

Research Article

Remote Sensing–Based Urban Green Space Detection Using Marine Predators Algorithm Optimized Machine Learning Approach

Nhat-Duc Hoang ^{1,2} and Xuan-Linh Tran ^{1,2}

¹Institute of Research and Development, Duy Tan University, Da Nang 550000, Vietnam

²Faculty of Civil Engineering, Duy Tan University, Da Nang 550000, Vietnam

Correspondence should be addressed to Nhat-Duc Hoang; hoangnhatduc@duytan.edu.vn

Received 2 March 2021; Revised 9 April 2021; Accepted 14 May 2021; Published 24 May 2021

Academic Editor: Qingzheng XU

Copyright © 2021 Nhat-Duc Hoang and Xuan-Linh Tran. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Information regarding the current status of urban green space is crucial for urban land-use planning and management. This study proposes a remote sensing and data-driven solution for urban green space detection at regional scale via employment of state-of-the-art metaheuristic and machine learning approaches. Remotely sensed data obtained from Sentinel 2 satellite in the study area of Da Nang city (Vietnam) are used to construct and verify an intelligent model that hybridizes Marine Predators Algorithm (MPA) and support vector machines (SVM). SVM are employed to generalize a decision boundary that separates features characterizing statistical measurements of remote sensing data into two categories of “green space” and “nongreen space”. The MPA metaheuristic is used to optimize the SVM training phase by identifying an appropriate set of the SVM’s hyperparameters including the penalty coefficient and the kernel function parameter. Experimental results show that the proposed model which processes information provided by all of the Sentinel 2 satellite’s spectral bands can deliver a better performance than those obtained from the model based on vegetation indices. With a good classification accuracy rate of roughly 93%, an F1 score = 0.93, and an area under the receiver operating characteristic = 0.98, the newly developed model is a promising tool to assist local authority to obtain up-to-date information on urban green space and develop plans of sustainable urban land use.

1. Research Background and Motivation

In many regions around the globe, fast pace of urbanization leads to various problems including traffic congestion, poor air quality, and noise pollution. As pointed out by Xian et al. [1], urbanization significantly transforms the landscape from the natural surface types to impervious surface such as housing, commercial building, and infrastructures. This transformation is happening in a large spatial extent and an increasing speed due to a burgeoning pressure on additional housing and commercial/industrial areas. Moreover, developments of urban lands consume areas of green land areas and bring about negative impacts on the urban environments [2–4]. Worsening living environment caused by lack of green space is a major issue for human health because roughly 54% of people in the world are living in urban areas [5].

Urban green space is generally defined as green infrastructure that contains vegetated spaces including urban parks, road, and workplace green space [6]. Previous works have recognized the crucial role of green space for reducing the adverse impacts of urbanization in both aspects of urban ecosystem and socioeconomics [7–15]. The World Health Organization has identified that green spaces are innovative methods for enhancing the quality of urban environments via improvements of local resilience and promotions of sustainable lifestyles [15]. Turaga et al. [16] state that urban green spaces become a critical asset because they deliver various benefits including aesthetics enhancement, pollution reduction, positive effects on physical and mental health of citizens, urban heat island reduction, and groundwater recharge. Therefore, there is a raising societal support for protection and development of green space in urban areas.

Therefore, up-to-date spatial information regarding the current status of urban green space is crucial for urban land-use planning and management. This information has become increasingly difficult to obtain via conventional landscape surveying approaches since green spaces have been constantly modified, fragmented, and dispersed due to the fast pace of urbanization. Moreover, surveying tasks at a regional scale are daunting because of both time and labor consumptions required for field data acquisition, processing, and report. Thus, there is a pressing need for advanced methods to automate the green space surveying task.

Recently, medium-resolution imagery coupled with advanced machine learning methods has provided effective solution for urban landscape survey [17–22]. Remote sensing data used with geographic information system (GIS) can be used to generate thematic maps to assess green vegetation cover at a regional scale. Data extracted from such map can be helpful for further data analyzing processes regarding the size, shape, and other landscape pattern of urban green space [23–28].

Rafiee et al. [13] relies on Landsat Thematic Mapper images to study the patterns of green areas; this study employs combined techniques of remote sensing image classification, landscape metrics assessment, and vegetation indices. El Garouani et al. [29] employ the maximum likelihood supervised classification to analyze data obtained from Landsat's bands with a spatial resolution of 30 m; the authors investigate the relationship between urbanization and land use changes as well as the effect of the increase in impervious surface areas. Urban green space distribution has been modeled in [7] with the use of remote sensing, GIS technology, and normalized difference vegetation index. Do et al. [3] relies on Landsat 8 OLI (Operational Land Imager) image datasets provided the United States Geological Survey to study green space patterns; the support vector machine (SVM) has been used by the authors for the task of image data classification; the overall accuracy of the proposed machine learning model is 82.70%.

Li et al. [30] construct land-use and land-cover maps including green spaces using Landsat Operational Land Imager (OLI) and Enhanced Thematic Mapper Plus (ETM+) imagery; convolutional neural network, random forest, and SVM are the employed machine learning models used for image data classification; this study reports a classification accuracy of 84.40% on the testing dataset. Dinda et al. [31] construct an integrated model for studying urban growth and associated green space loss; the model relies on maximum likelihood classifier, artificial neural network, and SVM for performing pattern recognition task; the SVM model has attained the most desired classification accuracy and the area under the receiver operating characteristic curve (0.906).

It is noted that besides SVM, deep neural networks (DNNs) have also been successfully applied in remote sensing-based land-use classification [32–34]. DNNs are highly appropriate for image categorization due to its convolution operator based autonomous feature extraction phase [35]. However, successful implementations of DNNs often require a large number of training image samples. The

computational expense of DNNs is generally significant due to the time-consuming training process used to fine-tune the networks' weights. Moreover, because deep learning models have a quiet large number of hyperparameters (e.g., the number of hidden layers, the number and the size of convolution operations, the size of the pooling operations, etc.), the process of identifying a suitable network architecture can be tedious. Moreover, since the extracted features are represented as numerical data in this study, the application of SVM can be highly appropriate. It is because SVM has been proven to be a capable tool for classifying extracted numerical datasets [19, 36–39].

Based on literature review, there is an increasing trend of applying machine learning in remote sensing-based urban green space study. Since the problem of interest is challenging due to the involvement of multivariate and non-linear data analysis, other advanced machine learning solutions need to be investigated to improve the urban green space detection accuracy. Moreover, the current literature also points out that individual machine learning methods are the commonly employed approach. Hybrid machine learning models that harness advantages of various computational intelligence techniques are rarely investigated to construct urban green space detection models. Specifically, previous studies have mainly relied on the individual machine learning approach [3, 13, 29, 31], and the employment of metaheuristic algorithms used for optimizing machine learning based remote sensing data classification has rarely been proposed and investigated. Therefore, the original contribution of the current work is proposing a hybridization of SVM machine learning and metaheuristic optimization used for remote sensing-based urban green space detection.

SVM [40] is considered to be a capable pattern recognizer with excellent generalization capability. It is due to the fact that the model structure of this machine learning method is learnt via the framework of structural risk minimization which is resilient to overfitting and noisy data [41]. Nevertheless, the model construction phase of a SVM model requires a proper setting of its two hyperparameters including the penalty coefficient and the kernel function parameter. The former specifies the amount of penalty imposing on data samples having classification errors. The later determines the locality of the employed kernel function which directly influences the generalization of the constructed model.

The task of determining hyperparameters of a machine learning model is known as model selection [41] and can be modeled as an optimization problem. For the case of a SVM model, this is a challenging task because of several reasons. First, the landscape of the objective function is unknown and not differentiable. Second, the hyperparameters must be searched in continuous space; therefore, there is an indefinite number of feasible solutions. This fact means that an exhaustive search on the hyperparameters is infeasible. Therefore, various scholars have resorted to metaheuristic algorithms for dealing with the model selection problems. The role of metaheuristic algorithms in the task of hyperparameter setting (also called model selection) is indeed

crucial. These algorithms are used to optimize the performance of machine learning model to achieve a balance between model accuracy and model generalization.

The employed metaheuristic approaches include symbiotic organisms search [42], particle swarm optimization [43, 44], the forensic-based investigation optimization [45], equilibrium optimization [20], Harris hawks optimization [46], simulated annealing [47], social spider optimization [48, 49], gray wolf optimization [38, 50], teaching-learning-based algorithm [51], salp swarm algorithm [52, 53], artificial bee colony [54], pigeon-inspired optimization [55], cuckoo search optimization [56], imperialist competitive algorithm [57], moth flame optimization [58], and cuckoo search algorithm [59]. Those previous works have demonstrated the effectiveness of metaheuristic algorithms in optimizing machine learning models and solving complex tasks in various application domains.

Marine Predators Algorithm (MPA), first introduced in [60], is a recently proposed nature-inspired metaheuristic inspired from the foraging strategy of marine predators. This metaheuristic is characterized by a novel combination of Lévy and Brownian movements used for enhancing the optimization performance. The capability of MPA has been demonstrated by various optimization tasks [60]. Nevertheless, the performance of this metaheuristic used in optimizing machine learning models has rarely been investigated. Hence, this study proposes to hybridize MPA and SVM to form an integrated intelligent model used for remote sensing-based urban green space detection.

Remote sensing data obtained from Sentinel 2 satellite in the study area of Da Nang city is used to train and verify the MPA-SVM hybrid model. In this work, the MPA optimized SVM model trained by remote sensing data with all of the Sentinel 2's spectral bands is compared with the models that use commonly employed vegetation indices including normalized difference vegetation index (NDVI) [61], normalized difference water index (NDWI) [62], soil-adjusted vegetation index (SAVI) [63], and MERIS terrestrial chlorophyll index (MTCI) [64].

The rest of the article is organized as follows: Section 2 reviews the research methodology and material. Section 3 presents the proposed MPA optimized SVM used for remote sensing-based urban green space detection. Experimental results are reported in Section 4. Concluding remarks of the current study are summarized in Section 5.

2. Research Methodology and Material

2.1. General Description of the Study Area and Remote Sensing Data. As mentioned earlier, urban green spaces play a significant role in the urban living environment; they serve a variety of functions including climatic modification, aesthetics, recreation, and physical/mental health improvement. Nevertheless, due to the physical expansion of Da Nang city (Vietnam), certain areas of green spaces have been replaced by impervious surface such as buildings and roads. Therefore, the current status of urban green space in this city needs to be updated in a timely manner and this city has been selected as the study area of this research work.

Da Nang is a crucial coastal city located in Central Vietnam. Da Nang's location is at 15°55' to 16°14' North and 107°18' to 108°20' East [3]. It is the third largest city within the nation with a population of about 1 million. Da Nang urban center (refer to Figure 1) is located in the eastern section of the area and consists of six districts: Hai Chau, Cam Le, Thanh Khe, Lien Chieu, Ngu Hanh Son, and Son Tra [65]. The rural districts of Hoa Vang and Hoang Sa also belong to Da Nang city but are not included in the Da Nang urban center; therefore, these two rural districts are excluded from the study area.

To survey the urban green space status of Da Nang city, remote sensing data in form of spectral bands have been collected from Sentinel 2 on July 16, 2020. These spectral bands (see Table 1) are provided openly by USGS [66]; they can be processed and analyzed by Sentinel Application Platform (SNAP) software package [67] as well as ENVI software package [68]. Using the open-accessed tools of SNAP, the original Sentinel 2's spectral bands are converted to TIF format via the geometric operation of resampling. Moreover, it is noted that the used map projection of the obtained images is Universal Transverse Mercator (UTM) within Zone 48N–Datum World Geodetic System (WGS) 84. Images of the Sentinel-2 spectral bands are demonstrated in Figure 2. This figure demonstrates the 13 spectral bands obtained from the Sentinel-2 on July 16, 2020. These bands are coastal aerosol, blue, green, red, red-edge 1, red-edge 2, red-edge, near infrared, near infrared narrow, water vapour, shortwave infrared/cirrus, shortwave infrared 1, and shortwave infrared 2. The wavelength range and resolution of each spectral band are provided in Table 1.

2.2. Remote Sensing-Based Vegetation Indices. In remote sensing field, vegetation indices have been widely used to extract vegetation biophysical information from satellite image data [69]. Previous works have demonstrated the effectiveness of vegetation indices in remote sensing-based green space mapping [7, 13, 70–72]. Therefore, this study relies on such conventional indices as a means of urban green space detection. The employed vegetation indices include normalized difference vegetation index (NDVI) [61], soil-adjusted vegetation index (SAVI) [63], normalized difference water index (NDWI) [62], and MERIS terrestrial chlorophyll index (MTCI) [64]. These indices can be computed as follows [61, 69, 73]:

$$\text{NDVI} = \frac{\rho_{\text{nir}} - \rho_{\text{red}}}{\rho_{\text{nir}} + \rho_{\text{red}}}, \quad (1)$$

where ρ_{red} and ρ_{nir} represent the red reflected radiant flux (Sentinel 2's band 4) and near-infrared radiant flux (Sentinel 2's band 8a), respectively.

$$\text{SAVI} = \frac{(\rho_{\text{nir}} - \rho_{\text{red}})}{(\rho_{\text{nir}} + \rho_{\text{red}} + L)} (1 + L), \quad (2)$$

where $L = 0.5$ denotes the soil brightness correction factor [74].

