sustainability

MDPI

Article

Classification of Urban Green Space Types Using Machine Learning Optimized by Marine Predators Algorithm

Jiayu Yan, Huiping Liu *, Shangyuan Yu, Xiaowen Zong and Yao Shan

Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China
* Correspondence: hpliu@bnu.edu.cn

Abstract: The accuracy of machine learning models is affected by hyperparameters when classifying different types of urban green spaces. To investigate the impact of hyperparametric algorithms on model optimization, this study used the Marine Predators Algorithm (MPA) to optimize three models: K-Nearest Neighbor (KNN), Support Vector Machines (SVM), and Random Forest (RF). The feasibility of the algorithm was illustrated by extracting and analyzing park green space and attached green spaces within the fifth-ring road of Beijing. A dataset of urban green space type labels was constructed using SPOT6. Three optimized models, MPA-KNN, MPA-SVM and MPA-RF, were constructed. The optimum hyperparameter combination was chosen based on the accuracy of the validation set, and the three optimized models were compared in terms of the Area Under Curve (AUC) value, accuracy on the test set, and other indicators. The results showed that applying MPA improves the accuracy of the validation set of the KNN, SVM, and RF models by 4.2%, 2.2%, and 1.2%, respectively. The MPA-RF model had an AUC value of 0.983 and a test set accuracy of 89.93%, indicating that it was the most accurate of the three models.

Keywords: urban green space; model optimization; Marine Predators Algorithm; machine learning

check for updates

Citation: Yan, J.; Liu, H.; Yu, S.; Zong, X.; Shan, Y. Classification of Urban Green Space Types Using Machine Learning Optimized by Marine Predators Algorithm. *Sustainability* **2023**, *15*, 5634. https://doi.org/10.3390/su15075634

Academic Editors: Xiaoguang Zhou, Yijiang Zhao and Fei Chen

Received: 31 January 2023
Revised: 11 March 2023
Accepted: 21 March 2023
Published: 23 March 2023

1. Introduction

As living standards rise and people become more concerned with their quality of life, there is a growing need for green space [1–4]. Green spaces may serve different functions. For instance, parks primarily offer recreational activities, and the attached green spaces play a role in improving and beautifying artificial building environments. The identification of different urban greenspaces is crucial for improving the quality of life for urban residents, providing knowledge and reference for sustainable urban development.

Many studies on the detection of urban green spaces have been conducted recently, but few have classified urban green space types. Machine learning methods are popular in research on urban green space detection via remote sensing images, owing to their simplicity and high classification accuracy [5]. Traditional supervised machine learning techniques, such as maximum likelihood estimation, K-nearest neighbor (KNN), artificial neural network, support vector machine (SVM), and random forest (RF), are frequently used in remote sensing image classification [6–8]. To distinguish between green and non-green spaces, Zhang et al. [9] introduced a multi-scale guided filtering feature (AMGF), by optimizing multi-scale features using kernel principal component analysis (KPCA), and then used an SVM classifier to detect green and non-green spaces. This method highlights the characteristics of ground objects and significantly improves texture information in an image. Degerickx et al. [10] used data from lidar remote sensing, hyperspectral imaging, and high-spatial-resolution imaging to extract different types of urban green functions. Thanh Noi et al. [11] used three different machine learning models to classify six land cover types in Sentinel-2 images, and found that the SVM classifier was superior to RF and KNN in terms of overall accuracy and had a lower sensitivity to training sample size. Although different classifiers have different overall classification accuracies, they all may be effective in extracting certain land-cover types [12].

The accuracy of classification in a study using a machine learning model to extract urban green space types is not only related to the model structure and selection of image features, but is also affected by the hyperparameters. The model parameters in the machine-learning model are learned and estimated from the data. However, the tuning parameters, or hyperparameters, are defined. To improve the performance and effect of the model, it is necessary to artificially set hyperparameters, or rely on an optimization process to produce a set of optimal hyperparameters. In recent years, automatic machine learning frameworks have gradually emerged that include automatic hyperparameter tuning [13,14]. Currently, there are three main hyperparameter optimization methods: batch sampling, Bayesian optimization, and metaheuristic algorithms. Batch sampling is simple to operate, and obtains a local optimal solution, but it is difficult to determine the preset number of hyperparameter combinations [15]. As a Sequential Model-Based Global Optimization algorithm, Bayesian optimization is a commonly used hyperparameter tuning method, but it has some limitations in high-dimensional hyperparameter space and parallelization [14]. The metaheuristic algorithm is a heuristic search method for global optimization in the machine-learning field, which is based on intuitive or empirical construction. It can provide a feasible solution to issues in high-dimensional hyperparameter space and parallelization at an acceptable cost in terms of computing time and space. Commonly used metaheuristic algorithms include particle swarm optimization, simulated annealing (SA), and the artificial bee colony algorithm [16]. Previous studies have proven the effectiveness of metaheuristic algorithms in optimizing machine learning models and solving complex tasks in various fields [17–19].

Xu et al. [20] optimized SVM parameters using adaptive mutation particle swarm optimization and applied this to construct a landslide displacement prediction coupling model. Ding et al. [21] proposed a new housing price evaluation algorithm that combines the random forest algorithm and SA to optimize the RF hyperparameter (n_estimators, max_features, min_samples_split, min_samples_leaf, max_leaf_nodes, max_depth). Compared to models optimized using the grid search and random search algorithm, the model optimized using the SA had improved running time and accuracy.

The Marine Predators Algorithm (MPA) is a new metaheuristic algorithm based on the foraging strategies of marine predators. Compared with batch sampling and Bayesian optimization, MPA can quickly converge to the global optimal solution and is ideally suited to optimization challenges involving multidimensional variable spaces. Faramarzi et al. [22] used mathematical test functions, engineering benchmarks, and engineering design challenges to evaluate and compare MPA with other optimization techniques. According to the results, MPA is a high-performance optimizer with significantly superior algorithms compared to GA, PSO, GSA, CS, SSA, and CMA-ES, whereas its performance is statistically similar to that of SHADE and LSHADE-cnEpSin. Peng [23] increased the semantic segmentation accuracy of COVID-19 images by optimizing a PSPNet network based on an improved MPA. This experiment shows that a model with optimized hyperparameters, using the modified MPA, outperforms the model without hyperparameter optimization. Moreover, it is substantially faster than manually trying all hyperparameter combinations. Hoang et al. [24] used Sentinel 2 satellite remote sensing data with a spatial resolution of 10 m to detect urban green and non-green spaces using a combination model of MPA and SVM, and obtained a classification accuracy of approximately 93%. The result shows the effectiveness of combining the MPA metaheuristic algorithm and the machine learning model to classify various types of urban green spaces based on remote sensing data.

In this study, the optimal hyperparameter combination of the MPA and machine learning models was used to classify the types of urban green spaces within the fifth-ring road of Beijing. MPA was used to optimize the hyperparameters. The optimal combinations of hyperparameters for the KNN, SVM, and RF models were obtained by comparing the evaluation indices of the models. This study aimed to construct a model that improves the efficiency and accuracy of classifying types of urban green space.

2. Materials

2.1. Study Area

The study area encompasses that within the fifth-ring road of Beijing (39.8° N–40.0° N, 116.2° E–116.5° E), covering an area of 667 km^2, which is located in the plain area of Beijing (Figure 1). Beijing places a high value on the development of green spaces, emphasizing ecological responsibility and people-oriented design. Beijing actively supports the creation of pocket parks and small green spaces, thereby increasing the public's sense of happiness, promoting greening, beautifying towns and villages, and increasing the total amount of greening resources. This is exemplified by the mantra "green for the people" used in Beijing.

