

Article

Mapping Heterogeneous Urban Landscapes from the Fusion of Digital Surface Model and Unmanned Aerial Vehicle-Based Images Using Adaptive Multiscale Image Segmentation and Classification

Mohamed Barakat A. Gibril ^{1,2} , Bahareh Kalantar ^{3,*} , Rami Al-Ruzouq ^{1,4} , Naonori Ueda ³,
Vahideh Saeidi ⁵, Abdallah Shanableh ^{1,4}, Shattri Mansor ² and Helmi Z. M. Shafri ²

¹ GIS & Remote Sensing Center, Research Institute of Sciences and Engineering, University of Sharjah, 27272 Sharjah, UAE; mbgibril@sharjah.ac.ae (M.B.A.G.); ralruzouq@sharjah.ac.ae (R.A.-R.); shanableh@sharjah.ac.ae (A.S.)

² Department of Civil Engineering, Faculty of Engineering, Universiti Putra Malaysia, Serdang, 43400 Selangor, Malaysia; shattri@upm.edu.my (S.M.); helmi@upm.edu.my (H.Z.M.S.)

³ RIKEN Center for Advanced Intelligence Project, Goal-Oriented Technology Research Group, Disaster Resilience Science Team, Tokyo 103-0027, Japan; naonori.ueda@riken.jp

⁴ Department of Civil and Environmental Engineering, University of Sharjah, 27272 Sharjah, UAE

⁵ Department of Mapping and Surveying, Darya Tarsim Consulting Engineers Co. Ltd., 1457843993 Tehran, Iran; saeidi@daryatarsim.com

* Correspondence: bahareh.kalantar@riken.jp; Tel.: +81-362-252-482

Received: 3 March 2020; Accepted: 24 March 2020; Published: 27 March 2020



Abstract: Considering the high-level details in an ultrahigh-spatial-resolution (UHSR) unmanned aerial vehicle (UAV) dataset, detailed mapping of heterogeneous urban landscapes is extremely challenging because of the spectral similarity between classes. In this study, adaptive hierarchical image segmentation optimization, multilevel feature selection, and multiscale (MS) supervised machine learning (ML) models were integrated to accurately generate detailed maps for heterogeneous urban areas from the fusion of the UHSR orthomosaic and digital surface model (DSM). The integrated approach commenced through a preliminary MS image segmentation parameter selection, followed by the application of three supervised ML models, namely, random forest (RF), support vector machine (SVM), and decision tree (DT). These models were implemented at the optimal MS levels to identify preliminary information, such as the optimal segmentation level(s) and relevant features, for extracting 12 land use/land cover (LULC) urban classes from the fused datasets. Using the information obtained from the first phase of the analysis, detailed MS classification was iteratively conducted to improve the classification accuracy and derive the final urban LULC maps. Two UAV-based datasets were used to develop and assess the effectiveness of the proposed framework. The hierarchical classification of the pilot study area showed that the RF was superior with an overall accuracy (OA) of 94.40% and a kappa coefficient (K) of 0.938, followed by SVM (OA = 92.50% and K = 0.917) and DT (OA = 91.60% and K = 0.908). The classification results of the second dataset revealed that SVM was superior with an OA of 94.45% and K of 0.938, followed by RF (OA = 92.46% and K = 0.916) and DT (OA = 90.46% and K = 0.893). The proposed framework exhibited an excellent potential for the detailed mapping of heterogeneous urban landscapes from the fusion of UHSR orthophoto and DSM images using various ML models.

Keywords: unmanned aerial vehicle; urban LULC; GEOBIA; multiscale classification

1. Introduction

Land use/land cover (LULC) maps play an indispensable part in gaining comprehensive insights into coupled human–environment systems, socioecological concerns, resource inventories, ecosystem management, planning activities, change monitoring, emergency response, and decision making. For instance, high-quality thematic LULC information is an essential input for versatile local and regional applications, such as natural disasters [1,2], agriculture [3], sustainable development [4], and land use suitability and management [5]. Therefore, producing accurate, up-to-date, and cost-efficient detailed LULC maps is crucial for resource managers, scientists, decision makers, and city planners.

Remote sensing technologies have been extensively used to retrieve LULC information using comprehensive options of platforms and sensors with versatile spatial, spectral, and temporal resolutions. Satellite and airborne remotely sensed data are usually expensive and constrained by the inability to deliver adequate spatial and temporal resolutions compared to drone-based data. Nowadays, unmanned aerial vehicles (UAVs) are used to collect remotely sensed data in a cost-effective manner at low altitudes below the cloud cover with ultrahigh spatial (UHSR) spectral and temporal resolutions. These advantages make the UAV system a powerful tool that can be used to fulfil the rapid monitoring and assessment during a natural disaster and real-time monitoring applications [6,7]. A plethora of studies have successfully used UAV platforms to acquire remotely sensed data for LULC applications [7–13].

Geographic object-based image analysis (GEOBIA), a paradigm that imitates the human visual perception of real-world targets by addressing the spectral variability amongst classes, has been a preferable classification approach because of its advantages over pixel-based classification [14,15]. The limitations of pixel-based approaches, such as misclassification and salt-and-pepper effects, are addressed through the hierarchical/multiscale (MS) exemplification of image objects, representation of image objects across single/multiple images at MS levels, and incorporation of spatial, spectral, textural, geometrical, elevation, backscattering, and contextual information in LULC classification [16,17]. GEOBIA has been widely used along with advanced machine learning (ML) algorithms for analyzing and classifying drone-based images in various applications. De Castro et al. [18] suggested an automatic GEOBIA approach using a random forest (RF) classifier for site-specific weed management with UAV-based images. Their results helped the farmers with timely decision making for crop optimization and management. Komárek et al. [19] utilized a three-level GEOBIA system with a support vector machine (SVM) algorithm to identify individual plant species from multispectral and thermal drone-based images. Kamal et al. [20] introduced a GEOBIA approach for mangrove canopy delineation using UAV-based data. The results showed that the UAV red, green, and blue (RGB) images are valuable inputs for GEOBIA regardless of the limit of spectral information. Mishra et al. [21] presented the potential for achieving species-level mapping from multispectral UAV data through GEOBIA. White et al. [22] proposed a GEOBIA approach to identify sapling Jak pine forests after wildfire. A MicaSense RedEdge 3 multispectral camera onboard a quadcopter UAV platform was used for data acquisition. The highest classification accuracy was achieved by including the red and near-infrared spectra. The following section reviews the various elements affecting the GEOBIA's overall quality, including image optimization of image segmentation parameters, several feature selection (FS) approaches, and machine learning (ML) algorithms.

Related Studies

GEOBIA is constructed on the basis of the conception of creating a meaningful representation of real-world targets (i.e., LULC types, such as buildings, roads, and vegetation) by generating homogenous regions from image pixels, and this procedure is referred to as image segmentation. This procedure groups the raw pixels into homogeneous segments that are jointly exhaustive and mutually disjointed, which are then used as the primary element for interpretation, classification, and modelling [23,24]. The overall performance of the GEOBIA's succeeding phases (i.e., feature computation, extraction, and classification) is immensely influenced by image segmentation quality [25].

Several image segmentation algorithms have been adopted to segment remotely sensed data in the GEOBIA domain. The four commonly used image segmentation algorithms are watershed [26,27], region-based [28], mean-shift [29], and hybrid segmentation [30]. Amongst them, the region-based and multiresolution segmentation algorithms have been widely adopted in various remote sensing applications because of their competency to produce meaningful image objects [31]. The scale parameter (SP), which explains the degree or the density level where a specific phenomenon can be presented, is the main parameter in all segmentation algorithms that require fine-tuning depending on the application [28]. The image segmentation on UHSR images with a single-scale (SS) value results in creating image objects that are either small (oversegmented) or large (undersegmented). Different issues should be considered when selecting image segmentation parameters [32]. Firstly, different urban LULC classes can differ in terms of size and structure and may require several optimum segmentation levels. Secondly, image objects that belong to the same class might correspond to different optimal scales because of their different surrounding contrast. Finally, different components within an object might require analysis at multiple scales. Therefore, finding the optimal SP(s) prior to the analysis is a crucial step in the GEOBIA framework because different objects can only be analyzed accurately on the basis of the scale(s) corresponding to their granularity [33].

Various supervised and unsupervised image segmentation quality evaluation techniques have been proposed to determine the optimal single or MS segmentation. Image segmentation results are usually assessed through the supervised image segmentation quality measures by evaluating the disparity between the manually digitized objects and the generated image objects from an image segmentation algorithm, whereas the unsupervised image segmentation quality measures evaluate MS segmentation results using various statistical-based image segmentation quality measures [34]. Considerable attention has been given to unsupervised segmentation quality measures [35,36]. The vast majority of the unsupervised optimization methods determine the optimum segmentation parameters to evaluate the segmentation outputs by computing the between-object heterogeneity and the within-object homogeneity metrics [37–44]. Moreover, other unsupervised techniques have been adopted in various applications to determine the optimum single or MS segmentation parameters. Xiao et al. [33] proposed a MS segmentation optimization using the unsupervised optimization technique to determine the optimum scales suitable for the extraction of urban green cover from the high-spatial-resolution dataset. Kamal et al. [45] mapped mangrove species using the MS GEOBIA approach from multiple images with a varied spatial resolution (i.e., WorldView 2, LiDAR, ALOS AVNIR-2, and Landsat TM).

FS, which is regarded as a crucial task that influences the GEOBIA classification accuracy, specifies the most relevant features to increase the effectiveness of the adopted classification approach and expedites the processing time by minimizing irrelevant or redundant features [46]. Various FS algorithms have been incorporated with the GEOBIA approach in various applications, and these methods include RF [47,48], SVM [47], ant colony optimization (ACO) [49,50], artificial bee colony [51], hybrid particle swarm optimization [52], correlation-based FS (CFS) [49,53], and chi-square [54]. Ridha and Pradhan [49] applied three FS methods, namely, CFS, RF, and ACO, to discriminate several types of landslides from LiDAR data. The results showed that CFS performs the best with 89.28% accuracy, followed by RF with 85.59% accuracy and ACO with 86.74% accuracy. Al-Ruzouq et al. [50] adopted ACO for feature reduction and identification of the most crucial features for date palm mapping from very-high-resolution aerial imageries. The results showed that ACO and CFS are superior to other algorithms, including principal component analysis, SVM, information gain, gain ratio, and chi-square. The effects of various feature importance evaluation methods, including gain ratio, chi-square, SVM-recursive feature elimination, CFS, Relief-F, SVM, and RF were investigated by Ma et al. [47] within the GEOBIA environment to map agricultural areas from UAV data. The results showed that CFS dominates other feature importance evaluation methods.

Recently, GEOBIA has been integrated with various ML methods to classify UAV-based images for various LULC applications. Ma et al. [47] used two ML classification approaches, namely, SVM and RF, to classify the UAV data into six categories, namely, building, crop, bare land, road, water,

and woodland in Deyang, China. RF exhibits a higher classification accuracy compared to SVM. Cao et al. [55] applied two classification algorithms, namely, SVM and k-nearest neighbors (KNN), in the GEOBIA domain to map mangrove species from a UAV hyperspectral image. The result showed that SVM performs better with 89.55% accuracy compared to KNN with 81.70% accuracy. Akar [56] compared various ML algorithms to perform LULC classification using the UAV images collected from urban and rural areas. The results showed that rotation forest (92.52%) outperforms RF (90.52) and gentle AdaBoost (87.52%). Liu et al. [57] proposed a SVM-deep belief network (restricted Boltzmann machine) method to extract eight land cover classes, namely, tree, building 1, road, grass, river, building 2, building 3, and bare-land, using the fusion of LiDAR data and UAV images. Their proposed technique shows an overall accuracy (OA) of 92.16% and a kappa (K) value of 0.904%.

In this study, an adaptive MS segmentation and classification approach was adopted to classify heterogeneous urban areas through the fusion of the UHSR orthophoto and digital surface model (DSM). The main objectives of the current study are to (1) develop an adaptive MS-optimized image object approach for detailed urban LULC mapping from UAV-based data, (2) investigate the effects of MS segmentation on FS computation (CFS and SVM) and its impact on classification accuracy, (3) compare the performance of three mature ML classification algorithms, namely, RF, SVM, and decision tree (DT), at MS levels, and (4) assess the transferability of the adopted framework. The remainder of this paper is organized as follows. Section 2 outlines the geographical location of the study area and describes the ground truth (GT) data. Section 3 presents a generic overview of the methodological framework and detailed information about image processing, image segmentation optimization, FS, MS classification, and evaluation metrics. Section 4 describes the results, and Section 5 discusses the experimental findings. Section 6 provides the conclusions.

