ANALYSIS OF THE SPATIOTEMPORAL FLUCTUATIONS IN MANGROVES USING MACHINE LEARNING FOR HEALTH ASSESSMENT

A PROJECT REPORT

Submitted by

MADHUMITHA P S (2021503520)

ACHYUT PRASAD D C (2021503002)

SANDHIYA S (2021503552)

COURSE CODE: CS6611

COURSE TITLE: CREATIVE AND INNOVATIVE PROJECT

DEPARTMENT OF COMPUTER TECHNOLOGY

ANNA UNIVERSITY, MIT CAMPUS

CHENNAI – 600044

MAY 2024

DEPARTMENT OF COMPUTER TECHNOLOGY

ANNA UNIVERSITY, MIT CAMPUS

CHROMEPET, CHENNAI – 600044

BONAFIDE CERTIFICATE

Certified that this project report "**Analysis of the Spatiotemporal fluctuations in Mangroves using Machine Learning for Health Assessment**" is the work of **Ms. Madhumitha P S (2021503520), Mr. Achyut Prasad D C (2021503002), Ms. Sandhiya S (2021503552)** in the Creative and Innovative Project Laboratory subject code CS6611 during the period January to May 2024.

SIGNATURE

Dr. Kathiroli R
SUPERVISOR
Assistant Professor (Sl. Gr)
Department of Computer Technology
Anna University, MIT Campus
Chromepet – 600 044

SIGNATURE

Dr. Kottlingam K
COURSE IN – CHARGE
Assistant Professor
Department of Computer Technology
Anna University, MIT Campus
Chromepet – 600 044

ABSTRACT

The health and evolution of mangrove forests are critical for ecosystem preservation. However, these forests face significant threats from rising water levels and changing weather patterns. Traditional monitoring methods are often inconsistent and slow, particularly in large areas like Pichavaram.

To address this, this project utilizes satellite imagery and machine learning techniques to analyze the spatiotemporal fluctuations in mangroves. Specifically, the project map's locations within Pichavaram monitors changes in mangrove area over time, and estimates mangrove density. A robust machine learning model will be developed to differentiate and map mangroves, non-mangroves, and water bodies in satellite imagery, quantifying temporal changes in mangrove extent from 2014 to 2024.

The project will explore the potential of incorporating multi-temporal satellite data to assess seasonal variations in mangrove extent and health, providing a more comprehensive understanding of the dynamics of these ecosystems throughout the year.

By establishing a robust methodology for mangrove mapping and monitoring using satellite imagery and machine learning, the project sets a foundation for long-term, scalable monitoring efforts that can be replicated in other mangrove regions globally, contributing to broader conservation initiatives. The proposed CNN model tested in Pichavaram presented an overall accuracy of 91% indicating that the model performed well in classifying the images.

ACKNOWLEDGEMENT

We take this humble opportunity to thank the Dean, MIT Campus, Anna University, Dr. Ravichandran K, and Dr. Jayashree P, Professor & Head, Department of Computer Technology, MIT Campus, Anna University for providing all the lab facilities in pursuit of this project.

Undertaking this project has helped us learn a lot, and we would like to express our gratitude towards our supervisor Dr. Kathiroli R, Assistant Professor, Department of Computer Technology, MIT, Anna University, whose guidance, and directions helped shape this project perfectly. The feedback from the supervisor was very instrumental in the successful completion of the project work.

We acknowledge the efforts and feedback of the panel members Dr. Chithra S, Associate Professor, Department of Computer Technology, MIT, Anna University, Dr. Jayachitra V P, Assistant Professor, Department of Computer Technology, MIT, Anna University, and Dr. Kottilingam K, Assistant Professor, Department of Computer Technology, MIT, Anna University, in reviewing our work, providing constant valuable comments and encouraging us to view the different aspects of the project in successful implementation of the project.

We thank all the teaching and non-teaching members of the Department of Computer Technology, MIT, Anna University whose appreciable help, either directly or indirectly helped in the smooth completion of the Creative and Innovative Project in a limited time frame.

Madhumitha P S

Achyut Prasad D C

Sandhiya S

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

S. No.	ABBREVIATION	DEFINITION
1	RGB	Red Blue Green
2	QA	Quality Assessment
3	CNN	Convolutional Neural Network
4	OLI	Operational Land Imager
5	MSI	Multispectral Instrument
6	QA_PIXEL	Pixel Quality Assessment band
7	QA60	Quality Spatial Resolution 60 metres
8	AOI	Area of Interest
9	ReLU	Rectified Linear Unit
10	TIFF	Tagged Image File Format
11	GeoTIFF	Geographic Tagged Image File Format
12	GEE	Google Earth Engine

CHAPTER 1

INTRODUCTION

Machine learning algorithms can effectively analyze multispectral satellite imagery to map and monitor mangrove forests over time, providing valuable insights into their spatial and temporal dynamics.

1.1 MACHINE LEARNING:

Machine learning (ML) is a subfield of artificial intelligence (AI) that focuses on the development of computer algorithms that can learn and improve from data without being explicitly programmed. It is a statistical approach to studying and making inferences about data that utilizes a variety of algorithms suited for answering complex questions and making predictions [2]. At its core, machine learning enables computers to learn and make decisions based on data, rather than following a set of pre-programmed instructions Machine learning is a process in which computing systems learn from data and use algorithms to execute tasks without being explicitly programmed. This is achieved by feeding the training data into the selected model, allowing the algorithm to learn the underlying patterns and relationships in the data [5]. The performance of ML algorithms adaptively improves with an increase in the number of available samples during the 'learning phase'.

There are three main types of machine learning: supervised learning, unsupervised learning, and reinforcement learning. Supervised learning involves training the model on labeled data, where the inputs and their corresponding outputs are provided [4]. Unsupervised learning involves training the model on unlabeled data, and it tries to find inherent patterns or structures within the data. Reinforcement learning involves the model learning by interacting with an environment and receiving feedback (rewards or penalties) for its actions.

Machine learning is a powerful tool that is driving innovation and automation across a wide range of industries, making it an increasingly important technology in the modern world.

1.2 MULTISPECTRAL IMAGES

A multispectral image is a type of image that captures information from a scene at multiple wavelengths of light. This can include wavelengths from the visible spectrum as well as wavelengths from the infrared and ultraviolet spectrum [1]. Multi-spectral images consist of multiple monochrome images of the same scene, each captured by a different sensor, with each image referred to as a "band." These images are commonly used in remote sensing applications, with satellites capturing images from various frequency bands in the visual and non-visual range. Techniques exist to process multi-spectral images, such as classifying pixels based on their intensities in different bands to form a feature vector. While multi-spectral images offer valuable insights, processing them requires more computation time and memory due to the additional data involved [2]. Some applications of multispectral imaging include environmental vegetation coverage detection, geological surveys, environmental monitoring.

Multi-spectral images capture data across multiple spectral bands, typically ranging from 3 to 15 bands, though some can have up to 36 bands. The spectral bands used in multi-spectral imaging can include visible light (blue, green, red), as well as non-visible portions of the electromagnetic spectrum like near-infrared, short-wave infrared, mid-wave infrared, and long-wave infrared [3]. The different spectral bands provide complementary information about the imaged scene, enabling the differentiation of features like vegetation, soil, water, and man-made structures that may not be easily distinguished in just the visible spectrum. Multi-spectral imaging has been widely used in remote sensing applications, such as satellite-based Earth observation, as well as in document and artwork analysis, where the different spectral bands can reveal hidden information.

1.2.1 ADVANTAGES OF MULTISPECTRAL IMAGES

1. Improved object identification and classification: Multispectral images capture reflectance data across multiple wavelength bands, allowing for better differentiation and identification of various earth surface features like vegetation, minerals, water bodies, etc. compared to single-band or RGB images.
2. Enhanced vegetation and crop monitoring: The different spectral bands, especially the near-infrared and red-edge bands, provide valuable information about plant health, stress, and productivity that cannot be obtained from visible-only imagery.
3. Ability to detect changes over time: By analyzing the spectral changes in different bands, multispectral data enables effective monitoring of land cover changes, deforestation, urbanization, and other environmental transformations.
4. Improved accuracy in land use/land cover classification: The additional spectral information in multispectral images leads to more accurate and detailed classification of land cover types compared to using RGB or single-band data.
5. Versatility in applications: Multispectral remote sensing has diverse applications ranging from precision agriculture, forestry, and environmental monitoring, to military surveillance and medical diagnostics.

1.2.2 PROPERTIES OF MULTISPECTRAL IMAGES

Multi-spectral imaging involves multiple spectral bands, typically ranging from visible to infrared wavelengths. In this imaging technique, various spectral bands are utilized to capture specific information about the object being imaged. The spectral bands commonly used in multi-spectral imaging include visible spectra, near-infrared spectra, short-wave infrared, mid-wave infrared, long-wave infrared. Multi-spectral imaging systems can consist of a minimum of 3 to a maximum of 15 spectral bands, allowing for detailed analysis and extraction of information from the captured images. Landsat 8 has 11 spectral bands, while Sentinel-2 has 13 spectral bands, allowing them to capture information in the visible, near-infrared, and shortwave infrared regions of the electromagnetic spectrum. This multispectral data provides more detailed information about the composition and properties of the imaged objects, such as vegetation, soils, water bodies, etc., compared to a simple optical (RGB) image. The multiple spectral bands in Landsat 8 and Sentinel-2 enable applications like mapping mangrove forests, detecting plant stress, and monitoring environmental changes that would not be possible with just a standard optical image.

1.3 IMAGE PROCESSING

I. Cloud Masking:

Cloud masking is a crucial preprocessing step for satellite imagery, as clouds can obscure the Earth's surface and introduce errors in subsequent analysis. The raw satellite data, such as from Sentinel-2 and Landsat 8, can be affected by radiometric distortions and atmospheric interference, including the presence of clouds and cloud shadows. Cloud masking involves identifying and removing pixels affected by clouds and cirrus clouds using the Quality Assessment (QA) band in the satellite data. Effective cloud masking is essential for creating accurate and reliable composite images, as it ensures that the analysis is performed on clear, cloud-free pixels.