Figure 1. The location of the study area in Beijing.

2.2. Data Sets

SPOT6 images from September 2019 were selected, with a spatial resolution of 6 m. Radiometric calibration, atmospheric correction, and geometric correction were completed using ENVI 5.3. The geometric correction accuracy was controlled within one pixel. The initial images contained four bands: blue, green, red, and near infrared. OpenStreetMap network data were used to extract the fifth-ring road boundary, and the SPOT6 images were mosaicked and clipped.

3. Methods

The classification process of urban green space types in this study (Figure 2) primarily included the following five procedures: (1) remote sensing image preprocessing and construction of feature space, (2) construction of the sample dataset for urban green space types, (3) hyperparameter optimization of the model, (4) evaluation and comparison of the three models, and (5) generation of a map of urban green space type based on the classification results. The main operating software used were ArcGIS10.6 and ENVI5.3. The programming language was Python 3.

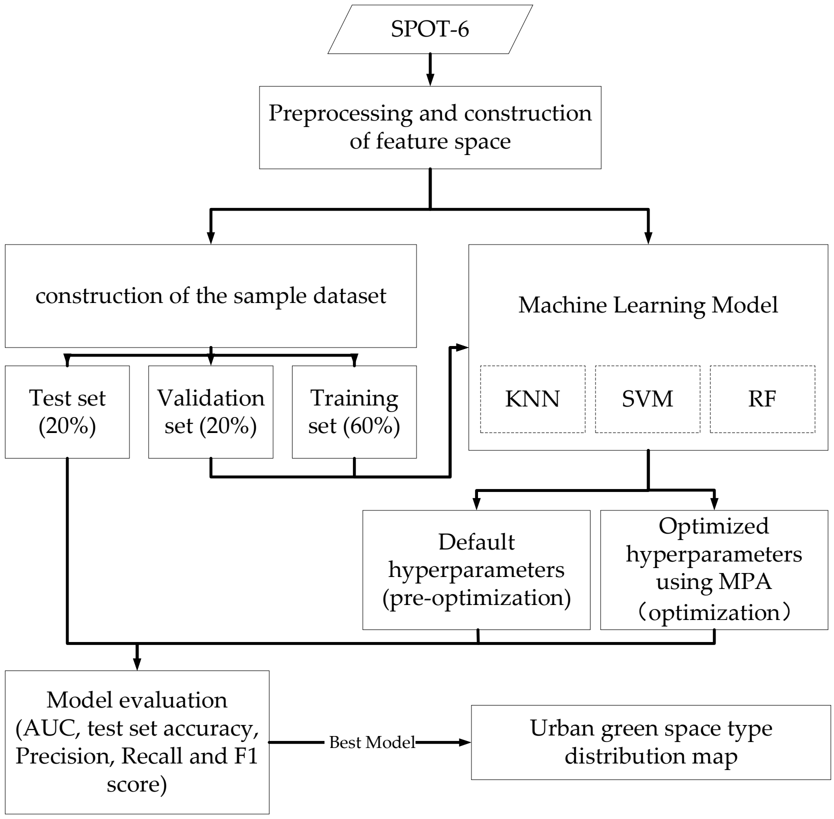

Figure 2. Overall workflow of this study.

3.1. Construction of the Feature Space and Sample Dataset

The selection of feature space is important when using remote sensing images for classification. Vegetation indices are widely used to extract biophysical information on vegetation from satellite images. Previous studies have fully verified the validity of vegetation indices in urban green space mapping based on remote sensing data [25,26]. In this study, the normalized difference vegetation index (*NDVI*) was selected as a feature space. The equation is as follows:

$$NDVI = \frac{NIR - RED}{NIR + RED}, \tag{1}$$

where *NIR* and *RED* represent the near-infrared radiant flux and the red-reflected radiant flux of SPOT6, respectively. The *NDVI* range is (-1,1), where vegetation is generally represented by values greater than 0.

SPOT6 images have a high spatial resolution and rich texture information. To fully explore the features of park green spaces and attached green spaces in the images, four texture indices, homogeneity, contrast, entropy, and correlation, were selected as the features for classification. The selection of these four texture indexes considered the texture features of the two types (park green spaces and attached green spaces) and related literature [27–29]. The five bands of blue, green, red, NIR, and NDVI were normalized before calculating the textures to ensure that all the feature spaces were dimensionless. The four texture indices under these five bands were then calculated separately. The equations and meanings of the texture indices are shown in Figure 3.

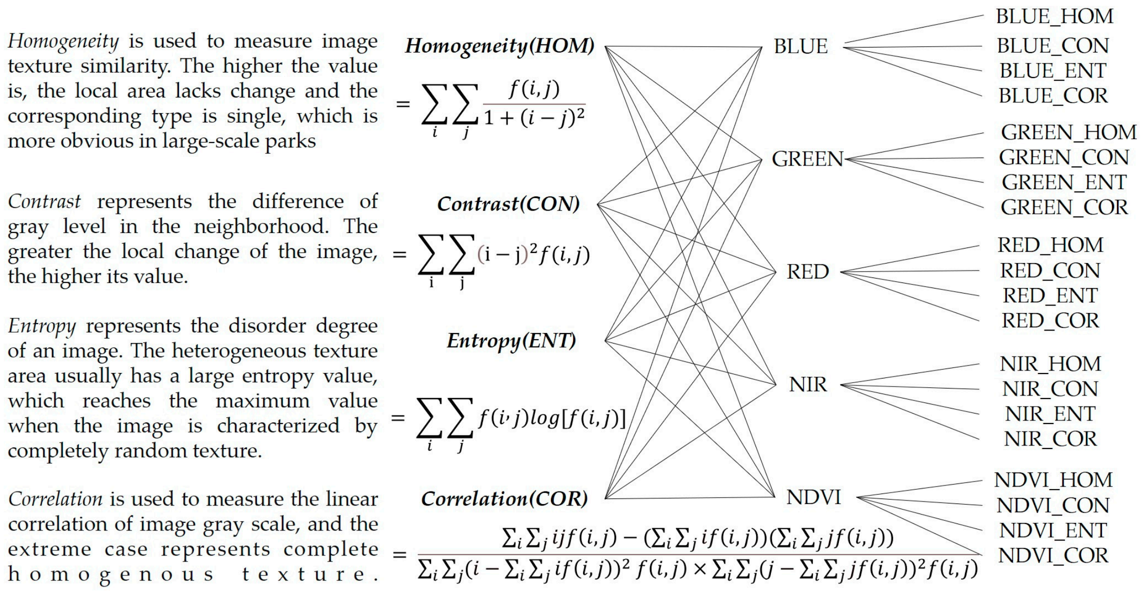

Figure 3. Equations and meanings of the texture indices. In the equation, $f(i,j)$ represents the frequency of the pixel pair under a specific position relation, and $(i-j)$ is the gray difference of the pixel pair.

A total of 25 feature spaces containing the BLUE, GREEN, RED, NIR, and NDVI bands, and the four texture indices corresponding to these five bands, were constructed.

Machine learning requires sufficient sample data for the target category for training and learning; therefore, it is necessary to create an urban green space type label dataset manually. Three categories of urban green space type labels were created: park green space, attached green space, and other land, with reference to the standards for the classification of urban green space [30]. In this study, park green space refers to green space of a suitably large scale (more than 1 ha), and green space covering at least 65% of the land. This corresponds to comprehensive parks, community parks, and specific parks, in the code for the classification of urban land use and planning standards for development land [31]. Attached green space refers to green space for environmental protection, and green spaces attached to urban construction areas, of small scale or irregular shapes. This corresponds to protective green areas, attached green areas, pleasure gardens, etc., in the Standard for the Classification of Urban Green Space. A sample set was constructed based on this dataset. To ensure the computational efficiency of the model, 5000 pixels were selected from park green spaces, attached green spaces, and other land as the classification sample set, respectively. The sample set was divided into training, validation, and test sets at a ratio of 6:2:2.