2. Study Area and Materials

2.1. Study Area

The location of the study area is geographically positioned at the University of Science, Malaysia (USM) campus, Penang, Malaysia. The study area represents an urban area of Penang island with different LULCs, including vegetation, water bodies, buildings, roads, and bare soil. The RGB images were acquired on February, 2018, (Figure 1) using a Canon PowerShot SX230 HS (4000 × 3000 resolution) boarding on a UAV from an altitude of 353 m. The ground resolution of the orthomosaic is approximately 10 cm, with an 8-bit radiometric resolution. The first dataset was a subset of 2.24 km² from the produced orthomosaic photos, located between 100°18'7.43"E, 5°21'51.574"N and 100°19'2"E, 5°21'8.143"N. A DSM with 0.8 m resolution was generated from 3500 points using Agisoft PhotoScan Professional (version 1.3.4, <http://www.agisoft.com>). The second subset (with the coordinates of 100°17'29.341"E, 5°21'38.6"N and 100°18'9.622"E, 5°21'5.345"N), covering an area of 1.27 km², was selected for investigating the transferability of the methodology.

2.2. GT Data

In the first study area, a total of 1177 GT samples for various urban LULC classes were prepared through field surveys with the aid of Google Earth images. Twelve different classes were identified, including water bodies, bare soil, grass, trees, clay tiles type 1, clay tiles type 2, metallic roofs type 1, metallic roofs type 2, concrete, dark concrete roofs, asbestos cement roofs, and asphalt. The training and testing GT samples were prepared as vector points and meticulously selected to ensure that all the available urban LULC classes are well represented. Table 1 presents several types of LULC classes available in the UAV-based images. Amongst the collected GT samples, 70% of each class was utilized in the training of ML models, and 30% was dedicated for testing them. For the second study area, sample statistics derived from the training samples of the first study area were used in the classification models, and 305 GT testing samples were used to evaluate the classification results.

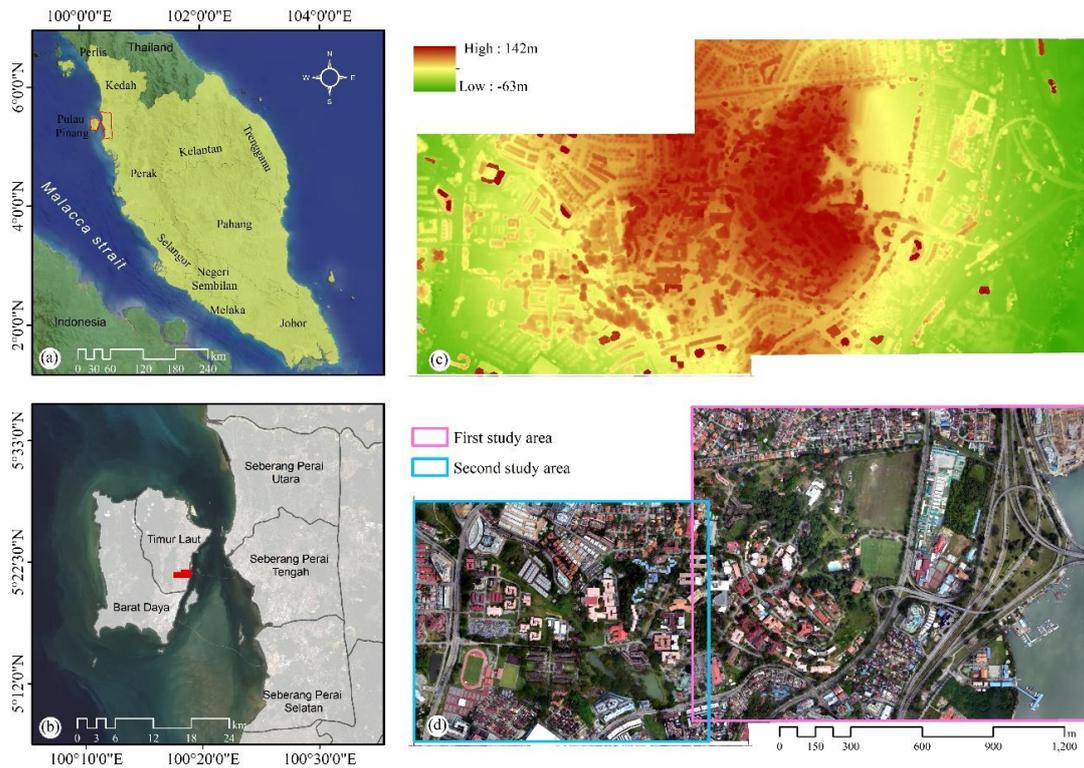


Figure 1. General location of the study sites: (a) Malaysian states; (b) location map; (c) digital surface model (DSM) of the study sites; (d) unmanned aerial vehicle (UAV) images of the study sites.

Table 1. Detailed description of different land use/land cover (LULC) classes available in the UAV-based images.

LULC Type	Images	Description
Water bodies		Water bodies with light blue and green colors
Trees and grass		Various tree species and grass thickness
Bare soil		Exposed soil with different colors
Concrete roofs		Concrete slab with bright white color

Table 1. Cont.

LULC Type	Images	Description
Dark concrete roofs		Residential and industrial buildings with dark brown color
Clay tiles type 1		Roofing material with different structural shapes and red color
Clay tiles type 2		Roofing material with different structures and bright peach color
Asbestos cement roofs		Roofs with regular shape and grey color
Metallic roofs type 1		Metal deck with blue color
Metallic roofs type 2		Rooftops with turquoise color
Roads		Urban roads with grey color

3. Methodology

3.1. Overview

In this study, MS image segmentation optimization, MS feature computation and evaluations, and supervised hierarchical ML models were conducted for accurate detailed mapping of a heterogeneous urban landscape from UAV-based images. As depicted in Figure 2, the adopted methodology comprises five main phases. Firstly, drone-based images were acquired and preprocessed to generate the orthophoto and the DSM. Secondly, the optimum MS segmentation parameters were identified

using unsupervised segmentation quality metrics. Thirdly, the most significant features were selected at MS levels on the basis of CFS and SVM wrapper approaches. Fourthly, adaptive MS segmentation optimization and classification were conducted for detailed urban LULC mapping using the RF, SVM, and DT algorithms. Finally, the transferability of the proposed methodology to a different study area was investigated.

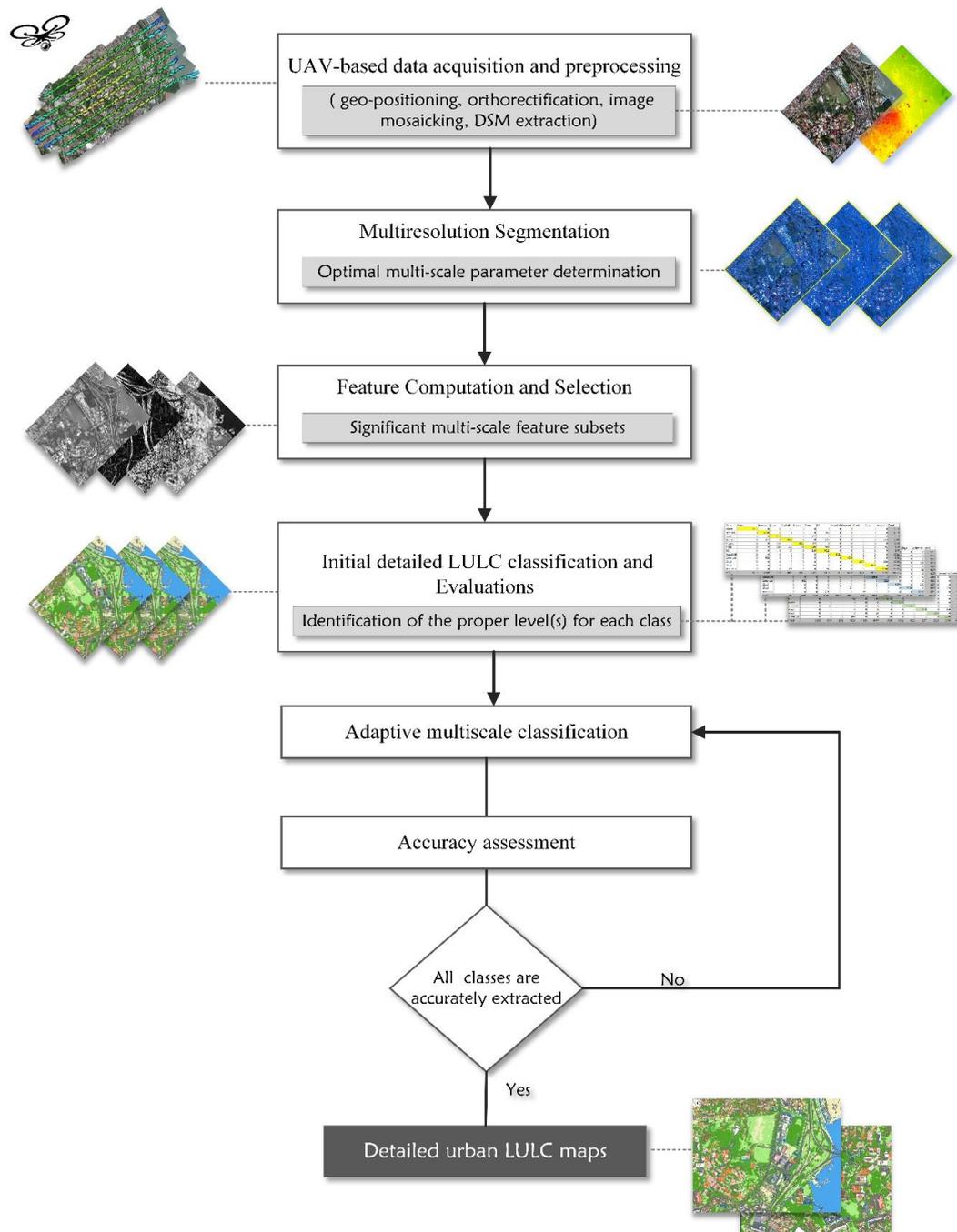


Figure 2. Framework of the proposed methodology.

3.2. Image Preprocessing

Various photogrammetric steps, such as interior, relative, and absolute orientations, have been conducted to establish the mathematical relationship between the image and the ground and subsequently generate the digital elevation model and the orthophoto (an image with the same

characteristics of the map, where distortions caused by relief displacement are removed and the image has a uniform scale). Throughout this process, image matching, automatic aerial triangulation, geopositioning, orthorectification, and image mosaicking were performed to create the orthomosaic image and the DSM from the UAV data using Agisoft PhotoScan and ArcGIS 10.4.1. The process commenced by estimating the exterior and interior orientation parameters that estimate the positions of the camera in each image and the camera calibration parameters. The RGB images were geometrically corrected and geotagged to the WGS1984 (world geodetic reference system) using the files extracted from the Global Positioning System units in the drone and the ground reference station. The images were projected to a Universal Transverse Mercator coordinate system (zone 38 North) and converted from JPEG to GeoTiff format. The following steps, such as aligning images, building field geometry, and orthophoto generation, were implemented to create a DSM (a 3D polygon mesh representing surface ground) and an orthomosaic. The DSM was generated with the nearest-neighbor interpolation method and resampled to the same resolution of the orthomosaic. The spatial resolution of the final orthomosaic for the two study areas was 10 cm, and the spatial resolution of the DSM was 80 cm.

3.3. MS Image Segmentation Optimization

The optimal segmentation level is defined in most of the unsupervised methods as the level that maximizes the between-object heterogeneity (i.e., adjacent objects can be distinguished from their surroundings) and the within-object homogeneity (i.e., pixels belonging to the same objects are similar) [40,41]. The likeness between each image and its neighbors, known as the undersegmentation metric, is determined through spatial autocorrelation (Moran's I (MI)) [58], whereas the internal homogeneity of an image object, known as the oversegmentation metric, is determined through the area-weighted variance (WV) [41].

