Urbanization Process: A Simulation Method of Urban Expansion Based on RF-SNSCNN-CA Model

Minghao Liu 1,2,*, Xiangli Liao 1,2 and Chun Chen 3,4

1 School of Computer Science and Technology, Chongqing University of Posts and Telecommunications, Chongqing 400065, China
2 Spatial Information Research Center, Chongqing University of Posts and Telecommunications, Chongqing 400065, China
3 School of Architecture and Urban Planning, Chongqing Jiaotong University, Chongqing 400074, China
4 Research Center for Ecological Human Settlements and Green Transportation, Chongqing Jiaotong University, Chongqing 400074, China

* Correspondence: liumh@cqupt.edu.cn

Abstract: One of the focal points in Geographic Information Science (GIS) is to uncover the spatial distribution patterns of geographical phenomena. In response to the insufficient spatial feature learning concerning neighborhoods in traditional machine learning-based Cellular Automata (CA) models for land use change, this study couples the Random Forest (RF) model and the Spatially Non-Stationary Convolutional Neural Network (SNSCNN) model to the CA model. The resulting RF-SNSCNN-CA model considers the issue of spatial non-stationarity by incorporating attention mechanisms. Using observed urban land change data from 2010 to 2017 in the 21 districts of Chongqing’s main city as an example, two sets of experiments comprising eight scenarios were designed to verify the neighborhood effects. The results demonstrate that the proposed RF-SNSCNN-CA model achieves an Overall Accuracy (OA) of 97.82%, Kappa of 0.7683, and Figure of Merit (FoM) of 0.3836. The study reveals the following findings. Firstly, the RF-SNSCNN-CA model integrates the dual advantages of traditional machine learning and deep learning models, in which SNSCNN improves by the combined effect of channel and spatial attention mechanisms improves the learning of neighborhood features; secondly, the machine learning-like urban sprawl CA modeling process, regardless of the approach taken to obtain development suitability, cannot completely replace the learning of the neighborhood part; lastly, the use of traditional neighborhood modeling methods may produce suppression of simulation results and make the model inadequately learn spatial features.

Keywords: RF-SNSCNN-CA coupling model; neighborhood effect; attention mechanism; urban expansion; convolutional neural network

1. Introduction

Land use change is one of the main research components of global environmental change and sustainable development [1]. Land use change simulation has become an important tool for land use change research with a wide range of applications, including the urbanization process, ecological environmental protection, and agricultural development [2–4]. The urban sprawl simulation model can be considered a special case of land use simulation model, focusing the simulation object on the urban boundary.

At present, there are three main aspects of research on land use change. One is the study of the drivers of land use change [5–7]. By analyzing a large number of experimental cases, researchers study the drivers of land use change by cutting from the most important factors, such as natural and economic multiple types. Second, the study of land use evolution mechanisms using relevant models to predict land use changes and future trends [8]. In this process, researchers have used a series of models, including statistical models and dynamic models, to better understand the mechanisms and trends of land use...
change. Third, multi-scenario land use simulation [9–11]. In this study, researchers used land use change models to simulate future land use based on different scenarios. These scenarios include economic development, population growth, policy changes, etc., and aim to predict the impacts of land use changes on ecosystems, economy, and society and to propose effective management strategies.

Currently, land use change simulations mainly simulate regional land use changes by coupling quantitative and spatial prediction models [12]. Among them, quantitative prediction models include logistic regression models, gray prediction models, Markov models, and system dynamics models [13–16], while spatial prediction models include CA models, PLUS (Patch-generating Land Use Simulation) models, and multi-intelligence models, etc. [17–19]. Although each of these two models has certain limitations, coupled quantitative and spatial prediction models can predict future land use conditions more accurately. Therefore, more and more scholars tend to adopt this coupled approach for their research.

Since the CA model was used for the evolution of complex systems in the 1970s, it has now been widely applied to simulate the spatial and temporal dynamics of land use and has achieved very good results [20]. Compared with other discrete system models, such as the Petri-net model and linear programming model, CA models can generate complex global behaviors and structures through simple local rules and interactions in a “bottom-up” manner [21,22]. The CA model consists of a rule grid in which each cell has discrete states, cells update their states at discrete time steps, and the new state of a cell depends on its own and neighboring cells’ current states, as well as on predefined transfer rules [23]. A large number of scholars have introduced some new driving factor feature mining methods to mine the transition rules of the model by improving the traditional CA model. In land use simulation, the metacell conversion rules and land use change drivers are closely related [24], and scholars have used a variety of different algorithms to extract metacell conversion rules in order to fully explore their intrinsic mechanisms.

Among the methods for extracting metacell conversion rules, many studies focus on mining the conversion rules by analyzing the driving factors of the metacell itself. BAIG et al. used an artificial neural network algorithm coupled with a CA model to predict land use change in Selangor, Malaysia, between 2031–2051 [25]. KARIMI et al. designed a support vector machine model to deal with the nonlinear relationship between driving factors and land use change and used it for urban sprawl prediction in Guilford County, North Carolina [26]. HAGENAUER et al. conducted a comparison of land use simulations using 38 machine learning methods and finally found that random forests outperformed most methods overall [27].

In contrast to the methods based on the tuple’s own driving factors to extract transformation rules, the neighborhood effect considers the influence of the tuple’s neighbors on the central tuple and therefore has been widely used and studied. Many scholars have improved the neighborhood effect to improve the extraction of metacell conversion rules. Since CNNs can be used to extract high-level features of raster data, CNNs and their improved models are widely used to extract neighborhood drivers. HE et al. first used CNNs to mine the neighborhood features of drivers. HE argued that previous studies did not sufficiently consider neighborhood effects in the rule mining process, and they used Convolutional Neural Network for United Mining (CNN) in their study. Neural Network for United Mining (UMCNN) to solve this problem [28]. WU et al. combined multi-labeling with convolutional neural networks and coupled CA models to simulate the complex evolution of land use [29]. XIAO et al. proposed the CNN-GRU (Convolutional Neural Network-Gated Recurrent Unit) model for mining the temporal and spatial characteristics of multi-period land use changes [30].