3.2. Machine Learning

In this study, three popular machine learning models, KNN, SVM, and RF, were chosen for analysis and comparison, both before and after optimization. KNN is a classification algorithm proposed by Cover and Hart in 1968. The fundamental mechanism is to determine a measure of proximity, and data samples with a higher similarity are considered to be in the same category. The measure of proximity is indicated by the Euclidean distance, similarity measure of binary data, and cosine similarity. The basic principle behind SVM [32] is to discover the separation hyperplane (hypersurface), maximize the distance from the nearest point to the hyperplane, and achieve multiple classifications of linearly indivisible data using a penalty coefficient and kernel function. This method was originally applied to binary classification, and multiclassification problems were generally solved using pairs of classifications. RF is a classifier that uses forests established by multiple decision trees to train and predict the samples [33]. The use of multi-tree training voting can effectively solve the “overfitting” phenomenon caused by the low anti-interference ability of a single tree. Typically, the larger the number of decision trees, the more robust the algorithm and the greater the classification accuracy. However, the classification accuracy tends to be stable after a certain number of trees. This method has high classification accuracy and a wide application range; therefore, it is a research hotspot in the field of remote sensing classification.

In this study, Python was used to construct three machine learning models with the help of scikit-learn. The models, constructed according to default hyperparameters, were recorded as pre-optimization models.

3.3. Marine Predators Algorithm

MPA [22] was proposed in 2020. It is a heuristic global optimization algorithm that imitates the natural rules for optimizing the foraging and encounter rate strategies between predators and prey in marine ecosystems. Brownian motion and Lévy flight are two of the best foraging strategies for marine predators [34], simulating natural laws of the ocean, namely, the optimal foraging and encounters rate strategies, the behavior change caused by the environment, and memory strategy, to construct a corresponding model. This model was used to propose the MPA. The procedure was divided into four parts.

(1) Initialization of the *Elite* matrix and *Prey* matrix.

A prey matrix (*Prey*) was constructed according to n (the number of search agents) and d (the dimensions of the solution of the problem to be solved). The initialization method for each element in *Prey* was as follows:

$$Prey = \begin{bmatrix} X_{1,1} & X_{1,2} & \cdots & X_{1,d} \\ X_{2,1} & X_{2,2} & \cdots & X_{2,d} \\ X_{3,1} & X_{3,2} & \cdots & X_{3,d} \\ \vdots & \vdots & \ddots & \vdots \\ X_{n,1} & X_{n,2} & \cdots & X_{n,d} \end{bmatrix}, \quad X_{i,j} = X_{min} + rand(X_{max} - X_{min}) \tag{2}$$

where X_{min} and X_{max} are the lower and upper bounds for the variables, respectively, and *rand* refers to a uniform random vector within (0,1). In practical applications, the initial value of an element can be adjusted based on actual data.

According to the survival of the fittest theory, the fitness solution was calculated by the objective function and nominated as a top predator to construct a matrix called *Elite*.

$$Elite = \begin{bmatrix} X^I_{1,1} & X^I_{1,2} & \cdots & X^I_{1,d} \\ X^I_{2,1} & X^I_{2,2} & \cdots & X^I_{2,d} \\ \vdots & \vdots & \ddots & \vdots \\ X^I_{n,1} & X^I_{n,2} & \cdots & X^I_{n,d} \end{bmatrix}, \tag{3}$$

where X^I is the top predator vector, which is copied n times to construct the *Elite* matrix. *Elite* was updated at the end of each iteration if the top predator was replaced by a better predator (a better solution).

(2) Optimization scenarios

The optimization process is the key to this algorithm, as it simulates the optimal foraging and encounter rate strategies between the marine predator and prey. According to the velocity ratio of the prey and predator (V), the trajectory types of the prey and predator under the corresponding circumstances were simulated, which can be divided into three main phases, where a specific iteration period was set for each phase. These phases are as follows:

In the first phase, the prey moves faster than the predator or has a high velocity ratio (V $\geq$ 10), which is applied to the early phase of the search iteration (*Max_Iter*) (while

$Iter < \frac{1}{3}Max\_Iter$). In this scenario, the predator does not move, and the prey position equation is updated as follows:

$$\overrightarrow{Prey}_i = \overrightarrow{Prey}_i + P.\vec{R} \otimes \left(\vec{R}_B \otimes \left(\overrightarrow{Elite}_i - \vec{R}_B \otimes \overrightarrow{Prey}_i\right)\right)\ i = 1, 2, \ldots, n. \tag{4}$$

where $\vec{R}_B$ is produced by the Brownian motion vector of normal distribution of random numbers, $\otimes$ is an entry-wise multiplication notation. P = 0.5, where R is a vector of uniformly distributed random numbers in the range (0,1).

The second phase occurs when the prey and predator move at similar speeds or have a velocity ratio of approximately 1, which is applied in the middle phase of the search iteration (*Max_Iter*) (while $\frac{1}{3}Max\_Iter < Iter < \frac{2}{3}Max\_Iter$). In this scenario, both prey and predator are looking for prey; half of the group are involved in exploration and the other half for exploitation. Brownian motion is the predator's best strategy when the prey moves in Lévy flight. One-half of the population adopts Lévy flight as its foraging strategy, and the corresponding update position equation is as follows:

$$\overrightarrow{Prey}_i = \overrightarrow{Prey}_i + P.\vec{R} \otimes \left(\vec{R}_L \otimes \left(\overrightarrow{Elite}_i - \vec{R}_L \otimes \overrightarrow{Prey}_i\right)\right)\ i = 1, 2, \ldots, n/2. \tag{5}$$

where $\vec{R}_L$ is based on the Lévy distribution vector. The second half of the population updates its position according to the Brownian motion of the predator, as follows:

$$\begin{gathered}\overrightarrow{Prey}_i = \overrightarrow{Elite}_i + P.CF \otimes \left(\vec{R}_B \otimes \left(\vec{R}_B \otimes \overrightarrow{Elite}_i - \overrightarrow{Prey}_i\right)\right)\ i = n/2, \ldots, n, \\ CF = \left(1 - \frac{Iter}{Max\_Iter}\right)^{\left(2\frac{Iter}{Max\_Iter}\right)},\end{gathered} \tag{6}$$

where *CF* refers to the adaptive parameter of the predator's moving step size.

The third phase occurs when the predator moves faster than the prey or when the velocity ratio is low (V = 0.1), which is applied to the later phase of the search iteration (*Max_Iter*) (when $Iter > \frac{2}{3}Max\_Iter$). In this scenario, the optimal movement strategy of the predator is Lévy flight, and the updated prey position equation is:

$$\overrightarrow{Prey}_i = \overrightarrow{Elite}_i + P.CF \otimes \left(\vec{R}_L \otimes \left(\vec{R}_L \otimes \overrightarrow{Elite}_i - \overrightarrow{Prey}_i\right)\right)\ i = 1, 2, \ldots, n. \tag{7}$$

(3) Eddy formation and *FADs* effect solution

Environmental factors lead to changes in the behavior of marine predators. The simulation of this process can help the algorithm recognize local optimal solutions in the iteration to achieve greater accuracy. The calculation is as follows.

$$\begin{gathered}\overrightarrow{Prey}_i = \begin{cases}\overrightarrow{Prey}_i + CF\left[\vec{X}_{min} + \vec{R} \otimes (X_{max} - X_{min})\right] \otimes \vec{U} & if\ r \leq FADs \\ \overrightarrow{Prey}_i + [FADs(1-r) + r]\left(\overrightarrow{Prey}_{r1} - \overrightarrow{Prey}_{r2}\right) & if\ r > FADs,\end{cases} \\ \vec{U}_i = \begin{cases}0 & if\ random \leq FADs, \\ 1 & if\ random > FADs\end{cases}\end{gathered} \tag{8}$$

where *FADs* = 0.2 and represents the probability of influencing the optimization process, $\vec{U}$ is a random binary vector, *random* is a random number in (0,1), *r* is a uniform random number within (0, 1), and *r*1 and *r*2 are two random subscript indices, where $1 \leq r1, r2 \leq$ n.