An adaptive segmentation optimization approach that integrates unsupervised quality measures, namely, the F -measure, accompanied with a machine learning classification model was adopted in this study to identify the optimal scale(s) for each urban LULC class. The F -measure quality measure [39] was utilized to determine the hierarchical scale values from a set of given segmentation outputs. The F -measure value for estimating the optimum MS of an application can be computed using Equation (1).

$$F\text{-measure} = (1 + \varphi^2) \frac{MI_{norm} \times WV_{norm}}{\varphi^2 \cdot MI_{norm} + WV_{norm}}, \quad (1)$$

where WV_{norm} and MI_{norm} represent the normalized area-WV (oversegmentation metric) and the normalized Moran's I (undersegmentation metric), respectively. The relative weights of WV_{norm} and MI_{norm} are controlled through a scene-independent factor (φ). The φ values are selected to ensure that the generated segmentation levels vary considerably in terms of the within-object homogeneity and between-object heterogeneities. For instance, $\varphi = 3$ signifies that triple weighting is assigned to WV_{norm} , $\varphi = 0.5$ indicates half weighting for WV_{norm} , and $\varphi = 1$ denotes that equal weighting is considered for WV_{norm} and MI_{norm} . Additional details about the unsupervised parameter optimization can be found in [39,59]. The levels defined by the F -measure are used in the second phase to perform a single scale (SS) classification of each defined segmentation scale. The class-specific accuracy (F -measure) is used to evaluate the accuracy of each class at multiple levels, as shown in Section 4.3. Then, the optimal scale(s) for extracting each class is determined and used for subsequent analysis.

3.4. Feature Computation and Selection

Considering the spectral similarity between the various urban LULC classes in the UHSR RGB images, various features, including spectral values, color invariants, and geometrical textural features, were computed and assessed at multiple scales, as shown in Table 2. Selecting the significant features prior to classification is necessary to minimize the computational time by excluding the redundant attributes and enhance the accuracy of an ML classifier [47]. In this study, CFS and SVM as wrapper FS techniques were utilized to identify the most relevant MS features of image objects from UAV datasets.

Table 2. Detailed description of the evaluated attributes (features).

Feature Type	Tested Feature Name	Description	Reference
Spectral	Mean	The mean intensity values computed for an image segment of the RGB channels and the DSM	[60]
	Standard deviation	The standard deviation values computed for an image segment of the RGM channels and the DSM.	[60]
	Max_ difference	The maximum difference between the RGB channels.	[60]
	Brightness	The average of means of the RGB channels.	[60]
	NDRG	$\frac{Red-Green}{Red+Green}$	[61]
	NDGB	$\frac{Green-Blue}{Green+Blue}$	[61]
	NDBG	$\frac{Blue-Green}{Blue+Green}$	[61]
	NDRB	$\frac{Red-Blue}{Red+Blue}$	[61]
	NDBR	$\frac{Blue-Red}{Blue+Red}$	[61]
	NDGR	$\frac{Green-Red}{Green+Red}$	[61]
	RB	$\frac{Blue}{Red}$	[61]
	Ratio-R	$\frac{Green+Blue+Red}{Green}$	[61]
	Ratio-G	$\frac{Green+Blue+Red}{Blue}$	[61]
	Ratio-B	$\frac{Green+Blue+Red}{Red}$	[61]
	V	$\frac{4}{\pi} \cdot \arctan \left(\frac{Green-Blue}{Green+Blue} \right)$	[62]
S	$\frac{4}{\pi} \cdot \arctan \left(\frac{1 - \sqrt{Red^2 + Green^2 + Blue^2}}{1 + \sqrt{Red^2 + Green^2 + Blue^2}} \right)$	[63]	
Texture	Mean	The grey level co-occurrence matrix (GLCM) mean sum of all directions determined for each band from the RGB channels and the DSM.	[64]
	Homogeneity	The GLCM homogeneity sum of all directions determined for each band from the RGB channels, and the DSM.	[64]
	Contrast	The GLCM contrast sum of all directions determined for each band from the RGB channels, and the DSM.	[64]
	Entropy	The GLCM and GLDV entropy sum of all directions determined for each band from the RGB channels, and the DSM.	[64]
	Correlation	The GLCM correlation sum of all directions determined for each band from the RGB channels, and the DSM.	[64]
	Standard deviation	The GLCM standard deviation sum of all directions determined for each band from the RGB channels, and the DSM.	[64]
	Dissimilarity	The GLCM dissimilarity sum of all directions determined for each band from the RGB channels, and the DSM.	[64]
	Angular second moment	The GLCM angular second-moment sum of all directions determined for each band from the RGB channels, and the DSM.	[64]
Geometric	Length\Width	The ratio between the length and width.	[60]
	Rectangular Fit	A ratio that is based on how well an image object fits into a rectangle.	[60]
	Shape_index	A ratio that defines border smoothness of image objects and can be computed by dividing the border length of an image object by four times the square root of its area.	[60]
	Density	It can be computed by dividing the area covered by an image object by its radius.	[60]
	Elliptic_fit	A ratio based on how well an image object can fit into an ellipse.	[60]
	Compactness	It is expressed as the ratio of the area of an image object to the area of a circle with a similar perimeter.	[60]

The seventy features listed above were computed for three optimized MS image objects, and two efficient FS methods, namely, CFS and SVM, were used to find the relevant feature subset for each optimized image object level.

3.4.1. CFS

CFS performs fast processing to appropriately select the optimal feature subset [53,65,66]. It uses a search algorithm that heuristically assesses each attribute's predictive capability and the degree of intercorrelation between the attributes [67]. In other words, this evaluating mechanism calculates the correlations between the features and classes to classify highly correlated features to the target class whilst considering the low correlations and low level of redundancy amongst the features [68]. The estimations of the correlation between the subset of attributes and target classes are performed using Equation (2).

$$R_s = \frac{s\bar{r}_{ci}}{\sqrt{s + s(s-1)\bar{r}_{ii}}} \quad (2)$$

where s denotes the number of features, \bar{r}_{ci} represents the correlation average between the subset features and the class variable, and \bar{r}_{ii} denotes the intercorrelation average between the subset features. Accordingly, the high correlation coefficients between the feature attributes and the target labels are considered to be relevant to the respective class characterization with a high level of association, whilst lower intercorrelation (\bar{r}_{ii}) is desired [68].

3.4.2. SVM

SVM is a widely applied regression algorithm with a nonparametric supervised statistical learning task and is highly suitable for GEOBIA FS and classification tasks [51,69]. This algorithm seeks an optimal separating hyperplane using the training dataset of so-called support vectors that can effectively separate the input features (datasets) into target classes with a minimum misclassification and a maximum margin amongst the target classes [70–72]. When the task is linearly separable, the hyperplane can be represented using Equation (3):

$$y_i(w \cdot x_i + b) \geq 1 - \delta_i, \quad (3)$$

where w indicates the coefficient vector that determines the orientation of the hyperplane in the feature space. The offsets of the hyperplane from the original and positive slack variables are represented by b and δ_i , respectively [73]. Equation (4) determines the optimized hyperplane, where many hyperplanes can be designed to distinguish between classes.

$$\text{Minimise } \sum_{i=1}^n a_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n a_i a_j y_i y_j (x_i x_j), \quad (4)$$

$$\text{Subject to } \sum_{i=1}^n a_i y_i = 0, \quad 0 \leq a_i \leq C, \quad (5)$$

where a_i denotes the Lagrange multipliers and C is the penalty.

3.5. Supervised MS Image Object Classification

Image classification is the final phase in GEOBIA, and the common classification methods used in this phase are supervised ML models or rule-based methods. In this study, the MS image object classification was implemented using three supervised classification algorithms, namely, RF, SVM, and DT. The classification models were trained using the sample statistics derived from the GT dataset of the first study area. The object-based classification outcomes at different scales were used to quantitatively evaluate the MS segmentation results and select the optimum scale for each urban LULC class. Then,

the classification scheme started with a single classification of each optimized image-level using the selected feature subsets for each level. After acquiring the proper information about the optimal scale(s) for each class, ML models were used to initially classify large objects at large SP. The classification results were then copied to a new level, where the unclassified objects were only resegmented to a fine segmentation level, and the ML models were then used to classify the resegmented objects on the basis of the selected significant features at that level. The process iteratively continued until all classes were accurately classified or no improvement was detected in the OA and class-specific accuracy (*F*-measure). The two of the aforementioned ML algorithms are briefly described in the following paragraphs.

A DT is a supervised and nonparametric ML technique that is operable without prior knowledge on data distribution, with easy interpretation and capability to model and handle the data complexity reduction and the relationships between variables [74–79]. It is a flexible, fast, and robust algorithm that can be used to control the nonlinearity between the input features and discrete classes [75]. DT hierarchically utilizes IF-THEN rules to label the variables of each class, where the tree structures, leaves, and end nodes represent the discrete class labels (decision), and the branches assist in assigning the labels on the basis of the attributes and majority voting [76]. A heuristic DT recursively partitions a dataset into homogenous subsets in conjunction with the attribute values at each branch or node in the single tree [77].

The RF algorithm is an ensemble of DT classifiers that improves the classification of variables with high accuracy, and its robustness against overfitting the training dataset along with insensitivity to nonnormal and noisy data makes it suitable for LULC classification [51,78,79]. RF is an ensemble method that exploits many DTs as a forest generated from bootstrap and utilizes each tree's vote to assign the most frequent class label to the input variables [78,80]. Each tree then randomly selects the predictors and object features from the input vector of every tree node to increase the generalization error [78,81]. The prediction of the samples is calculated on the basis of the majority votes amongst the trees [80,81]. The discrimination assignment is calculated using Equation (6):

$$H(x) = \operatorname{argmax}_Y \sum_{i=1}^k I(h_i(X, \theta_k) = Y), \quad (6)$$

where θ_k is a random vector for the k th tree, X is an input vector, $I(\cdot)$ is an indicator function, $h(\cdot)$ is a single DT, Y is an output variable, and argmax_Y denotes the Y value in the maximization of $\sum_{i=1}^k I(h_i(X, \theta_k) = Y)$.

3.6. Evaluation Metrics

The evaluation metrics of the classified images were generated through the frequently applied confusion matrix and its derivatives, including the OA, *K*, precision, recall, and *F*-measure. The error matrix (confusion matrix) evaluates the classification results versus the reference data in two dimensions as actual classes in rows and predicted classes in columns.

3.6.1. OA

The OA, which is a percentage indicator of the classification performance, can be defined as the sum of the correctly classified variables into discrete classes (true positives plus true negatives) to the total tested variables. OA can be computed from the confusion matrix by dividing the total number of correctly classified objects/pixels ($\sum D_{ij}$ or the sum of the major diagonal) with the total number of objects/pixels (N):

$$OA = \frac{\sum D_{ij}}{N}. \quad (7)$$

3.6.2. K Statistics

The K statistic is another statistical measure that defines the observed level of agreement or accuracy between a detailed map and reference data. The K value approaches +1 when the contribution of the chance of agreement diminishes and becomes negative when the effects of chance agreements increases. Conversely, a K value equaling 0 indicates no agreement, indicating that the classification is entirely conducted by chance or random assignment. A negative K value signifies that the agreement is worse than occurring by chance. The K statistic is computed using Equation (8):

$$K = \frac{N \sum_{i,j=1}^m D_{ij} - \sum_{i,j=1}^m R_i \times C_j}{N^2 - \sum_{i,j=1}^m R_i \times C_j}, \quad (8)$$

where m denotes the number of urban LULC classes in the confusion matrix, D_{ij} denotes the number of observations (objects/pixels) that are correctly classified in row i and column j , R_i denotes the total number of objects/pixels in row i , C_j denotes the total number of observations in column j , and N denotes the total number of objects/pixels.

3.6.3. Precision, Recall, and F-measure

The F -measure is the weighted average or harmonic mean of two ratios known as precision (p). Recall (r) metric is another performance measure used to assess the class-specific accuracy from retrieved information [82,83]. It can be computed using Equation (9) on the basis of the average of p and r . The F -measure value ranges from 0 (lowest) to 1 (highest).

$$F_{\text{measure}} = 2 \times \frac{p \times r}{p + r}. \quad (9)$$

The p or the confidence of a LULC class is determined by dividing the number of true positives (number of objects\pixels correctly belonging to the actual class) by the total number of objects categorized as the positive class (i.e., the sum of true positives and false positives, which are objects/pixels incorrectly categorized as belonging to the class). The r or the sensitivity shows the proportion of true positive objects/pixels that are correctly predicted and identified and can be defined as the number of true positives divided by the total number of objects/pixels that are members of the positive class (i.e., the sum of true positives and false negatives). p and r can be calculated using Equations (10) and (11), respectively. A perfect predictor's value for p and r would be described as 1.

$$p = \frac{\text{true positives}}{\text{true positives} + \text{false positives}}, \quad (10)$$

$$r = \frac{\text{true positives}}{\text{true positives} + \text{false negatives}}. \quad (11)$$

4. Results

This section summarizes the various outcomes of this study, including the MS image segmentation optimization and parameter selection, FS, and classification results.

