J. COMBIN. MATH. COMBIN. COMPUT. 127a (2025) 3395--3411

Journal of Combinatorial Mathematics and Combinatorial Computing
www.combinatorialpress.com/jcmcc

Optimising Territorial Spatial Planning Using Image Recognition Technology and Monitoring the Direction of Coordinated Development of Urbanisation

Lin Ge[1,✉]

[1] School of Public Administration, China University of Mining and Technology, Xuzhou, Jiangsu, 221116, China

ABSTRACT
With the accelerating process of urbanization development, it is urgent to optimize the national land spatial planning to promote the coordinated development of urbanization. Based on the image recognition technology, this study uses the kernel density gradient algorithm to segment the image samples of the national spatial layout and the GWO-SVM classification model to classify the land use types of the national spatial layout, and finally combines the Markov-FLUS model to predict the future planning of the existing national spatial layout. The research analysis found that the segmentation and classification accuracy of the kernel density gradient algorithm and the GWO-SVM classification model for the homeland spatial layout samples both reached more than 90%. The classification accuracy using the GWO-SVM classification model is improved to a greater extent than that of SVM, GA-SVM, etc. The Markov-FLUS model also maintains an accuracy of more than 80% for the prediction of future territorial spatial planning. In terms of land use types, the Markov-FLUS model shows that the proportion of residential land and industrial land will decrease after 10 years compared with 5 years, while the proportion of public facilities land will increase by about 8% after 10 years compared with 5 years. The optimization of national spatial layout is of great significance to the development of urbanization in China, and the research in this paper will promote the development of national spatial layout planning in a more reasonable direction.

Keywords: National land spatial planning, image recognition, GWO-SVM model, Markov-FLUS model

1. Introduction

Territorial spatial planning is a strategic guarantee to promote the "unity of multiple planning" and strengthen the management of natural resources, and it is also a key issue that needs to be solved in the new era of development, which is of great significance for improving China's spatial governance

✉ Corresponding author.
E-mail address: m13965881056@163.com (L. Ge).
Received 20 February 2024; Revised 15 April 2024; Accepted 20 November 2024; Published Online 15 April 2025.
DOI: 10.61091/jcmcc127a-193

system and coordinating the relationship between high-quality economic development and green development [1-3].

Territorial spatial planning is a national will-orientation to optimize ecological environmental protection, guarantee food security, and promote the intensive use of territorial resources, and it is a key initiative to promote the construction of ecological civilization [4-6]. Since the reform and opening up, the level of urbanization and industrialization in China has been increasing, and as a country with scarce natural resource endowment and rapid economic development, it faces a series of sustainable development problems while making world-renowned achievements [7-9]. With the deepening of the overall understanding of different types of natural resources, the intrinsic connection between the allocation of urban and rural resource elements and the strategic significance of integrated development, the re-cognition of the spatial conflict and coordinated development mechanism of production, life and ecology, and the re-balancing of the relationship between comprehensive development and utilization of land and the recuperation of resources and ecological restoration, it is all required to appropriately expand the scope of the management of the country's natural resources, and to establish a holistic, spatial, and inclusive national land resource management system [10-13]. Therefore, the State has established the Ministry of Natural Resources according to the needs of the new era, giving it the important responsibility of "establishing a spatial planning system and supervising its implementation", pointing out that it is necessary to strengthen the control of spatial planning in the national territory and improve the efficiency of spatial control [14-15].

Urbanization refers to the process of population influx to cities. This concept can be divided into two aspects, on the one hand, the number of cities and towns will increase significantly under this trend. The other aspect is that the number of people in towns and cities will continue to increase under this trend [16-17]. And with the accelerated development process of urbanization, the proportion of the primary sector will decrease significantly, and the proportion of the secondary and tertiary sectors will increase significantly, because the complete development system of the country is composed of both urban and rural areas, and an increase in the number of urban population means a decrease in the number of rural population.

The research in this paper starts from the image of the national land spatial layout, using the kernel density gradient algorithm to segment the image, and then the segmented image is successfully divided into different land use types through the GWO-SVM model, and finally, in order to further analyze how to achieve the optimization of the planning of the national land spatial layout, this paper adopts the Markov-FLUS coupled model to simulate the existing national land spatial layout, and get the the prediction results of the future spatial layout of the national territory. Finally, we analyze and draw conclusions according to the actual research needs.

2. Method

2.1. Image segmentation

2.1.1. Kernel density estimation and its properties.First, the concept of kernel function is introduced. The term "kernel" is often used in statistics. Here, kernel is used to refer to any smooth function K such that it satisfies $K \geq 0$ as well:

$$\int_{K}(x)dx=1,\int_{XK}(x)dx=0,\sigma_{K}^{2}\equiv\int_{x^{2}K}K(x)dx>0 \tag{1}$$

Here are some commonly used cores:

$$K(x)=\frac{1}{2}I(x) \tag{2}$$

$$K(x)=\frac{1}{\sqrt{2\pi}}e \tag{3}$$

$$K(x)=\frac{3}{4}\left(1-x^2\right)I(x) \tag{4}$$

$$K(x)=\frac{70}{81}\left(1-|x|^3\right)^3 I(x) \tag{5}$$

Here:

$$I(x)=\begin{cases}1, & |x|<1\\ 0, & |x|>1\end{cases} \tag{6}$$

In kernel density gradient segmentation algorithms, two types of kernel functions are often used and they are, unit uniform kernel functions:

$$K_U(x)=\begin{cases}1 & \|x\|<1\\ 0 & \|x\|\geq 1\end{cases} \tag{7}$$

Unit Gaussian kernel function:

$$K_N(x)=\exp\left(-\frac{1}{2}\|x\|^2\right) \tag{8}$$

Up to this point, a statistically meaningful definition of kernel density estimation can be given, such that $X_1, X_2, \cdots X_n$ denotes the observations, which come from a sample of f. According to the definition of a kernel function as any smooth function K, there is the following definition

Definition 2.1: Given a kernel K and a positive number h called bandwidth, the kernel density estimator is defined as:

$$\hat{f}_n(x)=\frac{1}{n}\sum_{i=1}^{n}\frac{1}{h}K\left(\frac{x-X_i}{h}\right) \tag{9}$$

The kernel density estimation effectively extends a smooth packet by assigning a weight of size $1/n$ on each data point X_i.