II. Composite Image Creation:

Compositing is a technique used to create a composite image from multiple satellite scenes acquired over a specific period. The process involves selecting the median value of each pixel across the multiple input scenes, effectively reducing the impact of outliers and noise in the data. Median compositing can help to create a more representative and consistent image, as it minimizes the influence of factors like cloud cover, atmospheric interference, and sensor artifacts. The resulting composite image can then be used for further analysis and interpretation, although it may not accurately represent the land surface at any single point in time.

III. Pixel Value Normalization:

Pixel value normalization is a common preprocessing technique used to ensure that the pixel values are within a consistent range, typically between 0 and 1. For Sentinel-2 and Landsat 8 data the raw pixel values can be normalized by dividing them by 10,000, effectively scaling the values to the desired range. Normalization helps to standardize the pixel values, which can be important for certain analysis techniques, such as machine learning algorithms, that may be sensitive to the scale of the input data.

1.4 OBJECTIVE

1. To utilize satellite imagery from different sources and assess the temporal dynamics and health of mangrove ecosystems.
2. To develop a robust classification model utilizing a neural network for classifying it into different density classes.
3. To perform a time series analysis on the classified images to quantify temporal changes in mangrove extent from 2014 to 2024 and visualize the trends in the health dynamics.

CHAPTER 2

LITERATURE SURVEY

In [1], a novel method for mangrove identification, Remote sensing for Mangrove Identification (REMI), has been introduced. This approach leverages specific satellite bands to effectively capture distinctive mangrove characteristics. REMI exhibits instability in non-vegetation areas hindering its accuracy and applicability.

Chen. R et al in [2] proposed an approach that involves integrating GF-2, GF-3, and UAV-LiDAR data for comprehensive geospatial analysis and applying random forest regression to retrieve height information. However, limitations include constraints faced by Random Forest(RF) and Support Vector Machine(SVM) studies due to data availability challenges and geographic specificity.

The proposed technique by Sharfi. A et al in [3] evaluate the Random Forest's efficiency in mangrove mapping using Sentinel-1, and Sentinel-2 (with red-edge bands) data by comparing three scenarios: sentinel-1 only, sentinel-2 only, and the combined data. However, a limitation observed in related works tends to the ineffectiveness of Landsat images with 30m resolution in identifying and monitoring isolated mangrove forest patches smaller than 1 ha.

In [4], this paper introduces the Generalized Composite Mangrove Index (GCMI) and the utilization of Sentinel-2 time series data to improve the separability between mangroves and other land covers. It achieves this by compositing vegetation indices and water indices and using similarity trend distance measures to determine the optimal indices. Despite these advancements, limitations persist in the phenological information it contains within the time series data.

The Proposed technique by R. Barella et al in [5] involves using sentinel-1 and sentinel-2 for glacier mapping. By leveraging Sentinel-1 structural information and sentinel-2 spectral data, the method aims to achieve more accurate delineation of glaciers including debris-covered areas, and differentiate between snow and ice. However, a significant limitation surfaces with persistent cloud cover affecting the reconstruction of minimum snow cover, particularly when relying on optical data such as Sentinel-2.

The study [6] aims to map mangrove ecosystems by using Sentinel-1 and Sentinel-2 data. The integration of the Random Forest(RF) algorithm in Google Earth Engine ensures both precision and computational efficiency, promising advancements in ecosystem mapping. However, it recognizes a limitation in the moderate accuracy observed for the aerial roots class.

In [7], this introduces the integration of Object-Based Image Analysis (OBIA) and Random Forest (RF) for accurate mangrove mapping, utilizing Gaofen-2(GF) and Sentinel-2 data. A key objective is to compare dataset efficiency and accuracy and limitations are acknowledged, encompassing restricted scalability, difficulties in discerning small mangrove fragments and challenges specific to the detection within certain datasets.

Yancho, J.M.M. et a [8] proposed the novel Google Earth Engine Mangrove Mapping Methodology (GEEMMM) method for global mangrove mapping, incorporating cloud computing and tidal calibration to enhance precision. However, GEEMMM encounters hurdles when mapping larger areas, influencing its reliability.

The paper [9] elevates the analysis of Random Forest (RF) and Support Vector Machine (SVM), aiming for improved accuracy through the incorporation of temporal trends. One limitation is the pixel-wise approach of Random Forest and Support Vector Machine, which neglects spatial dependencies among adjacent pixels, potentially leading to suboptimal classification results.

In [10], this paper aims to assess the dynamic changes within mangrove ecosystems. This involves a comprehensive analysis of satellite imagery to understand the evolving landscape. Despite its potential, a limitation surfaces due to the inherent constraints in satellite image precision, placing restrictions on the overall accuracy of the assessment.

The related works were surveyed, and the limitations were identified to propose an organized work for our research, as described in the following elaborative headings.

CHAPTER 3

PROPOSED WORK

The proposed work focuses on understanding the health and dynamics of mangrove forests in Pichavaram, India, utilizing satellite imagery and machine learning techniques. The primary objective is to analyze the changes in mangrove extent and health from 2014 to 2024 and provide an assessment of the health status of mangrove forests based on the results derived from Convolutional Neural Networks (CNN).

3.1 INTRODUCTION:

The analysis of the health of mangroves is important because mangroves are vital ecosystems that provide a range of ecological and economic benefits, including protection against coastal erosion, support for fisheries. However, these forests are under threat due to various factors, including climate change, pollution, and human activities. By analyzing the spatiotemporal fluctuations in mangroves, this project aims to provide valuable insights into the health and dynamics of these ecosystems, which can help inform conservation and management strategies. By using machine learning algorithms, the project will be able to analyze large amounts of data and identify patterns and trends in mangrove health and dynamics. This information can then be used to inform conservation strategies and help protect these valuable ecosystems for future generations.

The Proposed Architecture diagram shown in Figure 3.1 includes

I. Data Collection:

This is the first step in the process of creating a composite image. In this case, the data is collected from Landsat 8 and Sentinel 2 satellites.

Architecture Diagram.

II. Data Preprocessing:

The data collected from the satellites is then pre-processed. This involves correcting for atmospheric conditions and masking out clouds.

III. Classification:

The pre-processed data is then classified. This involves classifying the pixels in the image into different density classes.

IV. Time Series Analysis:

The classified data is then analyzed over time. This involves looking for trends in the data, such as changes in mangrove density over time.

3.2 DATA COLLECTION

The data collection process for the mangroves in Pichavaram is establishing the specific coordinates that define the area of interest. The coordinates provided, (79.75897, 11.45440), (79.82489, 11.45575), (79.82935,11.39921), (79.75760, 11.39854), (79.75897, 11.45440) with various bands delineate the boundaries within which data will be collected and analyzed.

The selection of Landsat 8 Operational Land Imager(OLI), and Sentinel-2 Multispectral Instrument (MSI) for different periods ranging from 2014 to 2024 in Table 3.1 offers a comprehensive temporal perspective on the mangrove ecosystem in Pichavaram. The satellites provide researchers with a wealth of data captured through various spectral bands and resolutions, enabling them to track changes and trends over time with a high level of detail and accuracy.

Table 3.1 – Time Period and Satellite Image

Time period	Satellite image
2014 - 2016	Landsat 8 OLI
2017 - 2024	Sentinel – 2 MSI Level 2A

3.3 DATA PREPROCESSING

3.3.1 NEED FOR PREPROCESSING

Clouds inevitably contaminate optical satellite images due to the physical limitations of sensor imaging systems. The presence of clouds impedes optical satellites from acquiring useful information about the Earth's surface and affects the usability of images in different degrees.

Shadows projected by clouds on the ground surface also contaminate images. The missing information in images caused by clouds and their shadows leads to spatial and temporal gaps in satellite Earth observation data, impacting image usability in various analyses. To overcome this, we need to do cloud masking and composite images.

3.3.2 QUALITY BANDS

Each satellite has specific bands called Quality Assessment (QA) bands contain information specifically about the quality of data in the corresponding spectral bands. To extract information about cloudy pixels and mask them. The information contained in QA bands is stored as bitwise flags. The quality band for the satellites is given in Table 3.2.

Table 3.2 Satellite and QA band

Satellite	QA Band
Landsat 8	Pixel quality assessment band (QA_PIXEL)
Sentinel 2	Quality spatial resolution 60 metres (QA60)

3.3.3 BITMASK FOR QA BANDS

Bitmasks are used to store information about the quality of a pixel in a satellite image. Each bit in the bitmask represents a different quality flag. By setting a specific bit to 1, which denotes the presence of the cloud in that pixel, if the bit value is 0 then it indicates the cloud is absent in that pixel. The specific bits used in a bitmask can vary depending on the satellite sensor. The QA Band for each satellite image has specific bitmask values that are given in Table 3.3

Table 3.3 Quality bands and bitmask values

Quality Band	Bitmask value
QA_PIXEL	Bit 3: Cloud
	Bit 4: Cloud Shadow
QA60	Bit 10: Opaque Clouds
	Bit 11: Cirrus Clouds

3.3.4 ANALYTICS OF COMPOSITE IMAGE CREATION

Sentinel-2 imagery pre-processing involves focusing on bit 10 of the QA60 band to detect cloud presence. By using bitwise AND operations with cloud bit mask (1024), which isolates the 10th bit and effectively identifies cloud-obscured pixels. Pixels with a bitwise AND result of 0 represent clear regions for further analysis, while a result of 1 indicates the presence of clouds.

Table 3.4 Sentinel 2 Bitwise AND Operation

Decimal	Binary Representation										
QA_Value	0	1	1	0	0	0	0	0	0	0	0
CloudBitMask	1	0	0	0	0	0	0	0	0	0	0
Bitwise AND	0	0	0	0	0	0	0	0	0	0	0

In Landsat 8, the approach adapts to the sensor's QA data structure. It focuses on bit 3 for cloud and bit 4 for cloud shadow presence in the QA_PIXEL band. Through bitwise AND operations with the respective bit masks, the pre-processing accurately identifies and masks cloud-affected pixels.