(4) Ocean memory.

Marine predators are able to remember successful foraging sites, and the MPA simulates this process by updating the *Elite* Matrix. The fitness of $Prey_i$ in the *Prey* matrix was calculated according to the objective function. If the fitness is better than the fitness of the corresponding position in the *Elite*, $Elite_i$ is updated. Whether the matrix is updated depends on the comparison between the optimal fitness of the *Elite* matrix in the current iteration and the optimal fitness of the previous iteration. If it meets the requirements, the optimal solution is updated; otherwise, the iteration continues.

This study used MPA to optimize the machine learning algorithm process, as shown in Figure 4. The optimization process followed five steps:

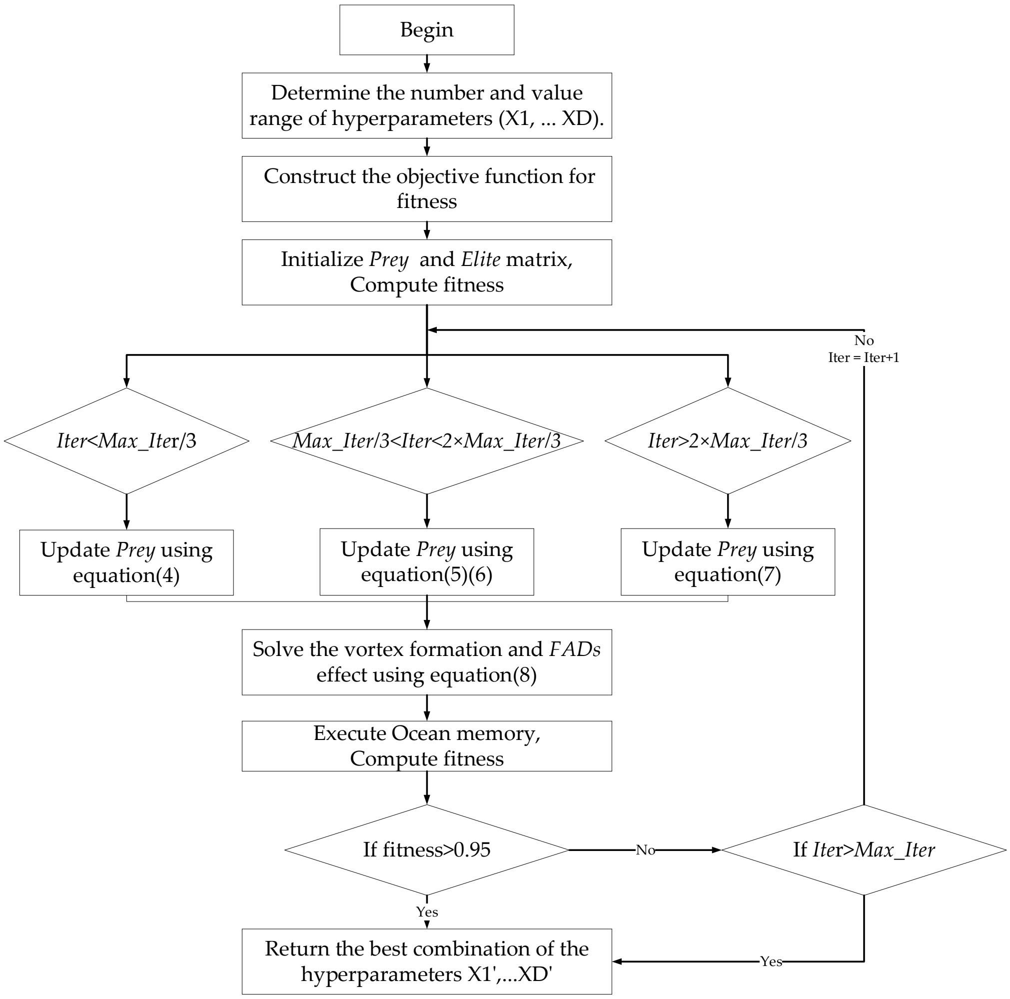

Figure 4. Marine Predators Algorithm (MPA) optimization process for hyperparameters.

1. Determination of the number and value range of hyperparameters (X1, . . . XD).

The number and value range of the hyperparameters were defined according to the ma chine-learning model to be optimized. The number of hyperparameters was dimension d.

2. Construction of the objective function for fitness.

To obtain a classification model with a higher accuracy, the fitness was set as the accuracy of the validation set, the input of fixed parameters was constructed as the training and validation sets, and the independent variable was the hyperparameter to be solved. Finally, the objective function for the accuracy of the validation set was calculated. Higher fitness values corresponded to better hyperparameter combinations.

3. Initialization.

To initialize this study, we set the number of search agents to n = 30 and the maximum number of iterations to 100 to construct the *Prey* matrix with a random combination of hyperparameters. According to the hyperparameter combination $Prey_i$ in the current prey matrix, the machine learning model trained to obtain the accuracy of the validation set, which refers to the fitness in the MPA. X^I, corresponding to the highest fitness value, was copied n times to construct the *Elite* matrix.

4. Completion of optimization.

The main process of the MPA was realized, including optimization scenarios, eddy formation and FADs effect solution, and marine memory

5. Determine whether the current adaptation degree satisfies the requirements and iterations.

In this study, the goal was to reach a fitness value greater than 0.95. If the goal was satisfied, the current hyperparameter combination was successfully optimized, and the program was terminated; if not, the current optimal hyperparameter combination was saved and the number of iterations was determined. If the threshold number of iterations was reached, the program was terminated and the current optimal combination was the output. If the specified number of iterations was not reached, the updated *Prey* and *Elite* were used as inputs, and steps 4–5 were repeated until the fitness requirement was met or the specified number of iterations was reached. The final output was the best combination of the hyperparameters.

3.4. Evaluation Index

1. ROC curve

In the model optimization process, the objective function was set as the accuracy of the validation set, and the optimal solution was the corresponding hyperparameter combination that maximized the accuracy of the validation set. The corresponding optimized models were built based on the optimal hyperparameter combination. The results of the model were compared using the Area Under Curve (AUC) values and test set accuracy. The cumulative distribution functions of the true positive rate (*TPR*) and false positive rate (*FPR*) were used to plot the ROC curve, which is often used to evaluate the overall performance of a model.

$$\begin{aligned} TPR &= \frac{TP}{TP+FN}, \\ FPR &= \frac{FP}{FP+TN}, \end{aligned} \tag{9}$$

where P represents a positive sample, N represents a negative sample, T is a correct prediction and F is a wrong prediction. *TP* and *FN* are a True Positive and False Negative, respectively, both of which indicate that the samples are actually positive. *FP* and *TN* are a False Positive and True Negative, respectively, both of which indicate that the sample is actually negative. The AUC value was obtained by calculating the area under the ROC curve. The AUC value ranges from zero to one, which can be divided into five grades: poor (0.5–0.6), average (0.6–0.7), good (0.7–0.8), very good (0.8–0.9), and excellent (0.9–1) [13].

