimensional data is generated for each grid. In addition, when there is a severe imbalance in the number of classes in the dataset, additional processing methods are required to reconstruct the dataset to make the class relatively balanced.

The essential algorithm development part aims to realize the spatial layout generation of a single research grid, including the steps of artificial neural network construction, model training optimization, and result generation and display. The expansion algorithm development part starts from two ideas to realize the generation of multigrid results: They are the solution space optimization search idea based on Monte-Carlo Tree Search (MCTS) [29, 30] and the iterative elimination idea based on adjacent grid filling. The MCTS algorithm uses the complete single-grid model of the neighborhood grid generated by the basic algorithm. The adjacent grid filling algorithm also needs to train several single-grid models with incomplete neighborhood grids. These incomplete models require sample augmentation to the original dataset to ensure sufficient data volume.

The above three parts of the work constitute the entire method content, and the framework flow is shown in Figure 1.

2.2. Data Preparation

2.2.1. Data Collection. This research aims to use machine learning technology to provide a reference for the generation of urban spatial layout. After selecting a target city, to avoid the tendency of result generation to be too evident and single, it will try to delineate the cities with more mature development and more mixed conditions. Areas are used as the research scope, and data acquisition work is carried out so that the data can cover more complex situations and enhance the generality of the method. For a target city, after the training is completed, the model can learn the layout logic of the city; that is, when it is practically applied, the generated reference layout scheme will present a “style” or “feature” similar to the city.

After determining the research scope, to describe the concept of urban spatial layout more comprehensively, this research screened out three elements that are closely related to spatial layout and have a substantial impact: land use function, urban morphology, and traffic connections, and collect the required data from these three aspects as the input to the machine learning algorithm.

(1) Functional Elements of Land Use. The function is one of the most intuitive elements to describe the content of urban land, guiding different human activities. The nature of land use generally refers to the functional use of a piece of land at the planning level, which can be directly obtained from planning data or indirectly obtained through the classification of planned or completed building functions on the plot.
The land use properties in the traditional sense are obtained through the urban land use status map. Still, the drawing of these maps has a significant delay and does not fully conform to the status, and the classification is not precise enough to show the mixed state of the region. In contrast, Point of Interest (POI) is another way of describing land use functions. Each point contains information such as name, latitude and longitude coordinates, and functional classification; the classification is detailed and updated quickly. However, POI also has a big problem: For functions with a single nature but a large area, such as schools and the industrial regions, one point obviously cannot cover the actual range, and the error is significant. The study integrates land use status and various real-time POI data crawled to generate more accurate data on land use properties. Land use status data and POI data are obtained using open map APIs like Baidu Maps and Tencent Maps.

(2) Functional Elements of Urban Morphology. Urban morphology is another key element in describing urban space. Applying machine learning requires using a quantitative, discretized representation of urban form so that this element can be smoothly fed into the model in numerical form. At the macroscale, common quantitative expression methods of the urban form include space syntax, Formtax, Morpho, LCZ theory, etc. Among them, LCZ theory covers more comprehensive quantitative indicators of urban conditions than other methods. While establishing qualitative and quantitative description indicators, it also provides a more applicable method for urban form’s quantitative description and research.

The LCZ theory divides urban forms into 17 classes; the built environment includes 10 classes: high-density high-rise, high-density mid-rise, high-density low-rise, low-density high-rise, low-density mid-rise, low-density low-rise, and light-weight low-rise, mass low-rise, scattered buildings, industrial plants. The natural environment includes 7 classes: dense trees, sparse trees, bushes, low vegetation, hard paving, bare land, and water. Since some classifications are rare and inapplicable in China’s urban areas, many studies have proposed revisions to the original classes. For example, high-density middle-rise buildings, scattered buildings, dense trees, and bushes were deleted in some studies, and 13 classes were retained. 11 classes were included after research reduction and integration, including 7 built class types and 4 nature class types. Considering that this study focuses on urban built-up areas, too detailed morphological classification of the natural environment will lead to too few samples per class, affecting the model’s training effect. Therefore, based on the 11-class classification, this study further simplified the natural environment class, and only the vegetation class and water surface are retained.

