



Hochschule für nachhaltige Entwicklung Eberswalde

Research Project Report

On

Significance Analysis of Normalized Indices in Machine Learning-Based Land Use and Land Class Classification in an Urban Landscape

Submitted by

Kazi Jahidur Rahaman

Matrikelnummer 21213984

Supervise by

Professor Dr. Jan-Peter Mund

Professor for GIS and Remote Sensing

Co-Head of the M.Sc. Study Programme Forest Information Technology (FIT)

Date of Submission

23 February 2024

Winter Semester 2023/2024

Contents

1	Introduction.....	5
1.1	Indices based on LULC classification	5
1.2	ML and DL-based LULC classification	6
1.2.1	Support Vector Machine.....	7
1.2.2	Artificial Neural Network	7
2	Aim and Objectives.....	8
2.1	Aim	8
2.2	Specific Objectives.....	8
2.3	Research Questions	8
2.4	Materials and Methods	8
2.5	Study Area	8
2.6	Data.....	9
3	Methodology	12
3.1	Data Acquisition.....	12
3.2	Data Preprocessing	13
3.2.1	Cloud Masking and Gap Filling.....	13
3.2.2	Resampling.....	14
3.2.3	Band selection	14
3.3	Indices Generation.....	14
3.4	Defining Training and Testing Data	15
3.5	Model Construction and Training.....	15
3.5.1	Support Vector Machine.....	15
3.5.2	Artificial Neural Network	16
3.6	Model Training and Result Analysis	17
3.7	Hybrid Selection Approach	18
4	Results.....	19
4.1	Proposed Hybrid Selection Approach.....	22
5	Discussion	26
6	Conclusion	28
7	References.....	29

Figure 1: Classification through hyperplane in SVM('Understanding Support Vector Machines', 2017)	7
Figure 2: Use of conventional NN algorithms in RS data(Jiang et al., 2020).....	7
Figure 3: Study area and location	9
Figure 4: Sample from train data	12
Figure 5: A generalized representation of the workflow	12
Figure 6: Cloud masking of the selected image tile.....	13
Figure 7: Histogram of the selected bands.....	14
Figure 8: SVM hyperparameter grid for tuning	16
Figure 9: ANN Model Structures	17
Figure 10: Calculation Parameters of Accuracy Metrics(F1 Score in Machine Learning, n.d.)	18
Figure 11: Detailed Methodology of the study	19
Figure 12: Algorithms Performance.....	20
Figure 13: Training and Validation Loss Curves for ANN Models. Left top and bottom (ANN No Indices accuracy and loss curve), Right top and bottom ANN with 4 indices accuracy and loss curves.	21
Figure 14: Class wise predictions by the algorithms.	22
Figure 15: Confusion matrix for proposed selection approach.	22
Figure 16: Model fitting time for different hyperparameter combinations and their accuracy with and without indices data.	23
Figure 17: Classified Outputs from trialled SVM models	24
Figure 18: Classified Outputs from trialled ANN models	25
Figure 19: Classified Outputs from proposed hybrid selection model.....	26
Figure 20: Correlation among the bands.....	27

Abstract: Urbanization and industrialization have been rapidly transforming land use and cover, impacting various aspects of the environment. To address resulting challenges, up-to-date land use and land cover (LULC) mapping is crucial. Recent developments in satellite remote sensing, including advanced satellite sensors and drones, offer higher resolutions and spectral bands. The current state-of-the-art involves utilizing machine learning (ML) and deep learning (DL) algorithms to automate land cover mapping, facilitating faster and more accurate analyses of remotely sensed data. The proven effectiveness of spectral indices like NDVI, NDBI, NDWI, and BSI in assessing diverse features such as vegetation, built-up, water, and open surfaces has prompted researchers to leverage their potential in ML-based multiclass LULC classification. This is done by incorporating the indices with the regular satellite bands. However, the significance of this incorporation has yet been understudied. Does the inclusion improve the performance of ML algorithms? This research focuses on empirically investigating the impact of integrating spectral indices with satellite bands in LULC classification, considering both classification accuracy and computation time complexity.

The research applied representative algorithms of discriminative, and neural network categories using solely satellite bands and satellite bands in conjunction with spectral indices for Phnom Penh city in Cambodia. The findings indicated that, for the specific urban area, Artificial Neural Network (ANN) and Support Vector Machine (SVM) models trained without indices exhibited higher overall accuracies (88.25% and 88%, respectively) compared to one trained with indices. However, the inclusion of indices significantly enhanced model training times in discriminative and while increasing fitting epoch time in the ANN. Additionally, the study generated an updated LULC map for the year 2023 in the specified area.

The research makes a valuable contribution to assessing the feasibility of integrating index-based data into machine learning models for enhanced LULC classification. Additionally, it provides insights for future studies to carefully choose ML/DL algorithms by weighing the trade-offs between class-wise accuracies, overall accuracies, and computational complexities. Although the study focused on four specific indices, the same methodology can be applied to explore the significance of a broader range of indices and auxiliary data in similar investigations in LULC mapping and change detection.

Keywords: Artificial Neural Network; Support Vector Machine; Land Use Land Cover Mapping; Normalized Indices; Google Earth Engine

1 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 computation time complexity. This approach contributes to future research in determining the suitability of incorporating index-based data into machine learning models for improved LULC classification.

Rapid urbanization and industrialization have accelerated the 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.

1.1 Indices based on LULC classification

In the literature satellite images have been leveraged in mostly 2 ways, either to identify single land cover using index-based approaches or using 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).

1.2 ML and DL-based LULC classification

At its core, the essence of machine learning algorithms lies in their ability to recognize patterns within entirely new data, either 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 the 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 literature both shallow and DL approaches were found in the field of RS classification, where the shallow algorithms dominantly used a 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 conditions and Middle Eastern warm conditions 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 a humid palm tree dominated area in Malaysia but the overall accuracy was 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 focus on detecting specific types of vegetation in Sweeden. The result of adding the additional bands was 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 a 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 any 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. To comprehend the impact of incorporating spectral indices with satellite bands, we intend to apply the same algorithms to the multiple combinations of datasets of one area to ensure the same spatial, temporal, and atmospheric characteristics of the datasets and investigate the outcomes.

1.2.1 Support Vector Machine

Support Vector Machine (SVM) is a supervised machine learning algorithm used for classification and regression tasks. In the context of classification, SVM finds the hyperplane that best separates data points belonging to different classes. The optimal hyperplane is the one that maximizes the margin between classes, making it robust to outliers. SVM can handle both linear and non-linear relationships in the data by using different kernel functions.

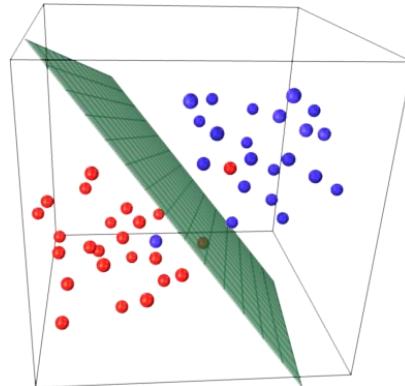


Figure 1: Classification through hyperplane in SVM('Understanding Support Vector Machines', 2017)

1.2.2 Artificial Neural Network

In summary, the Artificial Neural Network (ANN) is a computational model inspired by the human brain, with a Sequential model commonly used for its simplicity. It comprises layers, including an input layer, hidden layers, and an output layer. Neurons, the basic computation units, form dense layers with connections to every neuron in adjacent layers. Activation functions, like LeakyReLU and SoftMax, introduce non-linearity for learning complex patterns. The Sequential model facilitates the sequential addition of layers. ‘LeakyReLU’ provides flexibility in learning, while SoftMax is used in the output layer for multi-class classification. Optimizers adjust weights during training, and Dropout prevents overfitting by excluding random neurons. During training, early stopping helps prevent overfitting by terminating training when the model’s performance on the training data starts to deteriorate. Such structure empowers ANNs to learn intricate relationships, making them valuable across diverse applications.

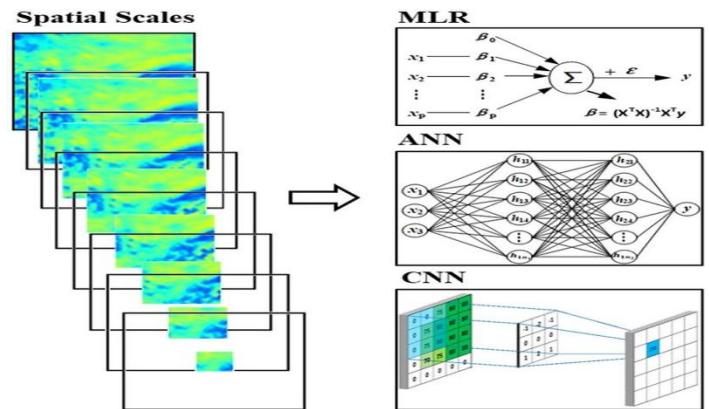


Figure 2: Use of conventional NN algorithms in RS data(Jiang et al., 2020)

2 Aim and Objectives

2.1 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.

2.2 Specific Objectives

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

2.3 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 different 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 outcomes for distinct land classes.

2.4 Materials and Methods

2.5 Study Area

The focus of the study was part of Phnom Penh, the capital city of Cambodia. Over the past two decades, Phnom Penh has experienced a 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 679Km² which hosts 2,281,951 people. (Pheakdey et al., 2023) The exact study area shown in Figure 3 includes Phnom Penh and its periphery.