Table 3.5 Landsat Bitwise AND Operation

Decimal	Binary Representation										
QA_Value	0	1	1	0	0	0	0	0	0	0	0
CloudBitMask	0	0	0	0	0	0	0	1	0	0	0
Bitwise AND	0	0	0	0	0	0	0	0	0	0	0

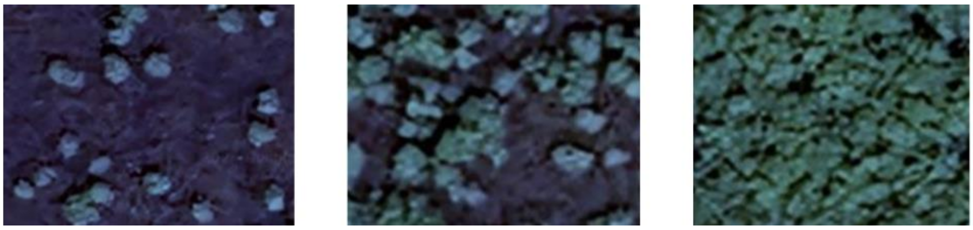
Hence, by performing Bitwise AND operations of QA_Value and CloudBitMask as all resulting bit values are 0, it confirms the absence of clouds in the pixels. The Median Composite is calculated by sorting pixel values from a set of images and selecting the median value for each pixel position, resulting in a composite image highlighting common features.

$$Median\ Composite\ =\ Sort([Image_1, [Image_2, \ldots, [Image_n])\left(\frac{(n+1)}{2}\right) \quad (3.1)$$

3.4 FEATURE EXTRACTION

Feature extraction is a crucial step in the process of analyzing satellite imagery for mangrove classification. It involves extracting relevant information from the image data that can be used to distinguish between mangroves and other land cover types. Features are collected by defining Area of Interest (AOI) for different classes within the satellite image. Each class is assigned a unique colour and spectral information is extracted from these designated areas.

To account for variability within a class, random samples are generated from each AOI. These samples are then split into training and testing sets. By extracting spectral information from the satellite imagery at the sample locations, features that differentiate mangroves from other land cover types can be identified. The classes included for the feature extraction are in Figure 3.2.



(a) (b) (c)

Fig 3.2 Mangrove Density Classes

(a) Sparse / Open Canopy (b) Moderate Canopy (c) Dense / Closed Canopy

3.5 CLASSIFICATION

The satellite imagery data needs to be further prepared for the Convolutional Neural Network (CNN) classification model. The CNN architecture is designed to effectively extract and learn features from satellite imagery data. The prepared satellite imagery data is split into training and validation sets. The CNN model is trained on the training data with the validation set used to monitor the model's performance during training. Typical CNN architectures for image classification tasks include layers such as convolutional layers, dense layers, and dropout layers shown in Figure 3.3.

I. Convolutional Layers:

These layers extract spatial features from the image. They use filters that slide across the image, detecting edges, shapes, and textures. Multiple convolutional layers with increasing complexity can learn higher-level features by combining simpler ones.

II. Pooling Layers:

These layers reduce the dimensionality of the data by down sampling the feature maps. Techniques like max pooling or average pooling can be used.

CNN diagram

III. Dropout Layer:

Dropout layers are used to prevent overfitting in neural networks, especially in deep architectures with many layers. Overfitting occurs when a model becomes too attuned to the training data and performs poorly on unseen data.

IV. Dense Layer:

Dense layers are known as fully connected layer of neural networks, performing the actual classification or prediction tasks. They combine the information extracted from previous layers to make a final decision.

V. Activation Function:

Dense layers typically employ activation functions like Rectified Linear Unit (ReLU) and softmax to introduce non-linearity. This allows the network to learn complex relationships between the features and make non-linear decisions.

VI. Evaluation:

Once trained, the model's performance is evaluated on a separate test set using metrics like accuracy, precision, recall, and F1-score. These metrics measure how well the model classifies different categories in the data. There are four parameters for calculating the metrics True Positive (TP), True Negative (TN), False Positive (FP), False Negative (FN).

Accuracy score is the metrics that is defined as the ratio of true positives and true negatives to all positive and negative observations.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{3.2}$$

Precision is defined as the ratio of correctly classified positive samples to a total number of classified positive samples.

$$Precision = \frac{TP}{TP+FP} \tag{3.3}$$

The recall is calculated as the ratio between the numbers of Positive samples correctly classified as Positive to the total number of Positive samples.

$$Recall = \frac{TP}{TP+FN} \tag{3.4}$$

The F1 score is an important evaluation metric that is commonly used in classification tasks to evaluate the performance of a model. It combines precision and recall into a single value.

$$F1 = \frac{2*Precision*Recall}{Precision+Recall} \tag{3.5}$$

3.6 TIME SERIES ANALYSIS

The time series module plays a crucial role in monitoring the health and dynamics of mangrove ecosystems using satellite imagery. Time Series module involves the process of analyzing changes in mangrove cover over a period of time using satellite imagery. Time series analysis is a statistical method used to analyze data points collected sequentially over time. In environmental science, time series analysis is valuable for understanding how ecological systems change over time, including variations in vegetation cover, land use, and climate variables.

In the context of mangrove conservation and monitoring, time series analysis offers insights into the temporal dynamics of mangrove ecosystems. Analyze data points collected at regular intervals to understand how land coverage changes over time. By analyzing the extent of each class across the years.

CHAPTER 4

IMPLEMENTATION

The implementation of this project involves a series of structured steps to effectively analyze the changes in mangrove extent and health, leveraging satellite imagery and machine learning algorithms.

4.1 PLATFORM USED:

1. Google Colab

Google Colab is a cloud-based Jupyter Notebook environment that allows users to write and execute Python code interactively. It provides free access to GPU and TPU resources, making it suitable for training deep learning models. Colab is well-suited for machine learning, data science, and education tasks. It requires no setup to use and integrates with Google Drive, making it easy to access and share notebooks.

2. Google Earth Engine

Google Earth Engine is a cloud-based platform that combines a vast catalog of satellite imagery and geospatial datasets with powerful analysis capabilities. It enables users, including scientists, researchers, and developers, to detect changes, map trends, and quantify differences on the Earth's surface. The platform provides an API in Python and JavaScript, making it easy to leverage Google's cloud infrastructure for geospatial analysis.

4.2 TOOLS USED

1. Rasterio:

Rasterio is a powerful Python library specifically designed for working with geospatial raster data. Rasterio excels at reading and writing a wide variety of geospatial raster formats, including commonly used ones like Geographic Tagged Image File Format (GeoTIFF) files. It provides a user-friendly and efficient API for accessing and manipulating raster data as NumPy arrays, making it easy to integrate with other Python libraries for scientific computing and data analysis.

2. EarthPy:

EarthPy is a free and open-source Python package designed to simplify working with spatial data, specifically focusing on raster and vector data commonly used in Earth science applications. It builds upon the functionalities of existing powerful libraries like rasterio (for raster data) and geopandas (for vector data) and aims to make common geospatial tasks in Python more user-friendly and efficient.

3. Pandas:

Pandas is a powerful and popular open-source Python library specifically designed for data analysis and manipulation. It offers high-performance, easy-to-use data structures and data analysis tools, making it a fundamental building block for many data science tasks in Python.

4. NumPy:

NumPy (Numerical Python) is a fundamental library for scientific computing in Python. It provides powerful tools for working with multidimensional arrays and matrices, making it essential for various tasks involving numerical computations, data analysis, and machine learning.

5. Keras:

Keras is a high-level, open-source deep learning API written in Python that runs on top of other deep learning frameworks like TensorFlow, JAX, and PyTorch. By leveraging Keras, developers can easily implement these different types of neural networks, ranging from simple sequential models to more complex architectures like CNNs and RNNs, making it a versatile tool for deep learning tasks. Keras provides a building-block approach, offering pre-built neural network layers. These layers can be easily stacked together to form complex neural network architectures.

6. Scikit-learn:

Scikit-learn's metrics module provides a collection of functions for evaluating and comparing the performance of machine learning models. It offers various metrics for different tasks, allowing you to assess how well your model is generalizing to unseen from scikit-learn are valuable tools for evaluating the performance of classification machine learning models.

7. Matplotlib:

Matplotlib is a powerful and versatile Python library for creating static, animated, and interactive visualizations. It's a cornerstone for data visualization in Python, offering a wide range of functionalities for generating various plot types and customizing their appearance. Matplotlib integrates well with other scientific Python libraries like NumPy and Pandas.

4.3 PSEUDO CODE FOR COMPOSITE IMAGE CREATION

Input: Geometry of Area of Interest (AOI) representing the region to be analyzed
Output: Composite satellite image
Steps:

1. begin
2. satellites = {Landsat 8, Sentinel 2}
3. QA_bands = {QA_PIXEL, QA60}
4. for each satellite in satellites

5. imageCollection[satellite] = find_images(AOI, satellite, startDate, endDate)

6. for each imageCollection[satellite] as image

7. imageCollection[satellite] = CloudMask(image)

8. function CloudMask(image)

9. cloud_shadow_bit , cloud_bit <- bit values

10. qa <- image.select('QA Band')

11. mask <- qa.bitwise_and(1<<cloud_shadow_bit).eq(0)

.and(qa.bitwise_and(1<<cloud_bit).eq(0))

12. return image.updateMask(mask).divide(10000)

13. medianComposite <- imageCollection[satellites].median()

14. for each satellite from 1 to satellites.length - 1

15. medianComposite <- medianComposite.addBands(imageCollection [satellites[i]].median())

16. return medianComposite

17. end

4.4 PSEUDO CODE FOR FEATURE EXTRACTION

Input: Area of Interest (AOI) as a single geometry and feature collection curated for each class used for classification

Output: a CSV file with necessary class labels and train/test samples.