To obtain LCZ data, it is necessary to download and get satellite remote sensing image data such as Landsat-8 from the official website of the United States Geological Survey (USGS) and then use the WUDAPT method to process and generate it. The WUDAPT method refers to using the band value of the remote sensing image as the classification basis for water, vegetation, non-vegetation, and other classes and then manually sampling each class in Google Earth, inputting the sample boundary position into the GIS software and comparing it with the satellite image. Get each class of samples based on satellite imagery. Based on these training samples, the random forest algorithm is used to further classify and discriminate the grids in the study area, supplemented by appropriate manual error correction operations, to obtain the final LCZ classification results.

(3) Functional Elements of Traffic Connection. In addition to the “content” of the plots, the relationship between plots, that is, traffic, needs to be described in urban design. Many factors affect the traffic connection in urban land use. This study selects point data based on the bus, subway, and other stations and line data related to road networks, such as road grades and road integration. After comprehensive calculation, it is used to characterize the traffic connection strength value of the grid. The site data is obtained through POI crawling, and the road network data is downloaded from the OpenStreetMap (OSM) open-source map website.

In summary, the data used in this study include five types of functional POI, land use status, Landsat-8 satellite remote sensing map, traffic station POI, and urban road network. They correspond to the three dimensions of land use, LCZ,
and traffic intensity of the dataset. Among them, Landsat-8 satellite images and road network data are downloaded from the corresponding website; the rest of the geographic data was obtained from the open map API using web crawling technology (Table 1).

2.2.2. Data Division and Processing. After obtaining the raw data, the following processing steps are performed to generate the dataset results.

1) Grid Division. The study area is divided into grids in the GIS software, and the data of each dimension is superimposed on the grid at the corresponding location using the overlay analysis tool of GIS.

The following two criteria determine the grid size: one needs to match the grid size of the LCZ data as much as possible, and the other is to conform to the common urban block plot size. The grid size of LCZ data is generally between 200 m and 500 m in the early period, and 100 m is more common in recent years. The common size of urban blocks is between 50 m and 200 m. The grid size used in this study is determined according to the specific conditions of different urban instances.

2) Data Processing. To ensure the effectiveness of the machine learning model, the data of the three dimensions need to be normalized.

(a) Land use data. This study divides land use properties into five basic classes: business, industrial, public, residential, and landscape, covering most functional land use classes. See Table 2 for specific descriptions and code abbreviations used in the text.

Integrate the crawled functional POIs into the above five classes and stack them on the grid. Considering the hybridity of functions, in order not to lose too much information, this study chooses to calculate the proportion data of five types of land in the grid to represent the land use properties of the grid area, rather than just taking the maximum value to obtain a single dominant class. Based on this, relative proportions of five classes of POIs are generated for each grid calculation. The continuous proportion data are discretized into 6 levels of 0%, 20%, 40%, 60%, 80%, and 100%. For the blank areas not covered by POI, use the current land use data as a supplement to reduce the number of blank grids as much as possible. Finally, a one-dimensional array of length 5 is generated for each grid, and the numeric type is float.

(b) LCZ data. The 11 classes of LCZ data generated by the WUDAPT method were integrated into the 9 classes defined for this study, superimposed into the grid, and the LCZ data of each grid was stored in the form of one-hot encoding (One-Hot), and the numerical type was an integer. One-hot encoding is an effective encoding method for dealing with discrete categorical features. For this study, it means initializing a one-dimensional all-zero array of length 9: [0, 0, 0, 0, 0, 0, 0, 0, 0], and change the value of the corresponding LCZ type position to 1; for example, LCZ 3 can be expressed as [0, 0, 1, 0, 0, 0, 0, 0, 0].

(c) Traffic intensity data. This study defines the concept of “traffic intensity,” which describes the traffic situation of a grid by integrating the four factors of subway station distance, bus station distance, road level, and road integration degree. The distance data of two classes of stations are calculated and generated by the corresponding POI through the Euclidean distance tool in GIS. The road level is classified by the road property identification in the OSM road network data, which is divided into 4 levels; the integration degree is calculated based on the road network using the DepthMapX tool generated. The latter two are converted from line to raster data by kernel density weighted calculation.

The four indicators are reclassified according to certain standards, and the traffic intensity level of each grid is obtained after weighted summation. The specific calculation method and standard are shown in Table 3. Among them, the subway and bus stations are reclassified according to a certain radiation distance interval, and the radiation distance interval of the subway station is slightly larger than that of the bus station to reflect the more significant influence of the subway. The road level and road integration degree are reclassified based on the natural breakpoint classification method. In the weighted summation process, the station index and the road index account for 50% of the weight, and the weight of defining the subway station is slightly larger than that of the bus. The weight of the road level is somewhat more significant than the integration degree, which is 30% and 20%, respectively.

