

OPTIMIZATION PATH OF LANDSCAPE ARCHITECTURAL DESIGN BASED ON PCA-SSA-BP ALGORITHM

Siwen DOU[✉]

Anshan Normal University, 114007 Liaoning, China

Highlights:

- The study uses principal component analysis to downscale the multidimensional data of landscape architecture;
- The findings demonstrate that the sparrow algorithm-optimized BP neural network can efficiently handle complicated multidimensional data;
- The study provides a scientific, efficient and data-driven optimization method for landscape architecture design.

Article History:

- received 4 June 2025
- accepted 4 December 2025

Abstract. With the acceleration of urbanization, the importance of landscape architecture design in improving the quality of urban ecological environment and the quality of life of residents is becoming more and more prominent. To enhance the scientificity and practicability of the landscape design scheme, the study uses principal component analysis to downscale the multidimensional data of landscape architecture. It also combines Sparrow Search Algorithm (SSA) to optimize the initial weights and biases of Back Propagation (BP) neural network, so as to realize the accurate evaluation and adjustment of the design scheme. Meanwhile, the performance of the algorithm is verified on the Garden Buildings and WHU Building datasets. The experimental results showed that the mean square error of the BP neural network optimized by the sparrow algorithm was 0.1228, and the mean absolute error was 0.2321. Meanwhile, the accuracy and recall rate of this algorithm after convergence were significantly higher than those of PSO-BP, GA-BP and BP, with values of 0.946 and 0.951, respectively. In addition, the algorithm performed well in terms of vegetative coverage enhancement, space utilization improvement, and visitor experience enhancement. After the specific optimization of this algorithm, the vegetation coverage rate increased by 14.5%, and the space utilization rate increased by 13.2% and 14.8%, respectively. The findings demonstrate that the sparrow algorithm-optimized BP neural network can efficiently handle complicated multidimensional data, optimize the landscape design scheme, and improve user experience and ecological value. The study provides a scientific, efficient and data-driven optimization method for landscape architecture design, which helps to promote the digital transformation and sustainable development of the landscape design industry.

Keywords: PCA-SSA-BP, landscape architectural design, data dimensionality reduction, L2 regularization, performance evaluation.

[✉]Corresponding author. E-mail: dousiwenn@163.com

Notations

μ – mean vector
 σ – standard deviation vector
 X – original data
 C – covariance matrix
 X^T – transposition matrix
 n – sample size
 $K(x, y)$ – Kernel function
 K_{ij}^- – centralized core matrix
 α_i – feature vector
 $z^{(t+1)}$ – weighted input
 $W^{(l)}$ – weight matrix
 $a^{(l)}$ – activation value
 $b^{(l)}$ – bias vector
 f – Activation function
 η – learning rate
 lb – lower bound of the variable

ub – upper bound of variables
 t – current number of iterations
 v – step size factor
 R_2 – warning value
 ST – safety value
 Q – random numbers with a normal distribution
 x_{best}^t – the current optimal individual position
 x_{worst}^t – the current worst individual position
 A^+ – random matrix
 $fitness$ – fitness function
 B – boundary processing function
 y_i^- – model predicted value
 y_i – true value
 λ – regularization coefficient
 ξ – the quantity of weights
 $\frac{\partial Loss}{\partial W}$ – the gradient of the loss function against the bias

1. Introduction

With the acceleration of urbanization, the data dimensions that landscape architecture design needs to handle are becoming increasingly complex, including but not limited to vegetation distribution, climatic conditions, topography and landforms, soil types, sunshine duration, and tourist flow. The complex correlations among these multi-dimensional data could lead to a significant increase in computational complexity if processed directly (Chen & Stouffs, 2023; Freschlin et al., 2024). In recent years, the application of data analysis and machine learning technologies to improve landscape design has gradually become a major research hotspot. However, existing machine learning algorithms are clearly limited in their ability to handle the complexity and multi-dimensionality of landscape data (Goodarzi et al., 2023; Groumpos, 2023). For example, the pArticle swarm optimization-back propagation neural network (PSO-BP) model combines the global search ability of the pArticle swarm optimization algorithm with the local optimization ability of the Back Propagation (BP) neural network. However, it is prone to premature convergence, which results in suboptimal solutions. Although the genetic algorithm-back propagation neural network (GA-BP) improves the performance of BP neural networks (NPNs) to a certain extent, its computational efficiency is relatively low and it is difficult to meet the rapid design requirements in practical applications. Furthermore, the lack of comprehensive data integration and preprocessing steps also limits the accuracy and reliability of these models (Ardhianto et al., 2023; Fernberg & Chamberlain, 2023).

At this stage, some scholars have also conducted research on the optimization of landscape architectural design. Liu et al. (2024) proposed a landscape design model based on PSO-BP to address the challenges faced by urban virtual environment landscape design in the process of urbanization. The model combined the global search capability of pArticle swarm optimization algorithm and the local optimization capability of BPNN, and optimized the design of the color and layout of the urban landscape through data-driven approach. The results indicated that after applying the PSO-BPNN model, the public satisfaction of the urban landscape was significantly improved. However, the PSO-BP algorithm tended to fall into a Local Optimum (LO) when dealing with small samples of data. It was also relatively sensitive to the selection of initial parameters, which limited its ability to generalize in practical applications (Liu et al., 2024). To address the existing state of environmental landscape design and planning, Chen (2023) proposed a method based on computer vision and deep learning. The system utilized deep neural networks and anthropocentric artificial intelligence techniques to optimize landscape design through a data-driven approach. The results indicated that the system demonstrated high accuracy in the rationality assessment of landscape design. The judgment rate of its 3D image simulation continued to improve with the expansion of the scope, which was significantly better than the traditional

method. However, this system relied on a large amount of labeled data, which could be difficult to obtain in practical applications, especially in some regions with limited resources. In addition, the adaptability of this system in dealing with complex ecosystems still needed to be improved. Hong et al. (2023) proposed an aesthetic style conversion method based on deep neural networks to address the lack of aesthetic significance of Chinese traditional garden virtual scenes from historical periods in digital preservation. The method consisted of style prediction and style conversion, by embedding the pre-trained network into the scene rendering pipeline. Moreover, combined with the screen post-processing method, the stylized representation of the virtual scene is realized. The outcomes proved that the technique could successfully achieve the virtual scene's aesthetic style conversion. However, it had a high demand for computing resources and was less efficient when dealing with large-scale datasets. Wu et al. (2023) proposed an ecological quality assessment model based on multi-source data fusion in view of the current situation of ecological quality in scenic spots in China. The results showed that from 2000 to 2017, the regional ecological quality of approximately 60.3% in China improved. Although the overall ecological quality increased over the past 18 years, the level remains relatively low. Chu et al. (2026) proposed a multi-objective optimization model based on the Improved Whale Optimization Algorithm (IWOA) in order to solve the contradiction between water supply and demand in multi-objective regions, which was used when considering multiple objectives such as economy, society, and environment. The model made rational allocations of regional water resources. The results showed that this model could effectively improve the utilization efficiency of water resources and achieve a good balance among economic, social and environmental benefits. It provided a new method and idea for solving the contradiction between water supply and demand.

Although the above studies have achieved better outcomes in the design of landscape architecture, there are problems such as limited model Generalization Ability (GA) and insufficient consideration of complex ecosystems. Based on this, the study proposes a Principal Component Analysis (PCA) downscaling and Principal Component Analysis-Sparrow Search Algorithm-Back Propagation Neural Network (PCA-SSA-BP). The research aims to address the complexity of landscape architectural design data as well as the optimization efficiency. The innovation of the research lies in proposing a hybrid algorithm framework (PCA-SSA-BP) to optimize the parameters of a BPNN. This framework uses PCA for dimension reduction and the Sparrow Search Algorithm (SSA) to optimize landscape architectural design. Compared with existing optimization algorithms, SSA has a unique "warning mechanism" that effectively prevents the algorithm from falling into a LO. This significantly improves the global search ability and convergence speed. In addition, this study introduced L2 regularization technology, combining it with SSA-BP to further enhance the model's GA and stability

while effectively preventing overfitting. This collaborative innovation enhances the scientific and practical aspects of landscape design schemes and provides new technical support for the digital transformation and sustainable development of the landscape design industry.

