sustainability

MDPI

Article

Transfer-Ensemble Learning: A Novel Approach for Mapping Urban Land Use/Cover of the Indian Metropolitans

Prosenjit Barman [1], Sheikh Mustak [1,*], Monika Kuffer [2] and Sudhir Kumar Singh [3]

[1] Department of Geography, Central University of Punjab, Bathinda 151401, India; prosenjitkm147@gmail.com
[2] Faculty of Geo-Information Science and Earth Observation, ITC, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands; m.kuffer@utwente.nl
[3] K. Banerjee Centre of Atmospheric and Ocean Studies, IIDS, University of Allahabad, Prayagraj 211002, India; sudhirinjnu@gmail.com
* Correspondence: sk.mustak@cup.edu.in; Tel.: +91-7047680305

Abstract: Land use and land cover (LULC) classification plays a significant role in the analysis of climate change, evidence-based policies, and urban and regional planning. For example, updated and detailed information on land use in urban areas is highly needed to monitor and evaluate urban development plans. Machine learning (ML) algorithms, and particularly ensemble ML models support transferability and efficiency in mapping land uses. Generalization, model consistency, and efficiency are essential requirements for implementing such algorithms. The transfer-ensemble learning approach is increasingly used due to its efficiency. However, it is rarely investigated for mapping complex urban LULC in Global South cities, such as India. The main objective of this study is to assess the performance of machine and ensemble-transfer learning algorithms to map the LULC of two metropolitan cities of India using Landsat 5 TM, 2011, and DMSP-OLS nightlight, 2013. This study used classical ML algorithms, such as Support Vector Machine-Radial Basis Function (SVM-RBF), SVM-Linear, and Random Forest (RF). A total of 480 samples were collected to classify six LULC types. The samples were split into training and validation sets with a 65:35 ratio for the training, parameter tuning, and validation of the ML algorithms. The result shows that RF has the highest accuracy (94.43%) of individual models, as compared to SVM-RBF (85.07%) and SVM-Linear (91.99%). Overall, the ensemble model-4 produces the highest accuracy (94.84%) compared to other ensemble models for the Kolkata metropolitan area. In transfer learning, the pre-trained ensemble model-4 achieved the highest accuracy (80.75%) compared to other pre-trained ensemble models for Delhi. This study provides innovative guidelines for selecting a robust ML algorithm to map urban LULC at the metropolitan scale to support urban sustainability.

Keywords: land use/land cover; machine learning; remote sensing; transferability; ensemble learning

Citation: Barman, P.; Mustak, S.; Kuffer, M.; Singh, S.K. Transfer-Ensemble Learning: A Novel Approach for Mapping Urban Land Use/Cover of the Indian Metropolitans. *Sustainability* **2023**, *15*, 16593. https://doi.org/10.3390/su152416593

Academic Editors: Yao-Min Fang, Thanh-Van Hoang and Duc-Dung Tran

Received: 12 September 2023
Revised: 21 October 2023
Accepted: 10 November 2023
Published: 6 December 2023

1. Introduction

Land use/land cover (LULC) classifications are essential in several fields, such as monitoring climate change [1] regional urban planning development [2], and policy generation [3]. According to the United Nations report (2019), the world's 4.3 billion urban population will reach around 9.8 billion by 2050. More than twice as many people, approximately 6.7 billion, will live in cities [4]. The Census of India (2011) revealed that out of a total of 1027 million population, 285 million (27.1%) lived in urban areas, while 742 million lived in rural areas. Overall, there is a vast decadal increase in the urban population, up to 31.16 percent [5]. Urbanization affects the surrounding valuable natural landscapes like wetlands, open spaces, and green spaces [6]. The conversion of impervious surfaces impacts the ecosystem, biological diversity, climate, etc., creating adverse effects like heat islands [7]. The provision of information for the dynamic monitoring and management of the Earth depends heavily on the real-time availability and accuracy of

land use maps [8]. Feature extraction and Machine Learning (ML) approaches in land use classification created new opportunities for more precise and extensive land use mapping [9]. To monitor and evaluate urban development, up-to-date and detailed information on multiple land uses in an urban area is required [10]. Updated land use and land cover information and the measure of built-up area growth are crucial inputs for measuring the urban Sustainable Development Goal (SDG 11). Classical ML is commonly used to map built-up growth, with recent attention on evaluating the model's transferability [11]. Robust and transferable ML approaches are essential for supporting various SDGs related to environmental management, flood risk modelling, urban planning development, and sustainable development [12,13].

LULC classification (including waterbodies, vegetation, built-up areas, crops, etc.) [14] that employs Earth Observation (EO) data is proficient with various classical techniques like maximum likelihood pixel-based classification [15,16] object-based classification [17,18], ML [19–22] and deep learning techniques [23]. EO datasets with robust ML algorithms play a significant role in LULC classification [20]. The commonly used ML algorithm in LULC classification was categorized as (a) fast naïve Bayes Statistical learning algorithm, (b) perceptron-based method based on Window and Perceptron, (c) Random Forests, CART and Gmo Max Entropy algorithm of logic-based, and finally, (d) a support vector machine (SVM) based on Margin SVM, Voting SVM, Pegasus, and IKPamir [19,24]. The ML algorithms such as SVM [15], artificial neural Network [20], random forest [25], etc., were applied in different disciplines. ML is widely used in various fields, such as landslide susceptibility mapping [26,27], gully network detection [28], groundwater potentiality mapping [29,30], and LULC prediction [31], etc.

Integrating Night-time light data, land surface temperature, and socioeconomic datasets can improve the accuracy of land use maps [32]. Malarvizhi et al. (2016) [33] proposed an alternative of high-resolution images, i.e., Google Earth, the "Elshayal smart Open-source software" for ward-wise land-use mapping and urban change detection analysis in the Vellore district of Tamil Nadu state. Feizizadeh et al. (2021) [19] applied different ML algorithms to classify land use and land cover based on a time series Landsat of images using Google Earth Engine (GEE). In this study, ML algorithms such as the SVM, RF, Classification, and regression tree were applied and compared to determine accuracy using the Dempster-Shafer theory (DST). The result showed that the SVM outperformed the other ML algorithms. In addition, ensembles of ML algorithms produce a higher accuracy than single ML algorithms. Thus, ML algorithms are robust to solve the complex land use and land cover classification problem.

Chen et al. (2021) [34] proposed a robust and cost-effective framework to map the essential land use using various public domain datasets in a block unit. The high-resolution satellite image and open-source data were integrated for the LULC classification. ML algorithms such as SVM, RF, LightGBM, CatBoost, Neural network, and automatic ensemble ML algorithms were applied where the ensemble of ML algorithms produces higher accuracy than a single ML algorithm. Commonly, high-resolution data and socioeconomic data are unavailable in the Global South. To overcome this limitation, time series image classification, feature extraction, open street datasets, algorithm parameter tuning, and domain adaptation can be an alternative to improve land use classification accuracy. Ma et al. (2017) evaluated advanced feature selection method's effects on well-known supervised classifiers like SVM and RF. Li et al. (2014) applied segmentation-based classification based on Landsat TM data and concluded that the SVM algorithm successfully classified images with a limited training sample [35,36].

Prasad et al. (2022) [37] evaluated different satellite sensors such as Landsat, Sentinel, LISS III, and LISS IV. For example, they combined Sentinel 1 and 2 in land use classification using the various ML algorithms. The result also shows that feature extraction plays a vital role in improving the classification accuracy. Combining high spectral and spatial resolution satellite sensors gives a higher accuracy than the single high spatial resolution. However, to achieve high accuracy levels in large-area (e.g., regional scale) mapping, parameter

tuning is a critical factor. For the algorithm's robustness, the transferability assessment of a model is perilous. The robustness of the model helps to generalize, model consistency, and efficacy [8]. Basheer et al. (2022) [38] suggested several robust classification methods and quality datasets for LULC mapping, e.g., SVM, RF, CART, ArcGIS Pro, and Google Earth Engine to map LULC using Landsat TM, Sentinel-2, and PlanetScope datasets. The result showed that the SVM algorithm and the PlanetScope data outperformed other datasets and algorithms. The study also showed that the fusion and ensemble of ML algorithms improved classification accuracy over a single ML algorithm [8,34].

Transfer learning (e.g., ensemble learning) is a popular remote sensing-based LULC classification strategy. This strategy improves the classification accuracy over complex datasets and outcomes [39]. For example, Praveen et al. (2019) [40] assessed the transferability of ML algorithms (SVM, RF, CART) in agriculture LULC mapping. In this study, feature extraction, feature selection, and parameter tuning were performed using the cloud-based platform GEE. Naushad et al. (2021) [41] stressed that the computational cost of the transfer learning in LULC classification outperforms the single ML-based LULC classification. Farda (2017) [24] explored the ML algorithm and its accuracy level in multi-temporal land use mapping. The extracted features such as principal component analysis (PCA), vegetation indices, and Gray-Level Co-Occurrence Matrix (GLCM), timeseries Landsat datasets, and ten selected ML algorithms were applied using the GEE to map LULC classification.

The reviewed literature shows that ML techniques have been extensively used in land use classification, land use prediction, and transferability analysis. The ML technology has become more prevalent in land use classification because of its transferability, robustness, and optimum computation cost. It gives a new paradigm in land use classification using remote sensing technology. Most of the studies were focused on land use classification using simple ML algorithms in global north cities, and very few limited studies were carried out on ensembles and transfer ML models to map the Indian complex cities, especially, Kolkata and Delhi metropolitan cities. To improve the land use and land cover classification accuracy for the complex cities in the temporal and spatial domains, the ensemble and transfer learning models outperform the ML algorithm. The ML algorithms (e.g., SVM, RF, etc.) are primarily affected by the feature dimension, unbalanced samples, and computation cost, which can be resolved in the ensemble and transfer learning model by employing robust feature selection, parameter tuning, k-old cross validation and transferability analysis. Therefore, the main objective of this study is to assess the performance of machine and ensemble-transfer learning algorithms to map the LULC of two metropolitan cities of India using EO datasets. Outcomes will help diverse land user mappers, scientists, and land use policy makers to implement a proposed robust model to address SDGs (e.g., SDG-11, 13, 15).

The entire research article has been structured into six sections. After the introduction, the following section provides an overview of the study area. Section 3 describes the datasets and methodology, the Section 4 provides the analysis of the results, the Section 5 is a discussion of the results while the last section provides the main conclusions.

2. Study Area

Kolkata Metropolitan Area (KMA) is the largest metropolis in eastern India, the second largest in India, and the tenth largest globally. The oldest metropolitan area in India is Kolkata, which has been inhabited for 300 years (Kolkata Metropolitan Development Authority, 2005) [42]. The latitudinal and longitudinal extension of the Kolkata metropolitan area is 88°02′ E to 88°32′ E, and 22°19′ N to 23°01′ N. The Kolkata Metropolitan Area spreads over 1851.41 sq. km and consists of three Municipal Corporations, 39 municipalities, and 28 Panchayat Samity within Kolkata Metropolitan Development Authority (KMDA) (Figure 1). According to the 2011 census, the population of KMA was 14.60 million, and its density was 8000 per sq. km. After 2011, the Haringhata Municipality of the KMA was established. As a result, neither the Census nor the KMA statistics from 2011 have information on Haringhata municipality. The last Municipal Corporation of KMA is Bidhannagar, which was formed by merging Bidhannagar and Rajarhat Gopalpur Municipality in 2015. The two amalgamated municipalities that comprise the current Bidhannagar Municipal Corporation (BMC) would

be counted individually in the Census of 2011. Additionally, the data for the newly established BMC is still pending. KMA encompasses six South Bengal districts: Kolkata, Nadia, North 24-Paraganas, South 24-Paraganas Haora, and Hugli (KMDA, 2017) [43]. The surrounding rural areas of the municipalities cover more than 40 percent of the Kolkata metropolitan area. The rural area has changed drastically in demography, transportation network, economy, etc. [44]. Delhi is the national capital of India, and it extends between 76°84′ E, 28°41′ N to 77°35′ E, 28°88′ N (Figure 1). According to the 2011 Census of India, Delhi's city population was over 11 million, while the NCT's population was 16.8 million. Delhi covers an area of 1483 sq. km. It falls under the Delhi metropolitan city. As per the 2011 Census of India, Delhi comprises nine administrative districts [45]. Delhi territory has a boundary line with Uttar Pradesh and Haryana state, and it is located on the right bank of the river Yamuna [46].