Current land use simulation models portray the patterns of land use change from different perspectives; however, spatial non-smoothness in the process of land use change is a phenomenon that cannot be ignored, and spatial non-smoothness refers to the existence of spatial data heterogeneity at different scales [31]. Spatial non-smoothness has important
research value in many fields, such as earth science, ecology, and economics [32–34]. Geographically weighted regression methods and partitioned modeling methods are typical applications of the strategy of “seeking common ground while reserving differences” for spatial non-stationarity in spatial modeling from non-neighborhoods [35–37]. With the rapid development of big data, deep learning, and other technologies, many new methods have emerged in spatial nonsmoothness modeling in recent years. Some studies have used attention mechanisms to build deep learning models that adapt to the nonsmoothness of spatial data [38]. The attention mechanism can help the model automatically select the input features that need attention, thus improving the accuracy of the model. In simple terms, an attention mechanism is a weighted averaging method that assigns different weights to different input features based on their importance [39]. CBAM (Convolutional Block Attention Module) incorporates both channel and spatial attention and can be embedded into a convolutional neural network (Convolutional Neural Network (CNN)) [40]. By using CBAM, the network can better focus on the important features, thus improving accuracy and efficiency. LI et al. found that CNN-CBAM-based simulation models were able to extract features from spatial data more accurately than CNN models with traditional structures [41]. In general, CNN and its coupled model in deep learning are more suitable methods for extracting features of spatial data, but the coupled model of CNN and CBAM has not been applied to the mining of CA-like model neighborhood driving factor features.

Based on this, this paper tries to construct an RF-SNSCNN-CA model, that is, on the basis of using the traditional machine learning model RF to obtain the probability of site suitability, the SNSCNN method with the introduction of attention mechanism is used to optimize the learning of neighborhood features to improve the spatial heterogeneity modeling effect in order to better reveal the urban development evolution law. Several sets of comparative experimental scenarios are designed to verify the performance of the RF-SNSCNN-CA model during the model implementation.

2. Materials and Methods

2.1. Study Area and Data

Chongqing is one of the important central cities in China, the economic center of the upper reaches of the Yangtze River, the core city of the twin-city economic circle in the Chengdu-Chongqing region, an important advanced manufacturing center in China, a financial center in western China, an international comprehensive transportation hub in western China, and an international gateway hub in China.

Chongqing’s primary urban metropolitan area, known as District 21, stands as a remarkable exemplar of a “mountain city” characterized by its undulating and mountainous topography. Situated within the parallel valley of eastern Sichuan, Chongqing’s main metropolitan area shares borders with Sichuan province and encompasses a sprawling expanse spanning 28,700 square kilometers. This vibrant urban region encompasses several distinct zones, including the core area known as the “two rivers and four banks,” the central urban area, four pioneering co-city development zones, four strategically pivotal cities, and four bridgehead cities. Each of these zones is endowed with unique development objectives, as illustrated in Figure 1.

The study area is located in the upper reaches of the Yangtze River and belongs to the humid subtropical climate zone. Due to the complex topography and large differences in terrain elevation, there are obvious vertical variations in the climate of the region. The average annual temperature in the urban area is about 18 °C, the annual precipitation is above 1000 mm, and the rainy season is mainly concentrated in summer, while the winter is relatively dry.

In terms of land use, the land types in the study area are relatively diverse. Mountains are the main topography of the region, occupying most of the total area of the municipality. Due to the large difference in elevation, the mountainous terrain is complex, and the mountains at different elevations have different land types, such as alpine meadows and deciduous broadleaf forests. In addition, the region has land types such as hills, plains,
and river valleys. In recent years, with the acceleration of urbanization, the area of urban construction land has been expanding year by year.

Figure 1. Twenty-one districts of Chongqing main city.

The intricate interplay of diverse natural conditions and socio-economic factors within Chongqing’s primary metropolitan area has established a robust regional framework for effectively modeling spatial heterogeneity.

The land use data of the study area include two periods of 2010 and 2017 (Figure 2), both from the Institute of Geographical Sciences and Resources, Chinese Academy of Sciences (https://www.resdc.cn, accessed on 4 May 2021). This study focuses on the expansion of urban land use and therefore focuses on the conversion of non-urban land use to urban land use, so ArcGIS is used to reclassify the land use data into urban land, non-urban land, and water area according to the first-level classification of the study area.

Figure 2. 21 districts of Chongqing main city.

The driving forces behind land use change typically encompass a range of factors, including natural, economic, and social factors [7]. In this study, we selected elevation and slope data as representative of natural factors. Economic factors were represented by density data of hospitals, shopping centers, schools, and attractions. Social factors were represented by distances to various amenities and infrastructure, such as airports, town
centers, bus stations, subway stations, railroad stations, highways, water areas, and light intensity at night. The slope data was derived through DEM (Digital Elevation Model) calculations, while DEM and light data were obtained from WorldPop. Water body and road network data were sourced from OpenStreetMap, and points of interest were collected using a crawler via AMAP's API. To facilitate the model construction, the collected data were transformed into a uniform 100 × 100 raster format. Furthermore, the data underwent normalization, and the resulting raster format was utilized as input data, as illustrated in Figure 3.