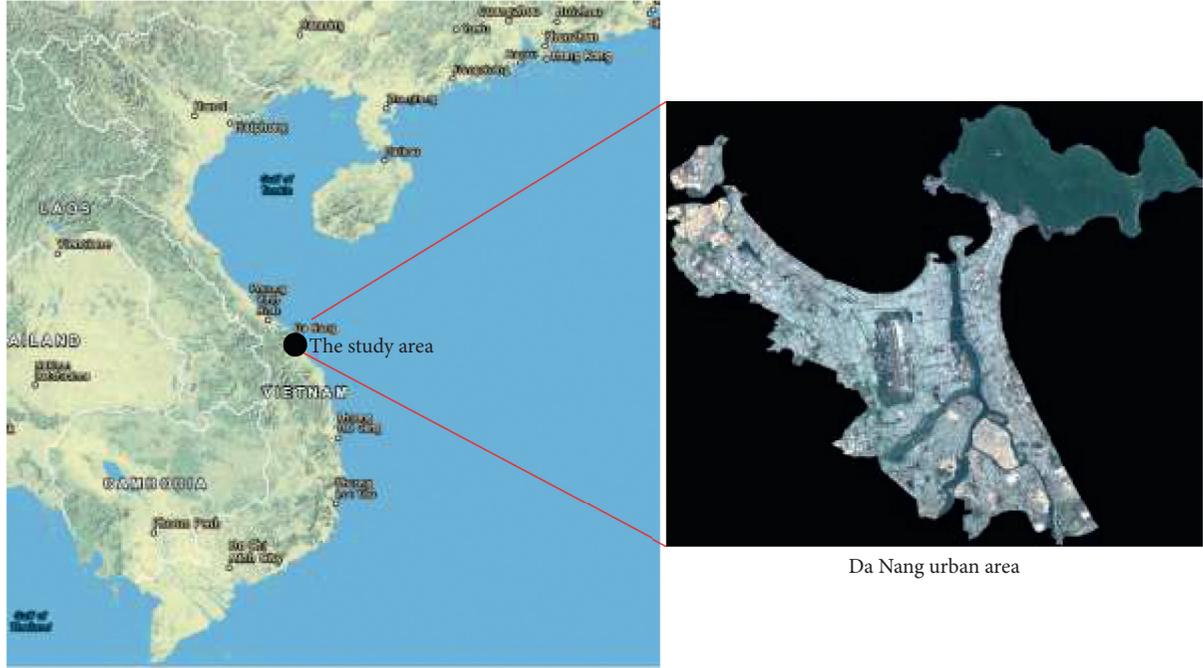


FIGURE 1: Da Nang urban area.

TABLE 1: The Sentinel-2 spectral bands.

Band number	Description	Wavelength range (nm)	Resolution (m)
B1	Coastal aerosol	433–453	60
B2	Blue	458–523	10
B3	Green	543–578	10
B4	Red	650–680	10
B5	Red-edge 1	698–713	20
B6	Red-edge 2	733–748	20
B7	Red-edge	773–793	20
B8	Near infrared (NIR)	785–900	10
B8a	Near infrared narrow (NIRn)	855–875	20
B9	Water vapour	935–955	60
B10	Shortwave infrared/Cirrus	1360–1390	60
B11	Shortwave infrared 1 (SWIR1)	1565–1655	20
B12	Shortwave infrared 2 (SWIR2)	2100–2280	20

$$NDWI = \frac{\rho_{nir} - \rho_{swir}}{\rho_{nir} + \rho_{swir}}, \quad (3)$$

where ρ_{red} and ρ_{swir} represent the near-infrared radiant flux (Sentinel 2's band 8a) and shortwave infrared (Sentinel 2's band 11), respectively.

$$MTCI = \frac{B6 - B5}{B5 - B4}, \quad (4)$$

where ρ_{red} is Sentinel 2's B4; ρ_{red_edge} is Sentinel 2's B5; ρ_{nir} is Sentinel 2's B6 [74, 75].

2.3. The Used Metaheuristic and Machine Learning Approaches

2.3.1. Marine Predators Algorithm. Marine Predators Algorithm (MPA), first introduced in [60], is a stochastic

global optimization algorithm inspired from the widespread foraging strategy of marine species such as sharks and tunas. The foraging strategy of these marine species effectively utilizes Lévy and Brownian movements along with optimal encounter rate policy in biological interaction between predator and prey [76–78].

The searching process of MPA consists of three phases considering three scenarios: (i) high velocity ratio when a prey is moving faster than a predator, (ii) unit velocity ratio when the rates of movement of a prey and a predator are similar, and (iii) low velocity ratio when the rate of movement of a predator is higher than that of a prey. The searching operation of the MPA metaheuristic is demonstrated in Figure 3. Let XE be the position of predators and XP be the position of preys within a marine ecosystem. The 1st phase aims at search space exploration and is applied for the first one-third of the searching iteration number; the

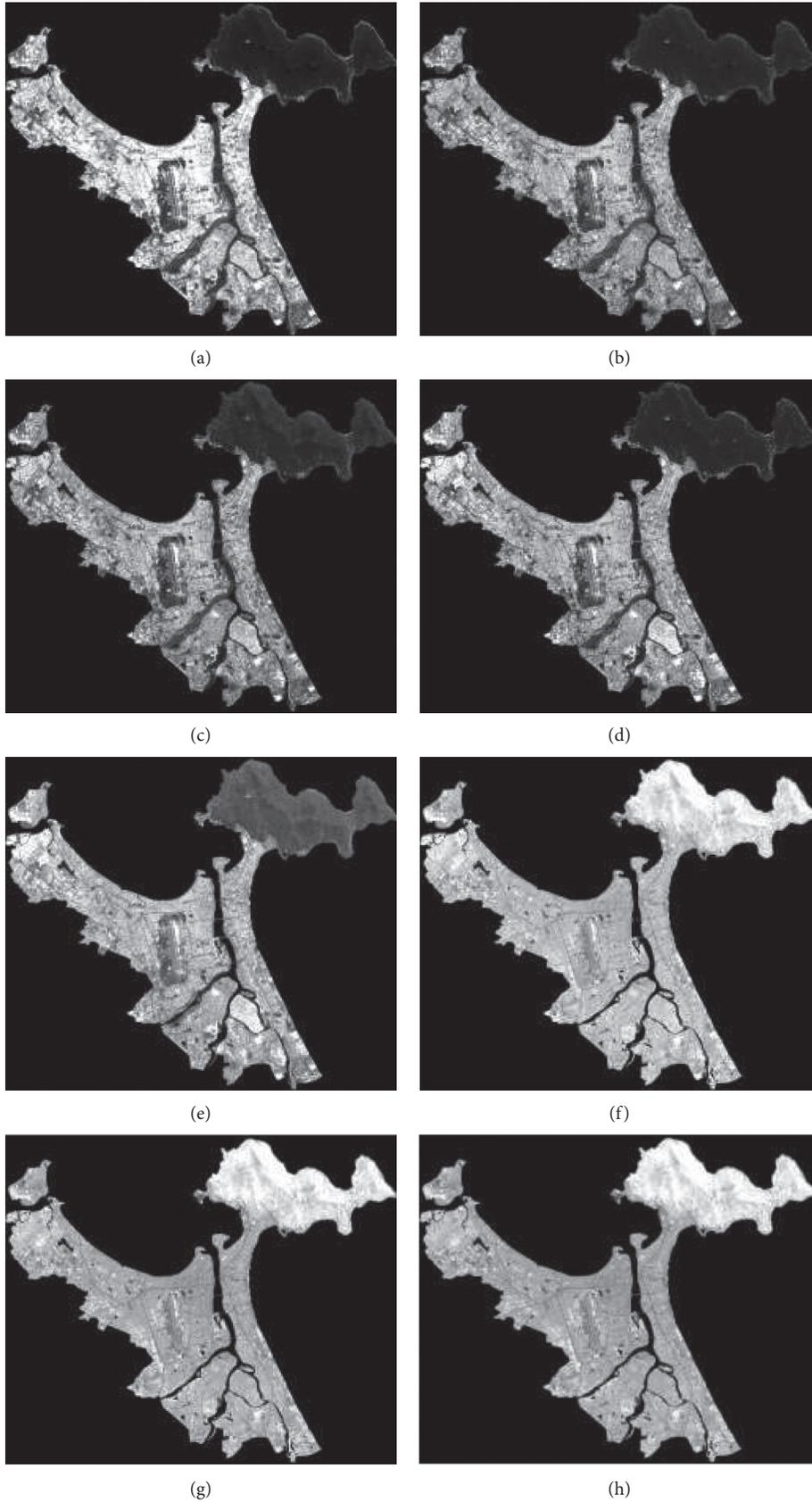


FIGURE 2: Continued.

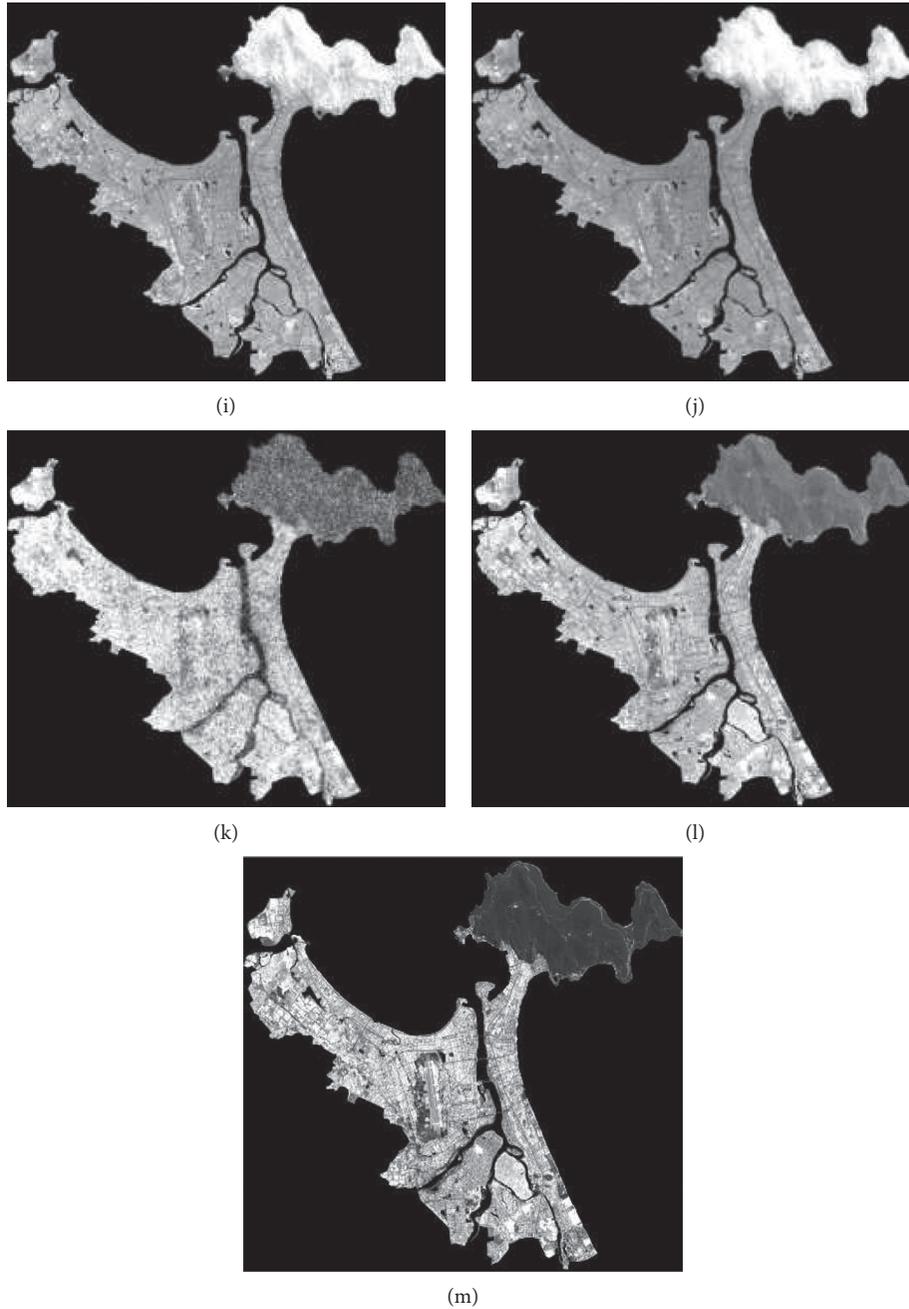


FIGURE 2: Gray-scale image demonstrated the Sentinel 2's spectral bands: (a) B1, (b) B2, (c) B3, (d) B4, (e) B5, (f) B6, (g) B7, (h) B8, (i) B8a, (j) B9, (k) B10, (l) B11, and (m) B12.

mathematical equation used to revise the prey position is given by

$$XP_i = XP_i + 0.5 \otimes R \otimes R_B \otimes (XE_i - R_B \otimes XP_i), \quad (5)$$

$$i = 1, 2, \dots, NP,$$

where \otimes is an entry-wise multiplication operator. R_B is a vector including random numbers generated from a normal distribution which mimics the Brownian motion. i denotes the index of population members. R represents a vector of

uniform random number within $[0, 1]$. NP is the number of population members.

