Significance Analysis of Normalized Indices in Land Use and Land Class Classification in an Urban Landscape Using Google Earth Engine and Machine Learning

Introduction

The proven effectiveness of spectral indices like NDVI, NDBI, NDWI, BSI in assessing diverse features such as vegetation, water, built up and open surfaces has prompted researchers to leverage their potential in machine learning (ML) based multiclass land use and land cover (LULC) classification. However, the significance of this incorporation yet been understudied. This research focuses on investigating the impact of integrating spectral indices with satellite bands in ML based LULC classification, considering both classification accuracy and computational complexity. This approach contributes to future research in determining the suitability of incorporating index-based data into machine learning models for improved LULC classification.

The rapid urbanization and industrialization have accelerated transformation of LULC, leading to extensive impacts on landscapes, ecosystems, biodiversity, temperature, hydrology, and air quality at local and global levels (Mitchell, 2003). Tackling the ensuing challenges require informed decision-making, wherein up to date LULC mapping plays a pivotal role, which was cost, time and resource intensive process in the past(K.-A. Nguyen & Liou, 2019). The inception of the initial artificial Earth Observation (EO) satellite in 1972 propelled the advancement of land cover mapping, a progress that has been subsequently enhanced through technological strides, encompassing increased computational capabilities and sophisticated ML algorithms. Resulting to satellite image processing being one of the most important tools used by contemporary researchers for generating LULC maps (Mohajane et al., 2018; Yuh et al., 2023). Recent advancements in remote sensing (RS) technology, including satellites with advanced sensors and drones, have revolutionized imagery by offering higher spatial and temporal resolutions and additional spectral bands. The state-of-the-art employs machine learning and deep learning algorithms to automate the land cover mapping using these enhanced remotely sensed data for faster, precise, and accurate analyses.

In the literature satellite images have been leveraged in mostly 2 ways, either to identify single land cover using index based approaches, or using the machine learning algorithms for multi class classification. (C. Li et al., 2021) compared Normalized Built-up Area Index, combinational build-up index, perpendicular impervious surface index, Otsu's method, manual method, and Iterative Self-Organizing Data Analysis Technique Algorithm (ISODATA) classification method for impervious surface area extraction from multi-seasonal Sentinel-2 images in Wuhan, China. (Aryal et al., 2022) applied NDVI thresholding approach for urban green space mapping using Sentinel-2A data, (Madasa et al., 2021) analysed land-use/cover changes in South Africa's Welkom - Virginia Goldfields using geospatial indices, i.e., Normalized Difference Built-up Index (NDBI), the Normalized Difference Vegetation Index (NDVI), the Normalized Difference Soil Index (NDSI) and the Normalized Difference Water Index (NDWI) with maximum likelihood algorithm.

Indices-based classification has demonstrated success in RS as it does not need a real time training data (H. Li et al., 2017); however, it is not exempt from limitations. A significant concern is that the classification of an object can fluctuate based on the threshold value, (Bhandari et al., 2012)which is difficult to determine and relies on the researcher's area of interest(AOI) and literature study, introducing the potential for human error and bias. (Chen et al., 2006; H. Li et al., 2017) Additionally, spectral indices are influenced by atmospheric conditions and other environmental factors, leading to variations

in index values that may not solely reflect LULC changes (Ihuoma & Madramootoo, 2019). Considering the constraints posed by indices-based classification and the advancements in computational power, coupled with the successful implementation of machine learning techniques in various scientific domains, researchers are progressively embracing machine learning methods for LULC classification (Talukdar et al., 2020).

At its core, the essence of machine learning algorithms lies in their ability to recognize patterns within entirely new data, drawn through unsupervised learning from unattended information or from past knowledge(supervised) much like the human cognitive process. Hence, this approach offers cost-effective ways to enhance our understanding of causal connections between satellite data reflectance and physical appearance of land cover elements, enabling the prediction of land covers by leveraging existing datasets to train models (Woldemariam et al., 2022). The big domain of supervised ML algorithms used in RS classification, can be categorized broadly into two groups: shallow machine learning algorithms (i.e., Decision Trees, k-Nearest Neighbors (k-NN), Naive Bayes, Maximum Likelihood Classifier, Support Vector Machines (SVM), Random Forest (RF) etc.)(Liu et al., 2021) and the more complex neural network-based, commonly referred to as deep learning(DL) algorithms (i.e. Artificial Neural Network(ANN), Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), Fully Connected Networks (FCN) etc.)(Cheng et al., 2020; Krizhevsky et al., 2017). In the literatures the both shallow and DL approaches were found in the field of RS classification, where the shallow algorithms dominantly used pixel based approach on the other hand the DL algorithms followed the path of object based image analysis(OBIA) (Hossain & Chen, 2019; Jia & Qin, 2022; Yan et al., 2022). (De Luca et al., 2022; Nasiri et al., 2022) combined vegetation indices with Sentinel 1 and 2 images and European warm condition and Middle Eastern warm condition in Iran respectively and achieved very good overall accuracy using RF, on the other hand (Ju et al., 2021) also combined several vegetation indices and water index to identify land classes, in an humid palm tree dominated area in Malaysia but the overall accuracy were average. Similarly, (Hosseiny et al., 2022) leveraged both shallow ML(SVM, XGB, RF) and DL(CNN, ResNet50)algorithms to examine the performance addition of NDVI and other auxiliary data with Sentinel 2 bands to focusing to detect specific type of vegetation in Sweeden. The result of adding the additional bands were not satisfactory as they claimed the inclusion of additional data did result in higher OA. Furthermore, (Tassi & Vizzari, 2020) argued that integrating textural information (NDVI, BSI) results in improved classification accuracy for both Sentinel 2 and PlanetScope imagery when employing RF and SVM algorithms. Similarly, (Yousefi et al., 2022) also used support vector machine (SVM) algorithm for generating land use/cover maps from Sentinel-2 satellite imagery in humid and arid regions of Iran, (Rumora et al., 2020) explored the performance of SVM, RF, and XGBoost algorithm on Sentinel 2 images, with different atmospherically corrected images and found promising result. (Thanh Noi & Kappas, 2017) also used SVM, RF, and kNN for examining the land cover classification performance of the balanced and imbalanced dataset of differing sizes using Sentinel 2 images. From these scenarios, it becomes evident that the inclusion of spectral indices does not always yield consistent classification results. Nevertheless, this consistency aspect was not further emphasized in either of these cases.

This study aims to highlight this specific issue of machine learning-based classification of satellite images. While existing literature on both indices-based single-class classification and ML-DL algorithms-based multiple landcover classification has reported promising overall results, the introduction of spectral indices has led to inconsistent outcomes in various locations and data compositions. In order to comprehend the impact of incorporating spectral indices with satellite bands, we intend to apply same algorithms to the multiple combinations of datasets of one area to ensure same spatial, temporal, and atmospheric characteristics of the datasets.

Aim and Objectives

Aim

The primary goal of this study is to comprehensively evaluate the impact of integrating spectral indices into land use and land cover (LULC) classification.

Specific Objectives

i. To compare the overall and class-wise accuracy of different combinations of data layers and algorithms.
ii. To analyse the computational complexity associated with various data layers and algorithm combinations.
iii. To identify the optimal combinations of data layers and algorithms for effective LULC classification.
iv. To generate an updated LULC map for the specified AOI.

Research Questions

To achieve this goal, the study will address the following research questions:

i. How does the inclusion of spectral indices impact the overall accuracy and accuracy for individual land cover classes across various data layers and algorithm combinations?
ii. What are the variations in computational complexity observed when employing different data layers and algorithms for LULC classification?
iii. Which combinations of data layers and algorithms yield the best performance in terms of optimal LULC classification results?

These research questions are designed to provide insights into the efficiency of incorporating spectral indices, as well as to identify the most suitable data and algorithm combinations for achieving accurate and efficient LULC classification outcome for distinct land classes.

Materials and Methods

Study Area

The focus of the study was part Phnom Penh, the capital city of Cambodia. Over the past two decades, Phnom Penh has experienced remarkable urban expansion in terms of both its physical size and population. This growth has been fuelled by the construction of various industrial and residential areas, which have emerged as a result of substantial foreign investments in the city. (Thanh Son et al., 2022) The current area of the city is 679Km^2 which hosts 2,281,951 people. (Pheakdey et al., 2023) The exact study area showed in Figure 1 includes Phnom Penh and its periphery.

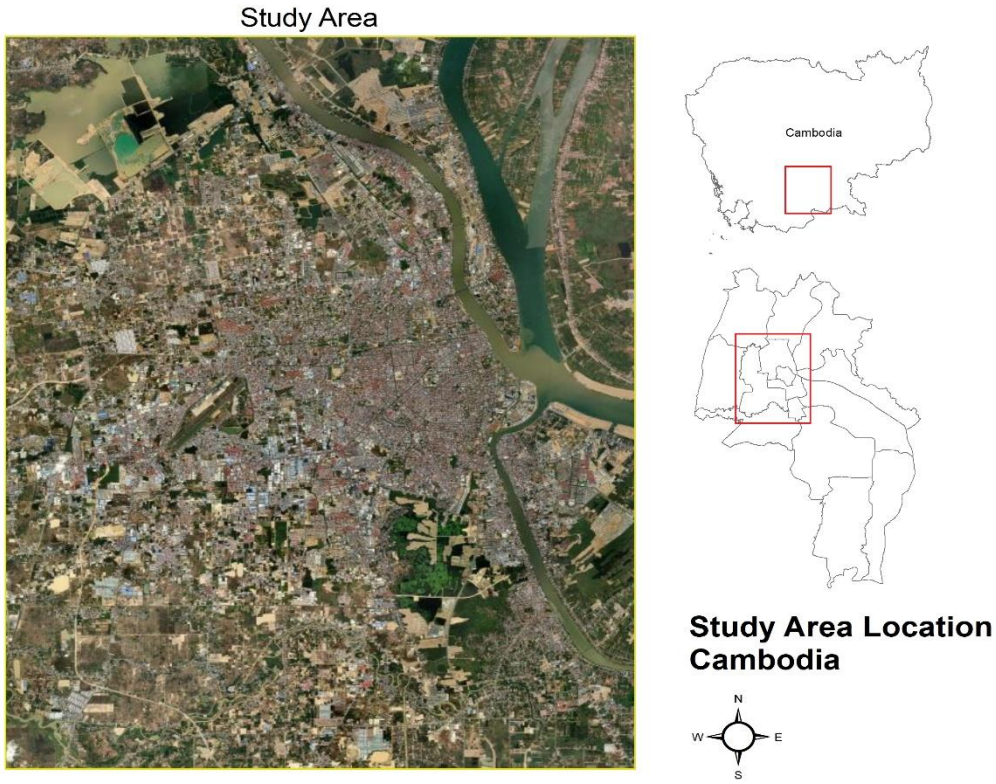


Figure 1: Study area and location

Data

In the literature, different satellite images have been used to classify land class and land uses (LULC) including Landsat 8, Sentinel-2, WorldView 2, 3 etc (E. D. Chaves et al., 2020; Jia et al., 2014; Xian et al., 2019). The choice of the imageries depends on the trade-off among spatial-temporal resolution, spectral availability, ease of open access, specific use case and study area etc (Burke et al., 2021). Despite the significantly enhanced resolution of very high-resolution (VHR) imagery like WorldView and PlanetScope from DigitalGlobe, challenges associated with open access limitations hinder their widespread usability. In this respect, the Sentinel 2 imageries have been predominantly used by the researchers from around the world for the LULC works because of high resolution, spatial-temporal coverage, and easy availability on platforms like Google Earth Engine which also provides sophisticated computational power (Macarringue et al., 2022). From critical perspective, another reason for preferring Sentinel 2 images is, it comes with a 10m resolution for blue (B2), green (B3), red (B4), and near-infrared (B8) channels have, while the Landsat, which is also popular for open accessibility, comes with a 30m resolution. In an urban environment, the spatial distribution of landcovers changes more frequently (Sun et al., 2017)in these sense that an 900m^2 may contain multiple types of land use whereas land cover maps using Landsat data will result in single landcover prediction for the total area (Zhang et al., 2022), which demands for the higher resolution images, where Sentinel 2 has three-fold capability than the Landsat satellites as well as VHR satellites in terms of open data availability. Taking these factors into consideration, the LULC classification was conducted using Sentinel-2 Harmonized Multi Spectral Instrument Level-2A data. This data, obtained from pre-processed and atmospherically corrected Sentinel-2 raw imagery, was offered within the Google Earth Engine (GEE) platform.