![Figure 3. Driving factors of 21 districts in Chongqing city.](image)

2.2. Methods

2.2.1. Research Idea and Test Scheme Design

As depicted in Figure 4, the technical workflow of the RF-SNSCNN-CA model in this study encompasses four significant modules: the data processing module, the model training module, the iterative calculation module, and the model application module.

The data processing module entails the acquisition and processing of multi-source heterogeneous data. This encompasses three types of driving factor data, namely natural factors, economic factors, and social factors, along with two phases of land use data. Following essential cleaning and processing steps, the data undergo vector to raster conversion, rasterization, and resolution normalization to create a dataset with consistent resolution and coordinates. In the model training module, a partial sampling approach is employed on the dataset for training purposes. In the iterative calculation module, predictions are made using the complete dataset.
The simulated results are then compared with the actual land use distribution data from probability, specifically for the classification of 1, as the neighborhood effect. In the model validation phase, the trained RF-SNSCNN-CA model is employed to simulate the predicted land use changes for the final year (2017) based on the base year data (2010). The land use change data are generated by overlaying the 2010 and 2017 land use data, resulting in a new layer. A value of 1 in this layer indicates a raster that has transformed from non-urban to urban, while 0 represents a non-urban raster that has not undergone a transformation. To address the imbalanced distribution of urban and non-urban rasters, the sample dataset consists of 10,000 samples labeled as 1 and 30,000 samples labeled as 0. The development suitability and neighborhood features are then extracted using the driver data as input for the development suitability extractor and the neighborhood effect extractor, respectively. SNSCNN serves as the neighborhood effect extractor, and the Softmax classifier of the SNSCNN model extracts the classification probability, specifically for the classification of 1, as the neighborhood effect.

The iterative computation module consists of model validation and model application. In the model validation phase, the trained RF-SNSCNN-CA model is employed to simulate the predicted land use changes for the final year (2017) based on the base year data (2010). The simulated results are then compared with the actual land use distribution data from 2017 to validate the effectiveness of the model proposed in this paper. Following the validation, the model application phase utilizes the 2017 land use data as input, and the established model is utilized to predict the land use status for the year 2024. In Figure 4 we use gray to show non-urban areas and red to show urban areas.

Figure 4. RF-SNSCNN-CA model technology flow chart.

The model training module consists of two main components: RF-based development suitability extraction and SNSCNN neighborhood feature extraction. Initially, the driver data and land use change data from the observation period are sampled to create a training dataset. This dataset is then used as input for training the RF and SNSCNN models, which construct the development suitability extractor and the neighborhood effect extractor, respectively. The land use change data are generated by overlaying the 2010 and 2017 land use data, resulting in a new layer. A value of 1 in this layer indicates a raster that has transformed from non-urban to urban, while 0 represents a non-urban raster that has not undergone a transformation. To address the imbalanced distribution of urban and non-urban rasters, the sample dataset consists of 10,000 samples labeled as 1 and 30,000 samples labeled as 0. The development suitability and neighborhood features are then extracted using the driver data as input for the development suitability extractor and the neighborhood effect extractor, respectively. SNSCNN serves as the neighborhood effect extractor, and the Softmax classifier of the SNSCNN model extracts the classification probability, specifically for the classification of 1, as the neighborhood effect.

2.2.2. Attention Module

Convolutional neural networks (CNNs) have shown excellent performance in extracting features from grid data. However, traditional CNNs treat all feature maps in each channel equally, disregarding the important differences among them. This approach fails to account for the varying densities of information contained in each channel. Additionally, the traditional pooling operation assumes equal importance for all spatial locations in the matrix, which may compromise model accuracy and efficiency by neglecting the importance of spatial differences.
To address these limitations, the attention mechanism is introduced. It assigns different weights to different pieces of information, emphasizing important information with larger weights and de-emphasizing unimportant information with smaller weights [21]. By incorporating the attention module into the CNN architecture, the model can capture more comprehensive and targeted information in a more sensible manner, leading to improved accuracy and efficiency.

In this study, the importance of driving factors is captured using channel attention through the neighborhood effect, while the variability of neighborhood influence is captured using spatial attention. This integration of attention mechanisms enables the model to effectively capture and utilize information, as illustrated in Figure 5.

**Figure 5.** Channel attention and spatial attention mechanism.

Its calculation formula is shown as

\[ F' = M_c(F) \odot F \]  \hspace{1cm} (1)

\[ F'' = M_s(F') \odot F' \]  \hspace{1cm} (2)

where \( F \) is the input data of size \( C \times H \times W \), i.e., \( C \) channels with each channel data of size \( H \times W \), which in this paper represents the neighborhood size of \( 51 \times 51 \).

CBAM first aggregates the spatial information of the feature map using the average pooling operation and the maximum pooling operation to generate two different spatial context descriptors: \( F_{avg} \) and \( F_{max} \), which represent the average pooling feature and the maximum pooling feature, respectively. The two descriptors are then passed to the weight sharing network to generate the channel attention matrix \( M_c(F) \). Element-level summation is used to merge the output feature vectors. As shown, the channel attention is calculated as

\[ M_c(F) = \sigma(\text{MLP(AvgPool}(F)) + \text{MLP(MaxPool}(F))) \]  \hspace{1cm} (3)

where \( \sigma \) is the sigmoid function. The intermediate data of \( F' \), which is also of size \( C \times H \times W \), is generated after channel emphasis by the channel attention layer.