The 2nd phase serves as an intermediate phase and occurs within the second one-third of the searching iteration number. The positions of the first half of the population members are updated as follows:

$$XP_i = XP_i + 0.5 \otimes R \otimes R_L \otimes (XE_i - R_L \otimes XP_i), \quad (6)$$

$$i = 1, 2, \dots, \frac{NP}{2},$$

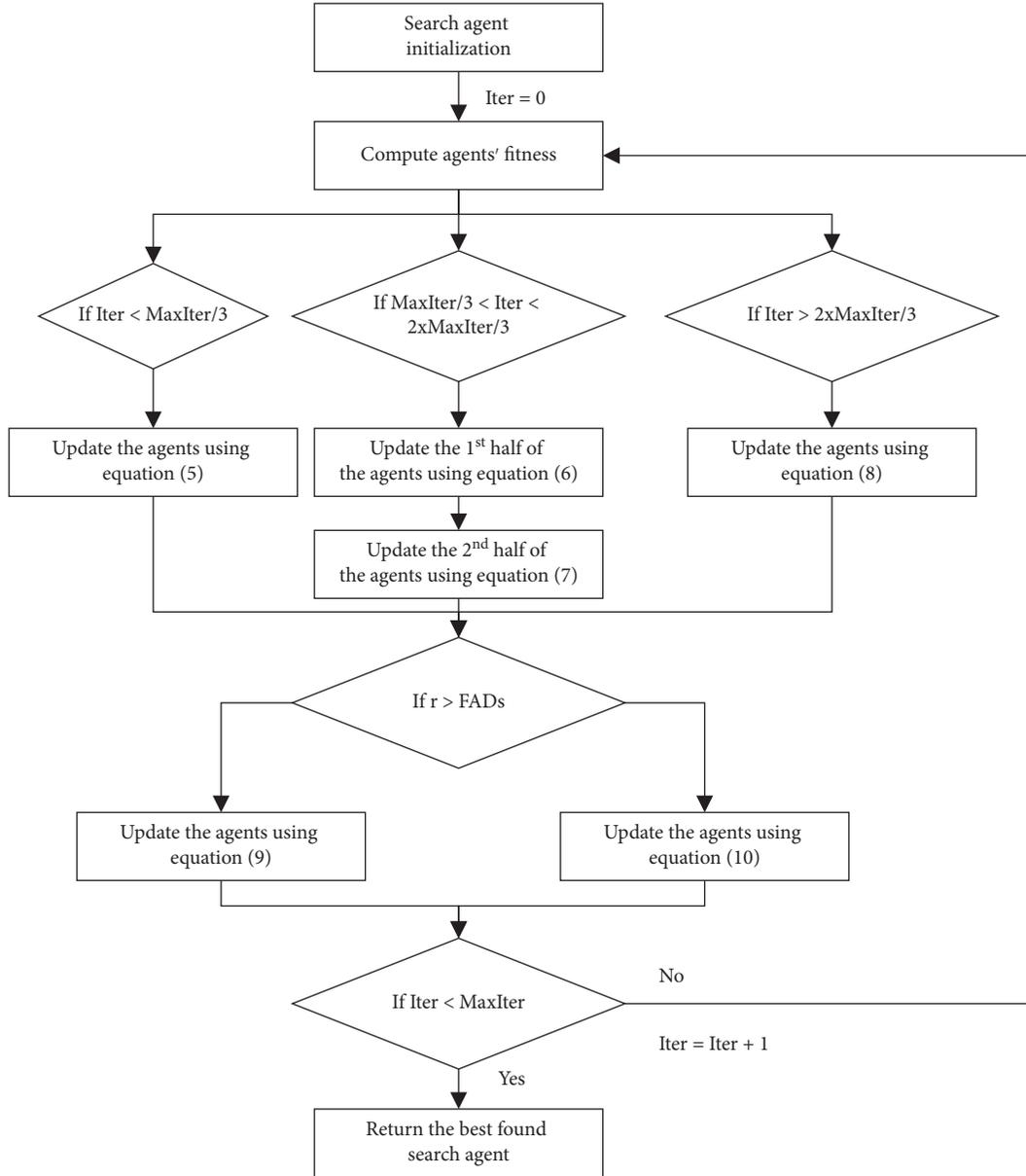


FIGURE 3: Flowchart of the MPA algorithm.

where R_L denotes a vector of random numbers generated from the Lévy distribution which represents the Lévy movement.

The positions of the second half of the population members are updated as follows:

$$XP_i = XE_i + 0.5 \times CF \otimes R_B \otimes (R_B \otimes XE_i - XP_i),$$

$$i = \frac{NP}{2}, \frac{NP}{2} + 1, \dots, NP, \quad (7)$$

where $CF = (1 - (\text{Iter}/\text{MaxIter}))^{2(\text{Iter}/\text{MaxIter})}$; Iter and MaxIter are the current iteration count and the maximum number of iterations, respectively.

The last phase of the optimization process aims at exploitation of the search space. The population members' positions are updated in the following equation:

$$XP_i = XE_i + 0.5 \times CF \otimes R_L \otimes (R_L \otimes XE_i - XP_i),$$

$$i = 1, 2, \dots, NP. \quad (8)$$

In addition, to model behavior shift in marine predators according to the eddy formation or Fish Aggregating Devices (FADs) effects [79], the MPA metaheuristic employs the following operation:

$$XP_i = XP_i + CF[LB + R \otimes (UB - LB)] \otimes U \text{ if } r \leq \text{FADs},$$

$$XP_i = XP_i + [\text{FADs}(1 - r) + r][XP_{r_1} - XP_{r_2}] \text{ if } r > \text{FADs}, \quad (9)$$

where $\text{FADs} = 0.2$ denotes the probability of the FADs effect. U represents a random binary vector. r is a random number within $[0, 1]$. LB and UB are vectors of lower and upper

boundaries of the searched variables. $r1$ and $r2$ denote two random indices.

2.3.2. Support Vector Machine. Introduced by Vapnik [40], support vector machines (SVM) have gained attentions of the academic community and have become a preeminent pattern recognition approach [55, 80–90]. Given a data sample set S drawn from a data universe X_U , a hidden target function $f: X \rightarrow \{0, 1\}$, we first create a labeled training dataset D , where $D = \{(x, y) | x \in S \text{ and } y = f(x)\}$. The SVM machine learning can be used to estimate the target function $f(x)$ by constructing a function $\hat{f}(x): X \rightarrow \{0, 1\}$ based on data samples stored in D so that $\hat{f}(x) \cong f(x)$ for all x in X . Herein, for the task of urban green space detection, the data label can be modeled as “0” = nongreen space (the negative class) and “1” = green space (the positive class). The input data X are properties of the Sentinel 2’s spectral bands.

To construct a SVM model, it is required to solve the following constrained optimization problem [91]:

$$\text{Minimize } J_p(w, e) = \frac{1}{2}w^T w + C \frac{1}{2} \sum_{k=1}^N e_k^2,$$

$$\text{Subjected to } y_k(w^T \varphi(x_k) + b) \geq 1 - e_k \quad k = 1, \dots, N, \quad e_k \geq 0,$$
(10)

where $w \in R^n$ and $b \in R$ are used to construct a classification hyperplane used for pattern classification. e is the vector of slack variables. C denotes the penalty coefficient. $\varphi(x)$ denotes a nonlinear data mapping used for dealing with data that cannot be linearly separated.

One advantage of the SVM method is that the explicit formula of $\varphi(x)$ is not required. To construct a SVM model, only the dot product of $\varphi(x)$ is necessary. The dot product of two data samples x_k and x_l is represented as a kernel function $K(x_k, x_l)$:

$$K(x_k, x_l) = \varphi(x_k)^T \varphi(x_l). \quad (11)$$

For multivariate and nonlinear data classification problem, the radial basis kernel function (RBKF) is commonly utilized:

$$K(x_k, x_l) = \exp\left(-\frac{\|x_k - x_l\|^2}{2\sigma^2}\right), \quad (12)$$

where σ denotes a hyperparameter of the RBKF.

By solving a Lagrangian dual of the aforementioned constrained optimization problem and using a quadratic programming solver, the SVM model used for data classification can be expressed compactly as follows [91]:

$$y(x_l) = \text{sign}\left(\sum_{k=1}^{SV} \alpha_k y_k K(x_k, x_l) + b\right), \quad (13)$$

where α_k represents the solution of the optimization problem; SV denotes the number of support vectors.

3. The Proposed Marine Predators Algorithm Optimized Machine Learning Approach for Urban Green Space Detection

This section of the article is dedicated to describing the integrated model used for remote sensing-based urban green space detection. The core of the proposed model is a hybridization of the MPA metaheuristic and the SVM machine learning. These two methods work synergistically to analyze patterns hidden in a set of remotely sensed data collected for the study area of Da Nang urban center. In detail, SVM is used to construct a decision boundary that separates the input data space into two distinctive regions of “nongreen space” and “green space”.

To further enhance the performance of the SVM model, MPA is utilized to autonomously fine-tune the SVM training process by identifying a set of appropriate model hyperparameters. The optimized hyperparameters include the penalty coefficient and the RBKF parameter. In this study, the searching range of the penalty coefficient is [1, 100]; the searching range of the RBKF parameter is [0.1, 100].

These two hyperparameters strongly influence the learning and the predictive capability of the integrated urban green space detection model. A too large penalty coefficient or a too small RBKF parameter leads to overfitted models. On the other hand, a too small penalty coefficient and a too large RBKF parameter tends to construct underfitted models [92]. Therefore, the role of the MPA is to find a set of the penalty coefficient and the RBKF parameter which features a balance between predictive accuracy and modeling generalization. Accordingly, it is expected that the constructed model will not suffer from either overfitting or underfitting.

The overall model structure is presented in Figure 4. The 13 spectral bands are obtained from the Sentinel 2 satellite and processed via SNAP [67] and ENVI [68] software packages. To accelerate the computing process, the images of spectral bands are divided into blocks with the size 5×5 pixels. For the purpose of data classification, the average (μ_c) and the standard deviation (σ_c) of gray intensity of each image block are computed and used as numerical features by the integrated MPA-SVM model.

The average and standard deviation of gray intensity are given by [93]

$$\mu_c = \sum_{i=0}^{NL-1} I_{i,c} \times P_c(I),$$

$$\sigma_c = \sqrt{\sum_{i=0}^{NL-1} (I_{i,c} - \mu_c)^2 \times P_c(I)}, \quad (14)$$

where $I_{i,c} = 0, 1, 2, \dots, 255$. $NL = 256$ represents the number of discrete color values. $P(I)$ is the first-order histogram of an image block [94].

To construct the integrated MPA-SVM model for urban green space detection, it is necessary to prepare a training dataset with assigned ground truth labels. This study has

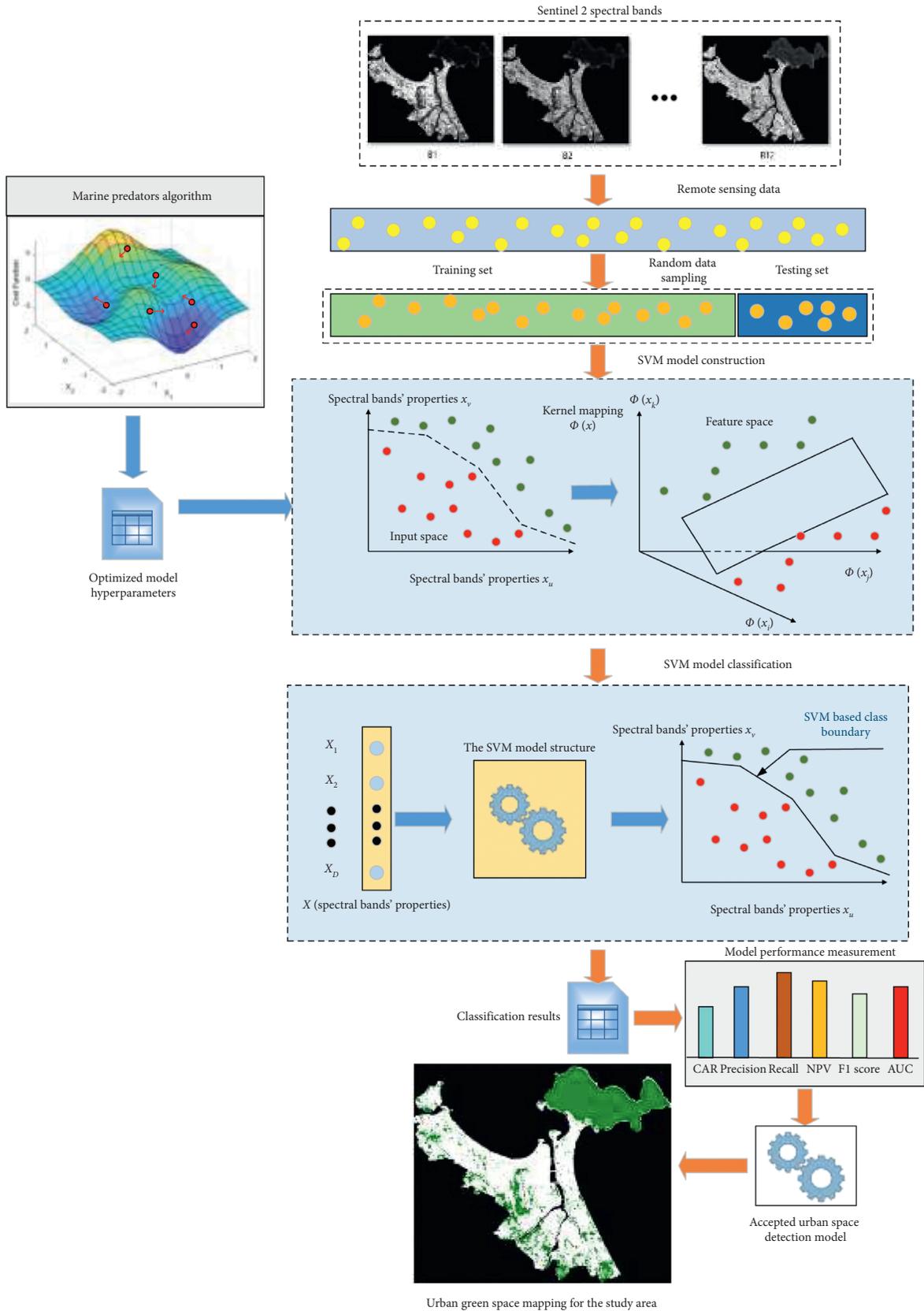


FIGURE 4: The proposed MPA optimized SVM for urban green space detection.

performed sampling process to collect data in the nongreen space and green space areas within the study area (demonstrated in Figure 5). It is noted that the ground truth label of each image data in this study has been verified by field trips and Google Earth Engine. Each block is sampled with the size of 25×25 pixels to generate nonoverlapped image patches with the size of 5×5 pixels. After the data sampling process, there are 1,000 image patches available for the feature extraction operator. Herein, each class label contains 500 image patches to guarantee a balanced classification. The collected dataset is illustrated in Table 2. Notably, each spectral band yields two statistical measurements (i.e., the mean and standard deviation) and there are 13 bands. Thus, the total number of features used for classification is $13 \times 2 = 26$.

It is worth noticing that the extracted dataset including the input features which characterize statistical properties of the spectral bands and the corresponding class labels has been randomly separated into a training (70%) dataset and a testing dataset (30%) [95]. The first set is used for model training and the second set is used to inspect the model predictive capability. In addition, to standardize the input data range, the Z-score equation is used as follows [96]:

$$X_Z = \frac{X_D - M_X}{STD_X}, \quad (15)$$

where X_Z and X_D denote the normalized and the original features, respectively. M_X and STD_X denote the mean value and the standard deviation of the features, respectively.