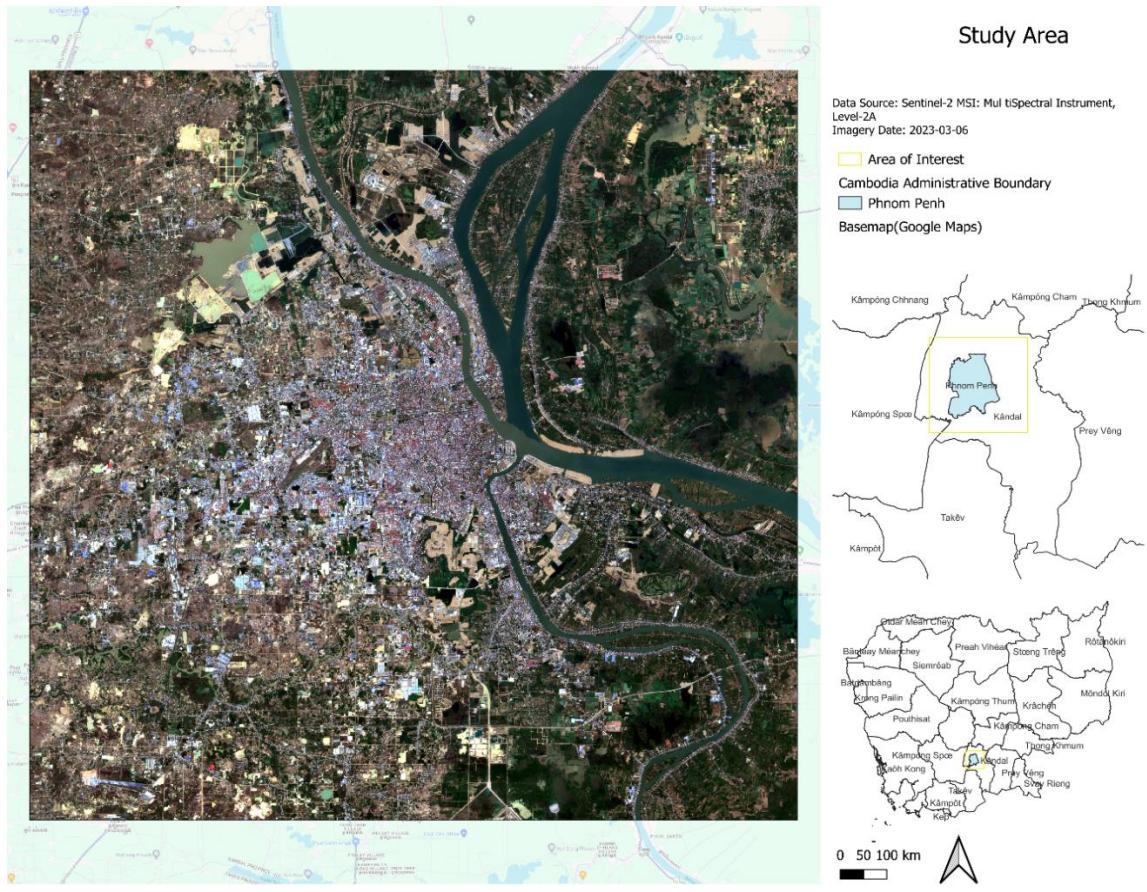


Figure 3: Study area and location

2.6 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 researchers from around the world for LULC works because of high resolution, spatial-temporal coverage, and easy availability on platforms like Google Earth Engine which also provides sophisticated computational power (Macarringué 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 900m² 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. Table 1 shows key information on the dataset used for the study.

Table 1: Information about the satellite data

Dataset Provider	European Union/ European Space Agency/ Copernicus Programme
SPACECRAFT NAME	Sentinel-2B
Data Accessing Medium	Google Earth Engine (GEE)
GEE Dataset Title	Sentinel-2 MSI: MultiSpectral Instrument, Level-2A
MGRS TILE	48PVT
Image ID	20230306T031619_20230306T033002_T48PVT
MEAN SOLAR AZIMUTH ANGLE	123.794241990723
MEAN SOLAR ZENITH ANGLE	29.7350215117703
SENSING ORBIT DIRECTION	DESCENDING
SENSING ORBIT NUMBER	118
Date Acquired	2023-03-06
CLOUDY PIXEL OVER LAND PERCENTAGE	1.190703
CLOUDY PIXEL PERCENTAGE	1.179621
DARK FEATURES PERCENTAGE	0.020201

These training samples were generated based on a prior GAF land use survey conducted in 2017. Nevertheless, given that one of the objectives of the study was to produce an 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 2324 visually confirmed samples of ‘Tree, Flooded, or Other Vegetation’, ‘Roads and Other Impervious Surfaces’, ‘Grassland’, ‘Building’, ‘Water’, ‘Agricultural Land, and ‘Bare Land, in an ESRI shape file.

In a similar process total of 800 testing samples were generated for testing purposes. These testing samples were not exposed to the training dataset. A short description of the classes can be found in Table 2.

Table 2: Training Data Description

Class (Numerical Class)	Description	Train Test Size
Tree, Flooded, or Other Vegetation (0)	Areas with dense vegetation, trees, urban greens, forests, or lowlands with both water and vegetation	(306, 117)
Roads and Other Impervious Surfaces (1)	Roads, highways, footpaths, parking spaces, and similar paved surfaces	(308, 104)
Grassland (2)	Open lands dominated by grasses, scattered trees, shrubs, unused real estate plots	(321, 117)
Building (3)	Human-made permanent residential, commercial, and industrial structures	(301, 131)
Water (4)	River, pond, canal, creek, post-mining opens flooded areas	(393, 114)
Agricultural Land (5)	Specifically used for farming and agricultural	(371, 116)
Bare Land (6)	Exposed soil, with no presence of vegetation	(317, 101)

The following plots in Figure 4 show 3 samples from each class.

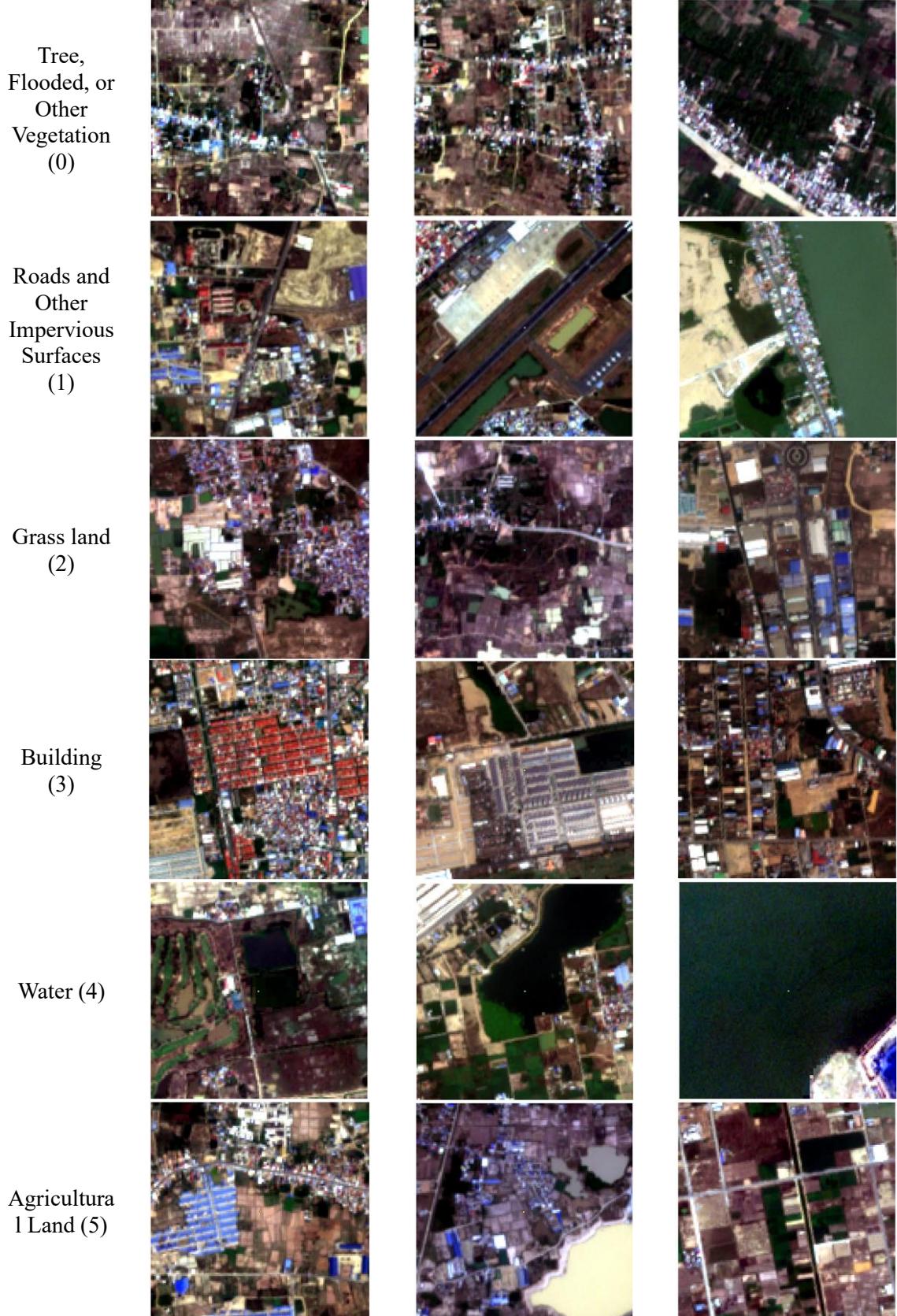




Figure 4: Sample from train data

Besides that, the AOI boundary was created using QGIS software. All the datasets 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, and 16 Gigabyte memory.

3 Methodology

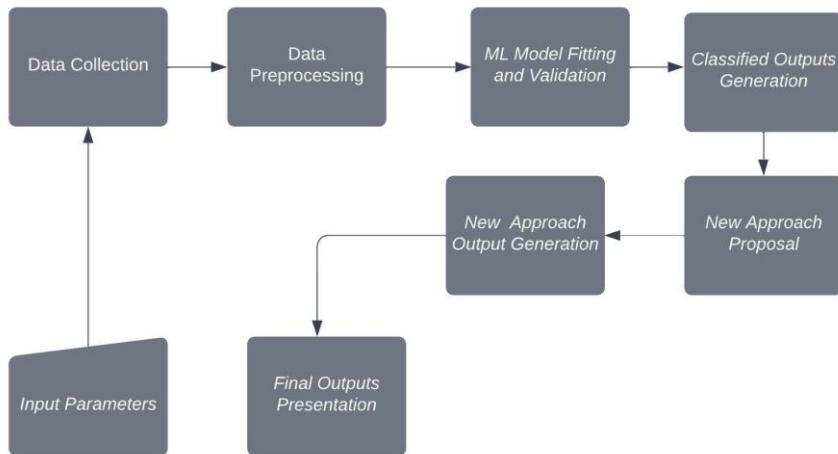


Figure 5: A generalized representation of the workflow

3.1 Data Acquisition

The sentinel 2 harmonized surface reflectance collection was derived from the GEE data catalog. GEE data catalog 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 therefore the filtering was done for 2023 from January to April. A cloud filter of 2% 'CLOUDY_PIXEL_PERCENTAGE' and 2% 'DARK_FEATURES_PERCENTAGE' was also imposed. The accumulated images were sorted by the cloud cover percentage and then the image with the best cloud cover was selected for further processing.