4.1. Results of MS Image Segmentation

In this study, the quantitative evaluation of image segmentation results at MS levels through unsupervised segmentation quality measures aims to determine the optimal SP that allows excellent delineation and extraction of urban LULC classes that may share a similar spectral response with each other and vary in structure, size, and their surrounding contrast. The oversegmentation (WV) and undersegmentation (MI) metrics were computed from the three RGB channels, and their mean values were normalized and used to compute the F -measure (for selecting the three optimum SPs),

as shown in Table 3. Three values, namely, 3, 1, and 0.33, of the scene-independent variables (φ) were selected to pinpoint the three SPs from the computation of Equation (1). These values were empirically selected and supported by the study of Johnson et al. [62] to ensure that the adopted segmentation levels vary remarkably from each other in terms of the between-object homogeneity and within-object heterogeneity. The highest values on the last three columns in Table 2 correspond to the optimal MS levels, and these scales are 200, 100, and 50. Figure 3a,b depict the image segmentation results of a small subset at the scale of 200, where large homogenous objects, such as water bodies, grass, bare soil, and some clay tiles, are well delineated. Figure 3c,d show the image segmentation results of a small subset at the scale of 100, where medium objects, such as some types of roofing materials, are well identified. Figure 3e,f display the image segmentation results of a small subset at the scale of 50, where large and medium objects are oversegmented but small roofing materials and trees are well distinguished.

Table 3. Results of the applied unsupervised segmentation quality measures at multiscale (MS) levels.

Scale	No of Objects	WV _{mean}	MI _{mean}	WV _{norm}	MI _{norm}	F-Measure		
						$\varphi = 3$	$\varphi = 1$	$\varphi = 0.33$
25	340731	78.606	0.548	1.000	0.000	0	0	0
50	104840	132.924	0.452	0.858	0.242	0.684	0.377	0.260
75	53011	178.661	0.395	0.739	0.385	0.677	0.506	0.404
100	33217	217.177	0.344	0.639	0.514	0.624	0.570	0.524
125	22978	253.258	0.314	0.545	0.588	0.549	0.566	0.584
150	16878	288.418	0.286	0.453	0.658	0.468	0.537	0.630
175	12887	322.978	0.250	0.363	0.748	0.383	0.489	0.674
200	10181	354.309	0.229	0.281	0.801	0.301	0.416	0.678
225	8216	384.925	0.217	0.202	0.833	0.218	0.325	0.637
250	6841	412.203	0.195	0.130	0.888	0.143	0.227	0.566
275	5738	439.515	0.169	0.059	0.953	0.065	0.112	0.384
300	4961	462.250	0.150	0.000	1.000	0	0	0



Figure 3. Optimized image objects of a subset UAV image mosaic: (a,b) Scale 200; (c,d) scale 100; (e,f) scale 50.

4.2. Results of FS

Following the optimization of segmentation SPs, several spectral, geometrical, and textural features were computed at MS levels for FS, as shown in Table 1. Two wrapper approaches, namely,

CFS and SVM, combined with the KNN algorithm, were used to assess all features as a part of classification. Table 4 compares the OA, K , and other relevant features selected by SVM and CFS at scales of 50, 100, and 200. The results of CFS and SVM exhibited significant differences in terms of the number and type of selected features in each scale. However, the two methods eliminated 60% from the total number of features, whereas less than 40% of the features contributed to achieving high accuracy. CFS attained a slight improvement in terms of the OA and number of selected features, as presented in Table 2, and was selected for subsequent processing.

4.3. Classification Results

The detailed mapping of impervious surfaces in a heterogeneous urban area from UAV-based images is particularly challenging when only three spectral channels, RGB, are used because of the spectral similarity of various urban LULC classes. In such a case, a successful extraction of urban objects should consider the information of the variation in size and the surroundings of the different types of LULC that exist in the image. For instance, asbestos cement and dark concrete roofs or cemented pavements may share similar spectral responses because of the presence of cement in their contents. To minimize the confusion between different LULC classes, the information of the suitable scale(s) that provides the best accuracy and ensures the strong differentiation between classes is necessary to obtain a holistic view and to perform hierarchical classification.

The initial stage of classification in this study is to find the optimum level for extracting each class, which can be achieved using ML models, followed by a class-specific accuracy measure. Three standard classification algorithms, namely, RF, SVM, and DT, were used to classify the first study area at the selected optimal scales (SP 200, SP 100, and SP 50). The accuracy of each classification level was evaluated on the basis of OA, K , and F -measure. Figure 4a–c show the SS classification results of RF, Figure 4d–f display the SS classification results of SVM, and Figure 4g–i show the SS classification results of DT. Table 5 shows the SS classification results for the first study area. The highest SS classification results were obtained by SS-RF at scale 50, with an OA of 92.2 and a K of 9.14, followed by SS-SVM at scale 100 with an OA of 90.5 and a K of 0.896 and SS-DT at scale 50 with an OA of 88.1% and a K of 0.87. Finding the optimum scale for extracting the LULC in heterogeneous urban areas can vary on the basis of the adopted classification algorithm by comparing the class-specific accuracy measures of SS-RF, SS-SVM, and SS-DT classification results. For instance, the SS-RF classification results showed that the SP 50 exhibited the highest OA for the extraction of water bodies, trees, grass, dark concrete, type 2 clay tiles, and type 2 metallic roofs, whereas the SP 200 showed enhanced extraction of bare soil, asphalt, type 1 metallic roofs, concrete, type 1 clay tiles, and asbestos cement roofs. The classification results of SS-SVM showed better extraction for clay tiles (types 1 and 2) at the largest optimized SP, whereas the smallest optimized SP was optimal for extracting water bodies, bare soil, trees, and metallic roofs (types 1 and 2). The previous step was adopted prior to the hierarchical classification approach to provide a diagnostic result where SP is suitable for extracting 12 urban LULC classes and for ensuring reasonable discrimination between the classes.

Table 4. Results of correlation-based feature selection (CFS) and support vector machine (SVM) at MS levels.

Scale	Feature Type	CFS			SVM				
		Selected Features	No	OA	K	Selected Features	No	OA	K
50	Spectral	Red, Blue, DSM, SD_DSM, Vegetation, Ratio_G, Ratio_B, NDRG, NDGR, NDBR, Max_diff	16	91.63	0.93	Red, Green, Blue, DSM, Vegetation, Ratio_G, Ratio_B, NDGR, NDBR, NDRB, NDGB, NDBG, Shadow, RB, Max_diff	21	91.61	0.91
	Textural	GLCM_Mean_DSM, GLCM_Entropy_Green, GLCM_Dissimilarity_Blue, GLCM_SD_Blue				GLCM_Entropy_R, GLCM_Entropy_Blue, GLCM_Entropy_Green, GLCM_Homogeneity_Red, GLDV_Mean_Blue			
	Geometrical	Shape_index				Length/Width			
100	Spectral	Red, Green, Blue, DSM, SD_DSM, SD_R, Vegetation, Ratio_R, Ratio_G, NDRG, NDBR, NDBG, Shadow, RB, Max_diff	22	93.3	0.92	Green, Red, Blue, DSM, Vegetation, Ratio_G, Ratio_Blue, NDRG, NDGR, NDGB, RB, NDBR, NDRB, Shadow, NDBG, Max_diff, Brightness	28	92.11	0.914
	Textural	GLCM_Mean_Red, GLCM_Mean_DSM, GLCM_SD_Green, GLCM_Correlation_DSM, GLCM_Dissimilarity_Blue, GLCM_Ang 2nd moment_Green				GLCM_Mean_DSM, GLCM_Mean_Blue, GLCM_Entropy_Blue, GLCM_Entropy_Red, GLCM_Entropy_Green, GLCM_Homogeneity_Red, GLDV_Mean_Blue, GLDV_Conrast_Blue, GLDV_Entropy_Red.			
	Geometrical	Shape_index				Shape_index, Length/Width			
200	Spectral	Red, Green, Blue, DSM, SD_DSM, Vegetation, Ratio_R, Ratio_G, Ratio_B, NDRG, NDGR, NDBR, NDGB, NDRB, Shadow, Max_diff	27	93.78	0.93	Red, Green, Blue, DSM, SD_DSM, Vegetation, Ratio_G, Ratio_B, NDGR, NDBG, Shadow, RB, Max_diff	21	92.3	0.92
	Textural	GLCM_Mean_Red, GLCM_Mean_Blue, GLCM_SD_DSM, GLCM_SD_Green, GLCM_Homogeneity_DSM, GLCM_Ang 2nd moment_Blue, GLCM_Dissimilarity_Blue, GLCM_Correlation_Rede, GLCM_Homogeneity_Red				GLCM_Mean_Blue, GLCM_Homogeneity_Green, GLCM_Entropy_R, GLCM_Correlation_Blue, GLDV_Mean_Blue, GLDV_Entropy_DSM			
	Geometrical	Shape_index, Length/Width				Length/Width, Compactness			

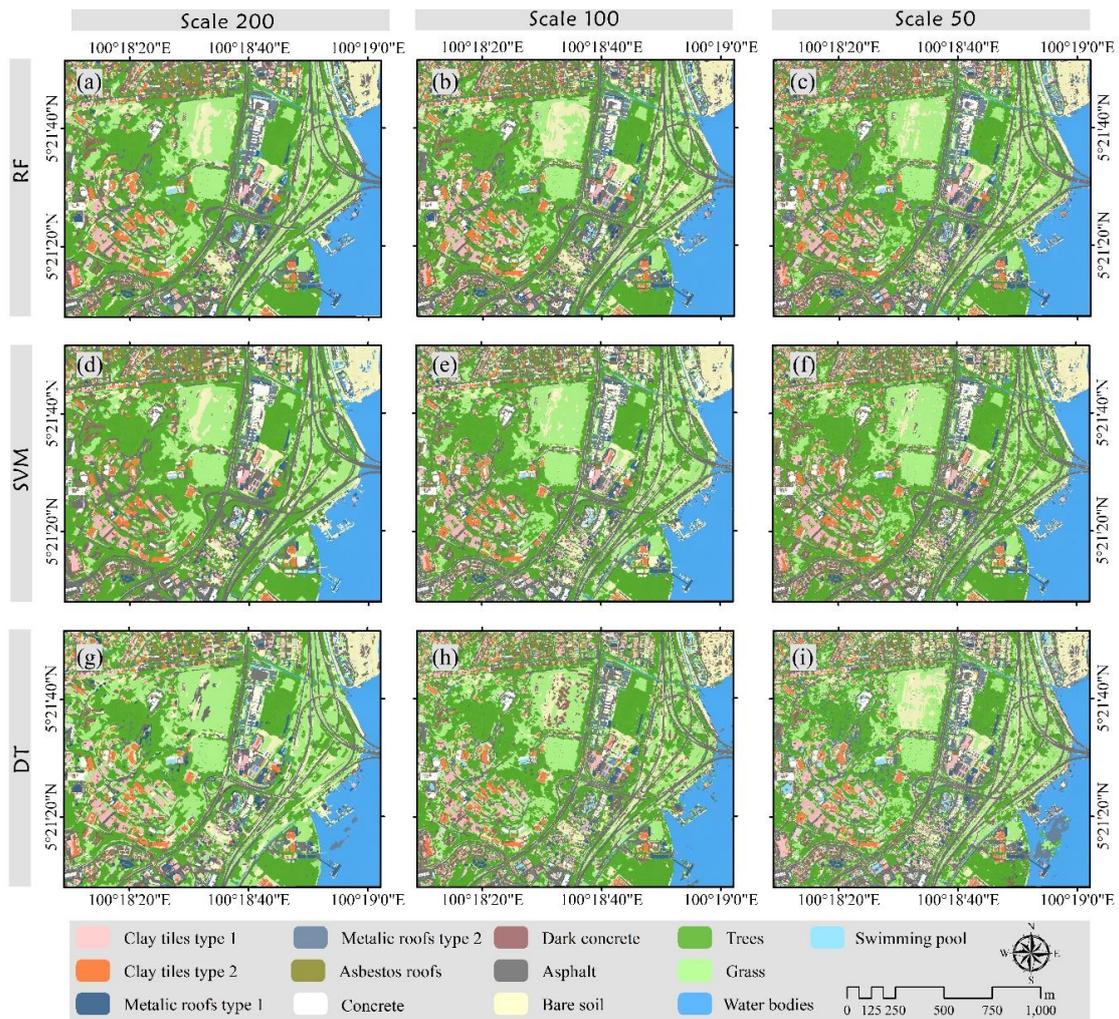


Figure 4. Single-scale (SS) classification results: (a–c) Random forest (RF) classification at scales of 200, 100, and 50, respectively; (d–f) SVM classification at scales of 200, 100, and 50, respectively; and (g–i) decision tree (DT) classification at scales of 200, 100, and 50, respectively.