To optimize the SVM model used for urban green space detection, the objective function of the MPA metaheuristic has employed a 5-fold cross validation process and the indices of false negative rate (FNR) and false positive rate (FPR). This objective function (OF) is described as follows [96]:

$$OF = \frac{\sum_{k=1}^5 (FNR_k + FPR_k)}{5}, \quad (16)$$

where FNR_k and FPR_k denote FNR and FPR computed in the k th run, respectively.

The FNR and FPR indices are given by [96]

$$\begin{aligned} FNR &= \frac{FN}{FN + TP}, \\ FPR &= \frac{FP}{FP + TN}, \end{aligned} \quad (17)$$

where FN, FP, TP, and TN are the false negative, false positive, true positive, and true negative data samples, respectively.

In this study, the source code of the MPA metaheuristic is provided by Faramarzi et al. [60]. For the purpose of model optimization, the integrated MPA-SVM has been constructed in MATLAB. The SVM model with an optimized set of hyperparameters is then developed in Visual C#

.NET framework 4.7.2 to process and analyze remote sensing data. The SVM model in Visual C# .NET has been built with functions provided by the Accord.NET Framework [97]. Moreover, the program has been implemented with the ASUS FX705GE-EW165T (Core i7 8750H and 8 GB Ram) platform.

4. Experimental Results and Discussion

In this section, a set of performance measurement indices is used to express the model predictive accuracy. This set includes classification accuracy rate (CAR), precision, recall, negative predictive value (NPV), F1 score, and area under the receiver operating characteristic curve (AUC) [98, 99]. The calculation of AUC is described in [99]. The formulas used to compute CAR, precision, recall, NPV, and F1 score are given by

$$CAR = \frac{N_C}{N_A} 100\%, \quad (18)$$

where N_C and N_A are the numbers of correctly predicted data and the total number of data, respectively.

$$\begin{aligned} \text{Precision} &= \frac{TP}{TP + FP}, \\ \text{Recall} &= \frac{TP}{TP + FN}, \\ \text{NVP} &= \frac{TN}{TN + FN}, \\ \text{F1 Score} &= \frac{2TP}{2TP + FP + FN}. \end{aligned} \quad (19)$$

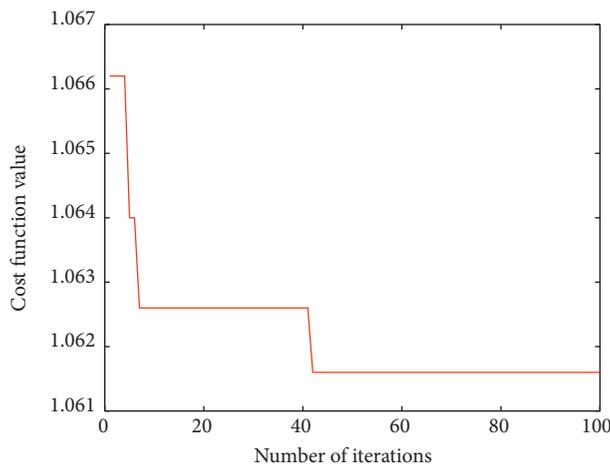
Besides the MPA-SVM model which utilizes information provided by 13 spectral bands, this study has employed the MPA-SVM models using the aforementioned vegetation indices as benchmark models. The MPA-SVM employing all of the Sentinel 2's bands is denoted as MPA-SVM-13B. The benchmark models that use the NDWI, NDVI, SAVI, and MTCI are denoted as MPA-SVM-NDWI, MPA-SVM-NDVI, MPA-SVM-SAVI, and MPA-SVM-MTCI, respectively. The MPA-SVM-13B utilizes the statistical information obtained from all of the 13 spectral bands (i.e., the mean and the standard deviation of each band). Meanwhile, the MPA-SVM-NDWI, MPA-SVM-NDVI, MPA-SVM-SAVI, and MPA-SVM-MTCI employ the statistical information of the vegetation indices of NDWI, NDVI, SAVI, and MTCI, respectively. Therefore, the feature extraction phase of the benchmark models is similar to that of the MPA-SVM-13B. This feature extraction phase also computes the two indices of mean and standard deviation of image patches. The model optimization processes of the constructed models are demonstrated in Figure 6. Herein, the maximum number of searching iteration (MaxIter) of the MPA metaheuristic has been set to be 100; the number of population members (NP) is fixed to be 20. The detailed



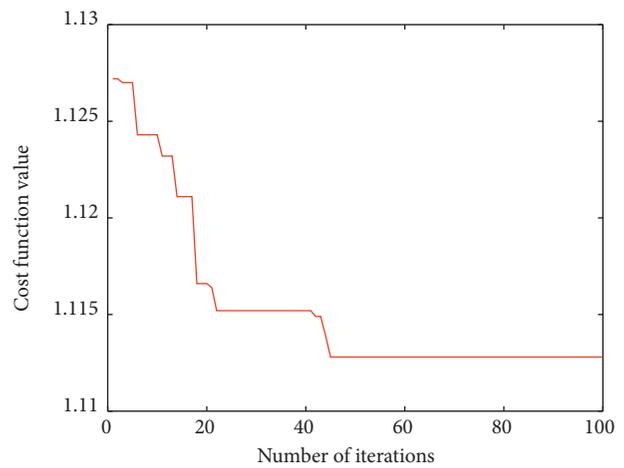
FIGURE 5: Demonstrations of the collected image samples using natural color composite: (a) nongreen space class and (b) green space class.

TABLE 2: Demonstration of the extracted dataset.

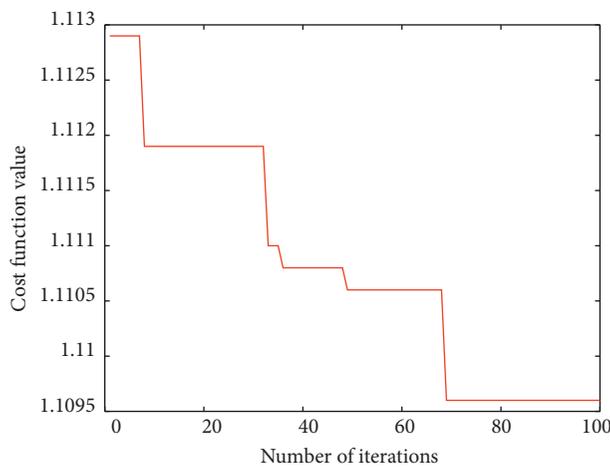
Sample	X1	X2	X3	X4	X5	...	X21	X22	X23	X25	X26	Y
0	165.96	152.68	145.48	168.60	168.72	15.58	13.33	5.21	35.76	20.36	20.32	0
1	152.96	101.28	95.32	138.36	158.08	17.27	7.10	9.08	31.78	12.11	16.31	0
2	133.12	90.92	83.68	143.56	160.76	14.02	11.95	6.57	21.24	29.14	31.10	0
3	122.40	91.84	84.52	140.20	168.60	12.01	5.77	3.92	18.13	22.55	26.75	0
4	119.40	90.52	76.00	104.08	124.00	13.80	11.45	9.80	0.00	21.92	21.31	0
...
996	23.08	24.68	25.76	26.20	67.60	16.22	6.67	11.00	21.87	10.96	7.93	1
997	25.16	25.56	31.16	25.64	69.84	16.18	9.55	9.90	18.37	6.81	3.75	1
998	25.24	30.76	30.16	35.68	72.16	8.07	5.14	8.14	18.49	7.19	5.39	1
999	16.80	19.52	21.00	18.12	63.16	14.17	16.83	5.39	18.13	12.39	6.39	1
1000	19.20	20.80	22.00	19.68	58.92	15.80	12.60	0.98	0.00	3.04	1.51	1



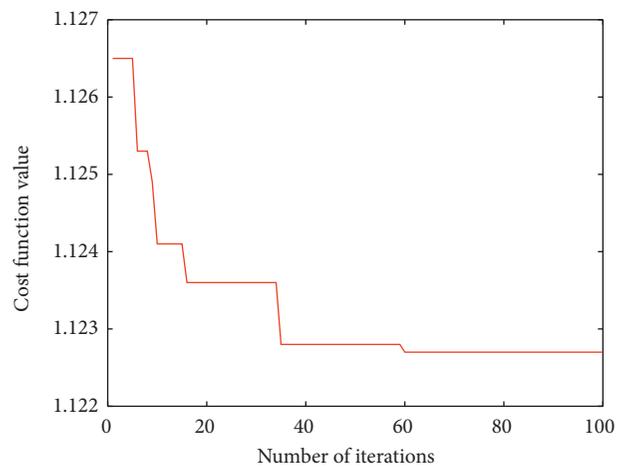
(a)



(b)



(c)



(d)

FIGURE 6: Continued.

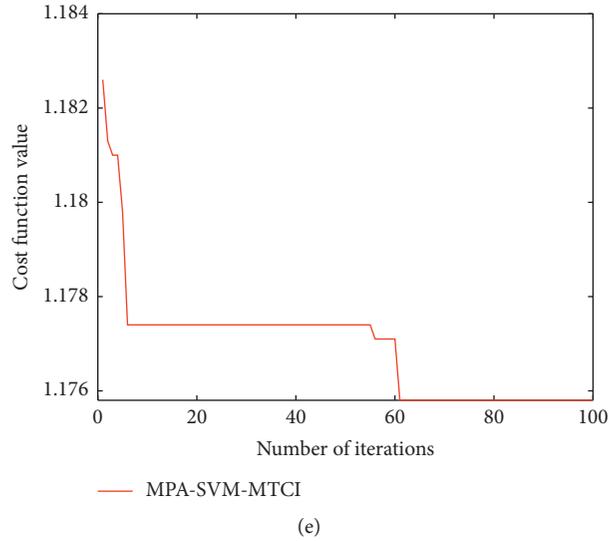


FIGURE 6: Optimization process of the MPA metaheuristic: (a) MPA-SVM-13B, (b) MPA-SVM-NDWI, (c) MPA-SVM-NDVI, (e) MPA-SVM-SAVI, and (e) MPA-SVM-MTCI.

TABLE 3: MPA-based optimization results.

Hyperparameters	MPA-SVM-13B	MPA-SVM-NDWI	MPA-SVM-NDVI	MPA-SVM-SAVI	MPA-SVM-MTCI
Penalty coefficient	79.782	3.938	89.886	9.605	32.495
RBFK	10.244	0.438	4.065	1.423	1.622

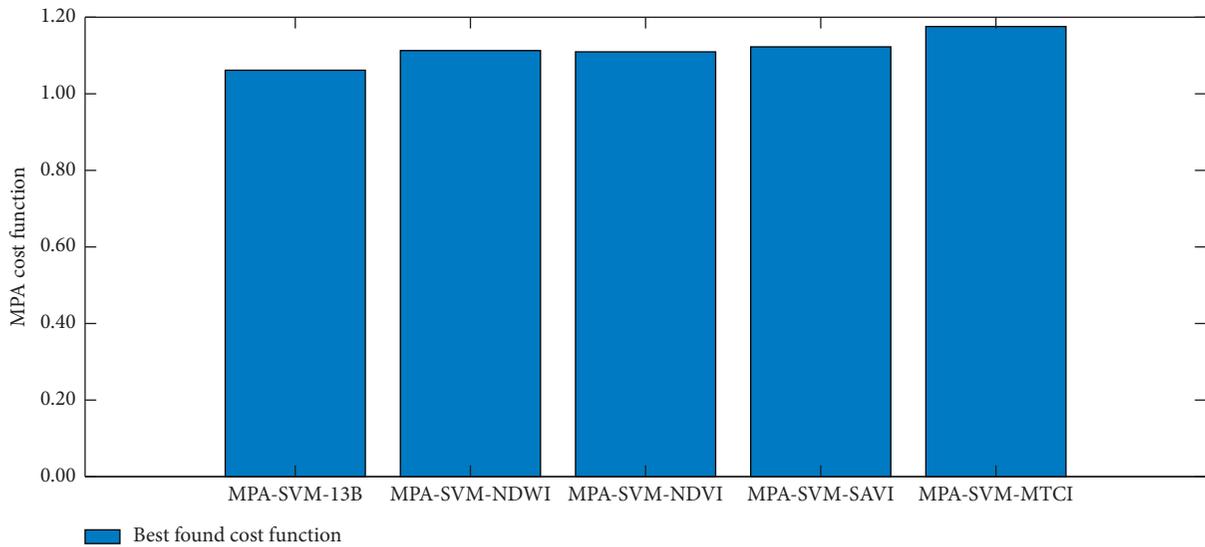


FIGURE 7: Cost function values optimized by MPA.

optimization results are reported in Table 3 which shows the best penalty coefficient and RBFK parameters for each model used for urban green space detection. In addition, the best found cost function values of the MPA-SVM-13B, MPA-SVM-NDWI, MPA-SVM-NDVI, MPA-SVM-SAVI, and

MPA-SVM-MTCI are provided in Figure 7. It can be observed from Figure 7 that the MPA-SVM using all of the 13 bands results in the lowest value of cost function.

As stated earlier, the constructed dataset has been randomly divided into a training set (70%) and a testing set

TABLE 4: Experimental result comparison.

