DELINEATION OF SMALL WATER BODIES USING SENTINEL 2 DATA

Abhinash Anilkumar

MECHANICAL ENGINEERING DEPARTMENT

INDIAN INSTITUTE OF TECHNOLOGY

KHARAGPUR – 721302

APRIL 2024

DELINEATION OF SMALL WATER BODIES USING SENTINEL 2 DATA

Project Report submitted to

Indian Institute of Technology Kharagpur
for the partial fulfillment of the requirements
for the award of the degree of

Master of Technology

in

Manufacturing Science and Engineering

by

Abhinash Anilkumar

(Roll No. : 20ME31056)

Under the guidance of

Prof. Chandranath Chatterjee

MECHANICAL ENGINEERING DEPARTMENT

INDIAN INSTITUTE OF TECHNOLOGY

KHARAGPUR – 721 302

APRIL 2024

DELINEATION OF SMALL WATER BODIES USING SENTINEL 2 DATA

Thesis submitted by

Abhinash Anilkumar

(Roll No.:20ME31056)

Approved by

Prof. Chandranath Chatterjee ***Prof. Madan Kumar Jha***

C. Chatterjee

(Supervisor) ***(Head of the Department)***

AGRICULTURAL AND FOOD ENGINEERING DEPARTMENT

INDIAN INSTITUTE OF TECHNOLOGY

KHARAGPUR – 721 302

ACKNOWLEDGEMENTS

I want to express my sincere gratitude to my Bachelor's thesis project supervisor, **Prof. Chandranath Chatterjee,** Agricultural and Food Engineering Department, Indian Institute of Technology Kharagpur, for his immense guidance and support throughout the thesis project. His precious insights proved to be extremely beneficial to me throughout my project.

I would like to express my heartfelt gratitude to Prof. Madan Kumar Jha, Head of the Agricultural and Food Engineering Department, Indian Institute of Technology Kharagpur, for providing the necessary facilities to complete this research work, as well as for all of his thoughtful considerations and timely interventions.

My special thanks to all the faculty of the department of Mechanical Engineering, Indian Institute of Technology, Kharagpur for the knowledge they have bestowed upon me during my academic career. Without their motivation, support, and teachings, I would not have been able to complete this project. I would like to acknowledge the invaluable contributions of **Saidutta Mohanty**, whose collaborative spirit, constructive discussions, and shared knowledge have enhanced the depth and scope of this project. I would also like to express my gratitude to the panel evaluating my thesis project for giving me the opportunity to present my work in front of such an esteemed panel. Their valuable feedback will be extremely beneficial for my future endeavours.

In the end I would like to thank the Agricultural and Food Engineering Department, Indian Institute of Technology, Kharagpur for providing me with the opportunity to work on this Project.

LIST OF CONTENTS

LIST OF TABLES		
Table	**Title**	**Page no.**
3.2	Band description of Sentinel-2 MSI, Level-1C	17-18
4.1	Accuracy assessment on test set	27

LIST OF FIGURES

ABSTRACT

Freshwater resources are crucial yet vulnerable, especially in regions like Nayagarh, facing threats due to various factors. This analysis delves into the detection of water bodies within the Odagaon block of Nayagarh district in Odisha, employing the advanced Deep learning model for segmentation called as the Nested-UNet or UNet++, in conjunction with the Machine Learning algorithm, XGBoost and spectral index, Modified Normalized Difference Water Index (MNDWI). The study rigorously assesses the effectiveness of these methodologies in accurately identifying water bodies and compares their performance based on the prediction on test data. The findings prominently demonstrate the superior accuracy of Deep learning over any other method, illustrating their remarkable efficiency in delineating water features within the region. In contrast, MNDWI showcases a moderate performance, identifying water bodies with a relatively smaller coverage compared to the UNet++ model.

The XGBoost model tends to show results that are similar to those of MNDWI, even though it should have performed better than MNDWI.

CHAPTER-1

INTRODUCTION

Water bodies, the lifeblood of our planet, encompass a diverse array of aqueous environments that shape and sustain ecosystems worldwide. From serene lakes to meandering rivers and expansive oceans to trickling streams, these bodies of water play a vital role in supporting life, regulating climate, and nurturing biodiversity. Water bodies play a central role in our planet's ecosystems, serving as crucial habitats for numerous species. They support a diverse range of life forms and facilitate the flourishing of intricate ecological networks. Natural environments are essential to helping human societies by providing indispensable resources encompassing sustenance, transportation, recreation, and a wellspring of natural beauty and inspiration. The sheer variety of water bodies is astonishing, each with its unique characteristics, ecological significance, and interactions with surrounding environments. Exploring these diverse water systems unveils a tapestry of interconnected ecosystems, demonstrating their profound impact on our planet's health.

The recent advances in satellite remote sensing technology, particularly the launch of the Sentinel-2 mission by the European Space Agency (through the European Copernicus programme), have provided unprecedented opportunities for mapping and monitoring water bodies at a global scale. The Sentinel-2 satellite system provides high spatial resolution (10 m) images with high temporal revisit time (5 days to a few weeks). The Sentinel-2 satellite is the second satellite in the Copernicus Programme of the European Commission (EC) and the European Space Agency (ESA) (Drusch et al., 2012). The main task of the Sentinel-2 satellite is to realize global land surface resolution multispectral imaging. Compared with the

Landsat-8 image, the Sentinel-2 image has a higher spatial and temporal resolution. In terms of water extraction using the water body indices (NDWI, MNDWI), the accuracy of the Sentinel-2 image is better than the Landsat-8 image, especially in river extraction (Yang et al., 2017).