Then:

\[ x = x_{\text{Subway}} \cdot 0.3 + x_{\text{Bus}} \cdot 0.2 + x_{\text{Road level}} \cdot 0.3 + x_{\text{Road integrity degree}} \cdot 0.2. \]  

(1)

In formula (1), \( x \) is the comprehensive traffic intensity value, and \( x_{\text{Subway}} \), \( x_{\text{Bus}} \), \( x_{\text{Road level}} \), \( x_{\text{Road integrity degree}} \) represent the index value of the subway station distance, bus station distance, road grade, and road integration index after classification.

The summation results are divided into 4 levels, normalized to the 0-1 interval (Table 4). The final traffic intensity level data is obtained, represented by a one-dimensional array with a length of 1. The numerical type is a floating-point type.

2.2.3. Dataset Generation. Based on the first law of geography, this study attempts to explore the correlation between unknown grids and their surrounding known grids. Hence, a neighborhood range needs to be determined before data input. According to the definition of Moore’s Neighborhood 99, 3 x 3 is the minor neighborhood and is suitable as the basic unit of research. In contrast, the 5 x 5 neighborhood is one more grid distance than the 3 x 3 neighborhood, and it is...
also a more comfortable theoretical walking distance, worthy of study.

This study chooses to "slice" the original grid based on the two neighborhood ranges of $3 \times 3$ and $5 \times 5$. It defines the 8 grids in the $3 \times 3$ neighborhood of a grid as an "inner circle"; the 16 grids extended from the $5 \times 5$ neighborhood are called "outer circle."

After combining the data of the above three dimensions, express the spatial layout properties of each grid as a one-dimensional array of length 15; for example the array [0.2, 0, 0, 0.8, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0.67] means that 80% of the grid plot is residential land; 20% is commercial land, mainly high-density mid-rise buildings, and high traffic intensity (Table 5).

We slice the matrix grid data into $3 \times 3$ and $5 \times 5$ ranges using Python and remove the slices that contain blank grids without data to obtain a slice set that can be used for training. Extract the data of 8 ($3 \times 3$ neighborhood) or 24
(5 × 5 neighborhood) grids around the slice, and connect them into a one-dimensional array as learning data; extract the data of 1 grid in the middle as a label data to compare with learning results to evaluate training accuracy. For the 5 × 5 neighborhood, it is necessary to extract the data of the 8 grids in the inner circle and bring it to the front to be consistent with the input order of the 3 × 3 neighborhood.

Each slice is transformed into a 1D array of length 15 × 9 = 135 (3 × 3 neighborhood) or 15 × 25 = 375 (5 × 5 neighborhood), and the first 120 (3 × 3 neighborhood) or 360 (5 × 5 neighborhood) values are learning data. The last 15 are labels, which are processed in turn, and the final dataset results of 3 × 3 neighborhood and 5 × 5 neighborhood are obtained based on the slice set.

### 2.2.4. Analysis and Handling of Dataset Imbalance

The scope of data collection selected in this study is urban built-up areas, which will inevitably encounter the problem of unbalanced dataset categories. For example, the proportion of business, public, and residential land use is obviously higher than industry and landscape. The proportion of medium- and high-rise building types corresponding to these types of land will also be higher than that of low-rise buildings, vegetation, and water surfaces.

Dataset imbalance is a common problem in machine learning. When the number of samples is not much different, it can be ignored, but if the gap is relatively large, such as class A: class B = 100:1, it will lead to a poor model training effect. Neural networks largely tend to predict outcomes as class A. This model looks very accurate, even reaching 99%, which fully meets the requirements of high scores, but in fact, it completely ignores a few classes, and the model is almost ineffective.

To solve this problem, on the one hand, it is necessary to choose a new evaluation criterion to accurately analyze the training and prediction of each class of samples to determine whether the high accuracy of a model is the result of being “cheated” by an unbalanced dataset. The confusion matrix was used as an additional evaluation criterion in this study. A confusion matrix is an effective tool for evaluating the accuracy of multiclassification problems in supervised learning and can calculate multiple evaluation metrics suitable for multiclassification problems.