Phase	Metrics	MPA-SVM-13B		MPA-SVM-NDWI		MPA-SVM-NDVI		MPA-SVM-SAVI		MPA-SVM-MTCI	
		Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std
Training	CAR (%)	95.429	0.586	89.564	0.516	89.514	0.755	88.929	0.439	84.686	0.785
	Precision	0.941	0.009	0.882	0.008	0.885	0.011	0.881	0.008	0.819	0.009
	Recall	0.970	0.005	0.914	0.006	0.907	0.013	0.901	0.009	0.894	0.015
	NPV	0.968	0.005	0.911	0.007	0.905	0.011	0.898	0.007	0.882	0.015
	F1 score	0.955	0.006	0.897	0.006	0.896	0.008	0.891	0.005	0.855	0.008
	AUC	0.991	0.002	0.940	0.007	0.943	0.006	0.936	0.006	0.916	0.006
Testing	CAR (%)	93.100	1.411	89.400	1.356	89.300	1.729	88.983	1.062	83.850	1.698
	Precision	0.916	0.024	0.881	0.024	0.894	0.029	0.886	0.017	0.803	0.022
	Recall	0.947	0.018	0.911	0.017	0.895	0.020	0.893	0.023	0.888	0.030
	NPV	0.947	0.019	0.908	0.020	0.893	0.022	0.895	0.017	0.881	0.031
	F1 score	0.931	0.014	0.896	0.013	0.894	0.017	0.889	0.012	0.843	0.018
	AUC	0.979	0.006	0.926	0.014	0.946	0.013	0.937	0.012	0.910	0.016

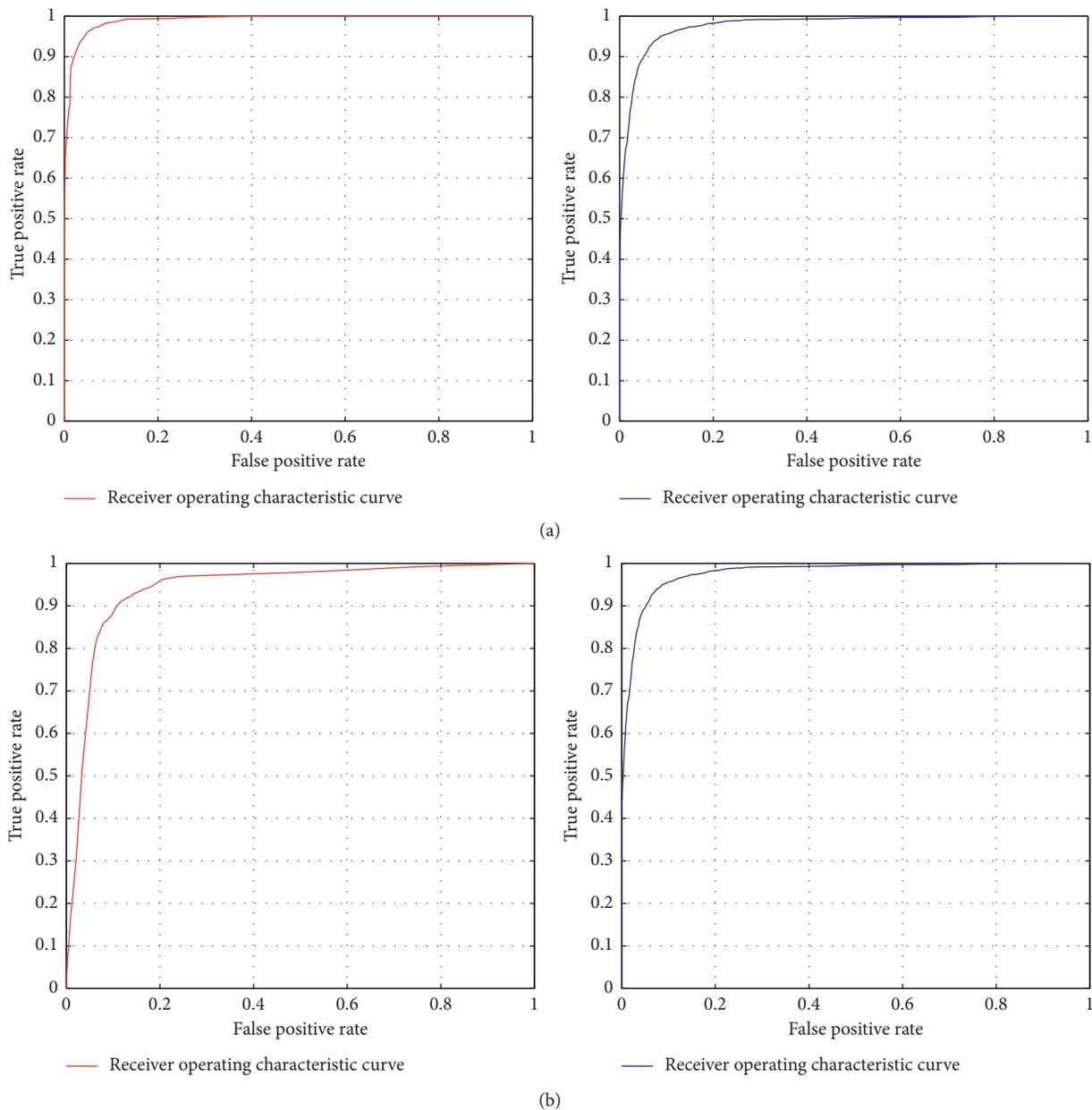
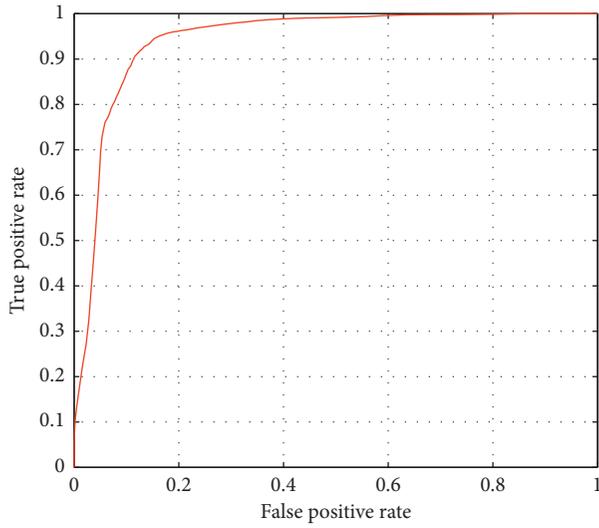
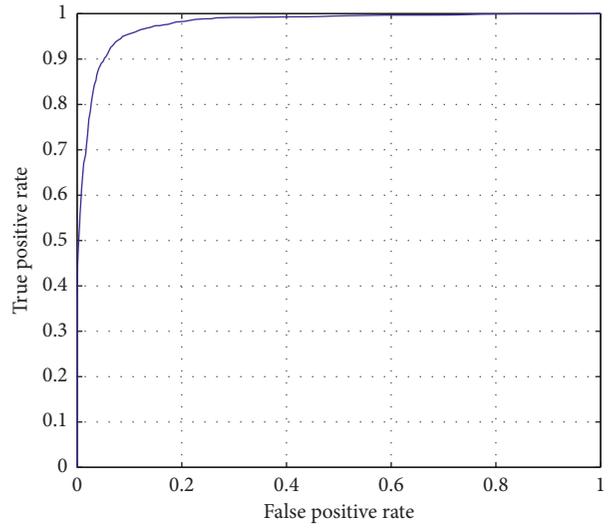


FIGURE 8: Continued.

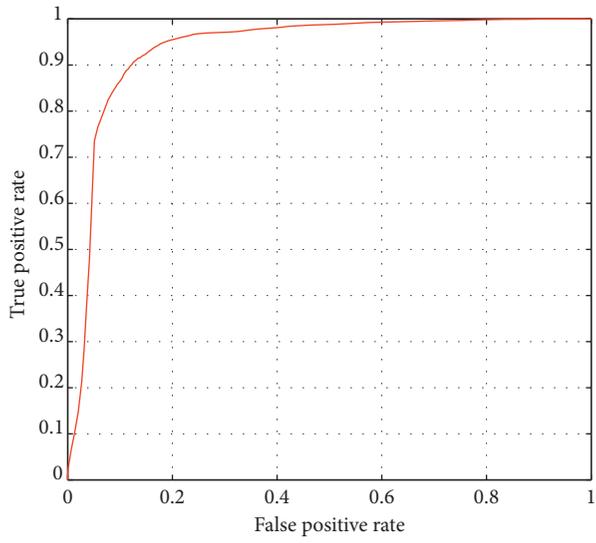


— Receiver operating characteristic curve

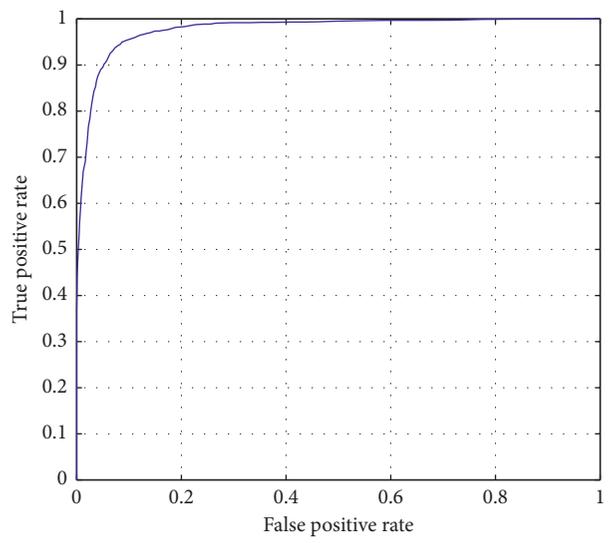


— Receiver operating characteristic curve

(c)



— Receiver operating characteristic curve



— Receiver operating characteristic curve

(d)

FIGURE 8: Continued.

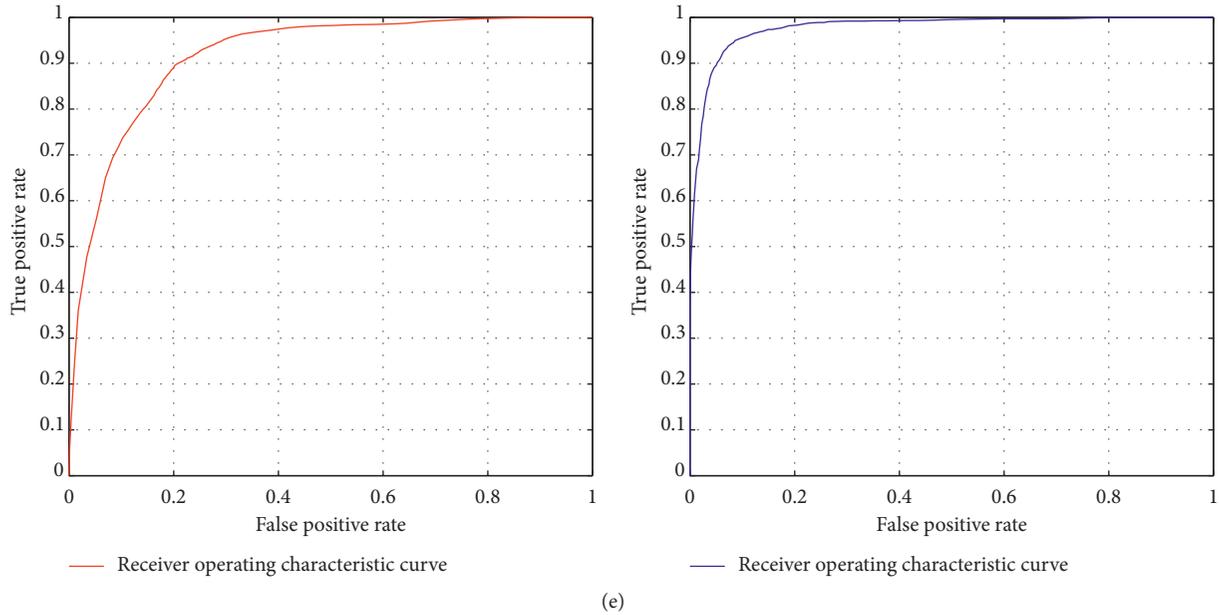


FIGURE 8: ROCs of the prediction models: (a) MPA-SVM-13B, (b) MPA-SVM-NDWI, (c) MPA-SVM-NDVI, (d) MPA-SVM-SAVI, and (e) MPA-SVM-MTCI.

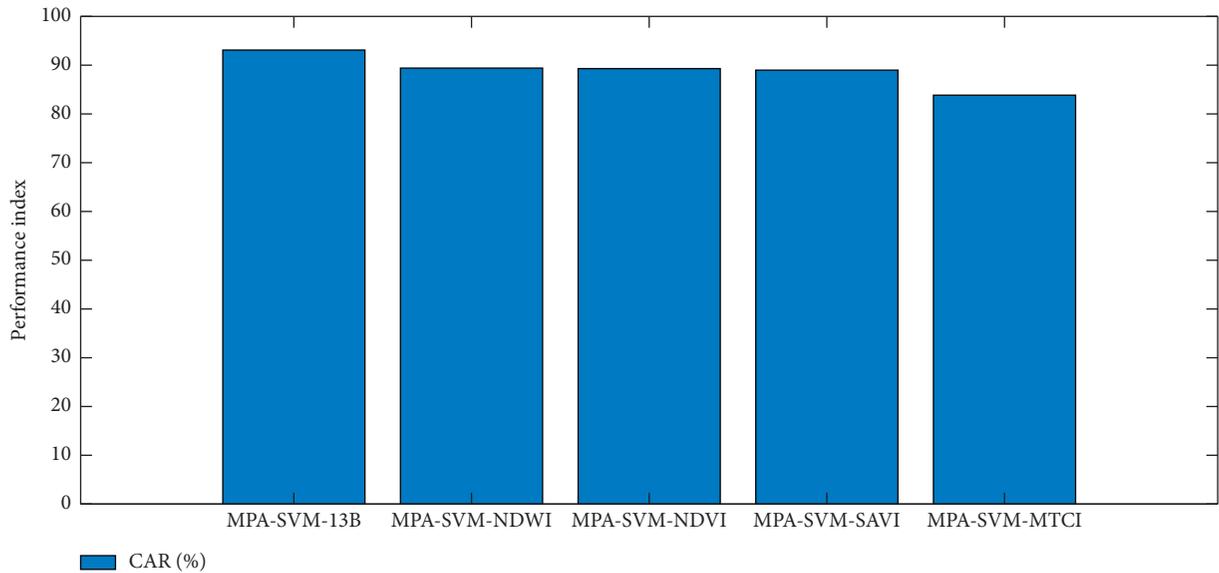


FIGURE 9: Result comparison in terms of CAR.