Accurately identifying small water bodies is always a difficult problem because of the complicated terrain, classification methods and remote sensing data. Because of its simplicity and convenience, the water index is the most commonly used water identification method. Among them, the Normalized Difference Water Index (NDWI) (Mcfeeters, 1996), the Modified Normalized Difference Water Index (MNDWI) (Xu, 2006), and the Automated Water Extraction Index (AWEI) (Feyisa et al., 2014) are most widely used indices. The NDWI normalizes the green and near-infrared bands to enhance the information to separate water pixels better from other features. However, it was found to have large errors in the case of urban areas. MNDWI ameliorates this problem by using mid-infrared bands. What these water indices have in common is that they all use differences in the reflectivity of water at different wavebands to enhance water information. The water pixel is then classified by setting a threshold.

The utilization of Machine Learning (ML) and Deep Learning (DL) models in identifying water bodies within satellite imagery represents significant progress in remote sensing technology. These models demonstrate remarkable capabilities in automating the process of identifying, classifying, and analysing water bodies with notable precision and effectiveness. When provided with large sets of labelled satellite imagery data, machine learning algorithms can identify intricate patterns and features related to different categories of water bodies. They demonstrate proficiency in extracting significant data from these images, thereby enabling the identification of bodies of water such as lakes, rivers, reservoirs, and other water features through the analysis of acquired attributes such as colour, texture, shape, and spatial arrangement.

1.1 OBJECTIVES OF THIS STUDY

- A comparative analysis between 3 methods for small water body identification. A UNet++ model and an XGBoost model has been trained to identify water bodies. Labels for these models have been prepared using shapefiles with water body polygons using softwares like Google Earth Engine (GEE), QGIS, Google Colab (using python).
- Identification and prediction of small water bodies and their evaluation using metrices like Accuracy, F1-Score, and IoU.

CHAPTER – 2

REVIEW OF LITERATURE

This section contains the reviews of the literature works used as a reference for this work.

Many indices have been developed to extract surface water areas or flood inundation extent. Mcfeeters et al. (1996) presented the NDWI (Normalized Difference Water Index) which helps delineate open water and enhance its presence in remote sensing images. NDWI is calculated as Equation:

$$\boldsymbol{NDWI = (G - \mathrm{N}IR)/\ (G + \mathrm{N}IR)} \qquad (2)$$

Where G is a green band and NIR is a near-infrared band. Though the NDWI could suppress and remove non-water features to a large degree, it failed to suppress built-up land signals efficiently. As a result, the extracted features could be a mixture of water and built-up land noises (Xu 2006).

Short-wave infrared (SWIR) band was found to be able to reflect some subtle characteristics of water, and so the NIR band in NDWI was replaced by the SWIR band when the modified normalized difference water index (MNDWI) was proposed by Xu (2006). It is widely accepted that MNDWI is more stable and reliable than NDWI, because the SWIR band is less sensitive to concentrations of sediments and other optical active constituents within the water than the NIR band.

$$\boldsymbol{\mathrm{MNDWI} = (G - \mathrm{SW}IR)/\ (G + \mathrm{SW}IR)} \qquad (3)$$

Machine learning feature pixel-based pattern recognition analysis, including supervised and unsupervised classification techniques. The supervised methods mainly include neural network, support vector machine (SVM), logistic regression, and random forest, and the unsupervised classification methods mainly include K-means clustering and Iterative Self-Organizing Data Analysis Technique (ISODATA) clustering methods. The machine learning algorithm has been widely used in remote sensing water extraction due to its high accuracy.

Due to the limitations of pixel-based classification methods, such as the salt and pepper phenomenon in classification results, object-based classification techniques have been increasingly applied in remote sensing classification in recent years. Many successful cases of water body extraction using OBIA methods have been reported. Given that urban functional zones (UFZs) are composed of diverse geographic objects, Du et al. presented a novel object-based UFZ mapping method using Very high-resolution (VHR) remote sensing images. Based on object-oriented analysis technology and multi-source data, Guo et al. proposed a multi-level classification scheme based on goals and rules to study the changes in glacier environments.

In addition, some studies have also used synthetic aperture radar (SAR) data to monitor the surface dynamics, because these data are insensitive to clouds; the surface water area can be extracted from SAR data based on textural analysis, change detection, automatic segmentation, and classification.

The state-of-the-art models for image segmentation are variants of U-Net and fully convolutional networks (FCN). Despite their success, these models have two limitations:

(1) Their optimal depth is apriori unknown, requiring extensive architecture search or inefficient ensemble of models of varying depths;

(2) Their skip connections impose an unnecessarily restrictive fusion scheme, forcing aggregation only at the same scale feature maps of the encoder and decoder sub-networks. To

overcome these two limitations, UNet++ was proposed, a new neural architecture for semantic and instance segmentation, by

(1) alleviating the unknown network depth with an efficient ensemble of U-Nets of varying depths, which partially share an encoder and co-learn simultaneously using deep supervision; (2) redesigning skip connections to aggregate features of varying semantic scales at the decoder sub-networks, leading to a highly flexible feature fusion scheme; and (3) devising a pruning scheme to accelerate the inference speed of UNet++.