On the other hand, after determining that the effect of a model has indeed been affected by the problem, action needs to be taken for it. The treatment of dataset imbalance problem is generally considered from two aspects: algorithm and data. The former refers to using some optimization methods in the machine learning algorithm to incorporate the difference in the number of classes into the impact on the parameters as much as possible. The latter refers to trying to change the imbalance in the dataset itself. This study chose to approach this question from a data perspective.

Without considering the fundamental adjustment of the research scope and classification method, the proportion of the original categories of the dataset cannot be changed. It is necessary to adjust the number of various types of samples through resampling to realize the reconstruction of the dataset. Resampling methods include undersampling (reducing the number of samples) for large classes and oversampling (increasing the number of samples) for small classes. Still, in general, undersampling will result in the loss of more information for the class of samples. Therefore, this study only adopts the oversampling method to expand the number of small class samples to narrow the gap between classes and solve the imbalance problem. Of course, this method will change the distribution of the original category, and further, it will forcefully reverse the preferences of the original model, which will undoubtedly cause new problems. As a compromise choice, this study will not resample a total balance of the number of each class when reconstructing the dataset. Still, there is a relative balance so that the proportion gap between the classes is not too large.

Taking a noncomplete LCZ model that needs to be added to the adjacent grid filling algorithm as an example, the reconstruction process of the dataset is briefly explained, and the sample data is shown in Table 6. As an incomplete model, the corresponding dataset needs to be expanded in quantity to ensure that the total number of samples is not too small. Assuming that the original number of 9 classes of LCZ samples is unbalanced, the initial total number of samples is the sum of 9 types of 2475, and the original expansion ratio is 6 times, the total number of samples that need to be expanded is 14850. Suppose we want to achieve absolute equality of the sample sizes of all types. In that case, 1650 of the total is the target value of each class of samples. Before the correction, the resampling ratio can be obtained by dividing the target value by the original number. A specific manual correction is made to this group of magnification values in the original distribution of various proportions. There is a specific difference between the multiplied resampling quantities, and the multiple classes are relatively balanced. After resampling, the total number of samples is 12850, which is not much different from the total number after the original augmentation.

As for the specific source of sample expansion, a sample slice with a complete neighborhood grid, 90° rotation, mirroring, etc. can generate new samples, and one sample can be changed to a maximum of 8; that is, the maximum magnification is 8 times; for slices with incomplete neighborhood grids, the expansion space depends on the degree of “incompleteness”; for example, any one of the 8 grids in the inner circle lacks any one of the corresponding \( C_3^2 = 8 \) possibilities, lack any 2 corresponding \( C_3^2 = 28 \) possibilities, and so on, different incomplete cases have different expandable magnifications, and the minimum is 8 times. When the dataset is unbalanced, this study will set the corresponding resampling ratio for each dataset based on the above process. Reconstruct the distribution of the dataset to reduce the negative impact of imbalance and train to obtain a better model effect.
2.3. Development of Single-Grid Basic Algorithm. ANN is generally composed of the input, hidden, and output layers; each layer contains several neuron nodes, and directed weighted arcs are used to connect each node. Build an ANN algorithm model and load the processed dataset for training. The input data is the three-dimensional (land property, LCZ, traffic intensity) data in the neighborhood of the unknown grid, and the output result and the corresponding label are the two-dimensional (land property, LCZ) data of the unknown grid. In reality, road traffic often exists before the construction of urban land. The subsequent urban design generally does not change the traffic conditions around the design land, so the traffic intensity data is only used as an input and not an output. In addition, to avoid the output results being too complicated to parse, this study separates the output of land use and LCZ and trains two types of models. The two datasets of $3 \times 3$ neighborhood and $5 \times 5$ neighborhood finally correspond to four models.

For the convenience of discussion, the numbering method of defining the dataset is neighborhood range_number of grids in the inner circle (number of grids in the outer circle), the same below. Among them, the dataset corresponding to the $3 \times 3$ neighborhood is numbered by the first two groups of numbers: neighborhood range_inner circle grid number. For example, $3_8$ means that, in the $3 \times 3$ neighborhood, the data of the 8 grids in the inner circle is used as the input dataset; the dataset corresponding to the $5 \times 5$ neighborhood uses 3 sets of numbers: neighborhood range_number of grids in the inner circle_number of grids in the outer circle; for example, $5_8_16$ means that, in a $5 \times 5$ neighborhood, the data of 8 grids in the inner circle plus 16 grids in the outer circle, a total of 24 grids are used as input dataset.