Bandwidth h controls the degree of smoothing. When $h\to 0$, the kernel density estimate $\hat{f}_n(x)$ contains many spikes, one for each data point. When $h\to 0$, the height of the spikes tends to infinity. When $h\to\infty$, $\hat{f}_n(x)$ tends to a uniformly distributed density function. Thus, the kernel density function is sensitive to the choice of bandwidth. Small bandwidths give very coarse estimates, while large bandwidths give smoother estimates.

2.1.2. Nuclear density gradient estimation.Kernel density estimation, also known as Parzen window estimation in pattern classification, is performed to analyze the probability density of the data by assuming that the feature dimension is d and the set of sampling points for the probability density function $f(x)$ is $x=\{x_1,\cdots,x_N\}\subset R^d$. The probability density estimate obtained at point x using the kernel function $K(x)$ and a window of size h is then:

$$\hat{f}(x)=\frac{1}{nh^4}\sum_{i=1}^{n}K\left(\frac{x-x_i}{h}\right) \tag{10}$$

Assuming that each sample point in this all sample point x_i has a different importance, a weighting factor is introduced for each sample. This expands the basic kernel density estimate as:

$$\hat{f}(x)=\frac{\sum_{i=1}^{n}K\left(\frac{x_i-x}{h}\right)s(x_i)}{h^d\sum_{i=1}^{n}s(x_i)} \tag{11}$$

where: $s(x_i)\geq 0$ is the value of the weight function of the sample point x_i, and $K(x)$ is a kernel density function which satisfies $\int k(x)dx=1$. Assume that $k(x)$ is used to represent the profile function of the original kernel density function $K(x)$, and make the

$$K(x)=k\left(\|x\|^2\right) \tag{12}$$

If the negative derivative of the profile function $k(x)$ is denoted by $g(x)$, i.e. $g(x)=-k(x)$, its corresponding kernel function is denoted by $G(x)$ as:

$$G(x)=g\left(\|x\|^2\right) \tag{13}$$

Thus the gradient $\nabla f(x)$ of the kernel density function of Eq. (11) is:

$$\nabla f(x)=\nabla\hat{f}(x)=\frac{2\sum_{i=1}^{n}(x-x_i)k\left(\left\|\frac{x_i-x}{h}\right\|^2\right)s(x_i)}{h^{d+2}\sum_{i=1}^{n}s(x_i)} \tag{14}$$

Computation of Eq. (14) from the negative derivative function $g(x)$ and the new kernel functions $G(x)$, $g(x)=-k'(x)$, $G(x)=g\left(\|x\|^2\right)$ is done:

$$\begin{aligned}\nabla f(x)&=\frac{2\sum_{i=1}^{n}(x_i-x)G\left(\left\|\frac{x_i-x}{h}\right\|^2\right)s(x_i)}{h^{d+2}\sum_{i=1}^{n}s(x_i)}\\&=\frac{2}{h^2}\left[\frac{\sum_{i=1}^{n}G\left(\frac{x_i-x}{h}\right)s(x_i)}{h^d\sum_{i=1}^{n}s(x_i)}\right]\left[\frac{\sum_{i=1}^{n}(x_i-x)G\left(\left\|\frac{x_i-x}{h}\right\|^2\right)s(x_i)}{\sum_{i=1}^{n}G\left(\frac{x_i-x}{h}\right)s(x_i)}\right]\end{aligned} \tag{15}$$

The second part of the product term of Eq. (15) is the kernel density gradient mean drift vector, and the first part of the product term is the estimate of the probability density function $f(x)$ expressed in terms of the kernel function $G(x)$, which is simplified and denoted as $\hat{f}_G(x)$. In order to differentiate between the estimate of the probability density function $f(x)$, expressed by the function $K(x)$, $\hat{f}(x)$, then Eq. (11) is expressed as $\hat{f}_K(x)$, so that (15) is changed into:

$$\nabla f(x)=\nabla\hat{f}_K(x)=\frac{2}{h^2}\hat{f}_G(x)M_h(x) \tag{16}$$

By a simple transformation, it is possible to express from Eq. (16) the local density maxima about the place searched by the mean drift vector of the kernel density gradient, essentially the locally flat location where the density is maximal and the modulus of the gradient is minimal, which is expressed as the following equation:

$$M_h(x)=\frac{1}{2}h^2\frac{\nabla\hat{f}_K(x)}{\hat{f}_G(x)} \tag{17}$$

The computational expression for the mean drift vector of the kernel density is obtained by separating the vector x from the second product term in Eq. (15), whereupon the specific mean drift vector is:

$$M_h(x)=\frac{\sum_{i=1}^{n}G\left(\frac{x_i-x}{h}\right)s(x_i)x_i}{\sum_{i=1}^{n}G\left(\frac{x_i-x}{h}\right)s(x_i)}-x \tag{18}$$

Using $m_h(x)$ to denote the first term of the kernel density gradient mean drift vector, we then have

$$m_s(x)=\frac{\sum_{i=1}^{n}G\left(\frac{x_i-x}{h}\right)s(x_i)x_i}{\sum_{i=1}^{n}G\left(\frac{x_i-x}{h}\right)s(x_i)} \tag{19}$$

According to the basic principles of the kernel density gradient mean drift algorithm, given a sample set of kernel density functions $G(x)$, initial vector points x, and cyclic control errors ξ, the kernel density gradient mean drift algorithm can in turn be expressed in the following three steps:
1) Calculate the sample weighted mean $m_h(x)$ of equation (19)
2). Assign the sample weighted mean $m_h(x)$ to the last initial mean point x.
3) Then determine the end condition of the loop $\left\|m_h(x)-x\right\|<\xi$, if the above end condition is not satisfied, then continue to perform steps 1-2.

According to Eq. (18) can be derived from the kernel density of the mean drift vector of the loop iteration formula $m_h(x)=x+M_h(x)$, there is a concept here is the mean drift vector of the local optimal search step, which refers to the mean drift vector of the mode of the size of the step is obviously dependent on the size of the sample distribution of the density of the size of the density of the local area, density of large areas, according to Eq. (17) it can be seen that the mode of the mean drift vector becomes smaller, that is, the step size On the contrary, in the place where the density is small, the step size becomes larger. In particular, it should be noted that under certain conditions, the mean drift algorithm will converge to the maximum point near the point.