To compute the spatial attention, the average pooling and maximum pooling operations are first applied along the channel axis, and they are concatenated to generate a valid feature descriptor. On the connected feature descriptors, a convolutional layer is applied to generate a spatial attention matrix \( M_s(F) \). This matrix emphasizes or suppresses the spatial locations by assigning weights. The spatial attention is calculated as

\[ M_s(F) = \sigma(f_{7 \times 7}([\text{AvgPool}(F); \text{MaxPool}(F)])) \]  \hspace{1cm} (4)
where $\sigma$ is a sigmoid function and $f^{7 \times 7}$ denotes a convolution operation with a convolution kernel size of $7 \times 7$.

2.2.3. SNSCNN Model Construction and Training

To address the issue of convolutional neural networks’ inability to selectively focus on high-density information regions, the integration of an attention module with CNN architecture allows the network to prioritize and extract features from areas with higher information density. By emphasizing important features and suppressing irrelevant ones, the accuracy and efficiency of CNN can be improved.

In this study, a CNN model is constructed, as illustrated in Figure 6. To capture the spatial non-smoothness of the driver data, attention layers are connected to the input data of the driver. This enables the model to effectively extract relevant spatial information and enhance its ability to identify important features.

![Figure 6. Structure of SNSCNN model.](image)

(1) Each of the 15 driver layers is cut into subgraphs of size $51 \times 51$, which are enhanced and used as the input of the CNN using the attention mechanism.

(2) Multi-scale features are obtained after several convolution and pooling operations. First, $49 \times 49 \times 16$ data are obtained by $2 \times 2$ convolution of the first layer, and then a $24 \times 24 \times 16$ intermediate layer is obtained by $2 \times 2$ pooling; second, a $22 \times 22 \times 32$ intermediate layer is obtained by $3 \times 3$ convolution, and then a $10 \times 10 \times 32$ intermediate layer is obtained by $3 \times 3$ pooling operation; then, a $2 \times 2$ convolution operation is used, while a 20% Dropout operation is applied to it to avoid overfitting.

(3) A fully-connected layer of $84 \times 1$ is obtained by using a fully-connected operation. The final layer is obtained by using the fully-connected operation with $2 \times 1$ data, and the mapping relationship between the driver and the label is established using the SoftMax classification function, which is the final classification result.

The SNSCNN model is trained using fifteen driver layers as input. These layers are cropped into multiple $51 \times 51$ matrices centered on the label coordinates, and they are used along with the labels for training. The model achieves a training set accuracy of 0.8894.

2.2.4. CA Model Construction

The CA model is comprised of five fundamental components: tuple and tuple space, tuple state, neighborhood, time, and transformation rules. The acquisition of transformation rules is the central aspect of CA model construction. A comprehensive probabilistic model based on CA typically consists of four elements: land development suitability, neighborhood constraint, restrictive constraint, and random perturbation, as depicted in Equation (5):

$$P = d_{ij} \times \Omega_{ij} \times r_{ij} \times t_{ij}$$

where $P$ is the integrated probability of raster $(i,j)$; $d_{ij}$ is the development suitability of raster $(i,j)$; $\Omega_{ij}$ is the neighborhood constraint of the current raster; $r_{ij}$ is the global
constraint of the current raster, which is set to watershed without conversion to urban in this study; $t_{ij}$ is the random perturbation, which is set to a random number from 0 to 1 in this study. In this model $d_{ij}$ and $\Omega_{ij}$ are the keys to learning the conversion rules.

Previous studies have demonstrated that RF stochastic classifiers outperform other machine learning classifiers and regression methods in terms of classification and fitting accuracy. Therefore, in the CA iterative model presented in this paper, RF is employed as the development suitability extractor to calculate the suitability of land for urban development.

In addition, the SNSCNN model is introduced to calculate the neighborhood effect under multi-scale regional characteristics, and the overall conversion probability of all current non-urban rasters is obtained by combining the random perturbation and restriction constraints in the CA model. During the simulation, the simulation is performed based on the base period year land-use data, and the total number of urban rasters at the end of the land use year is used as the demand constraint to convert sites from high to low based on the overall probability until the total demand is met. After multi-period CA iterative simulation, the result of urban land expansion in the target year is obtained. The flow chart is shown in Figure 7.

\[
P_0 = \frac{s}{n} \tag{6}
\]

\[
P_c = \frac{(a_1 \times b_1 + a_0 \times b_0)}{n^2} \tag{7}
\]

\[
Kappa = \frac{(P_0 - P_c)}{(1 - P_c)} \tag{8}
\]

**Figure 7.** Flow chart of CA iterative model.

### 2.2.5. Evaluation Method of Urban Expansion Simulation Accuracy

In this paper, the Kappa coefficient, the FoM (Figure of Merit) value, and the OA are selected as evaluation indexes, where the Kappa coefficient is used for the consistency test of the model; FoM value is the quality factor; OA is the overall accuracy rate.

Kappa is calculated as follows:

\[
P_0 = \frac{s}{n} \tag{6}
\]

\[
P_c = \frac{(a_1 \times b_1 + a_0 \times b_0)}{n^2} \tag{7}
\]

\[
Kappa = \frac{(P_0 - P_c)}{(1 - P_c)} \tag{8}
\]
where the total number of raster is \( n \), the number of raster that are truly urban is \( a_1 \), the number of the non-urban raster is \( a_0 \), the total number of predicted conversions to urban is \( b_1 \), the number of predicted conversions to non-urban is \( b_0 \), and the number of the predicted accurate raster is \( s \).

\( \text{FoM} \) is calculated by Equation (9):

\[
\text{FoM} = \frac{B}{A + B + C + D}
\]

where \( A \) is the number of raster that actually changed but were predicted not to change, \( B \) is the number of raster that actually changed and were predicted to change, and \( C \) is the land use category that actually changed but was predicted to be wrong. Since this study only deals with the non-urban-urban transition, the value of \( C \) is 0, and \( D \) is the number of raster that actually did not change but were predicted to change.