Steps:

1. begin
2. class_values=[]

3. class_names=[]
4. class Palette =[]
5. features=[]
6. classValues= []
7. for each classLabel in classLabels
8. classValues[classLabel] <- get_class_value(classLabel)
9. featureCollection <- empty_feature_collection()
10. for each classLabel, classSampleCollection in classSamples
11. classFeatureCollection <- empty_feature_collection()
12. for each sample in classSampleCollection
13. pixelValues <-extract_pixels (studyArea,bands,samplegeometry)
14. sampleFeature <- create_feature(sample.geometry)
15. classFeatureCollection <-classFeatureCollection.add(sampleFeature)
16. trainingSet, testingSet <- split_collection(classFeatureCollection, 0.8)
17. featureCollection <- featureCollection.merge(trainingSet)
18. featureCollection <- featureCollection.merge(testingSet)
19. export_feature_collection_to_csv(featureCollection, "features.csv")
20. end

4.5 PSEUDO CODE FOR CONVOLUTIONAL NEURAL NETWORK

Input: a CSV file for training the model and images for classification

Output: A classified image plotted as a class map

Steps:

1. begin
2. image <- Rasterio.open(imagePath)
3. imageData <- image.read()
4. imageProperties <- { "bands": image.count, "shape": imageData.shape }

5.data <- pd.read_csv(csvFile)

6. x_train, x_test, y_train, y_test <- train_test_split(data.drop("class", axis=1), data["class"], test_size=0.2)

7. x_train <- x_train.reshape(x_train.shape[0], x_train.shape[1], 1)

8. x_test <- x_test.reshape(X_test.shape[0], x_test.shape[1], 1)

9. y_train <- to_categorical(y_train, num_classes=n_classes+1)

10. model.add(Conv1D(filters=32, kernel_size=3, activation="relu")

11. model.add(Dropout(0.2))

12. model.add(Conv1D(filters=64, kernel_size=3, activation="relu"))

13. model.add(Dropout(0.2))

14. model.add(GlobalMaxPooling1D())

15. model.add(Dense(units=128, activation="relu"))

16. model.add(Dense(units=64, activation="relu"))

17. model.add(Dense(units=n_classes+1, activation="softmax"))

18.model.compile(optimizer="adam",loss="categorical_crossentropy", metrics=["accuracy"])

19. model.fit(X_train, y_train, epochs=10, validation_data=(X_test, y_test), batch_size=32)

20. imageStack <- imageData.transpose(0, 2, 1)

21. predictedClasses = model.predict_classes(imageStack)

22. display_class_map(predictedClasses)

23. end

4.6 PSEUDO CODE FOR CALCULATING CLASS COVERAGE IN A CLASSIFIED IMAGE

Input: Mangrove classified image plotted as a class map

Output: Line plots for trend analysis

Steps:

```
1.  begin
2.  image = rasterio.open(imagePath)
3.  bandCount = image.count
4.  height = image.height
5.  width = image.width
6.  shape = (height, width)
7.  for each class in Classes
8.      targetClassPixels = prediction == targetClass
9.      totalPixels = prediction.shape[0] * prediction.shape[1]
10.     targetClassCount = targetClassPixels.sum()
11.     coverage = (targetClassCount / totalPixels) * 100
12.     append coverage in list
13. function PlotTimeSeries(years, data, xLabel, yLabel, title)
14. plt.xlabel(xLabel)
15. plt.ylabel(yLabel)
16. plt.plot(years, list)
17. plt.title(title)
18. plt.show()
19. end
```

CHAPTER 5

RESULTS AND ANALYSIS

5.1 RESULTS OF CLOUD MASKING AND CREATION OF COMPOSITE IMAGE:

After performing cloud masking, the missing pixels are filled by creating a median composite satellite image. The results obtained are given in Figure 5.1 and Figure 5.2.

Fig 5.1 Original Image

Fig 5.2 Composite Image

5.2 CLASS MAPS OBTAINED FROM CLASSIFICATION

After predicting the available classes within the satellite images, the following colour palette used is given in Figure 5.3 and class maps obtained are given in Figures 5.4 to 5.9.

1 - Others (Non – Mangroves Class)