The training dataset comprised 3930 samples across 7 different classes. These training samples were generated based on a prior land use survey conducted in 201. Nevertheless, given that one of the objectives of the study was to produce updated LULC map for the AOI and considering substantial

changes in LULC over the past five years, the survey data was superimposed onto Google Earth desktop software and subjected to visual inspection. The final training dataset includes 3930 visually confirmed samples of 'Trees', 'RoadsAndPavements', 'SavannahGrassland', 'Water', 'AgriculturalLand', 'BareLand', and 'Building' in an ESRI shape file. In the similar process total of 840 testing samples were considered for testing purpose taking 120 samples from each class, these testing samples were not exposed to the training dataset. A short description of the classes can be found in following Table 1.

Table 1: Training Data Description

Class (Numerical Class)	**Description**	**Train Test Size**
Trees (0)	Areas with dense vegetations, trees, urban greens, forests	(580, 120)
Roads And Pavements (1)	Roads, highways, footpaths, unpaved roads	(580, 120)
Savannah Grass land (2)	Open lands dominated by grasses, scattered trees, shrub, unused real estate plots	(580, 120)
Building (3)	Human-made permanent residential, commercial, industrial structures;	(580, 120)
Water (4)	River, pond, canal, creek, open post mining flooded areas	(580, 120)
Agricultural Land (5)	Specifically used for farming and agricultural	(580, 120)
Bare Land (6)	Exposed soil, with no presence of vegetation	(450, 120)

Besides that, the AOI boundary was created using QGIS software. All the dataset used were in EPSG:32648 - WGS 84 / UTM zone 48N coordinate system. The preprocessing of the data was done using the GEE Python API, and various other Python packages on a Microsoft Windows 11, 64-bit computer with a hardware specification of Intel Core i5 processor, 1 Terabyte storage, 16 Gigabyte memory.

Methodology

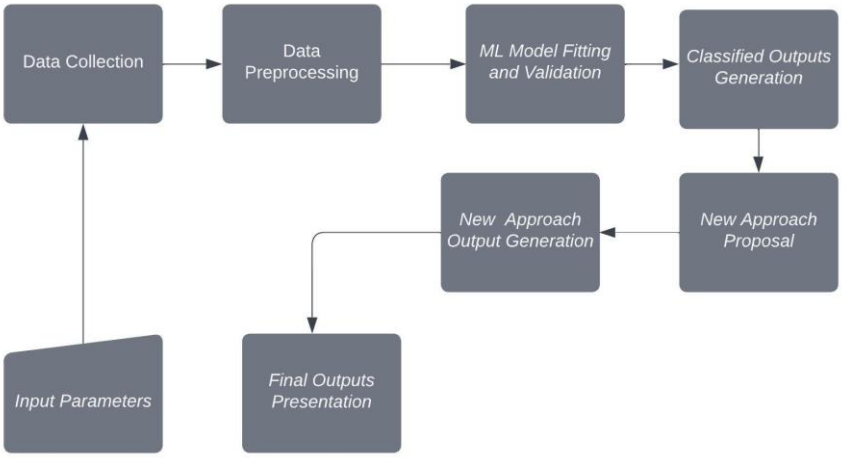


Figure 2: A generalized representation of the workflow

Data Acquisition

The sentinel 2 harmonized surface reflectance collection were derived from GEE data catalogue. GEE data catalogue holds petabytes of various data, among which 'COPERNICUS/S2_SR_HARMONIZED' index holds the imageries for Sentinel 2 Harmonized Surface reflectance data(Kumar & Mutanga, 2018). To skip the complexity of processing this gigantic data, we filtered the GEE S2_SR_HARMONIZED collection for our AOI using MGRS_TILE no '48PVT' which covers the AOI completely, other MGRS_TILEs touch the AOI but not cover completely. Further filtering was done for the time of the image, we were particularly interested in the images of 2023 there fore the filtering was done for 2023 from the month of January to December. A cloud filter of 2% 'CLOUDY_PIXEL_PERCENTAGE' and 1% 'DARK_FEATURES_PERCENTAGE' was also imposed. After applying the required filters, only one Sentinel 2 image was found.

Data Preprocessing

The Sentinel 2 SR Harmonized collection comes with atmospheric correction and radiometric correction (*User Guides - Sentinel-2 MSI - Level-2 Processing - Sentinel Online*, n.d.).

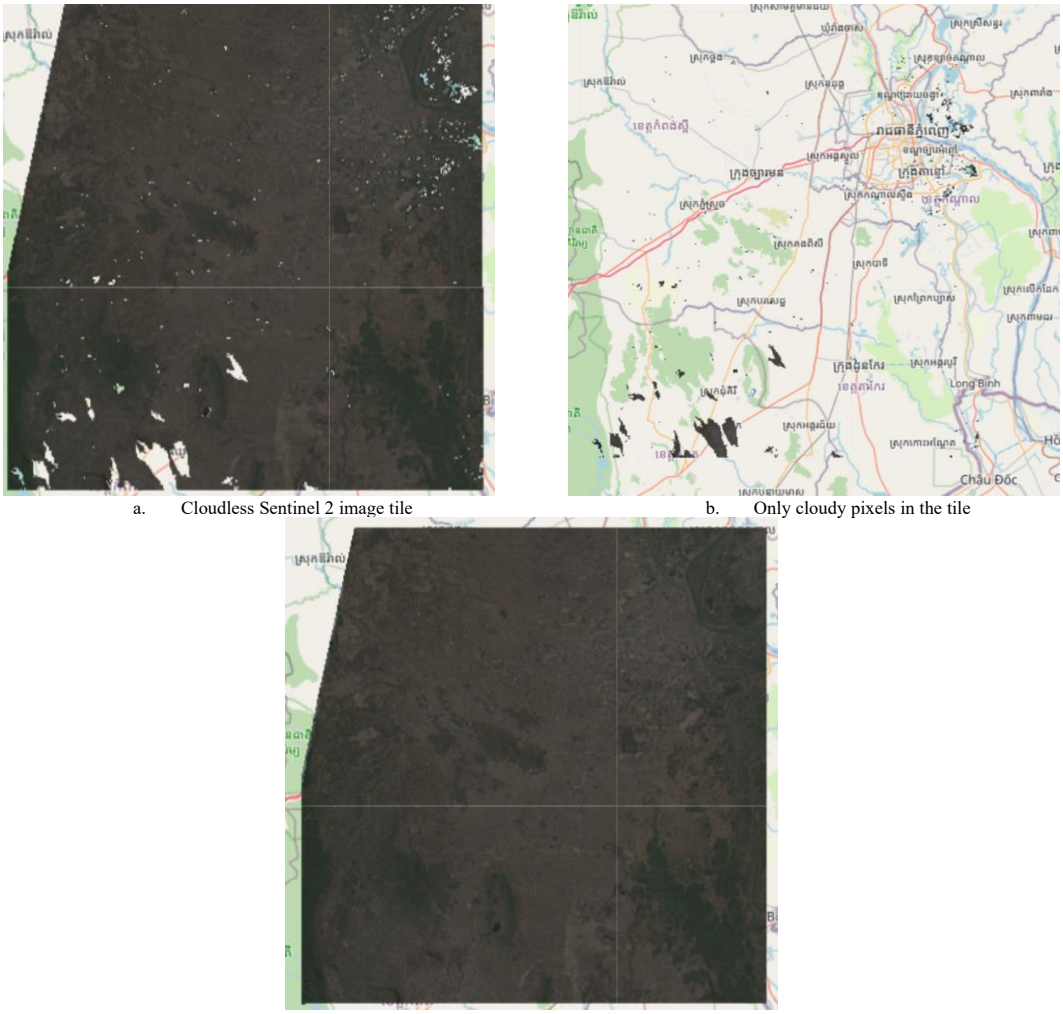
Cloud Masking and Gap Filling

Although a strict cloud filtering was imposed, the collected image was again cloud masked to remove the cloud shadow, low, high, and medium cloud, and cirrus pixels based on 'SCL' band. (Zekoll et al., 2021). In 'SCL' band the bitwise information is shown in table 2 below.

Table 2: Sentinel 2 SCL Class Table

Bit No	Description	Bit No	Description
1	Saturated or defective	**7**	Clouds Low Probability / Unclassified
2	Dark Area Pixels	**8**	Clouds Medium Probability
3	Cloud Shadows	**9**	Clouds High Probability
4	Vegetation	**10**	Cirrus
5	Bare Soils	**11**	Snow / Ice
6	Water		

For removing the mentioned objects, the bits 3, 7-10 were used. After the cloud masking, the masked gaps were filled by 21 image composites of the same timeframe with a cloud probability of 50%. The following figure 2 shows the cloudless, filtered cloud, and cloud filled images.



a. Cloudless Sentinel 2 image tile

b. Only cloudy pixels in the tile

c. Cloud filled Sentinel 2 image tile

Figure 3: Cloud masking of the selected image tile

Resampling

Sentinel 2 dataset comes with different spatial resolution for individual bands starting from 10 – 60 meters (Claverie et al., 2018). In an urban set up, the land cover changes rapidly in terms of spatial distribution, where bands having the smaller spatial resolution is better. In this case, the Sentinel 2 SR Harmonized dataset, comes with 10-meter resolution for Band 2, Band 3, Band 4, Band 8. The rest of the bands as well were resampled to 10 meters using 'Bilinear' approach. (S Malini & Patil, 2018; Wang, 2022)

Band selection

Sentinel 2 SR Harmonized collection comes with 23 bands out of those, Band 1, 2, 3, 4, 5, 6, 7, 8, 8A, 9, 11 and 12 were selected. This is to note that, QA10, QA20, and QA60 bands in Sentinel 2 SR are always empty.

Indices Generation

With the aim being to analyse the impact significance of indices incorporation, the objective of the task also includes generating an updated LULC map. To achieve this, in addition to the literature studies we considered characteristics of the AOI, and decided to include class-specific indices NDVI, NDBI, NDWI, BSI. From the literature it has been seen that, these indices help identifying land cover and vegetation characteristics, dynamics, and health, allowing assessment of land surface properties and changes. The Normalized Difference Vegetation Index (NDVI) evaluates vegetation health by analysing

NIR and red band reflectance, while the Normalized Difference Water Index (NDWI) identifies water bodies through NIR and SWIR reflectance. The Normalized Difference Built-up Index (NDBI) detects urban areas by contrasting SWIR and NIR reflectance. High NDVI values indicate healthy vegetation, high NDWI values indicate water presence, high NDBI values represent built-up areas, and the Bare Soil Index (BSI) assesses bare soil or non-vegetated areas by comparing red and SWIR bands. (Kuc & Chormański, 2019; McFEETERS, 1996) High BSI values indicate bare soil, while lower values correspond to vegetated regions. (C. T. Nguyen et al., 2021) Therefore, alongside 12 bands from sentinel 2 dataset, Normalized Difference Built Up index, Water Index, Vegetation index, and Bare Soil index is considered were created using the following formulas-

$$NDVI = \frac{NIR(B8) - Red(B4)}{NIR(B8) + Red(B4)} \dots\dots\dots\dots\dots\dots\dots\dots (1)$$

$$NDBI = \frac{SWIR(B11) - NIR(B8)}{SWIR(B11) + NIR(B8)} \dots\dots\dots\dots\dots\dots\dots\dots (2)$$

$$NDWI = \frac{Green(B3) - NIR(B8)}{Green(B3) + NIR(B8)} \dots\dots\dots\dots\dots\dots\dots\dots (3)$$

$$BSI = \frac{[SWIR2(B12) + Red(B4)] - [NIR(B8) + Blue(B2)]}{[SWIR2(B12) + Red(B4)] + [NIR(B8) + Blue(B2)]} \dots\dots\dots\dots\dots\dots\dots\dots (4)$$

Defining Training and Testing Data

After the creation of indices, the sentinel 2 selected bands(band1-12) and corresponding indices were merged and created an image of 16 bands. At this phase, the pixel values for each band against each training samples were identified in the image and pixel values were extracted. Thus, the final training dataset of 3930 samples was constructed where each training samples consist of the corresponding spectral values from Sentinel 2 bands, NDVI, NDBI, NDWI, BSI and the corresponding class. In the similar way, the pixel values for the testing dataset were also extracted and to construct the testing dataset. This is worth mentioning that the testing dataset was constructed separately, therefore no train test split was necessary. The training samples were shuffled, and positional indices were reset to remove any kind of positional bias and introduce randomness in the data. At the end, the final shape of training and testing dataset were (3816, 16) and (994, 16) respectively.

