Frontiers in Health Informatics www.healthinformaticsjournal.com
ISSN-Online: 2676-7104
2024; Vol 13: Issue 3 Open Access

Water Stream Segregation Using U-Net Architecture On Sentinel 2 A Dataset

Akanksha Singh[1*], Syed Wajahat Abbas Rizvi[1], Prabhat kr. Srivastava[2], Rashmi Saini[3]

[1*] Amity School of Engineering and Technology, Amity University, Lucknow, 226028, Uttar Pradesh, India. ankakaanjphd@gmail.com;

[1]Amity School of Engineering and Technology, Amity University, Lucknow, 226028, Uttar Pradesh, India. swarizvi@lko.amity.edu

[2]Department of Computer Science and Engineering, IMS Engineering College, Street, Ghaziabad, 201015, Uttar Pradesh, India. sri_prab@rediffmail.com;

[3], Department of Computer Science and Engineering, G. B. Pant Institute of Engineering and Technology, Street, Pauri Garhwal, 246194, Uttarakhand, India. 2rashmisaini@gmail.com

*Corresponding author(s). E-mail(s): ankakaanjphd@gmail.com;
Contributing authors: [1]swarizvi@lko.amity.edu; [2]sri_prab@rediffmail.com;
[3]rashmisaini@gmail.com;

Cite this paper as: Akanksha Singh, Syed Wajahat Abbas Rizvi, Prabhat kr. Srivastava, Rashmi Saini (2024) Water Stream Segregation Using U-Net Architecture On Sentinel 2 A Dataset. *Frontiers in Health Informatics, 13* (3), 10622-10640

Abstract.*Climate is a prolonged environmental style and its change is a natural process. As the climate is constantly changing, monitoring the surroundings globally is a need of the hour that requires attention. There's is need to monitor places that have suffered due to diverse climate changes like heavy rainfall, drought etc. at the time of harsh climate condition. First objective of this study deals with observing areas remotely that have experienced floods or water retention and particularly demarcate the boundaries of water from remote Sentinel-2A images. This segregation of water is essential for handling diverse climate changes. Using satellite image for this task provide huge coverage and fast and quick upgradation. Deep learning algorithm, U-Net is used in study which is aiding in automating the surrounding monitoring process, benefiting from sentinel satellite images. Another objective of this study is to effectively observe and train the machine to anticipate the region bordered by water, which is enhancing deep learning-grounded methodological framework. In order to accomplish our objective a U-Net construction was implemented that allows learning of the key water classes label features from input Sentinel-2A. The result shows encouraging and promising accuracy value of 99.4 %. Performance measures used in our research are Mean intersection over union using Jaccard coefficient which results as 0.91 and total loss using focal loss and dice loss which results to 0.53*

Keywords:*Deep learning, Mean IOU, Sentinel image, U-Net, sustainable, water bodies.*

Introduction

Hydrology is a growing field of study and it's one of the streams is to manage and develop water systems to address environmental problems such flood prevention and planning [1]. Floods can be devastating causing loss of structures and crop areas and these become soul reason for economical imbalance and spread of disease [2]. Now the question is are there sufficient researches for flood prevention? More than 140 studies were found to investigate flood detection and avoidance. Out of these 50 looked at using machine learning, 65 looked at using image processing, 15 looked at using remote sensing methods like GPS and GIS, and only 10 looked at using image processing and machine learning [3]. This region of research where remote sensing, ML and image processing are integrated together to predict flood or predict region surrounded by water is the least explored and serves as research gap for our study. What are the challenges associated with above integration? It is extremely tedious to accurately segregate water-bodies from very high resolution (VHR) pictures and major challenges are faced due to complex spectral combinations formed by aquatic flora, different colours of pond,

river tributaries and other water-bodies, silts along the bank. [4]. Good resolution remote satellite pictures are often indistinct. These corruptions are due to obstruction caused by oceanic vegetation, settled sediments near the shore and shadows of various tall plant [5]. Why accurate water body segmentation is crucial? a) accurate and in time flood prediction b) planning irrigation structure and systems c) urban development planning d) planning farming areas. Our objective in this study is to achieve and improved mechanism to detect water from satellite images. Deep learning has made very useful implication in reading satellite pictures for living beings [6]. Convolution neural networks (CNNs) have gained prominence around semantic segregation of remote sensing pictures in recent years [7]. The 2015 invention U-Net serves as an example of the advantages of precise segmentation. U-Net has more than 2500 research citations due to the sharp increase in the performance requirements for segmentation in medical imaging [8]. Our study is going to use U-Net for segmentation of aquatic boundaries from sentinel-2A satellite pictures and provide improved accuracy. Geographical area for our study is area near Brahmaputra tributaries (Teesta River and dharla river).

Related Work

Anusha Ch et al. [9] build a U-net model with integrated security feature for detecting bodies of water from sentinel-2 satellite pictures. A total dataset of 2841 images were collected along with their mask from Kaggle. This model used Nadam for optimization to get reduced loss along with parameters which are optimized as well like activation functions and nodes of various layers. The proposed model gives an accuracy of 94% with robustness.

Sarah Mazhar et al.[10] suggested a U-net, a network of deep learning for accurate water segregation. A fresh dataset for the mapping of water channels in the Pakistan was created. A poorly trained approach to extract water areas from high resolution images was employed instead of pixel-level waterbody annotations, relying solely on OpenStreetMap (OSM) waterways. This dataset was benchmarked using the most recent frameworks for semantic segmentation. The experimental findings show the suggested A U-net model's superior performance for segregation with less strengthen labels(supervised). It attained mean IOU value of 0.8791.

Min Xia et al.[11] suggested a multi-size network for image segmentation that uses Conditional Random Field (CRF) as post-processing and is built on dense attention and skip connection mechanisms. Dense skip connections and associated conventional convolution modules were appended to U-Net and called it DAU-NET, which is used to extract features from water. The attention module isolates, extracts, and combines the collected low-level and high-level information as two independent inputs in different dimensions to enhance the network. The data used was gathered from NASA's Landsat-8 and China's HJ-1A (HJ-1B) hyperspectral environmental remote sensing satellites. This approach significantly increases segmentation accuracy and detection time compared to earlier deep learning systems. The method can identify the mainstream and tributaries of rivers as well as process edge noise to segment the smaller tributaries, in addition to precisely segmenting the lake's salinity section.