A digital image that is commonly seen can actually be expressed as a d-dimensional feature domain space based on a two-dimensional airspace plane, where the dimension $d=1$ in the feature domain is defined as a grayscale map, and $d=3$ is defined as a color space, and so on when $d>3$, it is represented as a multispectral digital image map, so that the segmentation process is based on the space domain and the feature domain, the total dimension of this space is $d+2$, and a vector $x=(x^x,x^x)$ is used to express all the information features, then x^x represents the airspace feature vector, x' represents the eigendomain vector.
where the kernel function uses the expression $K_{h,h}$ of the joint probability density to estimate the distribution of vector x, and the form of the $K_{h,h}$ joint product kernel is as follows:

$$K_{h,h,}=\frac{C}{h_x^2h_r^d}k\left(\left\|\frac{x^2}{h_x}\right\|^2\right)k\left(\left\|\frac{x'}{h_r}\right\|^2\right) \tag{20}$$

where h_s, h_r is the window bandwidth parameter of the kernel density function and the constant c is a normalization constant.

Suppose the original and filtered images are denoted by x_i and z_i, $i=1,\cdots,n$. For each pixel point, the specific steps of image filtering with kernel density gradient mean drift algorithm are as follows:
1) Initialize $j=1$ and make $y_{i,1}=x_i$
2). Apply the kernel density gradient mean drift algorithm to compute $y_{i,j+1}$ until convergence. Record the converged value as $y_{i,c}$

3) Assign the value $z_i = \left(x_i', y_{i+}^r\right)$

Image segmentation based on kernel density gradient mean drift algorithm is consistent with the principle of image filtering, only need to converge to the same point of the data members into a class, and then assign the label of this class to these data members, in image segmentation sometimes need to remove the class that contains too few pixel points. Where the image segmentation process consists of two steps, i.e. image kernel density gradient filtering process and image extraction object process.

First is the image kernel density gradient filtering process:

1) Step 1: initialize $k=1$ and end condition ξ, initialize window center position $y_k = x_j$ with current image element point x_j;
2) Second step: calculate the new position y_{k+1} on the convergence path according to the kernel density gradient mean drift vector equation to get the vector value $M_k = y_{k+1} - y_k$;
3) Step 3: $k = k+1$, until $\|M_h\| < \xi$ then stop and note the convergence point as y_{jx};
4) Step 4: assign a new value $z_j = \left(x_j^*, y_{j,c}^r\right)$ to the j th image element point.

Next is the image extraction object process:

1) Step 1: Perform kernel density gradient filtering algorithm for each image element point to get $z_j = \left(x_j^*, y_{je}^r\right)$;
2) Step 2: After clustering each z_j, a set $\{C_p\}_{p=1,\cdots,N}$ containing N objects is obtained, and the rule of clustering and merging is that the image elements with a distance less than h_s in the null domain and less than h_r in the color space are the same object region;
3) Step 3: Based on the results of the objects extracted by clustering, assign each image element of the original image the label of the object to which it belongs.

2.2. Picture classification

2.2.1. Principles of the GWO algorithm.The GWO algorithm is inspired by modeling the leadership hierarchy and hunting mechanism of a gray wolf population. The algorithm has a simple structure, few control parameters, high robustness, and shows excellent performance in parameter optimization.

1) Leadership level.The leadership hierarchy of the gray wolf is defined as α, β, δ and ω wolves from high to low, in which α wolf is the alpha wolf, mainly responsible for the decision-making task, which is equivalent to the optimal solution of the function optimization process. β wolf is the deputy chief, responsible for assisting α wolf's management, and at the same time providing information feedback in the wolf pack, which is the sub-optimal solution in the function optimization process. The δ wolf carries out the orders of the α and β wolves, and is responsible for scouting and rounding up the prey, which is the sub-optimal solution in the optimization process.

2) Surrounding the prey.The gray wolf surrounds the prey in the roundup process, and this roundup behavior is represented by a mathematical model:

$$\begin{aligned} \vec{D} &= \left|\vec{C}\cdot\vec{X}_P(t) - \vec{X}(t)\right| \\ \vec{X}(t+1) &= \vec{X}_p(t) - \vec{A}\cdot\vec{D} \end{aligned} \tag{21}$$

where t represents the current number of iterations, $\vec{X}_p$ is the prey position vector, $\vec{X}$ is the gray wolf position vector, $\vec{A}$, $\vec{C}$ are the coefficient vectors $\vec{A}$, $\vec{C}$ are calculated as follows:

$$\vec{A} = 2\vec{a}\cdot(\vec{r}_1 - 1) \tag{22}$$

$$\vec{C} = 2\cdot\vec{r}_2 \tag{23}$$

where $\vec{a}$ is linearly decreasing, $\vec{a}\in[0,2]$; $\vec{r}_1$, $\vec{r}_2$ are random variables, and $\vec{r}_1$, $\vec{r}_2 \in [0,1]$.

3) Attacking prey.When the prey is surrounded, α, β, δ wolves have a strong ability to recognize the location of the prey, so they lead the pack in hunting during each iteration, and in each iteration,

the three best gray wolves (α,β,δ) in the population before the iteration are retained, and then the locations of the other search agents (ω) are updated based on their location information.

The update formula for the gray wolves approaching to the prey is as follows:

$$\begin{aligned}\vec{D}_\alpha &= \left|\vec{C}_1 \cdot \vec{X}_\alpha - \vec{X}_{(1)}\right| \\ \vec{D}_\beta &= \left|\vec{C}_2 \cdot \vec{X}_\beta - \vec{X}_{(1)}\right| \\ \vec{D}_\delta &= \left|\vec{C}_3 \cdot \vec{X}_\delta - \vec{X}_{(1)}\right|\end{aligned} \tag{24}$$

where $\vec{X}_\alpha$, $\vec{X}_\beta$, and $\vec{X}_\delta$ are the positions of α, β, and δ wolves relative to their prey, and $\vec{C}_1$, $\vec{C}_2$, and $\vec{C}_3$ are random perturbations of α, β, and δ.