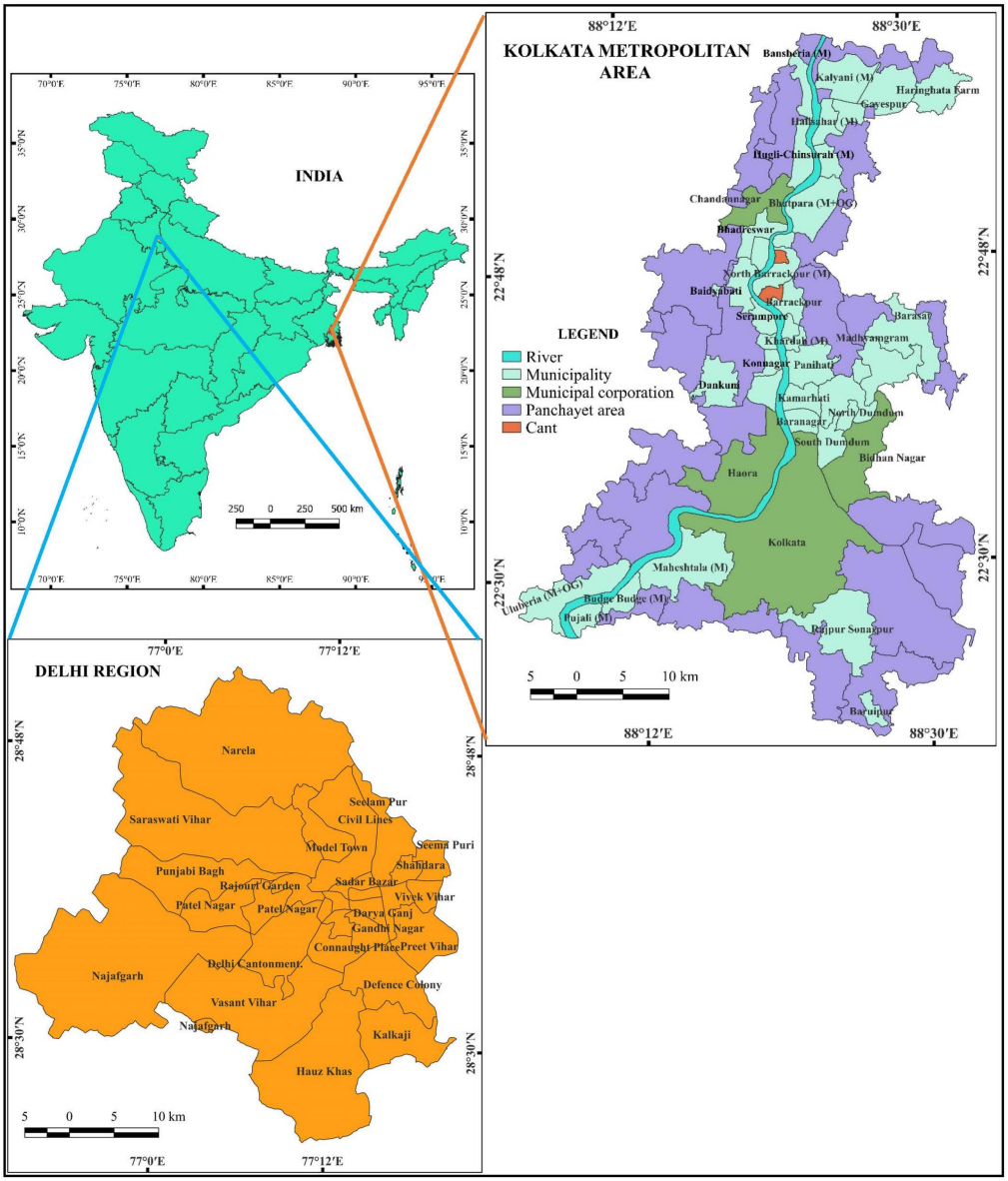


Figure 1. Location of the Kolkata and Delhi Metropolitan areas.

3. Datasets and Methodology

The transferability of the ML algorithms for the LULC mapping was assessed using EO datasets to develop a robust model. Classical ML algorithms were employed, fused,

compared, and validated to develop the best ML model for an accurate LULC map for the urban area (Figure 2).

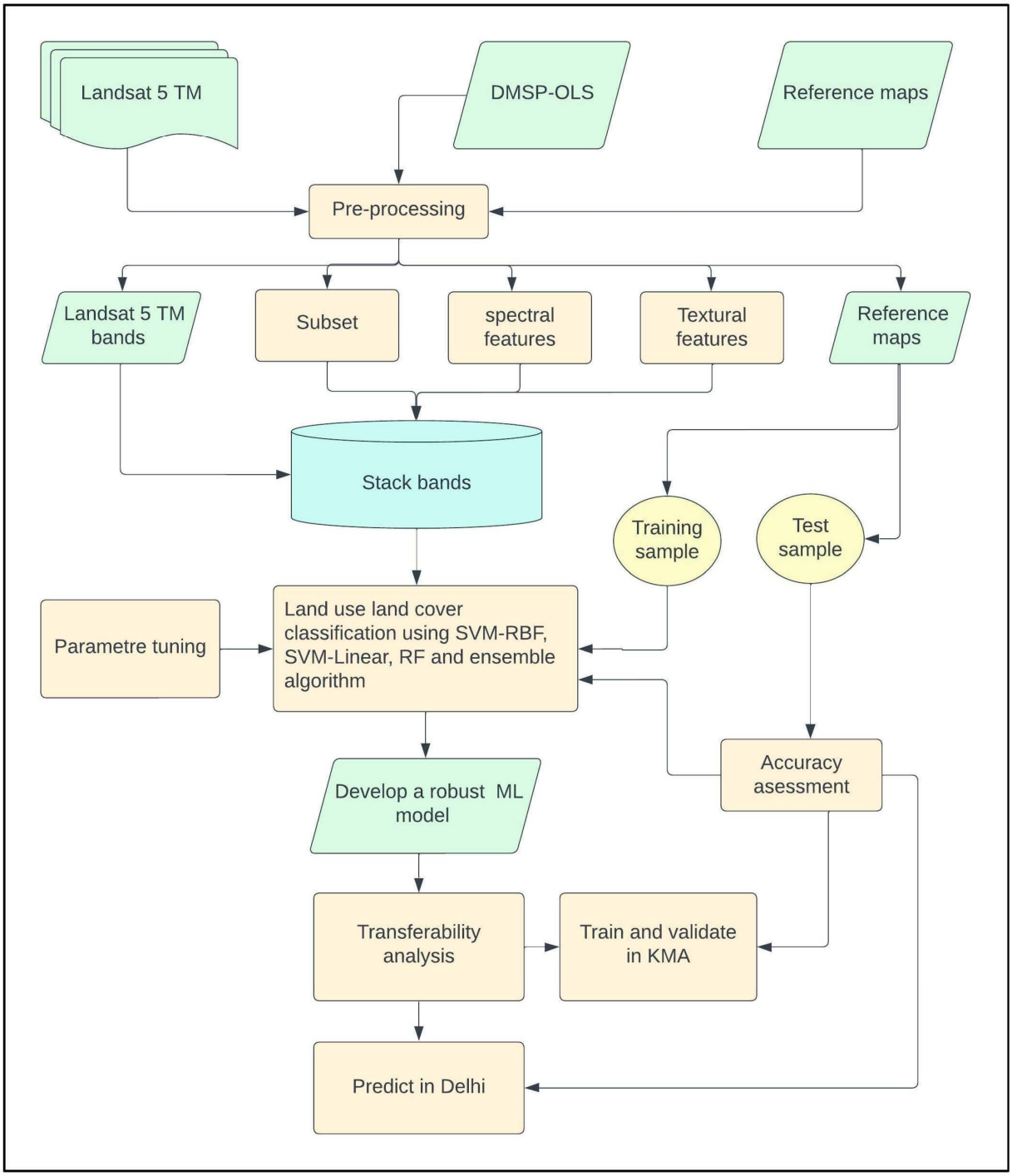


Figure 2. Methodological flowchart.

3.1. Datasets and Software Used

Landsat 5 TM (2011) was collected, using a cloud-based web platform (e.g., GEE), to analyse the transferability of the ML algorithms for LULC classification. Landsat TM was used because Defense Meteorological Satellite Program Operational Line-Scan System (DMSP-OLI) night-time light data is available from 1992 to 2013 [47]. In addition, the quality of the Landsat 7 ETM+ was not sufficient to map LULC in 2011. The Landsat 7 ETM+ data of the Kolkata metropolitan area was covered with fragmented strips and cloud issue for

the 2011 datasets. Therefore, Landsat 5 TM data of 30 m resolution have been used. Details of Landsat TM 5 satellite images are provided in Table 1.

Table 1. Landsat 5 TM (2011) bands have been used in the study (Kolkata and Delhi).

Bands	Spectral Bands	Spectral Range	Resolution		
			Spatial (Metre)	Temporal	Radiometric
B1	Blue	0.45–0.52 µm	30	16 days	7 bits
B2	Green	0.52–0.60 µm	30		
B3	Red	0.63–0.69 µm	30		
B4	Near-Infrared	0.76–0.90 µm	30		
B5	Near-Infrared	1.55–1.75 µm	30		
B7	Mid-Infrared	2.08–2.35 µm	30		

The DMSP-OLS provides cloud-free night-time light data, which were collected from the National Centres for Environmental Information (NOAA) website. The data provide night light data of average lights, derived from the average visible band digital number (DN) of light multiplied by the percent frequency of light detected [48]. The DMSP-OLS night light data improve the LULC classification accuracy, especially for the built-up area extraction [49]. In this study, open-source software, e.g., QGIS, and OTB (Orfeo Toolbox, https://www.orfeo-toolbox.org/ (accessed on 10 March 2023)) were used to assess the transferability of the ML algorithms. Other reference datasets such as Google Earth images, master plan map, and GPS field survey reference points were used to train and validate the ML algorithms.

3.2. Methodology

The methodological section is divided into pre-processing, feature extraction, normalization, parameter tuning, transferability analysis, LULC classification and validation (Figure 2).

3.2.1. Pre-Processing

Pre-processing is an essential step in data mining. Data pre-processing is a way to convert raw data into a much-desired form to obtain valuable information easily [50]. Techniques like Georeferencing and image subsets were pre-processed for the satellite images and mosaicked images into band sets. Atmospheric correction (e.g., dark-object substruction) and clipping, was applied to pre-process satellite images to develop high-quality and reliable information. Band stacking is another important step to create a false colour composite image to improve image interpretation quality. As explained below, the stack bands were used to extract several features.

3.2.2. Feature Extraction

Feature extraction is essential to contextualize the LULC information [51]. Integration of spectral bands with spectral, textural, morphological, and contextual features could improve classification accuracy because each land use/cover has a distinct shape, size, tone, and texture on the satellite image [51,52]. The vast number of variables involved in one sophisticated data analysis could also improve the classification accuracy [53]. In this study, a total 14 features (e.g., 02 spectral, 12 texture) were extracted and used for LULC classification, which has been explained below.

- Spectral features

The spectral features are important parameters to contextualize the spectral properties of the LULC in the image, which outperform a single spectral band of an image [52,54]. In this study, two spectral features, such as Normalized Difference Vegetation Index (NDVI) and Normalized Difference Built-up Index (NDBI), were extracted using the following Equations (1) and (2) [55].

$$\text{NDVI} = \frac{(\text{Band4} - \text{Band3})}{(\text{Band4} + \text{Band3})} \quad (1)$$

$$\text{NDBI} = \frac{(\text{Band5} - \text{Band4})}{(\text{Band5} + \text{Band4})} \quad (2)$$

The value of NDVI varies from +1 to −1 and +1 explains the existence of health greenspace while −1 shows waterbodies and other than greenspace. In addition, NDBI value varies from +1 to −1, which explains purely built-up (NDBI = +1) and non-built-up area (NDBI = −1) [55,56].

- Textural features

The Gray level co-occurrence matrix (GLCM) are second order's statistical texture characteristics from an image [53]. The GLCM are robust texture features widely used in improving the LULC classification [57]. The GLCM-mean and GLCM-variance features were selected and extracted from all the six bands of the Landsat 5 TM image, because GLCM-mean and variance are robust and outperforms the other GLCM features [58,59]. The GLCM features were extracted using the QGIS; see Equations (3) and (4). A total of 12 GLCM features (e.g., 6 GLCM-Mean and 6 GLCM-Variance) were extracted in this study.

$$\text{GLCM-Mean (MEA)} = \sum_{ij=0}^{N-1} i(p_{ij}) \quad (3)$$

$$\text{GLCM-Variance (VAR)} = \sum_{ij=0}^{N-1} p_{ij}(i - MEA)^2 \quad (4)$$

3.2.3. Normalization

Data normalization is one of the important steps in data science and ML because the performance of ML algorithms depends on how data has been normalized [60]. In this study, a simple max-min algorithm was used to normalize the stack bands using Equation (5), because this algorithm is robust and widely practised in ML-based LULC classification [61].

$$\text{Normalization} = \frac{(\text{Image} - \text{Image}_{\text{minimum}})}{(\text{Image}_{\text{maximum}} - \text{Image}_{\text{minimum}})} \quad (5)$$

The normalized bands were used in ML algorithms to classify the LULC of Kolkata and Delhi metropolitans. The value of the normalized image is varied from 0 to 1.

3.2.4. Selection of LULC

LULC classes were selected based on the literature review, the local LULC classification scheme, and the National Urban Information System (NUIS, 2006) [62]. In this study, six LULC classes were selected, i.e., built-up, cropland, fallow land, vegetation, open land, and waterbodies. These LULC classes are directly related to ecosystem services, planning, and management of the study area to support urban sustainability [63]. The selected LULC classes establish a baseline for the change analysis, the distribution and pattern of LULC classes in a city region [64]. Table 2 shows the selected LULC classes in this study.

Table 2. Selection of LULC classes.