Manohara Pai M. M et al.[12] proposed two models—Vanilla U-Net and Transfer U-Net. The Vanilla U-Net construct was coached using the SAR data, while the other construct was coached using the 2015 ISBI Cell Tracking challenge. The effectiveness of suggested framework was assessed using a portion of the SAR data. This technique produces significant performance with 0.98-pixel accuracy and 0.99 F1 value, according to experimental data.

Francois Waldner et al.[13] proposed ResU-Net a fully connected UNET backbone that helps to Identify extent of the field and its boundaries. Model performance and generalization is considerably enhanced by algorithm where three correlated outputs are reconstructed. Data for this study was collected by primary and secondary sites.

H Gao et al. [14] In their study, improved water body extraction strategy and put forth and tested utilizing Landsat-8 OLI images to extract water bodies. A normalized difference water index and Tasselled Cap transform are used to produce binary pictures, which collaborated to produce a mask. The Landsat-8 OLI pictures are then used to mask aquatic bodies, which are subsequently refined by removing erroneous areas using supervised classification. The generated water body maps demonstrate the effective removal of terrain-related shadows in images as well as the accurate identification of man-made ditches and river tributaries (up to 94%).

Sourav Karmaker et al.[15] determined the precise water flow utilizing spatial data and determined the current watershed condition of some Region of Bangladesh, by employing GIS and remote sensing techniques. Using a digital elevation model (DEM) of the research region, the correct watershed flow direction, network density, and confluence thresholds were determined using the Hydrology Toolset. The well-known D8 approach was used to compute the stream flow from each cell to its neighbour on the downslope. 100–1500 thresholds were then used to determine the stream flow directions and convert the streams into features to calculate the density of the watershed network. The findings demonstrated a relationship between the networks' length and density and their threshold.

KomeilRokni et al.[16] proposed work that analyse the spatiotemporal affect that Lake Urmia underwent within the span of three years. The NDWI was picked and used for this purpose after a comparative analysis. The findings indicated a sharp decline in the lake's surface area, particularly in the timeline from 2010 to 2013, when the lake's overall area decreased by about a third. It is highly possible that Lake Urmia will soon lose all its water surface if this downward trend in the lake persists.

Javier Del-Pozo-Velázquez et al.[17] developed and validated a technique that is quick in distinguishing between water and land zones. The Quadtree algorithm and image segmentation were the foundations of the technique used to process high-resolution satellite photos to find surface water. High-resolution satellite pictures obtained for free from the Open Aerial Map website were used to validate the strategy. Their results demonstrate that there is a distinct threshold wherewater and land can be separated. Here surface water detection had an average accuracy of 96.03%.

Gordana Kaplan et al.[18] researched on a 10 m-resolution Sentinel-2 satellite picture, a method that combines an object-based strategy with a pixel-based index. It takes use of water body extraction indicators such as the Normalized Difference Water Index (NDWI) using image segregation on a multispectral picture with 13 bands. A hilly region and an urban area, both in Macedonia, have been chosen as two research locations with contrasting characteristics here. The outcomes from pixel-based indexes like NDWI and object-based methods have been enhanced by a kappa value greater than 0.5.

GulcanSarp et al.[19] proposed work in which geographical alterations and water area change of Lake Burdur were detected, analyzed, and quantified using satellite image interpretation and GIS. This method is based on water spectral indexing like NDWI, MDWI, and AWEI and SVM classification. As per methods used, the lake's spatial-temporal changes demonstrate a substantial reduction in surface area between 1987 and 2011, especially from 1987 to 2000, when the pond lost around one fifth of surface area compared to that in 1987. The results show effectiveness of surface water variations detection techniques based on SVM and MNDWI, particularly when it comes to spotting changes between predetermined time intervals.

Dmytro Filatov et al.[20] suggested methods for viewing the area bounded by water and forest. An image segmentation model called U-Net model has been suggested to do this objective. For the segregation of areas captured by water and forest, our model's validation accuracy was 82.55% and 82.92%, respectively.

Fig. 1. Sentinel-2, level-2A natural colour composite images

Kunhao Yuan et al. [1]surpassed existing DCNN-based water body recognition methods, the multichannel water body detection network (MC-WBDN)which includesthree revolutionary components: a multichannel

function module, an Enhanced Spatial Pyramid Pooling module, and space-to-depth/depth-to-space operations. Their MC-WBDN model provided good water body recognition ability, which was robust to fluctuations in light and weather, and could more clearly differentiate tiny water bodies than existing DCNN models, according to conclusive experimental results.

H. Bao et al. [25] investigated Deeper-Feature Convolutional Neural Network (DFCNN), which can extract more and deeper features for the purpose of constructing semantic recognition, in response to the feature extraction limitations of classic convolutional neural network (CNN) models. Their method and model could effectively, quickly, and accurately recognize urban functional zones by combining building physical semantics and social functional semantics.

A detailed literature survey brings few points to highlight

I. Various segmentation techniques have been proposed but to achieve high efficiency is still tough task.
II. Few researchers have focused on the need of automation of system which forecasts the possibility of lands to be covered by water.

1 Study Area and Data Sources

Our area of study for this research is part of northeast India and some part of Bangladesh mainly covering areas near Brahmaputra tributaries (Teesta River and dharla river). In this area, two Sentinel-2, level-2A images were taken. The first image has the following coordinates 24.901035177555638, 90.06642564591849. The second image has coordinates 26.206000495572308, 89.00130194368923. The good data quality and the little cloud coverage served as the selection criteria for the pictures. They consist of a variety of different land coverage, such as concrete, pavement, water, plant, and mud, as shown in natural colors composite or true color composite represented in Figure 1. These images are very much similar to true images and are made up of red, green, blue bands. Both images picked as data source are in continuation to each therefore (area lie one above the other) are stacked together for complete pictorial representation and is shown in Figure 2(a). This figure consists of combined scene of two Sentinel-2, level-2A images. Together both images cover stretches of 10000 km^2 area on ground. On map Figure2(b) shows the location of images for better understanding. A thematic map of India is drawn above which picture from sentinel-2A are placed to show exact reference of the data.