Table 5. Performance of the extraction of LULC classes using RF, SVM, and DT at single scales.

Class	SS-RF								
	SP 200			SP 100			SP 50		
	Precision	Recall	F-Measure	Precision	Recall	F-Measure	Precision	Recall	F-Measure
Water bodies	1.000	0.980	0.990	1.000	0.997	0.999	1.000	0.997	0.999
Bare soil	0.879	0.918	0.898	0.653	0.404	0.499	0.850	0.942	0.893
Grass	0.790	0.352	0.487	0.871	0.878	0.874	0.871	0.994	0.928
Asphalt	0.855	0.757	0.803	0.674	0.626	0.649	0.771	0.569	0.655
Metallic roofs 2	0.992	0.862	0.922	1.000	0.924	0.961	1.000	1.000	1.000
Trees	0.580	0.830	0.683	0.790	0.999	0.883	0.999	0.973	0.986
Dark concrete	0.741	0.961	0.836	1.000	0.899	0.947	0.994	0.969	0.981
Metallic roofs 1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Concrete	1.000	0.925	0.961	1.000	0.906	0.951	0.983	0.914	0.947
Clay tiles type 2	1.000	0.964	0.981	0.980	0.874	0.924	0.985	0.983	0.984
Clay tiles type 1	1.000	0.996	0.998	0.474	0.902	0.621	0.977	1.000	0.989
Asbestos	0.913	0.951	0.932	0.743	0.838	0.788	0.624	0.849	0.719
OA	88.4%			84.25%			92.2%		
Kappa	0.873			0.827			0.914		

Table 5. Cont.

SS-SVM									
Class	SP 200			SP 100			SP 50		
	Precision	Recall	F-Measure	Precision	Recall	F-Measure	Precision	Recall	F-Measure
Water bodies	0.964	1.000	0.982	1.000	0.998	0.999	1.000	1.000	1.000
Bare soil	0.861	0.823	0.842	0.940	0.974	0.956	0.936	1.000	0.967
Grass	0.790	0.742	0.765	0.787	0.866	0.825	0.835	0.874	0.854
Asphalt	0.861	0.836	0.849	0.906	0.854	0.879	0.878	0.745	0.806
Metallic roofs 2	0.854	1.000	0.921	0.891	1.000	0.942	0.998	1.000	0.999
Trees	0.943	0.939	0.941	0.941	0.953	0.947	0.974	0.963	0.969
Dark concrete	0.835	0.624	0.714	0.999	0.651	0.788	0.977	0.633	0.768
Metallic roofs 1	0.969	0.808	0.881	1.000	0.848	0.918	1.000	1.000	1.000
Concrete	1.000	0.907	0.951	1.000	0.993	0.996	1.000	0.942	0.970
Clay tiles type 2	1.000	1.000	1.000	1.000	1.000	1.000	0.928	0.982	0.954
Clay tiles type 1	0.485	0.992	0.651	0.485	0.911	0.633	0.474	0.873	0.614
Asbestos	0.983	0.933	0.958	0.889	0.967	0.926	0.750	0.952	0.839
OA	88%			90.5%			89.7%		
Kappa	0.868			0.896			0.886		
SS-DT									
Water bodies	1.000	0.788	0.881	1.000	1.000	1.000	0.915	1.000	0.956
Bare soil	0.889	0.853	0.871	0.758	0.886	0.817	0.729	0.924	0.815
Grass	0.790	0.246	0.375	0.871	0.244	0.381	0.864	0.687	0.765
Asphalt	0.752	0.812	0.781	0.707	0.677	0.692	0.710	0.541	0.614
Metallic roofs 2	0.957	0.858	0.905	0.998	1.000	0.999	0.998	0.846	0.916
Trees	0.529	0.895	0.665	0.806	0.945	0.870	0.904	0.965	0.934
Dark concrete	0.732	0.966	0.833	0.946	0.843	0.891	0.943	0.925	0.934
Metallic roofs 1	1.000	0.964	0.982	1.000	1.000	1.000	1.000	1.000	1.000
Concrete	1.000	0.925	0.961	0.983	0.911	0.946	0.975	1.000	0.987
Clay tiles type 2	0.948	0.521	0.672	0.886	0.902	0.894	1.000	0.823	0.903
Clay tiles type 1	0.474	0.948	0.632	0.385	0.882	0.536	1.000	1.000	1.000
Asbestos	0.676	1.000	0.806	0.756	0.833	0.793	0.639	0.852	0.730
OA	79%			83.4%			88.1%		
K	0.771			0.819			0.869		

Utilizing the preliminary information acquired from the SS classification of the RF, SVM, and DT algorithms, the hierarchical classification was conducted for the first study area. The results are shown in Figure 5. Table 6 illustrates the OA, K, and class-specific accuracies of the first study area using the hierarchical RF, SVM, and DT classification algorithms. Similar to SS classification, the MS-RF classification was superior with an OA of 94.40% and a K of 0.938, followed by MS-SVM with an OA of 92.50% and a K of 0.917 and MS-DT with an OA of 91.60% and a K of 0.908.

Table 6. Class-specific accuracy measures for the MS RF, SVM, and DT of the first study area.

Class	MS-RF			MS-SVM			MS-DT		
	Precision	Recall	F-Measure	Precision	Recall	F-Measure	Precision	Recall	F-Measure
Water bodies	1.000	0.967	0.983	1.000	0.998	0.999	1.000	1.000	1.000
Bare soil	0.761	0.754	0.757	0.919	1.000	0.958	0.899	0.810	0.852
Grass	0.871	0.933	0.901	0.996	0.892	0.941	0.869	0.698	0.774
Asphalt	0.863	0.916	0.889	0.851	0.944	0.895	0.765	0.705	0.734
Metallic roofs 2	1.000	0.999	1.000	0.998	1.000	0.999	0.998	0.873	0.932
Trees	0.972	0.972	0.972	0.941	0.999	0.969	0.888	0.999	0.940
Dark concrete	0.832	0.963	0.893	0.996	0.646	0.783	0.768	0.950	0.849
Metallic roofs 1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Concrete	1.000	0.925	0.961	1.000	0.993	0.996	0.975	1.000	0.987
Clay tiles type 2	1.000	0.964	0.981	1.000	1.000	1.000	0.931	0.878	0.904
Clay tiles type 1	1.000	0.923	0.960	0.485	0.904	0.631	1.000	1.000	1.000
Asbestos	0.962	0.969	0.966	0.983	0.927	0.955	0.935	0.975	0.954
OA	94.40%			92.50%			91.60%		
K	0.938			0.917			0.908		

Compared to SS classification, the hierarchical classification results noticeably improved the extraction of urban LULC classes. For instance, an improvement of 2.24% in the OA was observed in the MS-RF algorithm, along with a significant improvement in the differentiation and extraction

of asbestos cement, concrete, and asphalt roofs. Similarly, the MS-SVM classification exhibited an enhancement in the class-specific accuracies, OA, and K of trees, grass, and asphalt classes. The OA accuracy of MS-DT showed an improvement with 3.57%, which achieved an overall improvement in the extraction of trees, grass, and asbestos cement roofs.

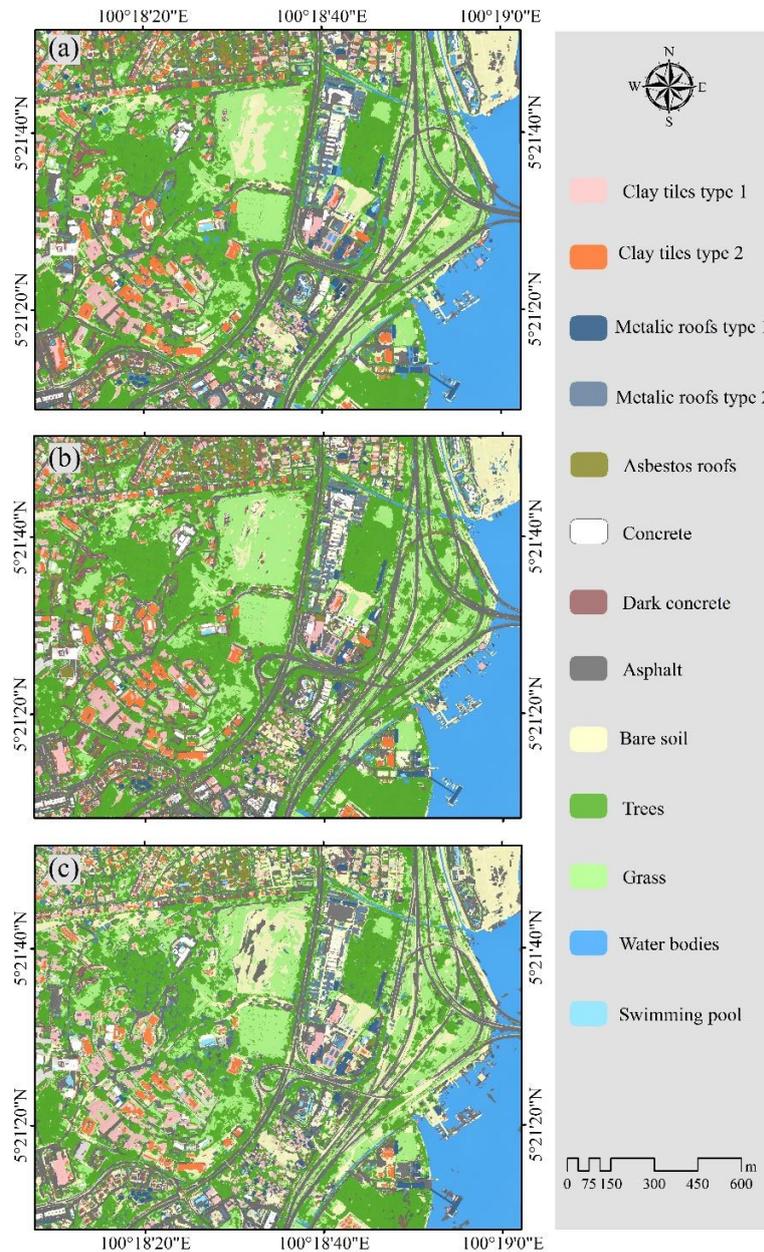


Figure 5. Results of MS classification using the integrated approach of the first dataset: (a) RF, (b) SVM, and (c) DT.

To validate the transferability of the hierarchical classification approach, the MS-RF, MS-SVM, and MS-DT classifications were applied in the second study area using the sample statistic file derived from the image of the first study area. Figure 6 and Table 7 show the classification results for the second study area. The results of the second dataset showed that the MS-SVM classification was superior with an OA of 94.45% and a K of 0.938, followed by MS-RF with an OA of 92.46% and a K of 0.916 and MS-DT with an OA of 90.46% and a K of 0.893. The proposed hierarchical classification approach demonstrates excellent potential for the detailed mapping of heterogeneous urban areas from RGB-UAV images and DSM. The proposed methodology can be adopted for various areas with different LULCs.

