**WATER BODY DETECTION AND ANALYSIS ON SATELLITE IMAGES**

*Bachelor thesis*

*by*

**Allu Vamsi Pavan Kumar**

**Vedant Gupta**

**Vineel Sai**

*Under the supervision of*

**Dr. Pratik Chakraborty**

**Dr. Oishila Bandhopadhyay**

****

*A thesis submitted to*

**Indian Institute of Information Technology, Kalyani**

*for the partial fulfillment of the degree of Bachelor of Technology*

*in*

**Department of Computer Science & Engineering**

**Autumn Semester 2021**

*Declaration*

We hereby declare that the work being presented in this thesis entitled “**Water body detection and analysis on Satellite images**” submitted to Indian Institute of Information Technology, Kalyani in partial fulfillment for the award of the degree of Bachelor of Technology in Computer Science and Engineering during the period from July,2021 to Nov, 2021 under the supervision of Dr. Pratik Chakraborty and Dr. Oishila Bandhopadhyay, Department of Computer Science and Engineering, Indian Institute of Information Technology, Kalyani, West Bengal 741235, India, does not contain any classified information.

Allu Vamsi Pavan Kumar (CSE/19010/437),

Vedant Gupta (CSE/19069/496),

Vineel Sai (CSE/19071/498)

Department: Computer Science & Engineering

Institute: Indian Institute of Information Technology, Kalyani

This is to certify that the above statement made by the candidate is correct to the best of my knowledge.

Signed :

Date :

*Certificate*

This is to certify that the thesis entitled “**Water body detection and analysis on Satellite images**” being submitted by Allu Vamsi Pavan Kumar (CSE/19010/437), Vineel Sai (CSE/19071/498) and Vedant Gupta (CSE/19069/496) an undergraduate student in the Department of Computer Science and Engineering, Indian Institute of Information Technology, Kalyani, West Bengal, India, for the award of Bachelor of Technology in Computer Science and Engineering is an original research work carried by them under my supervision and guidance. The thesis has fulfilled all the requirements as per the regulations of Indian Institute of Information Technology, Kalyani and in my opinion, has reached the standards needed for submission. The work, techniques and the result presented have not been submitted to any other University or Institute for the award of any other degree or diploma.

**Dr. Pratik Chakraborty,**

Assistant Professor

Department of Computer Science and Engineering

Indian Institute of Information Technology, Kalyani

**Dr. Oishila Bandhopadhyay,**

Assistant Professor

Department of Computer Science and Engineering

Indian Institute of Information Technology, Kalyani

Date:

*Acknowledgements*

We would like to express our sincere thanks and gratitude to our Super-visor Dr. Pratik Chakraborty, who gave us the golden opportunity to do this wonderful project and guided us immensely through the course of the project. We came to know about so many new things, all thanks to him.

We would like to express our sincere thanks to Dr. Oishila Bandhopadhyay, who guided us regarding the project and we were lucky that we have learnt a lots of concepts from her. We must thank to our classmates for their timely help and support for completion of this project. Last but not the least, we would like to thank all those who had helped directly or indirectly in this project.

Allu Vamsi Pavan Kumar (CSE/19010/437),

Vedant Gupta (CSE/19069/496),

Vineel Sai (CSE/19071/498)

Department: Computer Science & Engineering

Institute: Indian Institute of Information Technology, Kalyani

*Abstract*

The size, Shape, Connectivity of waterbodies can have drastic effects ecological communities and ecosystem processes. In this we have analyzed the water body detection capability of three NDWI models (SWIR1, SWIR2 and NIR) generated using Landsat-8 OLI multispectral satellite images on Hirakud Lake Region from the timestamp 25.05.2021. It analyses calculates the area of a particular Water body, and calculates the area and figures out the difference between the areas and change in the shape of the water bodies. This study mainly focuses on the changes in the Area of the waterbody occurred from past few years.

**Keywords :** Connected Components, Sobel Edge Detection, Normalized Differenced Waterbody Index (NDWI).

Contents

1. **Introduction ---------------------------------------------------------------- 7**
2. **Waterbody extraction techniques -------------------------------------------- 8**
   1. **Color Detection -------------------------------------------------------------- 8**
   2. **Multi Layer Perceptron ----------------------------------------------------- 8**
   3. **Maximum Likelihood ----------------------------------------------------- 9**
   4. **Normalized Differenced Water Index --------------------------------------- 9**
3. **NIR vs SWIR vs SWIR2 ---------------------------------------------------------- 9**
4. **Noise removal techniques ---------------------------------------------------------- 10**
   1. **Connected Components --------------------------------------------------------- 11**
   2. **Median Filter ------------------------------------------------------------ 11**
5. **Edge detection ------------------------------------------------------------- 12**
   1. **Sobel Edge Detection -------------------------------------------------------------- 12**
6. **Experimental results -------------------------------------------------------------------- 13**
7. **Discussion ------------------------------------------------------------------------------- 16**
8. **Future Scope ------------------------------------------------------------------------------ 17**
9. **References ---------------------------------------------------------------------------------- 17**
10. **Introduction**

Computer vision is a field of artificial intelligence that enables computers and systems to derive meaningful information from digital images, videos and other visual inputs and take actions or make recommendations based on that information. If artificial intelligence enables computers to think, computer vision enables them to see, observe and understand. We train computers on a massive amount of visual data computers process images, label objects on them, and find patterns in those objects.