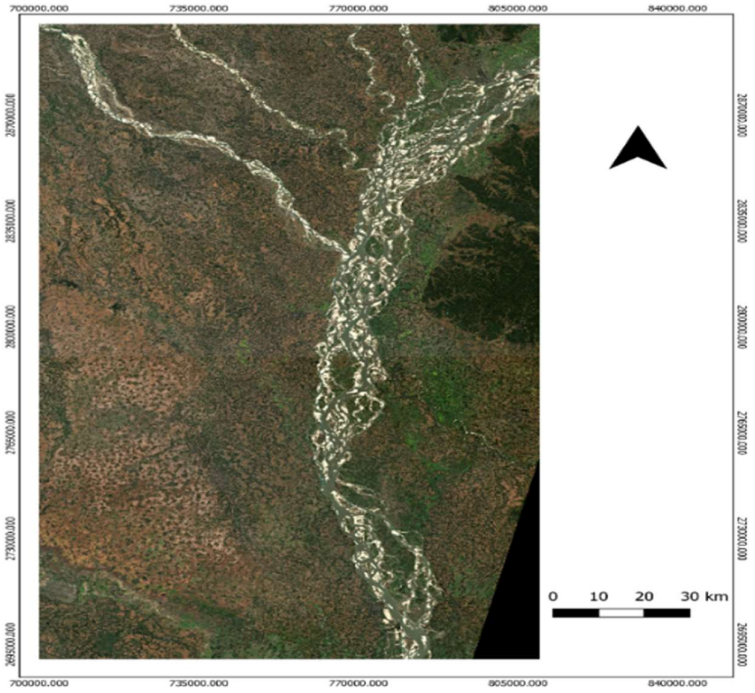


Fig. 2(a). Actual scene when combined natural composite images

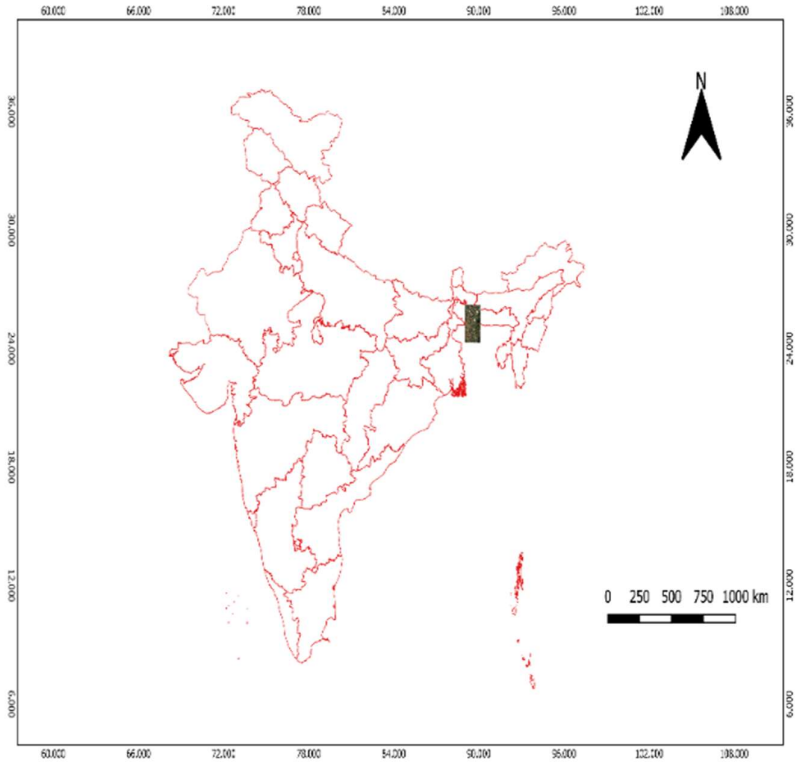


Fig. 2(b). Location with reference to map of India

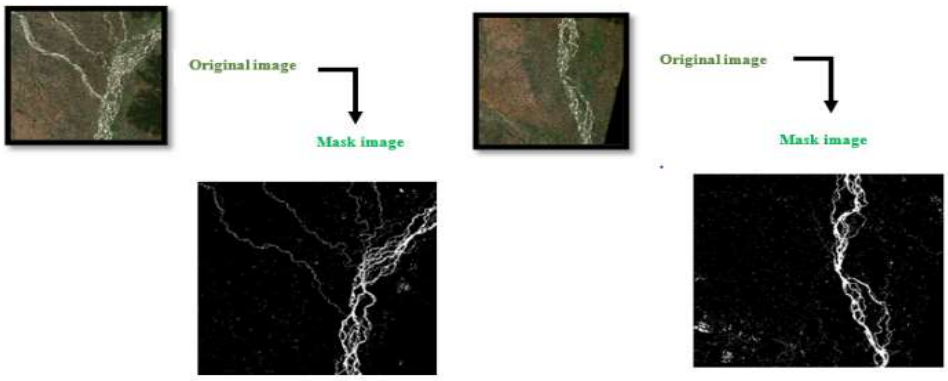


Fig. 3. Mask images of respective Sentinel-2 original image

Figure 3 elaborates the original image along with their mask images in one frame. Masking image is way of extracting feature from natural composite picture. Extraction of feature is a process to highlight pixel of one structure while diminishing the other pixel. Mask image of these sentinel-2A images in our study were created using the SNAP tool. Version 9.0.0 of sentinel toolbox (SNAP) was used for feature extraction and mask creation. Mask manager tool within SNAP allows us to create new mask for sentinel 2 pictures. General information like Image id, cloud coverage, sensing orbit number etc. about the selected data source has been discussed in table 1. The employed multispectral image is a 12-bit radiometric picture of size 10980 × 10980 pixels for R, G, B bands (10-m resolution) and of size 5490 × 5490 pixels for the SWIR band (20-m resolution). The image used in this article were taken from Sentinel-2 A in December 2022.

TABLE I. DETAILS OF SENTINEL-2A IMAGES

Processing Level	Product Type	Cloud Coverage Assessment	Sensing Orbit Number	Tile Id
Level 2-A	S2MSI2A	**0.004758**	33	T45RYH
Level 2-A	S2MSI2A	**0.004874**	33	T45RYJ

Multispectral imagery is crucial part of Sentinel 2A which delivers 12 spectral bands with resolution ranging from 10m to 60 m in accordance with size of pixel. Central wavelength of this bands ranges from 443 nm to 2190 nm. These spectral may have color information like shades of blue, red and green and some are categorized as visible near infrared and shortwave infrared. A detailed description of bands for Sentinel-2A is provided in Table 2.