3.2 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.). Therefore, no further atmospheric and radiometric correction was needed.

3.2.1 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 3 below.

Table 3: 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 12 image median composite of the same timeframe with a cloud probability of 50%. The following figure 6 shows the cloudless, filtered cloud, and cloud filled images.

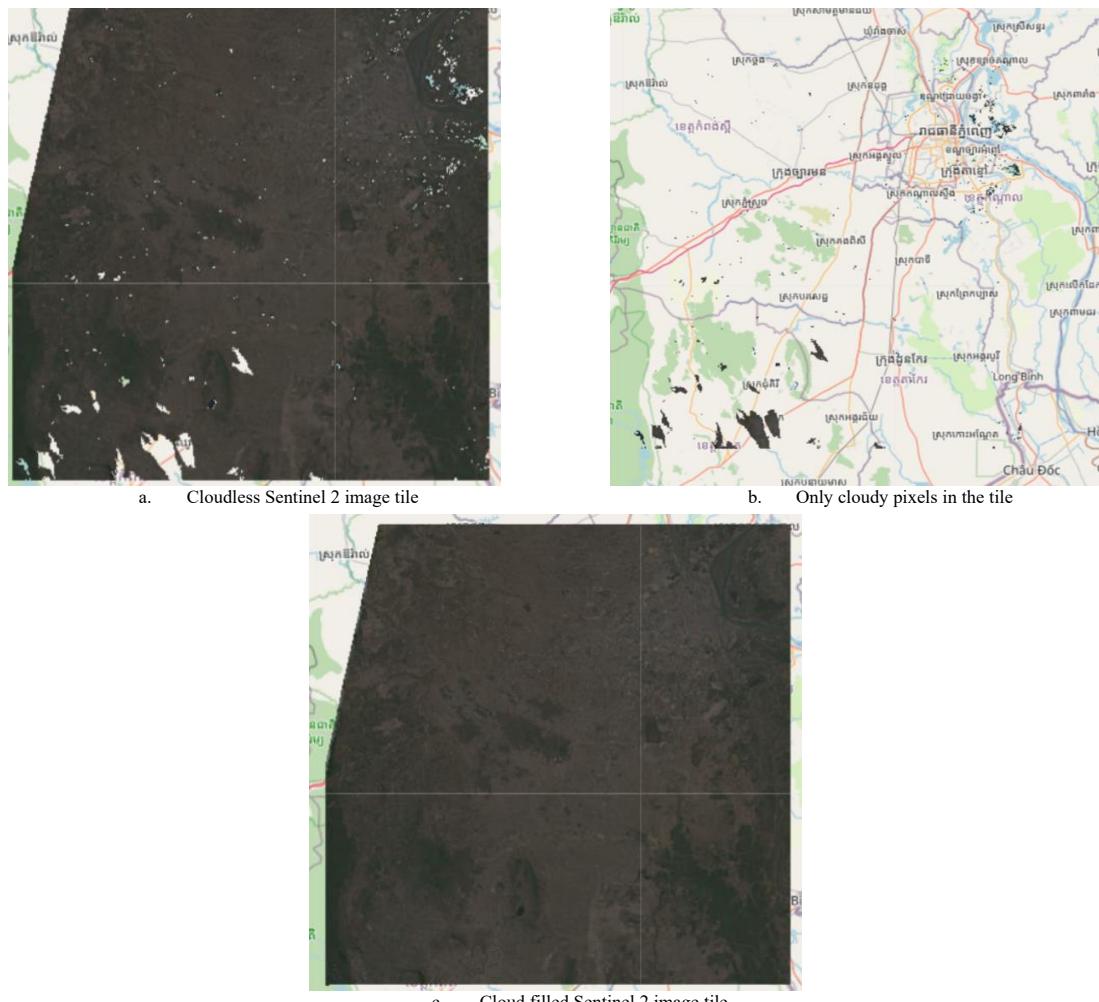


Figure 6: Cloud masking of the selected image tile

3.2.2 Resampling

Sentinel 2 dataset comes with different spatial resolutions 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 a 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)

3.2.3 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. These selected bands reflectance’s were then normalized to 0-1 scale. Figure 7 below shows the histograms of the normalized bands.

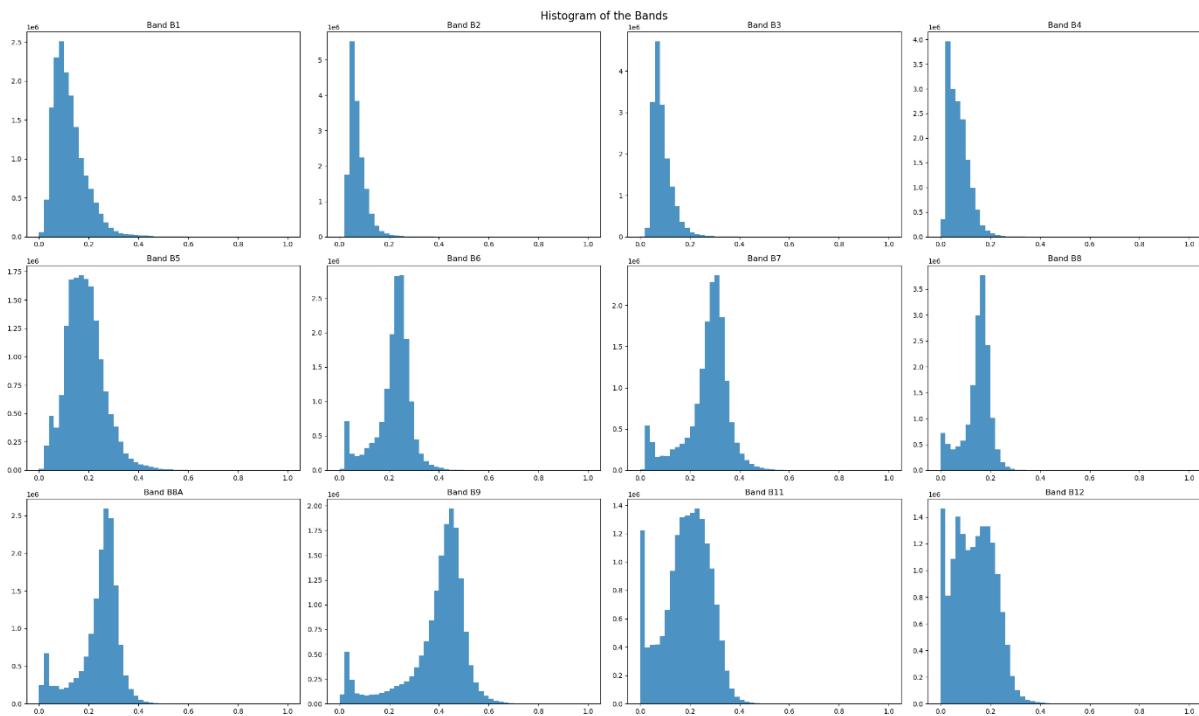


Figure 7: Histogram of the selected bands

3.3 Indices Generation

With the aim being to analyze 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 the characteristics of the study area, and decided to include NDVI, NDBI, NDWI, and BSI considering the relevance to the specific classes. From the literature it has been seen that these indices help identify 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 analyzing 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-

3.4 Defining Training and Testing Data

After the creation of indices, the sentinel 2 selected bands (Band 1-12) and corresponding indices were merged, and an image of 16 bands. At this phase, the pixel values for each band against each training sample were identified in the image, and pixel values were extracted. Thus, the final training dataset of 2324 samples was constructed where each training sample consisted of the corresponding spectral values from Sentinel 2 bands, NDVI, NDBI, NDWI, BSI, and the corresponding class. Similarly, the pixel values for the testing dataset were also extracted 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.

In the end, there were 2 different datasets: one with only 12 bands from Sentinel 2, and the other with 16 bands including the indices. The final shape of training and testing splits for both of those are summarized in Table 4 below.

Table 4: Training and testing data composition

Class (Numerical Class)	Train Shape		Test Shape	
	NoIndices	With4Indices	NoIndices	With4Indices
Tree, Flooded, or Other Vegetation (0)	(306, 12)	(306, 16)	(117, 12)	(117, 16)
Roads and Other Impervious Surfaces (1)	(308, 12)	(308, 16)	(104, 12)	(104, 16)
Grass land (2)	(321, 12)	(321, 16)	(117, 12)	(117, 16)
Building (3)	(301, 12)	(301, 16)	(131, 12)	(131, 16)
Water (4)	(393, 12)	(393, 16)	(114, 12)	(114, 16)
Agricultural Land (5)	(371, 12)	(371, 16)	(116, 12)	(116, 16)
Bare Land (6)	(317, 12)	(317, 16)	(101, 12)	(101, 16)

3.5 Model Construction and Training

3.5.1 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 grid of hyperparameter combinations. The remaining parameters were 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 for each of the dataset combinations.

```

GridSearchCV
GridSearchCV(cv=2, estimator=SVC(random_state=9),
            param_grid={'C': [2, 4], 'degree': [3, 4], 'gamma': [100, 200],
                        'kernel': ['poly', 'rbf']},
            verbose=3)

    +-- estimator: SVC
        SVC(random_state=9)
            +-- SVC
                SVC(random_state=9)

```

Figure 8: SVM hyperparameter grid for tuning

By 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 = 100, kernel = ('rbf')**], the optimal configuration for ‘With4Indices’ dataset was [**C = 2, 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.