2 - Sparse / Open Canopy

3 - Moderate Canopy

4 - Dense / Closed Canopy

Fig 5.3 Colour Palette for the different class labels

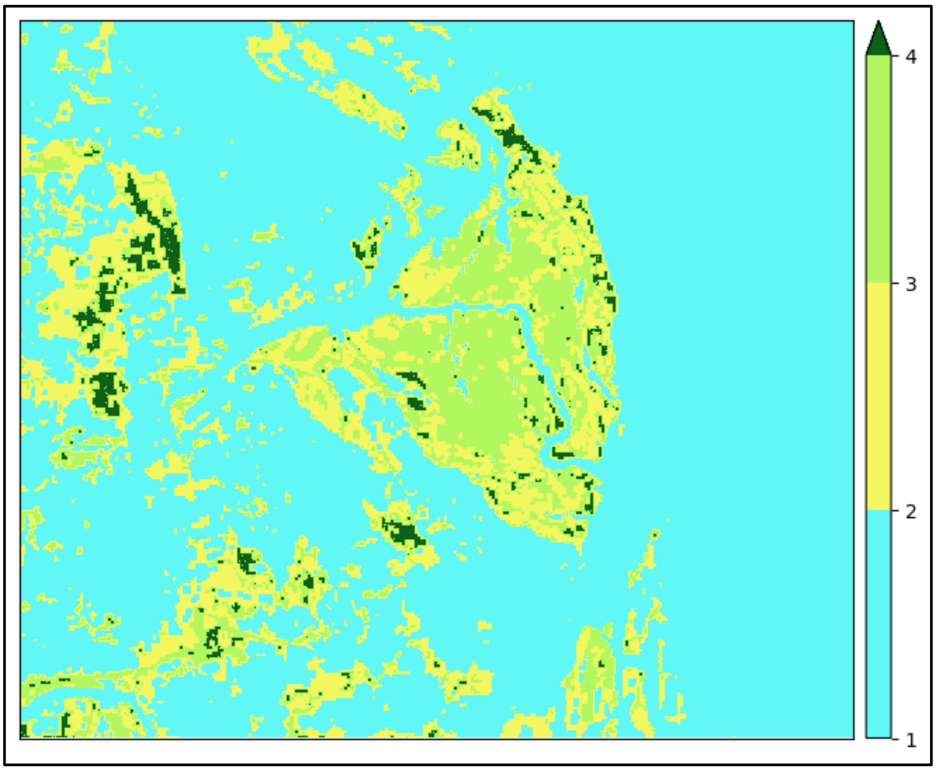


Fig 5.4 Class map for the year 2014

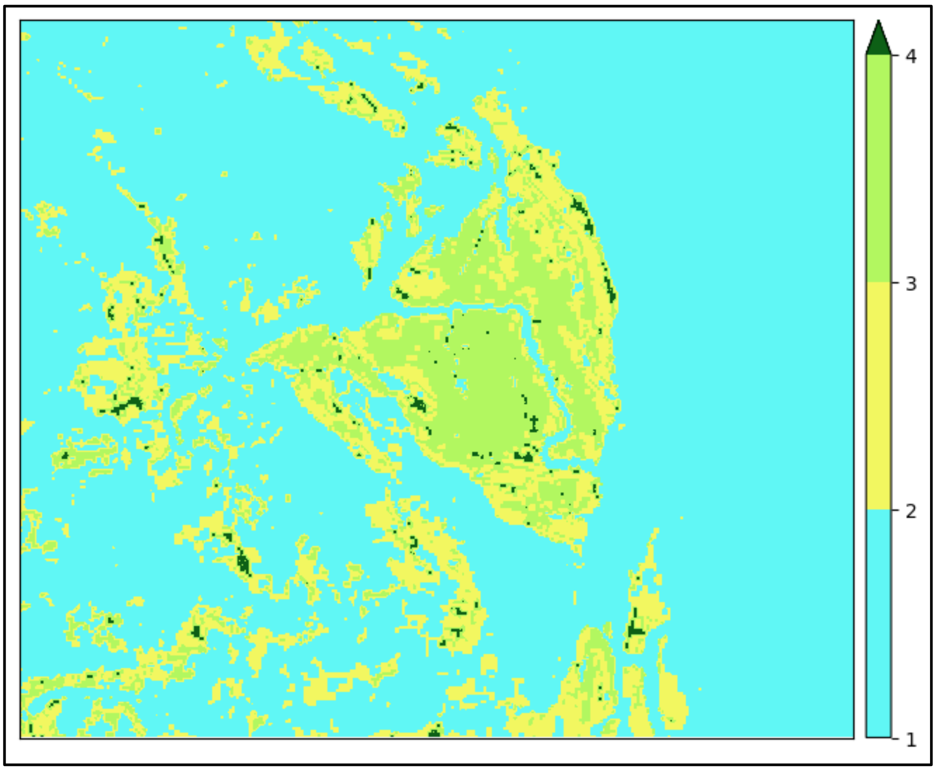


Fig 5.5 Class map for the year 2016

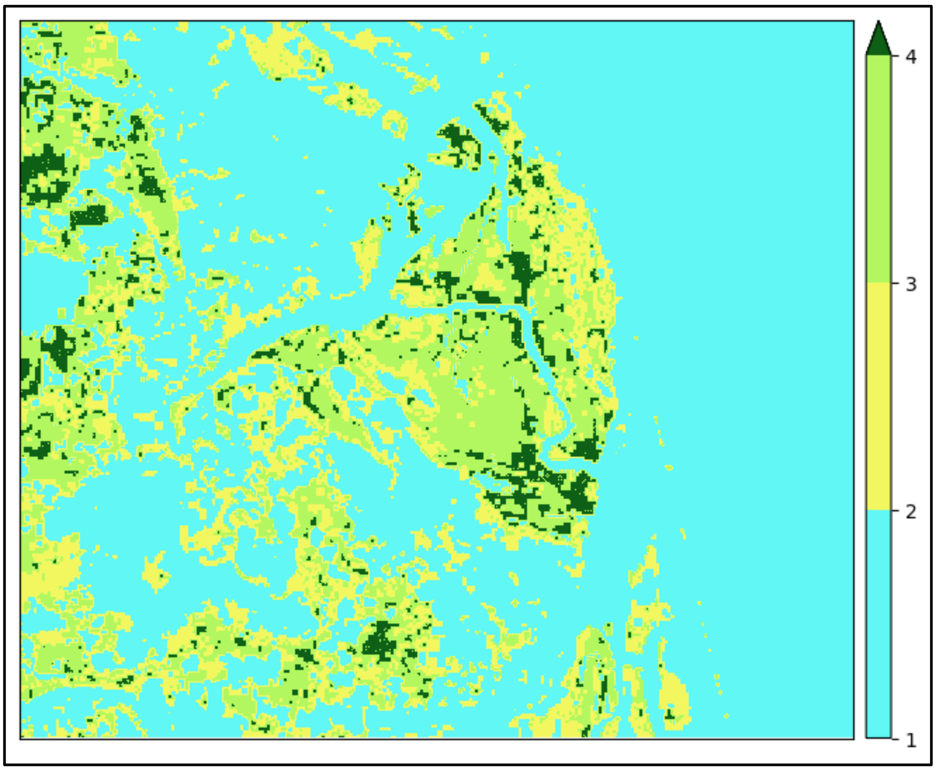


Fig 5.6 Class map for the year 2018

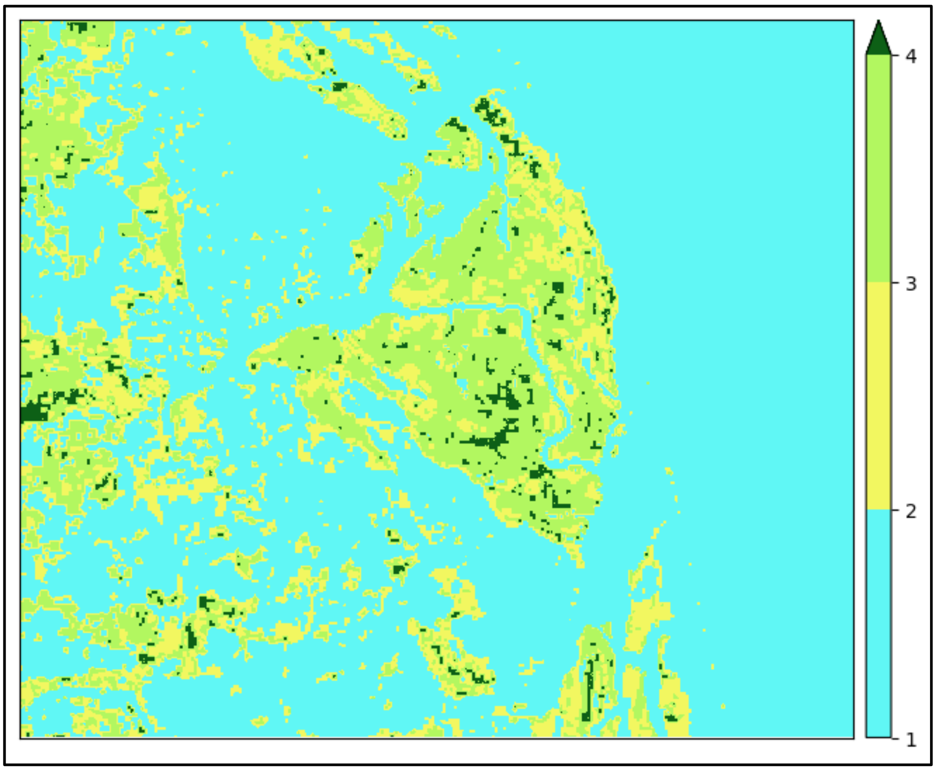


Fig 5.7 Class map for the year 2020

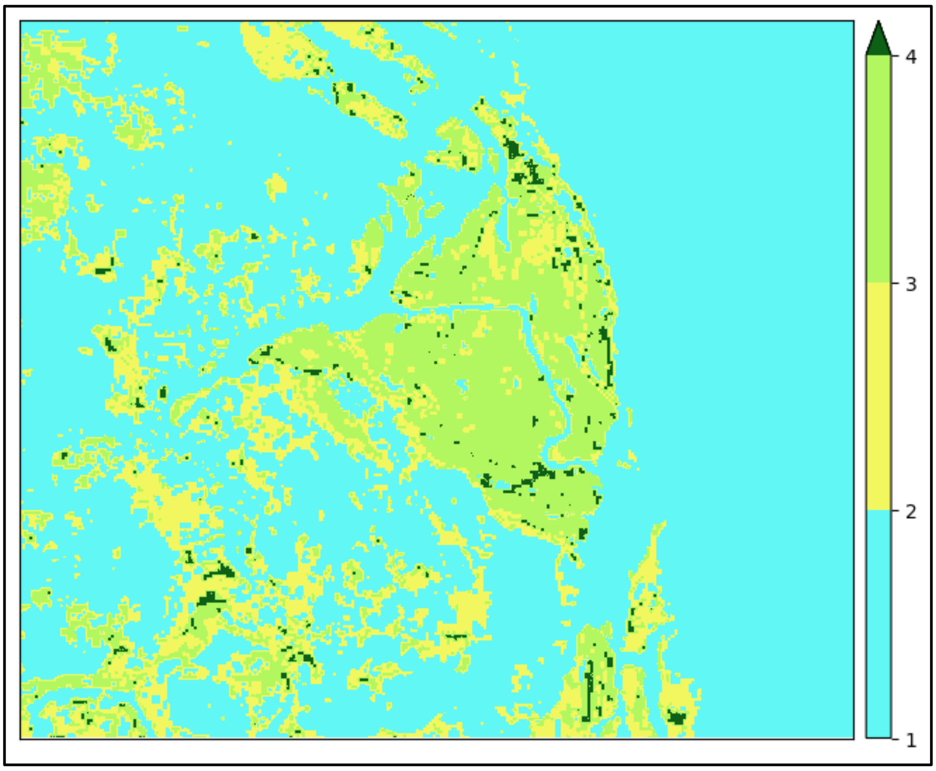


Fig 5.8 Class map for the year 2022

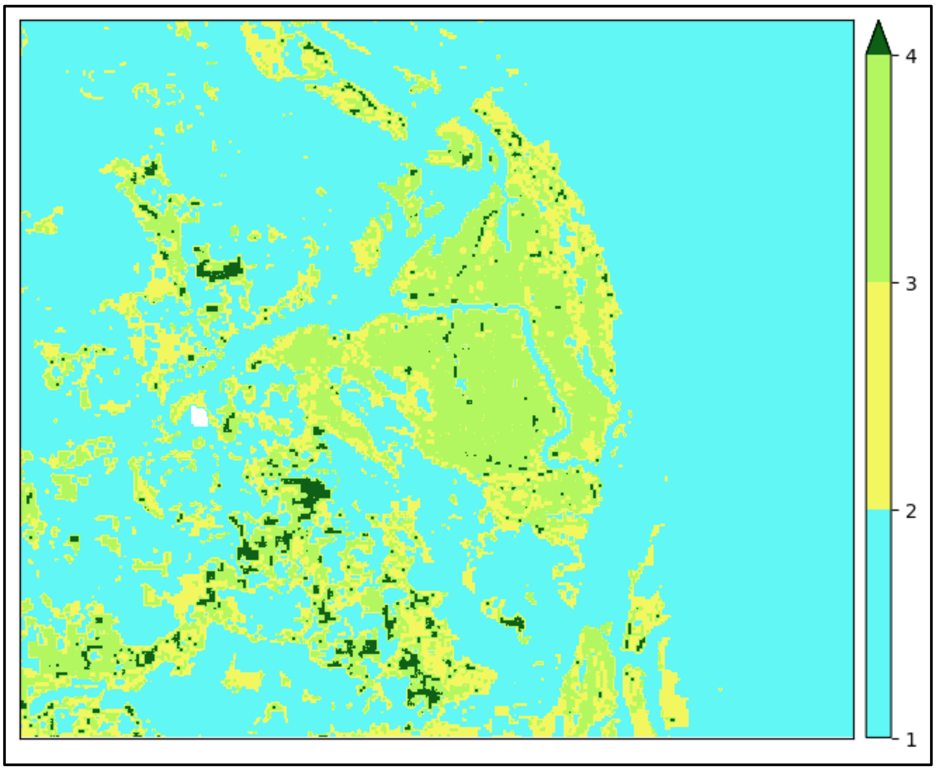


Fig 5.9 Class map for the year 2024

5.3 CHANGE IN THE EXTENT OF SPARSE / OPEN CANOPY COVER OVER TIME

A time series plot is created to visualize the extent of sparse /open canopy cover across a period of years. The plot allows us to identify trends and patterns in how the amount of sparse and open canopy has changed over time. The obtained time series plot is given in Figure 5.10.

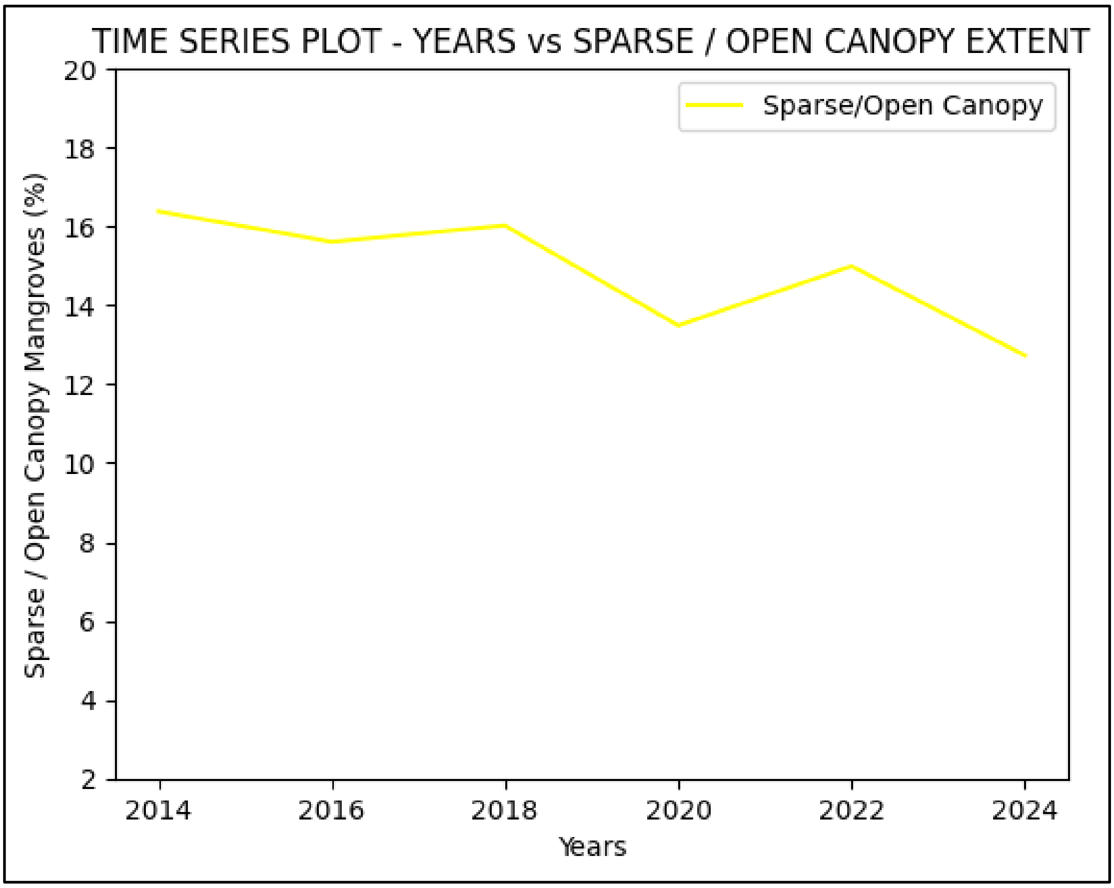


Fig 5.10 Time Series Plot – Years vs Sparse / Open Canopy Mangroves

5.4 CHANGE IN THE EXTENT OF MODERATE CANOPY COVER OVER TIME

To show the extent of moderate canopy cover over the years, a time series plot is plotted. This plot allows us to identify trends and patterns in how the amount of moderate canopy cover has changed over time. The resulting line plot can be seen in Figure 5.11.