Table 3: Training and testing data composition

Class (Numerical Class)	Train Shape		Test Shape	
	NoIndices	With4Indices	NoIndices	With4Indices
Trees (0)	(580, 12)	(580, 16)	(120, 12)	(120, 16)
Roads And Pavements (1)	(580, 12)	(580, 16)	(120, 12)	(120, 16)
Savannah Grass land (2)	(580, 12)	(580, 16)	(120, 12)	(120, 16)
Building (3)	(580, 12)	(580, 16)	(120, 12)	(120, 16)
Water (4)	(580, 12)	(580, 16)	(120, 12)	(120, 16)
Agricultural Land (5)	(580, 12)	(580, 16)	(120, 12)	(120, 16)
Bare Land (6)	(450, 12)	(450, 16)	(120, 12)	(120, 16)

Model Construction and Training

Support Vector Machine

The implementation utilized the ‘SVC()’ model from the scikit-learn(Pedregosa et al., 2011) Python package to perform C-Support Vector Classification using the 'libsvm' backend. As the algorithm's

performance significantly relies on well-chosen hyperparameters, a single value for hyperparameters might not yield optimal results. Thus, a comprehensive Grid Search (Liashchynskyi & Liashchynskyi, 2019) strategy was adopted for effectively tuning the regularization parameter (C) and the kernel coefficient (gamma) of the SVM algorithm. To conduct this search, a parameter space was defined, considering values from the ranges [2,4] for C and [3,4] for gamma. The SVM algorithm employed two different kernels, namely 'poly' and 'radial basis function (rbf)'. This generated a 2x2 grid of hyperparameter combinations. The remaining parameters was kept their default values, as provided by the scikit-learn Python library. This tuning procedure involved evaluating all potential candidates from these combinations in the hyperspace. Each model configuration was assessed through 2-fold cross-validation to estimate its performance. Hence, the iterative process resulted in a total of 16 model fittings.

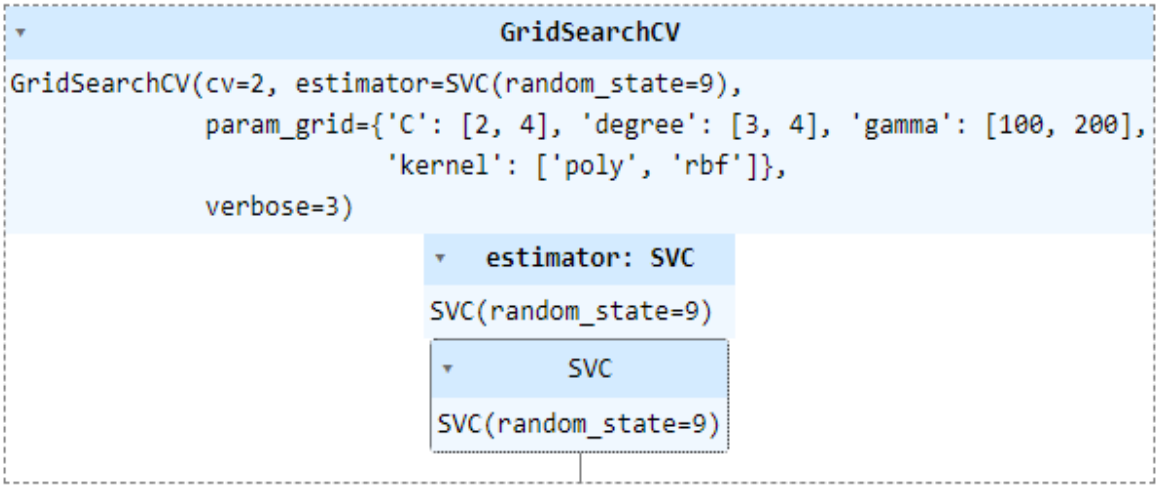


Figure 4: SVM hyperparameter grid for tuning

By methodically traversing this hyperparameter space and assessing models via cross-validation for both of the datasets (NoIndices, With4Indices), the configuration that showed the highest performance was selected as the best models for the classification for each dataset. The outcome of this process found that, while the optimal model associated with 'NoIndices' dataset had hyperparameters configuration of **[C = 4, gamma = 200, kernel = ('rbf')]**, the optimal configuration for 'With4Indices' dataset was **[C = 4, gamma = 100, kernel = ('rbf')]**. The finalized optimal model, representing the culmination of this exploration, was saved for subsequent deployment in the final imagery classification phase and subsequent result analysis.

Random Forest

The ensemble classifier architecture of Random Forest algorithm, which utilizes Bootstrap Aggregating, was applied using RandomForestClassifier() method of scikit-learn Python library . To determine the optimal configuration of the Random Forest model, here as well a Grid Search was conducted to explore the parameter space.

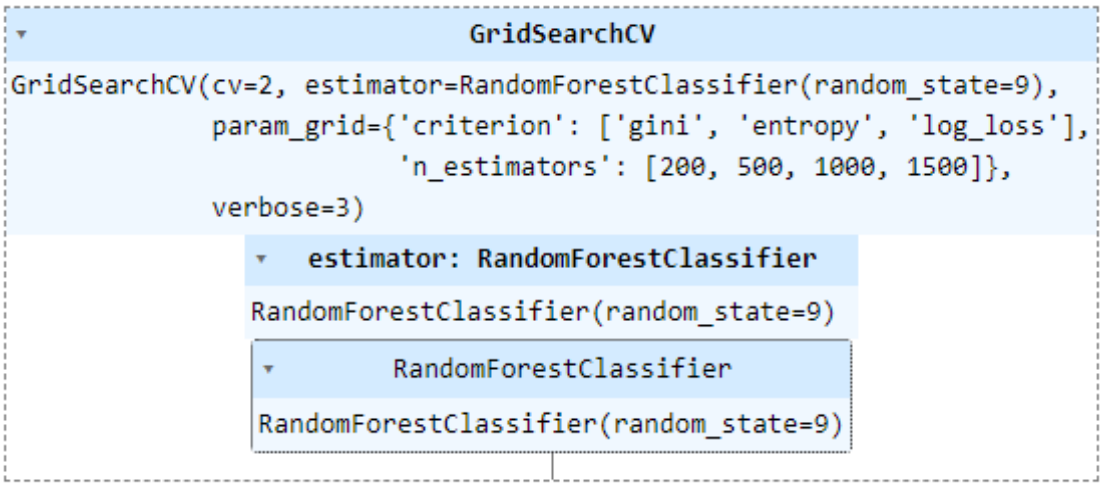


Figure 5: RF hyperparameter grid for tuning

The key hyperparameters considered were the number of estimators (200, 500, 1000, 1500), which determines the number of decision trees in the ensemble; function quality measuring criterion ('gini', 'entropy', and 'log_loss'), which defines the quality measure for splitting nodes. The rest of the parameters were let as default to the python scikit-learn package.

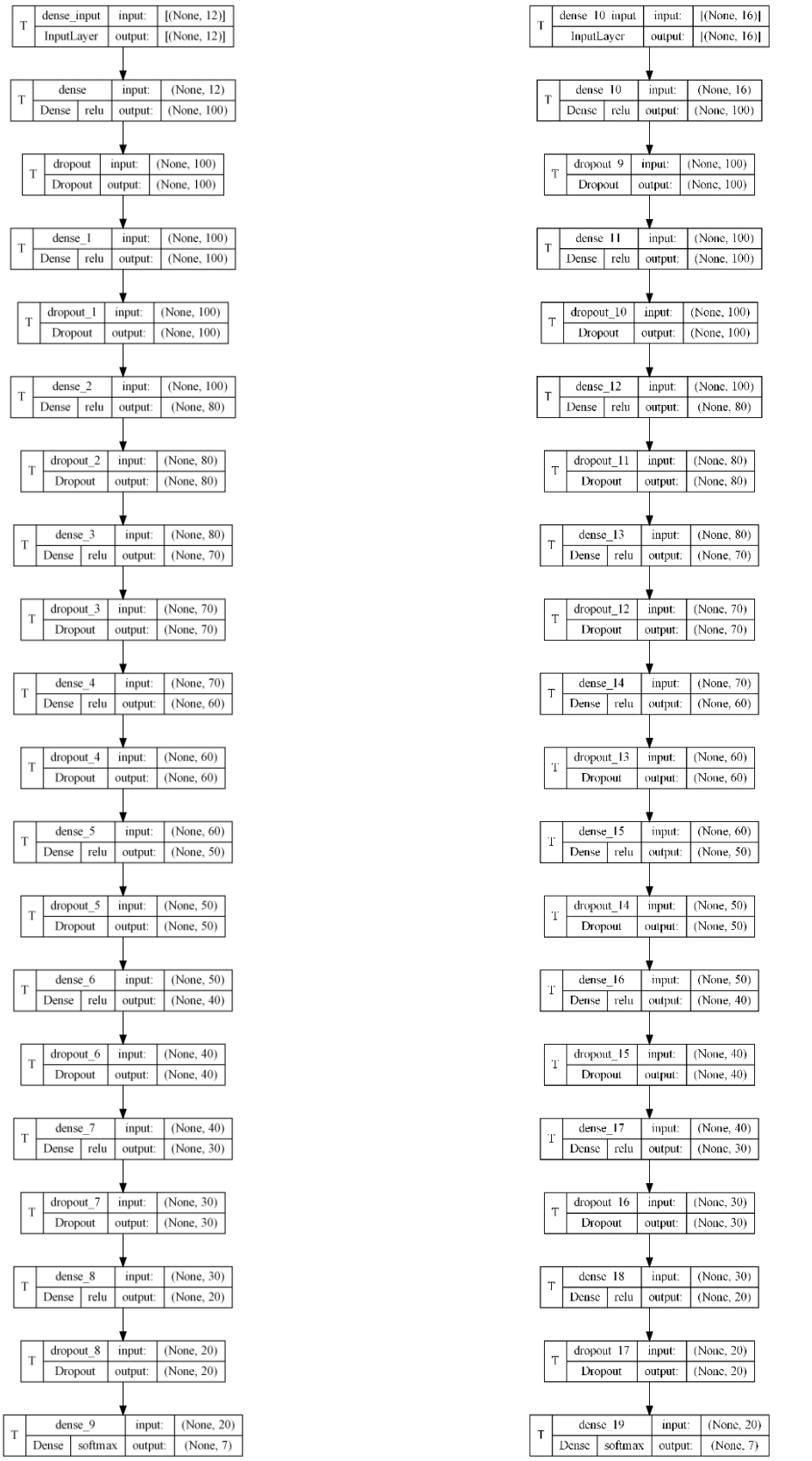
The configuration of the hyperspace, the grid tuning process checked all 12 candidates from the combinations, performance of which were validated using 2-fold cross validation, totalling fitting of model 24 times. At the end, the resulting optimal model had the configuration **['criterion': 'gini', 'n_estimators': 200]** for 'NoIndices' dataset; on the other hand, the optimal configuration for 'With4Indices' dataset was **['criterion': 'entropy', 'n_estimators': 1500]**.

Artificial Neural Network

The Artificial Neural Network (ANN) was implemented to construct 2 different models for 'NoIndices' in figure 6(a) and 'With4Indices' in figure 6(b) dataset. The key architecture was similar for both of the models, except the only difference being in the shape of first dense (input) layer between. The model architecture consists of multiple fully connected dense layers with varying numbers of neurons and dropout layers for regularization.

The first layer in the model is a dense layer with 100 neurons with 12 features for 'NoIndices' dataset and 16 features for 'With4Indices' dataset, which applies the rectified linear unit (ReLU) activation function. The ReLU activation function introduces non-linearity to the model, allowing it to learn complex patterns in the data. To prevent overfitting and improve generalization, a dropout layer with multiple dropout layers with rate of 0.2 is added after the dense layers. Dropout randomly sets a fraction of input units to 0 during training, which helps prevent the model from relying too heavily on specific input features. The subsequent layers follow a similar pattern: a dense layer with decreasing numbers of neurons and a dropout layer with a dropout rate of 0.2 after each dense layer. The purpose of these additional layers is to extract and capture higher-level features from the input data, gradually reducing the dimensionality and complexity of the representations. The model continues with several dense layers, including layers with 80, 70, 60, 50, 40, 30, and neurons, each followed by a dropout layer. These layers enable the network to learn increasingly abstract and discriminative representations of the data.