3.5.2 Artificial Neural Network

The Artificial Neural Network (ANN) was implemented to construct 2 different models for ‘NoIndices’ in and ‘With4Indices’ dataset with structures shown in figure 9(a) and 9(b) respectively. 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 neural network model is structured with an input layer consisting of 12 neurons, each consuming to the number of features in the input data. Following the input layer are three hidden layers with 16, 32, and 64 neurons, respectively. Leaky Rectified Linear Unit (ReLU) activation functions with a small negative slope (alpha=0.3) are employed in the hidden layers, introducing non-linearity to the model(Xu et al., 2015). Additionally, dropout layers with a dropout rate of 20% are inserted between the hidden layers to mitigate overfitting by randomly excluding a fraction of input connections during training. The output layer is designed with SoftMax activation function as the task is a multi-class classification. It converts raw output scores into probability distributions across different classes. The model is

compiled using categorical cross entropy as the loss function, AdamW as the optimizer with a learning rate of 0.015, and accuracy as the metric(Loshchilov & Hutter, 2017). This architecture is designed to keep a balance between model complexity and the prevention of overfitting. Selection of hyperparameters, such as the number of neurons and layers, dropout rate, and learning rate, and slope value was made based on the trial-and-error basis in this case.

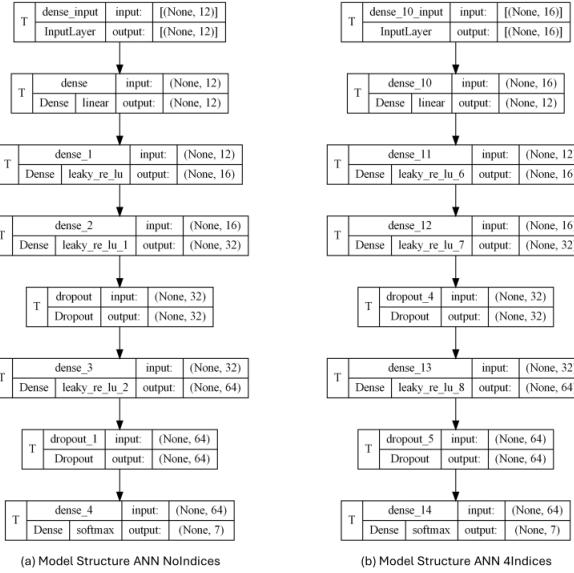


Figure 9: ANN Model Structures

The model is let to train for a maximum of 1000 epochs, with a batch size of 64. 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 300 consecutive of epochs.

3.6 Model Training and Result Analysis

After constructing the models and finding the optimal hyperparameters the models were trained individually for both of the training data combinations. During hyper parameter training the best models were picked based on the best cross validation accuracies. Then, with that model the test dataset which was not exposed to the model previously was used to predict and analyse the result. For evaluating the performance of the models' class wise and weighted Precision (also known as user accuracy), Recall (also known as producer accuracy), F1-Score and Overall accuracies and confusion metrices has been generated. For calculation the following equations has been used.

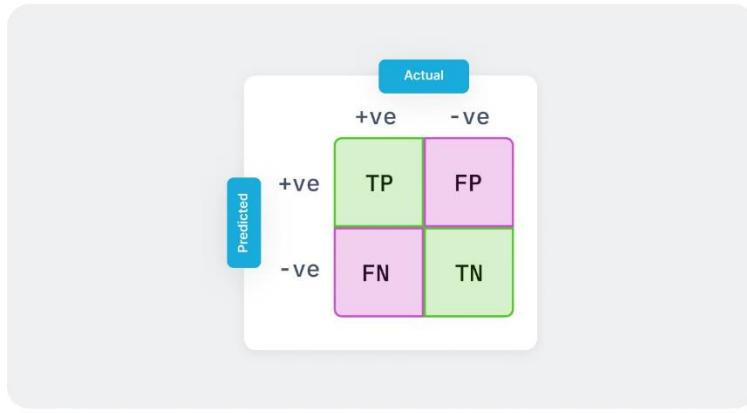
$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

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

$$\text{Accuracy} = \frac{\text{TP} * \text{TN}}{\text{TP} + \text{FN} + \text{TN} + \text{FP}}$$

$$\text{F1} = \frac{2 * \text{TP}}{2 * \text{TP} + \text{FP} + \text{FN}}$$

Where, TP = True Positive, FP = False Positive, FN = False Negative, TN = True Negative. The relationship between these values can be realized from the following diagram in figure 10.



v7

Figure 10: Calculation Parameters of Accuracy Metrics(F1 Score in Machine Learning, n.d.)

3.7 Hybrid Selection Approach

After fitting SVM, and ANN model with both with indices and without indices dataset, it was seen that different models are performing better in detecting particular classes but not giving a stable result for all the classes. The idea of the proposed hybrid selection approach is to combine the predictions from best performing models for each class.

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 F1-Score calculated by equation mentioned above. and stored in a matrix in all models. The matrix was later sorted to set the hierarchy among the classes. Based on 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 F1-Score.

The figure 11 represents the summary of the methodology of the study.

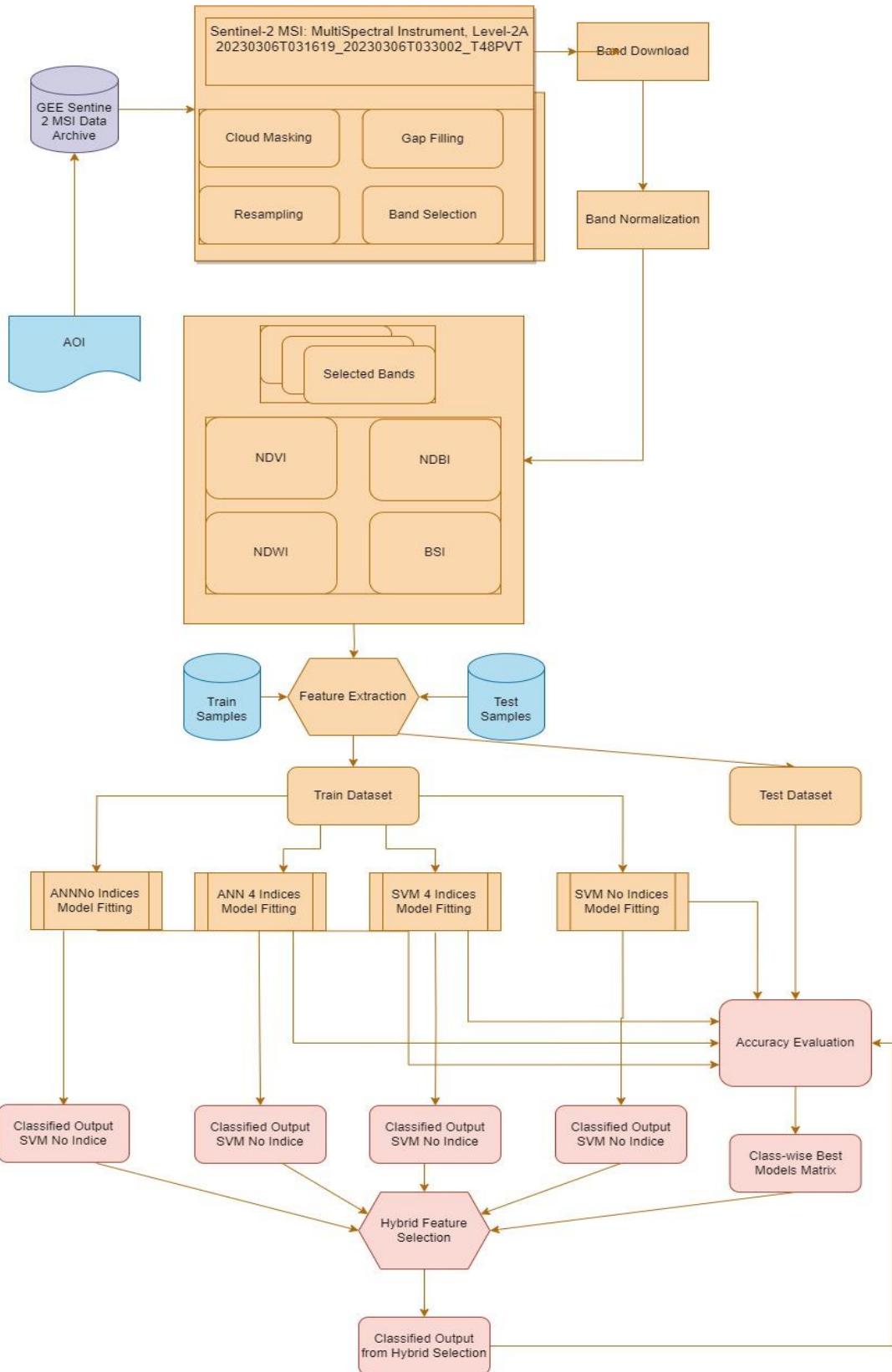


Figure 11: Detailed Methodology of the study

4 Results

For the performance evaluation total 800 samples were collected in the study area extent using random sampling method. These samples were not exposed to the model training. The ground landcovers for these locations were identified visually using the Google Earth images. The performance was as

measured in weighted F1 -score. The bar graph in figure 12 sums up the performances of the algorithms and data compositions.