CHAPTER – 3

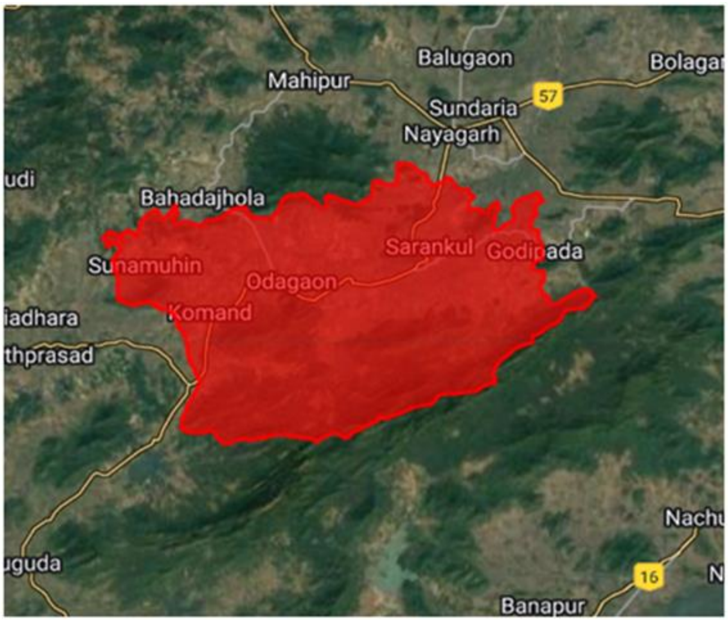
STUDY AREA AND METHODLOGY

3.1 STUDY AREA

An administrative District of Odisha, Nayagarh District was created on 1st April 1993 when the erstwhile Puri District was split into three distinct Districts. The District is bounded by Cuttack District on the North, Kandhamal District on the West, Ganjam District on the South and Khordha District on the East. Nayagarh district, situated southeast of Odisha (India), is a part of the Mahanadi River basin in eastern India. Odagaon is a block and Notified area council (geographical coordinates as 20°00'57.65"N 84°59'15.70"E) in the Nayagarh district of Odisha, India (Figure 3.1). It is around 100km away from the capital city of Bhubaneswar. As per Census of India, 2011, Odagaon Urban has the population of 5,401 with 1,240 households. As nearby rural areas have been included in new jurisdiction boundary, the population as of 2011 as by the Odagaon Notified Area Council (NAC) is 11,941 with 2,758 households. Current data from NAC shows that the population has been increased to 15,800 with 3,420 households. The current population has been used for the preparation of this SFD lite report. The Urban Local Body (ULB) has a total of 15 wards which is being administered by NAC for the provision of civic facilities. The total area under the jurisdiction is 15.14 sq. km. The NAC has been declared by ODF+ in 2020. The climate is generally dry tropical. The actual rainfall and the normal rainfall of the Nayagarh district recorded are found to be 1750.7mm and 1449.1 mm respectively. Odagaon south has a hilly terrain through which water is supplied to half of the households of NAC. The average elevation is 32 m. The water is supplied by NAC through community water taps to some of the households. Rest of the households are dependent upon hand pumps, wells and bore wells.

In summer months, NAC supplies water through water tankers where piped water supply is not there. The water is supplied for two hours in the morning. People here face issues related to water supply due to groundwater unavailability. The average depth of groundwater is around 30ft (9.14m). While post monsoon, water is available within 25ft (7.62m) depth as water level rises by 15.15%2. The water table is shallow in most part of the district. The pre-monsoon depth to water level values are generally in the range of (5m to 10m). Odagaon has around 15 number of water bodies, two nullah (large open drain) named Sagar and Bedha. The predominant occupation of residents of Odagaon is paddy and lentil cultivation. As per Central Ground Water Board (CGWB) report, there is no appreciable change in water levels for 10 years till 2013. Sanitation facilities are provided by the NAC all to the wards.

Figure 3.1: Odagaon



3.2 DATA

3.2.1 DATA PREPARATION

Sentinel-2 is an Earth observation mission from the Copernicus Programme that systematically acquires optical imagery at high spatial resolution (10 m to 60 m) over land and coastal waters. Sentinel-2 MSI: Multispectral Instrument, Level-1C was used to prepare the input data for the machine learning models. The Sentinel-2 data contains 13 spectral bands representing the top of atmosphere reflectance scaled by 10000. The bands and their spatial resolution are described in Table 3.2.

Table 3.2: Band description of Sentinel-2 MSI, Level-1C

Name	Pixel Size(meters)	Wavelength (S2A/S2B)	Description
B1	60	443.9nm / 442.3nm	Aerosols
B2	10	496.6nm / 492.1nm	Blue
B3	10	560nm / 559nm	Green
B4	10	664.5nm / 665nm	Red
B5	20	703.9nm / 703.8nm	Red edge 1
B6	20	740.2nm / 739.1nm	Red edge 2
B7	20	782.5nm / 779.7nm	Red edge 3
B8	10	835.1nm / 833nm	Near infrared (NIR)
B8A	20	864.8nm / 864nm	Red edge 4
B9	60	945nm / 943.2nm	Water vapour
B10	60	1373.5nm / 1376.9nm	Cirrus

B11	20	1613.7nm / 1610.4nm	Short wave infrared 1
B12	20	2204.4nm / 2185.7nm	Short wave infrared 2
QA10	10	-	Always empty
QA20	20	-	Always empty
QA60	60	-	Cloud mask

Among the bands, B2, B3, B4 and B11 were used in this study as the input features to the UNet++ model and XGBoost. November was chosen for data acquisition due to the reduced cloud cover this month.

3.2.2 DATA COLLECTION

Google Earth Engine (GEE) is a cloud-based platform developed by Google that provides a large-scale geospatial data repository and a set of tools for analysing and processing remote sensing data. It is designed to facilitate planetary-scale environmental data analysis.

The image of the study region from November 2023 was taken from Google Earth Engine. A median image was chosen from the collection of images in November 2023. A single satellite image covered the entire region of study, so mosaicking was unnecessary. Mosaicking refers to spatially assembling image datasets to produce a spatially continuous image.

3.2.3 DATA PREPROCESSING

Google Earth Engine offers built-in function to enhance data quality by identifying and masking out cloudy and cirrus regions within images using the Quality Assessment (QA) band. This preprocessing step is crucial as it ensures that undesirable cloud cover is removed, thereby improving the reliability and accuracy of subsequent analyses or visualizations. Additionally, after masking out these regions, the function further enhances the image quality by normalizing the pixel values.