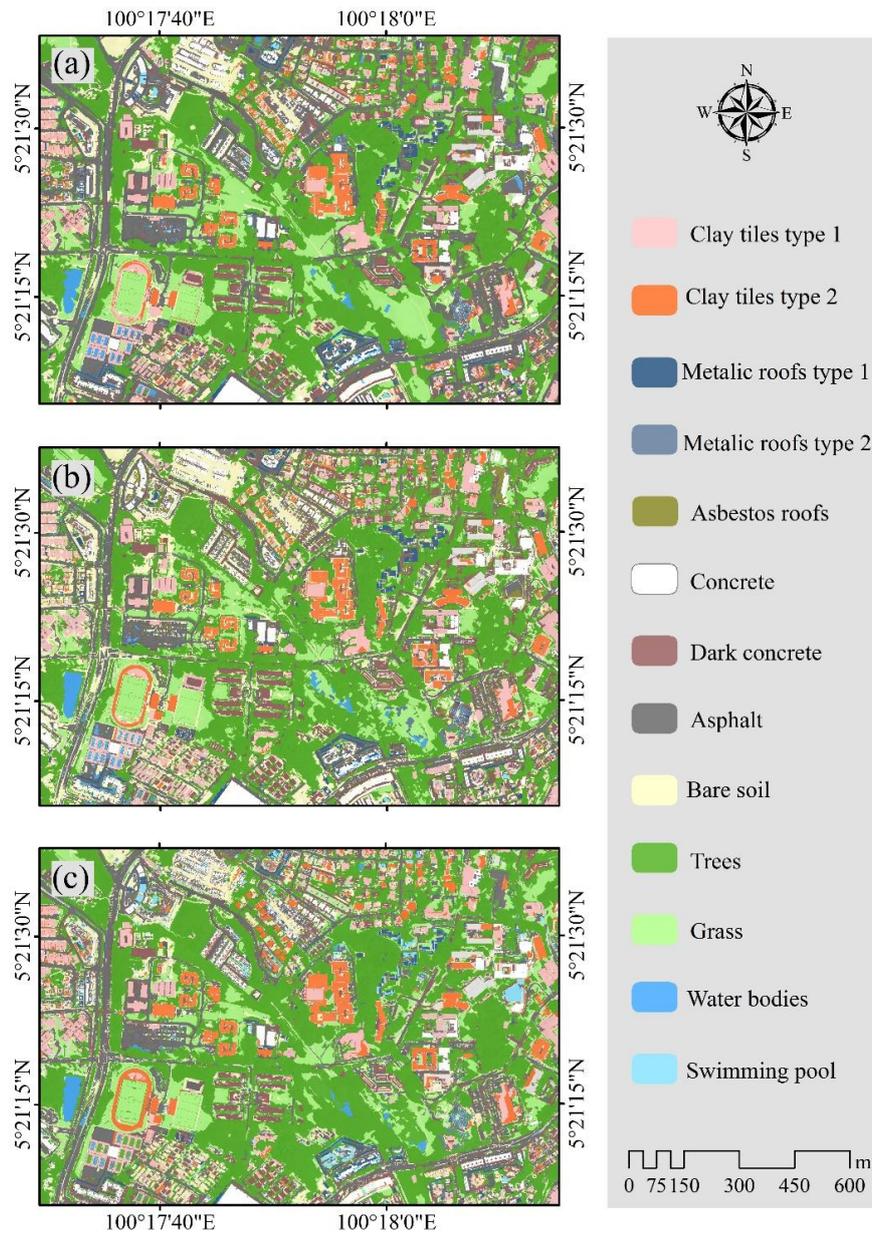


Figure 6. Results of MS classification using the integrated approach of the second study area: (a) RF, (b) SVM, and (c) DT.

Table 7. Class-specific accuracy measures for the MS RF, SVM, and DT of the second study area.

Class	RF			SVM			DT		
	Precision	Recall	F-Measure	Precision	Recall	F-Measure	Precision	Recall	F-Measure
Water bodies	0.531	0.989	0.691	0.594	0.969	0.737	0.585	0.974	0.731
Bare soil	0.960	0.884	0.921	0.982	0.960	0.971	0.953	0.788	0.862
Grass	0.972	0.650	0.779	0.984	0.786	0.874	0.912	1.000	0.954
Asphalt	0.925	1.000	0.961	0.976	1.000	0.988	0.973	1.000	0.986
Metallic roofs 2	1.000	1.000	1.000	0.969	0.680	0.800	0.745	1.000	0.854
Trees	1.000	0.959	0.979	1.000	0.984	0.992	1.000	0.618	0.764
Dark concrete	0.971	0.975	0.973	0.992	1.000	0.996	0.901	0.990	0.944
Metallic roofs 1	0.992	1.000	0.996	0.980	1.000	0.990	0.986	1.000	0.993
Concrete	0.953	1.000	0.976	1.000	1.000	1.000	0.953	1.000	0.976
Clay tiles type 2	1.000	0.961	0.980	1.000	0.983	0.991	1.000	0.856	0.923
Clay tiles type 1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
OA	92.46%			94.45%			90.46%		
K	0.916			0.938			0.893		

5. Discussion

Considering that segmenting UHSR UAV-based images of a heterogeneous and complex urban landscape is a challenging task in GEOBIA, the selection of the optimum SP(s) is an imperative step to ensure that different landscapes are well delineated at different scales. This study conducted a detailed mapping of a heterogeneous urban area, an area covered with various natural and impervious surfaces that vary in size and structure, from the fusion of the UAV-based orthophoto and DSM by improving the GEOBIA frameworks with different solutions to some of the issues stated in related studies section. An adaptive MS segmentation that assimilates an unsupervised image segmentation evaluation metric (i.e., *F*-measure) and ML algorithms were proposed to identify the optimal MS parameters for extracting the detailed urban LULC classes.

Although GEOBIA can leverage the computation and use various features in the classification process, adding many features can reduce the classification accuracy and increase the computational time. CFS and SVM were used in this study to select the most significant features computed for each level from the optimized three-scale levels. An object's spectral, geometrical, and textural feature values are different because the size of the generated image objects (i.e., roofing material and roads) varies on the basis of the selected scale level. CFS obtained a maximum OA of 93.78% ($K = 0.93$) at the scale level of 200 by selecting 27 significant features, whereas SVM obtained a minimum accuracy at the scale level of 50, with the value of OA = 91.61 ($K = 0.91$) by selecting 21 features. CFS and SVM selected a set of features that vary in terms of the number and type in each segmentation level. However, various spectral features, such as R, B, DSM, Ratio-G, Ratio-B, Vegetation, the normalized difference between the red and green channels (NDRG), the standard deviation of image objects derived from the DSM (SD-DSM), and the normalized difference between the blue and red channels (NDBR), were commonly selected at all levels. The selection and incorporation of DSM-derived features, along with other selected features, remarkably contributed in the differentiation of spectrally similar classes, such as asbestos cement roofs, dark concrete roofs, old pavements, and asphalt. In a complex landscape without height information, an ML model might erroneously categorize bare soil as a roofing material or the opposite in accordance with the parallel spectral and textural characteristics. Al-Najjar et al. [7] utilized the fusion of DSM and optical images to generate generic automatic LULC classes for a complex urban area.

As stated in Section 3, the SS RF, SVM, and DT classification models were initially examined in the first study area at the optimal scales, identified through the *F*-measure along with the significant features, and selected by CFS at each optimal scale level. The SS classification results varied from one level to another when each SS model was applied, and the OA classification accuracies ranged from 79% to 92.2%. The comparison of SS classification maps showed that clear misclassifications were obtained by the DT algorithm, especially on clay tiles, asphalt, grass, and trees.

Following the FS and SS classification, iterative adaptive MS classification models were developed in the first study area and were tested in the second study area. The adopted hierarchical classification scheme of the first study area using the RF algorithm showed outstanding performance (OA = 94.4%) on the extraction of urban LULC classes compared to SVM (OA = 92.5%) and DT (OA = 91.6%). However, slight confusion was observed between different classes, as shown in Figure 7. For instance, MS-SVM showed a minor confusion between metal type1, asphalt, bare soil, and grass, as demonstrated in Figure 7b,k. Similarly, MS-DT showed a great confusion amongst the asphalt, grass, and metal type 1, as shown in Figure 7c. In addition, a remarkable confusion was found between some asbestos cement roofs, dark concrete roofs, asphalt, clay tiles type 2, and bare soil classes in some areas when MS-DT was used, as shown in Figure 7f,i,l. MS-RF showed an outstanding performance but exhibited a minor confusion between some asphalt objects that mixed with shadows as water bodies, as depicted in Figure 7d,j. The comparison of SS and MS approaches showed that the accuracy of some classes (i.e., trees, type 1 clay tiles, and asbestos classes) clearly improved with the use of the proposed approach.



Figure 7. MS classification results for four different regions of the first study area using RF, SVM, and DT: (a–c) first region, (d–f) second region, (g–i) third region, and (j–l) fourth region.

The applicability of the adopted scheme in the second study area indicated that MS-RF and MS-SVM exhibited relatively similar classification results. However, the MS-SVM algorithm (OA = 94.45% and $K = 0.938$) was superior to RF (OA = 92.46% and $K = 0.816$), with a slight improvement in the OA and K values. All MS algorithms showed some degrees of confusion between some objects with grass and water bodies, as shown in Figure 8j–l, which may be attributed to the existence of new water objects that vary in spectral characteristics as the second study area was classified on the basis of the sample statistics derived from the first study area. As represented in Figure 8j, the water body was poorly classified using MS-RF and showed confusion with the grass class. In this scenario, the present water body in the second study area was a pond with extremely different reflectance from the training samples obtained in the first study area and was more obvious in the MS-RF classified map compared to other algorithms because of the RF algorithm sensitivity to training. RF is sensitive to the size of training samples and the selection of an accurate representative of each class for classification [84].

Moreover, utilizing MS-DT resulted in misclassification between the tree and dark concrete classes, and between the grass and tree classes (Figure 8c). DT demonstrated a minor confusion between the bare soil, type 2 clay tiles, dark concrete, and asphalt, as shown in Figure 8c,f,i,l. As shown in Figure 8d,e, most of the roof types were categorized in an extremely similar manner by utilizing RF and SVM. MS-SVM showed a minor confusion in some areas between the asphalt and dark concrete, as shown in Figure 8h, whereas MS-RF showed a relatively better differentiation between asphalt and dark concrete for the same area, as shown in Figure 8g.

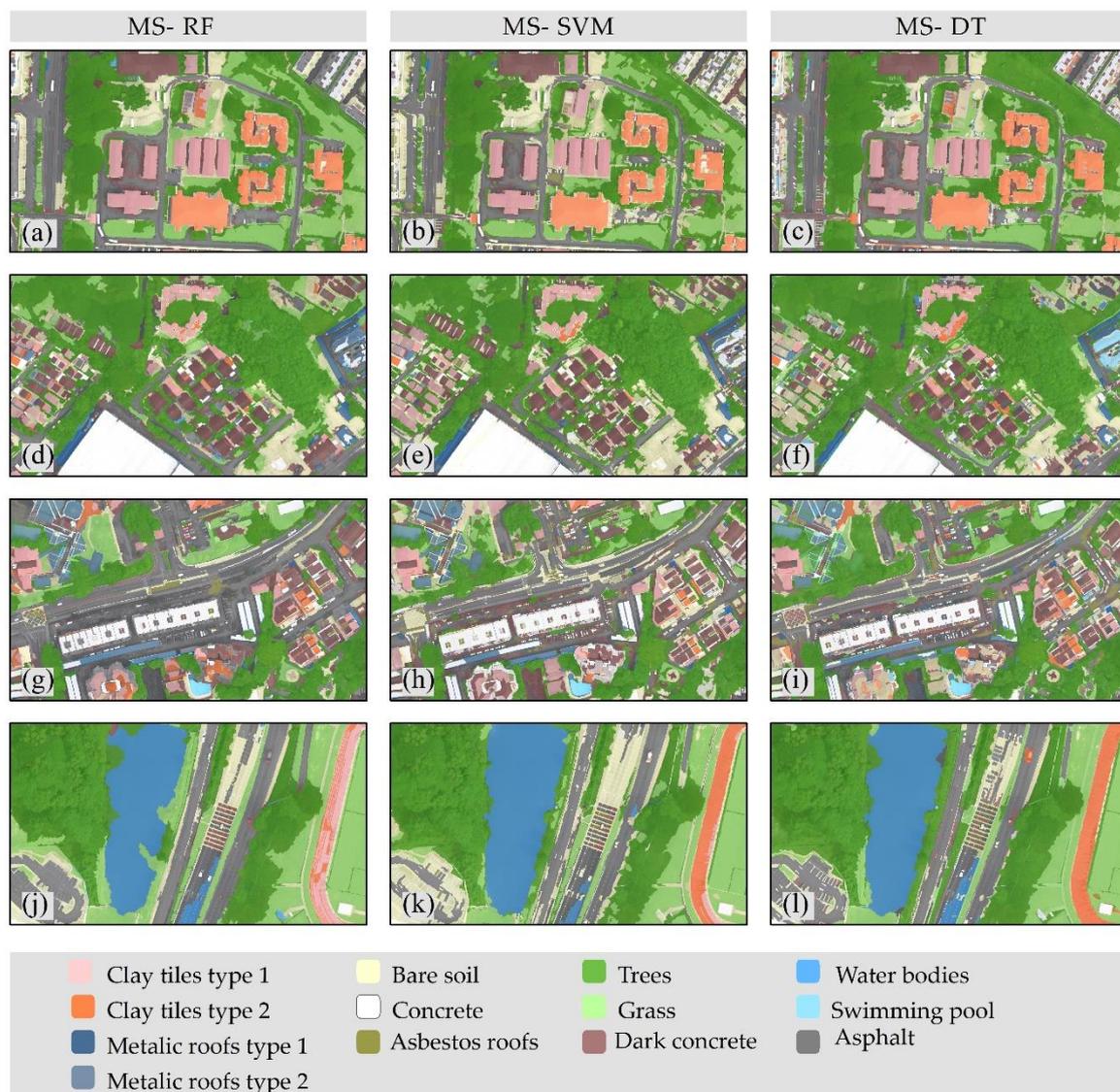


Figure 8. MS classification results of four different subsets of the second study area using RF, SVM, and DT: (a–c) first subset, (d–f) second subset, (g–i) third subset, and (j–l) fourth subset.