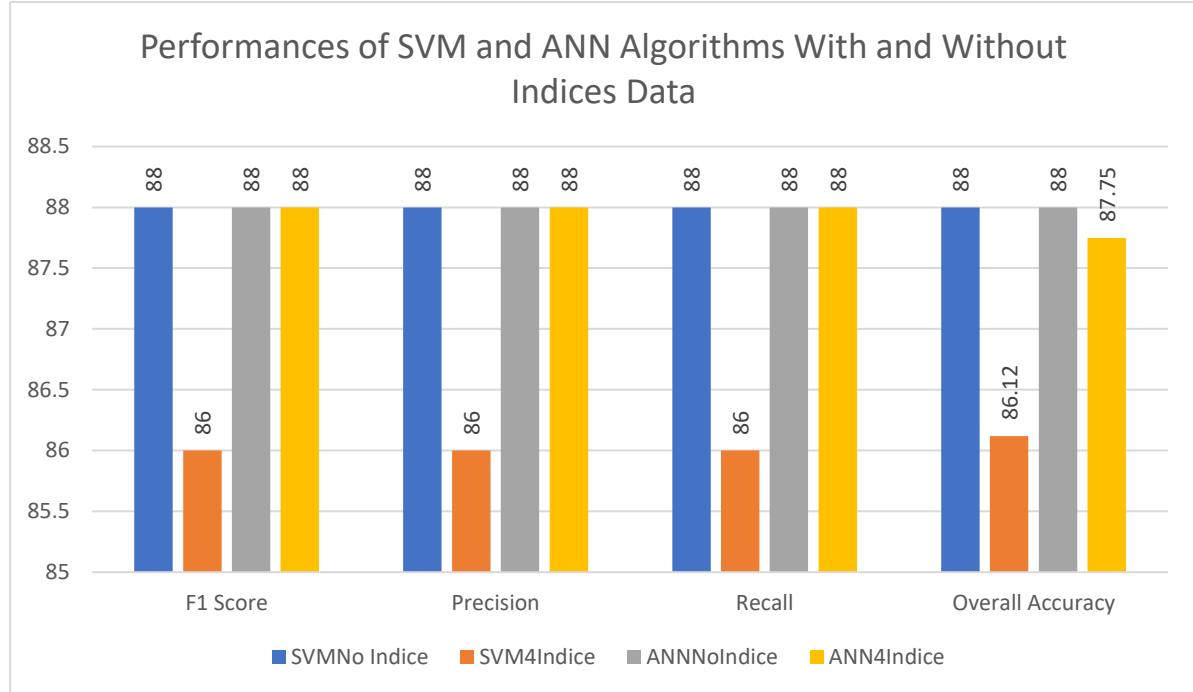


Figure 12: Algorithms Performance

Overall, all of the models had good F1-score of over 86%. From the classification results of the 4 trialled models, the F1-score of 88% was seen in 2, SVM without indices model and ANN with 4 indices model. Both of this model had precision and recall of weighted 88%. On the other hand, worst weighted F1-score of 86% was seen in SVM model with 4 indices. The trend of decrease with addition of indices was observed in overall accuracy too for both the models. The overall accuracy of SVM model without the indices was 88%, while addition of extra indices reduced it by 1.88%. The same trend was visible in the case of ANN too, but the margin was 0.25%.

Another crucial factor in measuring the performance of ANN is the validation and the loss curves to examine if there was any case of underfit or over fit. It is observed the models are not overfitting or underfitting in none of the cases. For No Indices model the validation accuracy was closer to the training accuracy than that of model with 4 indices. The validation loss curve stopped dropping after 400 epoch for the model with 4 indices, on the other hand, in no indices model the validation loss kept decreasing till 500 epochs. It is also noticeable that, the model without indices data took 700 iterations to obtain the optimal accuracy, whereas with the indices data, the model ran till 942 epochs to reach the optimal state.

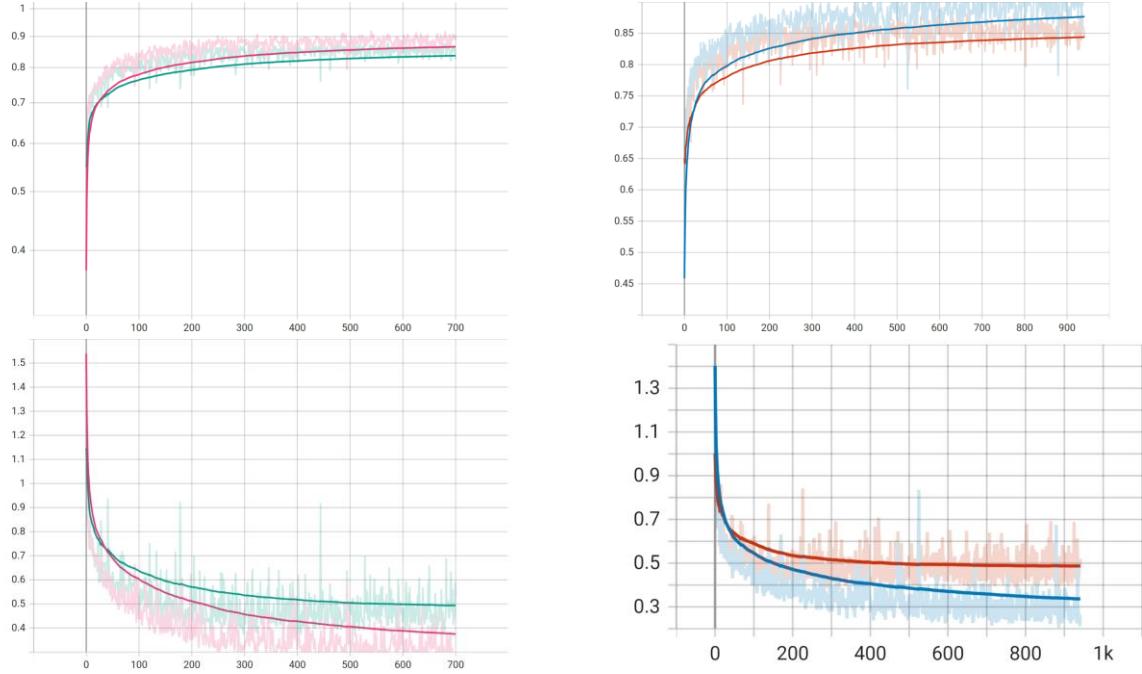
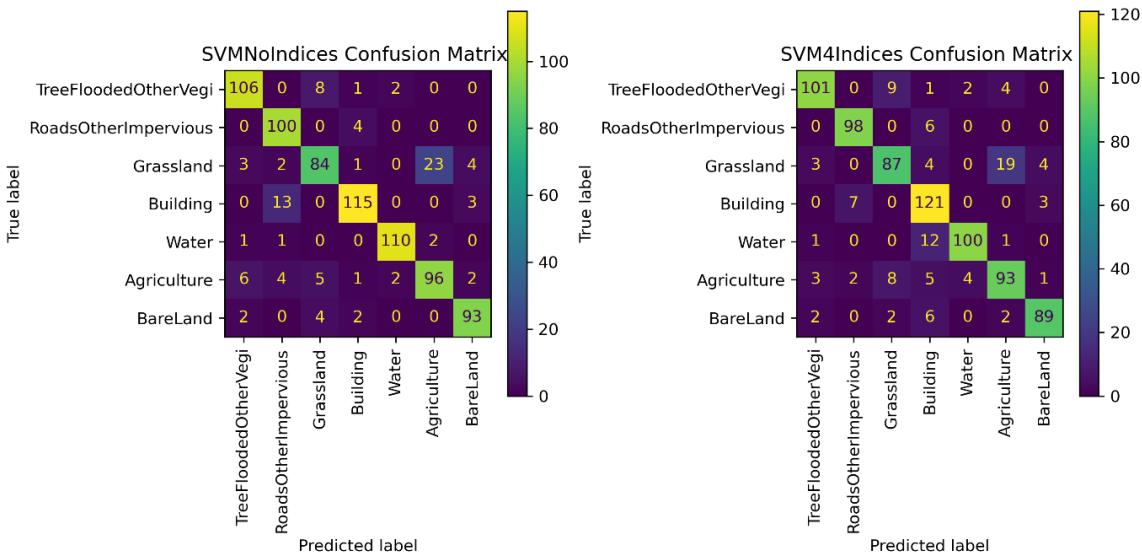


Figure 13: Training and Validation Loss Curves for ANN Models. Left top and bottom (ANN No Indices accuracy and loss curve), Right top and bottom ANN with 4 indices accuracy and loss curves.

The class wise correct predictions for the models are presented in the confusion matrices in figure 14. It is seen that, with the inclusion of indices, for 5 out of 7 classes number of correct predictions (true positive) decreases. Although the ANN model has also produced decreasing class wise accuracies with 4 indices 4 of the classes, but the classes Grassland, Building, Agriculture has had more correct predictions with 4 indices.



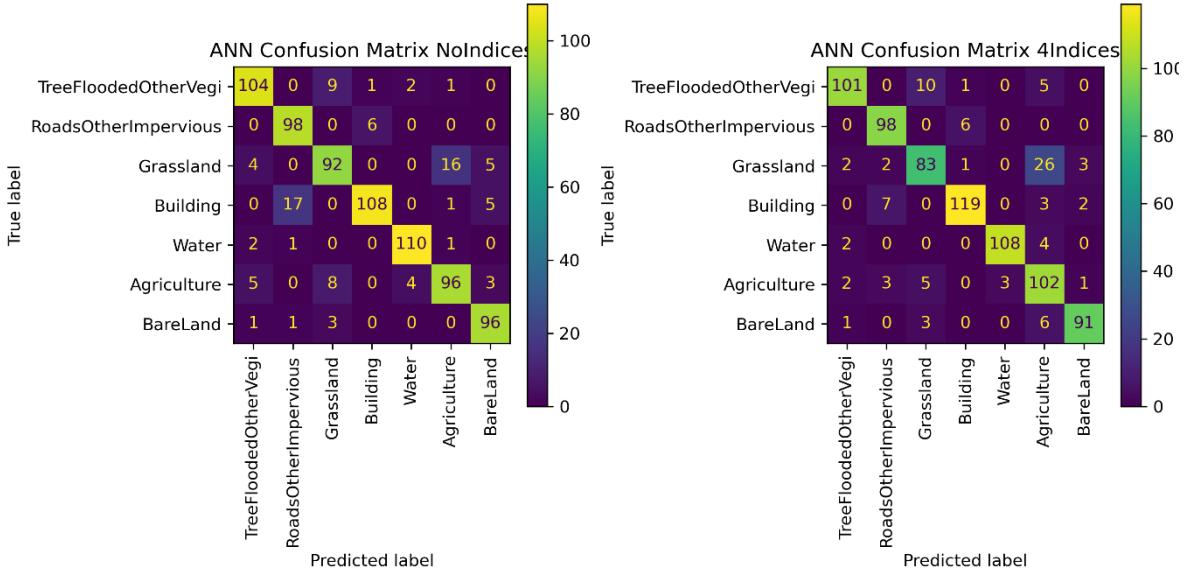


Figure 14: Class wise predictions by the algorithms.