3.2.4 DATASET PREPARATION

Given the substantial area under consideration, the data was segmented into tiles to facilitate processing by the models. Each tile was transformed into a numpy array to store the band values of the pixels within the tile. Moreover, to ensure data integrity and consistency, any NaN (Not a Number) values were removed from the numpy arrays. Subsequently, these preprocessed numpy arrays served as the input data for training the ML and DL models, enabling robust analysis and prediction tasks within the Google Colab environment.

Data augmentation involves enhancing a dataset by creating additional variations or instances from existing data. This process is commonly used in deep learning tasks to increase the diversity and quantity of available data, which can lead to improved model performance and generalization. Techniques such as rotation, flipping, cropping, scaling, and adding noise are applied to original data samples to generate augmented versions. As a result of this augmentation process, the dataset was organized into 1000 tiles of data. Each of these tiles represented a portion of the augmented dataset and was subsequently converted into a numpy array format, allowing for efficient storage and manipulation of the data for further analysis.

3.2.5 TRAINING LABELS

Preparing labels for a training dataset involves assigning target values or categories to data samples. In the case of detecting water bodies, this process can be done by marking water areas manually in tools like Google Earth Pro. Using Google Earth Pro, the study region's shapefile is projected onto the map. Water bodies within this region are manually labeled by outlining them with tools provided by Google Earth Pro. Once all water bodies are marked, the labels are exported as shapefiles from Google Earth Pro. Each labeled water body is represented as a polygon in the shapefile. In Google Colab, the shapefile is rasterized, converting vector data into a raster image where pixels within water body polygons are assigned a value of 1, while pixels outside are assigned 0. The rasterized image is split into tiles. Each tile corresponds to a portion of the study region. The rasterized shapefile is converted into a numpy array. Each pixel's value represents whether it falls within a water body (1) or not (0).

Figure 3.2 (a): Odagaon with labels

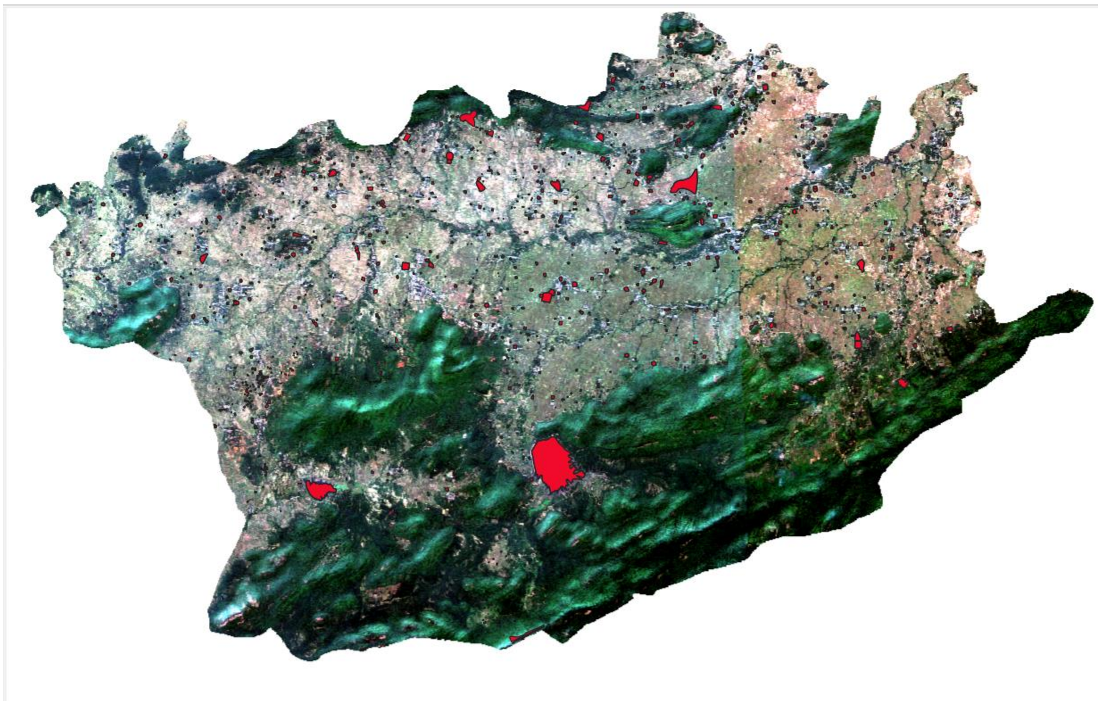


Figure 3.2 (b): Odagaon labels

Figure 3.2 (c): Odagaon part labels

3.3 METHODOLOGIES USED

The labels generated had two possible values, 0 and 1, indicating non-water and water, respectively. Hence, the models required were for a binary classification problem.

3.3.1 UNET++ / NESTED UNET

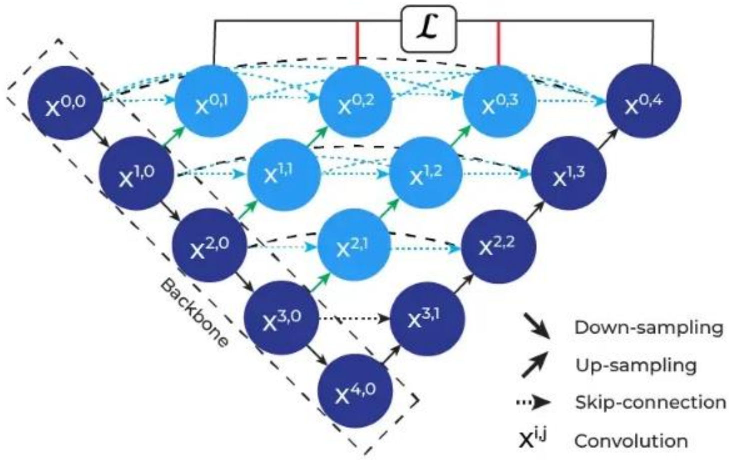
UNet++ or Nested UNet is a deep learning architecture that was introduced in 2019. In UNet, the encoder part captures high-level features from the input image through a series of convolutional and pooling layers, while the decoder part upsamples these features to generate a dense segmentation map. However, there can be a semantic gap between the encoder and decoder features, meaning that the decoder may struggle to reconstruct fine-grained details and produce accurate segmentation.