2.2.6. Landscape Index Evaluation Method

To investigate the spatial characteristics of urban land in the future urban pattern of Chongqing’s main city comprising 21 districts, the landscape pattern index is employed to depict the spatial distribution of urban land as a landscape. Drawing upon commonly used landscape indices and considering the specific conditions of Chongqing’s main city, this study selects indices that effectively capture landscape density, shape, connectivity, diversity, and aggregation at both the landscape and type levels. These selected indices include the number of patches (NP), maximum patch index (LPI), patch density (PD), landscape shape index (LSI), aggregation index (AI), and effective particle size (MESH). These six indices provide insights into the area-edge relationship, aggregation patterns, and shape characteristics of landscape patches. They are utilized to analyze the density, shape, connectivity, diversity, and aggregation of urban land in the 21 districts of Chongqing’s main city within the context of the future urban pattern. By comparing and analyzing the changes in landscape indices across the 21 districts, this study elucidates the spatial characteristics transformations within Chongqing’s main city.

(1) Number of patches (NP) counts the number of patches in this type of landscape.
(2) The maximum patch index (LPI) is used to determine the landscape dominance, and the larger LPI indicates the greater dominance of this type of landscape.

\[
\text{LPI} = \frac{\max(a_1, a_2, \cdots, a_n)}{A} \times 100
\]

(3) Patch density (PD) is the ratio of the number of land use patches to the total area, which reflects the degree of fragmentation and dispersal of the landscape.

\[
PD = \frac{N_p}{A}
\]

(4) Landscape shape index (LSI) is equal to the total edge length divided by the minimum possible class edge length of the largest aggregated class. It provides a standardized measure of the total edges adjusted to the size of the landscape.

\[
\text{LSI} = \frac{\beta E}{\sqrt{A}}
\]

where \( E \) represents the total length of the landscape edge, and \( \beta \) is a correction constant, usually 0.25.
(5) Aggregation index (AI) gives the frequency of different types of patch pairs adjacent to each other. It is scaled to take into account the maximum possible number of similar
adjacencies in any landscape composition. It is used to characterize the degree to which patches are clustered with each other; a larger $AI$ indicates greater clustering.

$$AI = \left( \frac{g_{ii \max} \rightarrow g_{ii}}{g_{ii}} \right) \times 100 \quad (13)$$

where $g$ indicates the number of connections between patches of the selected type, and $max \rightarrow g_{ii}$ refers to the maximum number of similar neighboring patches between patches $i$. $0 < AI \leq 100$.

(6) The effective size of particle size ($MESH$) reflects the fragmentation of this type of landscape in terms of area.

$$MESH = \frac{\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2}{A} \quad (14)$$

where $a_{ij}$ denotes the area of the patch $j$ of the class $i$ of landscape.

3. Results

3.1. Ablation Experimental Design and Results

3.1.1. Experimental Scheme Design

In this study, an approach is proposed to enhance the accuracy of neighborhood feature learning in early urban dynamic simulation integration. By combining the strengths of traditional machine learning and deep learning models, the deep learning SNSCNN method is introduced to improve the learning of neighborhood features in conjunction with the traditional RF model. The CA model considers that urban development is primarily influenced by both its own raster development suitability and neighborhood effects.

To demonstrate the effectiveness of the RF-SNSCNN-CA model in extracting neighborhood features, two scenarios with eight experiments are designed for comparison. The goal is to verify the impact of different approaches to extracting development suitability and neighborhood effects on the simulation results. This study aims to showcase how the RF-SNSCNN-CA model enhances the accuracy and effectiveness of neighborhood feature extraction.

It has been suggested that developmental suitability can be learned directly together with neighborhood features, so the experimental schemes for extracting developmental suitability are RF [27], SNSCNN, and CNN [28], and for extracting neighborhood effects are SNSCNN, CNN, TN (traditional neighborhood), and NN (no neighborhood), and the module combinations used are shown in Table 1.

<table>
<thead>
<tr>
<th>SNSCNN</th>
<th>CNN</th>
<th>TN</th>
<th>NN</th>
</tr>
</thead>
<tbody>
<tr>
<td>RF</td>
<td>I</td>
<td>II</td>
<td>III</td>
</tr>
<tr>
<td>SNSCNN</td>
<td>×</td>
<td>×</td>
<td>IV</td>
</tr>
<tr>
<td>CNN</td>
<td>×</td>
<td>×</td>
<td>V</td>
</tr>
</tbody>
</table>

Model I, which is the recommended model in this paper, utilizes the RF-SNSCNN-CA framework. The development suitability probability is obtained using the RF method. The neighborhood constraint component incorporates the SNSCNN model to extract neighborhood features at multiple scales. The central image element’s land type, with a specific window size, serves as the label data. The driver factor data, serving as independent variables, are used as input for training the SNSCNN model, which produces the probability for the neighborhood constraint factor. The combined probabilities are obtained by combining restrictive constraints and random perturbations. Taking into account the land use requirements during the observation period, the RF-SNSCNN-CA model is established as an urban land use dynamic simulation model.
This model stands out by incorporating the channel attention mechanism to capture spatial variations in the influence of driving factors. Additionally, it utilizes spatial attention to capture neighborhood driving factor characteristics. The integrated probability calculation model is represented as follows:

\[ P_1 = d_{ij}RF \times \Omega_{ij(SNSCNN)} \times r_{ij} \times t_{ij} \] (15)

Model II uses the RF-CNN-CA model. Suitability is extracted by RF, and neighborhood effects are extracted by CNN. This model is used to compare the effect of the use or not of the attention mechanism on the neighborhood effect. Its integrated probability calculation model is:

\[ P_2 = d_{ij}RF \times \Omega_{ij(CNN)} \times r_{ij} \times t_{ij} \] (16)