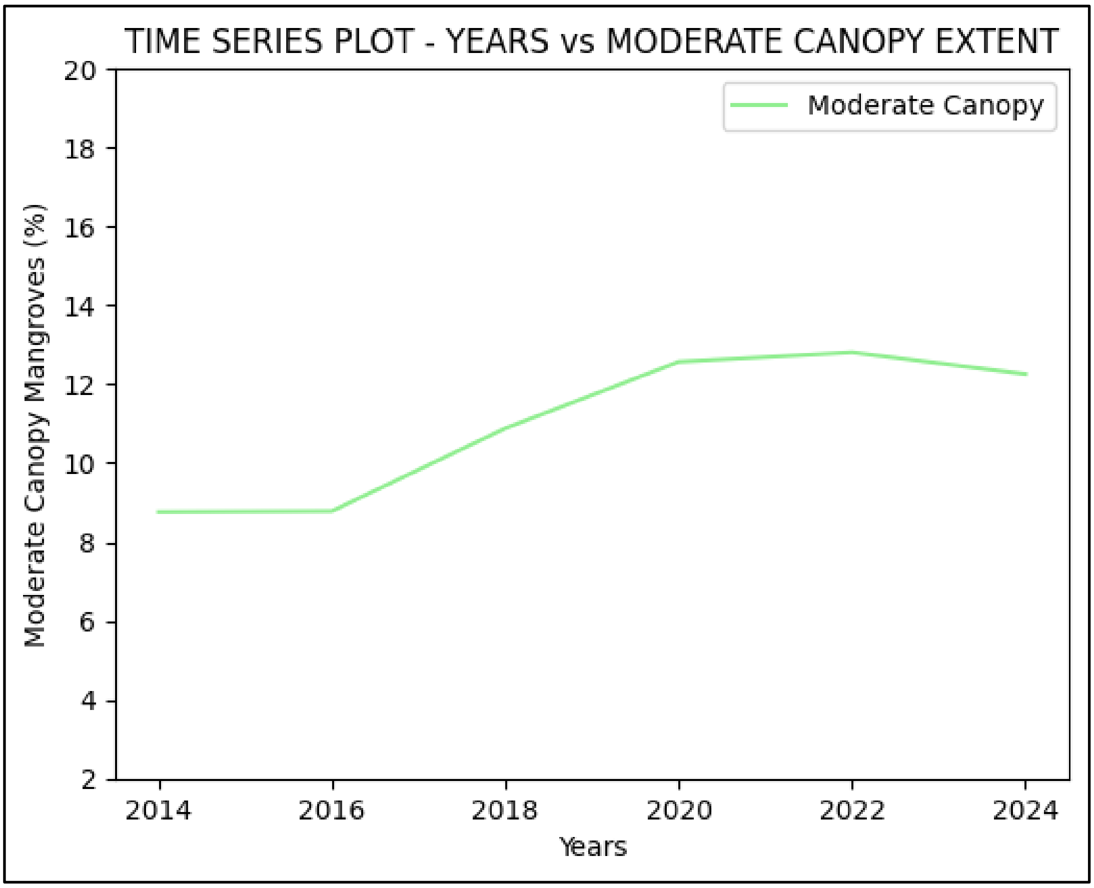


Fig 5.11 Time Series Plot – Years vs Moderate Canopy Mangroves

5.5 CHANGE IN THE EXTENT OF DENSE / CLOSED CANOPY COVER OVER TIME

A time series plot is created to visualize the extent of dense / closed canopy cover across a period of years. The plot allows us to identify trends and patterns in how the amount of dense / closed canopy has changed over time. The obtained time series plot is given in Figure 5.12.

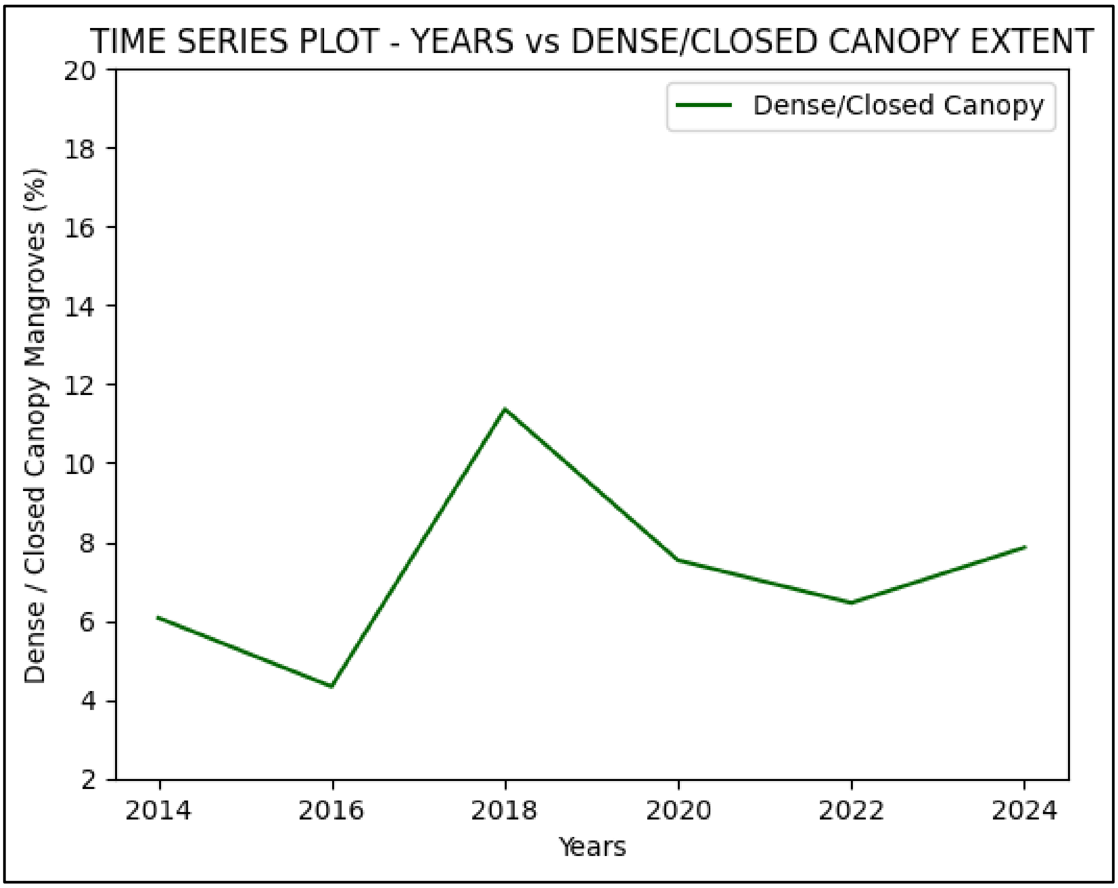


Fig 5.12 Time Series Plot – Years vs Dense / Closed Canopy Mangroves

5.6 CHANGE IN OVERALL MANGROVE EXTENT OVER TIME

Figure 5.13 presents a time series analysis depicting the overall extent of mangrove cover across a range of years. This analysis allows us to assess broad trends in the total mangrove cover within the study area, revealing potential changes in mangrove abundance over time.