TABLE II. BANDDETAILS OF SENTINEL-2A

Band	Resolution	Spectral information
B1	60 m	Ultra-blue
B2	10 m	Blue
B3	10 m	Green

B4	10 m	Red
B5	20 m	Visible and near Infrared
B6	20 m	Visible and near Infrared
B7	20 m	Visible and near Infrared
B8	10 m	Visible and near Infrared
B8a	20 m	Short Wave Infrared
B9	60 m	Short Wave Infrared
B10	60 m	Short Wave Infrared
B11	20 m	Short Wave Infrared
B12	20 m	Short Wave Infrared

These bands when come together in different combination create different color representation. Fusion of these bands provide a vivid picture which can highlight the concern object of study. A brief information about the fusion of these bands is given in Table 3. This tabular description clearly explains the implication of fusion of different bands.

TABLE III. BAND FUSION DISCRIPTION

Band fusion	Description for Band fusion	Implication of band fusion
B4, B3, B2	Natural color composite	Gives real life imagery
B8, B4, B3	Color Infrared	Segregates healthy and unhealthy vegetation
B12, B8a, B4	Short wave infrared	Segments vegetation in remote images
B11, B8, B2	Agriculture	Quantifies health of crops
B12, B11, B2	Geology	Highlights the geological structures
B4, B3, B1	Bathymetric	Used for Coastal Study
(B8-B4)/(B8+B4)	Vegetation Index	Quantifies Vegetation density
(B8a-B11)/(B8a+B11)	Moisture Index	Water content in vegetation

2 Data and Data Preparation

Two Sentinel-2A images were collected from Copernicus. European Union's space initiative has a component for earth observation, known as Copernicus, which studies our planet and its environment for the good of all European residents. It provides information services utilizing data from in-situ (non-space) sources and satellite Earth observation.

To improve the model's performance and accelerate the computation, data pre-treatment and Feature extraction techniques are used before model training. These include the following. Figure 4 is a detailed picture of various phases of data pre-processing.

Fig. 4. Data pre-processing steps

2.1 ORIGINAL SENTINEL-2A IMAGE

Original images are Sentinel-2, level-2A pictures. Two of them have following coordinates following coordinates 24.901035177555638, 90.06642564591849 and26.206000495572308, 89.00130194368923. The high picture quality and the less cloud coverage assessment were selection criteria.

2.2 IMAGE MASKING USING MNDWI

Green and SWIR bands are used by the Modified Normalized Difference Water Index (MNDWI) to highlight open water features. Additionally, it reduces elements of built-up areas that other indices frequently link with open water.

(Green - SWIR) / (Green + SWIR) is the MNDWI. [27]

Green: the values of the pixels in the green band

SWIR refers to the short-wave infrared band's pixel values.

The MNDWI calculation will yield several outcomes:

1) Since water absorbs more MIR light than NIR light, it will have higher positive results than in the NDWI.

(2) Soil and plant will still have negative values since plant reflect MIR light more than green light and soil reflects MIR light more than NIR light, according to Jensen (2004). Built-up land will also have negative values. As a result, when compared to the NDWI, the MNDWI's contrast between water and built-up land will be noticeably larger due to rising values for water features and falling, from positive to negative, values for built-up land. Since the built-up land, soil, and vegetation all have negative values andare thus noticeably muted and even deleted, the larger augmentation of water in the MNDWI-image will lead to a more accurate extraction of open water features. Then, to distinguish between these two classes in the visual data, a suitablethreshold of the index was applied basedon the spectral features therefore water was extracted.

2.3 PATCHIFYING MAGES AND MASK

The raster image is broken up (patch creation) into manageable parts to lessen the demanding computing and memory requirements and to permit parallel computation. The Natural Colour Composite picture and Mask, which are the inputs for our suggested model and are utilized for image classification, have a size of 256X256 pixels.

3 Methodological Framework

3.1 Tools

Spyder (anaconda 3), was used for Python code through which the model was developed for the study. Python libraries were used for the various tasks like

- Numpy(numerical python): was used for array operations. It performed various mathematical operations on array. It was preferred for its fast processing as compared to list.
- The patchify library: image splitting was done using this library for both sentinel-2A image and mask image. This library is used for converting original image into non-overlapping patches then merges those patches back into the original image during the prediction stage.
- Keras with TensorFlow: at backend is used for artificial intelligence and machine learning implications and for using the CNN model.
- Matplotlib: was used for making various plots and graphs like result plots.
- SNAP Tool and QGIS tools were used for the data pre-processing in this work. The production, modification, and visualization of geographic data are supported by QGIS, a free and open-source geographic information system. SNAP is an open-access Earth observation analysis tool that is used for MDWI and mask creation. In table 4 we have discussed the role of each tool with respect to our study
-

Sr.no	Tools	Roles
1	Python Libraries	Creating patches, reading images into array, Implementation of framework, plotting graphs
2	SNAP	Classification using MDWI and mask creation
3	QGIS	Production, modification, and visualization of geographic data

TABLE IV. TOOLS USED IN STUDY AND ITS SPECIFIC ROLE INOURSTUDY

3.2 METHODS

On a computer with 32GB of RAM, 8GB graphics processing unit (GPU), the model is trained. The 256x256x3 input pictures are provided to the U-Net network for training. Accuracy, loss, and intersection over union (IoU) coefficient matrices are tracked throughout training.

U-Net Model

U-net model is type of deep learning technique. It was originally developed around 2015 for segmentation in images related to medical. U-net performs extremely well for segmentation of image [8]. A gradual shift has been seen in usage of U-net model since 2015. Here we have presented the model usage for water segmentation from sentinel-2A picture.Our U-net model that is implemented in our study has beensummarized in figure 5(a) and 5(b) which clearly depicts the two stages of the algorithm a) encoder phase and the b) decoder phase. In total 3,167,205 parameters were trained. Figure clearly depicts about each layer and the entire working of the model. All layer's spatial dimensions are decreased in encoder phase, but the number of channels is increased.In decoder phase the channels are lowered while raising the spatial dimensions. At last, in order to produce a prediction for each pixel in the input image, the spatial dimensions are ultimately recovered.