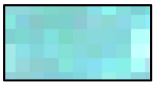
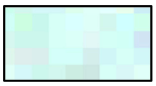
LULC Code	LULC Classes	Descriptions	Class Description Example (SFCC)
1	Built-up	Residential areas including urban, rural, industrial, all kinds of roads, and generally human made area	

Table 2. *Cont.*

LULC Code	LULC Classes	Descriptions	Class Description Example (SFCC)
2	Cropland	Land use for cultivation area covered by agricultural crops	
3	Fallow land	Agriculture lands without crops,	
4	Vegetation	Mostly dominated by forest, sparse vegetation, plantation	
5	Open land	Land use without vegetation, built up, and not any important object.	
6	Waterbody	Waterbodies like rivers, lakes, ponds, reservoir	

3.2.5. Collection of Training and Test Samples

A crucial component of image classification is the use of training and testing samples. This split helps to analyse the performance of ML models to predict the data. Training and test samples were assembled by digitizing polygons from visual classification and interpretation of various LULC classes on various band combinations of Landsat 5 TM satellite images, Google Earth images, and LULC reference maps [65]. In previous studies, the collected samples were divided into training and test samples in various ratios, including 60:40, 50:50, 67:33, and 80:20 [66,67]. Numerous techniques, including simple random sampling and the stratified sample strategy, were employed to acquire the training samples to avoid misclassification issues. Three diverse stratified sampling methods such as stratified equal random sampling, stratified proportional random sampling, stratified systematic sampling, and the binomial minimum fifty-sample rule, were employed to collect optimum training samples [40,68,69]. In this study, the binomial minimum fifty-sample rule with stratified random sampling was employed to collect a minimum of 50 samples per LULC class, dividing them into a 65:35 ratio as training (65%) and test samples (35%) (Table 3). Training samples were used to train the model, while test samples were used to validate the predicted result. The number of samples are less because the ground truth sample collection in large metropolitan cities is very complex, and the objective of this study is to develop a low-cost mapping system at city scale.

Table 3. Training and test samples.

LULC Code	LULC Classes	Kolkata Metropolitan Area		Delhi
		Training Samples	Test Samples	Test Samples
1	Built-up	66	34	56
2	Cropland	54	30	44
3	Fallow Land	48	30	30
4	Vegetation	58	32	42
5	Open Land	32	12	14
6	Waterbody	54	30	32
Total		312	168	218

3.2.6. Selection of ML Algorithms

ML is a field of study that turns empirical data into usable models using a computational algorithm. This field grew in traditional statistics and artificial intelligence communities [70]. In this study, classical ML like SVM, RF, etc., were used to map lower order LULC classes to reduce the computation cost over the large geographical area with lower-resolution data because deep learning requires huge computation costs and high-resolution data and is mostly suitable for higher order LULC classification [71–73].

- Support vector machine

The SVM is a data-driven method to solve the classification task. It gives a lower prediction error than other classification methods like Neural networks, especially in large datasets [74]. The most common method of SVM is the linear classifier, and SVM-RBF, in most cases outperformed linear classifier. The SVM created a hyperplane or set of hyperplanes to classify all inputs in a high dimensional [75].

In SVM classification, there exists a hyperplane that distinguishes the pattern. The geometric place of the separating hyperplane defined by the value of function f_{SVM}: $X \rightarrow R$ becomes null:

$$f_{SVM}(X) = (W, X) + b \tag{6}$$

where W represents the orthogonal vector to the separating hyperplane, f_{SVM} (X) = 0, and b defines the scalar value to represent the distance from the hyperplane to the origin of attribute space based on the equation $|b| / \|W\|$. The parameters W and b were obtained from the given equation (Equation (6)) to solve the optimization problem using the patterns of the training dataset [76].

In non-linear SVM classification, a nonlinear mapping function was utilised to separate two classes (Equation (9)) in the context of the relevant training sample from the hyperplane (H). The separated class was performed based on margin maximisation to solve the primal quadratic optimization problem [77]:

$$H : y_i \quad f(\varphi(x_i) = 0 \tag{7}$$

$$\min_{wb\xi} \quad \left\{ \frac{1}{2} \quad \|w\|^2 + \quad C\sum_{i=1}^{n} \xi_i \right\} \tag{8}$$

$$\text{Subject to :} \quad y_{i.}\left(w.\varphi \quad (x_i) + b\right) \geq 1 - \xi_{i,} i = 1, \ldots\ldots, n$$

$$\xi_i > 0, i = 1, \ldots\ldots, n$$

$$\text{Class separation subject to :} \; y_i f\left(\varphi \quad (x_i)\right) > 0 \Longrightarrow y_i = +1$$

$$\text{And } y_i f(\varphi(x_i)) < 0 \Longrightarrow y_i = -1 \tag{9}$$

where, w= weight vector, b = bias, and c = parameter correspond to the cost of wrong classification. C value assists in lowering the maximum marginal distance and provides a best fit hyperplane, which helps to solve the misclassification issue.

The primary quadratic optimization issue was resolved by employing the Lagrange multiplier to address the dual quadratic optimization problem. To build the SVM method, the kernel function was used in place of the nonlinear mapping function. In this study, SVM-RBF and SVM-linear kernel were replaced using Equations (10)–(13), respectively:

$$kRBF\left(x_{i,} \quad x_j\right) = e^{-\frac{\| \quad x_{i-} \quad x_j \quad \|^2}{2\sigma^2}} \tag{10}$$

$$\max_{n} \sum_{i=1}^{m} - \quad \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m} n_i \quad n_j \quad y_i \quad y_j \quad kRBF\left(x_{i,} \quad x_j\right) \tag{11}$$

$$\text{Subject to}: \begin{cases} 0 \leq & n_i & \leq C & i = 1, \ldots, m \\ & \sum_{i=1}^{m} n_i & y_i & = 0 \end{cases}$$

$$\text{kLinear}(x_i, \quad x_j) = \langle x_i, \quad x_j \rangle \tag{12}$$

$$\max_{n} \sum_{i=1}^{m} - \quad \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} n_i \quad n_j \quad y_i \quad y_j \quad \text{kLinear}(x_i, \quad x_j) \tag{13}$$

$$\text{Subject to}: \begin{cases} 0 \leq & n_i & \leq C & i = 1, \ldots, m \\ & \sum_{i=1}^{m} n_i & y_i & = 0 \end{cases}$$

where, n = Lagrange multiplier, and σ = bandwidth in kernel function is determined by the median distance of training sample. y = {−1, +1} is the set of class indicator [76–78].

- **Random forest**

A RF classifier is an ensemble classifier that produces multiple decision trees using a randomly selected subset of training samples and variables. This classifier has become popular in Remote sensing due to its classification accuracy [79]. In the RF algorithm, the margin function is used to measure the extent to which the average number of votes at X,Y for the right class exceeds that for the wrong class, and defines the margin function [80] as:

$$mg(X, Y) = av_k \quad I(h_k \quad ((x) = y) \quad -max_{j \neq Y} \quad av_k \quad I(h_k \quad (x)) \quad -j) \tag{14}$$

The larger the margin value, the higher the accuracy of the classification prediction, and the more confidence in the classification [80]. Parameter tuning was performed to develop the high accuracy model to classify LULC. The cost and gamma value were checked using the grid search method and select the best parameter.

3.2.7. Parameter Tuning and Feature Selection

In a ML application, parameter tuning is a crucial technique to optimize the classifier parameter and best-suited model fitting. Essential techniques for parameter tuning include k-fold (e.g., 10-fold) cross-validation with grid search [81]. K-fold cross-validation with grid search method is a robust and widely used method in SVM-RBF classification for parameter tuning. To avoid the overfitting issue in the SVM-RBF algorithm, the cost range (C) 0.1 to 1000 and gamma range (σ) 0.001 to 10 were considered for best parameter development [77]. Several classifier's parameter combinations were systematically tested with the training sample called the grid search approach [81]. As part of the cross-validation technique, 480 samples were randomly selected and divided into two portions as training and test samples. In this study, 65% (312) was used for training and 35% (168) for testing. The reserved 35% testing samples were used to evaluate the estimated trained models' performance. In this study, 312 training samples were split into different ratios to train and validate samples to select the best parameter. The parameter combination with the highest testing accuracy was considered the optimal classifier parameter in the classification process [40,81]. The cost (C) range was the only optimization in the SVM-linear technique for parameter adjustment [82]. Grid search and K-fold cross-validation were used to refine the SVM-linear algorithm's optimal parameter [83,84]. K-fold (for example, 10-fold) cross-validation approaches were used for parameter adjustment in the RF method. The nth no. of the tree's depth (e.g., range 5 to 50) is used in the RF method to maximize the highest parameter [40].

The feature selection approach is an essential procedure in image classification. This approach upgrades the classifier performance and reduces the complexity of datasets by removing redundant information [54,85]. Feature selection methods identify a minimum set of features by maintaining a class probability distribution as close as possible to the original distribution obtained using all features. Various feature selection methods were applied in LULC classification, such as filter methods like correlation, gain ratio, and relief, and another wrapper method like OOD-random forest. Each feature has a respective im-

portance value computed considering the metric used in the feature selection method [54]. Feature selection on lower feature dimensions may produce the cost of computation [54]. In this study, feature selection was averted to reduce the computation cost because of the few extracted features.

3.2.8. Ensemble and Transfer Learning

Ensemble learning is an essential technique in ML for land use and land cover mapping [86]. To avoid the biases of a single model in terms of performance (e.g., classification accuracy) of the estimated model, the application of an ensemble learning is crucial for land use and land cover classification [87]. It combines fine-tune ML algorithms to develop a robust predictive model [88]. In addition, ensemble models are proficient to classify the land use land cover to resolve the overfitting and misclassification issue [89]. In this study, four ensemble models were developed using majority voting algorithm to map LULC classification as follows:

$$\text{Model} - 1 = \text{f}(\text{SVM}_{\text{RBF}}, \text{SVM}_{\text{Linear}}) \quad (15)$$

$$\text{Model} - 2 = \text{f}(\text{SVM}_{\text{RBF}}, \text{Random Forest}) \quad (16)$$

$$\text{Model} - 3 = \text{f}(\text{SVM}_{\text{Linear}}, \text{Random Forest}) \quad (17)$$

$$\text{Model} - 4 = \text{f}(\text{SVM}_{\text{RBF}}, \text{SVM}_{\text{Linear}}, \text{Random Forest}) \quad (18)$$

where, f denotes majority voting function.

Transfer learning (TL) is commended for its connections to subsequent testing and its ability to produce findings quickly and accurately [90]. By reducing training time, memory requirements, and network design labour, TL is a common and very beneficial approach. TL was developed in two ways. Firstly, ML algorithms were fine-tuned on sample domain, e.g., KMA from where training samples were collected to achieve pre-trained model (e.g., domain on KMA). Secondly, the pre-trained model was employed on unsampled domain e.g., Delhi to classify LULC [91]. Building robustness of the model is facilitated by the developed model's applicability in another area. A pre-trained classification model was applied in another area for cost-effective analysis [40]. In this study, a fine-tuned model was developed in the Kolkata metropolitan area and applied on the Delhi metropolitan area to develop a robust spatial transferability model for land use and land cover classification.

3.2.9. Urban Land Use and Land Cover Classification

The land use land cover of the Kolkata and Delhi metropolitans were classified based on the selected land use and land cover classes (see Section 3.2.4), training and test samples (see Section 3.2.5), and best ML algorithm, transferable or ensemble model (see Sections 3.2.6–3.2.8). The best ML algorithm, transferable or ensemble model was only trained over Kolkata but used to predict (e.g., classified) both the Kolkata and Delhi metropolitan areas.

3.2.10. Accuracy Assessment

Accuracy assessment is an essential approach to assess the classification performance of ML algorithms [34]. Precision, recall, F-score, overall accuracy, and Kappa coefficient are robust and widely used indices for the accuracy assessment [92]. The precision (Equation (22)) value represents a positive prediction, which defines how many predicted settlement points are actual settlement points. It helps to define the reliability of the model. The recall (Equation (23)) value indicates the sensitivity or actual positive value, the number of true settlement points correctly predicted as settlement points. For an accurate classification, a perfect prediction result should have a perfect precision and recall value of one or 100%. The F-score (Equation (24)) is the harmonic mean of precision, which gives a combined idea about precision and recall. It is calculated as the weighted average of precision and recall [11,93]. The overall accuracy (Equation (19)) is the sum of several

correctly classified values divided by the total number of values. The kappa coefficient (Equation (20)) was calculated as an agreement between classification and truth values. A kappa value of 1 represents a perfect agreement, while 0 represents no agreement [94]:

$$\text{Overall acuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{19}$$

$$\text{Kappa} = \frac{\text{Observed agreement} - \text{chance agreement}}{1 - \text{chance agreement}} \tag{20}$$

$$\text{Chance agreement} = \left(\frac{((\text{TP} + \text{FN}) \times (\text{TP} + \text{FP}) + (\text{FP} + \text{TN}) \times (\text{FN} + \text{TN}))}{(\text{TP} + \text{TN} + \text{FP} + \text{FN})^2} \right) \tag{21}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TF} + \text{FP}} \tag{22}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{23}$$

$$\text{F-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{24}$$

where, TP = true positive, TN = true negative, FP = false positive and FN = false negative and observed agreement = overall accuracy.

4. Results and Discussion

Based on the objective of the study, several datasets and methods were applied to achieve the results as explained below.

4.1. Extracted Features

In this study, 21 features were extracted for LULC classification in both regions. These 21 features include six spectral bands, two spectral features, six textural features with GLCM mean and variance, respectively, and one nightlight band. The extracted features were stacked into a band stack image for parameter tuning, feature selection, ensemble and transfer learning, and LULC classification.