2. Methods and materials

2.1. Dimensionality reduction of multidimensional data of landscape garden architecture based on improved PCA

In landscape design path optimization, the complex correlation between the data makes direct processing may lead to a significant increase in computational complexity. Therefore, the study first introduces the PCA method by downscaling the multidimensional data of landscape garden architecture. Moreover, to preserve as much of the Original Data (OD)'s variance information as possible, the OD are split up into many Principal Components (PC). In Figure 1, the pertinent schematic is displayed.

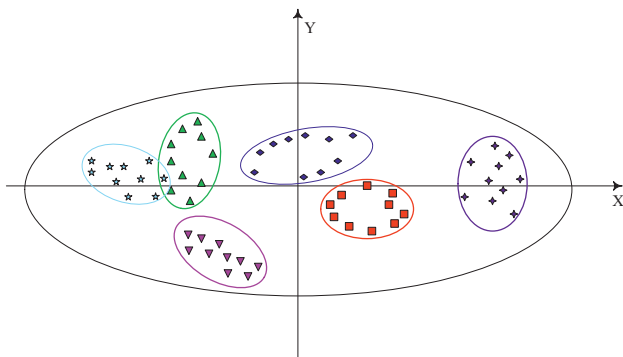


Figure 1. PCA schematic diagram

The fundamental principle of the PCA technique is to use orthogonal transformation to convert the OD into a new coordinate system in Figure 1. The main steps of PCA can be divided into data normalization, calculating Covariance Matrix (CM), solving eigenvalues as well as eigenvectors, selecting PCs, and data dimensionality reduction (You et al., 2023; Ahmadi et al., 2023). The related schematic is shown in Figure 2.

In Figure 2, the primary process of PCA is the standardization of data, defining the OD as X . By standardizing the OD of landscape architecture so that the mean is 0 and the variance is 1, the standardized data matrix can be obtained. Second, the CM is calculated, and the related mathematical expression is shown in Eq. (1).

$$\begin{cases} X^- = \frac{X - \mu}{\sigma} \\ C = \frac{1}{n-1} X^T X^- \end{cases} \quad (1)$$

In Eq. (1), μ is the mean vector of the data. σ is the standard deviation vector of the data. C is the CM. X^T is the transpose matrix. n is the sample size. In PCA, the OD is first standardized so that the mean of each feature is 0 and the standard deviation is 1. Second, the CM is computed to measure the linear relationships among different features. Subsequently, an eigenvalue decomposition is performed on this CM to obtain the eigenvalues and their corresponding eigenvectors. Finally, the study determines the number of retained PCs based on the cumulative variance contribution rate when selecting the PCs. Specifically, the study selects PCs with a cumulative variance contribution rate of 85%. This criterion ensures that the data after dimensionality reduction can retain most of the information of the OD, while significantly reducing the dimensions of the data and improving the efficiency of subsequent processing. For example, if the original dataset has 50 features, only 10 PCs need to be retained after PCA processing to achieve a cumulative variance contribution rate of 85% (Yang et al., 2023; Zhang et al., 2023). For instance, in a certain landscape design dataset, it includes the following features: vegetation coverage rate (%), average sunshine duration (hours), soil type (classification code), terrain slope (degree), annual precipitation (millimeters), number of visitors (person-times), regional area (square meters), green space area (square meters), building density (%), and noise level (decibels). The original dataset before dimensionality reduction contains 10 features and 100 sample points. The research found that by standardizing the data and analyzing the eigenvalues, the cumulative contribution rate of the first three features reached 86.5%.

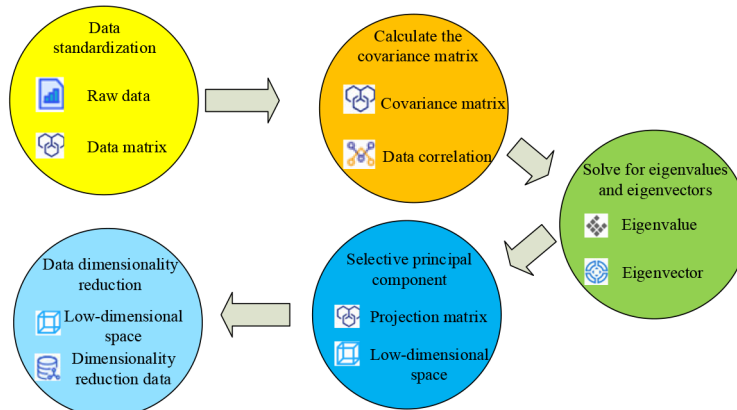


Figure 2. The main flow chart of PCA

This means that the first three PCs can explain more than 85% of the data. In the actual application of landscape architecture data, the appropriate number of PCs can be selected according to the specific needs to achieve the purpose of dimensionality reduction. Considering that the data of landscape architecture has a complex nonlinear structure, linear PCA may not be able to effectively extract the main features of the data, the study introduces a kernel function to optimize PCA. It achieves effective dimensionality reduction of nonlinear data by mapping the data to a high-dimensional feature space in which linear PCA is performed. To determine the σ parameter of the Gaussian Kernel (GK) function, the cross-validation method is adopted. Specifically, the σ value that optimizes model performance is selected by conducting multiple cross-validations on the training dataset. This method can effectively avoid overfitting and ensure the GA of the model on unseen data. The kernel function is the core of PCA analysis, and common kernel functions include GK, polynomial kernel, and linear kernel function. The GK function is selected for the study, and the related mathematical expression is shown in Eq. (2).

$$K(x, y) = \exp\left(-\frac{\|x - y\|^2}{2\sigma^2}\right). \quad (2)$$

In Eq. (2), the density of the kernel function can be controlled by controlling the parameter σ and thus the density of the kernel function. The original landscape architecture data is defined as $\{x_1, x_2, x_3, \dots, x_n\}$ while the kernel function is $K(x, y)$. The goal of the optimized PCA is mainly to find the PCs in the feature space. Firstly, the kernel matrix is calculated, and the elements of the kernel matrix K are $K_{ij} = K(x_i, y_i)$. Second, for the centering kernel matrix process as a way to eliminate the effect of the mean value of the data, the mathematical expression of the centered kernel matrix is shown in Eq. (3).

$$K_{ij}^- = K_{ij} - \frac{1}{n} \sum_{k=1}^n K_{ik} - \frac{1}{n} \sum_{k=1}^n K_{kj} + \frac{1}{n^2} \sum_{k,l=1}^n K_{kl}. \quad (3)$$

In Eq. (3), K_{ij}^- denotes the centered kernel matrix. Through the eigen-decomposition of this kernel matrix, the eigenvalues as well as the eigenvectors are obtained. Finally, the dimensionality reduced data is created by selecting the eigenvectors that correspond to the top d greatest eigenvalues. For the new data point x , its mathematical expression in the LDS is shown in Eq. (4).

$$\phi(x) = \sum_{i=1}^d \alpha_i K(x, x_i). \quad (4)$$

In Eq. (4), α_i denotes the feature vector.