6. Conclusions

Accurate and up-to-date urban LULC information is crucial for urban planning, management and environmental applications. UAVs allow the acquisition of remotely sensed data with UHSR, as high as 1 cm, in a flexible and inexpensive manner, significantly contributing to the initiation of a wide spectrum of applications. This study aimed to achieve an accurate and detailed urban LULC classification in a heterogeneous landscape using GEOBIA and ML models from UHSR drone-based images. Given the high-level details of UAV images and the limited amount of spectral information, a MS GEOBIA approach that integrates MS image segmentation evaluation, MS FS, and hierarchical

ML classification algorithms was used to generate detailed LULC urban maps from the fusion of orthophotos and DSMs. Two UAV-based images were used to implement and evaluate the efficiency of the proposed method. Three commonly used supervised ML models, namely, RF, SVM, and DT, were compared within the MS/hierarchical segmentation and classification approach. The MS-RF classification achieved the highest accuracy, with an OA of 94.40% and a K of 0.938, followed by MS-SVM with an OA of 92.50% and a K of 0.917 and MS-DT with an OA of 91.60% and a K of 0.908. The applicability of the proposed approach to the dataset of the second study area showed excellent performance when MS-SVM and MS-RF were used. The proposed framework exhibited enormous potential for the detailed mapping of heterogeneous urban areas from UHSR RGB and DSM images. The results obtained from this approach can serve as vital information and input for scientists, decision makers, and city planners.

Author Contributions: B.K. and S.M. acquired the UAV data; M.B.A.G., B.K. and R.A.-R. conceptualized and performed the analysis; M.B.A.G., B.K., R.A.-R. and V.S. wrote the manuscript; N.U. supervised the funding acquisition; M.B.A.G., R.A.-R., A.S. and H.Z.M.S. edited, restructured, and professionally optimized the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: The APC is supported by the RIKEN Centre for Advanced Intelligence Project (AIP), Tokyo, Japan.

Acknowledgments: The authors would like to thank the University of Sharjah, RIKEN Centre for AIP, Japan, and Universiti Putra Malaysia for providing all facilities during the research.

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

References

1. Glade, T. Landslide occurrence as a response to land use change: A review of evidence from New Zealand. *Catena* **2003**, *51*, 297–314. [[CrossRef](#)]
2. Scanlon, B.R.; Reedy, R.C.; Stonestrom, D.A.; Prudic, D.E.; Dennehy, K.F. Impact of land use and land cover change on groundwater recharge and quality in the southwestern US. *Glob. Chang. Biol.* **2005**, *11*, 1577–1593. [[CrossRef](#)]
3. Bonato, M.; Cian, F.; Giupponi, C. Combining LULC data and agricultural statistics for A better identification and mapping of High nature value farmland: A case study in the veneto Plain, Italy. *Land Use Policy* **2019**, *83*, 488–504. [[CrossRef](#)]
4. Acheampong, M.; Yu, Q.; Enomah, L.D.; Anchang, J.; Eduful, M. Land use/cover change in Ghana's oil city: Assessing the impact of neoliberal economic policies and implications for sustainable development goal number one—A remote sensing and GIS approach. *Land Use Policy* **2018**, *73*, 373–384. [[CrossRef](#)]
5. McDowell, R.W.; Snelder, T.; Harris, S.; Lilburne, L.; Larned, S.T.; Scarsbrook, M.; Curtis, A.; Holgate, B.; Phillips, J.; Taylor, K. The land use suitability concept: Introduction and an application of the concept to inform sustainable productivity within environmental constraints. *Ecol. Indic.* **2018**, *91*, 212–219. [[CrossRef](#)]
6. Lin, Y.; Jiang, M.; Yao, Y.; Zhang, L.; Lin, J. Use of UAV oblique imaging for the detection of individual trees in residential environments. *Urban For. Urban Green.* **2015**, *14*, 404–412. [[CrossRef](#)]
7. Al-najjar, H.A.H.; Kalantar, B.; Pradhan, B.; Saeidi, V. Land cover classification from fused DSM and UAV images using convolutional neural networks. *Remote Sens.* **2019**, *11*, 1461. [[CrossRef](#)]
8. Ahmed, O.S.; Shemrock, A.; Chabot, D.; Dillon, C.; Williams, G.; Wasson, R.; Franklin, S.E. Hierarchical land cover and vegetation classification using multispectral data acquired from an unmanned aerial vehicle. *Int. J. Remote Sens.* **2017**, *38*, 2037–2052. [[CrossRef](#)]
9. Kalantar, B.; Mansor, S.B.; Sameen, M.I.; Pradhan, B.; Shafri, H.Z.M. Drone-Based land-cover mapping using a fuzzy unordered rule induction algorithm integrated into object-based image analysis. *Int. J. Remote Sens.* **2017**, *38*, 2535–2556. [[CrossRef](#)]
10. Zhang, X.; Chen, G.; Wang, W.; Wang, Q.; Dai, F. Object-Based land-cover supervised classification for very-high-resolution UAV images using stacked denoising autoencoders. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, *10*, 3373–3385. [[CrossRef](#)]
11. Akar, Ö. Mapping land use with using Rotation Forest algorithm from UAV images. *Eur. J. Remote Sens.* **2017**, *50*, 269–279. [[CrossRef](#)]

12. Mancini, A.; Frontoni, E.; Zingaretti, P. A multi/hyper-spectral imaging system for land use/land cover using unmanned aerial systems. In Proceedings of the 2016 International Conference on Unmanned Aircraft Systems (ICUAS), Arlington, VA, USA, 7–10 June 2016; Institute of Electrical and Electronics Engineers (IEEE): Piscataway, NJ, USA, 2016; pp. 1148–1155.
13. Iizuka, K.; Itoh, M.; Shiodera, S.; Matsubara, T.; Dohar, M.; Watanabe, K. Advantages of unmanned aerial vehicle (UAV) photogrammetry for landscape analysis compared with satellite data: A case study of postmining sites in Indonesia. *Cogent Geosci.* **2018**, *4*, 1–15. [[CrossRef](#)]
14. Randall, M.; Fensholt, R.; Zhang, Y.; Jensen, M.B. Geographic object based image analysis of world view-3 imagery for urban hydrologic modelling at the catchment scale. *Water* **2019**, *11*, 1133. [[CrossRef](#)]
15. Mboga, N.; Georganos, S.; Grippa, T.; Lennert, M.; Vanhuysse, S.; Wolff, E. Fully convolutional networks and geographic object-based image analysis for the classification of VHR imagery. *Remote Sens.* **2019**, *11*, 597. [[CrossRef](#)]
16. Gu, H.; Han, Y.; Yang, Y.; Li, H.; Liu, Z.; Soergel, U.; Blaschke, T.; Cui, S. An efficient parallel multi-scale segmentation method for remote sensing imagery. *Remote Sens.* **2018**, *10*, 590. [[CrossRef](#)]
17. Gibril, M.B.A.; Idrees, M.O.; Yao, K.; Shafri, H.Z.M. Integrative image segmentation optimization and machine learning approach for high quality land-use and land-cover mapping using multisource remote sensing data. *J. Appl. Remote Sens.* **2018**, *12*, 1. [[CrossRef](#)]
18. De Castro, A.I.; Torres-Sánchez, J.; Peña, J.M.; Jiménez-Brenes, F.M.; Csillik, O.; López-Granados, F. An automatic random forest-OBIA algorithm for early weed mapping between and within crop rows using UAV imagery. *Remote Sens.* **2018**, *10*, 285. [[CrossRef](#)]
19. Komárek, J.; Klouček, T.; Prošek, J. The potential of unmanned aerial systems: A tool towards precision classification of hard-to-distinguish vegetation types. *Int. J. Appl. Earth Obs. Geoinf.* **2018**, *71*, 9–19. [[CrossRef](#)]
20. Kamal, M.; Kanekaputra, T.; Hermayani, R.; Juniansah, A. Geographic object based image analysis (GEOBIA) for mangrove canopy delineation using aerial photography. *IOP Conf. Ser. Earth Environ. Sci.* **2019**, *313*, 12048. [[CrossRef](#)]
21. Mishra, N.; Mainali, K.; Shrestha, B.; Radenz, J.; Karki, D. Species-Level vegetation mapping in a Himalayan treeline ecotone using unmanned aerial system (UAS) imagery. *ISPRS Int. J. Geo Inf.* **2018**, *7*, 445. [[CrossRef](#)]
22. White, R.; Bomber, M.; Hupy, J.; Shortridge, A. UAS-GEOBIA approach to sapling identification in jack pine barrens after fire. *Drones* **2018**, *2*, 40. [[CrossRef](#)]
23. Gonçalves, J.; Pôças, I.; Marcos, B.; Múcher, C.A.; Honrado, J.P. SegOptim—A new R package for optimizing object-based image analyses of high-spatial resolution remotely-sensed data. *Int. J. Appl. Earth Obs. Geoinf.* **2019**, *76*, 218–230. [[CrossRef](#)]
24. Blaschke, T.; Hay, G.J.; Kelly, M.; Lang, S.; Hofmann, P.; Addink, E.; Queiroz Feitosa, R.; Van der Meer, F.; Van der Werff, H.; Van Coillie, F.; et al. Geographic object-based image analysis—Towards a new paradigm. *ISPRS J. Photogramm. Remote Sens.* **2014**, *87*, 180–191. [[CrossRef](#)]
25. Ma, L.; Li, M.; Ma, X.; Cheng, L.; Du, P.; Liu, Y. A review of supervised object-based land-cover image classification. *ISPRS J. Photogramm. Remote Sens.* **2017**, *130*, 277–293. [[CrossRef](#)]
26. Vincent, L.; Soille, P. Watersheds in digital spaces: An efficient algorithm based on immersion simulations. *IEEE Trans. Pattern Anal. Mach. Intell.* **1991**, *13*, 583–598. [[CrossRef](#)]
27. Deren, L.; Zhang, G.; Wu, Z.; Yi, L. An edge embedded marker-based watershed algorithm for high spatial resolution remote sensing image segmentation. *IEEE Trans. Image Process.* **2010**, *19*, 2781–2787. [[CrossRef](#)]
28. Benz, U.C.; Hofmann, P.; Willhauck, G.; Lingenfelder, I.; Heynen, M. Multi-Resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *ISPRS J. Photogramm. Remote Sens.* **2004**, *58*, 239–258. [[CrossRef](#)]
29. Comaniciu, D.; Meer, P. Mean shift: A robust approach toward feature space analysis. *IEEE Trans. Pattern Anal. Mach. Intell.* **2002**, *24*, 603–619. [[CrossRef](#)]
30. Wang, Y.; Meng, Q.; Qi, Q.; Yang, J.; Liu, Y. Region merging considering within-and between-segment heterogeneity: An improved hybrid remote-sensing image segmentation method. *Remote Sens.* **2018**, *10*, 781. [[CrossRef](#)]
31. Kavzoglu, T.; Tonbul, H. A comparative study of segmentation quality for multi-resolution segmentation and watershed transform. In Proceedings of the 8th International Conference on Recent Advances in Space Technologies (RAST), Istanbul, Turkey, 19–22 June 2017; pp. 113–117.