4.2. Training and Test Sample

In this study, a total of 480 samples were collected from the Kolkata metropolitan area, and collected training samples were split into a 65:35 ratio. Sixty-five percent of the collected samples were considered as training samples, and the 35 percent were test samples. A total of 218 test samples were collected from Delhi to assess the transferability of the model; a pre-trained model was applied in the Delhi metropolitan.

4.3. Best Parameters

Parameter tuning was performed to develop a robust ML algorithm, ensemble, and transfer learning model for LULC classification. The best parameter was developed based on overall accuracy. The best parameter is the one that predicts higher accuracy as compared to others. The cost and gamma values were tuned in various combinations (based on grid search) for the selection of best parameter of the SVM-RBF ML algorithm. The combination of cost(c) value 0.5 and gamma(g) value 1 produced a high accuracy (91.99%) as compared to c-0.3, g-7 (91.31%); c-0., g-9 (91.56%); c-0.5, g-8 (91.41%) and c-0.5, g-1 (Table 4). The best cost value of the SVM-linear algorithm was extracted based on the iterative testing of cost value. The cost 0.5 is considered as the best parameter because it provides higher accuracy (86.30%) compared to c-0.3 (77.17%). The n'th number of tree depths (nTree) was used for parameter tuning in the RF algorithm. In this study, the no. of tree depth 10 to 500 ranges was used for parameter tuning. The 10th number is considered as the best number of trees as it is achieved higher accuracy (96.57%) (Table 4).

Table 4. Best parameters of SVM-RBF, SVM-Linear and Random Forest.

Name of the Algorithm	Parameters		
	Cost	**Gamma**	**nTree**
SVM-Linear	0.5	-	-
SVM-RBF	0.5	1	-
Random Forest	-	-	10

4.4. Land Use Land Cover Classification

The LULC map of 2011 was obtained using ML algorithms like RF, SVM-Liner, and SVM-RBF. Six LULC classes were considered such as built-up, cropland, fallow land, vegetation, open land and water body to classify the land use land cover of the KMA.

4.4.1. LULC Classification Using Machine Learning Algorithms, KMA

The accuracy assessment and parameter tuning has been performed and developed a best suited model to classify the LULC. The cost value 1 and gamma 0.5 of the SVM model provide the highest accuracy as compared to other combinations of the parameters. The cost value 5 has been selected as the best suited parameter in the SVM linear algorithm to produce the highest accuracy (Table 4). Parameter tuning of the model is an important strategy to develop an accurate LULC classification map. It plays a significant role for the improvement of the classification accuracy and the quality of the map as well. The depth of the tree in the random forest model was altered to improve the accuracy level. The depth of tree 10 gives high accuracy in the RF model as compared to other depths of trees (Table 4). The tree depth has been mentioned in a table with their accuracy level to obtain highly influential model parameters in LULC classification (Tables 4 and 5).

Table 5. Accuracy assessment of SVM-RBF, SVM-Linear, and RF.

ML Algorithm	Overall Accuracy	Kappa	Precision	Recall	F-Score
SVM-RBF	85.07	0.81	0.86	0.76	0.78
SVM-Linear	91.99	0.89	0.89	0.92	0.91
RF	94.43	0.93	0.89	0.95	0.91

The accuracy indices of the LULC map have been mentioned in Table 5. The models' overall accuracy, Kappa, precision, recall, and F-score have been computed. The table demonstrates that RF provides a greater accuracy (94.43%) than the other two independent algorithms, and RF's Kappa value is 0.93, which denotes a strong classification agreement with ground reality. Table 5 shows that the RF model's precision value is 0.89, which depicts that the model predicts all classes accurately. In addition, the recall value is 0.95 and shows that the model predicts LULC classes correctly as similar to the ground reality. At the same time, the F1-score also supports that the model is performed very well. Table 5 shows that the SVM-Linear model outperforms the SVM-RBF because simple ML algorithms like SVM-linear and RF can solve such types of classification problems.

The area level indices of different classes have been calculated in Kolkata metropolitan area (Table 6). The LULC classes area has been compared with the reference LULC map [95] (Table 6). Table 6 and Figure 3 shows that the area of most of the LULC classes (e.g., fallow land, vegetation, waterbody) are similar (more than 80%) to the LULC area produced by Random Forest as compared to SVM-linear and SVM-RBF. Next to RF, SVM-linear shows a very close similarity of LULC area to reference LULC.

Table 6. Land use and land cover classification, KMA using machine learning algorithms.

LULC Code	LULC Classes	SVM-Linear		SVM-RBF		Random Forest		* Reference LULC
		Area (ha)	Area (%)	Area (ha)	Area (%)	Area (ha)	Area (%)	Area (ha)
1	Built-up	647.60	35.57	642.18	35.27	519.34	28.53	713.67 (42.23%)
2	Cropland	510.87	28.06	640.99	35.21	625.14	34.34	317.30 (18.78%)
3	Fallow Land	197.18	10.83	157.89	8.67	195.66	10.75	151.86 (8.99%)
4	Vegetation	285.03	15.66	256.78	14.10	343.17	18.85	406.36 (24.05%)
5	Open Land	31.24	1.72	35.21	1.93	18.49	1.02	-
6	Waterbody	148.69	8.17	87.57	4.81	118.81	6.53	100.52 (5.95%)
Total		1820.61	100	1820.61	100	1820.61	100	1689.71

* (Ghosh et al., 2019) [95].

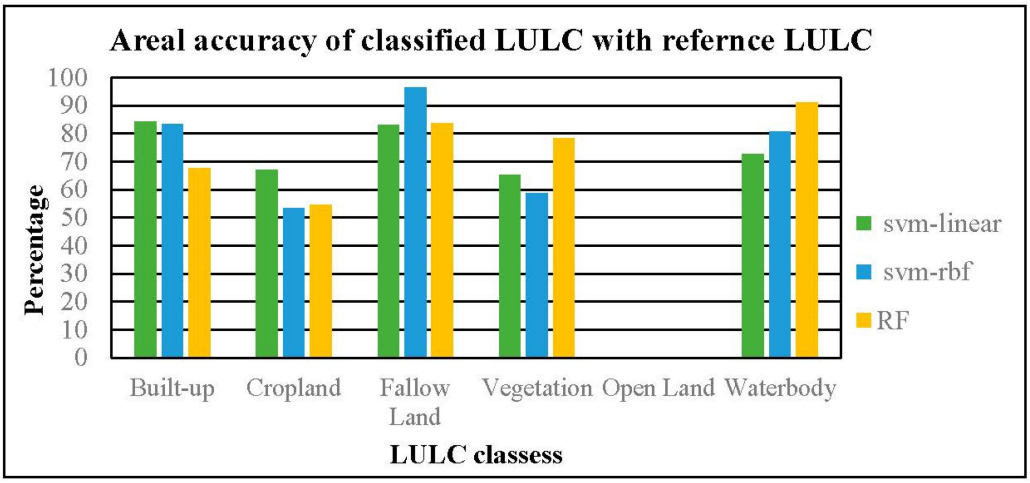


Figure 3. Areal accuracy of classified LULC with reference LULC, KMA.

The similarity of the area between the predicted class and reference class of waterbody, vegetation, built-up and crop land is better in the SVM-Linear model than the SVM-RBF. Open land class is not available in the reference. Consequently, open land class is not plotted in Figure 3. The SVM-Linear gives a higher overall higher accuracy than the SVM-RBF. Random Forest has performed very well in water body and vegetation extraction. The waterbody and vegetation classes of RF classification are close to the referenced classes, around 91% and 78%, respectively.

Overall, the RF model gives a higher accuracy (95%) than the SVM-linear and SVM-RBF models. Results have been plotted in a bar graph (Figure 3). The graph shows that the areal statistics of RF are closer with most of the LULC classes than the other two algorithms, giving an overall high accuracy in relation to the ground reality (Figure 4).

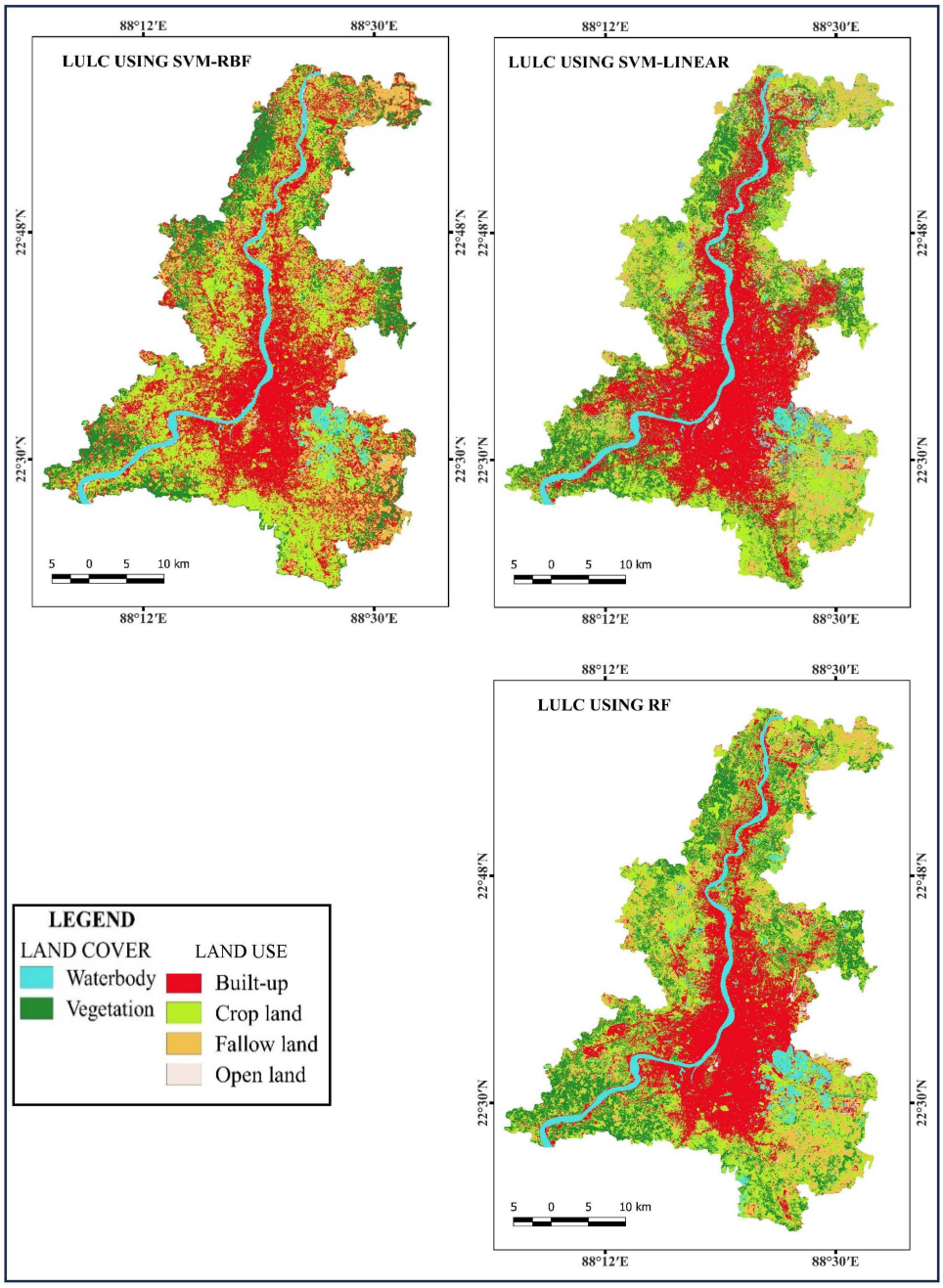


Figure 4. Land use land cover mapping using machine learning algorithms.

4.4.2. LULC Classification Using Ensemble Models, KMA

The ensemble of the ML algorithms is an important procedure to achieve a high accuracy LULC map. It reduces the biases in classification and gives an accurate information. The majority voting technique in QGIS software was used to fuse the algorithms. Four ensemble models were developed, showing varying accuracy in LULC mapping (Table 7, Figure 5).

Table 7. Accuracy assessment of ensemble models.