4.1 Proposed Hybrid Selection Approach

It was prevalent that, there is lack of consistency in the results when indices are added to the data. Because, cases when the model is showing weaker overall scores, but the class wise counts were increasing at the same time. The new approach which was applied to tackle this inconsistency had higher F1-Score, and accuracy. The approach showed a weighted F1 score of 90%, which is higher than all other models. At the same time the class wise correct predictions were also increased for Tree, Roads Breland and Water classes comparing to the other models.

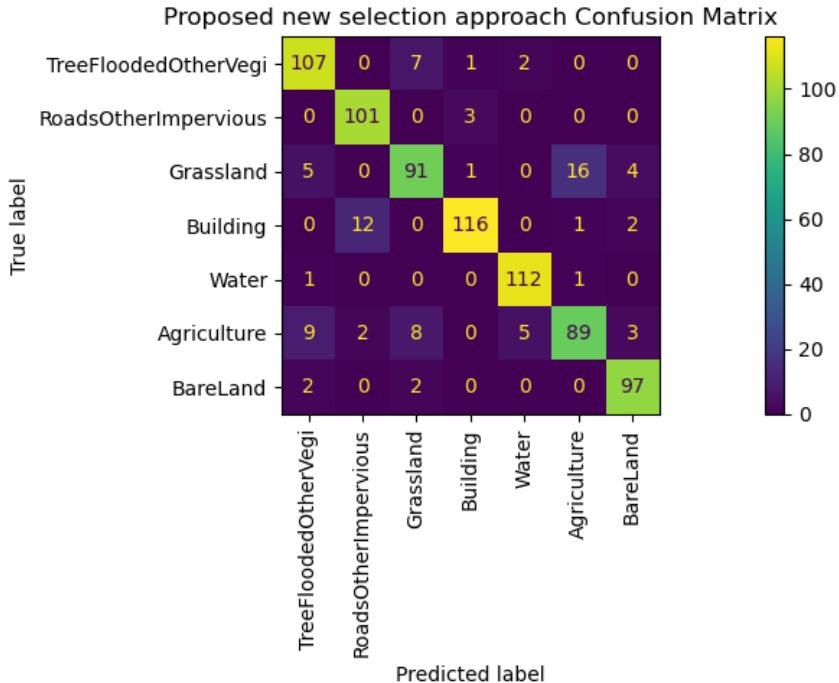


Figure 15: Confusion matrix for proposed selection approach.

Apart from prediction performances, the study also investigated the time taken to fit the model. As mentioned above when using the indices in training, ANN model with indices takes more epoch than the model without indices to reach optimum stage. However, it was different for the case of SVM.

While fitting the grid to tune the hyper parameter, the same grid took less time in the case of indices incorporation to fit. The total time to fit the grid with no indices data was 30.8021 minutes, where as with indices data it was only 0.6817 minute. The exact fitting times for individual hyper parameter combinations are shown in the following chart in figure.

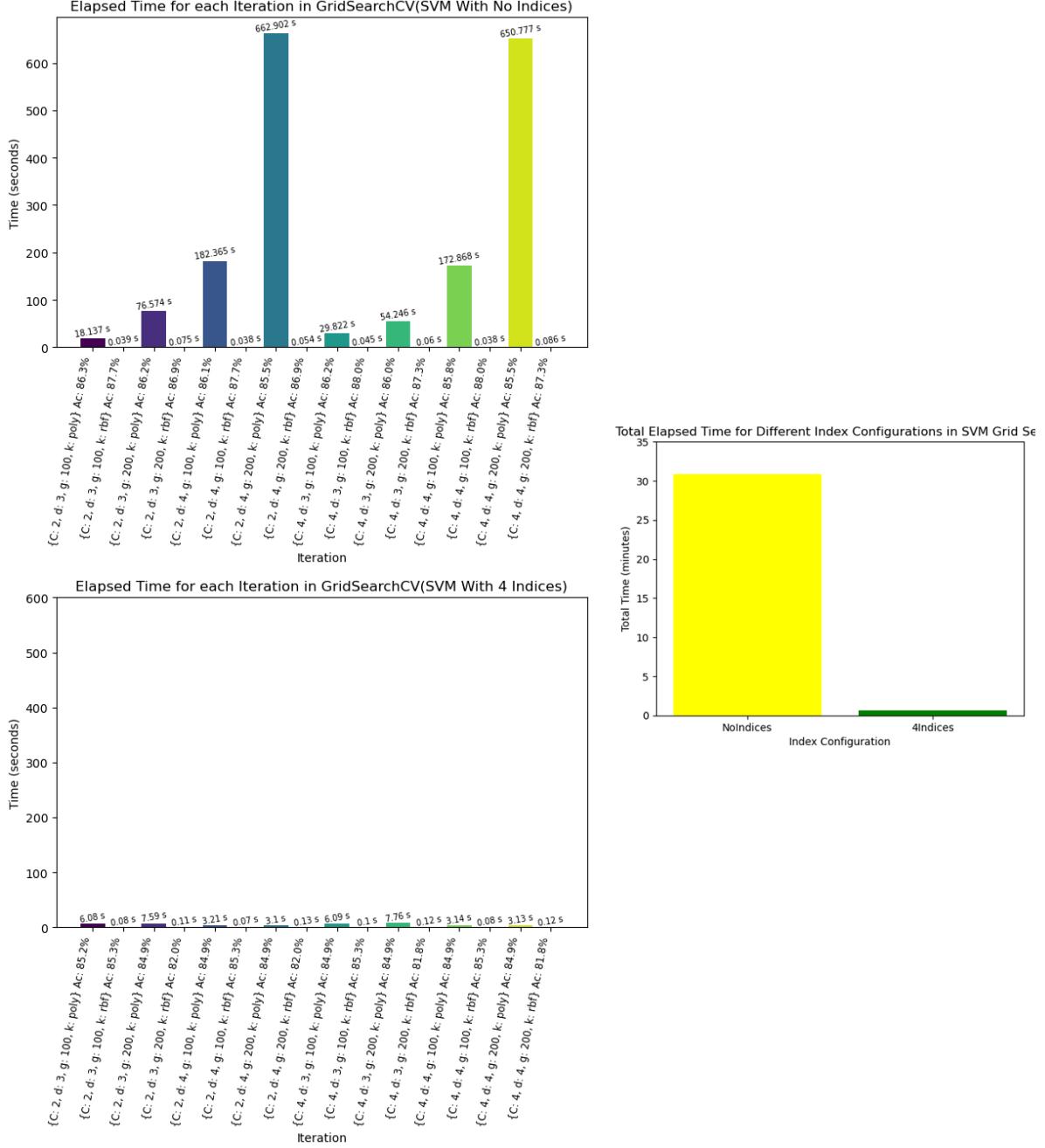


Figure 16: Model fitting time for different hyperparameter combinations and their accuracy with and without indices data.

From the figures it is also observed that, typically the SVM models with Polynomial kernels take more time to fit when only the bands are used. But when indices data are added in the training, the fitting time significantly reduces. The opposite happens for the Radial Bias Function (RBF) kernel.

Apart from investigating the attributes of the algorithms, one of the objectives of the study was to produce an up to date LULC map for Phnom Penh. The following maps present the 5 different LULC map of 2023 for the area. From all these maps it is observed that the core of the study area which is

Phnom Penh is highly dominated by buildings, while the surrounding areas are covered 3 different types of vegetation covers.

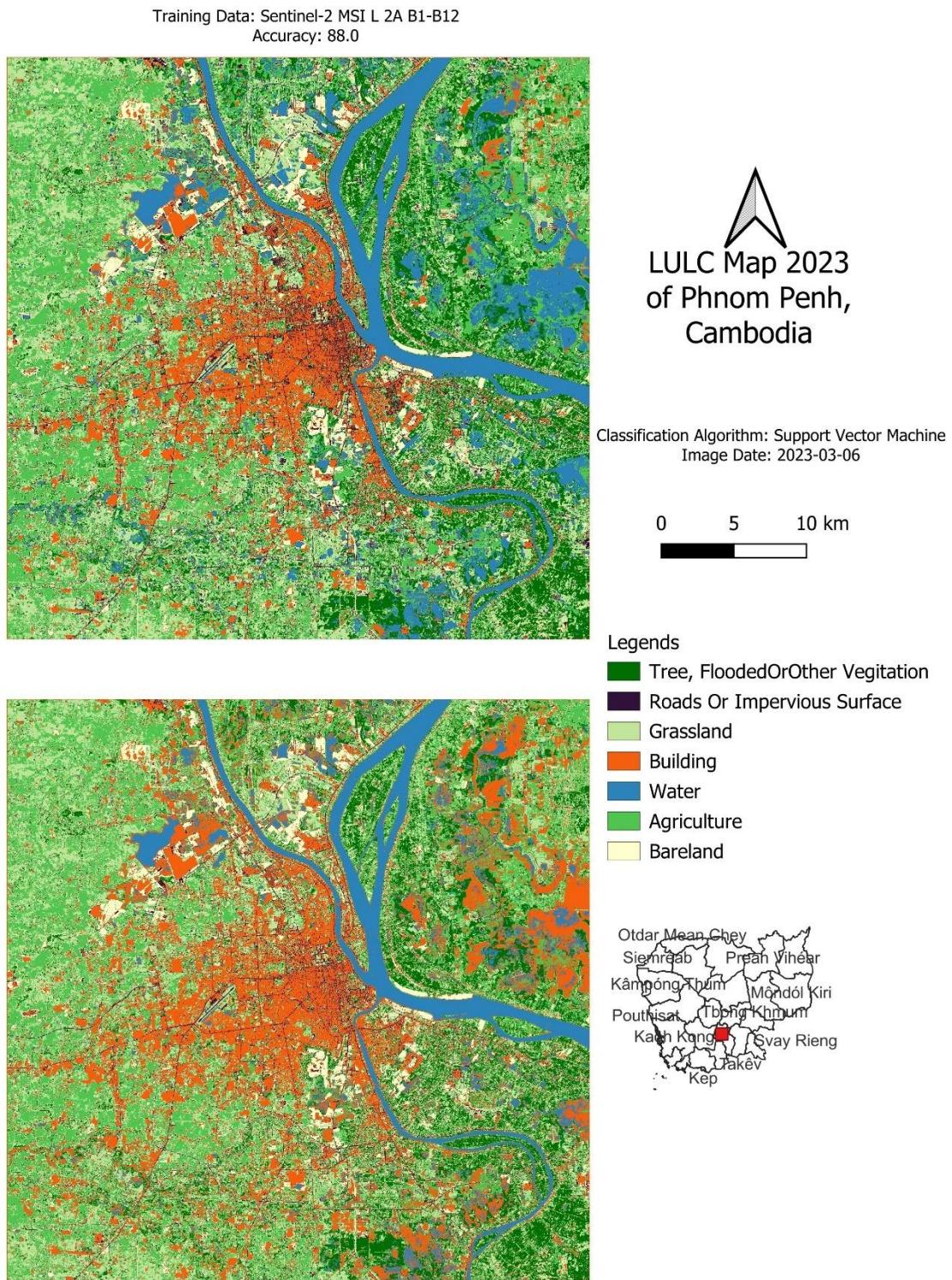
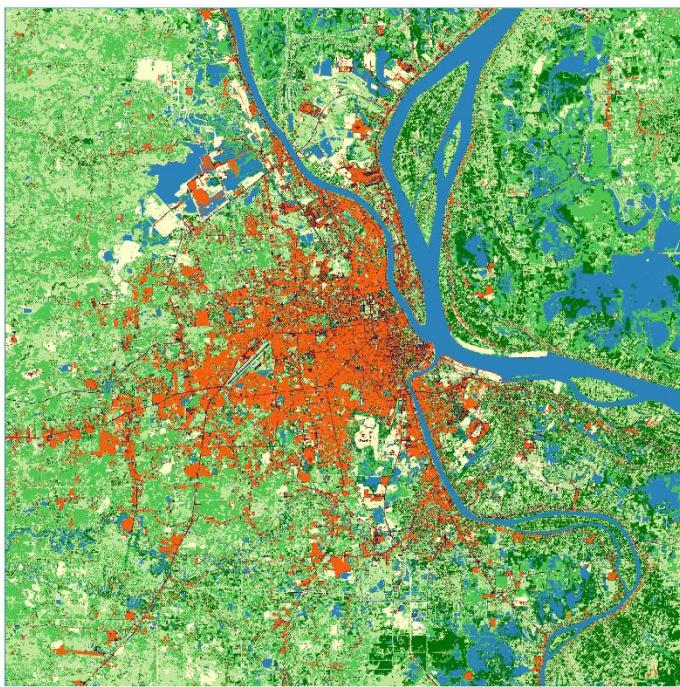