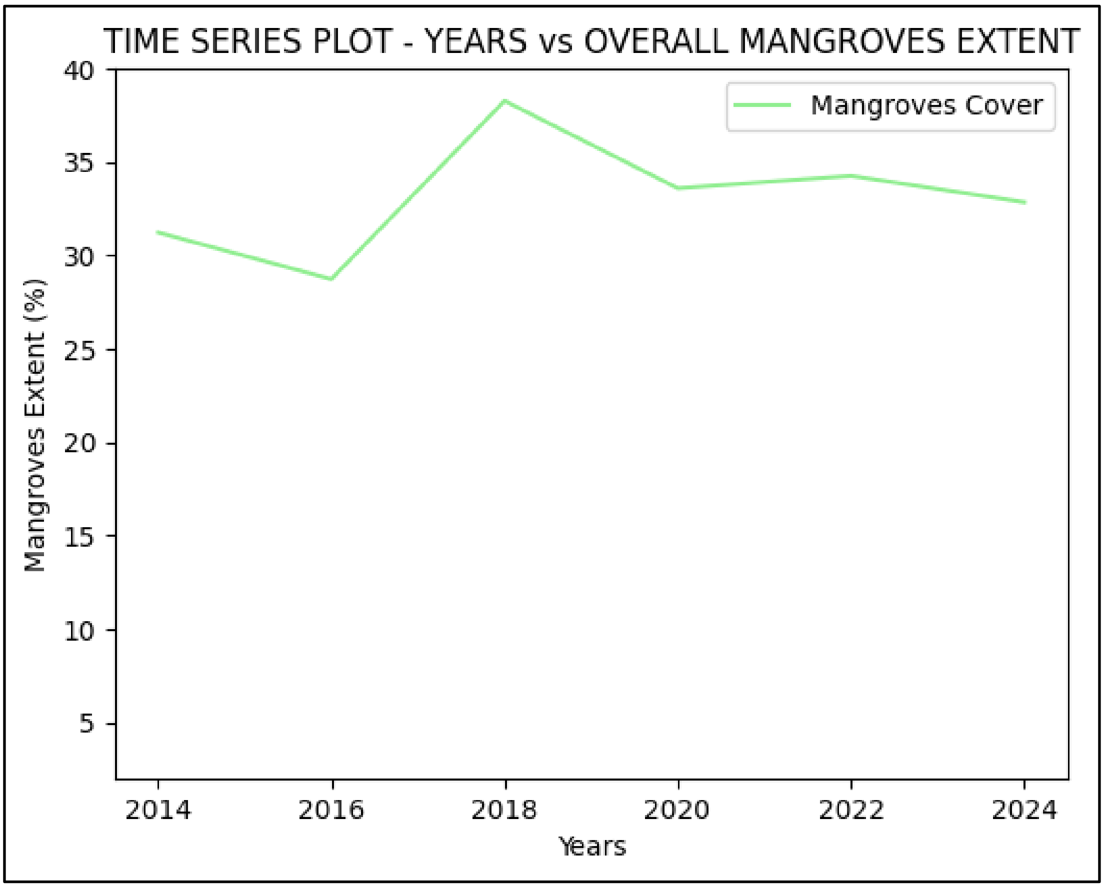


Fig 5.13 Time Series Plot – Years vs Overall Mangroves Extent

5.7 YEARLY TRENDS IN SPARSE, MODERATE, AND DENSE MANGROVE CANOPY COVER

A time series plot is created to visualize the extent of all three distinct types of canopy cover across a period of years. This analysis allows us to identify changes and patterns in the abundance of different canopy densities over time, revealing potential ecological shifts or management impacts. The obtained time series plot is given in Figure 5.14.

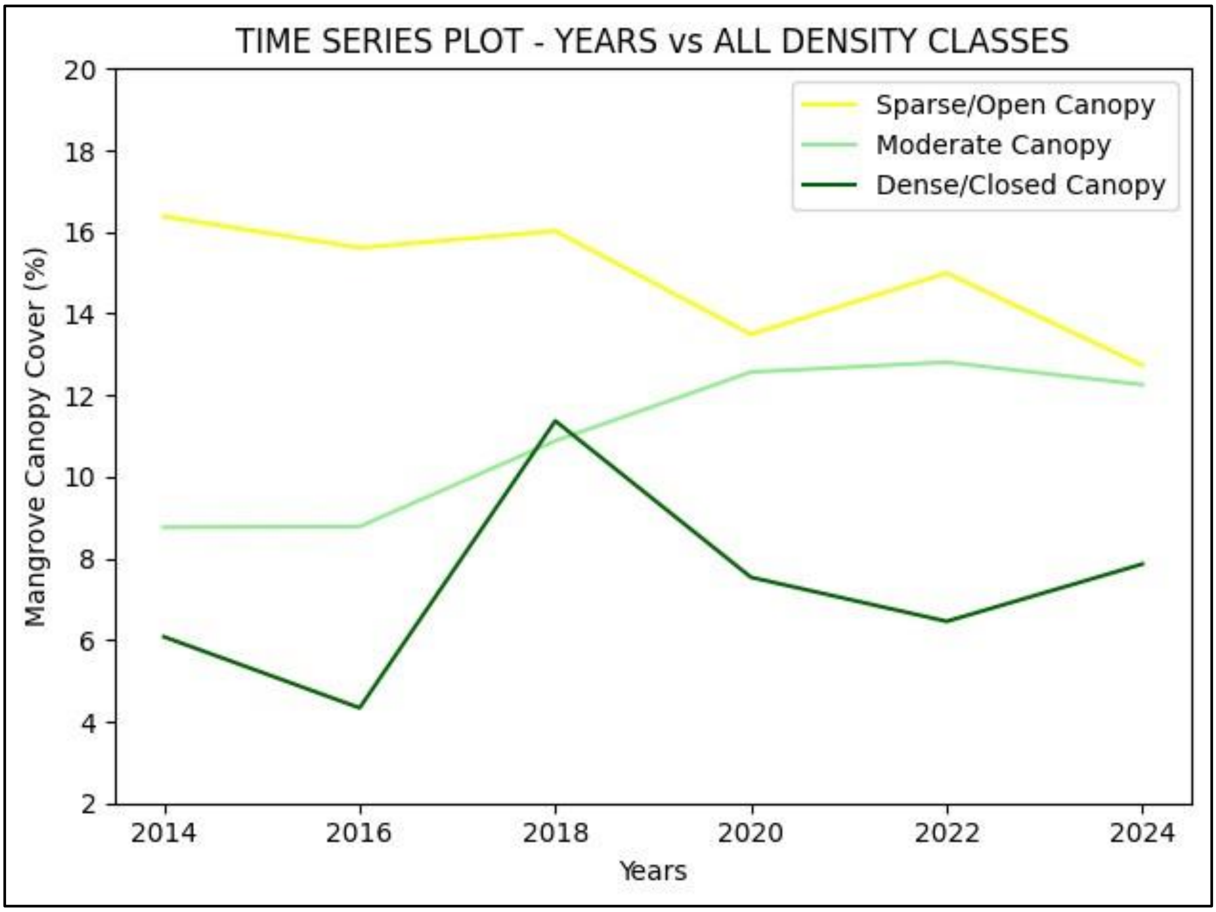


Fig 5.14 Time Series Plot – Years vs All density classes

5.8 EXTENT OF MANGROVES AND NON-MANGROVES BY YEAR

Figure 5.15 presents a time series analysis contrasting the extent of mangrove cover with the extent of non-mangrove cover across a range of years. The plot utilizes two lines: one representing the total mangrove cover and another representing the combined cover of all non-mangrove classes. This visualization allows for comparing and analyzing trends in mangrove abundance relative to non-mangrove cover over time, potentially revealing insights into habitat change dynamics within the study area.

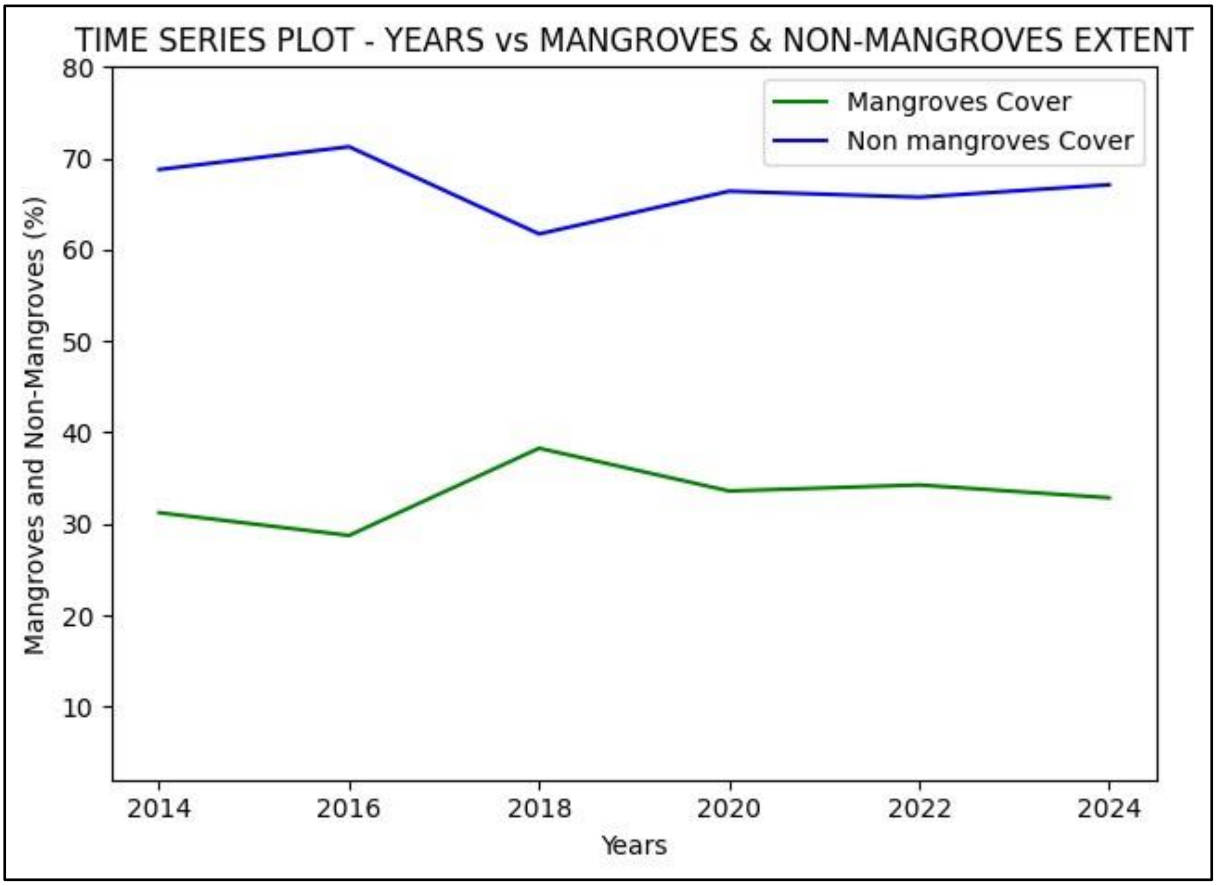


Fig 5.15 Time Series Plot – Years vs Mangroves and Non – Mangroves Extent

5.9 INTERPRETATIONS OF TEMPORAL DYNAMICS OF MANGROVES FROM 2014 TO 2024

I. Sparse Mangroves: In general, there is a decreasing trend in sparse canopy cover from 2014 to 2024, with slight increases in 2018 and 2022.

II. Moderate Mangroves: In overall, there is an increase in moderate cover from 2014 to 2024. However, from 2022 to 2024, a decrease in moderate cover is noted.

III. Dense Mangroves: Between 2014 and 2016, there was a steady decline in dense mangrove cover, which was followed by an increase between 2016 and 2018. However, from 2018 to 2024, there has been a decrease in dense mangrove cover. Nevertheless, from 2022 to 2024, there has been a slight upturn in the percentage of cover.

IV. Extent of Mangrove Cover: Between 2014 and 2016, there was a consistent decline in the coverage of mangrove forests. However, from 2016 to 2018, there was a smooth and steady increase in the coverage of these forests. After 2018, there was a gradual decrease in the percentage of mangroves with a slight increase in coverage during the year 2022.