Ensemble Model	Overall Accuracy	Kappa	Precision	Recall	F-Score
Model-1	86.02	0.82	0.91	0.78	0.80
Model-2	85.88	0.82	0.91	0.78	0.80

Table 7. *Cont.*

Ensemble Model	Overall Accuracy	Kappa	Precision	Recall	F-Score
Model-3	93.35	0.91	0.94	0.93	0.93
Model-4	94.84	0.93	0.95	0.95	0.95

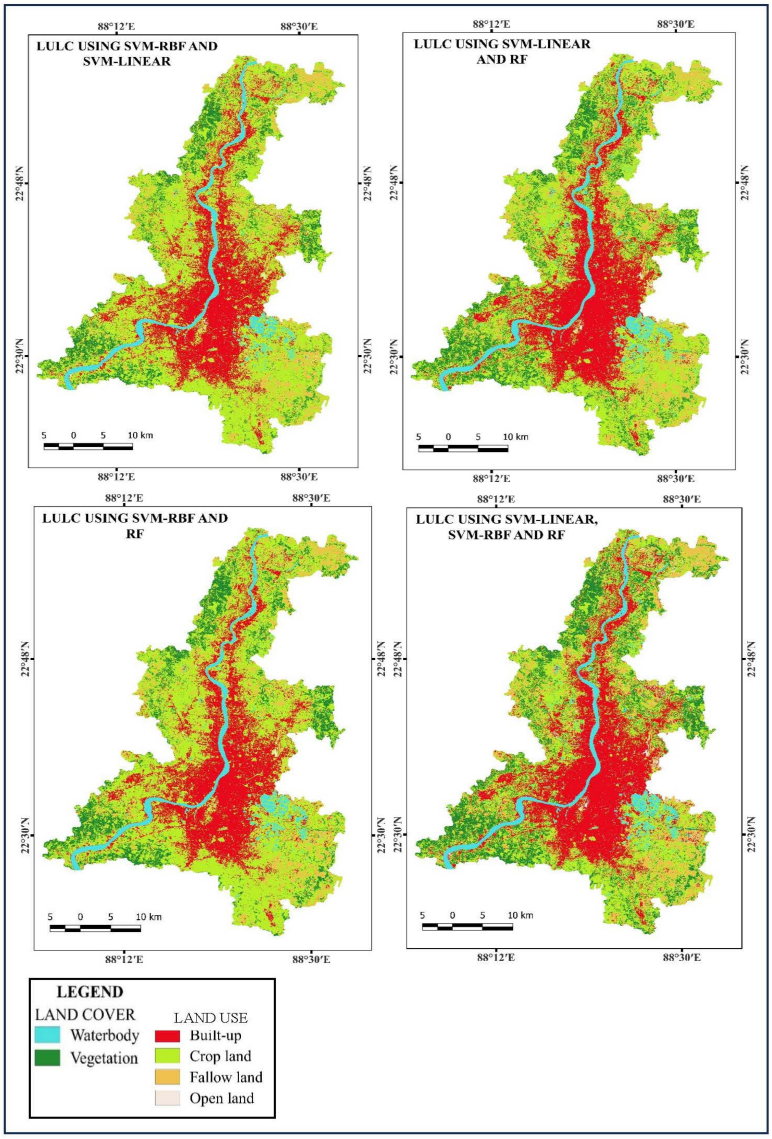


Figure 5. Land use land cover classification using ensemble models.

Table 7 shows that model-4 outperforms the other ensemble models in terms of Overall accuracy (94.84), Kappa (0.93), Precision (0.95), Recall (0.95), and F-score (0.95) because of the complex classification problems. The table shows that simple ensemble models are more affected by the complex classification problem compared to the higher order ensemble model.

The area statistics of the ensemble model have been computed (Table 8). The computed area of different classes generated from different ensemble models has been compared with the reference LULC classes of the Kolkata metropolitan area. The computed area and reference LULC area closeness have been plotted in Figure 6. The figure shows that the area of the LULC classes computed by the ensemble model-4 area is very similar (more than 75%) to the area of reference LULC classes. The similarity of the area between modelled LULC and reference LULC shows a good performance of the model.

Table 8. Land use and land cover classification, KMA using ensemble model.

LULC Code	LULC Classes	Model-1		Model-1		Model-1		Model-1		Reference LULC
		Area (sq. km)	Area (%)	Area (sq. km)	Area (%)	Area (sq. km)	Area (%)	Area (sq. km)	Area (%)	Area (sq. km)
1	Built-up	457.61	25.14	444.21	24.40	482.11	26.48	539.03	29.61	713.67 (42.23%)
2	Cropland	997.77	54.80	979.63	53.81	851.61	46.78	709.28	38.96	317.30 (18.78%)
3	Fallow Land	107.34	5.90	110.03	6.04	129.92	7.14	158.32	8.70	151.86 (8.99%)
4	Vegetation	157.84	8.67	189.54	10.41	242.13	13.30	284.96	15.65	406.36 (24.05%)
5	Open Land	17.21	0.95	10.32	0.57	13.96	0.77	23.60	1.30	-
6	Waterbody	82.85	4.55	86.89	4.77	100.88	5.54	105.43	5.79	100.52 (5.95%)
Total		1820.61	100	1820.61	100	1820.61	100	1820.61	100	1689.71

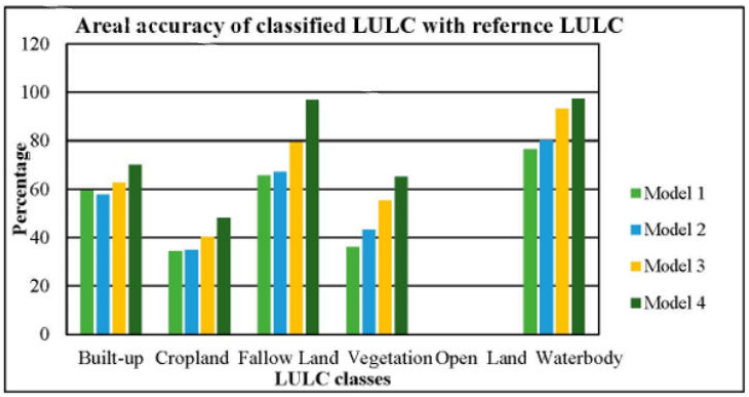


Figure 6. Areal accuracy of classified LULC with reference LULC, KMA.

4.5. Transferability Assessment

Transferability is a crucial term for developing the robustness of the model. It has been defined as the model is trained over one segment and tested over another region. It is typically used to generalize the model. When the model was trained in an area and tested over another area, it is called spatial transferability [96]. In this study, the model was trained in the Kolkata metropolitan area and tested in the Delhi region to assess the spatial transferability of ML models.

4.5.1. LULC Classification Using Pre-Trained ML Algorithms, Delhi

The pre-trained model of the ML algorithm was applied in the Delhi region for the transferability assessment. The trained algorithm having the highest accuracy has been considered for the transferability analysis in LULC classification. Similar image bands and features of KMA were developed for Delhi and used for LULC classification using pre-trained ML algorithms. Pre-trained ML algorithms SVM-RBF, SVM-linear, and RF were developed in KMA. The performance of the pre-trained ML algorithms was assessed based on the accuracy assessment indices and test samples collected from Delhi (Table 3). Table 9 shows that SVM-linear and RF (0.79) are more transferable as compared to SVM-RBF (0.71) based on the overall accuracy and other indices. The precision and F-score are higher in the SVM-linear (0.77, 0.75) than in the RF (0.76, 0.74) (Table 9). The performance of SVM-Linear outperforms the RF in terms of Precision (0.77) and F-score (0.75), which explains that SVM-Linear is more transferable in LULC classification as compared to SVM-RBF and RF.

Table 9. Accuracy assessment of pre-trained ML algorithms for LULC classification.

ML Algorithm	Overall Accuracy	Kappa	Precision	Recall	F-Score
SVM-RBF	71	0.64	0.70	0.64	0.65
SVM-Linear	79	0.74	0.77	0.75	0.75
RF	79	0.74	0.76	0.75	0.74

The area statistics of LULC classes produced by the ML algorithms have also been compared with the referenced LULC classes (Table 10) to assess the transferability of the ML algorithms. Figure 7 shows that the area of LULC classes of SVM-Linear is very similar to the reference LULC classes area followed by RF and SVM-RBF.

Table 10. Land use and land cover classification, KMA using pre-trained ML algorithms.

LULC Code	LULC Classes	SVM-Linear		SVM-RBF		Random Forest		* Reference LULC
		Area (sq. km)	Area (%)	Area (sq. km)	Area (%)	Area (sq. km)	Area (%)	Area (sq. km)
1	Built-up	570.28	38.55	594.15	40.16	497.59	33.63	623.96 (42.03%)
2	Cropland	436.42	29.50	489.98	33.12	458.00	30.96	335.73 (22.61%)
3	Fallow Land	164.52	11.12	117.89	7.97	166.70	11.27	214.89 (14.47%)
4	Vegetation	217.88	14.73	241.52	16.32	307.21	20.76	122.73 (8.27%)
5	Open Land	36.10	2.44	24.72	1.67	17.84	1.21	172.69 (11.63%)
6	Waterbody	54.29	3.67	11.23	0.76	32.15	2.17	14.67 (0.99%)
Total		1479.50	100	1479.50	100	1479.50	100	1689.71

* (Shahfahad et al., 2022) [97].

The results (e.g., accuracy indices, areal similarity, quality of the map, etc.) show that SVM-linear is more transferable, followed by RF and SVM-RBF to the ground reality (Table 10, Figure 8).

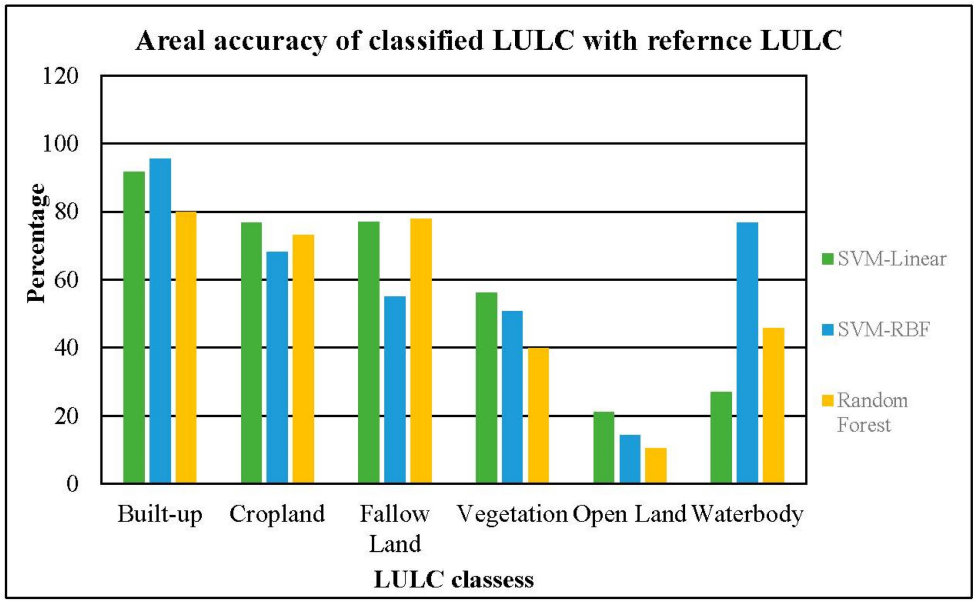


Figure 7. Areal accuracy of classified LULC with reference LULC, Delhi.

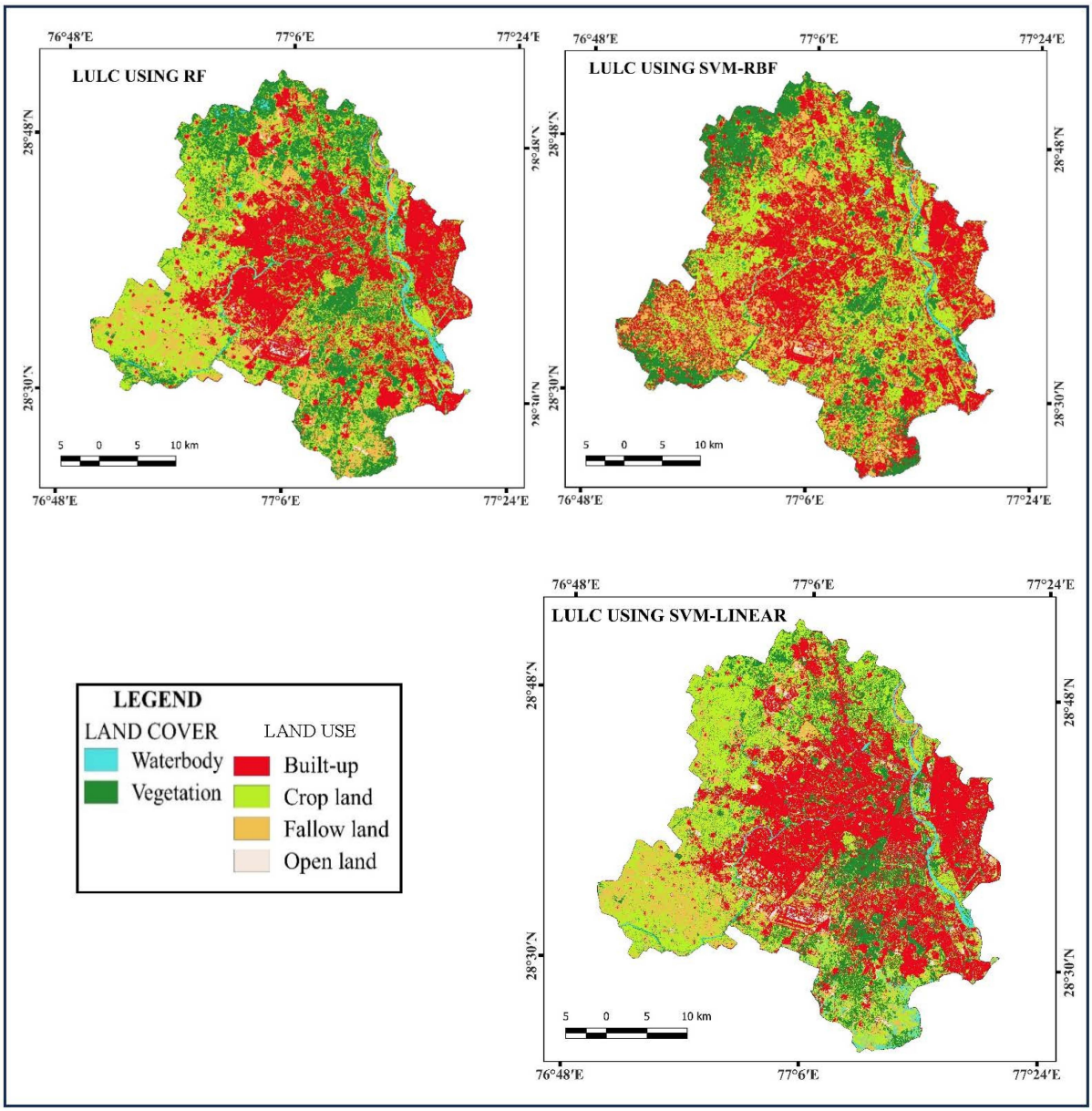


Figure 8. Land use land cover classification using pre-trained ML algorithms.