Finally, the output layer consists of a dense layer with 7 neurons, corresponding to the 7 land cover classes being classified. The activation function used in this layer is 'softmax', which outputs probabilities for each class, indicating the likelihood of the input data belonging to each category.



a. Model Structure ANN NoIndices

b. Model Structure ANN With4Indices

Figure 6: ANN Model Structures

During training, the Adam optimizer is utilized with a learning rate of 0. 001. Adam is an optimization algorithm that adapts the learning rate on a per-parameter basis, providing efficient and effective updates to the model weights during training. The categorical cross-entropy loss function is employed, which is

suitable for multi-class classification problems. The model is trained for a maximum of 500 epochs, with a batch size of 16. Nevertheless, early stopping callback was implemented based on the validation loss to monitor the loss value and stop training if it does not improve after a 20 consecutive number of epochs. Early stopping helps prevent overfitting by terminating training when the model's performance on the training data starts to deteriorate.

Proposed new approach

After the fitting of SVM, RF and ANN model with both form of dataset, it was seen that different models are performing better in detecting particular classes. The idea of the new proposed approach is to combine the predictions best performing models for each class.

$$Precision = \frac{TP}{(TP + FP)} \dots\dots\dots\dots\dots\dots\dots (5)$$

From the fitted models, the classification matrices and classified outputs were generated. Then, from the classification reports, for each of the classes corresponding best predicting models were identified based on class precision calculated by equation (5) and stored in a matrix in all models. The matrix was later sorted to set the hierarchy among the classes. Based this hierarchy for each class the best model's output raster were traversed and the pixels which contains value same as that class were picked and set in the new output image. The process was repeated for all the classes. At the end, the remaining unfilled pixels were identified and assigned the same class that was identified by the model with highest weighted average precision. The figure 6 represents the summary of the methodology of the study.

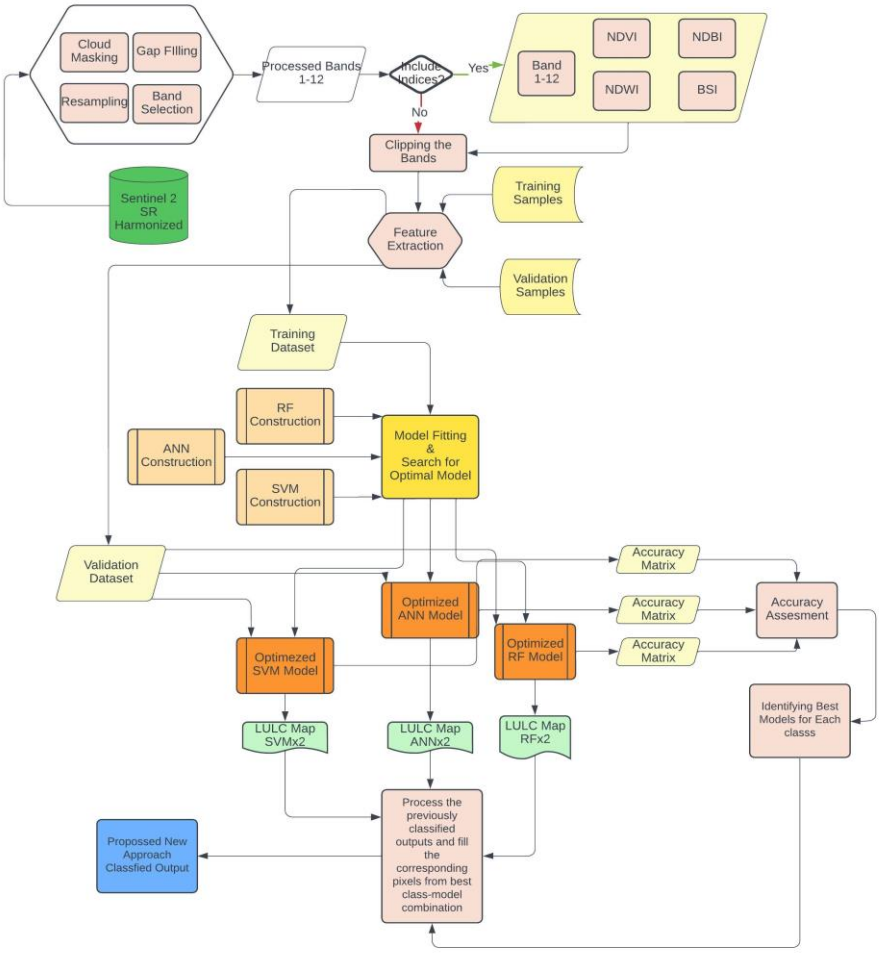


Figure 7: Detailed Methodology of the study

Results

For the performance evaluation total 840 samples were collected (120 random sample for each class) in the study area extent. These samples were not exposed to the model training. The landcovers for these locations were identified visually using the Google Earth images. For evaluation, the classification report, presenting precision, recall, and F1-score for each class, was generated to assess the model's performance.

Accuracies of the models

All the attempted models trained with Sentinel 2 data showed good overall accuracy (OA). The best and worst OA were shown by 'SVMNoIndices' (94.40%) and ANNNoIndices (88.33%) models respectively. Incorporating the normalized indices (NDVI, NDBI, NDWI, BSI) has distinct impacts on the LULC classification performance of the SVM, RF, and ANN models. For instance, RF and ANN performs considerably lower than the SVM model. However, combining the indices with the training data significantly influences the results, ameliorating the OA of ANN and RF models. With the additional bands, the accuracy of ANN increased to 90.12% (increase of 1.79%), which were previously the worst among all the experimented models. Although in small margin (0.48%), RF Models also gained OA from 89.5% and 90.0% because of indices. However, the OA deteriorated by 0.59% for SVM when the mentioned indices were added to train the model.

The observation of class wise accuracies shows that, the highest accuracy for 'Tree' class was 95.8% achieved exclusively by 'SVMWith4Indices' model. On the other hand, the best accuracies for 'Roads and Pavements', 'Water', 'Agricultural Land', and 'Bare Land' are 98.3%, 97.5%, 92.5%, and 95.8% respectively which is achieved by 'SVMNoIndices' model. The best (92.5%) accuracy for the 'Building' class was achieved by both, 'SVMNoIndices', and 'ANNWith4Indices' model. It is interesting that, the accuracy of 'Building' for ANN was as low as 85.8% for the without indices dataset which reached to the highest (92.5%) when indices were combined in the dataset. Similarly, the best accuracy (89.2%) for 'Grassland and Savanah' class was achieved by SVM models (NoIndices and With4Indices) in both combinations of dataset.

The overall impression of the index's incorporation is it has an impact on bot OA and class wise accuracies. These impacts vary based on the models and classes, which can be summed up in the following figure 7 and figure 8. Figure 7 shows the OA and the class wise accuracies for all the models and figure 8 summarizes the changes in the accuracies due to the addition of indices. It demonstrates that, the accuracy of- 'Tree' class always increases, class 'Grassland and Savanah' also either increases or remains unchanged, Water class always decreases, and the rest classes shows mixed changes when the indices are included in the training phase.

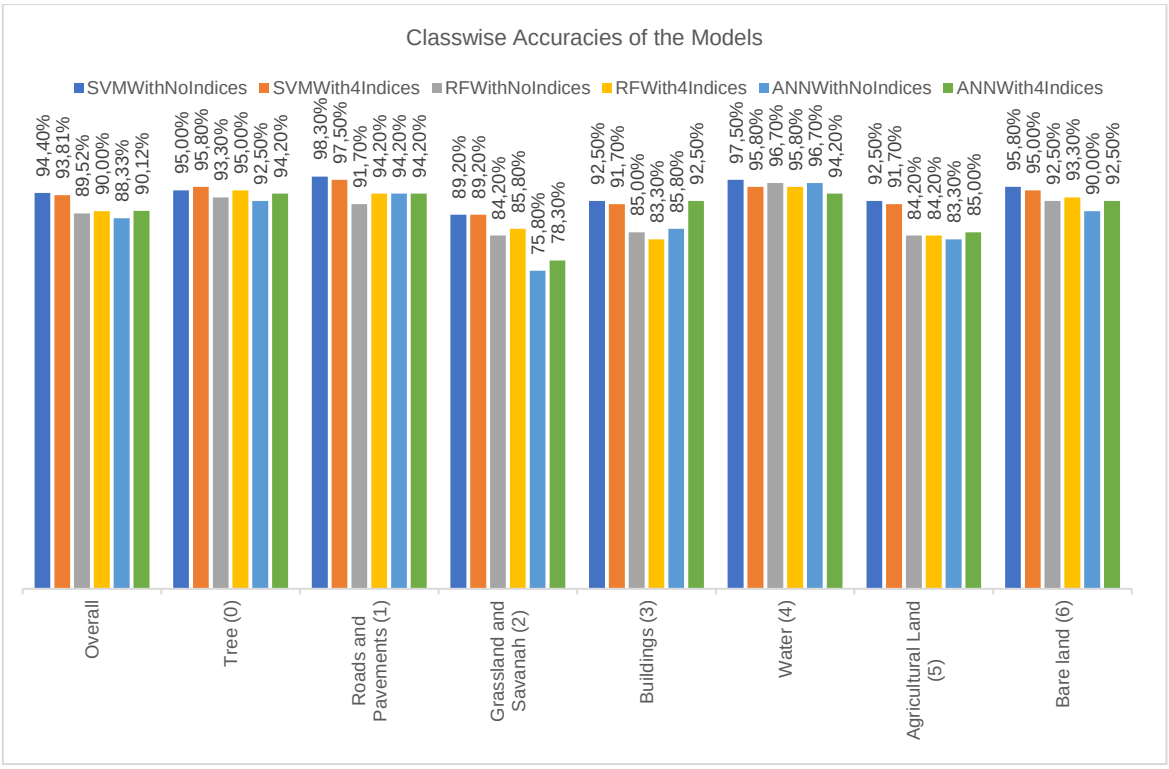


Figure 8: OA and class wise accuracies of the attempted models

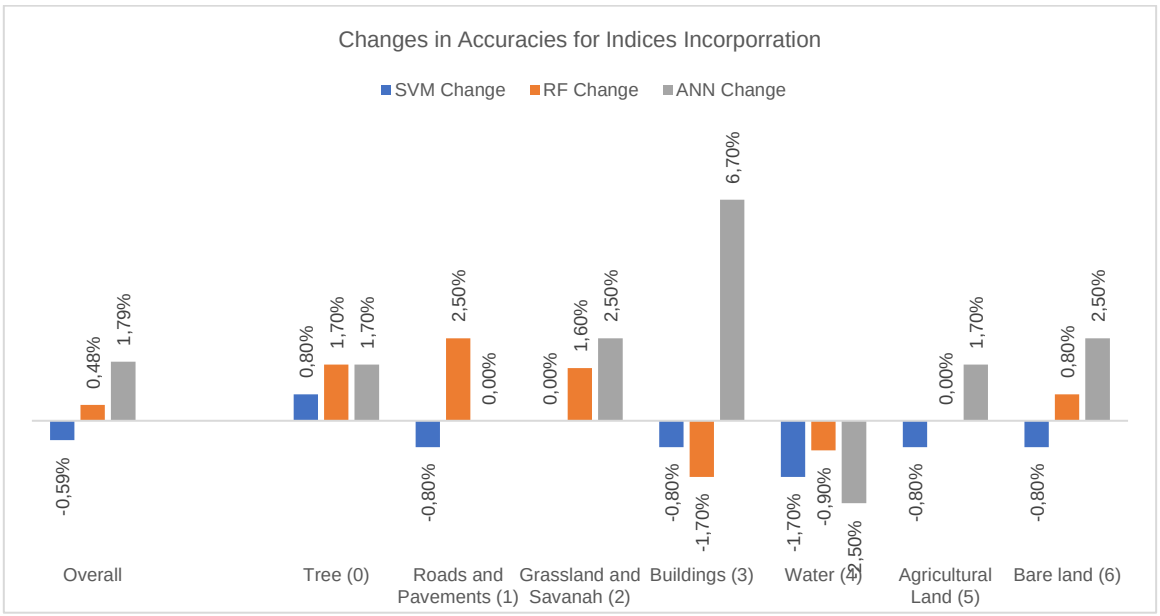
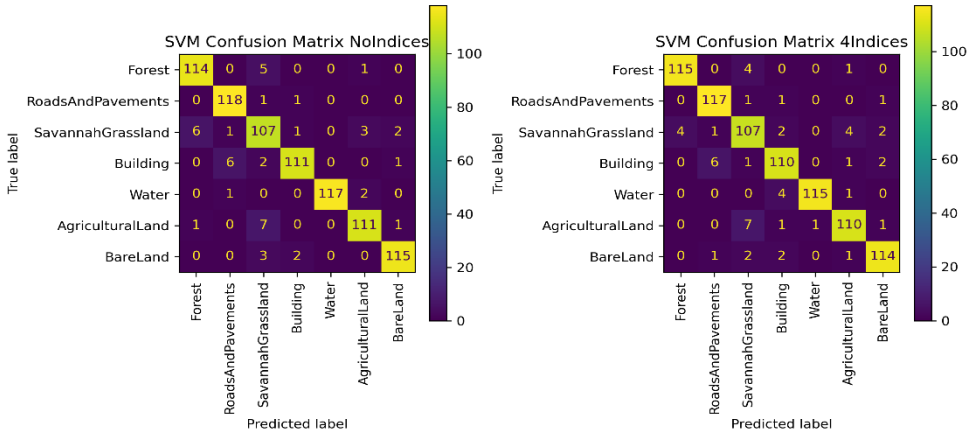