2. Statistical index

To evaluate the classification performance of different models for various green space types, this study selected three statistical indicators: *Precision*, *Recall and F*1 *score*. *Precision* represents the accuracy of the prediction for positive sample results, *recall* represents the proportion of positive samples predicted correctly, and the *F*1 *score* is the harmonic average of precision and recall. Each index was calculated as follows:

$$\begin{aligned} Precision &= \frac{TP}{TP+FP}, \\ Recall &= \frac{TP}{TP+FN}, \\ F1\ score &= \frac{2\times Precision\times Recall}{Precision+Recal}. \end{aligned} \tag{10}$$

4. Results

4.1. Hyperparameter Optimization Analysis

MPA was used to find the optimal set of hyperparameters for the KNN, SVM, and RF machine-learning models. The major hyperparameters of the three machine-learning models in this study, their value ranges and details are listed in Table 1.

Table 1. The meaning and value range of hyperparameters in the three machine learning models.

Model	Hyperparameter	Meaning	Value Range
KNN	n_neighbors	Number of proximity, when the target point is predicted, several nearby points are taken to predict	[2, 3, 4, . . . , 61]
	p	Distance parameter	[1, 2, 3, . . . , 60]
SVM	gamma	Kernel function coefficient	$[1 \times 10^{-8}, 1 \times 10^{-7}, 1 \times 10^{-6}, \ldots, 1 \times 10^{8}]$
	C	Coefficient of penalty	$[1 \times 10^{-8}, 1 \times 10^{-7}, 1 \times 10^{-6}, \ldots, 1 \times 10^{8}]$
RF	max_depth	Individual decision tree depth	[10, 15, 20, . . . , 105]
	max_features	Maximum number of features to consider for a single decision tree split	[0.05, 0.10, 0.15, . . . , 1.00]
	min_samples_split	Minimum number of samples to split an internal node (not a leaf node)	[2, 6, 10, . . . , 78]
	n_estimators	Number of decision trees	[100, 250, 400, . . . , 2950]

The accuracy of the validation set was set as the objective function, and the different hyperparameter combinations were the independent variables of the objective function. The default and optimal hyperparameter combinations of the three models, after iterative optimization by the MPA, are listed in Table 2. Table 3 shows the accuracy of the validation sets before and after the optimization of the three models.

Table 2. Optimal hyperparameter combinations for the three machine learning models.

Model	Hyperparameters	Default Combinations	Optimal Combinations
KNN	(n_neighbors, p)	(5, 2)	(14, 1)
SVM	(gamma, C)	(0.04, 1)	(0.01, 10,000)
RF	(max_depth, max_features, min_samples_split, n_estimators)	(None, 1, 1, 10)	(25, 0.55, 6, 250)

Table 3. Comparison of validation set accuracy before and after optimization of the three machine learning models.

Model	Validation Set Accuracy before Optimization (%)	Validation Set Accuracy after Optimization (%)	Range of Growth (%)
MPA-KNN	81.5	84.9	4.2
MPA-SVM	88.1	90.0	2.2
MPA-RF	89.9	91.0	1.2

In addition, the RF model was used as an example to analyze the time consumption and accuracy of the model under different combinations of hyperparameters. After sorting the models and combinations of hyperparameters by accuracy, the model training time and accuracy are shown in Figure 5. The graph shows that there is no linear correlation between the time and accuracy, that is, the shortening of the model training time does not necessarily result in lower model accuracy. Thus, a machine learning model with a short time consumption and high accuracy can be obtained by adjusting the combination of hyperparameters.

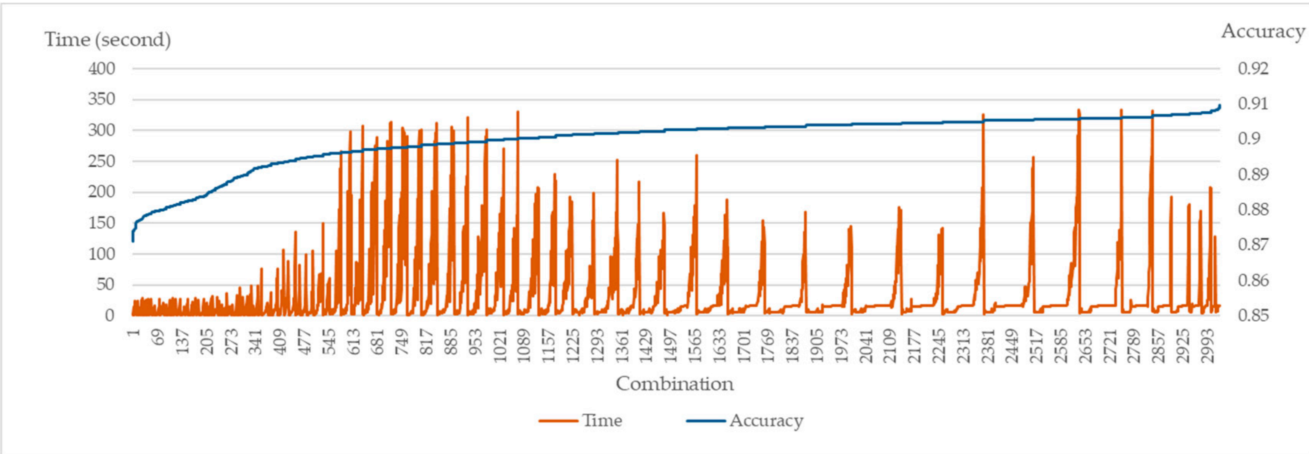

Figure 5. Relationship between running time and accuracy of the RF Model under different hyperparameter combinations. The horizontal axis represents the combination index, and the vertical axis represents the time (seconds) and accuracy, respectively.

4.2. Model Evaluation

Three machine learning models for urban green space type classification (MPA-KNN, MPA-SVM, and MPA-RF) were constructed according to the optimal combination of hyperparameters. The ROC curves are shown in Figure 6. A larger area between the curve and the FPR axis (denoted as AUC) indicates better performance. It can be seen from the figure that the performance of MPA-RF was the highest of the three models. To verify the optimization effect of the MPA on the model, we compared the AUC values of the model, built according to the default hyperparameter combination (before optimization) and the optimal hyperparameter combination (after optimization). As shown in Table 4, the AUC of the model after hyperparameter optimization was larger, which indicates that MPA hyperparameter optimization can improve the performance of the machine learning model.

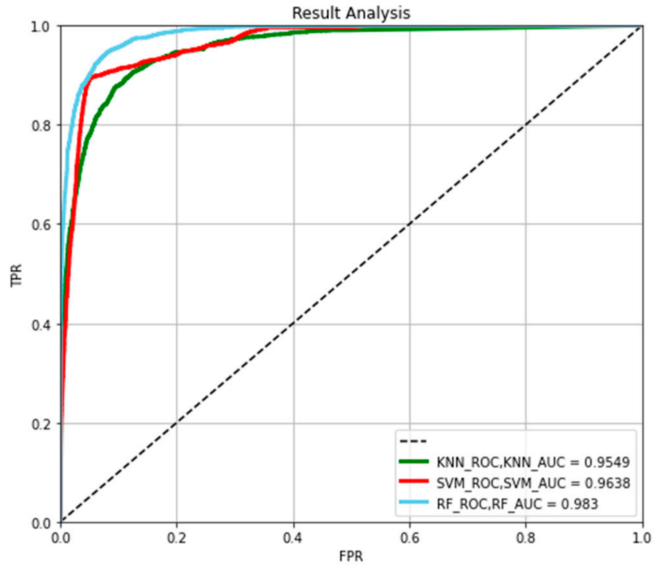

Figure 6. ROC curves and Area Under Curve (AUC) values of the three models after hyperparameter optimization.