4.5.2. LULC Classification Using Pre-Trained Ensemble Models, Delhi

The pre-trained ensemble models have been applied in the Delhi region to classify the land use land cover. Table 11 shows that model-4 shows the best performance and robustness compared to other pre-trained ensemble models for LULC classification (Figures 9 and 10).

Table 11. Accuracy assessment of pre-trained ensemble models for LULC classification.

Ensemble Algorithm	Overall Accuracy	Kappa	Precision	Recall	F-Score
Model-1	71.90	0.65	0.74	0.63	0.63
Model-2	73.96	0.66	0.76	0.65	0.66
Model-3	76.85	0.71	0.76	0.71	0.71
Model-4	80.75	0.76	0.78	0.76	0.76

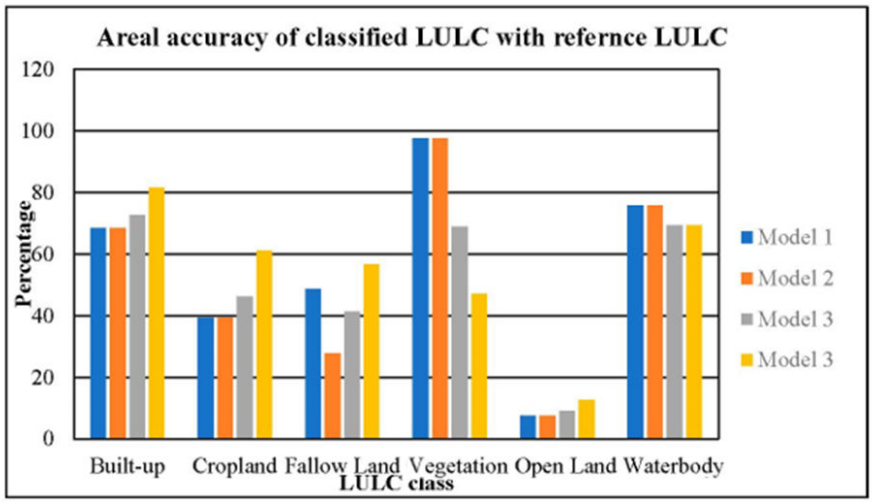


Figure 9. Areal accuracy of classified LULC with reference LULC, Delhi.

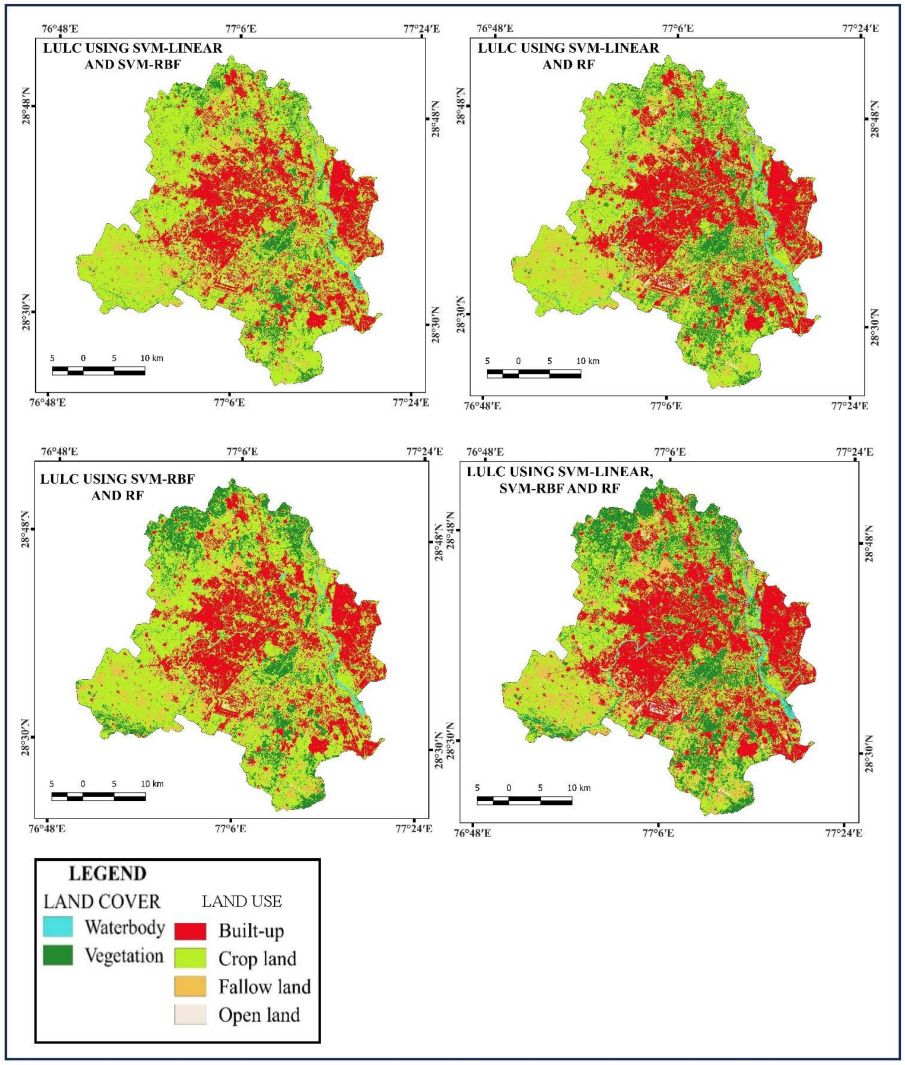


Figure 10. Land use land cover classification using pre-trained ensemble models.

Model-1 shows the lowest accuracy (71.90%), compared to other ensemble models. Model-3 produces a higher accuracy (76.85%) than the ensemble model-2 (73.96%), and

ensemble model-1 (71.90%) (Figure 9). The Kappa coefficient value of ensemble model-4 is 0.76, which shows a strong agreement with the ground reality of LULC classes. The precision value of model-4 is 0.78, which define that the model's predictive performance is 78 percent average of all classes, and the recall value is 0.76, which supports that the corrected classes of the model with ground reality are 76 percent. The F score value of ensemble model-4 is 0.76, which means that the performance of the model is good as compared to other pre-trained ensemble models in terms of spatial transferability.

The area statistics of pre-trained ensemble models have been highlighted in Table 12. The comparison of correlation of different models with reference LULC classes has been plotted in Figure 9, which shows that the computed area of LULC classes of model-4 model is highly matching with the reference LULC classes. The area of built-up, cropland, fallow land, open land mapped by the ensemble model-4 is strongly matching with reference LULC classes area as compared to the other pre-trained ensemble models (Figure 10).

Table 12. Land use and land cover classification, KMA using pre-trained ensemble models.

LULC Code	LULC Classes	Model-1		Model-2		Model-3		Model-4		Reference LULC
		Area (sq. km)	Area (%)	Area (sq. km)	Area (%)	Area (sq. km)	Area (%)	Area (sq. km)	Area (%)	Area (sq. km)
1	Built-up	426.16	28.80	426.16	28.80	452.38	30.58	507.83	34.32	623.96 (42.03%)
2	Cropland	850.63	57.49	850.63	57.49	724.27	48.95	547.64	37.02	335.73 (22.61%)
3	Fallow Land	59.43	4.02	59.43	4.02	88.40	5.97	121.00	8.18	214.89 (14.47%)
4	Vegetation	119.40	8.07	119.40	8.07	177.65	12.01	260.07	17.58	122.73 (8.27%)
5	Open Land	12.79	0.86	12.79	0.86	15.72	1.06	21.73	1.47	172.69 (11.63%)
6	Waterbody	11.09	0.75	11.09	0.75	21.08	1.43	21.23	1.43	14.67 (0.99%)
Total		1479.50	100	1479.50	100	1479.50	100	1479.50	100	1689.71

5. Discussion

In this study, LULC classification was performed using ML algorithms (SVM-RBF, SVM-linear and RF). Random Forest gave the highest accuracy followed by SVM-linear and SVM-RBF algorithm. Random forests classified the LULC classes and especially differentiated built-up areas from the other classes correctly. The random forest algorithm is a collection of many decision trees. Thus the algorithm can easily control complex patterns and data, while the linear model's performance is unsatisfactory comparatively [98]. Previous studies showed that ensemble models were the main contributor to improve the classification accuracy. For example, ensemble models may be combinations of many ML algorithms [89,93,99]. In this study, the combinations of three ML algorithms show 94.84% overall classification accuracy on sampled domain and 80% on the unsampled domain and mean classification accuracy in both the domains is 87.42%. The accuracy may be improved in the unsampled domain by incorporating others ML algorithms, which needs to be tested in future research. In this regard, model 2 ensemble algorithm LULC classification gives the lowest accuracy comparatively. Model 3 gave a higher accuracy than the previous one. Model 1 gave 88 percent accuracy, higher than the above two. The RF algorithm is a good performer individually, but the accuracy level was lower when the RF algorithm classification was ensembled with the SVM algorithm. The SVM-RBF and SVM-linear classification were ensembled then it performed better as compared to single application. However, the model 3 (SVM-RBF, SVM-linear and RF) ensembled LULC classification has a high accuracy compared to the other ensembled and individual algorithms classification. The Ensemble of ML algorithm increases the accuracy level 5 percent higher than the individual and fusion of only two algorithms for LULC classifications.

Transferable learning of the ML algorithm was applied to analyse the robustness of the model in LULC classification. Generally, the transferability of the models depends on the heterogeneity [34]. The ML model has been applied in Kolkata metropolitan area as a pretrained model and this pretrained model was applied in Delhi for the transferability analysis. The transferability analysis in different regions plays a significant role for a

robustness and cost effective model development [9]. In this study, model 4 ensembles are more accurate than the other ensemble models in transfer analysis. The fusion of three models has produced high accuracy in LULC classification. In transferability analysis, model 3 gave a higher accuracy in individual model performance than the SVM-RBF algorithm. The algorithm's performance in accuracy level analysis is the same order as the pre-trained mod but the overall accuracy level is low as compared to the pretrained model classification accuracy. In a previous study, the accuracy level was reduced in transferable model application [11]. The single algorithm classification gives lower accuracy as compared to the ensemble model. The single model has some limitations because single models cannot classify or differentiate similar objects from satellite images. The ensemble model reduces the limitations of individual algorithms and ensemble models produce higher accuracy. In this study, the ensemble model of three algorithms has produced the highest accuracy as compared to other singles as well as the ensemble of two algorithms [34, 40]. Thus, the ensemble models outperform in LULC classification than the single algorithm classification [89].

6. Conclusions

Land use and land cover (LULC) classification has been performed using ML algorithms (SVM-RBF, SVM-linear, and RF). The accuracy assessment employed robust accuracy assessment indices to assess the ML algorithms and developed models for LULC classification. The performance of ML algorithms on the sample domain (e.g., KMA) shows that RF gives higher accuracy and is competitive with SVM-linear and SVM-RBF. RF classifies the LULC classes, especially differentiating the built-up area from the other classes very correctly. The decision tree-based algorithm performs best in built-up classification compared to other classes. An RF is an ensemble algorithm itself with the collection of many decision trees. Thus, the algorithm can easily control complex patterns and data while the linear model's performance is unsatisfactory. In addition, a single ML algorithm is sometimes unable to perform very well for a complex LULC classification problem. In this regard, four ensemble models were developed and tested its performance on both sampled and unsampled domain. Regarding an ML algorithm that performs better on the sampled domain (e.g., from where the training sample collected, and the algorithm was trained), there is no guarantee that such an algorithm will achieve the same performance on the unsampled domain (a pre-trained model was used to classify LULC) like Delhi. The algorithm or model that performs better on both (sampled and unsampled domain) can be considered as a robust algorithm or model. The result shows that the pre-trained RF algorithm, which is competitive to SVM-Linear, shows higher accuracy as compared to SVM-RBF, and similarly higher order pre-trained ensemble models (e.g., model-4) provide better accuracy as compared to other pre-trained ensemble models for the LULC classification on an unsampled domain. Therefore, this study shows that the RF, which is competitive to SVM-linear and ensemble model-4 (combination of SVM-linear, SVM-RBF, and Random Forest), is robust and is expected to perform very well for the LULC classification of the other urban areas. The proposed study shows that ensemble and transfer learning models outperform the classical ML algorithms at complex urban land use and land cover pattern (based on the experience from the existing and past studies). The study provides reasonable land use and land cover accuracy with the experiment on two cities (e.g., Kolkata and Delhi). The limited experiment might partially be affected by the biased classification. The biased classification problem leads to questioning in sustainable urban land use development and planning. To achieve robust and optimum land use and land cover classification accuracy, deep learning algorithms (e.g., convolutional Neural Network (CNN)) could be employed and investigated over more cities. This study will be a promising guideline for the urban remote sensing scientist, land use planner and decision-maker and landscape ecologist to adopt the proposed state-of-the-art ML technology and EO dataset for the planning and management of the urban and landscape ecology.