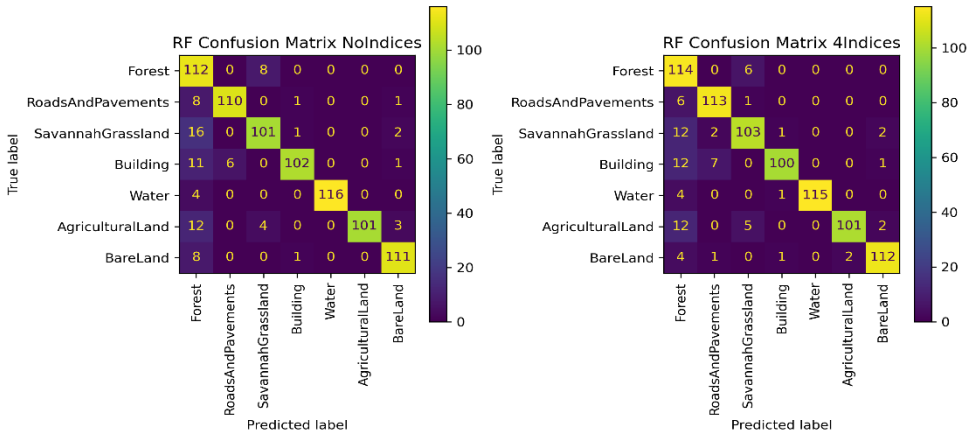
Figure 9: Impact on accuracy due to the index's incorporation.

In cases where the inclusion of index data caused fluctuations in accuracies, the extent of these changes varied. For SVM, similar to the OA, the class wise accuracies SVM also dropped for most of the classes. The highest drop in the accuracy was 1.7% for the 'Water' class while 'Roads and Pavements', 'Buildings', 'Agricultural Land', and 'Bare land' class also faced a 0.80% accuracy decrease. However, the 'Tree' class was better identified by the SVMWith4Indices model as the number of correctly identified samples were increased from 114 to 115, while the accuracy for 'Grassland and Savanah' class remained stable. Unlike the SVM model, the OA of the Random Forest classifier increased by 0.48% margin with the addition of the indices with the training data. The class wise accuracies enhanced by 1.7%, 2.5%, 1.5%, and 0.80% respectively for 'Tree', 'Roads and Pavements', 'Grassland and

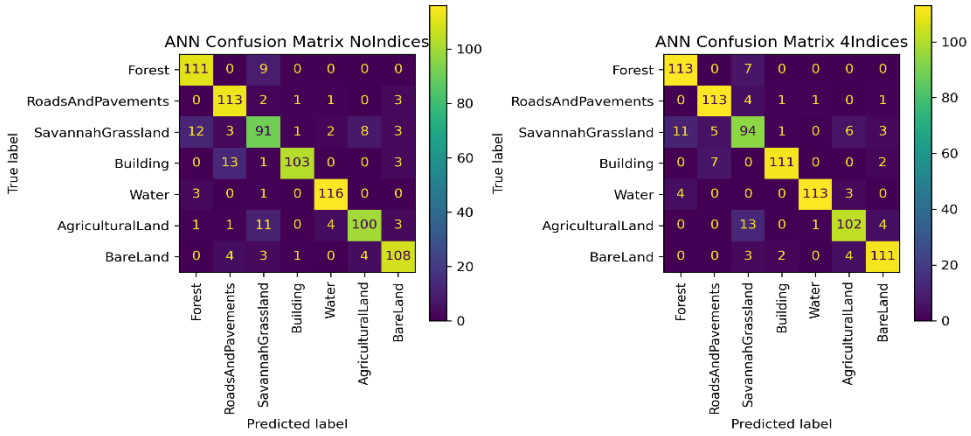
Savanah', and 'Bare land' classes. But the accuracy for 'Agricultural Land' class remained unchanged at 84.2%. The decrease in accuracies for 'Buildings' and 'Water' class were 1.7% and 0.9%. Among 3 of the experimented algorithms, most positive impact of incorporation indices with the training data was seen for ANN models. When only the bands were used for the classification, 'ANNNoIndices' model had the lowest OA of 88.33%, which was increased by 1.8% for 'ANNWith4Indices' model. The class wise accuracies for increased for Tree (1.7%), Grassland and Savanah (2%), Buildings (6.7%), Agricultural Land (1.7%), and Bare Land (2.5%). Similar to other models, the accuracy for 'Water' dropped by 2.5% with the inclusion of indices, but for 'Roads and Pavements' class it remained unchanged at 94.2%.



a. Confusion matrix of SVM models



b. Confusion matrix of RF models



c. Confusion matrix of ANN models

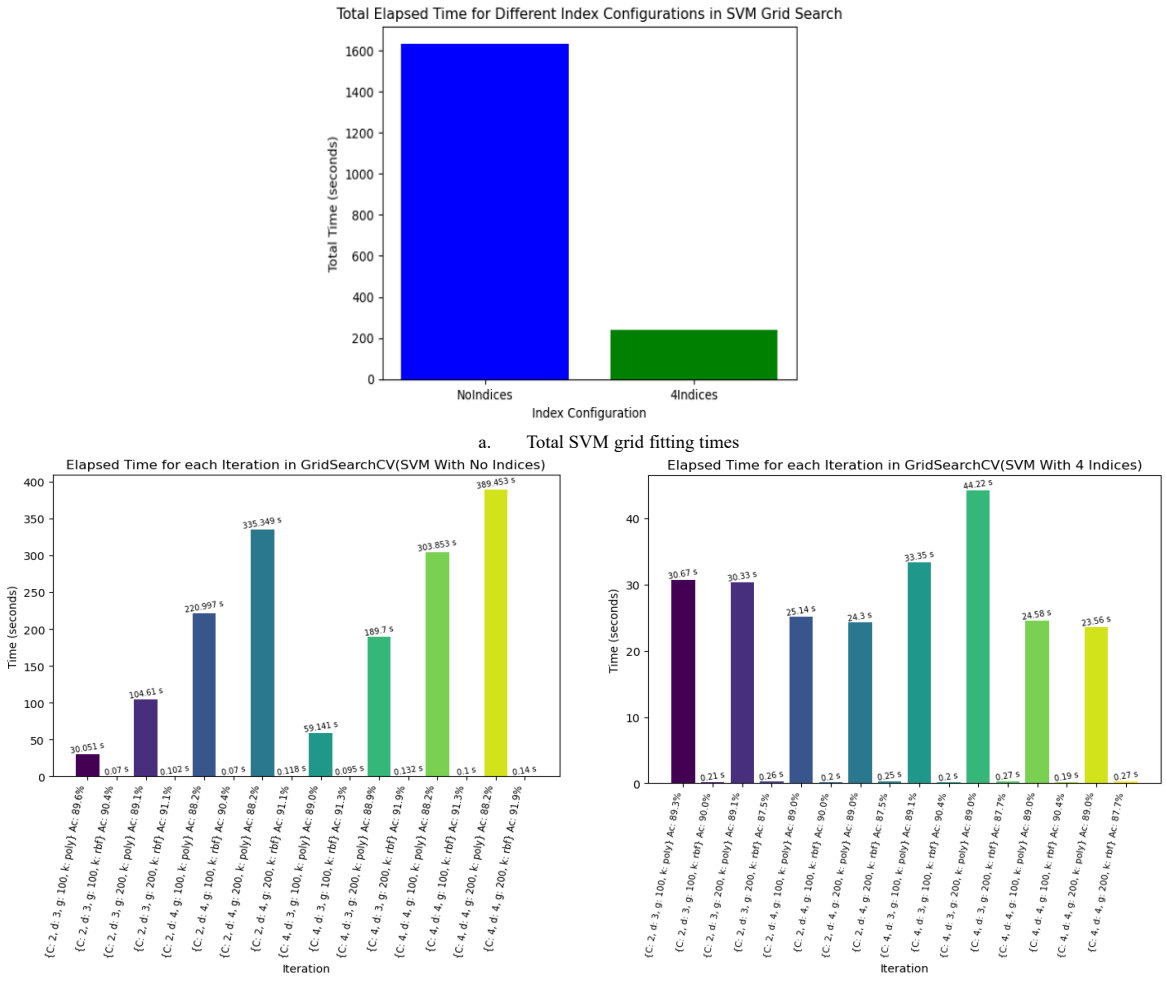
Figure 10: Confusion metrices of the attemted models

Based on above accuracy observations, it is evident that when the choice of algorithm is either Random Forest or ANN, the dataset 'With4Indices' allows models to better classify than the 'NoIndices' dataset. On the contrary, when SVM is employed for classification, the 'NoIndices' dataset shows better performance. Despite this drop in the OA, the data incorporation influences SVM model performance positively in terms of computation times. These performances can be further realized by the class wise true and false positive counts from confusion metrices presented in figure 9 above where figure 9a, 9b, and 9c represent the prediction statistics for SVM, RF, and ANN models respectively.

Time complexity of the ML models

In addition to the impact on classification accuracy, there was interesting observation of regarding computation time across the various training scenarios. As previously mentioned, it was observed that both overall accuracy (OA) and class-wise accuracy for specific classes declined when spectral indices were included in the training data for the SVM algorithm. Nevertheless, the computation time for the hyperspace was significantly dropped when the indices were added to the training data during training process of SVM model, which can be realized in the following graphs in figure 8.

The defined hyperspace had 16 different model fits. The total elapsed time for fitting these 16 model combinations was 1633.98 seconds for SVM algorithm when only the bands 1-12 were used. When additional data namely the 4 indices were added the total model fitting time for the same hyperspace was reduced to 237.99 seconds. This reduction is resulted from the drop of time complexity in fitting the SVM models which had Polynomial kernel.



a. Total SVM grid fitting times

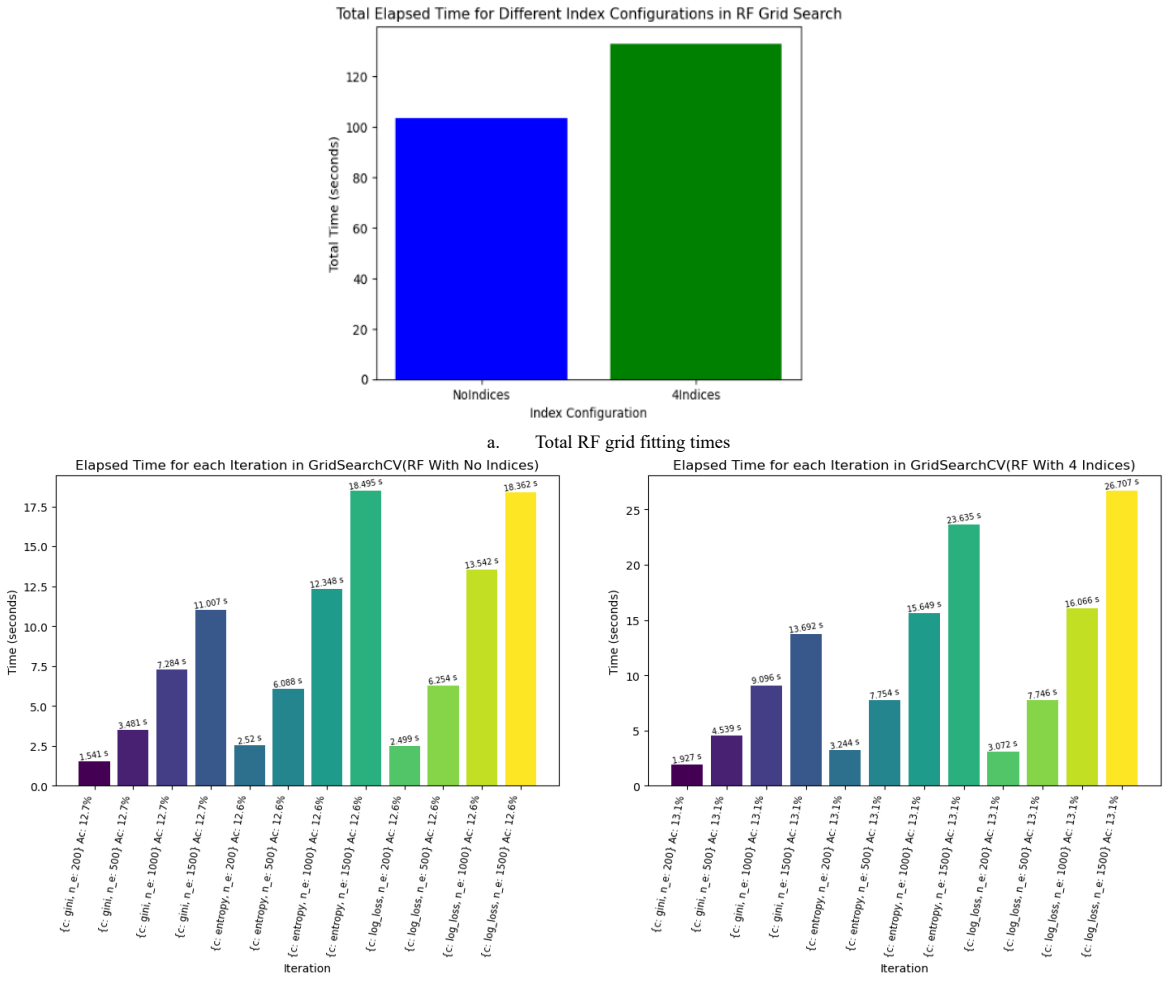
b. Indicvidual fitting time of SVM models in grid with no indices data