Table 4. AUC values before and after optimization of the three machine learning models.

Model	AUC Value before Optimization	AUC Value after Optimization
MPA-KNN	0.938	0.9549
MPA-SVM	0.9598	0.9638
MPA-RF	0.9817	0.983

The models were evaluated by comparing the statistical indices of each category after the classification of the test set, as shown in Table 5. RF showed the best performance in terms of the overall index AUC and test set accuracy. In the classification of park green space, SVM was the best, followed by RF and KNN. For the classification of attached green space and other land, RF indices were slightly higher than those of SVM, and significantly higher than those of KNN. In general, the classification performance of park green space and attached green space was good in the three models after hyperparameter optimization.

Table 5. Statistical indexes of model performance after hyperparameter optimization.

		MPA-KNN	MPA-SVM	MPA-RF
AUC		0.9541	0.9615	0.9828
Test set accuracy		84.53%	89.53%	89.93%
Park green space	PRECISION	0.89	0.91	0.9
	RECALL	0.83	0.84	0.84
	F1	0.86	0.87	0.87
Attached green space	PRECISION	0.75	0.82	0.83
	RECALL	0.84	0.88	0.88
	F1	0.79	0.85	0.85
Other land	PRECISION	0.91	0.96	0.97
	RECALL	0.87	0.97	0.97
	F1	0.89	0.96	0.97

4.3. Distribution of Green Space Types within the Fifth-Ring Road of Beijing

The MPA-RF model was selected to classify the urban green space types within the fifth-ring road of Beijing, as shown in Figure 7. Park and attached green spaces were accurately classified, and some urban pocket parks, that were difficult to extract, were included as park green spaces. As can be seen from the figure, the distribution of green space within the fifth-ring road of Beijing forms a network system with an annular, radial, and dot interweaving pattern, with Dongcheng and Xicheng districts as the center. There are landscape shelter belts around the fourth- and fifth-ring roads, and green channels are formed on both sides of the roads at all levels. The park and attached green spaces are evenly distributed within the fifth-ring road, which better meets the residents' demand for green spaces.

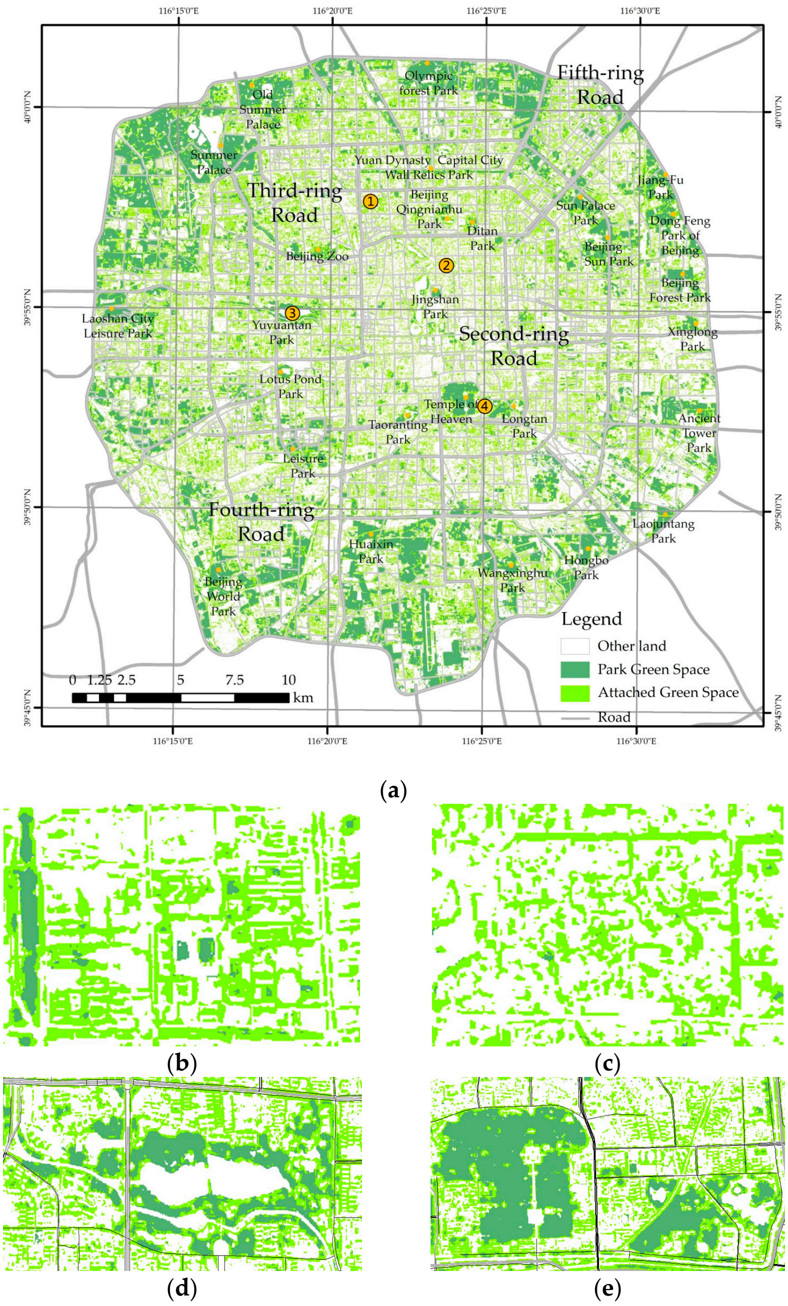

Figure 7. Maps of urban green space distribution in Beijing. (**a**) A map of urban green space distribution within the fifth-ring road of Beijing; (**b**) ① in (**a**), Beijing Normal University, Beijing University of Posts and Telecommunications; (**c**) ② in (**a**), South Luogu Alley; (**d**) ③ in (**a**), Yuyuantan Park; (**e**) ④ in (**a**), Temple of Heaven, Longtan park.

By comparing the data of the Dongcheng and Xicheng districts in the study area with the statistical data, the green space coverage rates of the Dongcheng and Xicheng districts were 37% and 29.8%, respectively, from classification. The Beijing Municipal Bureau of Landscaping data listed these areas as having green space coverage rates of 33.24% and 30.93%, with errors of 3.76% and 1.13%, respectively. The location vector data for parks in the Dongcheng and Xicheng districts of Beijing were created according to the list of parks in Beijing, published by the Beijing Municipal Bureau of Landscape. This was overlaid with the classification results, as shown in Figure 8. The park green space in this study was the same as that in the standard for classification of urban green space [30], except for petty street gardens, which correspond to historic parks, comprehensive parks, specific parks, and community parks in the park directory. The category of petty street gardens in the list is small in scale or diverse in shape, which is similar to the definition of a "pocket park" and belongs to the attached green space in this study.