Author Contributions: P.B. contributed to conceptualizing, data processing, land use mapping and drafting of the article. S.M. contributed to data mining, validation of land use maps and editing the article. M.K. contributed to developing the methodology and editing, reviewing, and conceptualising the article. S.K.S. contributed to developing methodology, discussion and editing of the article. All authors have read and agreed to the published version of the manuscript.

Funding: This research paper is a part of PhD thesis which is funded by the University Grants Commission (UGC), New Delhi, India (NTA Ref. No.190520221203), Dated: January 2020).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All the data available with the authors and such data will be submitted and shared to the publisher on demand based on freely available repositories, e.g., FigShare.

Acknowledgments: We are very grateful to the editor and anonymous reviewers for their valuable suggestions and helpful comments.

Conflicts of Interest: There is no conflict of interest among authors.

References

1. Li, Z.; Liu, S.; Tan, Z.; Sohl, T.L.; Wu, Y. Simulating the effects of management practices on cropland soil organic carbon changes in the Temperate Prairies Ecoregion of the United States from 1980 to 2012. *Ecol. Model.* **2017**, *365*, 68–79. [CrossRef]
2. Schneider, A. Monitoring land cover change in urban and peri-urban areas using dense time stacks of Landsat satellite data and a data mining approach. *Remote Sens. Environ.* **2012**, *124*, 689–704. [CrossRef]
3. Bernstein, M.J.; Franssen, T.; Smith, R.D.J.; de Wilde, M. The European Commission's Green Deal is an opportunity to rethink harmful practices of research and innovation policy. *Ambio* **2023**, *52*, 508–517. [CrossRef] [PubMed]
4. Henderson, V. The Urbanization Process and Economic Growth: The So-What Question. *J. Econ. Growth* **2003**, *8*, 47–71. [CrossRef]
5. Census of India 2011. Census Tables | Government of India. Ministry of Home Affairs. Available online: https://censusindia.gov.in/census.website/data/census-tables (accessed on 11 September 2023).
6. Xu, L.; Liu, X.; Tong, D.; Liu, Z.; Yin, L.; Zheng, W. Forecasting Urban Land Use Change Based on Cellular Automata and the PLUS Model. *Land* **2022**, *11*, 652. [CrossRef]
7. Xu, H. Extraction of Urban Built-up Land Features from Landsat Imagery Using a Thematicoriented Index Combination Technique. *Photogramm. Eng. Remote Sens.* **2007**, *73*, 1381–1391. [CrossRef]
8. Schulz, D.; Yin, H.; Tischbein, B.; Verleysdonk, S.; Adamou, R.; Kumar, N. Land use mapping using Sentinel-1 and Sentinel-2 time series in a heterogeneous landscape in Niger, Sahel. *ISPRS J. Photogramm. Remote Sens.* **2021**, *178*, 97–111. [CrossRef]
9. Rosier, J.F.; Taubenböck, H.; Verburg, P.H.; van Vliet, J. Fusing Earth observation and socioeconomic data to increase the transferability of large-scale urban land use classification. *Remote Sens. Environ.* **2022**, *278*, 113076. [CrossRef]
10. Yin, L.; Wang, L.; Li, J.; Lu, S.; Tian, J.; Yin, Z.; Liu, S.; Zheng, W. YOLOV4_CSPBi: Enhanced Land Target Detection Model. *Land* **2023**, *12*, 1813. [CrossRef]
11. Rudiastuti, A.W.; Lumban-Gaol, Y.; Silalahi, F.E.S.; Prihanto, Y.; Pranowo, W.S. Implementing Random Forest Algorithm in GEE: Separation and Transferability on Built-Up Area in Central Java, Indonesia. *JOIV Int. J. Informatics Vis.* **2022**, *6*, 74–82. [CrossRef]
12. Banzhaf, E.; Kabisch, S.; Knapp, S.; Rink, D.; Wolff, M.; Kindler, A. Integrated research on land-use changes in the face of urban transformations—An analytic framework for further studies. *Land Use Policy* **2017**, *60*, 403–407. [CrossRef]
13. Patino, J.E.; Duque, J.C. A review of regional science applications of satellite remote sensing in urban settings. *Comput. Environ. Urban Syst.* **2013**, *37*, 1–17. [CrossRef]
14. Riggan, N.D., Jr.; Weih, R.C., Jr. A Comparison of Pixel-Based versus Object-Based Land Use/Land Cover Classification Methodologies. *J. Ark. Acad. Sci.* **2019**, *63*, 145–152.
15. Kete, S.C.R.; Suprihatin; Tarigan, S.D.; Effendi, H. Land use classification based on object and pixel using Landsat 8 OLI in Kendari City, Southeast Sulawesi Province, Indonesia. *IOP Conf. Ser. Earth Environ. Sci.* **2019**, *284*, 012019. [CrossRef]
16. Das, S.; Angadi, D.P. Land use land cover change detection and monitoring of urban growth using remote sensing and GIS techniques: A micro-level study. *GeoJournal* **2022**, *87*, 2101–2123. [CrossRef]
17. Norman, M.; Shahar, H.M.; Mohamad, Z.; Rahim, A.; Mohd, F.A.; Shafri, H.Z.M. Urban building detection using object-based image analysis (OBIA) and machine learning (ML) algorithms. *IOP Conf. Ser. Earth Environ. Sci.* **2021**, *620*, 012010. [CrossRef]
18. Simanjuntak, R.M.; Kuffer, M.; Reckien, D. Object-based image analysis to map local climate zones: The case of Bandung, Indonesia. *Appl. Geogr.* **2019**, *106*, 108–121. [CrossRef]
19. Feizizadeh, B.; Omarzadeh, D.; Garajeh, M.K.; Lakes, T.; Blaschke, T. Machine learning data-driven approaches for land use/cover mapping and trend analysis using Google Earth Engine. *J. Environ. Plan. Manag.* **2021**, *66*, 665–697. [CrossRef]
20. Mishra, V.K.; Swarnkar, D.; Pant, T. A Modified Neural Network for Land use Land Cover Mapping of Landsat-8 Oli Data. In Proceedings of the 2021 IEEE International India Geoscience and Remote Sensing Symposium (InGARSS), Ahmedabad, India, 6–10 December 2021; pp. 65–69.

21. Talukdar, S.; Singha, P.; Mahato, S.; Shahfahad; Pal, S.; Liou, Y.-A.; Rahman, A. Land-Use Land-Cover Classification by Machine Learning Classifiers for Satellite Observations—A Review. *Remote Sens.* **2020**, *12*, 1135. [CrossRef]
22. Treitz, P.M.; Howarth, P.J.; Gong, P. Application of satellite and GIS technologies for land-cover and land-use mapping at the rural-urban fringe: A case study. *Photogramm. Eng. Remote Sens.* **1992**, *58*, 439–448.
23. He, D.; Shi, Q.; Liu, X.; Zhong, Y.; Zhang, X. Deep Subpixel Mapping Based on Semantic Information Modulated Network for Urban Land Use Mapping. *IEEE Trans. Geosci. Remote Sens.* **2021**, *59*, 10628–10646. [CrossRef]
24. Farda, N.M. Multi-temporal Land Use Mapping of Coastal Wetlands Area using Machine Learning in Google Earth Engine. *IOP Conf. Ser. Earth Environ. Sci.* **2017**, *98*, 012042. [CrossRef]
25. Nasiri, V.; Deljouei, A.; Moradi, F.; Sadeghi, S.M.M.; Borz, S.A. Land Use and Land Cover Mapping Using Sentinel-2, Landsat-8 Satellite Images, and Google Earth Engine: A Comparison of Two Composition Methods. *Remote Sens.* **2022**, *14*, 1977. [CrossRef]
26. Hussain, M.A.; Chen, Z.; Zheng, Y.; Shoaib, M.; Shah, S.U.; Ali, N.; Afzal, Z. Landslide Susceptibility Mapping Using Machine Learning Algorithm Validated by Persistent Scatterer In-SAR Technique. *Sensors* **2022**, *22*, 3119. [CrossRef] [PubMed]
27. Youssef, A.M.; Pourghasemi, H.R. Landslide susceptibility mapping using machine learning algorithms and comparison of their performance at Abha Basin, Asir Region, Saudi Arabia. *Geosci. Front.* **2021**, *12*, 639–655. [CrossRef]
28. Shahabi, H.; Jarihani, B.; Chittleborough, D.; Piralilo, S.T. Gully Networks Detection by Integration of Machine Learning and Geographic Object-Based Image Analysis. In Proceedings of the 8th International Symposium on Gully Erosion, Townsville, Australia, 21–27 July 2019.
29. Pourghasemi, H.R.; Sadhasivam, N.; Yousefi, S.; Tavangar, S.; Nazarlou, H.G.; Santosh, M. Using machine learning algorithms to map the groundwater recharge potential zones. *J. Environ. Manag.* **2020**, *265*, 110525. [CrossRef] [PubMed]
30. Sarkar, S.K.; Talukdar, S.; Rahman, A.; Shahfahad; Roy, S.K. Groundwater potentiality mapping using ensemble machine learning algorithms for sustainable groundwater management. *Front. Eng. Built Environ.* **2021**, *2*, 43–54. [CrossRef]
31. Mohammad, P.; Goswami, A.; Chauhan, S.; Nayak, S. Machine learning algorithm based prediction of land use land cover and land surface temperature changes to characterize the surface urban heat island phenomena over Ahmedabad city, India. *Urban Clim.* **2022**, *42*, 101116. [CrossRef]
32. Xia, N.; Cheng, L.; Li, M. Mapping Urban Areas Using a Combination of Remote Sensing and Geolocation Data. *Remote Sens.* **2019**, *11*, 1470. [CrossRef]
33. Malarvizhi, K.; Kumar, S.V.; Porchelvan, P. Use of High Resolution Google Earth Satellite Imagery in Landuse Map Preparation for Urban Related Applications. *Procedia Technol.* **2016**, *24*, 1835–1842. [CrossRef]
34. Chen, J.; Zhu, X.; Imura, H.; Chen, X. Consistency of accuracy assessment indices for soft classification: Simulation analysis. *ISPRS J. Photogramm. Remote Sens.* **2010**, *65*, 156–164. [CrossRef]
35. 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]
36. Li, C.; Wang, J.; Wang, L.; Hu, L.; Gong, P. Comparison of Classification Algorithms and Training Sample Sizes in Urban Land Classification with Landsat Thematic Mapper Imagery. *Remote Sens.* **2014**, *6*, 964–983. [CrossRef]
37. Prasad, P.; Loveson, V.J.; Chandra, P.; Kotha, M. Evaluation and comparison of the earth observing sensors in land cover/land use studies using machine learning algorithms. *Ecol. Inform.* **2022**, *68*, 101522. [CrossRef]
38. Basheer, S.; Wang, X.; Farooque, A.A.; Nawaz, R.A.; Liu, K.; Adekanmbi, T.; Liu, S. Comparison of Land Use Land Cover Classifiers Using Different Satellite Imagery and Machine Learning Techniques. *Remote Sens.* **2022**, *14*, 4978. [CrossRef]
39. Alem, A.; Kumar, S. Transfer Learning Models for Land Cover and Land Use Classification in Remote Sensing Image. *Appl. Artif. Intell.* **2022**, *36*, 2014192. [CrossRef]
40. Praveen, B.; Mustak, S.; Sharma, P. Assessing the transferability of machine learning algorithms using cloud computing and earth observation datasets for agricultural land use/cover mapping. *ISPRS Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2019**, *42*, 585–592. [CrossRef]
41. Naushad, R.; Kaur, T.; Ghaderpour, E. Deep Transfer Learning for Land Use and Land Cover Classification: A Comparative Study. *Sensors* **2021**, *21*, 8083. [CrossRef]
42. Kolkata Metropolitan Development Authority. *Vision 2025 Perspective Plan of CMA: 2025*; Kolkata Metropolitan Development Authority: Kolkata, India, 2005.
43. KMDA (Kolkata Metropolitan Development Authority). 2017. Available online: http://www.kmdaonline.org/home/ws_sector_info (accessed on 11 September 2023).
44. Mondal, B.; Samanta, G. Commuting and Metropolitan Development of Kolkata. *Hill Geogr.* **2018**, *33*, 61–79.
45. Jain, M.; Dawa, D.; Mehta, R.; Dimri, A.P.; Pandit, M.K. Monitoring land use change and its drivers in Delhi, India using multi-temporal satellite data. *Model. Earth Syst. Environ.* **2016**, *2*, 1–14. [CrossRef]
46. Pattanayak, S.P.; Diwakar, S.K. District-wise change analysis of land use-land cover in Delhi territory using remote sensing & GIS. *J. Urban Environ. Eng.* **2016**, *10*, 201–213. [CrossRef]
47. Li, X.; Zhou, Y.; Zhao, M.; Zhao, X. A harmonized global nighttime light dataset 1992–2018. *Sci. Data* **2020**, *7*, 1–9. [CrossRef] [PubMed]