2.2. Landscape architecture path optimization based on improved BPNN

After completing the dimensionality reduction processing of multi-dimensional data of landscape architecture, in an attempt to further improve the optimization effect

of landscape architecture design, the study introduces an improved BPNN. BPNN is a multilayer feed-forward neural network. In the process of BPNN forward propagation, the landscape architectural data downgraded by PCA is passed to the Hidden Layer (HL) through the Input Layer (IL) and then from the HL to the Output Layer (OL). In this case, the output mathematical expression of each layer is shown in Eq. (5).

$$\begin{cases} z^{(l+1)} = W^{(l)} \times a^{(l)} + b^{(l)} \\ a^{(l)} = f(z^{(l+1)}) \end{cases}. \quad (5)$$

In Eq. (5), $z^{(l+1)}$ denotes the weighted input of layer $l+1$. $W^{(l)}$ denotes the weight matrix of layers l through $l+1$. $a^{(l)}$ denotes the activation value of the l th layer. $b^{(l)}$ denotes the bias vector of layer l . f denotes the activation function. Among the activation functions the Sigmoid function is selected for the study. The BPNN updates the mesh parameters by BP algorithm in a way of calculating the error pair weights as well as the bias gradient. In the BP of error firstly for calculating the error in the OL. Second, the mathematical expression of the HL error is shown in Eq. (6).

$$\delta_j^{(l)} = \left(\sum_{k=1}^m \delta_k^{(l+1)} \times w_{jk}^{(l+1)} \right) \times f'(z_j^{(l)}). \quad (6)$$

In Eq. (6), $w_{jk}^{(l+1)}$ denotes the weight of layer $l+1$. Finally, the weights and biases are updated and the related mathematical expression is shown in Eq. (7).

$$\begin{cases} \Delta w_{ij}^{(l)} = -\eta \cdot \delta_j^{(l+1)} \cdot a_i^{(l)} \\ \Delta b_j^{(l)} = -\eta \cdot \delta_j^{(l+1)} \end{cases}. \quad (7)$$

In Eq. (7), η denotes the learning rate. The traditional BPNN is easy to fall into LO during the training process, and also has the problem of weak global optimization ability and long training time when dealing with nonlinear problems. In light of these issues, this work presents the SSA, which optimizes the BPNN's initial weights and bias to increase both the BPNN's safe search capability and convergence speed. Through the cooperative efforts of the discoverer, joiner, and detector, the SSA simulates the foraging and anti-predator actions of the sparrow in order to complete the optimization process (Zheng et al., 2023; Li et al., 2023). For the parameter settings of SSA, the population size is set to 30, which is based on a balanced consideration of the convergence speed of the algorithm and computing resources. Although a larger population size can improve the algorithm's ability to perform a global search, it will also increase the computational cost. The maximum number of iterations is set to 100. Through multiple experiments, the researchers found that, in most cases, 100 iterations can enable the algorithm to reach a stable convergence state while avoiding the unnecessary

waste of computing resources. In the individual update of SSA, if the current environment is safe, the discoverer will conduct a global search to explore new solution Spaces. If a predator is detected, the discoverer will perform a random walk to increase population diversity and avoid becoming trapped in a LO. For participants, if their fitness value is high, a local search will be conducted around the optimal individual to further optimize the solution. If the fitness value of the participant is poor, a random search will be conducted around the worst individual to increase the diversity of the population. For investigators, they assess the safety of the current environment based on the fitness function. If predators are detected, they will randomly move to a new location to increase the diversity of the population. During the SSA's parameter optimization process, the step size factor determines the search step size, affecting the algorithm's global and local search capabilities. This study compared the influence of different step size factor settings on algorithm performance through experimentation. Finally, 0.5 is selected as the step size factor. Its good balance is demonstrated in multiple experiments. This ensures the algorithm's global search ability and enhances the accuracy of local searches in later stages. The warning value is used to determine whether the algorithm converges prematurely and avoid falling into LO. Through sensitivity analysis and multiple experiments, the research determines 0.8 as the early warning value. This value can effectively prevent the algorithm from converging prematurely without increasing excessive computational load. The basic flow of SSA is shown in Figure 3.

In Figure 3, the population initialization is done by randomly generating the positions of the sparrows. The position of each sparrow is defined as an d -dimensional vector, which is denoted as $X = \{x_1, x_2, \dots, x_n\}$. Among them, n denotes the size of the population. Therefore, the mathematical expression of population initialization is shown in Eq. (8).

$$x_i = lb + rand \times (ub - lb). \quad (8)$$

In Eq. (8), lb and ub denote the variable lower and upper bounds. Subsequently, the discoverer's position is updated using Eq. (9), following the computation of the fitness value of each sparrow's associated position.

$$x_i^{t+1} = \begin{cases} x_i^t \times \exp\left(-\frac{i}{v \times iter_{max}}\right) & \text{if } R_2 < ST \\ x_i^t + Q & \text{if } R_2 \geq ST \end{cases} \quad (9)$$

In Eq. (9), t denotes the current iteration quantity. v denotes the step factor. R_2 is the warning value. ST is the safety value. Q denotes a random quantity obeying a normal distribution. When $R_2 < ST$, it indicates that the current environment is safe and the discoverer will perform a global search. The opposite indicates that a predator is detected and the discoverer will wander randomly (Liu et al., 2023; Wang et al., 2024). Eq. (10) is used to update the joiner's location after the discoverer's position has been modified.

$$x_i^{t+1} = \begin{cases} x_{best}^t + |x_i^t - x_{best}^t| \times A^+ & \text{if } i < \frac{n}{2} \\ x_{worst}^t \times \exp\left(\frac{x_{worst}^t - x_i^t}{i^2}\right) & \text{if } R_2 \geq \frac{n}{2} \end{cases} \quad (10)$$

In Eq. (10), x_{best}^t is the location of the current optimal individual. x_{worst}^t denotes the current location of the worst individual. A^+ denotes a random matrix that takes the value of ± 1 . When $i < \frac{n}{2}$, the joiner performs a local search around the optimal individual. Conversely, the joiner performs a random search around the worst individual. After the joiner position is updated, the position of the detector is updated based on Eq. (11). The related mathematical expression is shown in Eq. (11).

$$x_i^{t+1} = \begin{cases} x_{best}^t + rand \times (ub - lb) & \text{if } fitness(x_i^t) > fitness(x_{best}^t) \\ x_i^t + rand \times (x_{best}^t - x_i^t) & \text{otherwise} \end{cases} \quad (11)$$

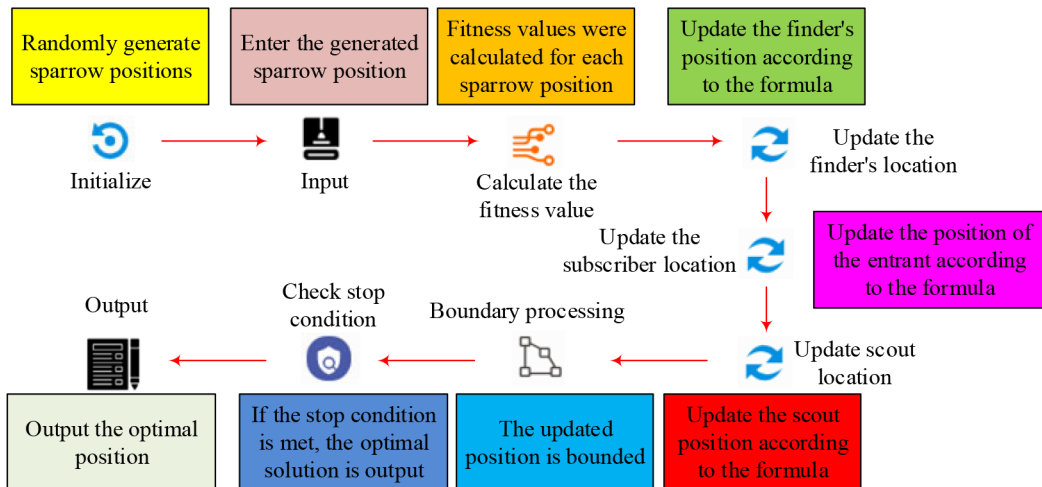


Figure 3. Basic flowchart of the SSA

In Eq. (11), *fitness* is denoted as the fitness function. The scout will randomly move to a new location when it finds a predator in order to increase the diversity of the population. Finally, to ensure that the sparrow's position lies within the search space, a boundary treatment of the unknown is required. The related mathematical expression is shown in Eq. (12).