UNet++ introduces the concept of nested skip pathways to bridge this semantic gap. It adds additional skip connections between the encoder and decoder blocks at multiple resolutions. These connections allow the decoder to access and incorporate both low-level and high-level features from the encoder, providing a more detailed and comprehensive understanding of the image.

UNet++ architecture:

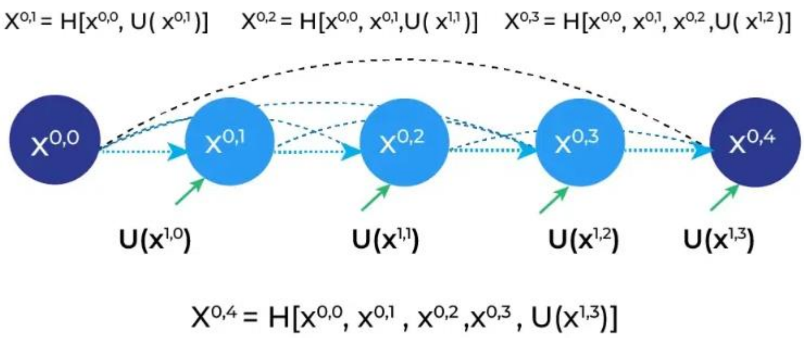
UNet++ architecture is a semantic segmentation architecture based on UNet. They introduced two main innovations in the traditional UNet, architecture namely, nested dense skip connections and deep supervision. In their research, they found that using nested dense skip connections bridges the semantic gap between encoder and decoder feature maps and improves the gradient flow. They also found that using deep supervision enhances the model performance by providing a regularization to the network while training.

Figure 3.3 (a): UNet++



The above figure illustrates the architecture design of UNet++. The figure illustrates the nested encoder and decoder architecture of the UNet++ architecture. We can notice that instead of a traditional skip connection, the feature map from the lower level is also convoluted with the upper-level feature and then the new combined feature data is then passed further. The basic idea behind UNet++ is to bridge the semantic gap between the feature maps of the encoder and decoder before the fusion. For example, the semantic gap between (X0, 0, X2, 2) is bridged using a dense convolution block with three convolution layers. The black dotted skip connection indicated the original skip connection present in U-Net architecture while the blue dotted skip connection indicated the newly introduced nested skip connection. It must be noted that before convoluting the lower level feature map it is upscaled to match the number of channels in that level. The figure also illustrated how deep supervision is also applied to the output of nodes X0,1, X0,2, X0,3, and X0,4 to improve the model learning while training. Deep supervision is an optimization technique where you optimize the model on the final as well as hidden layers (or nodes) in the model. This helps the model to generalize the problem in a better way. In U-Net++ architecture, they optimized the model on the output of X0,1, X0,2, X0,3, and X0,4 nodes by calculating the combined loss on the expected output based on the output of each of these nodes.

Figure 3.3 (b): UNet++ architecture



3.3.2 MODIFIED NORMALIZED DIFFERENCE WATER INDEX

Modified Normalized Difference Water Index (MNDWI) is a spectral index used to analyse water bodies such as rivers, lakes and dams, especially in built-up areas since it can reduce or even remove built-up land. This index uses two bands: green and short-wave infrared (SWIR). MNDWI values range from -1 to +1, where the negative values represent areas with no water bodies and positive values >0.5 represent water bodies.

$$MNDWI = (G - SWIR)/ (G + SWIR) \qquad (3.3.2)$$

G is the green band (B3 in Sentinel-2 MSI L1C) and SWIR is the short-wave infrared band. It is widely accepted that MNDWI is more stable and reliable than NDWI.