Models III, IV, and V use RF, SNSCNN, and CNN to extract land development suitability, respectively, combined with CA models for urban dynamic simulation. This class of models is used as an ablation experiment to compare the effects of traditional neighborhood extraction methods on simulation results. Their integrated probability calculation models are (corresponding to III, IV, and V in the table):

\[ P_3 = d_{ij}RF \times \Omega_{ij(TN_l)} \times r_{ij} \times t_{ij} \] (17)
\[ P_4 = d_{ij(SNSCNN)} \times \Omega_{ij(TN_l)} \times r_{ij} \times t_{ij} \] (18)
\[ P_5 = d_{ij(CNN)} \times \Omega_{ij(TN_l)} \times r_{ij} \times t_{ij} \] (19)

Models VI, VII, and VIII use RF, SNSCNN, and CNN to extract developmental suitability, respectively. Since SNSCNN and CNN have the extraction of neighborhood features in the learning of features itself, these two experiments are used to explore the impact of discarding the extraction of neighborhood factors in traditional CA on the final simulation accuracy. Experiment 6 serves as a control group experiment to test the effect of using the traditional development suitability algorithm RF on the simulation results without including any neighborhood effect factor, with the following integrated probability calculation model (corresponding to VI, VII, and VIII in the table):

\[ P_6 = d_{ij(RF)} \times r_{ij} \times t_{ij} \] (20)
\[ P_7 = d_{ij(SNSCNN)} \times r_{ij} \times t_{ij} \] (21)
\[ P_8 = d_{ij(CNN)} \times r_{ij} \times t_{ij} \] (22)

3.1.2. Results of the Experimental Scheme

Similar to the training process, the non-urban raster in the base year (2010) is used as a reference to crop the driving factor layer. The cropped data are then fed into the SNSCNN model, which produces a high-dimensional feature vector. This feature vector is subsequently input into a classifier for classification. The resulting classification probability represents the neighborhood probability of the raster and forms a neighborhood constraint layer. For development suitability extraction, the RF model is employed as the development suitability extractor to determine the development suitability of the current raster. This process generates the development suitability layer. By combining land development suitability, neighborhood effects, constraints, and random factors, the total non-urban to urban transition probability layer is obtained. Finally, the total number of non-urban rasters that have converted to urban rasters from 2010 to 2017 serves as a quantitative constraint for dynamic simulation. This constraint is used to generate the final predicted results for the year 2017. The actual situation in 2017 and the experimental prediction results can be observed in Figure 8.
3.1.3. Simulation Accuracy Results

In this paper, we designed a variety of experimental schemes as ablation experiments described in the previous section, and the experimental comparison results are shown in Figure 7, and the experimental accuracy comparison is shown in Table 2.

Table 2. Simulation accuracy comparison.

<table>
<thead>
<tr>
<th>Serial No.</th>
<th>Model Type</th>
<th>Kappa</th>
<th>FoM</th>
<th>OA</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>RF-SNSCNN-CA</td>
<td>0.7683</td>
<td>0.3836</td>
<td>0.9782</td>
</tr>
<tr>
<td>II</td>
<td>RF-CNN-CA</td>
<td>0.7663</td>
<td>0.3798</td>
<td>0.9780</td>
</tr>
<tr>
<td>III</td>
<td>RF-TN-CA</td>
<td>0.7278</td>
<td>0.3128</td>
<td>0.9744</td>
</tr>
<tr>
<td>IV</td>
<td>SNSCNN-TN-CA</td>
<td>0.7272</td>
<td>0.3118</td>
<td>0.9743</td>
</tr>
<tr>
<td>V</td>
<td>CNN-TN-CA</td>
<td>0.7257</td>
<td>0.3094</td>
<td>0.9742</td>
</tr>
<tr>
<td>VI</td>
<td>RF-NN-CA</td>
<td>0.7424</td>
<td>0.3374</td>
<td>0.9758</td>
</tr>
<tr>
<td>VII</td>
<td>SNSCNN-NN-CA</td>
<td>0.7354</td>
<td>0.3255</td>
<td>0.9751</td>
</tr>
<tr>
<td>VIII</td>
<td>CNN-NN-CA</td>
<td>0.7294</td>
<td>0.3155</td>
<td>0.9745</td>
</tr>
</tbody>
</table>

Based on Table 2, the comparison is illustrated in the following aspects: (1) the recommended model integrates the dual advantages of traditional machine learning...
and deep learning models, where SNSCNN improves the learning effect of neighborhood features through the combined effect of channel and spatial attention mechanisms; (2) the machine learning-like urban sprawl CA modeling process, regardless of the approach taken to obtain development suitability, cannot completely replace the learning of the neighborhood part; (3) when deep learning or decision tree learning models are used to extract developmental suitability, the use of traditional neighborhood modeling methods instead has an inhibiting effect on the model simulation results.

3.2. Explanatory Analysis of the Model

3.2.1. Interpretability Analysis of Attention Mechanism

During the extraction of neighborhood features using SNCSNN, the channel attention mechanism was employed to amplify the influential drivers and diminish the drivers with minimal impact. The significance of each driver in the model’s prediction process was extracted and represented as a driver importance distribution layer using ArcGIS software, as depicted in Figure 9. The analysis reveals several observations:

1. Different driving factors exhibit varying degrees of influence on urban land change, as also evident in Figure 10.
2. Spatial variations in the magnitudes and effects of driving factors on urban land change can be observed. For instance, elevation, distance to major roads, distance to major railroads, distance to airports, distance to town centers, nighttime light intensity, and shopping center point density display stronger importance trends at the edges and weaker trends at the center. Conversely, slope, distance to water bodies, distance to highway entrances and exits, distance to railway stations, distance to subway stations, hospital point density, school point density, and attraction point density exhibit weaker trends at the edges and stronger trends at the center.
3. The differences in importance among drivers with similar trends primarily occur at the urban edges. For example, the importance of shopping center point density versus town center point density varies, indicating that urban expansion tends to favor urban grids in edge areas, albeit with different impacts from each driver.