$$x_i^{t+1} = B(x_i^{t+1}, lb, ub). \quad (12)$$

In Eq. (12), *B* denotes the boundary processing function, which ensures that the value of each dimension is between the upper and lower bounds. To avoid overfitting of BPNNs, the study introduces L2 regularization to optimize BPNNs. By adding a penalty term to the loss function that is connected to the sum of squares of the weights, it keeps the weights minimal during training and keeps the network model from overfitting. Eq. (13) displays the loss function's mathematical expression for L2 regularization.

$$Loss = \frac{1}{2n} \sum_{i=1}^a (y_i^- - y_i)^2 + \lambda \sum_{j=1}^{\xi} W_j^2. \quad (13)$$

In Eq. (13), *a* denotes the number of training samples. y_i^- denotes the predicted value of the model. y_i denotes the true value. λ denotes the regularization coefficient, which is used to control the strength of regularization. For the regularization coefficient, the research adopts a grid search strategy to systematically traverse a series of values. Specifically, from 10⁻⁵ to 1, this range covers a range from almost no regularization to strong positive. The model for each regularization coefficient is trained and the performance metrics are recorded. W_j denotes the weights in the network. ξ denotes the number of weights. In the BPNN BP process, the weights are updated as shown in Eq. (14).

$$\Delta W = -\eta \left(\frac{\partial Loss}{\partial W} + 2\lambda W \right). \quad (14)$$

In Eq. (14), $\frac{\partial Loss}{\partial W}$ denotes the gradient of the loss function with respect to the weights. In addition the bias is updated as shown in Eq. (15).

$$\Delta b = -\eta \frac{\partial Loss}{\partial b}. \quad (15)$$

In Eq. (15), *b* denotes the bias. $\frac{\partial Loss}{\partial b}$ denotes the gradient of the loss function to the bias. Through L2 regularization, the BPNN can be effectively prevented from overfitting problems when dealing with multidimensional data of landscape architecture, and the GA of the model can be improved at the same time (Gao et al., 2023). In the BPNN, the number of neurons in the input layer, HL and OL is 20, 30, and 10, respectively. At the same time, the number of HLs is 1. The HL activation function uses the Sigmoid activation function, while the OL activation function uses the linear activation function. The learning rate is set to 0.05 and the SGD optimizer is selected as the optimizer. Finally, the landscape architecture design optimization path flow is shown in Figure 4.

In Figure 4, the landscape architecture design optimization path process is described as follows. First, it is the data collection and pre-processing process, which eliminates the influence of different scales as well as orders of magnitude by standardizing the data. Subsequently, it is the dimensionality reduction process of PCA. The BPNN's initial parameters are then adjusted for the SSA, with each SSA location corresponding to a set of BPNN starting weights and biases. Next, the BPNN is trained as well as optimized after introducing the L2 regularization technique in the training process. Finally, the optimized BPNN is applied to the evaluation of landscape architectural design schemes. The design scheme is dynamically adjusted according to the evaluation results, and the procedure is repeated until the optimal design scheme is obtained.

To enhance the efficiency and quality of the scheme formulation, the following optimizations are carried out in the process of embedding the algorithm into landscape design. During the initial design stage, it is important to communicate with landscape designers and relevant stakeholders to clarify design requirements and goals. Subsequently, PCA-SSA-BP is used to generate the preliminary design scheme. The model generates multiple feasible design schemes based on the input data and optimization objectives. Second, the generated design

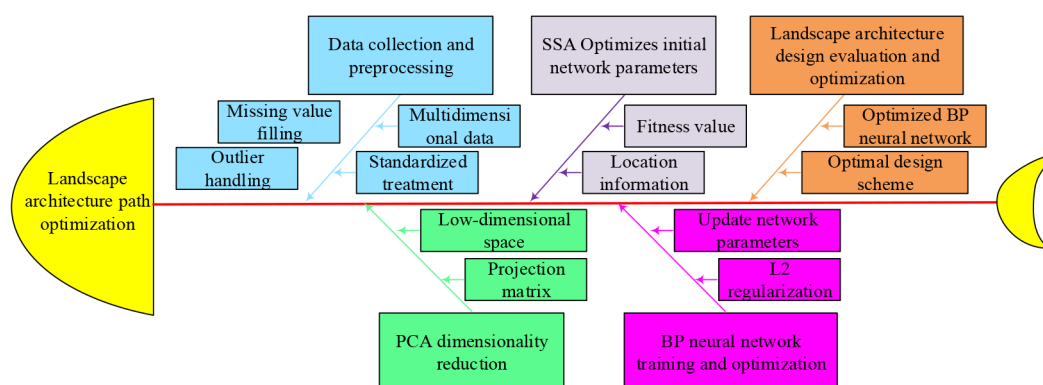


Figure 4. Landscape architecture design optimization path flow

scheme is evaluated, using indicators such as accuracy rate, recall rate, and F1 score to assess the performance of the scheme. At the same time, landscape designers and relevant experts are invited to review the plan and provide feedback. Then, based on the assessment results and expert feedback, adjustments and optimizations are made to the design plan. The adjusted plan is evaluated again through the model to ensure its scientificity and practicality. After multiple iterations and optimizations, the design plan is determined. The final plan is submitted to the relevant stakeholders for implementation and application.

3. Results

3.1. Performance evaluation of the PCA-SSA-BP algorithm

To make sure the reproducibility and scientific validity of the results, high performance computer equipment is used to ensure that the algorithms can operate in an efficient and stable environment. Table 1 displays the first study's experimental setup and algorithm parameters.

Based on the experimental environment in Table 1 and the parameter information of the algorithm, the study selects the GardenBuildings dataset and the WHU Building dataset in terms of datasets. The GardenBuildings dataset covers landscape architecture design projects in multiple cities. The data is sourced from on-site measurements, satellite remote sensing, and historical archive records. The WHU Building Dataset contains precise data on buildings

and natural features, including building heights, types of vegetation, types of land use, and areas of tourist activity. It is derived from field measurements, unmanned aerial vehicle aerial photography, and Geographic Information System (GIS) data. For the obtained data, the data are preprocessed by means of mean filling, Z-score, and standardization processing. Meanwhile, 85% of the cumulative variance contribution rate is set as the critical value for feature selection. Finally, the dataset is divided into the training set and the test set in a ratio of 7:3. To ensure the stability and repeatability of the experimental results, the study introduces randomness control. That is, each experiment is repeated multiple times, and random seeds are set to ensure the consistency of the results. Each experiment is repeated ten times to ensure the stability and reliability of the results. In terms of random seeds, a fixed initialization seed is used in the research of the SSA to ensure that the initial population of the algorithm is consistent each time it runs. In addition, the study includes the following variables: vegetation coverage rate, diversity of vegetation species, space utilization rate, path accessibility, and rationality of rest area layout. The vegetation coverage rate, for example, is the proportion of the area covered by plants and reflects the density of vegetation in landscape design. The diversity of vegetation species indicates the quantity of vegetation species and reflects the diversity of vegetation in landscape design. Space utilization rate represents the proportion of effectively utilized space and reflects the efficiency of space utilization in landscape design. Path accessibility indicates whether the location, quantity and

Table 1. Experimental environment and algorithm parameters

Configuration	Model number	BP neural network parameters	Parameter information	Set SSA parameters	Parameter information	The Garden-Buildings dataset	Numerical values and descriptions	WHU Building Dataset	Numerical values and descriptions
CPU	Intel Core i9-13900K	Number of neurons in the input layer	20	Population size	30	Sample size	1000	Sample size	800
GPU	NVIDIA GeForce RTX 4090	Number of hidden layer neurons	30	Maximum iterations	100	Feature dimension	50	Feature dimension	40
Internal memory	64GB	Number of neurons in the output layer	10	Step factor	0.5	Data source	Landscape design projects, on-site measurement and satellite remote sensing data collection	Data source	The Geographic Information System Laboratory of a certain university
Hard disk	1TB NVMe SSD	Learning rate	0.05	Prewarning value	0.8	/	/	/	/
Operating system	Windows 11 Professional	Optimizer	SGD	Safe value	0.7	/	/	/	/
Python	Python 3.8	/	/	/	/	/	/	/	/
Deep learning framework	TensorFlow 2.3.0	/	/	/	/	/	/	/	/

design of rest areas are reasonable, which reflects the passage experience of tourists in the landscape. The rationality of the rest area layout indicates whether its location, quantity, and design are reasonable. This reflects whether the layout meets tourists' needs. Traditional BPNN, PSO-BP, and GA-BP are chosen for comparison tests in this study. Figure 5 displays the four algorithms' accuracy and recall curves.