Figure 17: Classified Outputs from trialled SVM models

Training Data: Sentinel-2 MSI L 2A B1-B12
Accuracy: 88.0



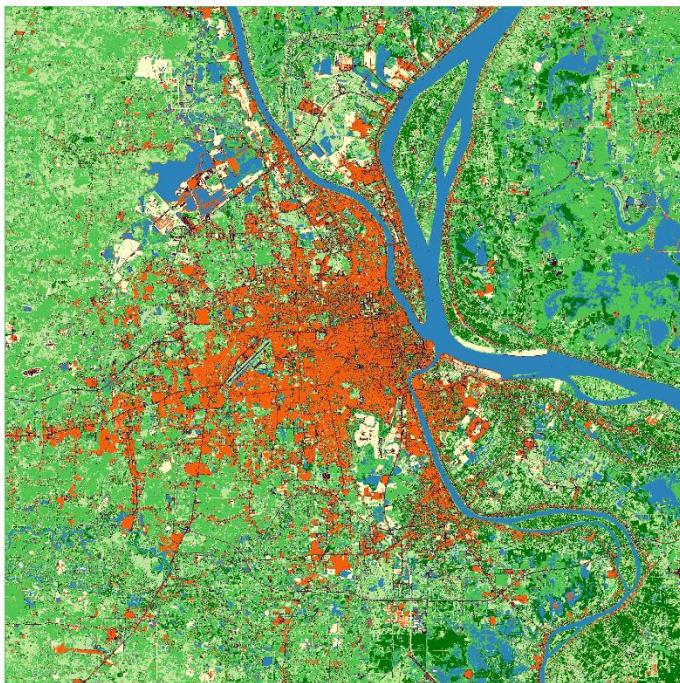
LULC Map 2023
of Phnom Penh,
Cambodia

Classification Algorithm: Artificial Neural Network
Image Date: 2023-03-06

0 5 10 km

Legends

- [Green] Tree, FloodedOrOther Vegetation
- [Dark Purple] Roads Or Impervious Surface
- [Light Green] Grassland
- [Orange] Building
- [Blue] Water
- [Dark Green] Agriculture
- [Yellow] Bareland



Otdar Mean Chey
Siem Reap Preah Vihear
Kampóng Chhnang Mondol Kiri
Pouthsat Tbong Khmum
Kaoh Kong Svay Rieng
Takeo Kep

Training Data: Sentinel-2 MSI L 2A B1-B12 + (NDBI, NDWI, NDVI, BSI)
Accuracy: 87.75

Figure 18: Classified Outputs from trialled ANN models

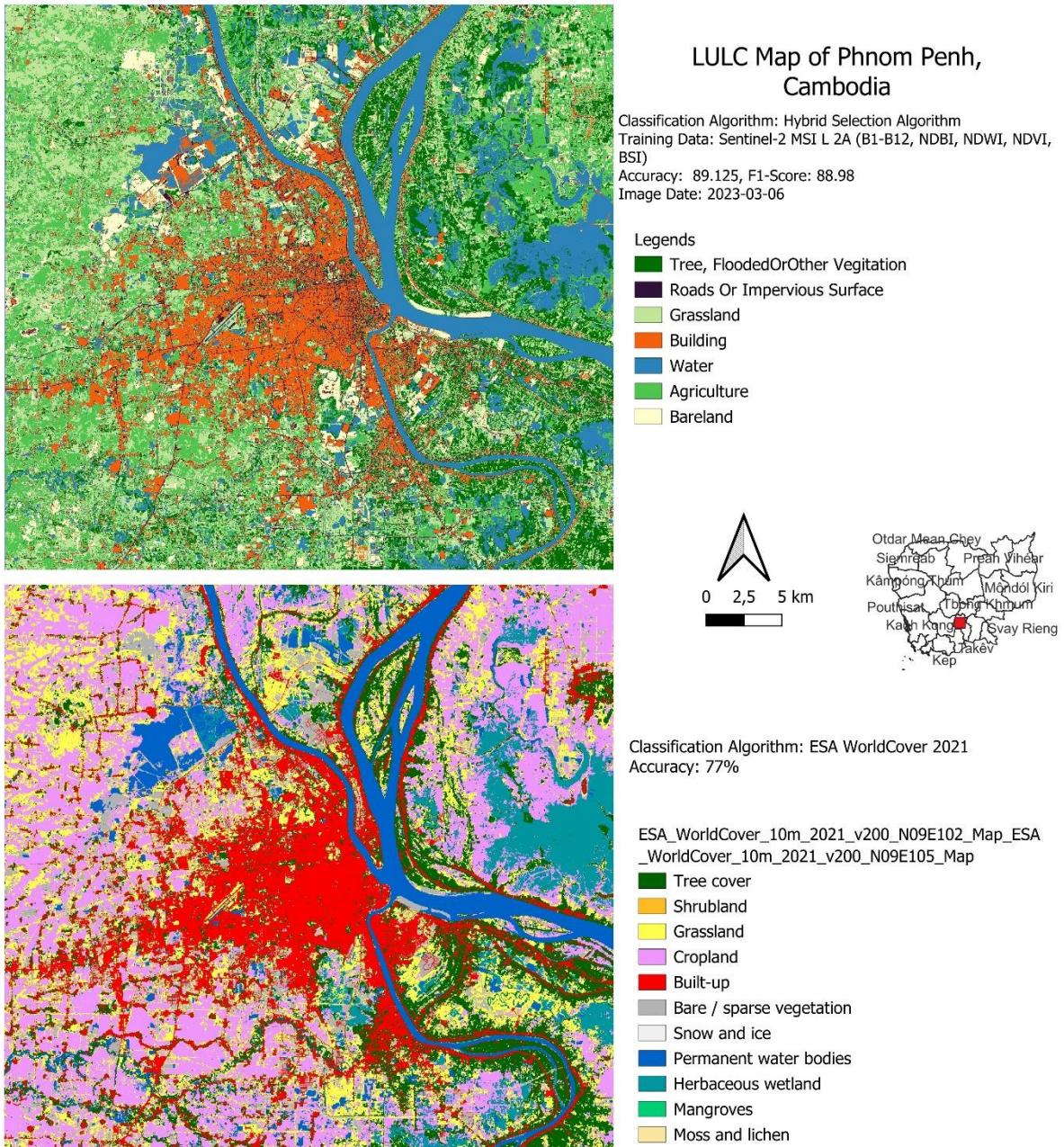


Figure 19: Classified Outputs from proposed hybrid selection model

In figure 17, 18, and 19 the classified outputs from SVM, ANN, and Hybrid Selection have been shown. An additional classified output has been provided from EAS World cover 2021 map (Zanaga et al., 2022). The EAS Worl cover has an overall accuracy of 77%. From the visual inspection it can be said that the map produced from hybrid selection approach and others has better classification ability. Although the EAS World cover has more micro classes, but they are comparatively generalized and less detailed. For instance, the class road has been included in the built-up area though it was not able to predict the pattern of the roads. In this sense, it can be said that the hybrid selection output could generalize better.

5 Discussion

From the systematic study it has been seen that, the inclusion of spectral indices does not significantly contribute to the classification accuracies. In the conducted studies it has been

seen that, in both SVM and ANN algorithm the F1-Score has either been unchanged or decreased. Here, F1-score is given preference over other accuracy measurements such as precision (users accuracy), recall (producers accuracy), and over all accuracy. Because overall accuracy measurements do not provide information about false positives and false negatives separately. F1 score gives a balance between precision (false positives) and recall (false negatives), offering insights into the model's ability to minimize both types of errors.