32. Zhang, X.; Du, S. Learning selfhood scales for urban land cover mapping with very-high-resolution satellite images. *Remote Sens. Environ.* **2016**, *178*, 172–190. [[CrossRef](#)]
33. Xiao, P.; Zhang, X.; Zhang, H.; Hu, R.; Feng, X. Multiscale optimized segmentation of urban green cover in high resolution remote sensing image. *Remote Sens.* **2018**, *10*, 1813. [[CrossRef](#)]
34. He, Y.; Weng, Q. *High Spatial Resolution Remote Sensing: Data, Analysis, and Applications*; CRC Press: Boca Raton, FL, USA, 2018; ISBN 9781498767682.
35. Shen, Y.; Chen, J.; Xiao, L.; Pan, D. Optimizing multiscale segmentation with local spectral heterogeneity measure for high resolution remote sensing images. *ISPRS J. Photogramm. Remote Sens.* **2019**, *157*, 13–25. [[CrossRef](#)]
36. Wang, Y.; Qi, Q.; Liu, Y.; Jiang, L.; Wang, J. Unsupervised segmentation parameter selection using the local spatial statistics for remote sensing image segmentation. *Int. J. Appl. Earth Obs. Geoinf.* **2019**, *81*, 98–109. [[CrossRef](#)]
37. Drăguț, L.; Tiede, D.; Levick, S.R. ESP: A tool to estimate scale parameter for multiresolution image segmentation of remotely sensed data. *Int. J. Geogr. Inf. Sci.* **2010**, *24*, 859–871. [[CrossRef](#)]
38. Drăguț, L.; Csillik, O.; Eisank, C.; Tiede, D. Automated parameterisation for multi-scale image segmentation on multiple layers. *ISPRS J. Photogramm. Remote Sens.* **2014**, *88*, 119–127. [[CrossRef](#)] [[PubMed](#)]
39. Johnson, B.; Bragais, M.; Endo, I.; Magcale-Macandog, D.; Macandog, P. Image segmentation parameter optimization considering within-and between-segment heterogeneity at multiple scale levels: Test case for mapping residential areas using landsat imagery. *ISPRS Int. J. Geo Inf.* **2015**, *4*, 2292–2305. [[CrossRef](#)]
40. Johnson, B.; Xie, Z. Unsupervised image segmentation evaluation and refinement using a multi-scale approach. *ISPRS J. Photogramm. Remote Sens.* **2011**, *66*, 473–483. [[CrossRef](#)]
41. Espindola, G.M.; Camara, G.; Reis, I.A.; Bins, L.S.; Monteiro, A.M. Parameter selection for region-growing image segmentation algorithms using spatial autocorrelation. *Int. J. Remote Sens.* **2006**, *27*, 3035–3040. [[CrossRef](#)]
42. Yang, J.; He, Y.; Weng, Q. An automated method to parameterize segmentation scale by enhancing intrasegment homogeneity and intersegment heterogeneity. *IEEE Geosci. Remote Sens. Lett.* **2015**, *12*, 1282–1286. [[CrossRef](#)]
43. Ming, D.; Li, J.; Wang, J.; Zhang, M. Scale parameter selection by spatial statistics for GeOBIA: Using mean-shift based multi-scale segmentation as an example. *ISPRS J. Photogramm. Remote Sens.* **2015**, *106*, 28–41. [[CrossRef](#)]
44. Yang, J.; Li, P.; He, Y. A multi-band approach to unsupervised scale parameter selection for multi-scale image segmentation. *ISPRS J. Photogramm. Remote Sens.* **2014**, *94*, 13–24. [[CrossRef](#)]
45. Kamal, M.; Phinn, S.; Johansen, K. Object-Based approach for multi-scale mangrove composition mapping using multi-resolution image datasets. *Remote Sens.* **2015**, *7*, 4753–4783. [[CrossRef](#)]
46. Pedernana, M.; Marpu, P.R.; Mura, M.D.; Benediktsson, J.A.; Bruzzone, L. A novel technique for optimal feature selection in attribute profiles based on genetic algorithms. *IEEE Trans. Geosci. Remote Sens.* **2013**, *51*, 3514–3528. [[CrossRef](#)]
47. Ma, L.; Fu, T.; Blaschke, T.; Li, M.; Tiede, D.; Zhou, Z.; Ma, X.; Chen, D. Evaluation of feature selection methods for object-based land cover mapping of unmanned aerial vehicle imagery using random forest and support vector machine classifiers. *ISPRS Int. J. Geo Inf.* **2017**, *6*, 51. [[CrossRef](#)]
48. Martha, T.R.; Kerle, N.; Van Westen, C.J.; Jetten, V.; Kumar, K.V. Segment optimization and data-driven thresholding for knowledge-based landslide detection by object-based image analysis. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 4928–4943. [[CrossRef](#)]
49. Ridha, M.; Pradhan, B. An improved algorithm for identifying shallow and deep-seated landslides in dense tropical forest from airborne laser scanning data. *Catena* **2018**, *167*, 147–159. [[CrossRef](#)]
50. Al-Ruzouq, R.; Shanableh, A.; Barakat, A.; Gibril, M.; AL-Mansoori, S.; Al-Ruzouq, R.; Shanableh, A.; Barakat, A.; Gibril, M.; AL-Mansoori, S. Image segmentation parameter selection and ant colony optimization for date palm tree detection and mapping from very-high-spatial-resolution aerial imagery. *Remote Sens.* **2018**, *10*, 1413. [[CrossRef](#)]
51. Hamedianfar, A.; Gibril, M.B.A. Large-Scale urban mapping using integrated geographic object-based image analysis and artificial bee colony optimization from worldview-3 data. *Int. J. Remote Sens.* **2019**, *40*, 6796–6821. [[CrossRef](#)]

52. Hamedianfar, A.; Gibril, M.B.A.; Pellikka, P.K.E. Synergistic use of particle swarm optimization, artificial neural network, and extreme gradient boosting algorithms for urban LULC mapping from WorldView-3 images. *Geocart. Int.* **2020**, 1–19. [[CrossRef](#)]
53. Shanableh, A.; Al-ruzouq, R.; Gibril, M.B.A.; Flesia, C. Spatiotemporal mapping and monitoring of whiting in the semi-enclosed gulf using moderate resolution imaging spectroradiometer (MODIS) time series images and a generic ensemble. *Remote Sens.* **2019**, *11*, 1193. [[CrossRef](#)]
54. Shahi, K.; Shafri, H.Z.M.; Hamedianfar, A. Road condition assessment by OBIA and feature selection techniques using very high-resolution WorldView-2 imagery. *Geocart. Int.* **2017**, *32*, 1389–1406. [[CrossRef](#)]
55. Cao, J.; Leng, W.; Liu, K.; Liu, L.; He, Z.; Zhu, Y. Object-based mangrove species classification using unmanned aerial vehicle hyperspectral images and digital surface models. *Remote Sens.* **2018**, *10*, 89. [[CrossRef](#)]
56. Akar, Ö. The Rotation Forest algorithm and object-based classification method for land use mapping through UAV images. *Geocart. Int.* **2018**, *33*, 538–553. [[CrossRef](#)]
57. Liu, L.; Liu, X.; Shi, J.; Li, A. A land cover refined classification method based on the fusion of LiDAR data and UAV image. *Adv. Comput. Sci. Res.* **2019**, *88*, 154–162. [[CrossRef](#)]
58. Fotheringham, A.S.; Brunson, C.; Charlton, M. *Quantitative Geography: Perspectives on Spatial Data Analysis*; SAGE: Thousand Oaks, CA, USA, 2000.
59. Grybas, H.; Melendy, L.; Congalton, R.G. A comparison of unsupervised segmentation parameter optimization approaches using moderate-and high-resolution imagery. *GISci. Remote Sens.* **2017**, *54*, 515–533. [[CrossRef](#)]
60. Trimble, T. *ECognition Developer 8.7 Reference Book*; Trimble Germany GmbH: Munich, Germany, 2011; pp. 319–328.
61. Gevers, T.; Smeulders, A.W.M. PicToSeek: Combining color and shape invariant features for image retrieval. *IEEE Trans. Image Process.* **2000**, *9*, 102–119. [[CrossRef](#)]
62. Sirmacek, B.; Unsalan, C. Building detection from aerial images using invariant color features and shadow information. In Proceedings of the 23rd International Symposium on Computer and Information Sciences, Istanbul, Turkey, 27–29 October 2008; pp. 1–5.
63. Cretu, A.M.; Payeur, P. Building detection in aerial images based on watershed and visual attention feature descriptors. In Proceedings of the International Conference on Computer and Robot Vision CRV 2013, Regina, SK, Canada, 28–31 May 2013; pp. 265–272. [[CrossRef](#)]
64. Haralick, R.M.; Shanmugam, K.; Dinstein, I. Textural features for image classification. *IEEE Trans. Syst. Man Cybern.* **1973**, *3*, 610–621. [[CrossRef](#)]
65. Hall, M.A.; Holmes, G. Benchmarking attribute selection techniques for discrete class data mining. *IEEE Trans. Knowl. Data Eng.* **2003**, *15*, 1437–1447. [[CrossRef](#)]
66. Shang, M.; Wang, S.; Zhou, Y.; Du, C.; Liu, W. Object-based image analysis of suburban landscapes using Landsat-8 imagery. *Int. J. Digit. Earth* **2019**, *12*, 720–736. [[CrossRef](#)]
67. Al-ruzouq, R.; Shanableh, A.; Mohamed, B.; Kalantar, B. Multi-scale correlation-based feature selection and random forest classification for LULC mapping from the integration of SAR and optical Sentinel images. *Proc. SPIE* **2019**, 11157. [[CrossRef](#)]
68. Hall, M.A.; Smith, L.A. Practical feature subset selection for machine learning. In Proceedings of the 21st Australasian Computer Science Conference ACSC'98, Perth, Australia, 4–6 February 1998; Volume 98, pp. 181–191.
69. Fong, S.; Zhuang, Y.; Tang, R.; Yang, X.; Deb, S. Selecting optimal feature set in high-dimensional data by swarm search. *J. Appl. Math.* **2013**, 2013. [[CrossRef](#)]
70. Raczko, E.; Zagajewski, B. Comparison of support vector machine, random forest and neural network classifiers for tree species classification on airborne hyperspectral APEX images. *Eur. J. Remote Sens.* **2017**, *50*, 144–154. [[CrossRef](#)]
71. Tzotsos, A. Preface: A support vector machine for object-based image analysis. In *Approach Object Based Image Analysis*; Springer: Cham, Switzerland, 2008; Volume 5–8. [[CrossRef](#)]
72. Taner San, B. An evaluation of SVM using polygon-based random sampling in landslide susceptibility mapping: The Candir catchment area (western Antalya, Turkey). *Int. J. Appl. Earth Obs. Geoinf.* **2014**, *26*, 399–412. [[CrossRef](#)]
73. Cortes, C.; Vapnik, V. Support-Vector networks. *Mach. Learn.* **1995**, *20*, 273–297. [[CrossRef](#)]

74. Pradhan, B.; Seeni, M.I.; Kalantar, B. Performance evaluation and sensitivity analysis of expert-based, statistical, machine learning, and hybrid models for producing landslide susceptibility maps. In *Laser Scanning Applications in Landslide Assessment*; Springer International Publishing: Cham, Switzerland, 2017; pp. 193–232.
75. Pal, M.; Mather, P.M. An assessment of the effectiveness of decision tree methods for land cover classification. *Remote Sens. Environ.* **2003**, *86*, 554–565. [[CrossRef](#)]
76. Hamedianfar, A.; Shafri, H.Z.M. Integrated approach using data mining-based decision tree and object-based image analysis for high-resolution urban mapping of WorldView-2 satellite sensor data. *J. Appl. Remote Sens.* **2016**, *10*, 025001. [[CrossRef](#)]
77. Pal, M.; Mather, P.M. Decision tree based classification of remotely sensed data. In *Proceedings of the 22nd Asian Conference on Remote Sensing, Singapore, 5–9 November 2001*; p. 9.
78. Eisavi, V.; Homayouni, S.; Yazdi, A.M.; Alimohammadi, A. Land cover mapping based on random forest classification of multitemporal spectral and thermal images. *Environ. Monit. Assess.* **2015**, *187*, 1–14. [[CrossRef](#)]
79. Berhane, T.M.; Lane, C.R.; Wu, Q.; Autrey, B.C.; Anenkhonov, O.A.; Chepinoga, V.V.; Liu, H. Decision-Tree, rule-based, and random forest classification of high-resolution multispectral imagery for wetland mapping and inventory. *Remote Sens.* **2018**, *10*, 580. [[CrossRef](#)]
80. Goetz, J.N.; Brenning, A.; Petschko, H.; Leopold, P. Computers & geosciences evaluating machine learning and statistical prediction techniques for landslide susceptibility modeling. *Comput. Geosci.* **2015**, *81*, 1–11. [[CrossRef](#)]
81. Tian, S.; Zhang, X.; Tian, J.; Sun, Q. Random forest classification of wetland landcovers from multi-sensor data in the arid region of Xinjiang, China. *Remote Sens.* **2016**, *8*, 954. [[CrossRef](#)]
82. Zerrouki, N.; Bouchaffra, D. Pixel-based or object-based: Which approach is more appropriate for remote sensing image classification? In *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics (SMC), San Diego, CA, USA, 5–8 October 2014; Volume 2014*, pp. 864–869.
83. Li, W.; Guo, Q. A new accuracy assessment method for one-class remote sensing classification. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 4621–4632. [[CrossRef](#)]
84. Millard, K.; Richardson, M. On the importance of training data sample selection in random forest image classification: A case study in peatland ecosystem mapping. *Remote Sens.* **2015**, *7*, 8489–8515. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).