Figure 5a displays the comparison of the accuracy curves of the four algorithms. According to the findings, the PCA-SSA-BP algorithm, which is employed in the study, has the highest accuracy value as the number of iterations increases. Its converged accuracy approaches to 0.946. The accuracy size is followed by PSO-BO, GA-BP, and traditional BP. The comparison of the four algorithms' recall curves is displayed in Figure 5b. The outcomes display that the recall of PCA-SSA-BOP algorithm is significantly higher than the rest of the algorithms as the iteration increases.

Its final recall converges to 0.951. The values of PSO-BP and GA-BP recall after convergence are closer. The traditional BP has the lowest recall after convergence with a value converging to 0.695. The study further compares the Mean-Square Error (MSE) and Mean Absolute Error (MAE) of the four algorithms. Figure 6 displays the findings.

In Figure 6, the PCA-SSA-BP algorithm proposed in the study shows low error values on both datasets. Each algorithm corresponds to two color bands, which represent the GardenBuildings dataset and the WHU Building Dataset, respectively. The outcomes display that the PCA-SSA-BP algorithm used in the study has small errors on both datasets. Its MSE is 0.1228 and 0.0982, and MAE is 0.2321 and 0.1987, respectively. It shows that its prediction accuracy is high and it can handle complex garden building data better. The next highest error is GA-BP, which significantly improves the performance of BPNN by optimizing the initial parameters through genetic algorithm.

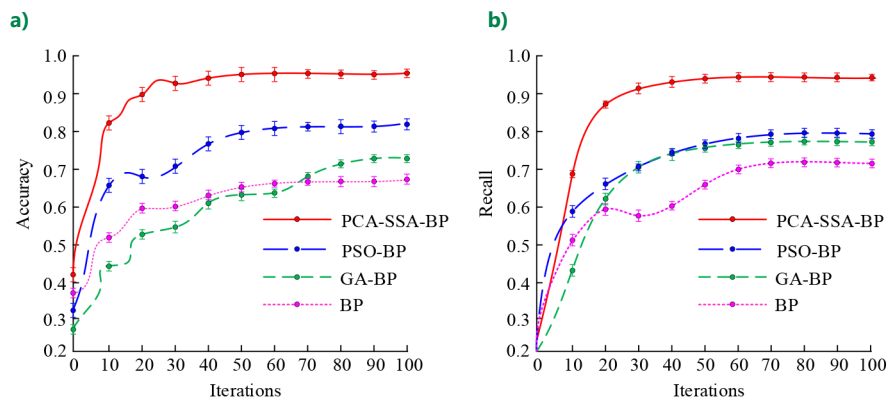


Figure 5. Accuracy curve and recall curve comparison: a) Accuracy curve comparison; b) Recall curve comparison

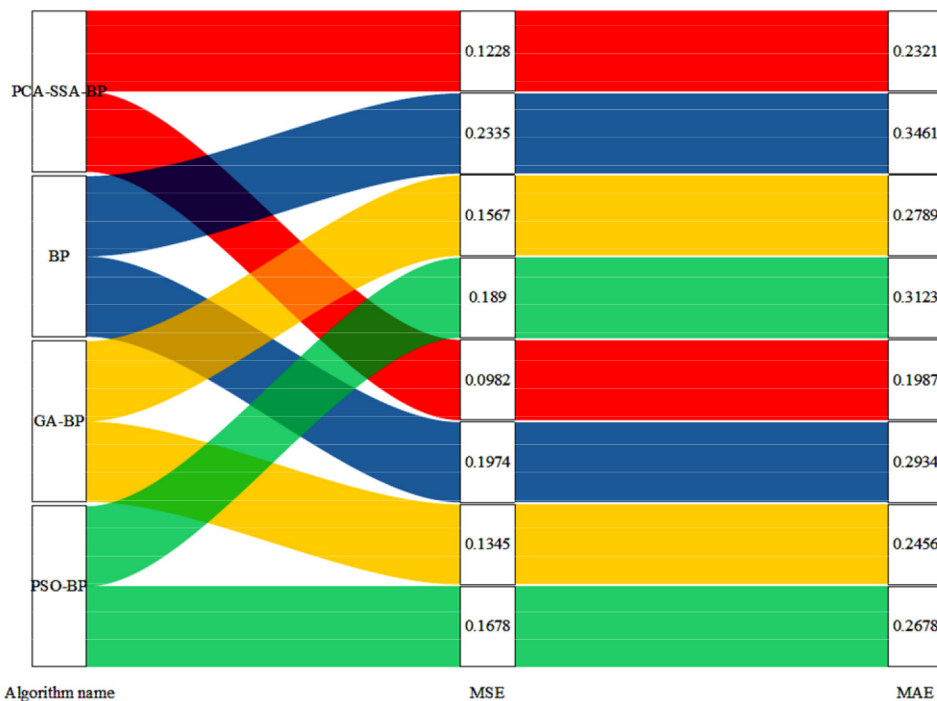


Figure 6. Comparison of MSE and MAE

The largest error is the traditional BPNN. Lower prediction accuracy results from LO's ease of occurrence and increased sensitivity to bias and initial weight selection. In summary, the PCA-SSA-BP algorithm has high prediction accuracy when dealing with complex multidimensional data. The study further cross-validates the four algorithms. The study further conducts cross-validation on the four algorithms. Among them, 5-fold cross-validation is a specific form of cross-validation, which divides the dataset into five equally sized subsets. Each time, one subset is selected as the test set and the remaining four are selected as the training set. This process is repeated five times, with each subset having one chance to be the test set. Figure 7 displays the error outcomes.

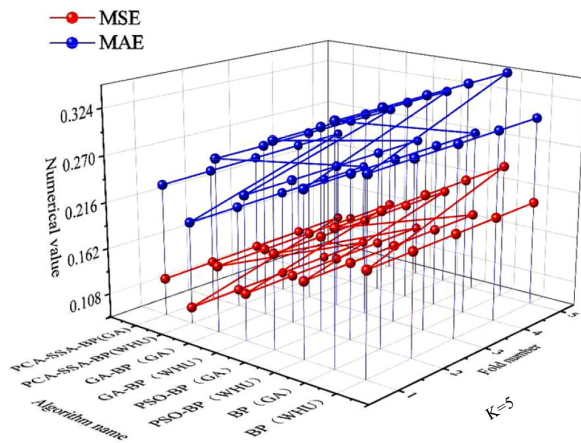


Figure 7. Cross validation result

Figure 7 shows the cross-validation results of the four algorithms on the two datasets. The results show that the changes in MSE and MAE of the PCA-SSA-BP algorithm at different folds are very small. Its MSE and MAE tend to 0.1250 and 0.2370, respectively, and tend to 0.1001 and 0.2015 on the WHU Building Dataset. The algorithms with the second-highest errors are GA-BP and PSO-BP in sequence, and the traditional BP has the largest error. The study further compares the computing times of each algorithm. Figure 8 displays the findings.