Layer (type)	Output Shape	Param #	Connected to
input_1 (InputLayer)	[(None, 256, 256, 3)]	0	[]
conv2d (Conv2D)	(None, 256, 256, 16)	448	['input_1[0][0]']
dropout (Dropout)	(None, 256, 256, 16)	0	['conv2d[0][0]']
conv2d_1 (Conv2D)	(None, 256, 256, 16)	2320	['dropout[0][0]']
max_pooling2d (MaxPooling2D)	(None, 128, 128, 16)	0	['conv2d_1[0][0]']
conv2d_2 (Conv2D)	(None, 128, 128, 32)	4640	['max_pooling2d[0][0]']
dropout_1 (Dropout)	(None, 128, 128, 32)	0	['conv2d_2[0][0]']
conv2d_3 (Conv2D)	(None, 128, 128, 32)	9248	['dropout_1[0][0]']
max_pooling2d_1 (MaxPooling2D)	(None, 64, 64, 32)	0	['conv2d_3[0][0]']
conv2d_4 (Conv2D)	(None, 64, 64, 64)	18496	['max_pooling2d_1[0][0]']
dropout_2 (Dropout)	(None, 64, 64, 64)	0	['conv2d_4[0][0]']
conv2d_5 (Conv2D)	(None, 64, 64, 64)	36928	['dropout_2[0][0]']
max_pooling2d_2 (MaxPooling2D)	(None, 32, 32, 64)	0	['conv2d_5[0][0]']
conv2d_6 (Conv2D)	(None, 32, 32, 128)	73856	['max_pooling2d_2[0][0]']
dropout_3 (Dropout)	(None, 32, 32, 128)	0	['conv2d_6[0][0]']
conv2d_7 (Conv2D)	(None, 32, 32, 128)	147584	['dropout_3[0][0]']
max_pooling2d_3 (MaxPooling2D)	(None, 16, 16, 128)	0	['conv2d_7[0][0]']
conv2d_8 (Conv2D)	(None, 16, 16, 256)	295168	['max_pooling2d_3[0][0]']
dropout_4 (Dropout)	(None, 16, 16, 256)	0	['conv2d_8[0][0]']
conv2d_9 (Conv2D)	(None, 16, 16, 256)	590080	['dropout_4[0][0]']
dropout_8 (Dropout)	(None, 256, 256, 16)	0	['conv2d_16[0][0]']
conv2d_17 (Conv2D)	(None, 256, 256, 16)	2320	['dropout_8[0][0]']
conv2d_18 (Conv2D)	(None, 256, 256, 2)	34	['conv2d_17[0][0]']

Total params: 1,941,122
Trainable params: 1,941,122
Non-trainable params: 0

Figure 5(a). U-Net Model (Encoder Stage) implemented

```
decoder_stage3a_conv (Conv2D)  (None, None, None,   36864       ['decoder_stage3_concat[0][0]']
                               32)

decoder_stage3a_bn (BatchNorma  (None, None, None,   128        ['decoder_stage3a_conv[0][0]']
lization)                      32)

decoder_stage3a_relu (Activati  (None, None, None,   0          ['decoder_stage3a_bn[0][0]']
on)                            32)

decoder_stage3b_conv (Conv2D)  (None, None, None,   9216        ['decoder_stage3a_relu[0][0]']
                               32)

decoder_stage3b_bn (BatchNorma  (None, None, None,   128        ['decoder_stage3b_conv[0][0]']
lization)                      32)

decoder_stage3b_relu (Activati  (None, None, None,   0          ['decoder_stage3b_bn[0][0]']
on)                            32)

decoder_stage4_upsampling (UpS  (None, None, None,   0          ['decoder_stage3b_relu[0][0]']
ampling2D)                     32)

decoder_stage4a_conv (Conv2D)  (None, None, None,   4608        ['decoder_stage4_upsampling[0][0]
                               16)                              ']

decoder_stage4a_bn (BatchNorma  (None, None, None,   64         ['decoder_stage4a_conv[0][0]']
lization)                      16)

decoder_stage4a_relu (Activati  (None, None, None,   0          ['decoder_stage4a_bn[0][0]']
on)                            16)

decoder_stage4b_conv (Conv2D)  (None, None, None,   2304        ['decoder_stage4a_relu[0][0]']
                               16)

decoder_stage4b_bn (BatchNorma  (None, None, None,   64         ['decoder_stage4b_conv[0][0]']
lization)                      16)

decoder_stage4b_relu (Activati  (None, None, None,   0          ['decoder_stage4b_bn[0][0]']
on)                            16)

final_conv (Conv2D)            (None, None, None,   290         ['decoder_stage4b_relu[0][0]']
                               2)

softmax (Activation)           (None, None, None,   0           ['final_conv[0][0]']
                               2)

==================================================================================================
Total params: 24,456,299
Trainable params: 3,167,205
Non-trainable params: 21,289,094
```

Figure 5(b). Glimpse of U-Net Model's (Decoder Stage) of this study

Cost Function and Metric calculation

Loss function, also called cost function, measures the error between the output value and target value.When the model is being learning from pre-labeled data, a loss function provides feedback on how well the model is approaching the ideal model expectation. It serves as a roadmap for a model as it looks for the approximation that converts input data into output data. In this study we have formulated Total loss metric function which is combination of two other loss function dice loss and focal loss.

Loss function could be categorized as distribution-based loss, region-based loss, or boundary-based loss. Most commonly used loss function are dice loss, focal loss and cross entropy loss. The dice loss function which is region-based loss function [20] is a frequently chosen loss function for model like U-net to minimize prediction bias. It is especially helpful when there is a class imbalance issue with the training data. To understand Focal loss separately which is generated from cross-entropy, examination of cross entropy (CE)loss is required.

$$CE = \begin{cases} -\log(p) | (for y = 1) \\ -\log(1-P) | (Elseforally) \end{cases} [26]$$

Where p is probability of class for Focal Loss. Focal Loss (FL) trains hard negative which is exception from cross entropy loss. FL is balanced by a modulating factor of (1-p) t) γ

$$FL(p_t) = -\alpha \cdot (1-P_t)^{\gamma} \log(P_t) [26]$$

Here γ > 0 but when γ = 1 FL is equal to CE and α here is hyperparameter which can range from 0 to 1.

Dice Loss (DL) word has been derived dice coefficient (DC)which meant to measure similarity index of pictures. Later it became a metric for measuring loss.

$$DC = \frac{2TP}{2TP + FP + FN} [26]$$

Where TP stands for true positive and FP, FN are false positive and false negative.

$$DL(y,p) = 1 - \frac{2yp + 1}{y + p + 1} [26]$$

For y = p = 0 function is undefined for edge case scenarios.