c. Indicvidual fitting time of SVM models in grid with 4 indices data

Figure 11: Time complexity of SVM models

It was seen that, the SVM models which uses polynomial function for higher-dimensional space mapping, they need significantly more time than the models which use RBF function when only bands were used. When the additional indices were included in the training data, the time drops sharply. The drop ranges from 43.61 % to 93.91% of the training time when only bands were used. Besides that, although in narrow margin, the increase in the accuracy for was also observed in these models while using indices. The only case when the computation time was increased with using indices were for {'C': 2, 'degree': 3, 'gamma': 100, 'kernel': 'poly'} model; however, the accuracy also showed decrease for this particular model. On the other hand, the indices incorporation resulted in reduction in the accuracy and as well as increase in the computation time. Therefore, it could be concluded that, the SVM models performance for LULC highly depends on the kernel selection; especially SVM models which use Polynomial kernel, including the indices improves both accuracy and computation time for those.

For RF hyperparameter tuning, 12 models with differing splitting criterion and estimator numbers. The total time of the hyperspace tuning has risen from 103.42 seconds to 133.13 seconds when additional indices were incorporated with the training data. The following graphs show the changes in time complexity for the hyperparameter combination for both 'NoIndices' and 'With4Indice' model combinations.



a. Total RF grid fitting times

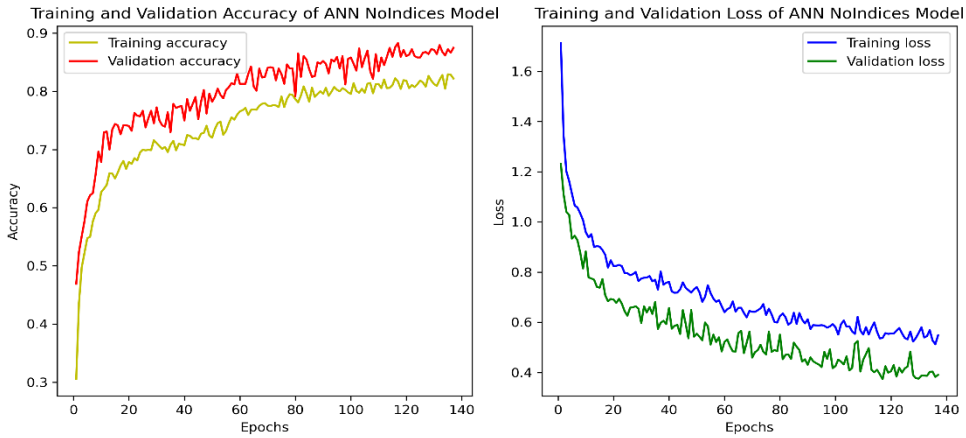
b. Indicvidual fitting time of RF models in grid with no indices data

c. Indicvidual fitting time of RF models in grid with 4 indices data

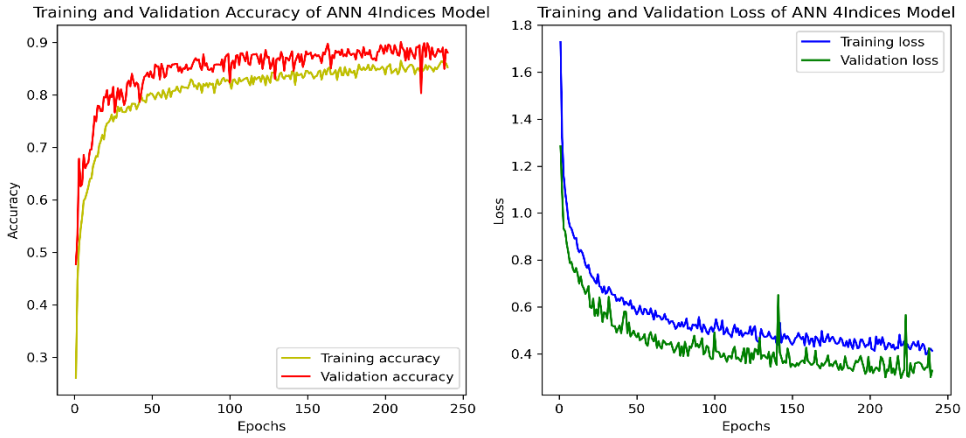
Figure 12: Time complexity of SVM models

In the graphs for individual iterations, we see that, all the models have gained both mean test accuracy and computation time with the addition of indices data with the training data. Changes in fitting time range from 0.026% to 0.23% of the training time with 'NoIndices' dataset.

The learning curves from the ANN models showed that, the inclusion of the indices data increases the training time for the model fitting as the training with no data takes 138 epochs, while the same architecture (only difference being the training data) with the indices data takes 241 epochs to reach the optimal state of validation loss. This also infers that; additional indices increased the training time for ANN model. The optimal state of validation was considered by 20 consecutive epochs with no reduction in validation loss. However, the generalization was better demonstrated in the ANNWith4Indices model as the validation accuracy is consistently higher and the loss is lower for the this for the epochs of this model. The learning curves in the below figure 10 also shows lower differences between training and validation lines for the model With4Indices signifying better convergence.



a. ANN accuracy and loss curve with no indices data



b. ANN accuracy and loss curve wit 4 indices data

Figure 13: ANN learning curves

A combined approach to find the optimal classification algorithm for the

One of the objectives of the study was to find most accurate algorithm for producing the LULC map. However, while choosing the algorithm with best OA, it has been seen that, some algorithm identifies specific classes better than the algorithm with best OA. Therefore, a new approach considering the best class wise performances among all the models has been trialled. It has been found that, the new approach overtrumps all other algorithms. The following graph demonstrates the confusion matrix for the proposed new method.

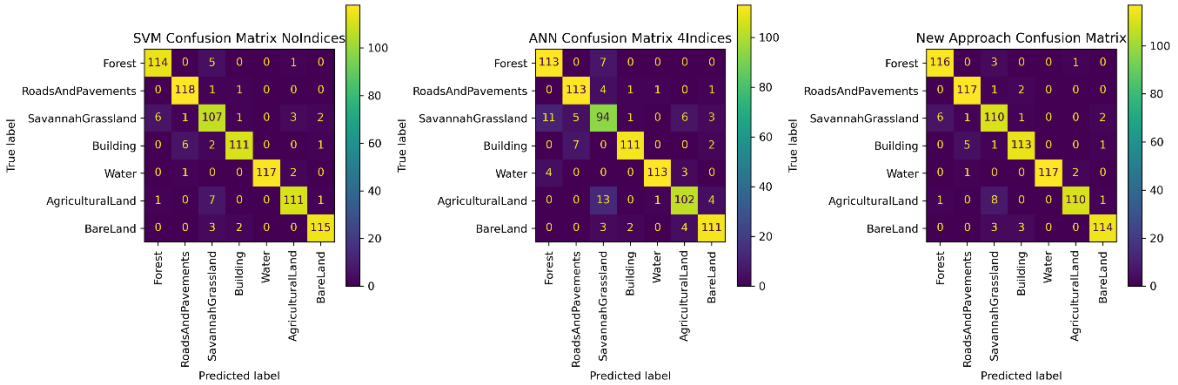


Figure 14: Comparison of correct prediction done by 2 best models(previous) and proposed new approach

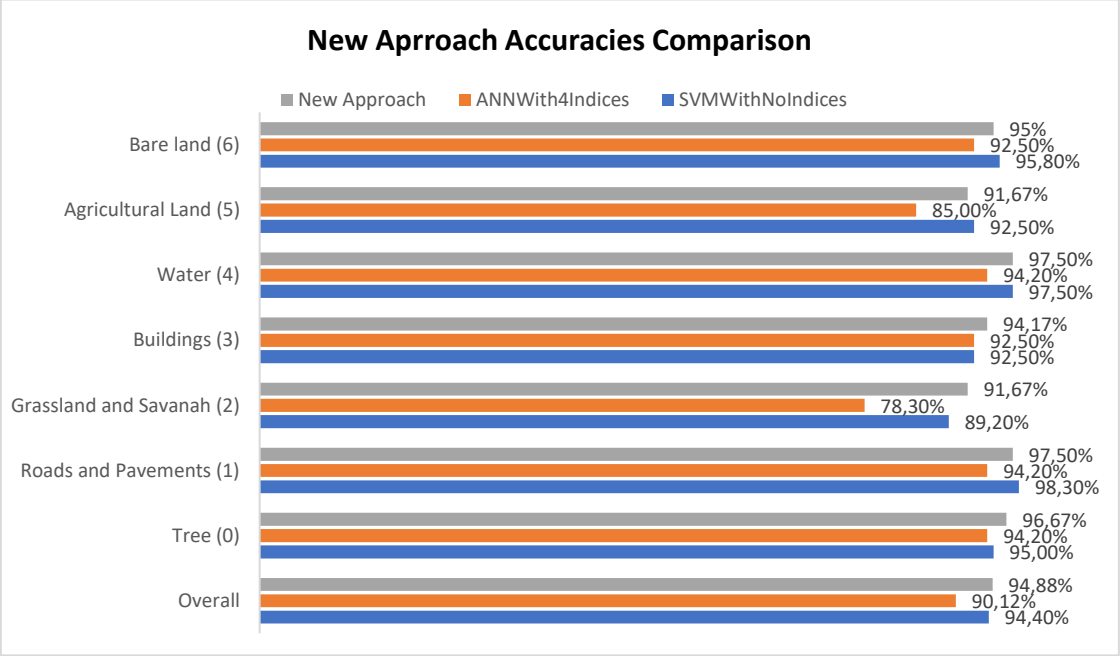
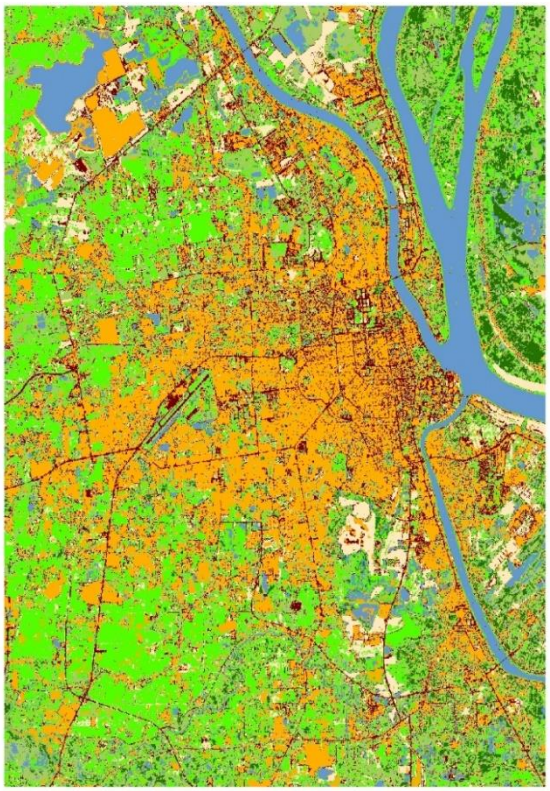
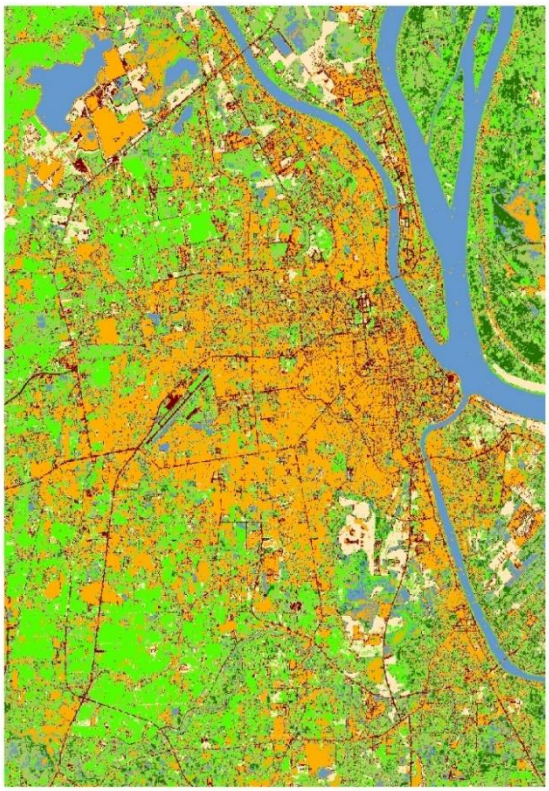


Figure 15: Comparison of class wise and overall accuracies of 2 best models(previous) and proposed new approach.