In Figure 8, the PCA-SSA-BP algorithm has the fastest convergence speed and the shortest computing time after convergence. The specific PCA-SSA-BP algorithm converges when the number of iterations is 40. At this time, the calculation times are 121.37 s and 107.27 s respectively. In contrast, BP has the slowest convergence speed and longest computing time. It converges after 80 iterations, with a final computing time of 353.67 and 322.73 s, respectively. The calculation times of GA-BP and PSO-BP converge between PCA-SSA-BP and BP, and their calculation times are 180.63 s, 159.37 s and 243.73 s, 206.80 s, respectively. In summary, the PCA-SSA-BP algorithm is significantly superior in terms of calculation time and provides an efficient, effective method for optimizing landscape architectural design. The study also compares the four algorithms' differences in various indicators on dense horticultural structures and predominantly natural landscapes. The results are shown in Table 2.

In Table 2, the PCA-SSA-BP algorithm shows significant advantages compared with the other three algorithms. Specifically, in the Dense Horticultural Structures dataset, PCA-SSA-BP achieves an increase of 14.2% in vegetation coverage rate, which is much higher than 10.5% of GA-BP, 9.3% of PSO-BP and 4.8% of BP. In terms of the improvement of vegetation species diversity, PCA-SSA-BP achieves an increase of 6 species, while GA-BP, PSO-BP, and BP increased by 4, 3, and 2 species, respectively. In terms of the improvement of space utilization rate, PCA-SSA-BP reaches 12.8%, which is superior to 11.2% of GA-BP, 10.5% of PSO-BP, and 7.5% of BP. In terms of path accessibility score and rest area layout score, PCA-SSA-BP achieves 8.4 and 8.7, respectively, both of which are higher than those of other algorithms. In the Predominantly Natural Landscapes dataset, PCA-SSA-BP achieves 14.8%, 6, 14.5%, 8.3, and 8.6 respectively in terms of vegetation coverage rate improvement, vegetation species diversity improvement, space utilization rate improvement, path accessibility score, and rest area layout score. It is also superior to GA-BP, PSO-BP and BP. The results demonstrate the effectiveness of the PCA-SSA-BP algorithm in improving vegetation coverage, diversity, space utilization, and tourist

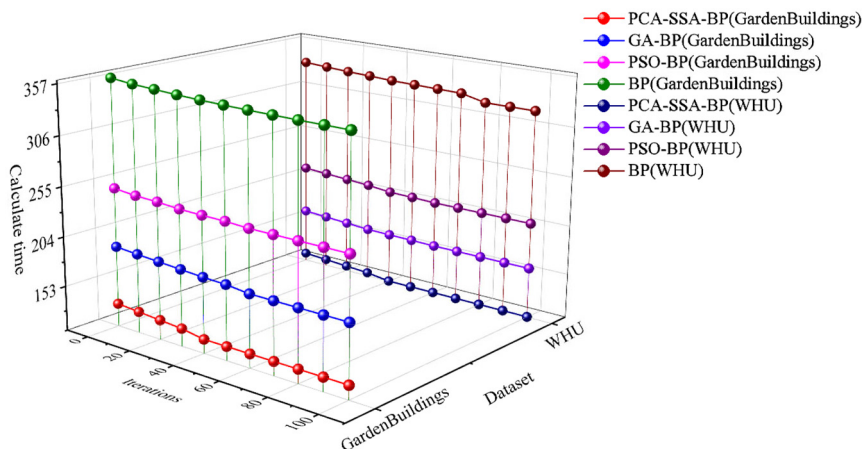


Figure 8. Computation time comparison

Table 2. Comparison of differences in various indicators on dense horticultural structures and predominantly natural landscapes

Dataset	Indicator	PCA-SSA-BP	GA-BP	PSO-BP	BP
Dense horticultural structures	Vegetation coverage rate increase (%)	14.2	10.5	9.3	4.8
	The diversity of vegetation species has increased	6	4	3	2
	Space utilization rate improvement (%)	12.8	11.2	10.5	7.5
	Path accessibility score	8.4	8	7.6	7.2
	Score for the layout of the rest area	8.7	8.3	7.9	7.5
Predominantly natural landscapes	Vegetation coverage rate increase (%)	14.8	10.3	9.5	6.7
	The diversity of vegetation species has increased	6	4	3	2
	Space utilization rate improvement (%)	14.5	12.1	11	8.6
	Path accessibility score	8.3	7.9	7.5	7.1
	Score for the layout of the rest area	8.6	8.2	7.8	7.4

experience in various landscape scenarios. This shows its strong adaptability and superiority in optimizing landscape architectural design. To further verify the performance of PCS-SSA-BP, the study finally compares it with GA, PSO, SVM, and RF. The results are shown in Table 3.

Table 3. Performance comparison results of five algorithms

Dataset	Algorithm	Accuracy (%)	Recall (%)	F1 score (%)
Garden-Buildings	PCA-SSA-BP	94.7	95.3	95
	GA	83.2	84.5	83.8
	PSO	82.9	83.7	83.3
	SVM	84.6	85.2	84.9
	RF	83.5	84.1	83.8
WHU Building Dataset	PCA-SSA-BP	95.4	96	95.7
	GA	83.8	84.8	84.3
	PSO	83.5	84.3	83.9
	SVM	84.9	85.5	85.2
	RF	84.2	84.7	84.5

In Table 3, in the performance comparison of the GardenBuildings Dataset and the WHU Building Dataset, the PCA-SSA-BP algorithm shows significant advantages over other algorithms such as GA, PSO, SVM, and RF.

Specifically, on the GardenBuildings dataset, the accuracy rate of PCA-SSA-BP reaches 94.7%, the recall rate is 95.3%, and the F1 score is 95.0%, which are much higher than 83.2%, 84.5%, and 83.8% of GA, 82.9%, 83.7%, and 83.3% of PSO, 84.6%, 85.2%, and 84.9% of SVM, as well as the 83.5%, 84.1%, and 83.8% of RF. On the WHU Building Dataset, the accuracy rate, recall rate, and F1 score of PCA-SSA-BP are 95.4%, 96.0%, and 95.7% respectively, which are also significantly superior to other algorithms. The results show that the PCA-SSA-BP algorithm provides higher prediction accuracy and optimization efficiency when dealing with complex, multidimensional data. This makes it a more scientific and effective solution for optimizing landscape architectural design.

3.2. Evaluation of the optimization effect of design solutions

After evaluating the performance of the PCA-SSA-BP algorithm, the study further evaluates the optimization effect of the design scheme in terms of vegetation layout optimization, space utilization efficiency, and visitor experience enhancement. In terms of the evaluation of tourist experience, the study quantifies the accessibility of the path. That is, 10 landscape design experts are invited to rate the accessibility of the path, with the scoring range being from 1 to 10 points. Among them, 1 to 3 points indicate that there are obvious obstacles on the path and it is difficult to pass through. A score of 4 to 6 indicates that the path is basically feasible, but there are some minor issues. A score of 7 to 10 indicates that the path design is reasonable and the passage is smooth. At the same time, the study also invited experts to rate the layout of the rest area. Among them, 1 to 3 points indicate that the layout of the rest area is unreasonable. A score of 4 to 6 indicates that the layout is generally good, though there is room for improvement. A score of 7 to 10 means the rest area's layout is good, there are enough spaces, and it provides a comfortable resting environment. The comparison of vegetation coverage and species based on before and after optimization is shown in Figure 9.

In Figure 9, the PCA-SSA-BP algorithm shows significant advantages on the vegetation coverage rate and vegetation species trees. Figure 9a shows the comparison of vegetation coverage before and after optimization. The results show that the vegetation coverage rates of the PCA-SSA-BP algorithm are 54.7% and 49.8%, respectively, with improvement rates of 14.5%, respectively. The values are significantly higher than those of the control algorithm. In Figure 9b, the number of vegetation types optimized by the PCA-SSA-BP algorithm reaches 16 and 14 respectively, with an increase of 6 for both. The values are significantly better than those of the comparison algorithms. The study further compares the space utilization efficiency before and after the optimization of the four algorithms. The results are shown in Figure 10.