In the specific training process, the dataset shall be randomly divided into training set and testing set in the ratio of $7:3$, and the built ANN shall be input for certain epoch times of training. The training set is used as the known data input to the neural network for learning, and the test set is used to evaluate the performance of the model on the unknown dataset. Through repeated training of known data, the network gradually adjusts and optimizes the weight value of the connection between nodes and learns the relationship between input and output. In addition to the weight parameters, there are also superparameters that need to be manually defined, such as the number of hidden layers and the number of neuron nodes in each layer. It is often impossible to find the parameter setting that can make the network play the best performance immediately when the network is initially built. It is necessary to evaluate the actual performance of the network on the dataset through the evaluation function. Based on this, the parameter setting of the network is repeatedly adjusted and optimized, and finally a model that can reflect the internal law of urban spatial layout in the corresponding area of the dataset is obtained.

The evaluation function generally refers to the loss function and accuracy function of the model. The worse the weight of the neural network is, the worse the performance of the neural network is calculated automatically. In this study, the cross-entropy error function is used as the loss function:

$$E = - \sum_i t_i \log y_i.$$  \hspace{1cm} (2)

In formula (2), $E$ is the cross-entropy error, $y_i$ is the neural network output of the $i$th sample, $t_i$ is the correct unlabeling of the $i$th sample, and $y_i$ and $t_i$ are represented by one-hot coding.

The calculation of accuracy can more intuitively show the quality of network performance. Different tasks will define different accuracy functions. For some simple problems, it can be calculated directly by judging whether the generated result is "equal to" label. This method is called “equivalence” evaluation in this study. When the result is more complex, and the judgment of the original result will be affected by simply summarizing it with the word “right” or “wrong,” the “Deviation” evaluation of the result can be obtained by calculating the difference between the generated result and the label data. This study combines the two standards of equivalence and deviation to score the two dimensions of output.

### 3. Analysis and Discussion

It is one of the important reasons for this study to explore the laws existing in urban spatial layout data from the perspective of computers. Given the use of an artificial neural network as the main model algorithm in this study, exploring the layout rule is equivalent to exploring the connection between the input and output data, that is, analyzing the possible relationships and laws between the current situation data of the surrounding environment input by the sample and the prediction results of the central grid output. Before conducting this analysis, three questions need to be clarified: What is the sample to be analyzed? What does surrounding data mean? What is the forecast result?

For question 1, it is evident that the model’s performance on the unknown test set can better reflect the real learning situation than the training set, so the rule analysis is mainly based on the test set samples of 90 models generated by the training. Moreover, no matter whether the score is high or low, the model’s prediction of the sample will follow certain internal logic and laws. Still, the laws embodied by the
samples with high scores conform to the real situation, while the samples with low scores do not. To explore the realistic part of the rules found by the model, in this study, the test set samples are divided into different levels according to the scores of the prediction results, and the in-depth analysis is performed based on the best performing part of the samples.

For question 2, the surrounding data refer to the grid’s land use, LCZ, and traffic intensity data in the neighborhood. For question 3, the prediction results refer to the land use properties and LCZ prediction results of the unknown grid in the middle. And the results of land use properties can be subdivided into 16 classes: there are only one dominant class (5 class), two dominant classes (10 class), and no obvious dominant class (1 class); LCZ prediction results can be subdivided into 9 classes, corresponding to 9 LCZ classes. Figure 2 shows the relationship between the above three answers and briefly illustrates the steps of law analysis.

After the analysis object is defined, the analysis is carried out according to the following steps: (1) divide the test set samples into different grades (grades) according to the prediction score, and briefly describe the general situation of samples at each grade. (2) Observe whether the samples with good performance have common neighborhood characteristics. (3) Take the best performing class A samples and briefly describe the general situation of the prediction results of each subdivision category. For (4) the nature of land use and (5) LCZ, deeply analyze the relationship between the surrounding current situation and the data of the prediction results in the middle, and summarize the possible laws in the urban spatial layout.

The test set samples are divided into different grades according to the result score, in which the land property results are divided into five grades, and LCZ has only two grades of correct or wrong prediction. Table 7 shows the specific number and proportion of samples at each level.