Figure 3.3 (c): Atmospheric transmission VS Wavelength

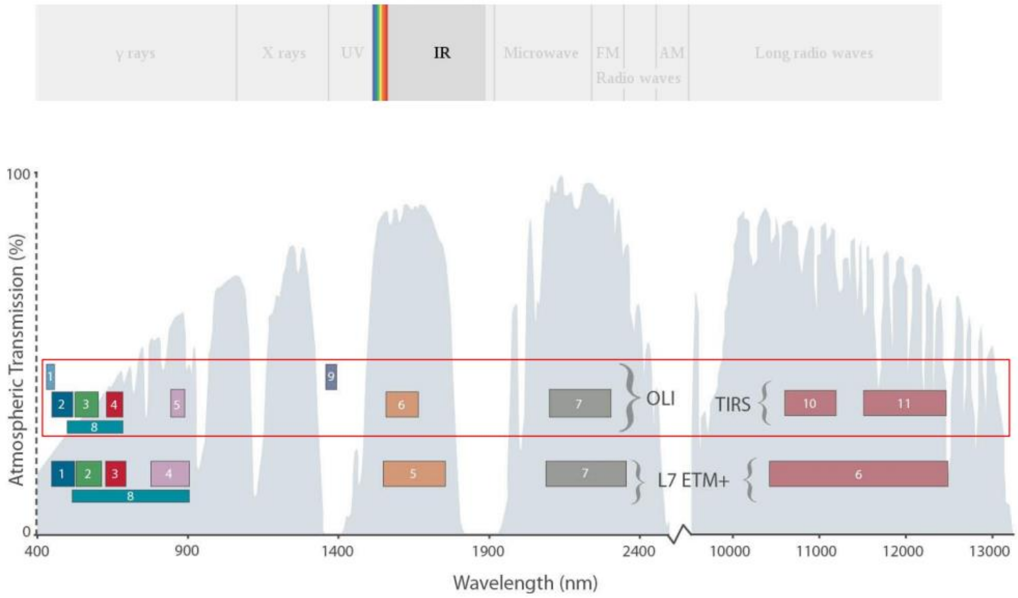
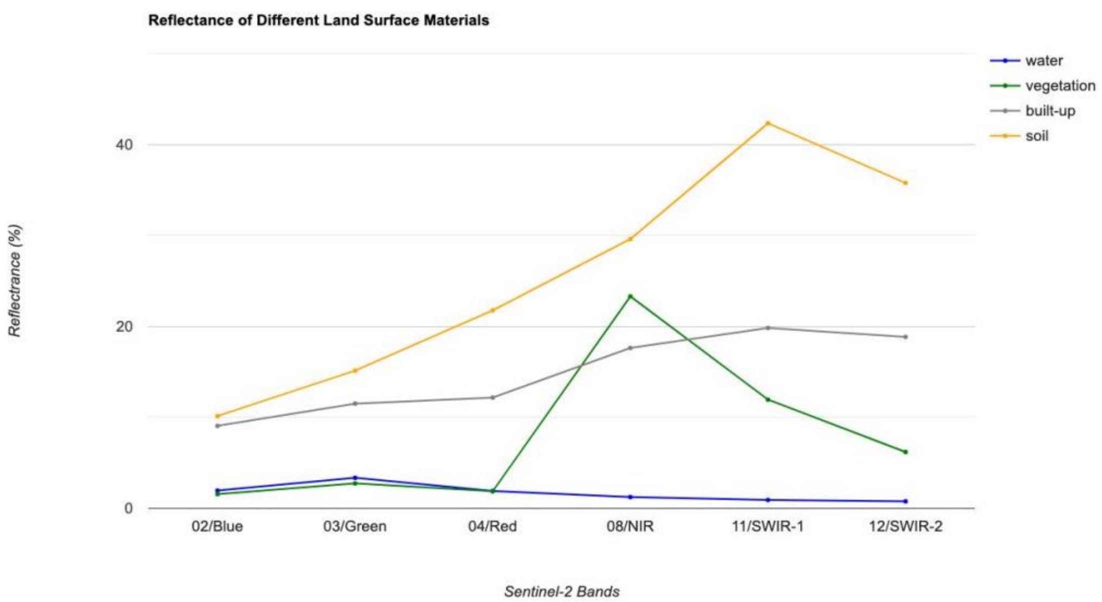


Figure 3.3 (d): Reflectance graph of Sentinel-2 bands



3.3.3 XGBOOST

XGBoost is an optimized distributed gradient boosting library designed for efficient and scalable training of machine learning models. It is an ensemble learning method that combines the predictions of multiple weak models to produce a stronger prediction. XGBoost stands for “Extreme Gradient Boosting” and it has become one of the most popular and widely used machine learning algorithms due to its ability to handle large datasets and its ability to achieve state-of-the-art performance in many machine learning tasks such as classification and regression. One of the key features of XGBoost is its efficient handling of missing values, which allows it to handle real-world data with missing values without requiring significant pre-processing. Additionally, XGBoost has built-in support for parallel processing, making it possible to train models on large datasets in a reasonable amount of time.

Decision Tree:

A Decision tree is a flowchart-like tree structure, where each internal node denotes a test on an attribute, each branch represents an outcome of the test, and each leaf node (terminal node) holds a class label. A tree can be “learned” by splitting the source set into subsets based on an attribute value test. This process is repeated on each derived subset in a recursive manner called recursive partitioning. The recursion is completed when the subset at a node all has the same value of the target variable, or when splitting no longer adds value to the predictions.

Bagging:

A Bagging classifier is an ensemble meta-estimator that fits base classifiers each on random subsets of the original dataset and then aggregate their individual predictions (either by voting or by averaging) to form a final prediction. Each base classifier is trained in parallel with a training set which is generated by randomly drawing, with replacement, N examples (or data) from the original training dataset, where N is the size of the original training set.

Random Forest:

The basic idea behind this is to combine multiple decision trees in determining the final output rather than relying on individual decision trees.

Random Forest has multiple decision trees as base learning models. We randomly perform row sampling and feature sampling from the dataset forming sample datasets for every model.