Figure 10a–b respectively show the comparison of space utilization on the GardenBuildings dataset and the WHU

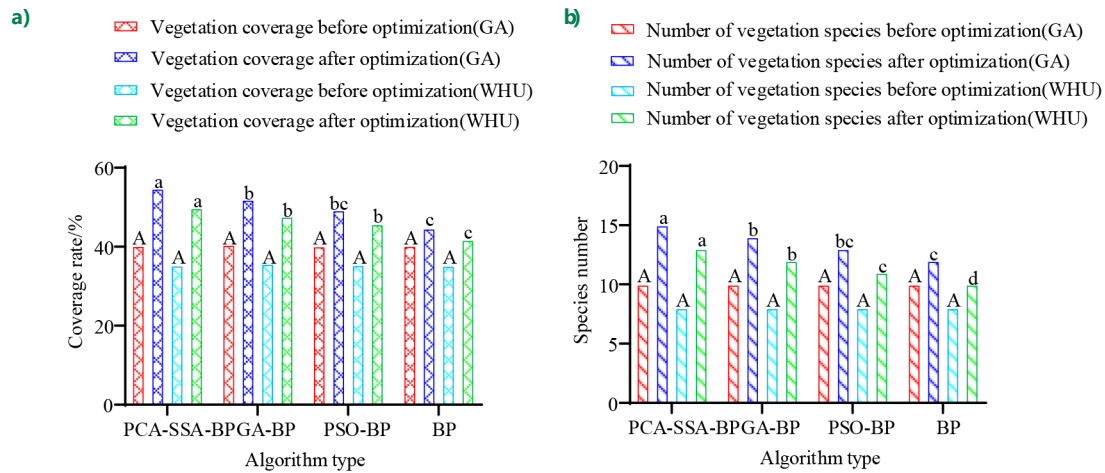


Figure 9. Comparison of vegetation coverage and species before and after optimization: a) Comparison of vegetation coverage before and after optimization; b) Comparison of vegetation species before and after optimization

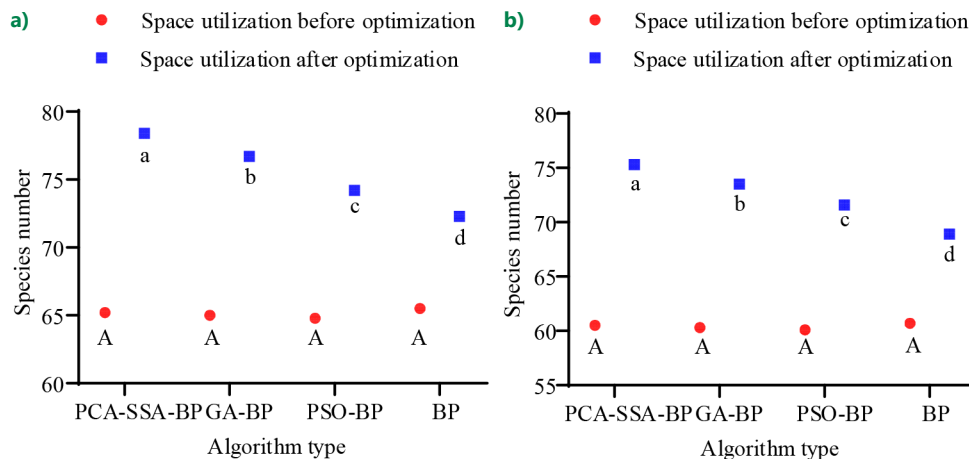


Figure 10. Comparison of space utilization before and after optimization: a) GardenBuildings data set space utilization on comparison; b) WHU Building Dataset data set space utilization comparison

Building dataset. In the result, PCA-SSA-BP has the best optimization effect on space utilization efficiency. The PCA-SSA-BP algorithm demonstrates the highest improvement in space utilization on both datasets. The optimized space utilization rates reaches 78.4% and 75.3%, respectively, and the improvement rates are 13.2% and 14.8%, respectively. The values are significantly better than those of the control algorithm. It indicates that the PCA-SSA-BP algorithm

has significant advantages in optimizing space utilization, which can effectively reduce redundant space and improve space utilization rate. The space utilization improvement of traditional BP on the two datasets is the smallest, and the optimized space utilization rates are 72.3% and 68.9%, respectively. The study concludes by comparing the reasonableness of rest area layout before and after optimization. Table 4 displays the findings.

Table 4. Comparison of path accessibility and rationality of rest area layout

Algorithm name	Data set	Path reachability score before optimization	Path reachability score after optimization	Optimize the layout score of the front lounge area	Optimized rest area layout score
PCA-SSA-BP	GardenBuildings	6.2	8.5	6.5	8.8
	WHU Building Dataset	5.8	8.2	6.0	8.5
GA-BP	GardenBuildings	6.1	8.1	6.4	8.4
	WHU Building Dataset	5.7	7.9	6.1	8.2
PSO-BP	GardenBuildings	6	7.7	6.3	8.0
	WHU Building Dataset	5.6	7.5	6.2	7.8
BP	GardenBuildings	6.3	7.3	6.6	7.6
	WHU Building Dataset	5.9	7.1	6.3	7.4

In Table 4, the PCA-SSA-BP algorithm shows the highest improvement in path accessibility and the rationality of rest area layout. The improvement amplitudes of the optimized path accessibility scores are 2.3 and 2.4, respectively. The increase rates of the layout score of the rest area are 2.3 and 2.5, respectively, indicating that the PCA-SSA-BP algorithm has significant advantages in improving the tourist experience. In summary, the PCA-SSA-BP algorithm shows significant superiority in enhancing the visitor experience. It can effectively optimize path accessibility and rest area layout, thus enhancing the overall visitor experience.

4. Discussion

The study initially used PCA for data dimensionality reduction in order to increase the efficacy and efficiency of landscape architectural design solutions. Subsequently, after optimizing the initial weights and bias of the BPNN based on the SSA, this was used to model and predict the landscape design results. In the research results, the PCA-SSA-BP algorithm performed well in terms of MSE and MAE, and it outperformed the traditional BPNN, GA-BP, and PSO-BP algorithms. This result was basically consistent with the findings of Cui and Wang (2025). The SSA combined with the self-organizing mapping algorithm was used to enhance the interactivity and ecological beauty in the design of park landscape environments. The results demonstrated that the MSE of the algorithm was only 0.01 and its accuracy was as high as 98.1% (Cui & Wang, 2025). In terms of design scheme optimization, the PCA-SSA-BP algorithm significantly improved vegetation coverage and species diversity. Currently Guo et al. (2024) constructed a potential disaster identification model for reservoir scenic spots based on 11 evaluation factors. The research results indicated that the neural network constructed by the SSA-BP algorithm was not as effective as the improved PSO-BP algorithm in recognizing potential hazards, although it had higher accuracy. The possible reason for this was caused by the limited adaptability of the SSA-BP algorithm in the case of certain complex data features and complex sample size (Guo et al., 2024). Yang et al. (2025) proposed a model based on recurrent neural networks and long short-term memory networks in view of the current situation of rural landscape planning and lithology identification in China. The results demonstrated that applying ecological principles to urban landscape design not only protected the natural environment, but also promoted community participation, social interaction, and environmental harmony. Leveraging neural network technology could enhance urban landscape design by providing data-driven insights. This allowed for more accurate and efficient identification of landscape features and prediction of ecological patterns, contributing to better urban planning and development. Consistent with the research

results, the accuracy rate of this method remained stable at over 90% after convergence (Yang et al., 2025). In summary, the PCA-SSA-BP algorithm provides an efficient and effective method for optimizing landscape architecture design. Its ability to process complex data, improve design solutions and enhance visitor experience makes it a powerful tool for landscape architects and designers. In addition, PCA-SSA-BP demonstrates good adaptability to high-dimensional nonlinear data. Specific PCA dimension reduction lowers the computational complexity of the model while retaining important information and reducing the data dimension. SSA optimizes the BPNN's weights and biases, enhancing its ability to model nonlinear relationships.