As can be seen in Table 7, the prediction performance of 5×5 neighborhood is better than 3×3 neighborhood, and the prediction performance of land property is better than LCZ, which is in line with the conclusions obtained from the analysis above.

Comparing the conclusions of land use property model and LCZ model, it is found that the samples with good prediction performance are the same, but some characteristics of their surrounding environment are opposite. For example, after the proportion of landscape land increases, the prediction effect of land use property is worse, but the prediction effect of LCZ is better. This shows that the model has great differences in the prediction law of land use properties and LCZ.

### 4. Conclusions

This study comprehensively analyzes and evaluates the artificial intelligence generation method of urban spatial layout from two aspects: algorithm and result. In the algorithm analysis, firstly, the 90 ANN models of the example city are analyzed and evaluated based on various scoring indicators, including the analysis of the overall performance of the model and the comparative analysis of the land use property model and LCZ model. The analysis
shows that the overall effect of the ANN algorithm is better, and the performance of the land use property model is better than LCZ model, and several factors affecting the performance of the model are summarized. The results show that the adjacent lattice algorithm is better than MCTS algorithm. In the result analysis, the three multigrid test results generated based on the adjacent lattice filling algorithm in the example application are briefly analyzed and interpreted, and then the sample data of the test set based on 90 ANN models are deeply discussed. The relationship between the surrounding current situation and the prediction results in the middle is analyzed, from which a series of more reasonable urban spatial layout laws in the research area are summarized. It is proved that this research method has certain practical value.

Data Availability

The raw data supporting the conclusions of this article can be obtained from the corresponding author upon request.

Conflicts of Interest

The authors declared that they have no conflicts of interest regarding this work.

References

[1] J. Hua, M. Cai, Y. Shi et al., “Investigating pedestrian-level greenery in urban forms in a high-density city for urban planning,” Sustainable Cities and Society, vol. 80, Article ID 103755, 2022.

[2] F. T. da Silva, J. M. Santos, N. C. Reis Jr, E. V. Goulart, and C. E. de Alvarez, “Influence of urban form on air quality: the combined effect of block typology and urban planning indices on city breathability,” Science of the Total Environment, vol. 814, Article ID 152670, 2022.

[3] V. Carter and C. Henríquez, “Can Strategic Environmental Assessment (SEA) contribute towards the implementation of biophilic urbanism in urban planning? The case of Chilean Municipal Regulatory Plans,” Environmental Impact Assessment Review, vol. 95, Article ID 106765, 2022.

[4] W. Wang, K. Liu, M. Zhang, Y. Shen, R. Jing, and X. Xu, “From simulation to data-driven approach: a framework of integrating urban morphology to low-energy urban design,” Renewable Energy, vol. 179, pp. 2016–2035, 2021.

[5] Y. J. Kim and R. D. Brown, “A multilevel approach for assessing the effects of microclimatic urban design on pedestrian thermal comfort: the High Line in New York,” Building and Environment, vol. 205, 2021.

[6] E. Ortega, B. Martin, M. E. López-Lambas, E. Maria, and J. A. Soria-Lara, “Evaluating the impact of urban design scenarios on walking accessibility: the case of the Madrid “Centro” district,” Sustainable Cities and Society, vol. 74, Article ID 103156, 2021.

[7] J. Huang and Z. Wang, “The use of scientific programming method under visual communication design concept in urban landscape design,” Scientific Programming, vol. 2022, Article ID 8661969, 10 pages, 2022.

[8] J. Wan and H. Shi, “Research on urban renewal public space design based on convolutional neural network model,” Security and Communication Networks, vol. 2021, Article ID 9504188, 9 pages, 2021.

[9] X. Zhang, W. Fan, and X. Guo, “Urban landscape design based on data fusion and computer virtual reality technology,” Wireless Communications and Mobile Computing, vol. 2022, Article ID 7207585, 14 pages, 2022.

[10] L. Liu, “Urban complex public space design method based on support vector machine,” Mathematical Problems in Engineering, vol. 2022, Article ID 9812223, 13 pages, 2022.

[11] J. He, “Landscape design method of urban wetland park using the building information model,” Wireless Communications and Mobile Computing, vol. 2022, Article ID 6228513, 10 pages, 2022.