![Figure 9. Spatial heterogeneity of neighborhood characteristics. (a) Elevation influence degree; (b) slope data; (c) hospital density; (d) shopping center density; (e) school site density; (f) attraction density data; (g) distance to airport; (h) distance to town center; (i) distance to bus station; (j) distance to subway station; (k) distance to train station; (l) distance to railroad; (m) distance to road; (n) distance to water. (o) nighttime light intensity.](image-url)
Figure 10. A (800, 800) and B (1200, 1200) two-point neighborhood impact difference map.

The model proposed in this paper incorporates the spatial attention mechanism, which highlights areas with high information and downplays areas with low information within the neighborhood sampling window. This allows for a differential portrayal of the regional impacts of each raster transformation. By extracting the spatial attention weights and visualizing the regional disparities in two study regions (coordinates A: 800, 800 and B: 1200, 1200), the effects were analyzed.

1. The overall influence of the raster around region A is weaker than that of region B.
2. The influence of region B is greater in the northeast and southwest and weakest in the southeast.
3. The influence of the northwest and southeast of area A is greater, and the central part of the east-west line that crosses itself is less influential.

3.2.2. Explainability Analysis of the Overall Driving Factor of the RF Model

Traditional machine learning methods, such as random forest, can provide us with an understanding of the overall importance of driving factors in the urbanization process of the study area. By using RF to extract development suitability, we can extract and visualize the importance coefficients of the driving factors. Figure 11 reveals the following findings:

(1) Nighttime light intensity, representing economic factors, holds the highest importance.
(2) Chongqing, being a “multi-center, cluster” development city, is influenced by multiple centers, and the distance to the town center has a lesser impact on urban development.
(3) Education and healthcare are prominent concerns in traditional Chinese thinking, resulting in school and hospital densities playing a significant role in urban development.

Coupling the attention mechanism with the decision tree RF method retains the advantage of the decision tree method in terms of the measurability of the global influence of each driver and also enhances the expressiveness of the spatial non-smoothness of the drivers of the attention mechanism, thus enhancing the overall interpretability of the coupled model.
3.3. Future Projections of Urban Expansion

3.3.1. Future Projection Results

In future urban projections, the Markov chain is utilized to forecast the additional urban land in 2024 and 2031 based on the land use data from 2010 to 2017. The total area of newly added urban land was 703.58 km² between 2010 and 2017, between 2017 and 2024 it was 633.13 km², and between 2024 and 2031 it was 651.98 km². By employing the RF-SNSCNN-CA model, simulations are conducted using the real land use data of 2010 as the initial state to predict the spatial patterns of urban expansion in 2024 and 2031. The final simulation results are shown in Figure 12. The findings of future urban development trends are as follows:

1. The overall results of the Markov chain prediction indicate a slowdown in the urban development rate, with the urban expansion rate in the period of 2017–2031 being lower than that of 2010–2017.
2. Discrete areas within the central city gradually become connected, forming a more cohesive urban landscape.
3. Peripheral urban areas exhibit a tendency to develop towards the central area.

3.3.2. Analysis of Urban Land Use Landscape Index Changes

The landscape pattern changes of urban land use in the 21 districts of the main city of Chongqing were analyzed by utilizing Fragstats 4.2 software. The real land use data for 2010 and 2017, as well as the projected land use data for 2024 and 2031, were inputted into the software. The changes in six types of landscape factors related to urban land use were compared to illustrate the alterations in the landscape pattern.

We conducted an analysis of the landscape pattern indices for urban sites in the study area from 2010 to 2031 (Table 3). The analysis reveals the following changes:

1. Number of patches (NP) and patch density (PD) increased from 2010 to 2017 and then gradually decreased from 2017 to 2031. This suggests that urban land development from 2010 to 2017 mainly occurred in multiple isolated patches, while from 2017 to 2031, these patches merged into larger contiguous areas.
2. The landscape shape index (LSI) gradually decreased, indicating a trend of more regular and organized development of urban land in the future.
3. The maximum patch index (LPI) increased over time, indicating the growing importance of urban land in the entire study area.
4. The effective particle size (MESH) and aggregation index (AI) showed an increasing trend. This suggests that while patches became more clustered, the overall fragmentation of urban land, in terms of area, also increased due to the development of additional urban areas.
3.3. Future Projections of Urban Expansion

3.3.1. Future Projection Results

In future urban projections, the Markov chain is utilized to forecast the additional urban land in 2024 and 2031 based on the land use data from 2010 to 2017. The total area of newly added urban land was 703.58 km² between 2010 and 2017, between 2017 and 2024 it was 633.13 km², and between 2024 and 2031 it was 651.98 km². By employing the RF-SNSCNN-CA model, simulations are conducted using the real land use data of 2010 as the initial state to predict the spatial patterns of urban expansion in 2024 and 2031. The final simulation results are shown in Figure 12. The findings of future urban development trends are as follows:

(1) The overall results of the Markov chain prediction indicate a slowdown in the urban development rate, with the urban expansion rate in the period of 2017–2031 being lower than that of 2010–2017.

(2) Discrete areas within the central city gradually become connected, forming a more cohesive urban landscape.

(3) Peripheral urban areas exhibit a tendency to develop towards the central area.

Table 3. Urban Landscape Index.