Although the research has achieved relatively good experimental results, it still has certain limitations. The study only utilizes two datasets, which mainly covers dense horticultural structures and landscapes dominated by natural landscapes. However, these datasets fail to cover landscape design scenarios of different climate zones or project scales. Future research will consider a broader dataset to enhance the universality of the research results. This dataset will cover different climate zones, project scales, and landscape types. For instance, landscape data can be collected and analyzed from various climate zones, such as tropical rainforests, deserts, and grasslands. Additionally, data can be collected from different project scales, including small community parks, large urban parks, and national parks.

5. Conclusions

In response to the demand for multi-dimensional data optimization in landscape architectural design, a hybrid algorithm of PCA-SSA-BP is studied and proposed. The results show that this algorithm can effectively reduce the data dimension and retain key information. Meanwhile, the algorithm performs outstandingly in enhancing ecological value, improving user experience and increasing space utilization rate, demonstrating its effectiveness and practicality as an optimization tool for landscape architectural design.

Funding

This research did not receive any specific funding.

Conflicts of interest

The authors declare no conflict of interest.

Data availability

The data supporting the findings of this study are available within the article.

References

- Ahmadi, M., Sharifi, A., Jafarian Fard, M., & Soleimani, N. (2023). Detection of brain lesion location in MRI images using convolutional neural network and robust PCA. *International Journal of Neuroscience*, 133(1), 55–66. <https://doi.org/10.1080/00207454.2021.1883602>
- Ardhianto, P., Santosa, Y. P., Moniaga, C., Utami, M. P., Dewi, C., Christanto, H. J., & Chen, A. P. S. (2023). Generative deep learning for visual animation in landscapes design. *Scientific Programming*, Article 9443704. <https://doi.org/10.1155/2023/9443704>
- Chen, J., & Stouffs, R. (2023). Deciphering the noisy landscape: Architectural conceptual design space interpretation using disentangled representation learning. *Computer-Aided Civil and Infrastructure Engineering*, 38(5), 601–620. <https://doi.org/10.1111/mice.12908>
- Chen, X. (2023). Environmental landscape design and planning system based on computer vision and deep learning. *Journal of Intelligent Systems*, 32(1), Article 20220092. <https://doi.org/10.1515/jisys-2022-0092>
- Chu, J., Wang, Z., Bao, X., Yao, Z., & Cui, X. (2026). Addressing the contradiction between water supply and demand: A study on multi-objective regional water resources optimization allocation. *Environment, Development and Sustainability*, 28, 11289–11317. <https://doi.org/10.1007/s10668-024-05214-z>
- Cui, B., & Wang, S. (2025). Interactive park landscape environment design by incorporating improved SOM algorithm. *Journal of Advanced Computational Intelligence and Intelligent Informatics*, 29(2), 241–255. <https://doi.org/10.20965/jaciii.2025.p0241>
- Fernberg, P., & Chamberlain, B. (2023). Artificial intelligence in landscape architecture: A literature review. *Landscape Journal*, 42(1), 13–35. <https://doi.org/10.3368/lj.42.1.13>
- Freschlin, C. R., Fahlberg, S. A., Heinzelman, P., & Romero, P. A. (2024). Neural network extrapolation to distant regions of the protein fitness landscape. *Nature Communications*, 15, Article 6405. <https://doi.org/10.1038/s41467-024-50712-3>
- Gao, Y., Wang, N., & Ma, Y. (2023). L2-SSA-LSTM prediction model of steering drilling wellbore trajectory. *IEEE Access*, 12(4), 450–461. <https://doi.org/10.1109/ACCESS.2023.3347611>
- Goodarzi, P., Ansari, M., Rahimian, F. P., Mahdavinnejad, M., & Park, C. (2023). Incorporating sparse model machine learning in designing cultural heritage landscapes. *Automation in Construction*, 155, Article 105058. <https://doi.org/10.1016/j.autcon.2023.105058>
- Groumpos, P. P. (2023). A critical historic overview of artificial intelligence: Issues, challenges, opportunities, and threats. *Artificial Intelligence and Applications*, 1(4), 181–197. <https://doi.org/10.47852/bonviewAIA3202689>
- Guo, J., Xi, W., Yang, Z., Shi, Z., Huang, G., Yang, Z., & Yang, D. (2024). Landslide hazard susceptibility evaluation based on SBAS-InSAR technology and SSA-BP neural network algorithm: A case study of Baihetan Reservoir Area. *Journal of Mountain Science*, 21, 952–972. <https://doi.org/10.1007/s11629-023-8083-9>
- Hong, S., Shen, J., Lü, G., Liu, X., Mao, Y., Sun, N., & Tang, L. (2023). Aesthetic style transferring method based on deep neural network between Chinese landscape painting and classical private garden's virtual scenario. *International Journal of Digital Earth*, 16(1), 1491–1509. <https://doi.org/10.1080/17538947.2023.2202422>
- Li, M., Gu, Y., Ge, S., Zhang, Y., Mou, C., Zhu, H., & Wei, G. (2023). Classification and identification of mixed gases based on the combination of semiconductor sensor array with SSA-BP neural network. *Measurement Science and Technology*, 34, Article 085110. <https://doi.org/10.1088/1361-6501/acd0cb>
- Liu, Y., Fan, L., & Wang, L. (2024). Urban virtual environment landscape design and system based on PSO-BP neural network. *Scientific Reports*, 14, Article 13747. <https://doi.org/10.1038/s41598-024-64296-x>
- Liu, Z., Yu, H., & Jin, W. (2023). Adaptive leakage protection for low-voltage distribution systems based on SSA-BP neural network. *Applied Sciences*, 13(16), Article 9273. <https://doi.org/10.3390/app13169273>
- Wang, J., Zeng, G., Xu, M., Wan, X., Wang, K., Mou, J., Hua, C., Fan, C., & Han, P. (2024). SSA-BP neural network model for predicting rice-fish production in China. *Journal of Applied Ichthyology*, Article 5739961. <https://doi.org/10.1155/2024/5739961>
- Wu, S., Cao, L., Xu D., & Zhao, C. (2023). Historical eco-environmental quality mapping in China with multi-source data fusion. *Applied Sciences*, 13(14), Article 8051. <https://doi.org/10.3390/app13148051>
- Yang, Z., Li, B., Wu, H., Li, M. H., Fan, J., Chen, M., & Long, J. (2023). Water consumption prediction and influencing factor analysis based on PCA-BP neural network in karst regions: A case study of Guizhou Province. *Environmental Science and Pollution Research*, 30, 33504–33515. <https://doi.org/10.1007/s11356-022-24604-2>
- Yang, M., Zhuang, J., & Wang, M. (2025). Leveraging recurrent neural networks for lithology identification and Chinese rural landscape planning in sustainable design. *Sustainability*, 17(7), Article 3078. <https://doi.org/10.3390/su17073078>
- You, K., Qiu, G., & Gu, Y. (2023). An efficient lightweight neural network using BiLSTM-SCN-CBAM with PCA-ICEEMDAN for diagnosing rolling bearing faults. *Measurement Science and Technology*, 34(9), Article 094001. <https://doi.org/10.1088/1361-6501/acd5ef>
- Zhang, K., Chen, Z., Yang, L., & Liang, Y. (2023). Principal component analysis (PCA) based sparrow search algorithm (SSA) for optimal learning vector quantized (LVQ) neural network for mechanical fault diagnosis of high voltage circuit breakers. *Energy Reports*, 9(S1), 954–962. <https://doi.org/10.1016/j.egyr.2022.11.118>
- Zheng, Y., Li, L., Qian, L., Cheng, B., Hou, W., & Zhuang, Y. (2023). Sine-SSA-BP ship trajectory prediction based on chaotic mapping improved sparrow search algorithm. *Sensors*, 23(2), Article 704. <https://doi.org/10.3390/s23020704>