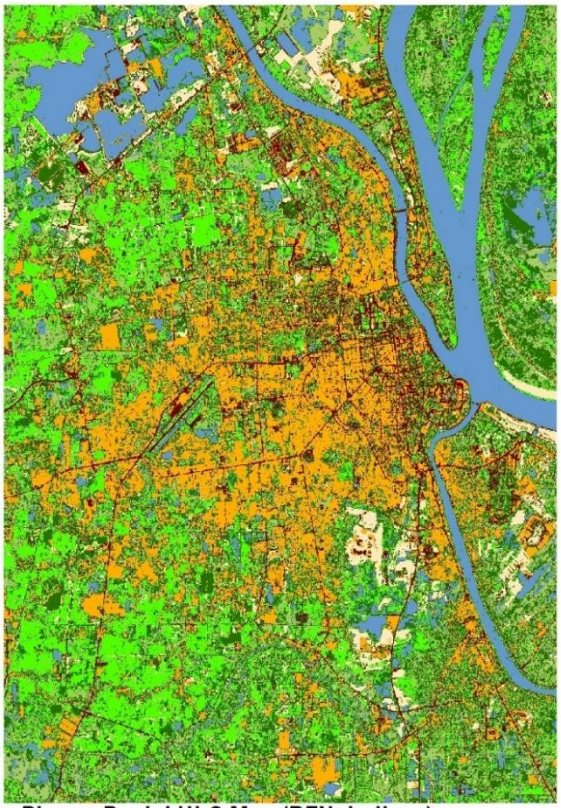
The objective of the study also seeks to produce updated LULC map for the area of interest. Using each of these approaches corresponding classified raster images were also generated for the area of interest. The following figure 16 demonstrates the classified images for area using all the experimented algorithms.



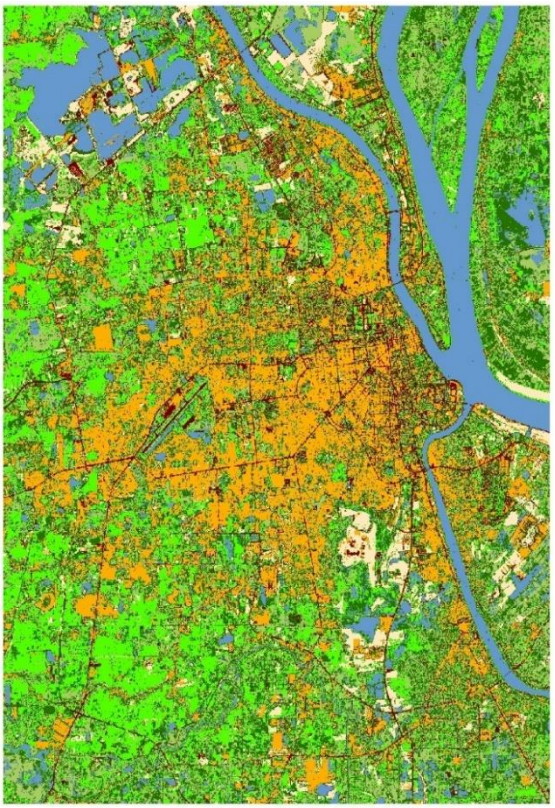
Phnom Penh LULC Map (SVMNoIndices)



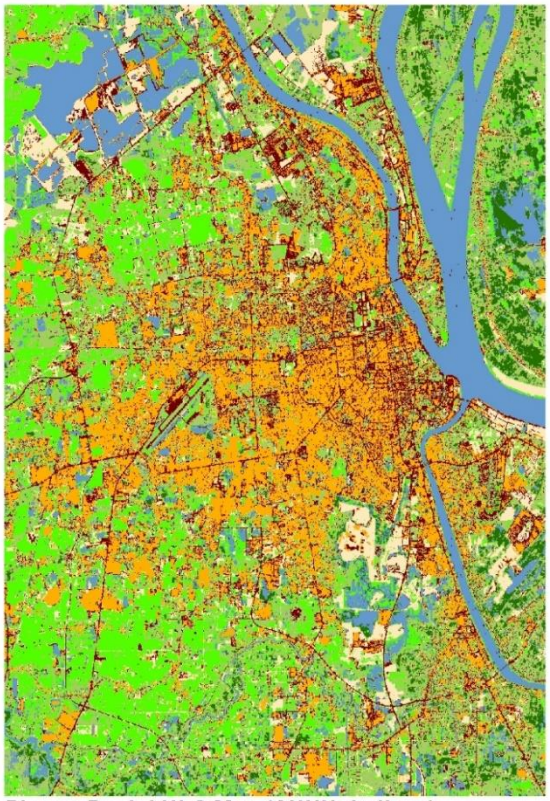
Phnom Penh LULC Map (SVMWith4Indices)



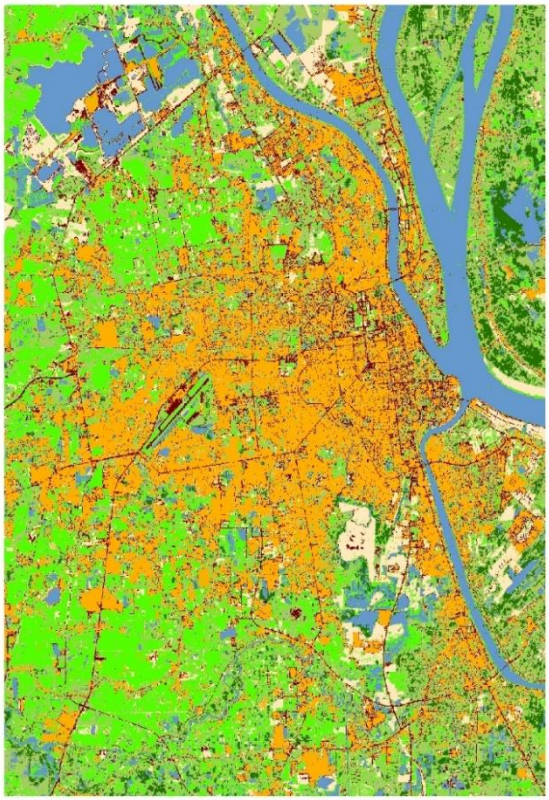
Phnom Penh LULC Map (RFNoIndices)



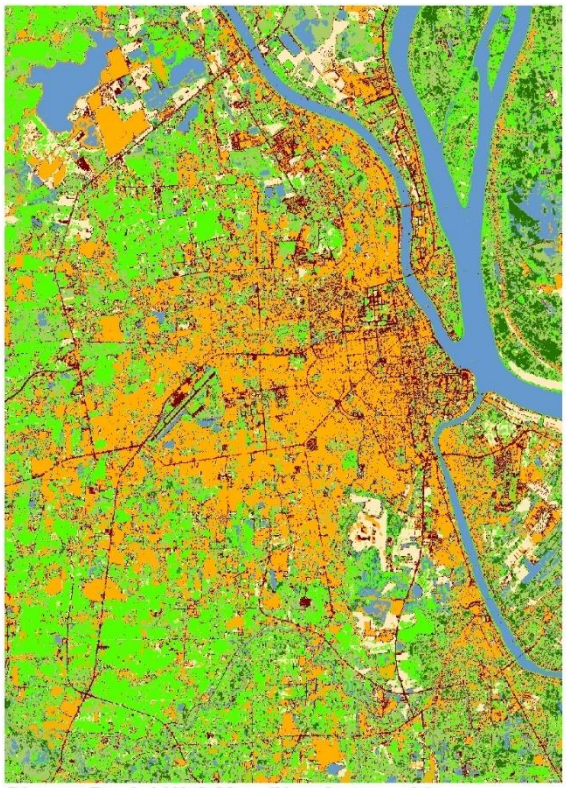
Phnom Penh LULC Map (RFWith4Indices)



Phnom Penh LULC Map (ANNNoIndices)



Phnom Penh LULC Map (ANNWith4Indices)



Phnom Penh LULC Map (NewApproach)

Data Source: Sentinel-2B
Product ID: S2B_MSIL2A_20230306T031619_
N0509_R118_T48PVT_
20230306T070117

N
W E
S

Legend

- Tree
- RoadsAndPavements
- SavannahGrassland
- Building
- Water
- AgriculturalLand
- BareLand

Figure 16: Classified Outputs from trialled models

References

Aryal, J., Sitaula, C., & Aryal, S. (2022). NDVI Threshold-Based Urban Green Space Mapping from Sentinel-2A at the Local Governmental Area (LGA) Level of Victoria, Australia. Land, 11(3), 351. https://doi.org/10.3390/land11030351

Bhandari, A. K., Kumar, A., & Singh, G. K. (2012). Feature Extraction using Normalized Difference Vegetation Index (NDVI): A Case Study of Jabalpur City. Procedia Technology, 6, 612–621. https://doi.org/10.1016/j.protcy.2012.10.074

Boyle, S. A., Kennedy, C. M., Torres, J., Colman, K., Pérez-Estigarribia, P. E., & De La Sancha, N. U. (2014). High-Resolution Satellite Imagery Is an Important yet Underutilized Resource in Conservation Biology. PLoS ONE, 9(1), e86908. https://doi.org/10.1371/journal.pone.0086908

Burke, M., Driscoll, A., Lobell, D. B., & Ermon, S. (2021). Using satellite imagery to understand and promote sustainable development. Science, 371(6535), eabe8628. https://doi.org/10.1126/science.abe8628

Chen, X.-L., Zhao, H.-M., Li, P.-X., & Yin, Z.-Y. (2006). Remote sensing image-based analysis of the relationship between urban heat island and land use/cover changes. Remote Sensing of Environment, 104(2), 133–146. https://doi.org/10.1016/j.rse.2005.11.016

Cheng, G., Xie, X., Han, J., Guo, L., & Xia, G.-S. (2020). Remote Sensing Image Scene Classification Meets Deep Learning: Challenges, Methods, Benchmarks, and Opportunities. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 13, 3735–3756. https://doi.org/10.1109/JSTARS.2020.3005403

Claverie, M., Ju, J., Masek, J. G., Dungan, J. L., Vermote, E. F., Roger, J.-C., Skakun, S. V., & Justice, C. (2018). The Harmonized Landsat and Sentinel-2 surface reflectance data set. Remote Sensing of Environment, 219, 145–161. https://doi.org/10.1016/j.rse.2018.09.002

De Luca, G., M. N. Silva, J., Di Fazio, S., & Modica, G. (2022). Integrated use of Sentinel-1 and Sentinel-2 data and open-source machine learning algorithms for land cover mapping in a Mediterranean region. European Journal of Remote Sensing, 55(1), 52–70. https://doi.org/10.1080/22797254.2021.2018667