[12] X. Ma, J. Li, and X. Zhang, “Data model of key indicators of urban architectural design based on long-and short-term convolutional memory network,” Mathematical Problems in Engineering, vol. 2022, Article ID 7607928, 13 pages, 2022.

[13] S. Liu, “Application of big data technology in urban greenway design,” Source: Security and Communication Networks, vol. 2022, Article ID 4826523, 10 pages, 2022.

[14] L. Li, B. Lei, and C. Mao, “Digital twin in smart manufacturing,” Journal of Industrial Information Integration, vol. 26, no. 9, Article ID 100289, 2022.

[15] A. Naseem, M. A. Rehman, and J. Younis, “A new root-finding algorithm for solving real-world problems and its complex dynamics via computer technology,” Complexity, vol. 2021, Article ID 6369466, 10 pages, 2021.

[16] L. Li, T. Qu, Y. Liu et al., “Sustainability assessment of intelligent manufacturing supported by digital twin,” IEEE Access, vol. 8, pp. 174988–175008, 2020.

[17] L. Li and C. Mao, “Big data supported PSS evaluation decision in service-oriented manufacturing,” IEEE Access, vol. 8, no. 99, pp. 154663–154670, 2020.

[18] H. Zhu, “The optimization function of computer image technology in processing oil painting creation,” Wireless Communications and Mobile Computing, vol. 2022, Article ID 3188527, 6 pages, 2022.

[19] L. Li, C. Mao, H. Sun, Y. Yuan, and B. Lei, “Digital twin driven green performance evaluation methodology of intelligent manufacturing: hybrid model based on fuzzy rough-sets AHP, multistage weight synthesis, and PROMETHEE II,” Complexity, vol. 2020, Article ID 3853925, 24 pages, 2020.

[20] Y. Wang and L. Chen, “Architectural and landscape garden planning integrated with artificial intelligence parametric analysis,” Security and Communication Networks, vol. 2022, Article ID 8577269, 9 pages, 2022.

[21] H. Dong, X. Liang, Y. Liu, and D. Wang, “5G virtual reality in the design and dissemination of contemporary urban image system under the background of big data,” Wireless Communications and Mobile Computing, vol. 2022, Article ID 8430186, 14 pages, 2022.

[22] G. Wu, Y. Miao, and F. Wang, “Intelligent design model of urban landscape space based on optimized BP neural network,” Journal of Sensors, vol. 2022, Article ID 9704287, 10 pages, 2022.

[23] N. Zhang, H. Chen, X. Chen, and J. Chen, “ELM meets urban big data analysis: case studies,” Computational Intelligence and Neuroscience, vol. 2016, Article ID 4970246, 10 pages, 2016.
[24] X. Yang and F. Bai, “Three-dimensional structure analysis of urban landscape based on big data technology and digital technology,” *Scientific Programming*, vol. 2021, Article ID 7970870, 10 pages, 2021.

[25] Li Zhan and K. Ning, “Minority tourist information service and sustainable development of tourism under the background of smart city,” *Mobile Information Systems*, vol. 2021, Article ID 6547186, 14 pages, 2021.

[26] K. Wang, F. Yan, Y. Zhang, Y. Xiao, and L. Gu, “Supply chain financial risk evaluation of small- and medium-sized enterprises under smart city,” *Journal of Advanced Transportation*, vol. 2020, Article ID 8849356, 14 pages, 2020.

[27] Y. Li and H. Mao, “Study on machine learning applications in ideological and political education under the background of big data,” *Scientific Programming*, vol. 2022, Article ID 3317876, 9 pages, 2022.

[28] Y. Zhang, F. He, J. Kong, Y. Zhu, and L. Wang, “Relationship between surface subsidence range and geological mining conditions using numerical simulation and machine learning,” *Scientific Programming*, vol. 2022, Article ID 8720831, 12 pages, 2022.

[29] X. Xia and Z. Tong, “A machine learning-based method for predicting urban land use,” in *Proceedings of the 25th International Conference on Computer-Aided Architectural Design Research in Asia (CAADRIA)*, pp. 21–30, Bangkok, Thailand, August 2020.

[30] R. Senington, B. Schmidt, and A. Syberfeldt, “Monte Carlo Tree Search for online decision making in smart industrial production,” *Computers in Industry*, vol. 128, Article ID 103433, 2021.