48. DMSP Data Collected by US Air Force Weather Agency. Earth Observation Group-Defense Meteorological Satellite Progam, Boulder. Image and Data Processing by NOAA's National Geophysical Data Center. 2018. Available online: https://ngdc.noaa.gov/eog/viirs/download_boat.html (accessed on 11 September 2023).
49. Goldblatt, R.; Stuhlmacher, M.F.; Tellman, B.; Clinton, N.; Hanson, G.; Georgescu, M.; Wang, C.; Serrano-Candela, F.; Khandelwal, A.K.; Cheng, W.-H.; et al. Using Landsat and nighttime lights for supervised pixel-based image classification of urban land cover. *Remote Sens. Environ.* **2018**, *205*, 253–275. [CrossRef]
50. Al-jabery, K.K.; Obafemi-Ajayi, T.; Olbricht, G.R.; Wunsch, D.C., II. Data preprocessing. In *Computational Learning Approaches to Data Analytics in Biomedical Applications*; Academic Press: Cambridge, MA, USA, 2020; pp. 7–27. [CrossRef]
51. Özkan, C.; Erbek, F.S. Comparing feature extraction techniques for urban land-use classification. *Int. J. Remote Sens.* **2006**, *26*, 747–757. [CrossRef]
52. Gong, P.; Marceau, D.J.; Howarth, P.J. A comparison of spatial feature extraction algorithms for land-use classification with SPOT HRV data. *Remote Sens. Environ.* **1992**, *40*, 137–151. [CrossRef]
53. Zadeh, F.A.; Ardalani, M.V.; Salehi, A.R.; Farahani, R.J.; Hashemi, M.; Mohammed, A.H. An Analysis of New Feature Extraction Methods Based on Machine Learning Methods for Classification Radiological Images. *Comput. Intell. Neurosci.* **2022**, *2022*, 3035426. [CrossRef] [PubMed]
54. Tolentino, F.M.; Galo, M.d.L.B.T. Selecting features for LULC simultaneous classification of ambiguous classes by artificial neural network. *Remote Sens. Appl. Soc. Environ.* **2021**, *24*, 100616. [CrossRef]
55. Zha, Y.; Gao, J.; Ni, S. Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. *Int. J. Remote Sens.* **2003**, *24*, 583–594. [CrossRef]
56. Zheng, Y.; Tang, L.; Wang, H. An improved approach for monitoring urban built-up areas by combining NPP-VIIRS nighttime light, NDVI, NDWI, and NDBI. *J. Clean. Prod.* **2021**, *328*, 129488. [CrossRef]
57. Mishra, V.N.; Prasad, R.; Rai, P.K.; Vishwakarma, A.K.; Arora, A. Performance evaluation of textural features in improving land use/land cover classification accuracy of heterogeneous landscape using multi-sensor remote sensing data. *Earth Sci. Inform.* **2019**, *12*, 71–86. [CrossRef]
58. Hall-Beyer, M. Practical guidelines for choosing GLCM textures to use in landscape classification tasks over a range of moderate spatial scales. *Int. J. Remote Sens.* **2017**, *38*, 1312–1338. [CrossRef]
59. Mishra, V.N.; Prasad, R.; Kumar, P.; Gupta, D.K.; Srivastava, P.K. Dual-polarimetric C-band SAR data for land use/land cover classification by incorporating textural information. *Environ. Earth Sci.* **2017**, *76*, 26. [CrossRef]
60. Singh, D.; Singh, B. Investigating the impact of data normalization on classification performance. *Appl. Soft Comput.* **2020**, *97*, 105524. [CrossRef]
61. Cao, F.; Yang, Z.; Ren, J.; Jiang, M.; Ling, W.-K. Does Normalization Methods Play a Role for Hyperspectral Image Classification? *arXiv* **2017**, arXiv:1710.02939.
62. NUIS. *National Urban Information System (NUIS): Design and Srandards*; Ministry of Urban Development: New Delhi, India, 2006.
63. Sharma, S.; Hussain, S.; Singh, A.N. Impact of land use and land cover on urban ecosystem service value in Chandigarh, India: A GIS-based analysis. *J. Urban Ecol.* **2023**, *9*, juac030. [CrossRef]
64. Singh, Y. *Significance of Land Use/Land Cover (LULC) Maps | SATPALDA*; Geospatial Insight: Birmingham, UK, 2013.
65. Ouma, Y.O.; Keitsile, A.; Nkwae, B.; Odirile, P.; Moalafhi, D.; Qi, J. Urban land-use classification using machine learning classifiers: Comparative evaluation and post-classification multi-feature fusion approach. *Eur. J. Remote Sens.* **2023**, *56*, 2173659. [CrossRef]
66. Jason, B. Train-Test Split for Evaluating Machine Learning Algorithms. Machine Learning Mastery. 2020. Available online: https://machinelearningmastery.com/train-test-split-for-evaluating-machine-learning-algorithms/ (accessed on 11 September 2023).
67. Van Leeuwen, B.; Tobak, Z.; Kovács, F. Machine Learning Techniques for Land Use/Land Cover Classification of Medium Resolution Optical Satellite Imagery Focusing on Temporary Inundated Areas. *J. Environ. Geogr.* **2020**, *13*, 43–52. [CrossRef]
68. Shetty, S.; Gupta, P.K.; Belgiu, M.; Srivastav, S.K. Assessing the Effect of Training Sampling Design on the Performance of Machine Learning Classifiers for Land Cover Mapping Using Multi-Temporal Remote Sensing Data and Google Earth Engine. *Remote Sens.* **2021**, *13*, 1433. [CrossRef]
69. Colditz, R.R. An Evaluation of Different Training Sample Allocation Schemes for Discrete and Continuous Land Cover Classification Using Decision Tree-Based Algorithms. *Remote Sens.* **2015**, *7*, 9655–9681. [CrossRef]
70. Edgar, T.W.; Manz, D.O. Machine Learning. *Res. Methods Cyber Secur.* **2017**, *2017*, 153–173. [CrossRef]
71. Campos-Taberner, M.; García-Haro, F.J.; Martínez, B.; Izquierdo-Verdiguier, E.; Atzberger, C.; Camps-Valls, G.; Gilabert, M.A. Understanding deep learning in land use classification based on Sentinel-2 time series. *Sci. Rep.* **2020**, *10*, 1–12. [CrossRef] [PubMed]
72. Sertel, E.; Ekim, B.; Osgouei, P.E.; Kabadayi, M.E. Land Use and Land Cover Mapping Using Deep Learning Based Segmentation Approaches and VHR Worldview-3 Images. *Remote Sens.* **2022**, *14*, 4558. [CrossRef]
73. Storie, C.D.; Henry, C.J. Deep Learning Neural Networks for Land Use Land Cover Mapping. In Proceedings of the International Geoscience and Remote Sensing Symposium (IGARSS), Valencia, Spain, 22–27 July 2018; pp. 3445–3448.
74. Byvatov, E.; Schneider, G. Support vector machine applications in bioinformatics. *Appl. Bioinform.* **2003**, *2*, 67–77.
75. Gove, R.; Faytong, J. Machine Learning and Event-Based Software Testing: Classifiers for Identifying Infeasible GUI Event Sequences. *Adv. Comput.* **2012**, *86*, 109–135. [CrossRef]

76. Negri, R.G.; Dutra, L.V.; Sant'anna, S.J.S. An innovative support vector machine based method for contextual image classification. *ISPRS J. Photogramm. Remote Sens.* **2014**, *87*, 241–248. [CrossRef]
77. Mustak, S. Evaluating the Performance of Machine Learning Algorithms for Urban Land Use Mapping Using Very High Resolution. Master's Thesis, University of Twente, Enschede, The Netherlands, April 2018.
78. Oommen, T.; Misra, D.; Twarakavi, N.K.C.; Prakash, A.; Sahoo, B.; Bandopadhyay, S. An Objective Analysis of Support Vector Machine Based Classification for Remote Sensing. *Math. Geosci.* **2008**, *40*, 409–424. [CrossRef]
79. Belgiu, M.; Drăguţ, L. Random forest in remote sensing: A review of applications and future directions. *ISPRS J. Photogramm. Remote Sens.* **2016**, *114*, 24–31. [CrossRef]
80. Liu, Y.; Wang, Y.; Zhang, J. New machine learning algorithm: Random forest. *Lect. Notes Comput. Sci.* **2012**, *7473*, 246–252. [CrossRef]
81. Tsai, Y.H.; Stow, D.; Chen, H.L.; Lewison, R.; An, L.; Shi, L. Mapping Vegetation and Land Use Types in Fanjingshan National Nature Reserve Using Google Earth Engine. *Remote Sens.* **2018**, *10*, 927. [CrossRef]
82. Soner, Y. Hyperparameter Tuning for Support Vector Machines—C and Gamma Parameters. Towards Data Science. 2020, pp. 1–7. Available online: https://towardsdatascience.com/hyperparameter-tuning-for-support-vector-machines-c-and-gamma-parameters-6a5097416167 (accessed on 11 September 2023).
83. Goel, A.; Srivastava, S.K. Role of Kernel Parameters in Performance Evaluation of SVM. In Proceedings of the 2016 Second International Conference on Computational Intelligence & Communication Technology (CICT), Ghaziabad, India, 12–13 February 2016; pp. 166–169.
84. Ren, Y.; Hu, F.; Miao, H. The optimization of kernel function and its parameters for SVM in well-logging. In Proceedings of the 2016 13th International Conference on Service Systems and Service Management (ICSSSM), Kunming, China, 24–26 June 2016; pp. 1–5.
85. Pedergnana, 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]
86. Phan, D.C.; Trung, T.H.; Truong, V.T.; Nasahara, K.N. Ensemble learning updating classifier for accurate land cover assessment in tropical cloudy areas. *Geocarto Int.* **2022**, *37*, 4053–4070. [CrossRef]
87. Du, H.; Li, M.; Xu, Y.; Zhou, C. An Ensemble Learning Approach for Land Use/Land Cover Classification of Arid Regions for Climate Simulation: A Case Study of Xinjiang, Northwest China. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2023**, *16*, 2413–2426. [CrossRef]
88. Inamdar, A. Ensemble Learning Techniques in Machine Learning. 2021. Available online: https://www.fireblazeaischool.in/blogs/ensemble-learning-techniques-in-machine-learning/ (accessed on 16 November 2023).
89. Benbriqa, H.; Abnane, I.; Idri, A.; Tabiti, K. Deep and Ensemble Learning Based Land Use and Land Cover Classification. *Lect. Notes Comput. Sci.* **2021**, *12951*, 588–604. [CrossRef]
90. Hosna, A.; Merry, E.; Gyalmo, J.; Alom, Z.; Aung, Z.; Azim, M.A. Transfer learning: A friendly introduction. *J. Big Data* **2022**, *9*, 102. [CrossRef] [PubMed]
91. Baker, N.A.; Zengeler, N.; Handmann, U. A Transfer Learning Evaluation of Deep Neural Networks for Image Classification. *Mach. Learn. Knowl. Extr.* **2022**, *4*, 22–41. [CrossRef]
92. Elmahdy, S.I.; Mohamed, M.M. Regional mapping and monitoring land use/land cover changes: A modified approach using an ensemble machine learning and multitemporal Landsat data. *Geocarto Int.* **2023**, *38*, 2184500. [CrossRef]
93. Li, N.; Sepúlveda, N.; Li, N. IEEE Xplore. In Proceedings of the 2011 IEEE International Conference on Robotics and Biomimetics, Karon Beach, Thailand, 7–11 December 2011; pp. 1343–1348.
94. L3Harris Geospatial Documentation Center. *Calculate Confusion Matrices*; L3Harris Geospatial Solutions, Inc.: Boulder, CO, USA, 2022.
95. Ghosh, S.; Das Chatterjee, N.; Dinda, S. Relation between urban biophysical composition and dynamics of land surface temperature in the Kolkata metropolitan area: A GIS and statistical based analysis for sustainable planning. *Model. Earth Syst. Environ.* **2019**, *5*, 307–329. [CrossRef]
96. Wieland, M.; Pittore, M. Performance Evaluation of Machine Learning Algorithms for Urban Pattern Recognition from Multi-spectral Satellite Images. *Remote Sens.* **2014**, *6*, 2912–2939. [CrossRef]
97. Shahfahad; Naikoo, M.W.; Das, T.; Talukdar, S.; Asgher, S.; Asif; Rahman, A. Prediction of land use changes at a metropolitan city using integrated cellular automata: Past and future. *Geol. Ecol. Landscapes* **2022**, 1–19. [CrossRef]
98. Rudiastuti, A.W.; Farda, N.M.; Ramdani, D. Mapping built-up land and settlements: A comparison of machine learning algorithms in Google Earth engine. In Proceedings of the Seventh Geoinformation Science Symposium (GSS 2021), Yogyakarta, Indonesia, 25–28 October 2021; pp. 42–52.
99. Kavzoglu, T.; Colkesen, I. An assessment of the effectiveness of a rotation forest ensemble for land-use and land-cover mapping. *Int. J. Remote Sens.* **2013**, *34*, 4224–4241. [CrossRef]