Update the offspring gray wolf position equation:

$$\begin{cases}\vec{X}_1 = \vec{X}_\alpha - \vec{A}_1 \cdot \vec{D}_\alpha \\ \vec{X}_2 = \vec{X}_\beta - \vec{A}_2 \cdot \vec{D}_\beta \\ \vec{X}_3 = \vec{X}_s - \vec{A}_3 \cdot \vec{D}_\delta\end{cases} \tag{25}$$

Final position of the zygotic gray wolf:

$$\vec{X}(t+1) = \left(\vec{X}_1 + \vec{X}_2 + \vec{X}_3\right)/3 \tag{26}$$

In the GWO algorithm, there are two most important parameters $\vec{A}$, $\vec{C}$. For dispersion modeling, when $|\vec{A}| \geq 1$, it makes the gray wolf move away from the prey and GWO performs a global search. When $|\vec{A}| < 1$, the gray wolf attacks the prey and the algorithm performs a local search. Where $\vec{C} \in [0,2]$ is a random value that provides random perturbations to the prey and wolves to prevent falling into a local optimal solution and improve the global search capability.

2.2.2. GWO-SVM classification models.The GWO-SVM classification model is a structure with support vector machine as a framework, which obtains the position of the optimal particles through GWO and finds the optimal parameter combination of SVM to build the optimal classifier, and then finds the optimal classification hyperplane. The model building process is as follows:

1) Initialize the number of wolves, the number of iterations, and set the range of penalty factor C and kernel function radius g.

2) SVM is trained and tested according to the initial parameters C, g with the goal of minimizing the error rate.

3) GWO is optimized with C, g as prey and outputs the global optimum of GWO when the maximum number of selected generations is reached. The optimal parameters C, g are used to build the recognition model.

The specific flow of GWO optimized SVM classification model is shown in Fig. 1:

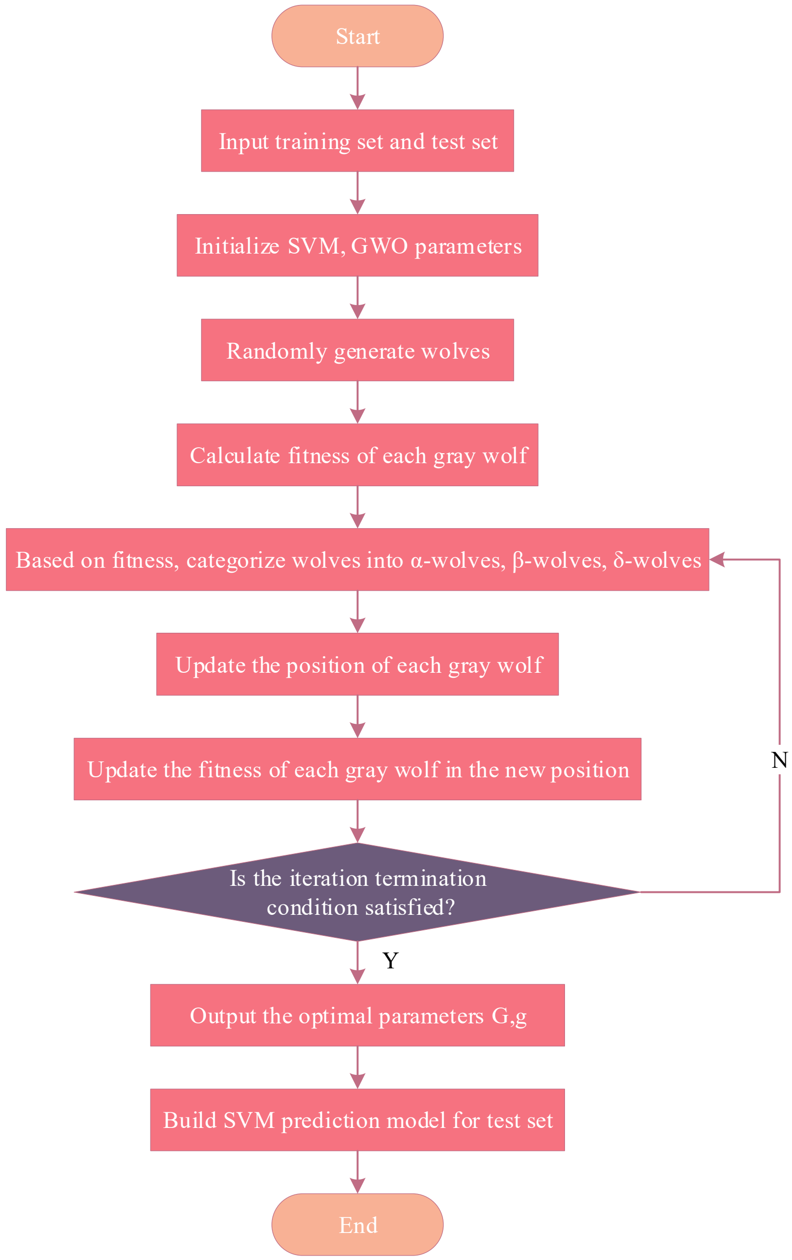


Fig. 1. Gwo optimized sVM classification model flow

2.3. Optimized prediction of land space based on image recognition technology

2.3.1. Overview of Coupled Markov-FLUS Model Construction and Principles.In this paper, the future spatial layout of the national territory is simulated by using the Markov-FLUS coupled model. The Markov model can only predict the quantitative changes of various types of land space and does not have the ability to simulate the changes of spatial layout, while the FLUS model, on the contrary, has a strong spatial computational ability and can deal with the complex spatial relationships among the metameres, but is only limited to dealing with the interactions among the metameres. The coupling of these two models brings together the advantages of Markov model in the prediction of the spatial quantity of the national territory and the characteristics of FLUS model in the simulation of spatial pattern changes, and realizes the goal of simulating the spatial layout of the national territory, and the coupling and simulation process is shown in Fig. 2.