(30%). The first set is used for model training and the second set is reserved for model validation. Moreover, in order to reliably evaluate the model predictive performance, this study has repeated the model training and prediction processes 20 times. It is noted that the training and testing datasets are resampled in each run. The statistical

measurements obtained from this multiple model construction and validation phases are used for model assessment. This repeated process aims at diminishing the variation caused by the randomness in data sampling. The model prediction outcomes are summarized in Table 4 which shows the mean and standard deviation (Std) of

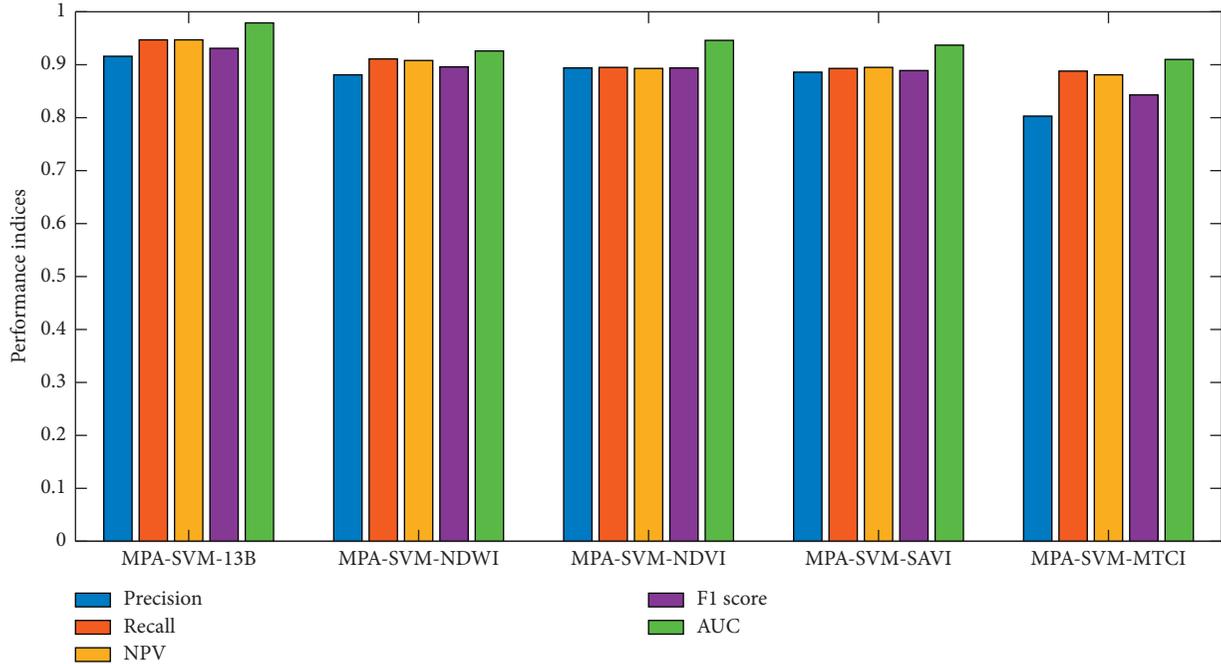


FIGURE 10: Result comparison in terms of precision, recall, NPV, F1 score, and AUC.

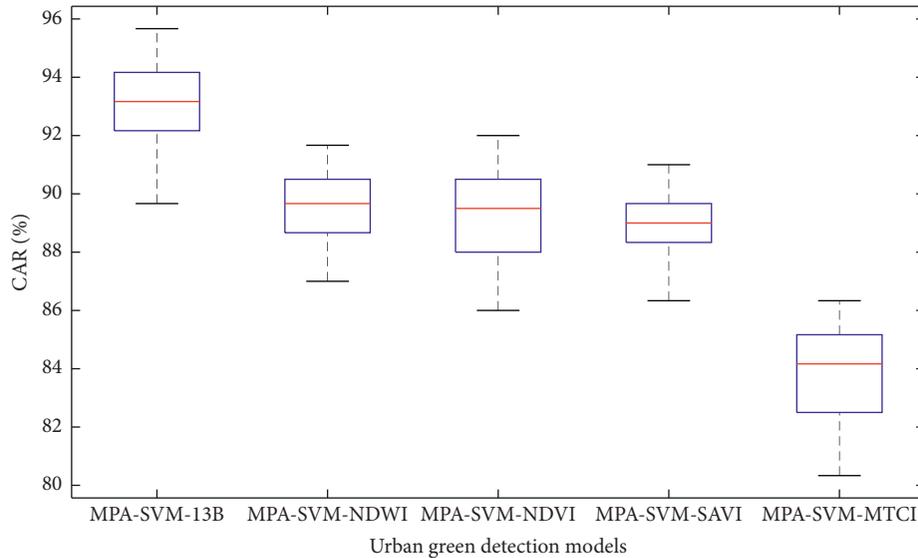


FIGURE 11: Box plot of CAR values obtained from the employed machine learning models.

the employed performance measurement indices. It can be observed that the MPA-SVM-13B with CAR=93.100%, precision=0.916, recall=0.947, NPV=0.947, F1 score=0.931, and AUC=0.979 has obtained the most desired urban green space detection performance. Compared with the proposed method, MPA-SVM-NDWI has gained a lower

CAR (89.400%) and F1 score (0.896), followed by MPA-SVM-NDVI (CAR= 89.300% and F1 score=0.894), MPA-SVM-SAVI (88.983% and F1 score=0.889), and MPA-SVM-MTCI (83.850% and F1 score=0.843).

In terms of AUC score, MPA-SVM-13B is the best model (AUC=0.979), followed by MPA-SVM-NDVI (AUC=

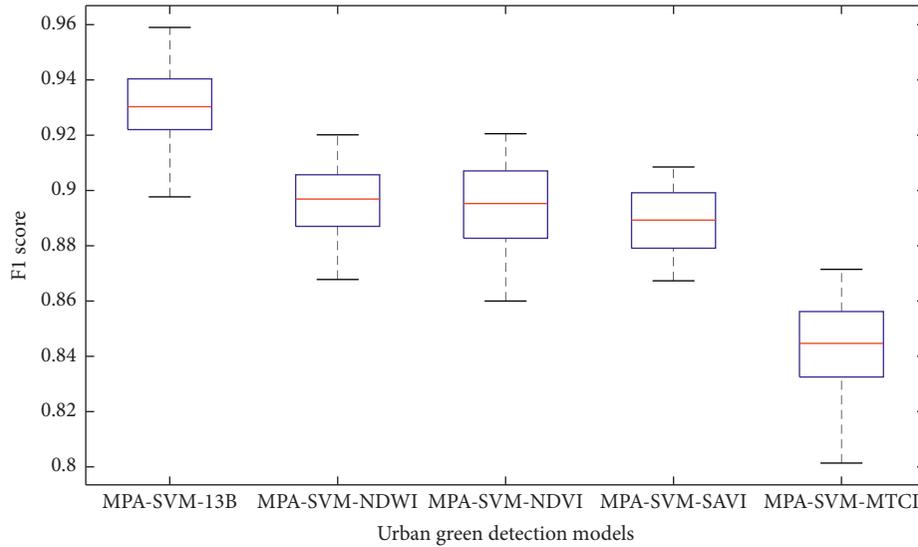


FIGURE 12: Box plot of F1 score values obtained from the employed machine learning models.

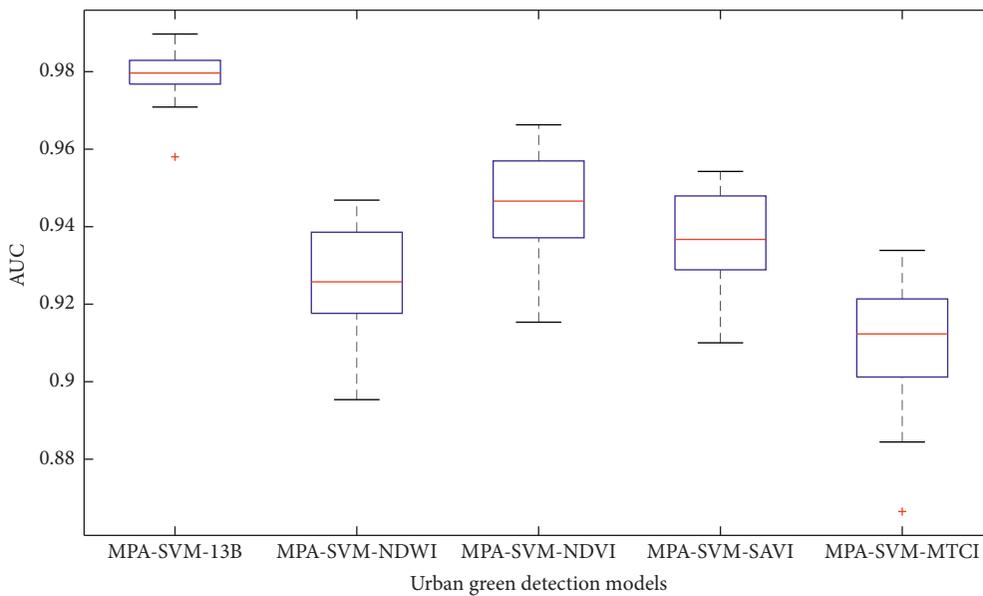


FIGURE 13: Box plot of AUC values obtained from the employed machine learning models.

TABLE 5: Wilcoxon signed rank test results with CAR index.

	MPA-SVM-13B	MPA-SVM-NDWI	MPA-SVM-NDVI	MPA-SVM-SAVI	MPA-SVM-MTCI
MPA-SVM-13B	0.00000	0.00010	0.00013	0.00013	0.00009
MPA-SVM-NDWI	0.00010	0.00000	0.76496	0.31380	0.00009
MPA-SVM-NDVI	0.00013	0.76496	0.00000	0.49869	0.00009
MPA-SVM-SAVI	0.00013	0.31380	0.49869	0.00000	0.00009
MPA-SVM-MTCI	0.00009	0.00009	0.00009	0.00009	0.00000

0.946), MPA-SVM-SAVI (AUC = 0.937), MPA-SVM-NDWI (AUC = 0.926), and MPA-SVM-MTCI (AUC = 0.910). The AUC values of the employed models used for urban green space detection are demonstrated in Figure 8. In

addition, the model result comparison in terms of CAR, precision, recall, NPV, and AUC is provided in Figures 9 and 10. The box plots of CAR, F1 score, and AUC are illustrated in Figures 11–13.

TABLE 6: Wilcoxon signed rank test results with F1 score index.

	MPA-SVM-13B	MPA-SVM-NDWI	MPA-SVM-NDVI	MPA-SVM-SAVI	MPA-SVM-MTCI
MPA-SVM-13B	0.00000	0.00009	0.00010	0.00010	0.00009
MPA-SVM-NDWI	0.00009	0.00000	0.70891	0.10843	0.00009
MPA-SVM-NDVI	0.00010	0.70891	0.00000	0.31346	0.00009
MPA-SVM-SAVI	0.00010	0.10843	0.31346	0.00000	0.00009
MPA-SVM-MTCI	0.00009	0.00009	0.00009	0.00009	0.00000

TABLE 7: Wilcoxon signed rank test results with AUC index.

	MPA-SVM-13B	MPA-SVM-NDWI	MPA-SVM-NDVI	MPA-SVM-SAVI	MPA-SVM-MTCI
MPA-SVM-13B	0.00000	0.00009	0.00009	0.00009	0.00009
MPA-SVM-NDWI	0.00009	0.00000	0.00132	0.02762	0.00642
MPA-SVM-NDVI	0.00009	0.00132	0.00000	0.06195	0.00010
MPA-SVM-SAVI	0.00009	0.02762	0.06195	0.00000	0.00010
MPA-SVM-MTCI	0.00009	0.00642	0.00010	0.00010	0.00000

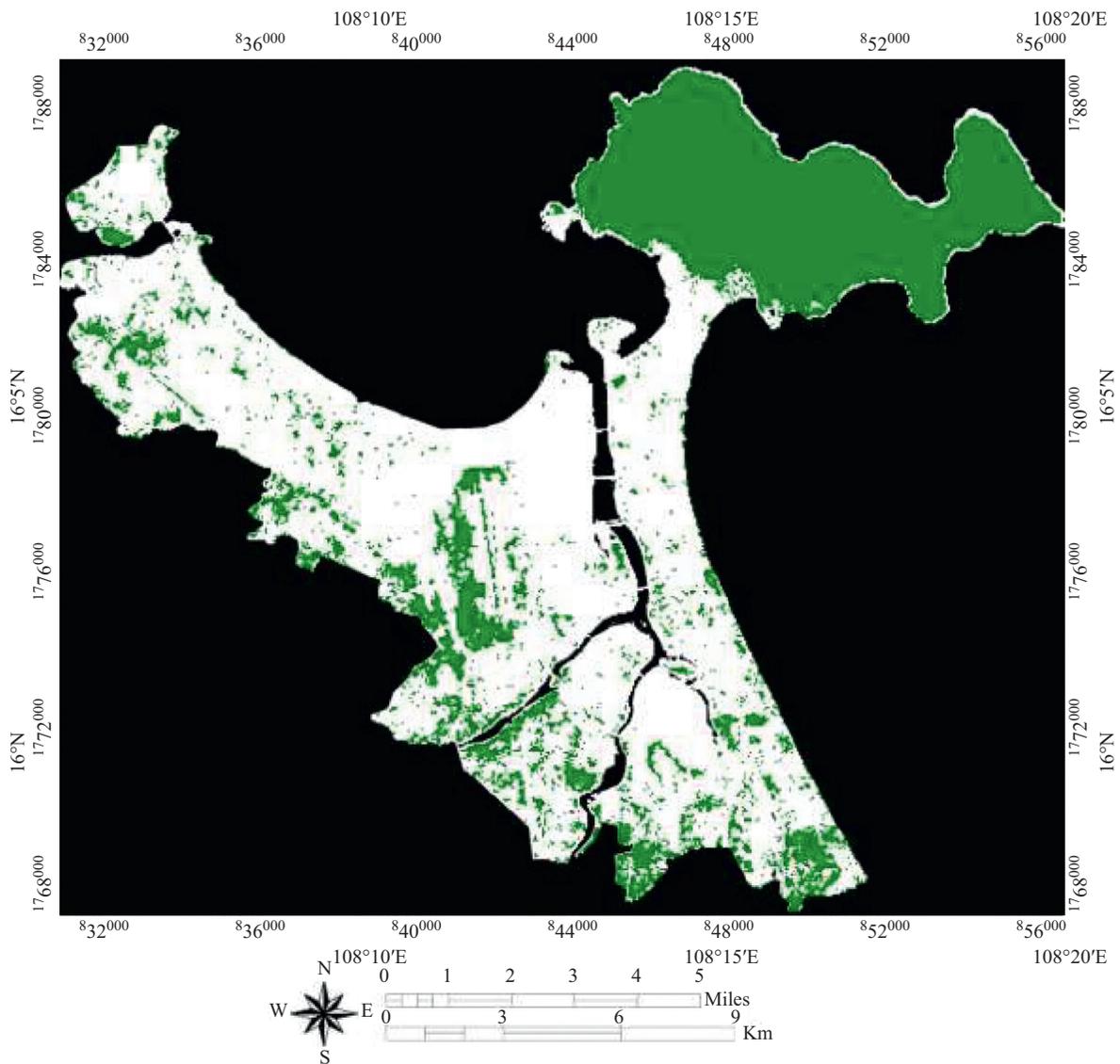


FIGURE 14: Urban green space detection map of the study area.