In our study we used dice loss and focal loss for calculating total loss which has been formulated as above. Our study found the total loss to be near 0.530.

```
Dice_loss = s.losses.Diceloss(Class_Weights)
Focal_loss=s.losses.CategoricalFocalLoss()
Total_loss= dice_loss+(1*focal_loss)
```

The model's most important component is performance analysis. This study uses Mean IOU for examining how well the model produces outcomes in a substantial way. It is widely used statistic for assessing segmentation models and also used to evaluate the model against other existing models. It is mostly used to determine the amount by which the expected mask and original output overlap.

$$MeanIOU = \frac{\text{overlap area}}{Unionarea}$$[9]

$$MeanIOU = \frac{AintersectionB}{AunionB}$$[9]

Mean IOU score > 0.5 is a good score.

Jacard_coef (Y_true, Y_pred)

Y_true = K.flatten(Y_true)

Y_pred =K.flatten(Y_pred)

Intersection=K.sum(Y_true*Y_pred)

Return(intersection+1.0)/(K.sum(Y_true) +K.sum(Y-pred)-intersection+1.0)

Mean IOU (Jaccard coefficient) in our study was calculated as 0.91640234 and has been evaluated using the above formula.

A. Optimizer

Optimizer Keeps updating the model in response to the output of cost function. Optimizer helps in reducing the cost function. Optimizer can be broadly classified as adaptive and gradient descent optimizers. The line of distinction can be drawn on the basis of feature, when it comes to adaptive algorithms, it automatically adjusts the learning rate, while in the case of gradient descent algorithms, demands manual tuning. Various optimizers are used in case of segregation and segmentation cases some popular ones are Adam, Nadam which lie under adaptive optimizer, and some belong to Gradient descent etc. The optimizer used in our research is Adam which stands for adaptive moment estimation. The The reason of this choice of Adam optimizer is because of its advantages.

- Good with sparse data
- It is not necessary to pay attention to the learning rate value.

The algorithm also requires less tuning than any other optimization algorithm, has a faster running time, needs less memory, and is simple to build.

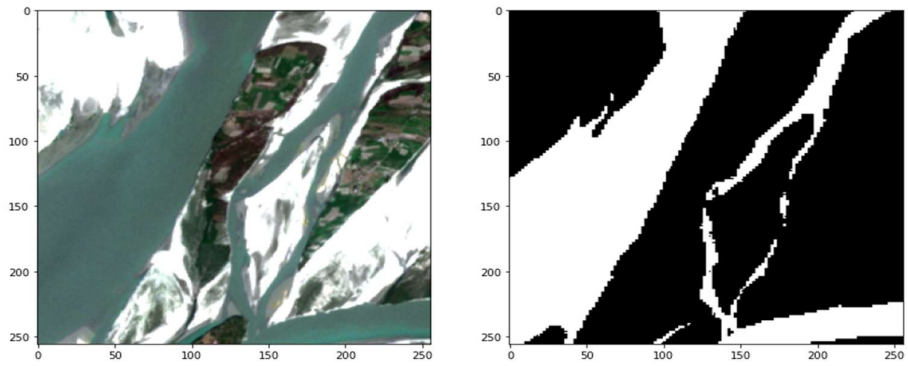
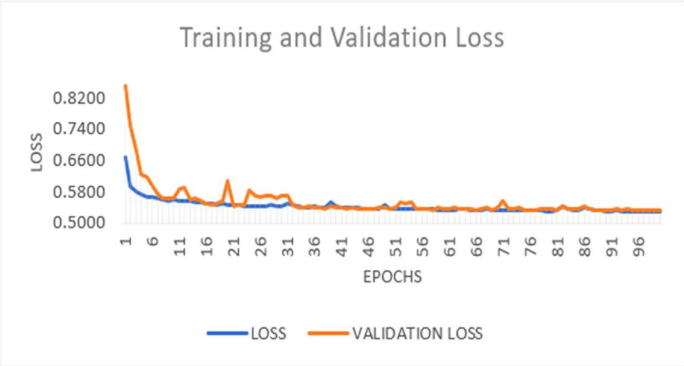


Fig. 6 Single Patch with Associated Mask Image of The Patch

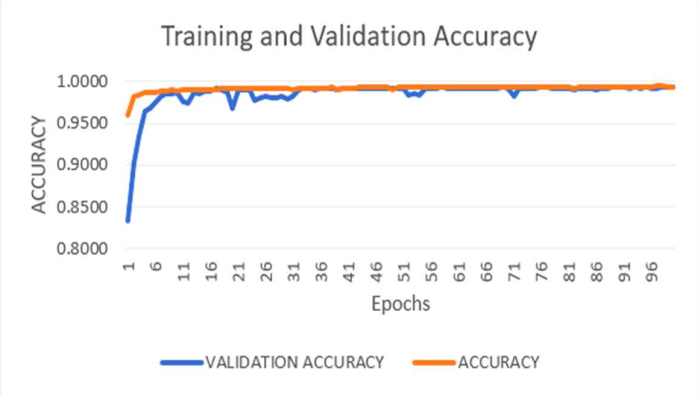
4 Experimental Results and Comparative Analysis

The suggested model uses satellite pictures to segregate water bodies. The Sentinel-2A image and associated original mask are shown in Figure 6 together. The water body is segregated from the satellite pictures using U-Net. The model underwent 100 epochs of training. Table displays the outcomes for the U-Net segregation model on datasets. The table demonstrates the U-Net model's successful use of the water bodies dataset. Masks were created using the normalized difference based on MNDWI for the water bodies dataset.