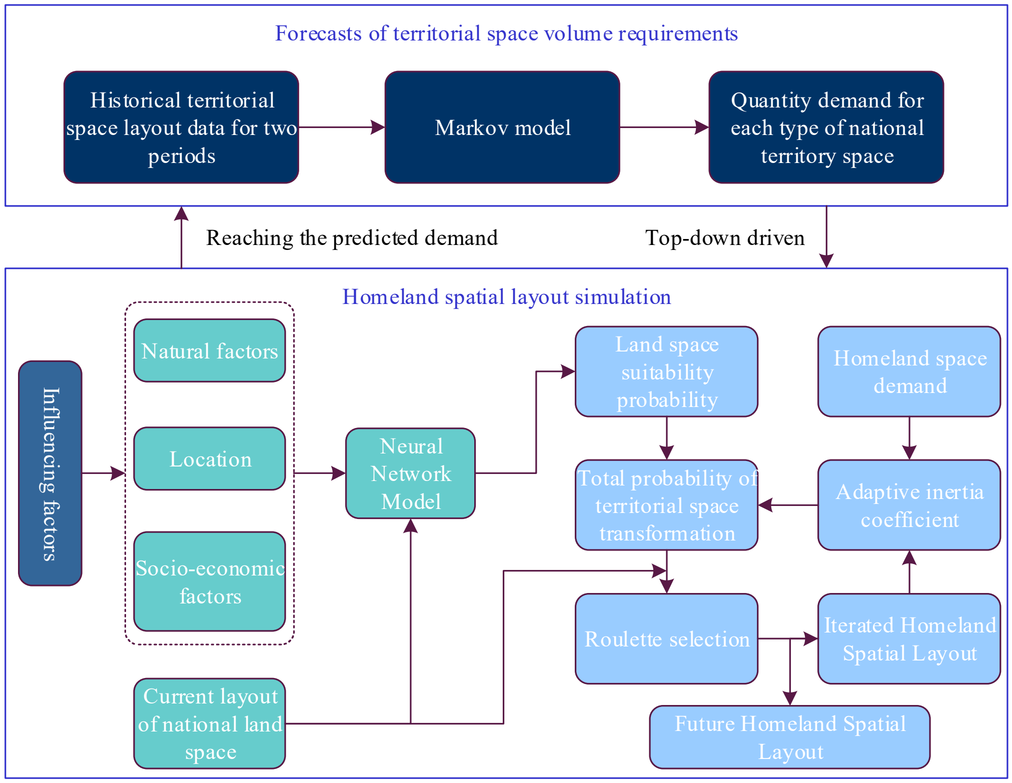


Fig. 2. Simulation process diagram of Markov-FLUS model

2.3.2. Principles of Markov modeling.Markov is a prediction method based on probabilistic transfer describing the process of change of a certain system from one state to another, which has two distinctive features, namely stability and no posteriority. The evolution process of land space also has stability and no posteriority, and these two characteristics are of great significance to the research on the transformation between land space types and the dynamic simulation of land space changes, therefore, Markov can be used to predict the change trend of various types of land space quantities.

The formula for predicting the amount of homeland space demand based on the Markov process is as follows:

$$S_{t+1} = S_t \times P_{ij} \tag{27}$$

where S_{t+1} denotes the state of the homeland space type at $t+1$ the future moment, S_t denotes the state of the homeland space type at the current t moment, and P_{ij} is the probability that the homeland space type i will be converted to the homeland space type j, which is expressed as follows:

The matrix satisfies the following two conditions:

$$\begin{cases} 0 \le P_{ij} \le 1 \quad (i, j = 1, 2, \cdots n) \\ \sum_{j=1}^{n} P_{ij} = 1 \quad (i = 1, 2, 3, \cdots n) \end{cases} \tag{28}$$

2.3.3. FLUS modeling principles.The FLUS model is mainly composed of a neural network-based suitability probability calculation module and a metacellular automaton module based on adaptive inertia mechanism. Firstly, the Phase I territorial space layout data with the selected various types of spatial driving factors are imported into the neural network calculation module, and the suitability probability of various types of territorial space is obtained by setting relevant parameters. Second, the total probability of transformation of the national space is obtained based on the set neighborhood influence factor, adaptive inertia coefficient and conversion cost, and combined with the suitability probability obtained in the previous step. Finally, based on the total probability of conversion, restriction of conversion area and other settings, the adaptive inertia competition mechanism of roulette choice is adopted to realize the conversion between each homeland space type and simulate the homeland space layout. The roulette selection competition mechanism makes up for the shortcomings of the previous model in simulating the competition relationship between the land space types, and the simulation accuracy is high.

1) Neural network-based suitability probability calculation. Artificial neural network is a kind of intelligent algorithm that simulates the neuron structure of human brain, and it has a strong ability to deal with nonlinear function relationship. Different from the traditional mathematical operation model, ANN has strong self-organization, self-adaptation and self-learning capabilities, which can maximally solve the problem of strong subjectivity in setting the weights of the spatial variables of previous models. In recent years, with the continuous development of LUCC, ANN is often used in land use change simulation.

The ANN in the FLUS model is a three-layer BP neural network, including input layer, hidden layer and output layer. Among them, the neurons in the input layer correspond to each driving factor, the neurons in the output layer correspond to each type of land space, and the hidden layer is determined based on the actual situation of the study area, the research results of related literature and the experience of experts. According to the one-to-one correspondence, the connection between various types of homeland space and the selected driving factors is established, and the suitability probability of various types of homeland space in each image element is obtained through training, and its expression is:

$$p(p,k,t)=\sum_{j} w_{j,k}\times sigmoid\left(net_j\left(p,t\right)\right) \tag{29}$$

$$sigmoid\left(net_j\left(p,t\right)\right)=\frac{1}{1+e^{-net_j(p,t)}} \tag{30}$$

$$net_j\left(p,t\right)=\sum_{i} w_{i,j}\times x_i\left(p,t\right) \tag{31}$$

Where, $p(p,k,t)$ denotes the suitability probability of the homeland spatial type k on image element p and time t, $w_{j,k}$ and sigmoid are the weights and correlation functions of the hidden layer and the output layer, respectively, $w_{i,j}$ is the weight between the input layer and the hidden layer, $net_j(p,t)$ is the signal received by the neuron j in the hidden layer on image element p and time t, and $x_i(p,t)$ is the neuron's input value on image element p and time t.

From the above, it can be seen that the suitability probability of each type of land space is not independent of each other, but is interconnected, and the growth of the suitability probability of a certain type of land space will inevitably make the suitability probability of another type of land space decrease. While traditional mathematical models (logistic regression, etc.) tend to calculate the relationship between various types of land space and driving factors individually, compared with them, ANN can better simulate the competitive relationship between various types of land space, and is more suitable for solving complex nonlinear problems, and has certain methodological advantages.