<table>
<thead>
<tr>
<th>Year</th>
<th>NP</th>
<th>PD</th>
<th>LPI</th>
<th>LSI</th>
<th>MESH</th>
<th>AI</th>
</tr>
</thead>
<tbody>
<tr>
<td>2010</td>
<td>5499</td>
<td>0.1912</td>
<td>0.7655</td>
<td>101.5996</td>
<td>179.0912</td>
<td>62.3411</td>
</tr>
<tr>
<td>2017</td>
<td>8420</td>
<td>0.2927</td>
<td>1.2015</td>
<td>109.1113</td>
<td>490.9877</td>
<td>71.254</td>
</tr>
<tr>
<td>2024</td>
<td>6274</td>
<td>0.2181</td>
<td>1.7121</td>
<td>75.3425</td>
<td>1107.279</td>
<td>83.5528</td>
</tr>
<tr>
<td>2031</td>
<td>4560</td>
<td>0.1585</td>
<td>2.2904</td>
<td>57.2767</td>
<td>2349.293</td>
<td>89.1649</td>
</tr>
</tbody>
</table>

In conclusion, the future development of urban areas is characterized by the gradual connection of patches, synergistic development of multiple areas, and a move towards more intensive land use.

4. Discussion

This study conducted two sets of scenarios comprising eight experiments to compare and verify the effects of different approaches for extracting development suitability and neighborhood effects on simulation results. By comparing with other models, it was observed that the RF-SNSCNN-CA-based urban sprawl simulation model can enhance the accuracy of urban sprawl simulation and exhibit a strong capability to model spatial heterogeneity.

The improvement in modeling spatial heterogeneity was achieved by considering the learning of both neighborhood and non-neighborhood features. Furthermore, the study demonstrated that random forest exhibits a relatively low computational cost while providing good performance with diverse interpretations of decision trees. On the other hand, deep learning models, such as neural networks, are computationally expensive and sacrifice the interpretability of features. However, by integrating attention mechanisms with traditional deep learning methods, the model’s interpretability can be enhanced while capturing spatial heterogeneity effectively.

5. Conclusions

Traditional neighborhood factors are inadequate in learning spatial features. To address these problems, this study takes 21 districts in the main city of Chongqing as an example, and improves the urban dynamic model from two aspects of neighborhood feature learning and non-neighborhood feature learning, respectively, and constructs the RF-SNSCNN-CA urban expansion simulation model. On the one hand, starting from
the improvement of the learning effect of neighborhood features, the learning effect of neighborhood features is optimized by introducing the SNSCNN method to improve the effect and capability of spatial heterogeneity modeling; on the other hand, starting from the improvement of the learning effect of non-neighborhood features, the traditional machine learning model RF is combined with the neighborhood features module of deep learning to improve the simulation accuracy while preserving the global interpretability of the random forest model. The main feature of the SNSCNN model is its ability to capture the spatial heterogeneity in the spatial problem by portraying the important change of the driver through channel attention and the differential influence of the neighborhood through spatial attention, so that the interpretability and accuracy of the model are improved, which provides a new way of thinking for studying similar spatio-temporal simulation problems. In terms of the application of the model in this paper, we find that the urban development of the 21 districts of the main city of Chongqing to 2031 in the future is infill agglomeration development, which is consistent with Chinese policies related to the rural revitalization strategy and other related policies. Despite the gradual integration of urban areas, the overall landscape fragmentation of land use increases.

There are still some shortcomings in this study, as shown in the following points: (1) The land use types only involve urban and non-urban land. The model can be used in subsequent studies for multiple land class simulations to try to convert multiple land use types into multiple classification task learning. (2) The use of a deep learning approach can at least partially replace the learning of neighborhood features, which is part of the reason why some of the literature ignores the neighborhood factor when deep learning is used to extract development suitability. As to whether it can completely replace the learning of neighborhood features, further experimental verification is needed. (3) At present, in the study of spatial phenomena, there are more evaluation indices for spatial relevance, such as the Moran index, but less for spatial heterogeneity. The evaluation parameters of spatial heterogeneity will be explored in the subsequent research. (4) Due to the uncertainty of policies, it is easy to change the original urban development pattern, and it is difficult to accurately predict the future urban development trend if only based on the historical trend. Scenario prediction can combine historical trends and relevant policies. (5) In this study, we mainly discuss the extraction of development suitability and neighborhood effects in the CA class urban expansion model. In future studies, we will discuss the restriction constraints to match policy constraints with natural constraints and the stochastic factors to improve the robustness of the model.

Author Contributions: Conceptualization, M.L. and X.L.; resources, software, and writing—original draft preparation, X.L.; writing—review and editing, M.L. and X.L.; Funding acquisition, M.L. and C.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China of funder grant number 42071218 and Natural Science Foundation of Chongqing, China of funder grant number cstc2019jcyj-msxmX0139.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References


12. Wang, Q.; Guan, Q.; Lin, J.; Luo, H.; Tan, Z.; Ma, Y. Simulating Land Use/Land Cover Change in an Arid Region with the Coupling Models. *Ecol. Indic.* **2021**, *122*, 107231. [CrossRef]


34. Huang, Z.; Li, S.; Peng, Y.; Gao, F. Spatial Non-Stationarity of Influencing Factors of China’s County Economic Development Base on a Multiscale Geographically Weighted Regression Model. *ISPRS Int. J. Geo-Inf.* 2023, 12, 109. [CrossRef]


41. Li, D.; Liu, J.; Zhao, Y. Prediction of Multi-Site PM2.5 Concentrations in Beijing Using CNN-Bi LSTM with CBAM. *Atmosphere* 2022, 13, 1719. [CrossRef]

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