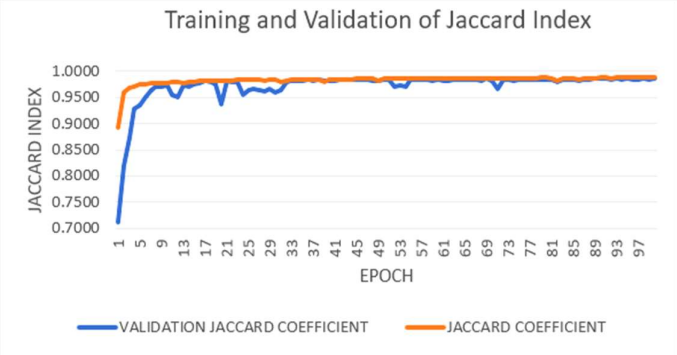
Figure 7 displays the learning curves for the aquatic body's datasets in terms of loss, accuracy, and IOU. The primary and important consideration is how many iterations the model was trained for. U-Net trained for 100 iterations on the water segmentation dataset. The accuracy increased until the tenth epoch, at which point the training and validation datasets nearly coincided. Additionally, thanks to the mask made using MNDWI, the water bodies dataset's validation accuracy curve is practically consistent. In addition, the validation IoU curve for the dataset of water bodies shows an increasing tendency from the beginning to the twentieth epoch. Table 5 discusses the result at 20th epoch,50th epoch and at 100th epoch to illustrate how the metrics have varied over a period time in our model.



(A)



(B)



(C)

Fig 7. Curves for (A) Loss(B) Accuracy (C) Iou

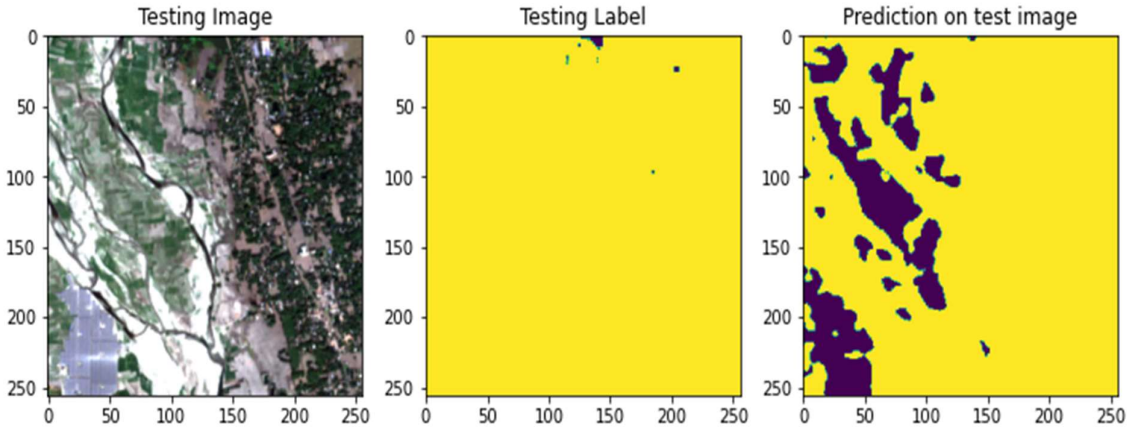


Fig 8. Predicted Mask for Water Segmentation

This demonstrates the significance of a properly labelled dataset. Final, segregation outcomes are shown in Figure 8. The actual picture (Testing picture), their ground truth (Testing Label), and anticipated mask (Prediction on test image) using U-Net are all displayed in the figure. As can be observed, the segmentation outcomes for the segmentation of water images are acceptable despite the loss value.

TABLE V. Results with rescpect to epochs

EPOCHES	DATA	LOSS	ACCURACY	IOU
20th EPOCH	VALIDATION DATA	0.6086	0.9681	0.9381
	TRAINING DATA	0.5470	0.9914	0.9821
50th EPOCH	VALIDATION DATA	0.5368	0.9923	0.9844
	TRAINING DATA	0.5378	0.9928	0.9851
100th EPOCH	VALIDATION DATA	0.5326	0.9929	0.9855
	TRAINING DATA	0.5306	0.9941	0.9877

TABLE VI. Comparison between sentinel-2 and Landsat 8 OLI.

IMAGE DATASET	MIOU	ACCURACY
LANDSAT 8[21]	65.09	81.93
Sentinel 2 A	98.77	99.41

TABLE VII. Comparison of U-Net with Other Methods.

SR.NO	AUTHOR	METHOD	EFFICIENCY
1	Fengyu Yang et al. [22]	R-CNN	90%
2	Hu Tao et al.[23]	CRF-CNN	73.9%
3	Yiqin Wang[24]	SE-ENet	88.4%
4	Hanqing Bao et al.[25]	DF-CNN	96.5%
5	Ours	U-Net	99.4%

The Results of the current suggested work have been compared with other published work based on accuracy and are shown in Table 6 and Table 7 to further confirm its efficacy. The effectiveness of the suggested technique has

been estimated considerable improvement over the comparable models, outperforming them by a heartening margin of 9% (highest) and 2% (lowest).

5 Conclusion

This study gives a deep learning model that uses U-Net and TensorFlow to identify water features in satellite images. Even in the event of tiny and blurry satellite photos, this model performs well. Due to U-Net's design, it also performs well when applied to water bodies found at the land-water interface. The significance of accurate datalabelling was also emphasized since it contributed to the model's good performance in the dataset for water segmentation. The model also makes use of the Adam Optimizer, which is the best available optimization tool. Additionally, great attention is paid to actions like scaling, utilizing cropping, making masks, etc. when doing data analysis. The model uses max pooling to increase performance and reduce image dimensionality. The model that comes next distinguishes the predicted region from the satellite images more precisely than other models presently in use.

6 Future Scope

This work can be further carried out for areas more prone to flood or regular water retention to analyse the yearly submerged area and also generate an improved mechanism for flood alarm utilizing the remote synthetic aperture radar images which would surpass the ground flood prediction difficulties like harsh weather etc. Prime motive of suggestion of utilizing SAR images is its zero cloud coverage images. Difficulties that would arise using SAR images would be high resolution processing and its size.

Conflict of Interest

The authors state that there are no conflicts of interest in the publishing of this literature survey study. Transparency and neutrality are critical to guaranteeing the integrity of this literature survey.

References

[1]K. Yuan, X. Zhuang, G. Schaefer, J. Feng, L. Guan, and H. Fang, “Deep-Learning-Based Multispectral Satellite Image Segmentation for Water Body Detection,” *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 14, pp. 7422–7434, 2021, doi: 10.1109/JSTARS.2021.3098678.

[2]Gayen, S., Villalta, I. V., & Haque, S. M. (2022). “Flood Risk Assessment and Its Mapping in PurbaMedinipur District, West Bengal”, India. *Water (Switzerland)*, *14*(7). https://doi.org/10.3390/w14071049

[3]H. Suliman, A. W. A. Hammad, and S. T. Waller, “A review on flood management technologies related to image processing and machine learning Automation in Construction A review on flood management technologies related to image processing and machine learning,” *Autom. Constr.*, vol. 132, no. December, p. 103916, 2021, doi: 10.1016/j.autcon.2021.103916.