In addition, to confirm the superiority of the proposed MPA-SVM model that employs all of the Sentinel 2's spectral bands, the Wilcoxon signed-rank test [100] with the significant level (p value) = 0.05 is utilized in this study to express the statistical significance of the model performance indices. The test outcomes of pair-wise model comparison with respect to CAR, F1 score, and AUC are shown in Tables 5–7, respectively. Observably, with p values <0.05, the null hypotheses of insignificant model performances can be rejected. Therefore, MPA-SVM-13B is confirmed to be the best model which provides the classification performance on the collected dataset. Accordingly, the MPA-SVM-13B model is employed to construct an urban green space map for the whole study area. The mapping outcome is demonstrated in Figure 14. Based on the constructed map, it can be found that the green space occupies roughly 34.40% of the study area. Nevertheless, the green space is not evenly distributed in Da Nang. The majority of the green space is located in Son Tra peninsula within the Son Tra district.

5. Concluding Remarks

Urban green space plays a crucial role in improving the living quality of urban environment and has a positive effect on citizens' physical/mental health. Nevertheless, few researches have been dedicated to detecting, locating, and quantifying green space in the study of Da Nang urban center. This study is an attempt to fill this knowledge gap by developing a remote sensing and data-driven approach for urban green space detection applied in the study area. Remotely sensed data obtained from the Sentinel 2 satellite are used to train and validate a hybrid metaheuristic-machine learning approach of MPA-SVM. This hybrid method is employed to construct a decision boundary that separates the input space into two distinctive regions of green space and nongreen space.

The experimental results supported by the Wilcoxon signed-rank test show that the MPA-SVM model employing all of the spectral bands is superior to those of the models relying on individual vegetation indices. Good green space detection results with CAR = 93.100%, precision = 0.916, recall = 0.947, NPV = 0.947, F1 score = 0.931, and AUC = 0.979 demonstrate that the proposed method is highly suited for the task at hand. Moreover, the MPA metaheuristic is confirmed to be a capable method for optimizing machine learning models. Accordingly, the green space mapping of the entire study area can be constructed by the proposed hybrid approach. The information provided by the newly developed model can be helpful for local authority to evaluate the status of green spaces in Da Nang city.

Although MPA-SVM has attained a good predictive performance in urban green space mapping in the study area, the proposed approach also has several limitations. The first limitation is that the MPA-SVM model has not been integrated with feature selection algorithms used for dimensionality reduction. In addition, although the RBFK is widely used for SVM-based pattern recognition, the effectiveness of other sophisticated kernel functions (e.g., hybrid

kernel functions [101, 102]) in urban green space detection should be investigated. Accordingly, future extensions of the current study may include the following:

- (i) Investigating other state-of-the-art metaheuristic algorithms used for optimizing data-driven urban green space detection
- (ii) Studying the effects of the maximum number of searching iterations and the number of population members on the performance of the SVM-based urban green space detection models
- (iii) Employing other advanced texture descriptors to further meliorate the detection accuracy
- (iv) Performing detection tasks at different time periods to inspect changes and trends in urban green space
- (v) Performing urban green space detection using high-resolution satellite images
- (vi) Incorporating advanced feature selection algorithms and kernel functions into the current model structure

Data Availability

The dataset used to support the findings of this study has been deposited in the repository of GitHub (<https://github.com/NDHoangDTU/MPA-SVM-UGSD>).

Conflicts of Interest

The authors confirm that there are no conflicts of interest.

Acknowledgments

This research was funded by the Vietnam National Foundation for Science and Technology Development (NAFOSTED) under grant no. 105.99-2019.339.

References

- [1] G. Xian, M. Crane, and D. Steinwand, "Dynamic modeling of Tampa Bay urban development using parallel computing," *Computers & Geosciences*, vol. 31, pp. 920–928, 2005.
- [2] T. Chen, W. Lang, and X. Li, "Exploring the impact of urban green space on residents' health in guangzhou," *China Journal of Urban Planning and Development*, vol. 146, Article ID 05019022, 2020.
- [3] D. T. Do, J. Huang, Y. Cheng, and T. C. T. Truong, "Da Nang green space system planning," *An Ecology Landscape Approach Sustainability*, vol. 10, p. 3506, 2018.
- [4] I. Jarvis, S. Gergel, M. Koehoorn, and M. van den Bosch, "Greenspace access does not correspond to nature exposure: measures of urban natural space with implications for health research," *Landscape and Urban Planning*, vol. 194, Article ID 103686, 2020.
- [5] Ritchie H., Roser M. (2019) Urbanization Our World in Data <https://ourworldindata.org/>.
- [6] M. A. Benedict and E. T. McMahon, *Green Infrastructure: Linking Landscapes and Communities Urban Land*, Island Press, Washington, DC, USA, 2006.
- [7] M. Atasoy, "Monitoring the urban green spaces and landscape fragmentation using remote sensing: a case study in

- Osmaniye, Turkey,” *Environmental Monitoring and Assessment*, vol. 190, no. 12, p. 713, 2018.
- [8] C. Bertram and K. Rehdanz, “The role of urban green space for human well-being,” *Ecological Economics*, vol. 120, pp. 139–152, 2015.
- [9] A. Callaghan, G. McCombe, A. Harrold et al., “The impact of green spaces on mental health in urban settings: a scoping review,” *Journal of Mental Health*, pp. 1–15, 2020, inpress.
- [10] J. Davtalab, S. P. Deyhimi, V. Dessi, M. R. Hafezi, and M. Adib, “The impact of green space structure on physiological equivalent temperature index in open space,” *Urban Climate*, vol. 31, Article ID 100574, 2020.
- [11] K. De Ridder, “An integrated methodology to assess the benefits of urban green space,” *Science of The Total Environment*, vol. 334–335, pp. 489–497, 2004.
- [12] K. K. Lwin and Y. Murayama, “Modelling of urban green space walkability: eco-friendly walk score calculator,” *Computers, Environment and Urban Systems*, vol. 35, pp. 408–420, 2011.
- [13] R. Rafiee, A. Salman Mahiny, and N. Khorasani, “Assessment of changes in urban green spaces of Mashad city using satellite data,” *International Journal of Applied Earth Observation and Geoinformation*, vol. 11, pp. 431–438, 2009.
- [14] D.-h Shin and K.-s Lee, “Use of remote sensing and geographical information systems to estimate green space surface-temperature change as a result of urban expansion,” *Landscape and Ecological Engineering*, vol. 1, pp. 169–176, 2005.
- [15] <https://www.euro.who.int/WHO> (2021) Urban Green Spaces: A Brief for Action World Health Organization.
- [16] R. M. R. Turaga, U. Jha-Thakur, S. Chakrabarti, and D. Hossain, “Exploring the role of urban green spaces in ‘smartening’ cities in India,” *Impact Assessment and Project Appraisal*, vol. 38, pp. 479–490, 2020.
- [17] E. Barbierato, I. Bernetti, I. Capecchi, and C. Saragosa, “Integrating remote sensing and street view images to quantify urban forest ecosystem services,” *Remote Sensing*, vol. 12, no. 2, p. 329, 2020.
- [18] K. Chapi, V. P. Singh, A. Shirzadi et al., “A novel hybrid artificial intelligence approach for flood susceptibility assessment,” *Environmental Modelling & Software*, vol. 95, pp. 229–245, 2017.
- [19] N.-D. Hoang, Q.-L. Nguyen, and D. T. Bui, “Image processing-based classification of asphalt pavement cracks using support vector machine optimized by artificial bee colony,” *Journal of Computing in Civil Engineering*, vol. 32, Article ID 04018037, 2018.
- [20] P.-T. T. Ngo, T. D. Pham, N. D. Hoang et al., “A new hybrid equilibrium optimized SysFor based geospatial data mining for tropical storm-induced flash flood susceptible mapping,” *Journal of Environmental Management*, vol. 280, Article ID 111858, 2021.
- [21] V.-H. Nhu, “A new hybrid firefly-PSO optimized random subspace tree intelligence for torrential rainfall-induced flash flood susceptible mapping,” *Remote Sensing*, vol. 12, p. 2688, 2020.
- [22] S. A. Yeprintsev, M. A. Klevtsova, L. A. Lepeshkina, S. V. Shekoyan, and A. A. Voronin, “Assessment of the dynamics of urbanized areas by remote sensing,” *IOP Conference Series: Earth and Environmental Science*, vol. 115, Article ID 012034, 2018.
- [23] X. Cao, A. Onishi, J. Chen, and H. Imura, “Quantifying the cool island intensity of urban parks using ASTER and IKONOS data,” *Landscape and Urban Planning*, vol. 96, no. 4, pp. 224–231, 2010.
- [24] W. Kuang and Y. Dou, “Investigating the patterns and dynamics of urban green space in China’s 70 major cities using satellite,” *Remote Sensing Remote Sensing*, vol. 12, p. 1929, 2020.
- [25] J. Li, C. Song, L. Cao, F. Zhu, X. Meng, and J. Wu, “Impacts of landscape structure on surface urban heat islands: a case study of Shanghai,” *China Remote Sensing of Environment*, vol. 115, pp. 3249–3263, 2011.
- [26] S. Oliveira, H. Andrade, and T. Vaz, “The cooling effect of green spaces as a contribution to the mitigation of urban heat,” *A Case Study in Lisbon Building and Environment*, vol. 46, pp. 2186–2194, 2011.
- [27] V. Sathyakumar, R. Ramsankaran, and R. Bardhan, “Geospatial approach for assessing spatiotemporal dynamics of urban green space distribution among neighbourhoods: a demonstration in Mumbai,” *Urban Forestry & Urban Greening*, vol. 48, Article ID 126585, 2020.
- [28] N. H. Wong, S. Kardinal Jusuf, A. Aung La Win, H. Kyaw Thu, T. Syatia Negara, and W. Xuchao, “Environmental study of the impact of greenery in an institutional campus in the tropics,” *Building and Environment*, vol. 42, pp. 2949–2970, 2007.
- [29] A. El Garouani, D. J. Mulla, S. El Garouani, and J. Knight, “Analysis of urban growth and sprawl from remote sensing data: case of Fez,” *Morocco International Journal of Sustainable Built Environment*, vol. 6, pp. 160–169, 2017.
- [30] W. Li, R. Dong, H. Fu, J. Wang, L. Yu, and P. Gong, “Integrating Google Earth imagery with Landsat data to improve 30-m resolution land cover mapping,” *Remote Sensing of Environment*, vol. 237, Article ID 111563, 2020.
- [31] S. Dinda, N. Das Chatterjee, and S. Ghosh, “An integrated simulation approach to the assessment of urban growth pattern and loss in urban green space in Kolkata, India,” *A GIS-Based Analysis Ecological Indicators*, vol. 121, Article ID 107178, 2021.
- [32] K. Bhosle and V. Musande, “Evaluation of deep learning CNN model for land use land cover classification and crop identification using hyperspectral remote sensing images,” *Journal of the Indian Society of Remote Sensing*, vol. 47, no. 11, pp. 1949–1958, 2019.
- [33] M. Campos-Taberner, F. J. García-Haro, B. Martínez et al., “Understanding deep learning in land use classification based on Sentinel-2 time series,” *Scientific Reports*, vol. 10, no. 1, p. 17188, 2020.
- [34] M. Carranza-García, J. García-Gutiérrez, and J. Riquelme, “A framework for evaluating land use and land cover classification using convolutional neural networks,” *Remote Sensing*, vol. 11, no. 3, p. 274, 2019.
- [35] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning (Adaptive Computation and Machine Learning Series)*, The MIT Press, Cambridge, MA, USA, 2016.
- [36] G. M. Hadjidemetriou, P. A. Vela, and S. E. Christodoulou, “Automated pavement patch detection and quantification using support vector machines,” *Journal of Computing in Civil Engineering*, vol. 32, Article ID 04017073, 2018.
- [37] Q. He, Q. Zhang, H. Wang, and C. Zhang, “Local similarity-based fuzzy multiple kernel one-class support vector machine,” *Complexity*, vol. 2020, Article ID 8853277, 12 pages, 2020.
- [38] X. Li, Y. Li, Y. Zhang, F. Liu, and Y. Fang, “Fault diagnosis of belt conveyor based on support vector machine and grey