2) Metacellular automata based on adaptive inertia mechanism. The type of homeland space for future conversion of a certain image element is not only determined by the suitability probability of each type of homeland space, but also affected by the competition and interaction between each type of homeland space 16%. Therefore, in this paper, we combine the suitability probability with the neighborhood influence, adaptive inertia coefficient and conversion cost to calculate the total probability of conversion of each image element to each type of homeland space.

①Neighborhood impact. Neighborhood influence reflects the role between each homeland space type within the neighborhood, and the expression is:

$$\Omega_{p,k}^{t}=\frac{\sum_{N\times N} con\left(c_{p}^{t-1}=k\right)}{N\times N-1}\times w_{k} \tag{32}$$

where $\Omega_{p,k}^{t}$ is the neighborhood influence factor, $\sum_{N\times N} con\left(c_{p}^{t-1}=k\right)$ is the number of image elements of the homeland spatial type k after the end of the previous iteration $(t-1)$ in the $N\times N$ window, and w_k is the weight of the neighborhood of each homeland spatial type. It is generally obtained after several tests according to the relevant expertise and experience of experts.

② Adaptive inertia coefficient. The adaptive inertia coefficient of each homeland space type is based on the difference between the actual number of image elements and the number of image elements required in the future, and is adjusted in the iteration to make the simulated number of each homeland space type closer to the target value. The adaptive inertia coefficient of homeland space type k at moment t is:

$$Inertia_{k}^{t}\begin{cases} Inertia_{k}^{t-1} & if\left|D_{k}^{t-1}\right|\le\left|D_{k}^{t-2}\right| \\ Inertiaa_{k}^{t-1}\dfrac{D_{k}^{t-2}}{D_{k}^{t-1}} & if D_{k}^{t-1}<D_{k}^{t-2}<0 \\ InertiaD_{k}^{t-1}\dfrac{D_{k}^{t-1}}{D_{k}^{t-2}} & if\, 0<D_{k}^{t-2}<0 \end{cases} \tag{33}$$

Where, D_k^{t-1}, D_k^{t-2} are the difference between the actual number of image elements and the number of demanded elements of the homeland space type k at $t-1$ and $t-2$ moments, respectively.

③Conversion cost. The conversion cost is defined by the transfer matrix, which refers to the degree of difficulty in converting each homeland space type to each other. The conversion cost matrix is a binary matrix, with 0 indicating that no conversion is allowed to occur and 1 indicating that conversion is allowed to occur.

④ Calculation of total probability of conversion of homeland space. Based on the calculation of the above steps, the total probability of homeland spatial transformation is calculated for each image element according to Equation 5.7, and the different homeland spatial types are assigned to the raster by iterating through the metacellular automata. The formula is:

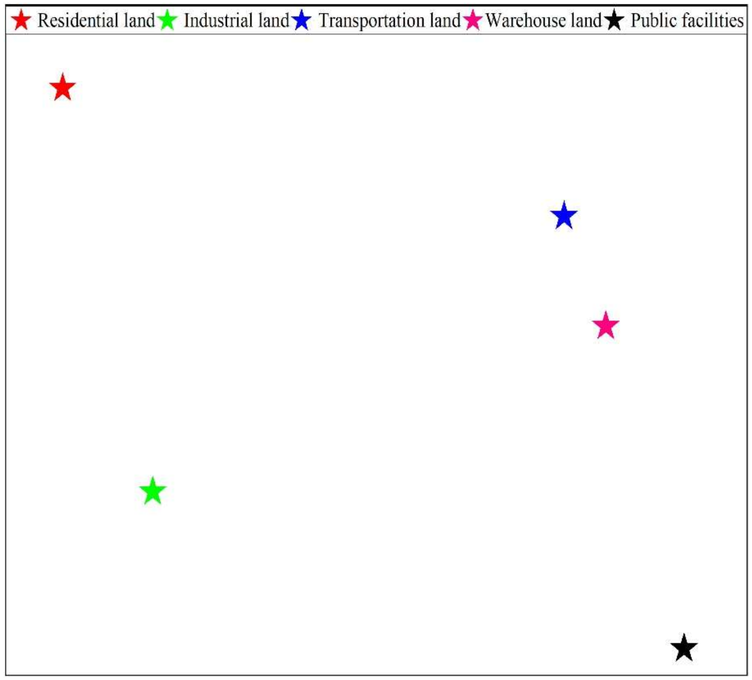
$$TP_{p,k}^{t}=p\left(p,k,t\right)\times\Omega_{p,k}^{t}\times\ Inertia\ _{k}^{t}\times\left(1-sc_{c\to k}\right) \tag{33}$$

where $TP_{p,k}^{t}$ is the total probability that an image element p is transformed in time t into a homeland spatial type k and $sc_{c\to k}$ is the homeland spatial type conversion cost.