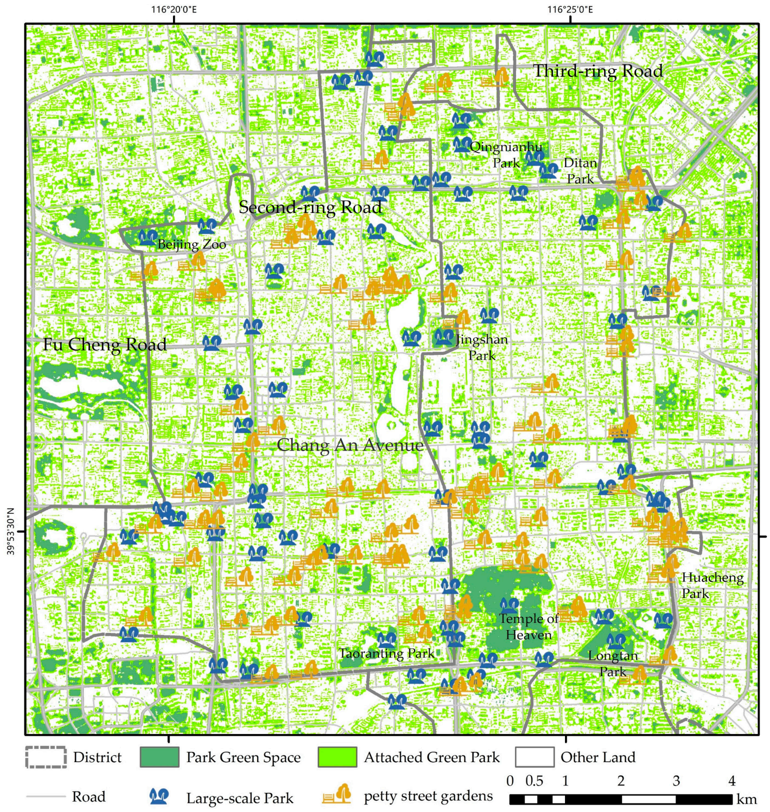

Figure 8. Beijing Dongcheng and Xicheng District Park distribution map.

The parks and their directory and classification types in the Dongcheng and Xicheng districts are listed in Table 6. The overall correct classification rate reached 82.5%. In this study, 80.8% of parks were correctly identified as park green spaces, and the proportions of historic, comprehensive, specific, and community parks in the Dongcheng and Xicheng districts are detailed in Table 6. Among them, the correct classification rate of historic, comprehensive, and specific parks reached 93.9% because of the larger area scales and proportion of greening. Of the petty street gardens, 83.9% were correctly identified as attached green spaces.

Table 6. Park classification type statistics in Dongcheng and Xicheng Districts.

Type of Park List	Type of Green Space in the Study	Identified as Park Green Space (PCS)	Identified as Attached Green Space (PCS)	Total (PCS)	Correct Classification Rate
Historic Park	park green space	9	1	10	90.0%
Comprehensive Park	park green space	14	0	14	100.0%
Specific Park	park green space	8	1	9	88.9%
Community Park	park green space	28	12	40	70.0%
Subtotal		59	9	73	80.8%
Petty Street Garden	attached green space	15	78	93	83.9%
Total				166	82.5%

5. Discussion

5.1. Machine Learning Optimized by MPA

In this study, the MPA optimized machine learning models were proposed, which achieved accurate and efficient classification of urban green space types. The results supported the reliability of MPA for machine learning hyperparameter optimization and the accuracy of the MPA-RF model in the classification of urban green space types. Compared with previous studies [9,24], this method achieved a more accurate and detailed identification of urban green space types by using remote sensing images. Subsequently, an urban green space database can be established for rapid assessment and field investigations, and to provide knowledge and reference for sustainable urban development.

This study used SPOT images with a spatial resolution of 6 m. Experiments on remote sensing images with varying spatial resolutions have not been performed. Low-resolution remote sensing images may be effective for extracting park green space, but due to the mixed pixels, these images may not be accurate for extracting the attached green space. Higher-resolution images contain more information. However, more information often leads to overfitting in classification. Therefore, comparing the classification accuracy of this optimized classification model in different spatial resolution images to determine the best model for extracting different types of urban green spaces and the most appropriate remote sensing data is a worthwhile research topic. The four texture indices of spectral bands and NDVI were calculated separately. Results of urban green space type classification were obtained with higher accuracy under the combined effect of these 25 feature spaces. However, the computational efficiency and possibility of data redundancy were ignored in the pursuit of high accuracy. In the subsequent study, a smaller (but more sensitive) number of feature spaces can be identified for classification by factor analysis and other dimensionality reduction methods to ensure classification accuracy with reduced classification time. This method is applicable to hyperparameter searching for machine learning classification models, but it is important to note that the choice of hyperparameters varies for different sample data. In this study, the selected hyperparameters and the ranges were determined by previous studies [11,13] and experience. For different classification objectives, it is necessary to determine the range of hyperparameters according to reality or pre-experiments.

5.2. Classification of Green Space Types

The classification results for the study area were consistent with the actual results. A high test set accuracy of 92.5% was obtained, using this method to classify SPOT images, in

2017. In Beijing, the scale of the park green space is generally large, whereas the scale of the attached green space is small. The texture difference between the two is large, as seen from the remote sensing image. Therefore, this method can be used to identify the park and attached green space for urban green space. However, in some small cities, the size of the park green space is similar that of the attached green space. It is necessary to verify whether the model has the same high accuracy in these cities.

6. Conclusions

To reduce the problems associated with the use and automation of urban green space type classification methods, MPA was applied to optimize machine learning and improve an existing machine learning model to classify urban greenspace types. The results show that:

1. The machine learning model optimized by the MPA was improved, mainly reflected in the increased accuracy of the validation set. After MPA optimization of the KNN, SVM, and RF, the accuracy of the validation sets increased by 4.2%, 2.2%, and 1.2%, respectively.
2. Compared with the other optimized models, the MPA optimized RF model had higher accuracy and efficiency. The AUC value was 0.9828, and the accuracy of the test set reached 89.93%. By comprehensively comparing the accuracy rate, recall rate, and F1 value of the different types of land use, it was found that MPA-RF had higher stability than the MPA optimized KNN and SVM models.
3. The MPA-RF was effective in extracting park and attached green spaces within the fifth-ring road of Beijing, indicating that it can be used to rapidly evaluate green space types in cities. The accuracy rate of identifying large parks in the Dongcheng and Xicheng districts of Beijing reached 80.8%.

In summary, the accuracy of the machine learning model was improved after optimization using MPA, and displayed excellent classification ability in urban green space type classification. These attributes are important for the accurate identification of urban green space types. In future work, urban green space types can be extracted from remote sensing images for many years to establish an urban green space database. The internal and external factors can be analyzed according to the change in areas in order to provide a basis for sustainable development planning in Beijing. In this study, six-meter resolution SPOT6 images were used to classify urban green space types. The classification results were satisfactory. However, the image data used in urban research is gradually dominated by the use of high-resolution images. In the future, the performance of the optimized classification method in higher resolution images should be compared to evaluate the adaptability of the method in future. In addition, the objective function chosen for hyperparameter optimization in current machine learning models primarily considers the accuracy values. In future studies, models should be obtained by designing an objective function that considers the accuracy and time consumption in order to obtain a more efficient model.