- wolf,” *Optimization Mathematical Problems in Engineering*, vol. 2020, Article ID 1367078, 10 pages, 2020.
- [39] Z. Zhao, T. Liu, and X. Zhao, “Variable selection from image texture feature for automatic classification of concrete surface voids,” *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 5538573, 10 pages, 2021.
- [40] V. N. Vapnik, *Statistical Learning Theory*, John Wiley & Sons, Inc, Hoboken, NJ, USA, 1998.
- [41] C. M. Bishop, *Pattern Recognition and Machine Learning (Information Science and Statistics)*, Springer, Berlin, Germany, 2011.
- [42] D. Prayogo, M.-Y. Cheng, Y.-W. Wu, and D.-H. Tran, “Combining machine learning models via adaptive ensemble weighting for prediction of shear capacity of reinforced-concrete deep beams,” *Engineering with Computers*, vol. 36, pp. 1135–1153, 2019.
- [43] N. Kardani, A. Zhou, M. Nazem, and S.-L. Shen, “Estimation of bearing capacity of piles in cohesionless soil using optimized machine learning approaches,” *Geotechnical and Geological Engineering*, vol. 38, pp. 2271–2291, 2020.
- [44] H. Nguyen, N.-M. Nguyen, M.-T. Cao, N.-D. Hoang, and X.-L. Tran, “Prediction of long-term deflections of reinforced-concrete members using a novel swarm optimized extreme gradient boosting machine,” *Engineering with Computers*, 2021.
- [45] J.-S. Chou and N.-M. Nguyen, “Metaheuristics-optimized ensemble system for predicting mechanical strength of reinforced concrete materials,” *Structural Control and Health Monitoring*, vol. 28, no. 5, p. e2706, 2021.
- [46] Z. Yu, X. Shi, J. Zhou, X. Chen, and X. Qiu, “Effective assessment of blast-induced ground vibration using an optimized random forest model based on a Harris hawks optimization,” *Algorithm Applied Sciences*, vol. 10, p. 1403, 2020.
- [47] W. Yong, J. Zhou, D. Jahed Armaghani et al., “A new hybrid simulated annealing-based genetic programming technique to predict the ultimate bearing capacity of piles,” *Engineering with Computers*, 2020.
- [48] A. Sina and D. Kaur, “Short term load forecasting model based on kernel-support vector regression with social spider optimization algorithm,” *Journal of Electrical Engineering & Technology*, vol. 15, pp. 393–402, 2020.
- [49] D. Tuan Vu, X.-L. Tran, M.-T. Cao, T. Cuong Tran, and N.-D. Hoang, “Machine learning based soil erosion susceptibility prediction using social spider algorithm optimized multivariate adaptive regression,” *Spline Measurement*, vol. 164, Article ID 108066, 2020.
- [50] N.-T. Ngo, H. A. Le, and T.-P.-T. Pham, “Integration of support vector regression and grey wolf optimization for estimating the ultimate bearing capacity in concrete-filled steel tube columns,” *Neural Computing and Applications*, 2021.
- [51] W. Chen, X. Chen, J. Peng, M. Panahi, and S. Lee, “Landslide susceptibility modeling based on ANFIS with teaching-learning-based optimization and Satin bowerbird optimizer,” *Geoscience Frontiers*, vol. 12, no. 1, pp. 93–107, 2021.
- [52] E. Li, J. Zhou, X. Shi et al., “Developing a hybrid model of salp swarm algorithm-based support vector machine to predict the strength of fiber-reinforced cemented paste backfill,” *Engineering with Computers*, 2020.
- [53] Z. M. Yaseen, H. Faris, and N. Al-Ansari, “Hybridized extreme learning machine model with salp swarm algorithm,” *A Novel Predictive Model for Hydrological Application Complexity*, vol. 2020, Article ID 8206245, 14 pages, 2020.
- [54] Z.-Z. Zhu, Y.-W. Feng, C. Lu, and C.-W. Fei, “Reliability optimization of structural deformation with improved support vector regression,” *Model Advances in Materials Science and Engineering*, vol. 2020, Article ID 3982450, 8 pages, 2020.
- [55] S. Fan, S. Cao, and Y. Zhang, “Temperature prediction of photovoltaic panels based on support vector machine with pigeon-inspired optimization,” *Complexity*, vol. 2020, Article ID 9278162, 12 pages, 2020.
- [56] A. Gholami, S. M. Seyedali, and H. R. Ansari, “Estimation of shear wave velocity from post-stack seismic data through committee machine with cuckoo search optimized intelligence models,” *Journal of Petroleum Science and Engineering*, vol. 189, Article ID 106939, 2020.
- [57] Q. Fang, H. Nguyen, X.-N. Bui, and T. Nguyen-Thoi, “Prediction of blast-induced ground vibration in open-pit mines using a new technique based on imperialist competitive algorithm and M5Rules,” *Natural Resources Research*, vol. 29, no. 9, 2019.
- [58] J. Zhou, Y. Qiu, S. Zhu et al., “Optimization of support vector machine through the use of metaheuristic algorithms in forecasting TBM advance rate,” *Engineering Applications of Artificial Intelligence*, vol. 97, Article ID 104015, 2021.
- [59] Z. Yu, X. Shi, J. Zhou et al., “Feasibility of the indirect determination of blast-induced rock movement based on three new hybrid intelligent models,” *Engineering with Computers*, vol. 37, pp. 991–1006, 2021.
- [60] A. Faramarzi, M. Heidarinejad, S. Mirjalili, and A. H. Gandomi, “Marine predators algorithm: a nature-inspired metaheuristic,” *Expert Systems with Applications*, vol. 152, Article ID 113377, 2020.
- [61] J. W. Rouse, R. H. Haas, J. A. Schell, and D. W. Deering, “Monitoring Vegetation Systems in the Great Plains with ERTS,” in *Conference Proceedings of the Third ERTS Symposium, NASA SP-351*, pp. 309–317, Washington DC, USA, December 1973.
- [62] B.-c Gao, “NDWI—a normalized difference water index for remote sensing of vegetation liquid water from space,” *Remote Sensing of Environment*, vol. 58, pp. 257–266, 1996.
- [63] A. R. Huete, “A soil-adjusted vegetation index (SAVI),” *Remote Sensing of Environment*, vol. 25, pp. 295–309, 1988.
- [64] J. Dash and P. J. Curran, “The MERIS terrestrial chlorophyll index,” *International Journal of Remote Sensing*, vol. 25, no. 23, pp. 5403–5413, 2004.
- [65] MPI (2020) Da Nang City Ministry of Planning and Investment <http://www.mpigovvn/Pages/tinhthanhhचितiet.aspx?idTinhThanh=41>.
- [66] USGS (2020) Earth Explorer US Department of the Interior (<https://earthexplorer.usgs.gov/>).
- [67] ESA (2020) Sentinel Application Platform (SNAP) The European Space Agency <http://stepesaint/main/toolboxes/snap/>.
- [68] ENVI (2021) ENVI Tutorials Harris Geospatial Solutions, <https://www13harrisgeospatial.com/docs/tutorialshtml>.
- [69] J. R. Jensen, *Introductory Digital Image Processing A Remote Sensing Perspective*, Prentice Hall Press, Upper Saddle River, NJ, USA, 2015.
- [70] K. Abutaleb, M. Freddy Mudede, N. Nkongolo, and S. W. Newete, “Estimating urban greenness index using remote sensing data: a case study of an affluent vs poor suburbs in the city of Johannesburg,” *The Egyptian Journal of Remote Sensing and Space Science*, 2020, in press.
- [71] W. Chen, H. Huang, J. Dong, Y. Zhang, Y. Tian, and Z. Yang, “Social functional mapping of urban green space using

- remote sensing and social sensing data,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 146, pp. 436–452, 2018.
- [72] L. Rumora, I. Majić, M. Miler, and D. Medak, “Spatial video remote sensing for urban vegetation mapping using vegetation indices,” *Urban Ecosystems*, vol. 24, pp. 21–33, 2021.
- [73] T.-X. Zhang, J.-Y. Su, C.-J. Liu, and W.-H. Chen, “Potential bands of sentinel-2A satellite for classification problems in precision agriculture,” *International Journal of Automation and Computing*, vol. 16, pp. 16–26, 2019.
- [74] K. W. Pałaś and J. Zawadzki, “Sentinel-2 imagery processing for tree logging observations on the Białowieża forest world heritage site,” *Forests*, vol. 11, p. 857, 2020.
- [75] I. Kamenova and P. Dimitrov, “Evaluation of Sentinel-2 vegetation indices for prediction of LAI, fAPAR and fCover of winter wheat in Bulgaria,” *European Journal of Remote Sensing*, vol. 54, pp. 1–20, 2020.
- [76] F. Bartumeus, J. Catalan, U. L. Fulco, M. L. Lyra, and G. M. Viswanathan, “Optimizing the encounter rate in biological interactions: Lévy versus brownian strategies,” *Physical Review Letters*, vol. 88, no. 9, Article ID 097901, 2002.
- [77] N. E. Humphries, N. Queiroz, J. R. M. Dyer et al., “Environmental context explains Lévy and Brownian movement patterns of marine predators,” *Nature*, vol. 465, pp. 1066–1069, 2010.
- [78] G. M. Viswanathan, E. P. Raposo, and M. G. E. da Luz, “Lévy flights and superdiffusion in the context of biological encounters and random searches,” *Physics of Life Reviews*, vol. 5, pp. 133–150, 2008.
- [79] J. D. Filmlalter, L. Dagorn, P. D. Cowley, and M. Taquet, “First descriptions of the behavior of silky sharks, *Carcharhinus falciformis*, around drifting fish aggregating devices in the Indian Ocean,” *Bulletin of Marine Science*, vol. 87, pp. 325–337, 2011.
- [80] K. O. Achieng, “Modelling of soil moisture retention curve using machine learning techniques: artificial and deep neural networks vs support vector regression models,” *Computers & Geosciences*, vol. 133, Article ID 104320, 2019.
- [81] W. L. Al-Yaseen, Z. A. Othman, and M. Z. A. Nazri, “Multi-level hybrid support vector machine and extreme learning machine based on modified K-means for intrusion detection system,” *Expert Systems with Applications*, vol. 67, pp. 296–303, 2017.
- [82] W. Chen, Z. Shang, and Y. Chen, “A novel hybrid network traffic prediction approach based on support vector machines,” *Journal of Computer Networks and Communications*, vol. 2019, Article ID 2182803, 10 pages, 2019.
- [83] S. Deng, X. Wang, Y. Zhu, F. Lv, and J. Wang, “Hybrid grey wolf optimization algorithm based support vector machine for groutability prediction of fractured rock,” *Mass Journal of Computing in Civil Engineering*, vol. 33, Article ID 04018065, 2019.
- [84] K.-W. Liao, N.-D. Hoang, and S.-C. Chang, “Estimating landslide occurrence via small watershed method with relevance vector machine,” *Earth Science Informatics*, vol. 13, pp. 249–260, 2019.
- [85] C. Liu, H. Li, and Q. Zhang, “Research on sewage monitoring and water quality prediction based on wireless sensors and support vector machines,” *Wireless Communications and Mobile Computing*, vol. 2020, Article ID 8852965, 10 pages, 2020.
- [86] A. Malik, Y. Tikhmarine, D. Souag-Gamane, O. Kisi, and Q. B. Pham, “Support vector regression optimized by metaheuristic algorithms for daily streamflow prediction,” *Stochastic Environmental Research and Risk Assessment*, vol. 34, 2020.
- [87] N. Naicker, T. Adeliyi, and J. Wing, “Linear support vector machines for prediction of student performance in school-based education,” *Mathematical Problems in Engineering*, vol. 2020, Article ID 4761468, 7 pages, 2020.
- [88] T.-D. Nguyen, T.-H. Tran, H. Nguyen, and H. Nhat-Duc, “A success history-based adaptive differential evolution optimized support vector regression for estimating plastic viscosity of fresh concrete,” *Engineering with Computers*, vol. 37, pp. 1485–1498, 2019.
- [89] P. Saha, P. Debnath, and P. Thomas, “Prediction of fresh and hardened properties of self-compacting concrete using support vector regression approach,” *Neural Computing and Applications*, vol. 32, pp. 7995–8010, 2019.
- [90] H. Tanyildizi, “Prediction of the strength properties of carbon fiber-reinforced lightweight concrete exposed to the high temperature using artificial neural network and support vector machine,” *Advances in Civil Engineering*, vol. 2018, Article ID 5140610, 10 pages, 2018.
- [91] L. H. Hamel, *Knowledge Discovery with Support Vector Machines*, John Wiley & Sons, Inc., Hoboken, NJ, USA, 2009.
- [92] M.-Y. Cheng and N.-D. Hoang, “Typhoon-induced slope collapse assessment using a novel bee colony optimized support vector classifier,” *Natural Hazards*, vol. 78, no. 3, pp. 1961–1978, 2015.
- [93] S. Theodoridis and K. Koutroubas, *Pattern Recognition*, Academic Press, Cambridge, MA, USA, 2009.
- [94] N.-D. Hoang, “Image processing based automatic recognition of asphalt pavement patch using a metaheuristic optimized machine learning approach,” *Advanced Engineering Informatics*, vol. 40, pp. 110–120, 2019.
- [95] N.-D. Hoang and D. T. Bui, “A novel relevance vector machine classifier with cuckoo search optimization for spatial prediction of landslides,” *Journal of Computing in Civil Engineering*, vol. 30, Article ID 04016001, 2016.
- [96] T. V. Dinh, H. Nguyen, X.-L. Tran, and N.-D. Hoang, “Predicting rainfall-induced soil erosion based on a hybridization of adaptive differential evolution and support vector machine classification,” *Mathematical Problems in Engineering*, vol. 2021, Article ID 6647829, 14 pages, 2021.
- [97] Accord (2019) Accord.NET Framework <http://accord-framework.net/>.
- [98] V. López, A. Fernández, S. García, V. Palade, and F. Herrera, “An insight into classification with imbalanced data: empirical results and current trends on using data intrinsic characteristics,” *Information Sciences*, vol. 250, pp. 113–141, 2013.
- [99] A. R. van Erkel and P. M. T. Pattynama, “Receiver operating characteristic (ROC) analysis: basic principles and applications in radiology,” *European Journal of Radiology*, vol. 27, pp. 88–94, 1998.
- [100] M. Hollander and D. A. Wolfe, *Nonparametric Statistical Methods*, John Wiley & Sons, Hoboken, NJ, USA, 1999.
- [101] D. Segera, M. Mbuthia, and A. Nyete, “Particle swarm optimized hybrid kernel-based multiclass support vector machine for microarray cancer data analysis,” *BioMed Research International*, vol. 2019, Article ID 4085725, 11 pages, 2019.
- [102] X. Wu, W. Tang, and X. Wu, “Support vector machine based on hybrid kernel function,” *Information Engineering and Applications*, Springer, London, UK, pp. 127–133, 2012.