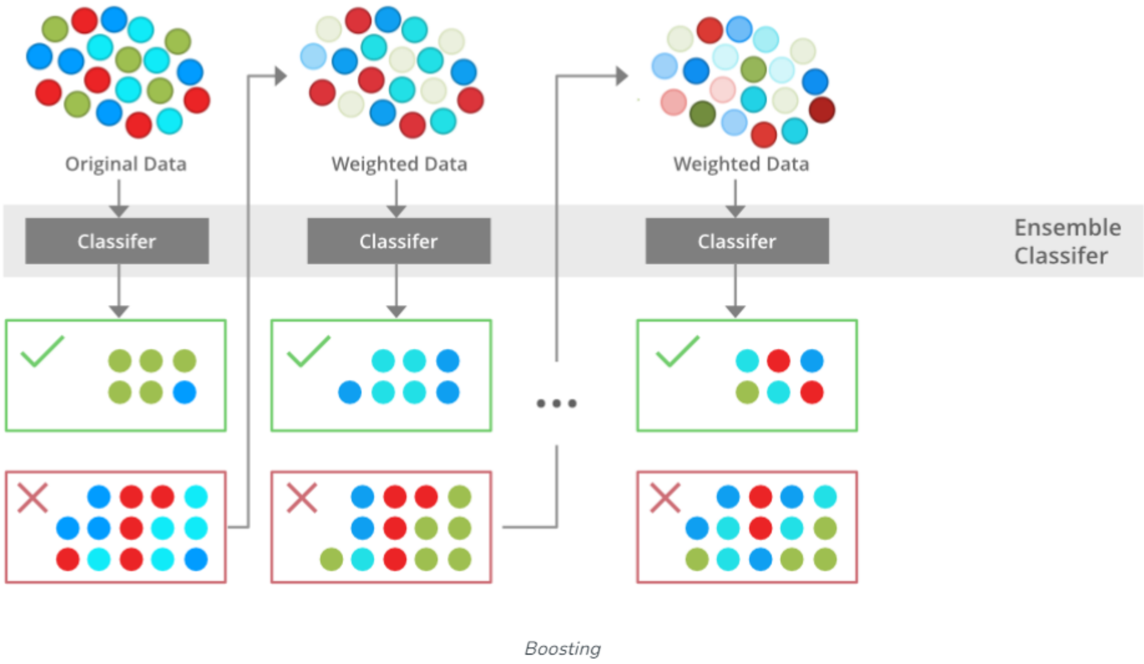
Boosting:

Boosting is an ensemble modelling, technique that attempts to build a strong classifier from the number of weak classifiers. It is done by building a model by using weak models in series. Firstly, a model is built from the training data. Then the second model is built which tries to correct the errors present in the first model. This procedure is continued and models are added until either the complete training data set is predicted correctly or the maximum number of models are added.

XGBoost:

In this algorithm, decision trees are created in sequential form. Weights play an important role in XGBoost. Weights are assigned to all the independent variables which are then fed into the decision tree which predicts results. The weight of variables predicted wrong by the tree is increased and these variables are then fed to the second decision tree. These individual classifiers/predictors then ensemble to give a strong and more precise model.

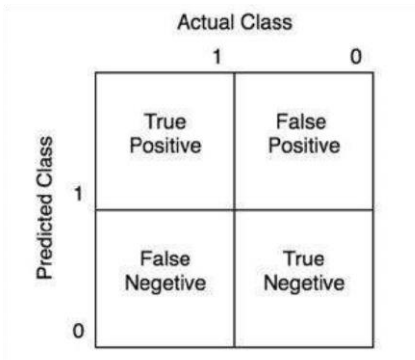
Figure 3.3 (e): XGBoost



3.3.4 ACCURACY METRICES

To calculate the efficiency, the confusion matrix for the validation set was created, and then four types of accuracy parameters were used, namely overall accuracy measure, F-score, precision, recall and IoU. Confusion Matrix, also known as an error matrix, is a specific table layout that allows visualization of the performance of an algorithm, typically a supervised learning one. Each row of the matrix represents the instances in an actual class, while each column represents the instances in a predicted class. A typical confusion matrix is shown in Figure 3.3.4 (a).

Figure 3.3 (f): Confusion Matrix



The overall accuracy is the percentage of labels correctly classified by the algorithm and is calculated as follows:

$$Overall\ Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \qquad (3.3.4(a))$$

Here,

TP = True positive, implying predicted class and actual class both were water

TN = True negative, implying predicted class and actual class both were non-water FP = False Positive, implying predicted class was water whereas actual class was nonwater (also called Type I error)

FN = False negative, implying predicted class was non-water whereas actual class was water (also called Type II error)

Precision (also called positive predictive value) is the fraction of relevant instances among the retrieved instances and is given by:

$$Precision = \frac{TP}{TP+FP} \qquad (3.3.4(b))$$

Recall (also known as sensitivity) is the fraction of relevant instances that were retrieved and is computed as:

$$Recall = \frac{TP}{TP+FN} \qquad (3.3.4(c))$$

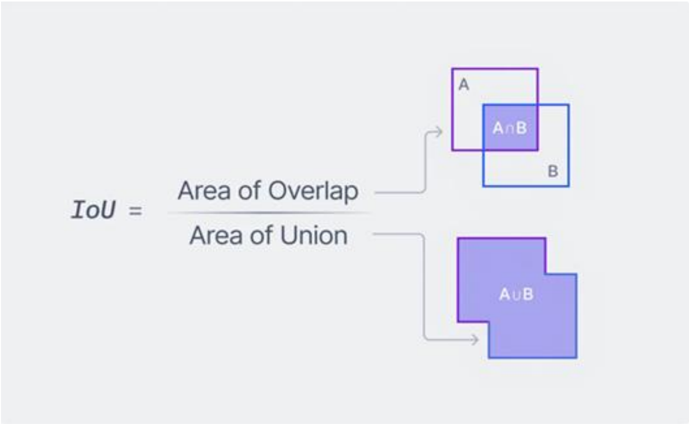
The F-score, also called the F1-score, measures a model's accuracy on a dataset. It is used to evaluate the binary classification systems, which classify examples into

'positive' or 'negative'. The F-score combines the precision and recall of the model, and it is defined as the harmonic mean of the model's precision and recall.

$$F - score \ = \ 2 * \frac{(Precision * Recall)}{Precision+Recall} \qquad (3.3.4(d))$$

Intersection over Union (IoU), also known as the Jaccard index, is the ratio of the 'area of intersection' to the 'area of the union' between the predicted and ground truth bounding boxes. Thus, the IoU meaning consists of the quantitative measurement of how well a predicted bounding box aligns with the ground truth bounding box. The IoU is calculated as

Figure 3.3 (g): IoU



$$Intersection\ over\ Union\ \left(IoU\right) = \frac{|\mathbf{A} \cap \mathbf{B}|}{|\mathbf{A}| \cup |\mathbf{B}|}$$

But for binary classification, it is written as:

$$Intersection\ over\ Union\ \left(IoU\right) = \frac{\mathbf{TP}}{\mathbf{TP} + \mathbf{FN} + \mathbf{FP}}$$

Where

-
- TP= True Positive.
- FN= False Negative
- FP= False Positive

CHAPTER – 4

RESULTS AND DISCUSSION

The labelled dataset was partitioned into separate sets to train and validate the machine learning models. The training set consisted of 600 images which accounts for around 60% of the total 1000 images in the dataset. The remaining 40% data were used for validation and testing (20% each). The Table 4.1 presents the evaluation metrics for RF and SVM models.