Author Contributions: Conceptualization, J.Y. and H.L.; methodology, J.Y. and H.L.; validation, J.Y., X.Z. and Y.S.; formal analysis, J.Y. and H.L.; investigation, J.Y. and S.Y.; resources, J.Y. and H.L.; writing—original draft preparation, J.Y. and H.L.; writing—review and editing, J.Y. and H.L.; visualization, J.Y.; project administration, H.L.; funding acquisition, H.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Natural Science Foundation of China Major Program, grant number No. 42192580, No. 42192584.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Dai, F.; Yang, C.; Chen, M. Research Progress in Urban Green Spaces in Recent 10 Years in China—Mapping Knowledge Domains Analysis Based on CiteSpace. *J. Chin. Urban For.* **2019**, *17*, 87–92. [CrossRef]
2. Dong, Y.; Liu, H.; Qi, J. Progress of Studies on the Relationship Between Urban Green Space and Public Health. *Urban Plan. Int.* **2020**, *35*, 70–79. [CrossRef]
3. Maas, J.; Verheij, R.A.; Groenewegen, P.P.; de Vries, S.; Spreeuwenberg, P. Green Space, Urbanity, and Health: How Strong Is the Relation? *J. Epidemiol. Commun. Health* **2006**, *60*, 587–592. [CrossRef]
4. Mitchell, R.J.; Richardson, E.A.; Shortt, N.K.; Pearce, J.R. Neighborhood Environments and Socioeconomic Inequalities in Mental Well-Being. *Am. J. Prev. Med.* **2015**, *49*, 80–84. [CrossRef]
5. Gong, J. Chances and Challenges for Development of Surveying and Remote Sensing in the Age of Artificial Intelligence. *Geomat. Inf. Sci. Wuhan Univ.* **2018**, *43*, 1788–1796. [CrossRef]
6. Ahmad, A.; Ahmad, S.R.; Gilani, H.; Tariq, A.; Zhao, N.; Aslam, R.W.; Mumtaz, F. A Synthesis of Spatial Forest Assessment Studies Using Remote Sensing Data and Techniques in Pakistan. *Forests* **2021**, *12*, 1211. [CrossRef]
7. Huang, C.; Yang, J.; Clinton, N.; Yu, L.; Huang, H.; Dronova, I.; Jin, J. Mapping the Maximum Extents of Urban Green Spaces in 1039 Cities Using Dense Satellite Images. *Environ. Res. Lett.* **2021**, *16*, 64072. [CrossRef]
8. Shirmard, H.; Farahbakhsh, E.; Müller, R.D.; Chandra, R. A Review of Machine Learning in Processing Remote Sensing Data for Mineral Exploration. *Remote Sens. Environ.* **2022**, *268*, 112750. [CrossRef]
9. Zhang, T.; Dai, Q.; Xu, W.; Dai, F.; Wang, L. Research on Extraction Method of Urban Green Space from High-Resolution Remote Sensing Image. *J. Southwest For. Univ. Sci.* **2020**, *40*, 105–114. [CrossRef]
10. Degerickx, J.; Hermy, M.; Somers, B. Mapping Functional Urban Green Types Using High Resolution Remote Sensing Data. *Sustainability* **2020**, *12*, 2144. [CrossRef]
11. Thanh Noi, P.; Kappas, M. Comparison of Random Forest, k-Nearest Neighbor, and Support Vector Machine Classifiers for Land Cover Classification Using Sentinel-2 Imagery. *Sensors* **2018**, *18*, 18. [CrossRef]
12. Du, P.; Xia, J.; Zhang, W.; Tan, K.; Liu, Y.; Liu, S. Multiple Classifier System for Remote Sensing Image Classification: A Review. *Sensors* **2012**, *12*, 4764–4792. [CrossRef]
13. Yang, C.; Liu, L.; Zhang, Y.; Zhu, W.; Zhang, S. Machine Learning Based on Landslide Susceptibility Assessment with Bayesian Optimized the Hyperparameters. *Bull. Geol. Sci. Technol.* **2022**, *41*, 228–238. [CrossRef]
14. Zhang, A.; Yang, Z. Hyperparameter Tuning Methods in Automated Machine Learning. *Sci. Sin. Math.* **2020**, *50*, 695–710. [CrossRef]
15. Bergstra, J.; Bengio, Y. Random Search for Hyper-Parameter Optimization. *J. Mach. Learn. Res.* **2012**, *13*, 281–305.
16. Boussaïd, I.; Lepagnot, J.; Siarry, P. A Survey on Optimization Metaheuristics. *Inf. Sci.* **2013**, *237*, 82–117. [CrossRef]
17. Song, H.; Triguero, I.; Özcan, E. A Review on the Self and Dual Interactions between Machine Learning and Optimisation. *Prog. Artif. Intell.* **2019**, *8*, 143–165. [CrossRef]
18. Tien Bui, D.; Hoang, N.D.; Samui, P. Spatial Pattern Analysis and Prediction of Forest Fire Using New Machine Learning Approach of Multivariate Adaptive Regression Splines and Differential Flower Pollination Optimization: A Case Study at Lao Cai Province (Viet Nam). *J. Environ. Manag.* **2019**, *237*, 476–487. [CrossRef]
19. Akinola, O.O.; Ezugwu, A.E.; Agushaka, J.O.; Zitar, R.A.; Abualigah, L. Multiclass Feature Selection with Metaheuristic Optimization Algorithms: A Review. *Neural Comput. Appl.* **2022**, *34*, 19751–19790. [CrossRef]
20. Xu, F.; Fan, C.; Xu, X.; Li, L.; Ni, J. Displacement Prediction of Landslide Based on Variational Mode Decomposition and AMPSO-SVM Coupling Model. *J. Shanghai Jiaotong Univ.* **2018**, *52*, 1388–1395+1416. [CrossRef]
21. Ding, Y.; Cao, H. Housing Prices Evaluation Using Random Forest Algorithm Combing with Simulated Annealing. *Appl. Res. Comput.* **2020**, *37*, 784–788.
22. Faramarzi, A.; Heidarinejad, M.; Mirjalili, S.; Gandomi, A.H. Marine Predators Algorithm: A Nature-Inspired Metaheuristic. *Expert Syst. Appl.* **2020**, *152*, 113377. [CrossRef]
23. Peng, X. Research on Semantic Segmentation of Medical Images Based on Improved Marine Predator Algorithm Optimized PSPNet. Master's Thesis, Northeast Forestry University, Harbin, China, 2021.
24. Hoang, N.-D.; Tran, X.-L. Remote Sensing–Based Urban Green Space Detection Using Marine Predators Algorithm Optimized Machine Learning Approach. *Math. Probl. Eng.* **2021**, *2021*, 1–22. [CrossRef]
25. Seidl, M.; Saifane, M. A Green Intensity Index to Better Assess the Multiple Functions of Urban Vegetation with an Application to Paris Metropolitan Area. *Environ. Dev. Sustain.* **2021**, *23*, 15204–15224. [CrossRef]
26. Chen, B.; Xu, B.; Gong, P. Mapping Essential Urban Land Use Categories (EULUC) Using Geospatial Big Data: Progress, Challenges, and Opportunities. *Big Earth Data* **2021**, *5*, 410–441. [CrossRef]
27. Ulaby, F.T.; Kouyate, F.; Brisco, B.; Williams, T.H.L. Textural Infornation in SAR Images. *IEEE Trans. Geosci. Remote Sens.* **1986**, *GE-24*, 235–245. [CrossRef]
28. Gao, C.; Hui, X. GLCM-Based Texture Feature Extraction. *Comput. Syst. Appl.* **2010**, *19*, 195–197.
29. Huang, X. Multiscale Texture and Shape Feature Extraction and Object-Oriented Classification for Very High Resolution Remotely Sensed Imagery. Ph.D. Thesis, Wuhan University, Wuhan, China, 2009.
30. *CJJ/T85-2017*; Standard for Classification of Urban Green Space. Ministry of Housing and Urban-Rural Development PRC: Beijing, China, 2017.

31. *GB 50137-2011*; Code for Classification of Urban Land Use and Planning Standards of Development Land. Ministry of Housing and Urban-Rural Development PRC: Beijing, China, 2011.
32. Cortes, C.; Vapnik, V. Support-Vector Networks. *Mach. Learn.* **1995**, *20*, 273–297. [CrossRef]
33. Leo, B. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32.
34. Bartumeus, F.; Catalan, J.; Fulco, U.L.; Lyra, M.L.; Viswanathan, G.M. Optimizing the Encounter Rate in Biological Interactions: Lévy versus Brownian Strategies. *Phys. Rev. Lett.* **2002**, *88*, 097901. [CrossRef]