V. Extent of Non – Mangroves Cover: From 2014 to 2016, there was an increase in non-mangrove cover. From 2016 to 2024, there was a gradual increase in the percentage of non-mangrove cover with slight decreases in 2018 and 2022.

CHAPTER 6

CONCLUSION AND FUTURE WORKS

This project utilized machine learning techniques and satellite imagery to analyze the temporal changes in mangrove forests in Pichavaram, India. A Convolutional Neural Network (CNN) was employed to accurately classify the density of mangroves, which helped to identify changes in the vegetation coverage in the study area from 2014 to 2024. Thus the proposed CNN model tested in Pichavaram presented an overall accuracy of 91% indicating that the model performed well in classifying the images. The findings of this study are of immense importance and can provide valuable insights for conservation efforts. Furthermore, the study can aid in the development of effective strategies aimed at preserving mangrove forests, which are crucial for the ecological balance of the region.

Some of the future works for this project include investigating the influence of external factors like climate change and human activities on the observed changes in mangrove cover. Integrating additional geospatial data sources, such as high-resolution imagery or LiDAR data, to improve the accuracy of mangrove classification and health assessments. Exploring more sophisticated deep-learning architectures might allow us to predict future trends in mangroves. Developing a web-based application for monitoring and management of the health of the mangroves in a real-time scenario.

CHAPTER 7

REFERENCES

[1] Z. Chen, M. Zhang, H. Zhang, and Y. Liu, ”Mapping Mangrove Using a Red-Edge Mangrove Index (REMI) Based on Sentinel-2 Multi spectral Images” in IEEE Transactions on Geo science and Remote Sensing, 2023, vol. 61, pp. 1-11.

[2] R. Chen, R. Zhang, Zhao, C. Wang, Z. Jia “High-Resolution Mapping of Mangrove Species Height in Fujian Zhangjiangkou National Mangrove Nature Reserve Combined GF-2, GF-3, and UAV-LiDAR” in Remote Sens, 15, 2023.

[3] A. Sharifi, S. Felegari, A. Tariq “Mangrove forests mapping using Sentinel-1 and Sentinel-2 satellite images” in Arab J Geosci 15, 2022.

[4] Z. Xue and S. Qian, ”Generalized Composite Mangrove Index for Mapping Mangroves Using Sentinel-2 Time Series Data,” in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 15, pp. 5131-5146, 2022.

[5] R. Barella et al.,”Combined Use of Sen nel-1 and Sentinel-2 for Glacier Mapping: An Application Over Central East Alps,” in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol.15, 2022.

[6] Ghorbanian,A.; Zaghian, S.; Asiyabi, R.M.; Amani, M.; Mohammadzadeh, A.; Jamali, S. ”Mangrove Ecosystem Mapping Using Sentinel-1 and Sentinel-2 Satellite Images and Random Forest Algorithm in Google Earth Engine” in Remote Sens. 13, 2565, 2021.

[7] R. Zhang et al., ”A Comparison of Gaofen-2 and Sentinel-2 Imagery for Mapping Mangrove Forests Using Object-Oriented Analysis and Random Forest,” in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 14, pp. 4185-4193, 2021.

[8] Yancho, J.M.M.; Jones, T.G.; Gandhi, S.R.; Ferster, C.; Lin,A.; Glass, L. ”The Google Earth Engine Mangrove Mapping Methodology (GEEMMM)” in Remote Sens. 12, 3758, 2020.

[9] M. Sheykhmousa, M. Mahdianpari, H. Ghanbari, F. Mohammadimanesh,P. Ghamisi and S. Homayouni, ”Support Vector Machine Versus Random Forest for Remote Sensing Image Classification: A Meta-Analysis and Systematic Review,” in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 13, pp. 6308-6325, 2020.

[10] Y. Zhu et al., ”Esmating and Mapping Mangrove Biomass Dynamic Change Using WorldView-2 Images and Digital Surface Models,” in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 13, pp. 2123-2134, 2020.

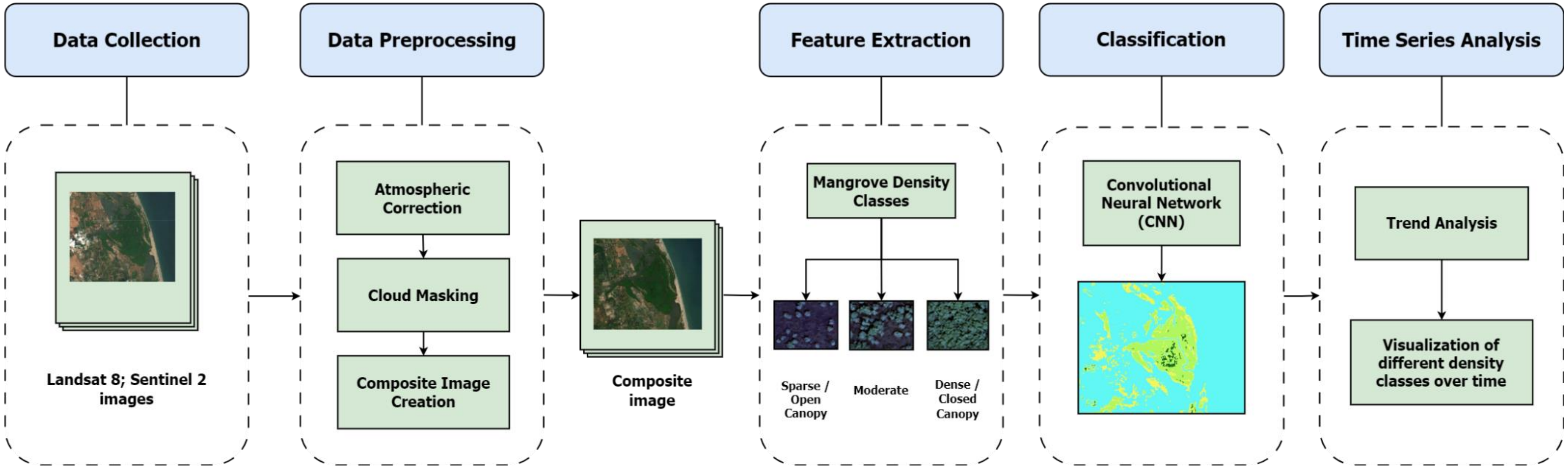


Fig 3.1 Proposed Architecture Diagram

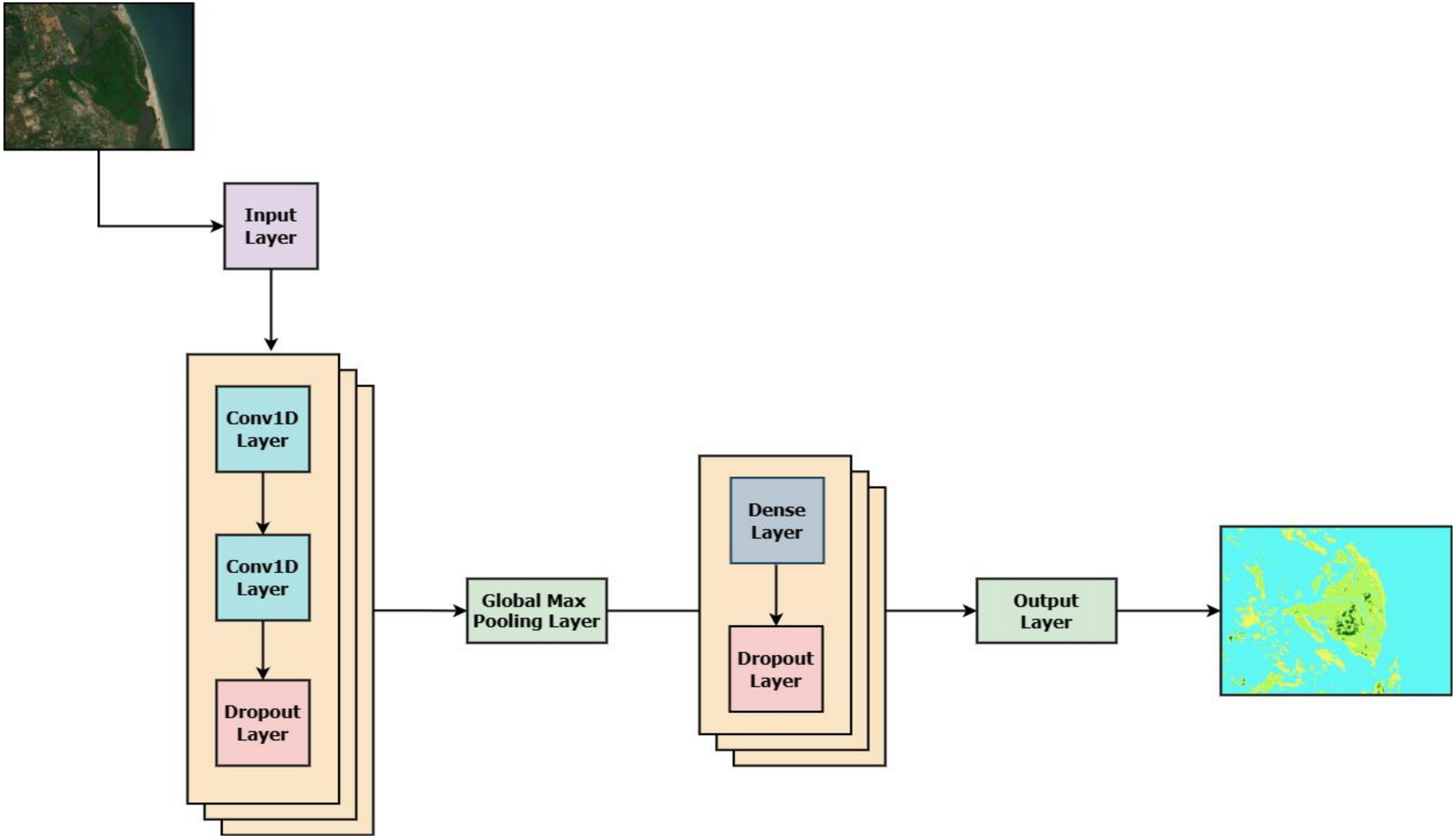


Fig 3.3 Proposed Convolutional Neural Network (CNN) Architecture