[4]Z. Zhang, M. Lu, S. Ji, H. Yu, and C. Nie, “Rich cnn features for water-body segmentation from very high-resolution aerial and satellite imagery,” *Remote Sens.*, vol. 13, no. 10, 2021, doi: 10.3390/rs13101912.

[5]B. S. Kumar, C. H. Sripriya, B. R. Devi, S. K. Reshma, and N. C. Lahari, “Advanced CNN Approach to Predict Water Body Segmentation over Satellite Images,” vol. 10, no. 6, pp. 698–704, 2022.

[6]Yadavendra and S. Chand, “Semantic segmentation and detection of satellite objects using U-Net model of deep learning,” *Multimed. Tools Appl.*, vol. 81, no. 30, pp. 44291–44310, 2022, doi: 10.1007/s11042-022-12892-2.

[7]X. Zhang, J. Li, and Z. Hua, "MRSE-Net : Multiscale Residuals and SE-Attention Network for Water Body," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 15, pp. 5049–5064, 2022, doi: 10.1109/JSTARS.2022.3185245.

[8]X. Yin, L. Sun, Y. Fu, R. Lu, and Y. Zhang, "U-Net-Based Medical Image Segmentation," vol. 2022,*Journal of Healthcare Engineering* (Vol. 2022). Hindawi Limited. https://doi.org/10.1155/2022/4189781

[9]A. Ch, R. Ch, S. Gadamsetty, C. Iwendi, T. R. Gadekallu, and I. Ben Dhaou, "ECDSA-Based Water Bodies Prediction from Satellite Images with UNet," pp. 1–22, 2022.

[10]S. Mazhar *et al.*, "AUnet : A Deep Learning Framework for Surface Water Channel Mapping Using Large-Coverage Remote Sensing Images and Sparse Scribble Annotations from OSM Data," pp. 1–18, 2022.

[11]M. Xia, Y. Cui, Y. Zhang, Y. Xu, J. Liu, and Y. Xu, "DAU-Net: a novel water areas segmentation structure for remote sensing image," *Int. J. Remote Sens.*, vol. 42, no. 7, pp. 2594–2621, 2021, doi: 10.1080/01431161.2020.1856964.

[12]M. P. M. M, "Automatic Segmentation of River and Land in SAR Images : A Deep Learning Approach," 2019, doi: 10.1109/AIKE.2019.00011.

[13]F. Waldner and F. I. Diakogiannis, "Deep learning on edge: Extracting field boundaries from satellite images with a convolutional neural network," *Remote Sens. Environ.*, vol. 245, 2020, doi: 10.1016/j.rse.2020.111741.

[14]E. Science, "An effective modified water extraction method for Landsat-8 OLI imagery of mountainous plateau regions An effective modified water extraction method for Landsat-8 OLI imagery of mountainous plateau regions," pp. 0–7, 2016, doi: 10.1088/1755-1315/34/1/012010.

[15]S.Karmaker, S. Sheikh, S. Islam, and R. S. Nisha, "Watershed delineation in South Bengal Ganges Delta Region of Bangladesh using satellite imagery and digital elevation model," pp. 1–10, 2022, doi: 10.1007/s12517-022-10203-5.

[16]K.Rokni, A. Ahmad, A. Selamat, and S. Hazini, "Water Feature Extraction and Change Detection Using Multitemporal Landsat Imagery," pp. 4173–4189, 2014, doi: 10.3390/rs6054173.

[17]J. Del-pozo-vel, P. Chamorro-posada, J. Manuel, and P. Casaseca-de-la-higuera, "Water Detection in Satellite Images Based on Fractal Dimension," 2022.

[18]G. Kaplan and U. Avdan, "Object-based water body extraction model using Sentinel-2 satellite imagery," *Eur. J. Remote Sens.*, vol. 50, no. 1, pp. 137–143, 2017, doi: 10.1080/22797254.2017.1297540.

[19]G.Sarp and M. Ozcelik, "Water body extraction and change detection using time series: A case study of Lake Burdur, Turkey," *J. Taibah Univ. Sci.*, vol. 11, no. 3, pp. 381–391, 2017, doi: 10.1016/j.jtusci.2016.04.005.

[20]D.Filatov and G. N. A. H. Yar, "Forest and Water Bodies Segmentation Through Satellite Images Using U-Net," 2022, [Online]. Available: http://arxiv.org/abs/2207.11222.

[21]W.Boonpook*et al.*, "Deep Learning Semantic Segmentation for Land Use and Land Cover Types Using Landsat 8 Imagery," *ISPRS Int. J. Geo-Information*, vol. 12, no. 1, 2023, doi: 10.3390/ijgi12010014.

[22]F. Yang, T. Feng, G. Xu, and Y. Chen, "Applied method for water-body segmentation based on mask R-CNN," *J. Appl. Remote Sens.*, vol. 14, no. 01, p. 1, 2020, doi: 10.1117/1.jrs.14.014502.

[23]H. Tao, W. Li, X. Qin, and D. Jia, "Image semantic segmentation based on convolutional neural network and conditional random field," *Proc. - 2018 10th Int. Conf. Adv. Comput. Intell. ICACI 2018*, pp. 568–572, 2018, doi: 10.1109/ICACI.2018.8377522.

[24]Y. Wang, "Remote Sensing Image Semantic Segmentation Algorithm Based on Improved ENet Network," Hindawi Scientific Programming vol. 2021.

[25]H. Bao, D. Ming, Y. Guo, K. Zhang, K. Zhou, and S. Du, "DFCNN-based semantic recognition of urban functional zones by integrating remote sensing data and POI data," *Remote Sens.*, vol. 12, no. 7, 2020, doi: 10.3390/rs12071088.

[26] Jadon, S. (2020) "*A survey of loss functions for semantic segmentation*"https://doi.org/10.1109/CIBCB48159.2020.9277638

[27] Du Y, Zhang Y, Ling F, Wang Q, Li W, Li X. "Water Bodies' Mapping from Sentinel-2 Imagery with Modified Normalized Difference Water Index at 10-m Spatial Resolution Produced by Sharpening the SWIR Band". Remote Sensing. 2016; 8(4):354. https://doi.org/10.3390/rs8040354