The dataset combinations that were investigated showed that these does not contribute to improving the accuracies. Rather, in specific cases it increases the time complexity. For instance, while using RBF kernel in SVM, it increases the time complexity. Although it was smaller in margin, but it was consistent. While applying for a larger dataset it would increase the time complexity more significantly. However, when Polynomial kernel is used in SVM algorithm, the indices data can significantly reduce the time complexity. Therefore, it can be said that, for improvement in overall classification it is not always the best to use the indices. For the indices used in this study, namely NDBI, NDVI, NDWI, and BSI, it can be related to their co-relationship with the LULC classes. The following co-relation matrix shows that the LULC classes does not have a strong correlation with the spectral index's bands.

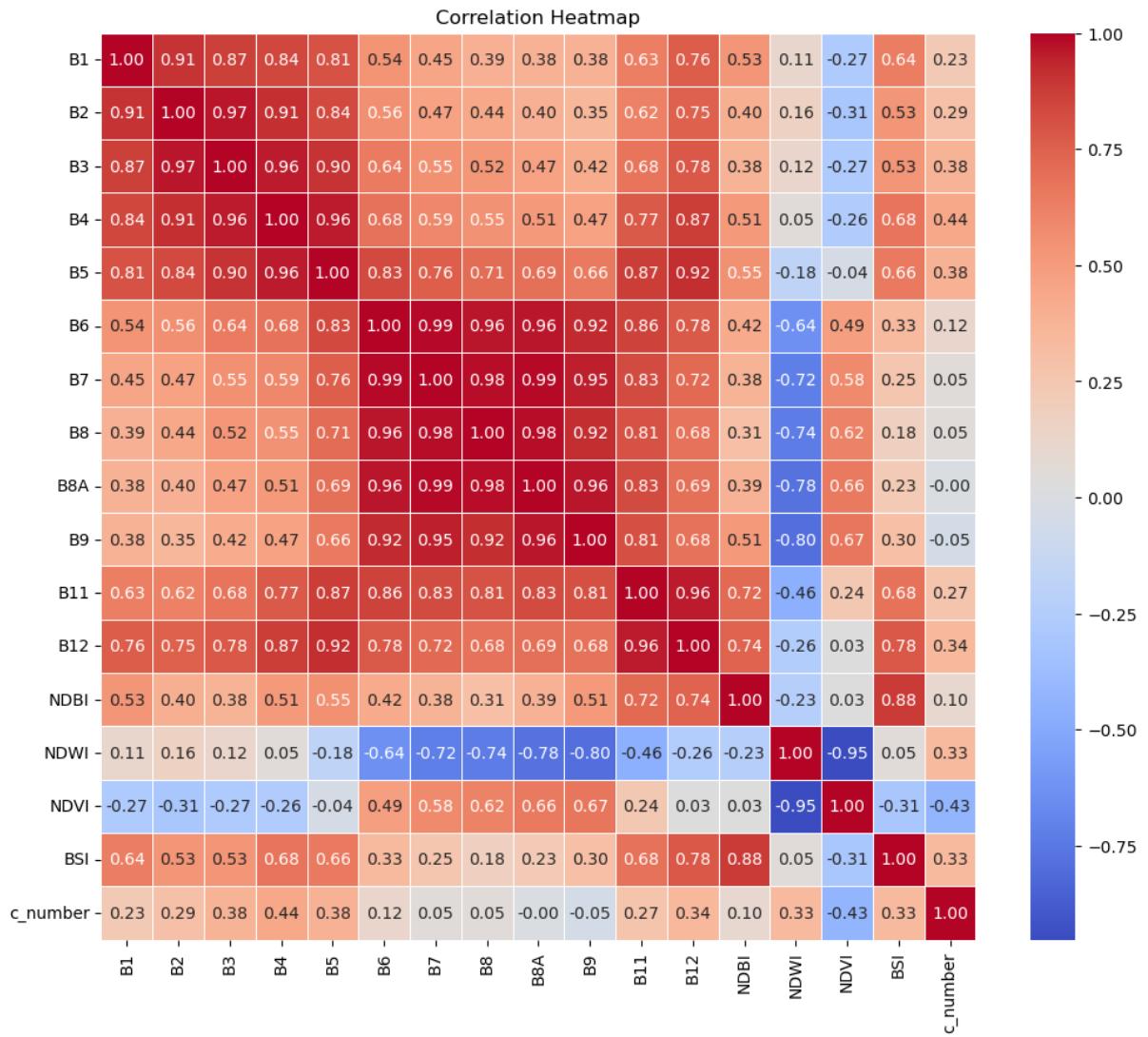


Figure 20: Correlation among the bands

This outcome also aligns with the previous literature from (Hosseiny et al., 2022). They also agreed that adding additional bands does not always improve the accuracy. Thus it can be questioned for (Ju et al., 2021)'s case, that if the additional bands were not incorporated how the performance of sophisticate deep learning models. It is true that, the deep learning architecture which is used in this study is not the most sophisticated one. Therefore, further studies could be done to investigate the significance of these additional bands in state-of-the art algorithms. Therefore, while choosing the band combinations for LULC tasks a thorough literature review or trial should be done to find the optimal band sets.

One option while choosing the best band set combination could be focusing on the classes included in the LULC task. For instance, the classes Buildings was always better identified with the indices data. The could other classes which has such tendencies. For such works the following table can be helpful.

Tree Flooded and Other Vegetation	Recommended Data and Algorithm
Roads and Other Impervious	SVM_NoIndices
Grassland	SVM_4Indices
Building	ANN_NoIndices
Water	ANN_4Indices
Agriculture	SVM_NoIndices
Bare Land	ANN_NoIndices
Tree Flooded and Other Vegetation	ANN_4Indices

Along with other insights, the study also produced up to date LULC maps for Phnom Penh area. According to those maps the land use statistics can be generated which can be useful for the policy makes. The best F1 score of 90% was achieved by the newly proposed hybrid selection approach which classifies each pixel based on the probabilities received from ANN and SVM models. According to the LULC map from the algorithm, the LULC scenario in the study area is as following.

Class	Pixel count	Area (m ²)
Tree Flooded and Other Vegetation	268050800	268050800
Roads and Other Impervious	967605	96760500
Grassland	3933015	393301500
Building	1940731	194073100
Water	2295655	229565500
Agriculture	3200180	320018000
Bare Land	1066410	106641000

These insights can be utilized in the policy making and further change detection analysis of the area. The limitation of the study is it used a dataset of nearly 3000 samples, this limited dataset is not always capable of generalizing intricate characteristics in ML DL algorithms. Therefore, these models can not be used as a generalized ML DL models for every location rather their scope is limited only to the area. Nevertheless, the focus of the study was to find the significance of the incorporation which the study successfully could investigate.

6 Conclusion

In conclusion, the study aimed to evaluate the performance of adding indices with sentinel 2 bands in different machine learning models, specifically SVM and ANN, for land use and land cover (LULC)

classification in the Phnom Penh area. The results indicated that all models achieved good overall performance, with F1-scores exceeding 86%. Notably, the SVM without indices and ANN with 4 indices models stood out with the highest F1-score of 88%, demonstrating balanced precision and recall. However, the addition of indices seemed to have a diminishing effect on overall accuracy, particularly for SVM, where the model without indices outperformed its counterpart with additional indices. A novel selection approach was proposed to address inconsistencies observed when incorporating spectral indices. This approach demonstrated a higher weighted F1-score of 90%, surpassing all other models. Additionally, class-wise correct predictions were increased for specific land cover classes compared to other models. The study also considered the time taken to fit the models, revealing that SVM with indices reduced fitting time during hyperparameter tuning, contrasting with the increased fitting time observed for the ANN model with indices. Beyond model evaluation, the study aimed to produce an up to date LULC map for Phnom Penh, highlighting the dominance of buildings in the city core and various vegetation covers in surrounding areas. The importance of considering the specific characteristics of the LULC task when choosing band combinations, suggesting recommendations for optimal band sets based on different land classes should also be acknowledged prior to data selection. Finally, the study provided valuable insights for policymakers, presenting LULC statistics and maps that can aid in decision-making and change detection analysis. However, the study recognized limitations, such as the use of a limited dataset that may not generalize well to other locations.

7 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>
- Bhatnagar, S., Gill, L., & Ghosh, B. (2020). Drone Image Segmentation Using Machine and Deep Learning for Mapping Raised Bog Vegetation Communities. *Remote Sensing*, 12(16), 2602. <https://doi.org/10.3390/rs12162602>
- 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>
- F1 Score in Machine Learning: Intro & Calculation. (n.d.). Retrieved 23 February 2024, from <https://www.v7labs.com/blog/f1-score-guide>, <https://www.v7labs.com/blog/f1-score-guide>
- 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>
- Jiang, H., Lu, N., Huang, G., Yao, L., Qin, J., & Liu, H. (2020). Spatial scale effects on retrieval accuracy of surface solar radiation using satellite data. *Applied Energy*, 270, 115178.
<https://doi.org/10.1016/j.apenergy.2020.115178>
- 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>

Loshchilov, I., & Hutter, F. (2017a). Decoupled Weight Decay Regularization. <https://doi.org/10.48550/ARXIV.1711.05101>

Loshchilov, I., & Hutter, F. (2017b, November 14). Decoupled Weight Decay Regularization. arXiv.Org. <https://arxiv.org/abs/1711.05101v3>

Macarringué, 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). INTERNATIONAL ENVIRONMENTAL AGREEMENTS: 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., Essahaoui, A., Oudja, 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>

Understanding Support Vector Machines: A Primer. (2017, March 9). Machine Learning in Action. <https://appliedmachinelearning.wordpress.com/2017/03/09/understanding-support-vector-machines-a-primer/>

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>

Xu, B., Wang, N., Chen, T., & Li, M. (2015, May 5). Empirical Evaluation of Rectified Activations in Convolutional Network. arXiv.Org. <https://arxiv.org/abs/1505.00853v2>

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>

Zanaga, D., Van De Kerchove, R., Daems, D., De Keersmaecker, W., Brockmann, C., Kirches, G., Wevers, J., Cartus, O., Santoro, M., Fritz, S., Lesiv, M., Herold, M., Tsednebazar, N.-E., Xu, P., Ramoino, F., & Arino, O. (2022). ESA WorldCover 10 m 2021 v200 (Version v200) [dataset]. Zenodo. <https://doi.org/10.5281/ZENODO.7254221>

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>