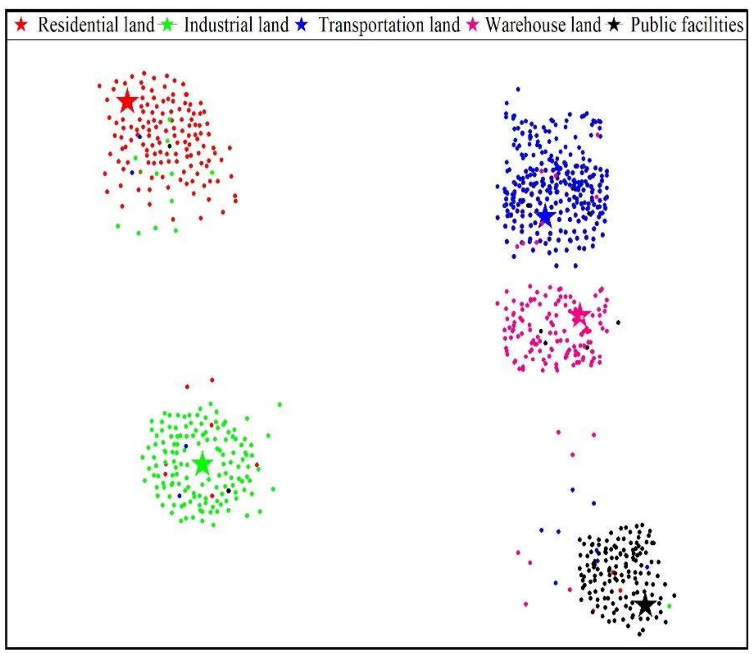
3. Results and Discussion

3.1. Analysis of image segmentation results under kernel density gradient method

This study further illustrates the effectiveness of the image segmentation algorithm proposed in this paper through visualization experiments. Specifically, a 2way-1shot image segmentation task is performed on the miniImagenet dataset and the different stages of the segmentation process are visualized using the algorithm proposed in this paper. The visualization is shown in Fig. 3. In Fig. 3, different colors represent the different categories segmented. Specifically, the five stars in Fig. 3(a) represent the five types of results obtained after the segmentation of all samples, and the same-colored dots surrounding the stars in Fig. 3(b), 150 for each category, represent the results of the algorithm's automatic classification of all the samples after the image segmentation, whereas the different-colored dots surrounding the stars represent the misclassification results in the classification process. From the figure, it can be seen that the image segmentation work has segmented five urban planning land types: urban residential land, industrial land, transportation land, storage land and public facility land, and the accuracy of the algorithm of this paper for the segmentation and classification of these five types of land can reach more than 90%.



(a) Image segmentation result



(b) Classification results after segmentation

Fig. 3. Results of image segmentation and classification

3.2. GWO-SVM model classification accuracy study

In this paper, the study on the classification accuracy of GWO-SVM is based on the image segmentation work of the samples, after the image segmentation is completed and the feature extraction is performed, the above features are inputted into the following classifier models: PSO-SVM, GS-SVM, GA-SVM, SVM, and so on, and the results of the classification and recognition of the classification and recognition of the classification model are compared and analyzed with those of the GWO-SVM classification model. Table 1 shows the classification recognition results of SVM optimized with different swarm intelligence algorithms and the optimal parameters C, g achieved when optimizing the SVM. As can be seen from the table, from the above table, all SVM classifiers optimized with swarm intelligence optimization algorithms classify the samples better than a single SVM classifier. The GWO-SVM classification model designed in this paper comes out on top of the classification performance of all SVM classifiers with 95.93% classification accuracy.

Table 1. Classification results and optimal parameters

Classification Method	Classification Accuracy(%)	Parameter C	Parameter g
SVM	87.63	2.03	1.26
PSO-SVM	91.49	6.02	0.31
GS-SVM	91.57	5.67	0.38
GA-SVM	90.28	8.43	0.88
GWO-SVM	95.93	27	0.13

3.3. Spatial optimization analysis of the national territory under the Markov-FLUS model

The land space layout samples obtained and completed the segmentation and classification tasks, as well as the processed and set up data and parameters such as future land space prediction demand, suitability probability atlas, conversion cost matrix, neighborhood factor, etc. were imported into the Markov-FLUS model algorithm designed in this paper, and then clicked Accept to start running the GeoSOS-FLUS software. After running, the simulation results of the future spatial layout of the country are generated as shown in Fig. 4 and Fig. 5. From Fig. 4, it can be seen that the Markov-FLUS model of this paper achieves a prediction accuracy of more than 80% for the five land use types, among which the prediction accuracy of residential land is 84%, the prediction accuracy of industrial land is 86%, the prediction accuracy of transportation land is 91%, the prediction accuracy of warehousing land is 92%, and the prediction accuracy of public facilities land is 93%. Among all the predictions, 5% of industrial land was incorrectly predicted as residential land and 8% of warehousing land was incorrectly predicted as industrial land. The simulation results in Figure 5 show that after 5 years, the share of all site types will increase, but after 10 years, the share of residential land use will decrease by about 3%, the share of industrial land use will decrease by about 2%, and the share of public facility land use will increase by about 8% compared to 5 years later. The reason for this projection may be related to the aging population, the digital economy and the public's pursuit of a better life.