E. D. Chaves, M., C. A. Picoli, M., & D. Sanches, I. (2020). Recent Applications of Landsat 8/OLI and Sentinel-2/MSI for Land Use and Land Cover Mapping: A Systematic Review. Remote Sensing, 12(18), 3062. https://doi.org/10.3390/rs12183062

Hossain, M. D., & Chen, D. (2019). Segmentation for Object-Based Image Analysis (OBIA): A review of algorithms and challenges from remote sensing perspective. ISPRS Journal of Photogrammetry and Remote Sensing, 150, 115–134. https://doi.org/10.1016/j.isprsjprs.2019.02.009

Hosseiny, B., Abdi, A. M., & Jamali, S. (2022). Urban land use and land cover classification with interpretable machine learning – A case study using Sentinel-2 and auxiliary data. Remote Sensing Applications: Society and Environment, 28, 100843. https://doi.org/10.1016/j.rsase.2022.100843

Hu, S., & Wang, L. (2013). Automated urban land-use classification with remote sensing. International Journal of Remote Sensing, 34(3), 790–803. https://doi.org/10.1080/01431161.2012.714510

Ihuoma, S. O., & Madramootoo, C. A. (2019). Sensitivity of spectral vegetation indices for monitoring water stress in tomato plants. Computers and Electronics in Agriculture, 163, 104860. https://doi.org/10.1016/j.compag.2019.104860

Jia, K., Wei, X., Gu, X., Yao, Y., Xie, X., & Li, B. (2014). Land cover classification using Landsat 8 Operational Land Imager data in Beijing, China. Geocarto International, 29(8), 941–951. https://doi.org/10.1080/10106049.2014.894586

Jia, Z., & Qin, A. (2022). Pixel-based versus object-based identification of scenic resources using Gaofen-2 images: A case study of Yesanpo National Park. PLOS ONE, 17(4), e0267435. https://doi.org/10.1371/journal.pone.0267435

Jin, Y., Liu, X., Chen, Y., & Liang, X. (2018). Land-cover mapping using Random Forest classification and incorporating NDVI time-series and texture: A case study of central Shandong. International Journal of Remote Sensing, 39(23), 8703–8723. https://doi.org/10.1080/01431161.2018.1490976

Ju, Z., Leong Tan, M., GeoInformatic Unit, Geography Section, School of Humanities, Universiti Sains Malaysia, 11800 USM, Penang, Malaysia, Samat, N., GeoInformatic Unit, Geography Section, School of Humanities, Universiti Sains Malaysia, 11800 USM, Penang, Malaysia, Kiat Chang, C., & River Engineering and Urban Drainage Research Centre (REDAC), Universiti Sains Malaysia, Engineering Campus, Nibong Tebal 14300, Penang, Malaysia. (2021). Comparison of Landsat 8, Sentinel-2 and spectral indices combinations for Google Earth Engine-based land use mapping in the Johor River Basin, Malaysia. Malaysian Journal of Society and Space, 17(3). https://doi.org/10.17576/geo-2021-1703-03

Jung, M., Henkel, K., Herold, M., & Churkina, G. (2006). Exploiting synergies of global land cover products for carbon cycle modeling. Remote Sensing of Environment, 101(4), 534–553. https://doi.org/10.1016/j.rse.2006.01.020

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. Communications of the ACM, 60(6), 84–90. https://doi.org/10.1145/3065386

Kuc, G., & Chormański, J. (2019). SENTINEL-2 IMAGERY FOR MAPPING AND MONITORING IMPERVIOUSNESS IN URBAN AREAS. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLII-1/W2, 43–47. https://doi.org/10.5194/isprs-archives-XLII-1-W2-43-2019

Kumar, L., & Mutanga, O. (2018). Google Earth Engine Applications Since Inception: Usage, Trends, and Potential. Remote Sensing, 10(10), 1509. https://doi.org/10.3390/rs10101509

Li, C., Shao, Z., Zhang, L., Huang, X., & Zhang, M. (2021). A Comparative Analysis of Index-Based Methods for Impervious Surface Mapping Using Multiseasonal Sentinel-2 Satellite Data. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 14, 3682–3694. https://doi.org/10.1109/JSTARS.2021.3067325

Li, H., Wang, C., Zhong, C., Zhang, Z., & Liu, Q. (2017). Mapping Typical Urban LULC from Landsat Imagery without Training Samples or Self-Defined Parameters. Remote Sensing, 9(7), 700. https://doi.org/10.3390/rs9070700

Liashchynskyi, P., & Liashchynskyi, P. (2019). Grid Search, Random Search, Genetic Algorithm: A Big Comparison for NAS. https://doi.org/10.48550/ARXIV.1912.06059

Liu, S., Jin, X., Nie, C., Wang, S., Yu, X., Cheng, M., Shao, M., Wang, Z., Tuohuti, N., Bai, Y., & Liu, Y. (2021). Estimating leaf area index using unmanned aerial vehicle data: Shallow vs. deep machine learning algorithms. Plant Physiology, 187(3), 1551–1576. https://doi.org/10.1093/plphys/kiab322

Macarringue, L. S., Bolfe, É. L., & Pereira, P. R. M. (2022). Developments in Land Use and Land Cover Classification Techniques in Remote Sensing: A Review. Journal of Geographic Information System, 14(01), 1–28. https://doi.org/10.4236/jgis.2022.141001

Madasa, A., Orimoloye, I. R., & Ololade, O. O. (2021). Application of geospatial indices for mapping land cover/use change detection in a mining area. Journal of African Earth Sciences, 175, 104108. https://doi.org/10.1016/j.jafrearsci.2021.104108

McCallum, I., Obersteiner, M., Nilsson, S., & Shvidenko, A. (2006). A spatial comparison of four satellite derived 1 km global land cover datasets. International Journal of Applied Earth Observation and Geoinformation, 8(4), 246–255. https://doi.org/10.1016/j.jag.2005.12.002

McFEETERS, S. K. (1996). The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. International Journal of Remote Sensing, 17(7), 1425–1432. https://doi.org/10.1080/01431169608948714

Mitchell, R. B. (2003). I NTERNATIONAL E NVIRONMENTAL A GREEMENTS: A Survey of Their Features, Formation, and Effects. Annual Review of Environment and Resources, 28(1), 429–461. https://doi.org/10.1146/annurev.energy.28.050302.105603

Mohajane, M., Essahlaoui, A., Oudija, F., El Hafyani, M., Hmaidi, A. E., El Ouali, A., Randazzo, G., & Teodoro, A. C. (2018). Land Use/Land Cover (LULC) Using Landsat Data Series (MSS, TM, ETM+ and OLI) in Azrou Forest, in the Central Middle Atlas of Morocco. Environments, 5(12), 131. https://doi.org/10.3390/environments5120131

Nasiri, V., Deljouei, A., Moradi, F., Sadeghi, S. M. M., & Borz, S. A. (2022). Land Use and Land Cover Mapping Using Sentinel-2, Landsat-8 Satellite Images, and Google Earth Engine: A Comparison of Two Composition Methods. Remote Sensing, 14(9), 1977. https://doi.org/10.3390/rs14091977

Nguyen, C. T., Chidthaisong, A., Kieu Diem, P., & Huo, L.-Z. (2021). A Modified Bare Soil Index to Identify Bare Land Features during Agricultural Fallow-Period in Southeast Asia Using Landsat 8. Land, 10(3), 231. https://doi.org/10.3390/land10030231

Nguyen, K.-A., & Liou, Y.-A. (2019). Global mapping of eco-environmental vulnerability from human and nature disturbances. Science of The Total Environment, 664, 995–1004. https://doi.org/10.1016/j.scitotenv.2019.01.407

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research, 12(85), 2825–2830.

Pheakdey, D. V., Quan, N. V., & Xuan, T. D. (2023). Economic and Environmental Benefits of Energy Recovery from Municipal Solid Waste in Phnom Penh Municipality, Cambodia. Energies, 16(7), 3234. https://doi.org/10.3390/en16073234

Phnom Penh Capital Hall. (2023). Land Area and Population. Phnom Penh Capital Hall. https://phnompenh.gov.kh/en/phnom-penh-city/facts/

Rumora, L., Miler, M., & Medak, D. (2020). Impact of Various Atmospheric Corrections on Sentinel-2 Land Cover Classification Accuracy Using Machine Learning Classifiers. ISPRS International Journal of Geo-Information, 9(4), 277. https://doi.org/10.3390/ijgi9040277

S Malini, Manjunatha., & Patil, M. (2018). Interpolation Techniques in Image Resampling. International Journal of Engineering & Technology, 7(3.34), 567. https://doi.org/10.14419/ijet.v7i3.34.19383

Sun, Y., Zhang, X., Zhao, Y., & Xin, Q. (2017). Monitoring annual urbanization activities in Guangzhou using Landsat images (1987–2015). International Journal of Remote Sensing, 38(5), 1258–1276. https://doi.org/10.1080/01431161.2016.1268283

Talukdar, S., Singha, P., Mahato, S., Shahfahad, Pal, S., Liou, Y.-A., & Rahman, A. (2020). Land-Use Land-Cover Classification by Machine Learning Classifiers for Satellite Observations—A Review. Remote Sensing, 12(7), 1135. https://doi.org/10.3390/rs12071135

Tassi, A., & Vizzari, M. (2020). Object-Oriented LULC Classification in Google Earth Engine Combining SNIC, GLCM, and Machine Learning Algorithms. Remote Sensing, 12(22), 3776. https://doi.org/10.3390/rs12223776

Thanh Noi, P., & Kappas, M. (2017). Comparison of Random Forest, k-Nearest Neighbor, and Support Vector Machine Classifiers for Land Cover Classification Using Sentinel-2 Imagery. Sensors, 18(2), 18. https://doi.org/10.3390/s18010018

Thanh Son, N., Thi Thu Trang, N., Bui, X. T., & Thi Da, C. (2022). Remote sensing and GIS for urbanization and flood risk assessment in Phnom Penh, Cambodia. Geocarto International, 37(22), 6625–6642. https://doi.org/10.1080/10106049.2021.1941307

User Guides—Sentinel-2 MSI - Level-2 Processing—Sentinel Online. (n.d.). Sentinel Online. Retrieved 17 July 2023, from https://copernicus.eu/user-guides/sentinel-2-msi/processing-levels/level-2

Wang, N. (2022). The use of bilinear interpolation filter to remove image noise. Journal of Physics: Conference Series, 2303(1), 012089. https://doi.org/10.1088/1742-6596/2303/1/012089

Woldemariam, G. W., Tibebe, D., Mengesha, T. E., & Gelete, T. B. (2022). Machine-learning algorithms for land use dynamics in Lake Haramaya Watershed, Ethiopia. Modeling Earth Systems and Environment, 8(3), 3719–3736. https://doi.org/10.1007/s40808-021-01296-0

Xian, G., Shi, H., Dewitz, J., & Wu, Z. (2019). Performances of WorldView 3, Sentinel 2, and Landsat 8 data in mapping impervious surface. Remote Sensing Applications: Society and Environment, 15, 100246. https://doi.org/10.1016/j.rsase.2019.100246

Yan, Z., Ma, L., He, W., Zhou, L., Lu, H., Liu, G., & Huang, G. (2022). Comparing Object-Based and Pixel-Based Methods for Local Climate Zones Mapping with Multi-Source Data. Remote Sensing, 14(15), 3744. https://doi.org/10.3390/rs14153744

Yousefi, S., Mirzaee, S., Almohamad, H., Al Dughairi, A. A., Gomez, C., Siamian, N., Alrasheedi, M., & Abdo, H. G. (2022). Image Classification and Land Cover Mapping Using Sentinel-2 Imagery: Optimization of SVM Parameters. Land, 11(7), 993. https://doi.org/10.3390/land11070993

Yuh, Y. G., Tracz, W., Matthews, H. D., & Turner, S. E. (2023). Application of machine learning approaches for land cover monitoring in northern Cameroon. Ecological Informatics, 74, 101955. https://doi.org/10.1016/j.ecoinf.2022.101955

Zekoll, V., Main-Knorn, M., Alonso, K., Louis, J., Frantz, D., Richter, R., & Pflug, B. (2021). Comparison of Masking Algorithms for Sentinel-2 Imagery. Remote Sensing, 13(1), 137. https://doi.org/10.3390/rs13010137

Zhang, Y., Skakun, S., Adegbenro, M. O., & Ying, Q. (2022). Leveraging the use of labeled benchmark datasets for urban area change mapping and area estimation: A case study of the Washington DC–Baltimore region. International Journal of Digital Earth, 15(1), 1169–1186. https://doi.org/10.1080/17538947.2022.2094001