Table 4.1: Accuracy assessment on test set

Method	**F1 Score**	**IoU**
UNet++	0.7697	0.6256
XGBoost	0.7553	0.6069
MNDWI	0.7581	0.6105

The evaluation metrics, namely accuracy, F1 score, and IoU, were computed for UNet++, XGBoost and MNDWI. The UNet++ model performed better in F1 score and IoU. The MNDWI has the 2nd best performance. XGBoost has the lowest performance. This is due to the reason that the model weren't trained enough due to lack of computational power.

Figure 4.1 (a): XGBoost prediction

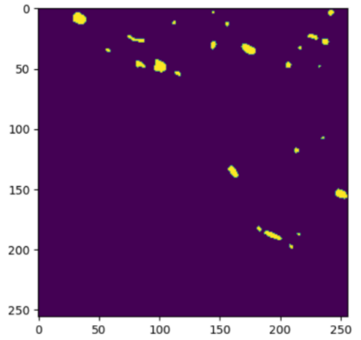


Figure 4.1 (b): MNDWI prediction

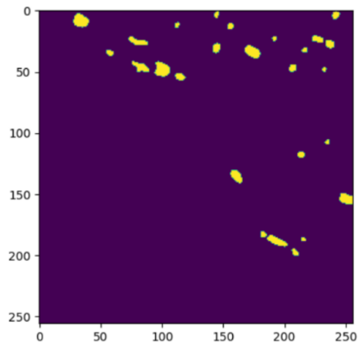


Figure 4.1 (d): Label

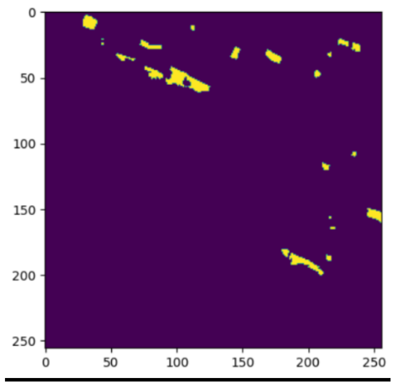
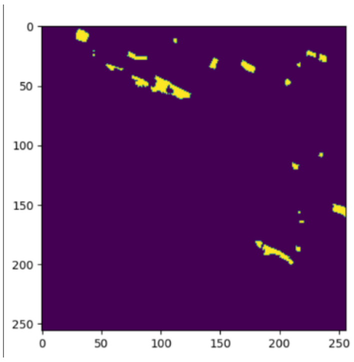


Figure 4.1 (c): UNet++ prediction



CHAPTER – 5

SUMMARY AND CONCLUSIONS

The analysis conducted using UNet++, XGBoost, Modified Normalized Difference Water Index (MNDWI provided valuable insights into water body detection within the Odagaon region. The UNet++ model exhibited the highest water area detection, closely followed by MNDWI and XGBoost. MNDWI, a commonly used spectral index for water detection, presented moderate results, indicating water bodies with more coverage than machine learning model XGBoost. This leads to fact that learning model should be well trained with accurate data to have better performance. In this analysis the training for both the learning model were constrained due to lack of computational power and time. Model such as the UNet++ are the state-of-the-art model but its performance were not as expected. In conclusion, the study provided insights into the strengths and limitations of different small water body detection methods. The results can help researchers and decision-makers choose the appropriate method for water body detection, depending on the specific environmental conditions and the size of the water bodies of interest.

CHAPTER – 6

REFERENCES

Deep Learning for Automatic Extraction of Water Bodies Using Satellite Imagery - *Reham Gharbia*

Comparative Analysis of Machine Learning Algorithms in Automatic Identification and Extraction of Water Boundaries - *Aimin Li , Meng Fan, Guangduo Qin , Youcheng Xu and Hailong Wang*

SFD Lite Report Odagaon, Odisha, India

Produced by: CSE, this SFD Lite Report was prepared by Centre for Science and Environment

UNet++: Redesigning Skip Connections to Exploit Multiscale Features in Image Segmentation - *Zongwei Zhou, Member, IEEE,*

Md Mahfuzur Rahman Siddiquee, Member, IEEE,

Nima Tajbakhsh, Member, IEEE, and

Jianming Liang, Senior Member, IEEE

Remote Sensing Inversion and Prediction of Land Use Land Cover in the middle reaches of the Yangtze River Basin, China - *Zhang, SQ (Zhang, Shengqing); Yang, P (Yang, Peng); Xia, J (Xia, Jun); Wang, WY (Wang, Wenyu) ; Cai, W (Cai, Wei); Chen, NC (Chen, Nengcheng); Hu, S (Hu, Sheng); Luo, XA (Luo, Xiangang); Li, J (Li, Jiang); Zhan, CS (Zhan, Chesheng)*

Deep-Learning-Based Multispectral Satellite Image Segmentation for Water Body Detection - *Kunhao Yuan , Xu Zhuang, Gerald Schaefer, Jianxin Feng , Lin Guan, and Hui Fang*