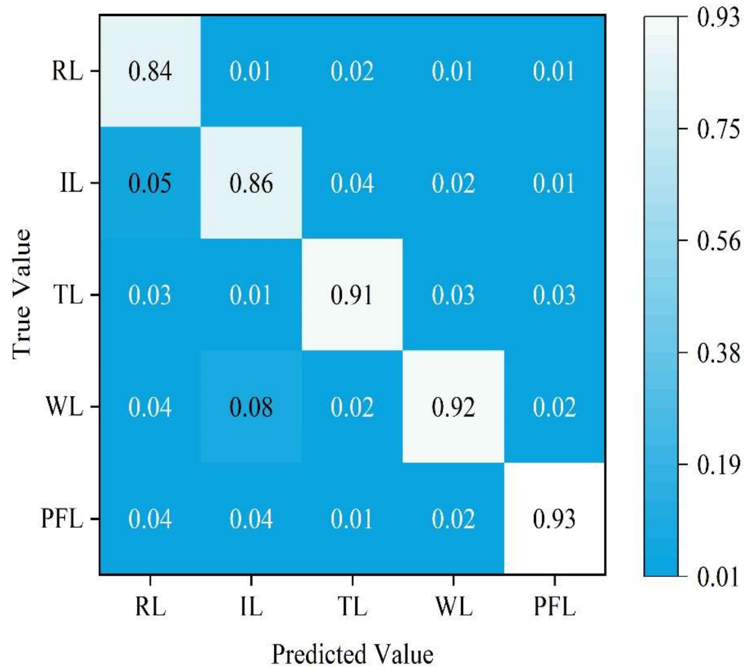


Fig. 4. Prediction accuracy of the mode

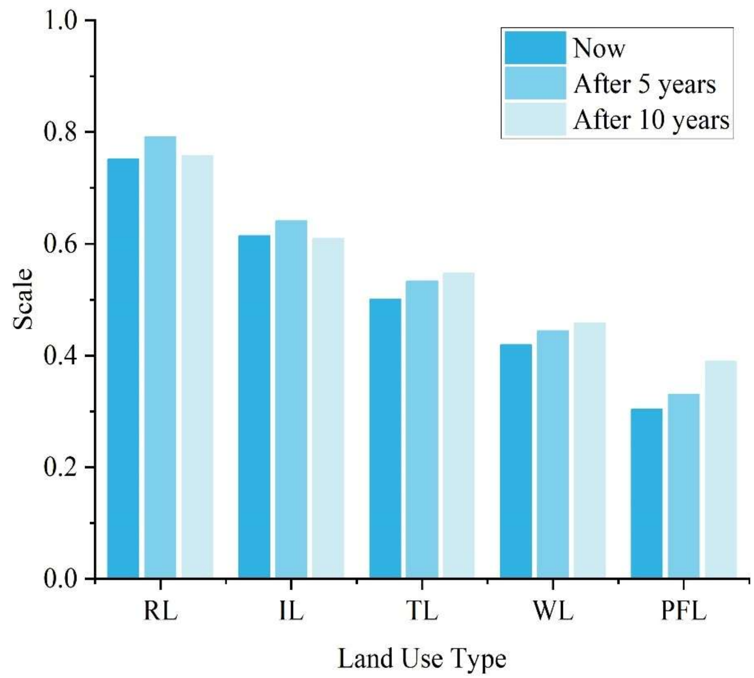


Fig. 5. Prediction of changes in the proportion of different land use types

4. Conclusion

With the research objectives of studying the optimization of national land spatial layout planning and promoting the coordinated development of urbanization, this paper combines the kernel density gradient algorithm, the GWO-SVM classification model, and the Markov-FLUS model and other technical means to carry out layer-by-layer research and analysis. Through the research, it is found that:

1) The kernel density gradient algorithm and the GWO-SVM model can segment and classify the sample of national land spatial layout to produce five types of urban planning land use, including urban

residential land, industrial land, transportation land, warehousing land, and public utility land, and the accuracy of segmentation and classification for each land use type can reach more than 90%. It shows that the image recognition and processing technology chosen in this paper can fully meet the practical needs of this research.
2) The SVM classification model optimized with the swarm intelligence optimization algorithm is better than the individual SVM classification model, and among the many SVM classification models optimized with the swarm intelligence optimization algorithm, the recognition accuracy of the GWO-SVM classification model designed in this paper is 95.93%, which is about 4.44% and 5.65% higher than that of the PSO-SVM classification model and the GA-SVM classification model, respectively. 4.44% and 5.65%, respectively. It indicates that the classification and recognition performance of GWO-SVM is more suitable for studying the optimization problem of land spatial layout planning.
3) Although the Markov-FLUS coupled model predicts that the proportion of each land use type will increase after 5 years, it should be noted that the model predicts that the proportion of residential land will decrease by about 3%, the proportion of industrial land will decrease by about 2%, and the proportion of public facility land will increase by about 8% after 10 years compared with 5 years. This result suggests that policymakers should plan the future spatial distribution of land more rationally, so as to promote the coordinated development of urbanization.

References

[1] Di Ludovico, D. F. . (2019). European cross-scale spatial planning and territorial frames in the italian median macroregion. European Planning Studies, 27(7).

[2] Zhang, Y. , Khan, S. U. , Swallow, B. , Liu, W. , & Zhao, M. . (2022). Coupling coordination analysis of china's water resources utilization efficiency and economic development level. Journal of cleaner production.

[3] Xu, Y. L. A. . (2019). Regional economic development coordination management system based on fuzzy hierarchical statistical model. Neural computing & applications, 31(12).

[4] Wu, K. , Chen, K. , Gao, Y. , Jiang, S. , & Huang, H. . (2022). Applying a set of potential methods for the integrated assessment of the marine eco-environmental carrying capacity in coastal areas. Sustainability, 14.

[5] Zhu, Y. , Wang, Z. , & Zhu, X. . (2023). Utilization effect of water-land resources under the evolution of chinese dietary patterns. Journal of Geographical Sciences, 33(4), 741-759.

[6] Lv, H. , Zheng, X. , & Chen, S. . (2022). Revealing the pattern of causality in processes of urbanization and economic growth: an evidence from china. Scientific programming(Pt.14), 2022.

[7] Miaomiao, Xingyu, Yining, Huang, & Zongwei. (2017). Epidemic transition of environmental health risk during china's urbanization. Science Bulletin, 62(2), 92-98.

[8] Yang, P. , Zhang, L. , & Tao, G. . (2023). Smart chemical industry parks in china: current status, challenges, and pathways for future sustainable development. Journal of loss prevention in the process industries, 83.

[9] Chen, Z. , Fu, W. , Bosch, C. V. D. , Pan, H. , Huang, S. , & Zhu, Z. , et al. (2019). National forest parks in china: origin, evolution, and sustainable development. Forests.

[10] Jinlang, Z. , & Qun, W. . (2017). Spatial analysis of chinese grain production for sustainable land management in plain, hill, and mountain counties. Sustainability, 9(3), 348.

[11] Zhang, Y. , Long, H. , Tu, S. , Ge, D. , & Ma, L. . (2019). Spatial identifcation of land use functions and their tradeos/synergies in china: implications for sustainable land management. Ecological Indicators, 107(4).

[12] Untenecker, J. , Bärbel Tiemeyer, Freibauer, A. , Laggner, A. , & Jürg Luterbacher. (2017). Tracking changes in the land use, management and drainage status of organic soils as indicators of the effectiveness of mitigation strategies for climate change. Ecological Indicators.

[13] Shrestha, B. , Cochrane, T. A. , Caruso, B. S. , Arias, M. E. , & Wild, T. B. . (2021). Sediment management for reservoir sustainability and cost implications under land use/land cover change uncertainty. Water Resources Research.

[14] Chaolin, G. U. , Lingqian, H. U. , & Cook, I. G. . (2017). China's urbanization in 1949–2015: processes and driving forces. Chinese Geographical Science, 27(006), 847-859.

[15] Yang, J. , Wu, T. , & Gong, P. . (2017). Implementation of china's new urbanization strategy requires new thinking. Science Bulletin, 62(2), 81-82.

[16] Meng, S. , Pham, T. , Dwyer, L. , & Grant, B. . (2021). Carbon price impacts on the chinese tourism industry:. Journal of Travel Research, 60(2), 370-383.

[17] Zhang, H. , Zhang, S. , & Liu, Z. . (2020). Evolution and influencing factors of china's rural population distribution patterns since 1990. PLoS ONE, 15(5), e0233637-.