A convolution Neural Network is most commonly applied to analyze visual images. These are based on the shared-weight architecture of the convolution kernels or filters that slide along input features and provide translation equivalent responses equivariant responses known feature maps.

A convolutional neural network consists of an input layer, [hidden layers](https://en.wikipedia.org/wiki/Multilayer_perceptron#Layers) and an output layer. In any feed-forward neural network, any middle layers are called hidden because their inputs and outputs are masked by the activation function and final [convolution](https://en.wikipedia.org/wiki/Convolution). In a convolutional neural network, the hidden layers include layers that perform convolutions.

1. **Waterbody Extraction Techniques**

Multiple methods including unsupervised classification, supervised classification, single-band threshold, inter spectrum relation method and water index method (normalized difference water index, modified normalized difference water index, and new water index) are used for the Water body analysis.

* 1. **Color Detection :**

We can perform color detection on the Satellite images by using functions in Open CV. In this method we convert the image frame in BGR to HSV (hue – saturation - value) color space, where hue describes a color in terms of saturation. Now we define the range off each color and create the corresponding mask and perform bitwise and operation between the image frame and mask to specifically detect that particular color and discard others.

* 1. **Multi Layer perceptron:**

In this method we first collect training samples and manually label them as water types and non- water types and perform training. With the trained model the probability of water and non-water types for each pixel are computed. The classification depends on the probability value.

* 1. **Maximum Likelihood:**

Maximum likelihood estimation (MLE) is a method of [estimating](https://en.wikipedia.org/wiki/Estimation_theory) the [parameters](https://en.wikipedia.org/wiki/Statistical_parameter) of an assumed [probability distribution](https://en.wikipedia.org/wiki/Probability_distribution), given some observed data. This is achieved by [maximizing](https://en.wikipedia.org/wiki/Mathematical_optimization) a [likelihood function](https://en.wikipedia.org/wiki/Likelihood_function) so that, under the assumed [statistical model](https://en.wikipedia.org/wiki/Statistical_model), the [observed data](https://en.wikipedia.org/wiki/Realization_(probability)) is most probable.

* 1. **NDWI:**

Normalized difference Water Index (NDWI) is the most commonly used technique for water body detection. Normalized difference water index(NDWI) is the most used index to enhance the water information.

1. **NIR vs SWIR1 vs SWIR2**

Normalized Difference Water Index (NDWI) may refer to one of at least two remote sensing indexes related to liquid water.

For NIR

NDWI =

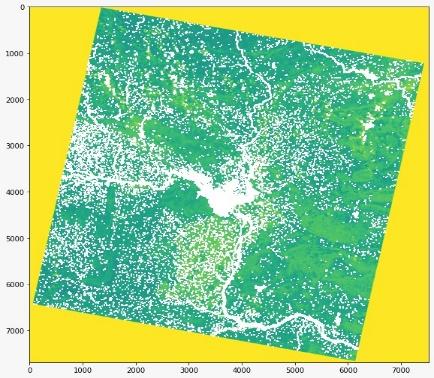
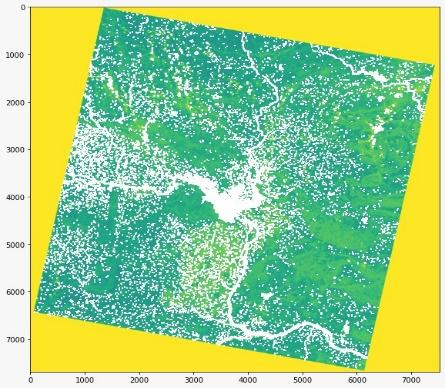
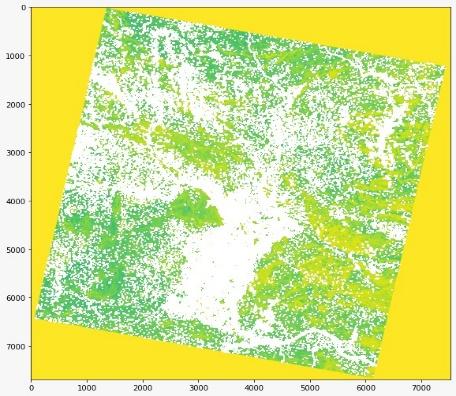
For SWIR1

NDWI =

For SWIR2

NDWI =

We have used NDWI model to extract the water bodies from the satellite images of the Hirakud dam area. The extraction is done in on NIR ,SWIR1 and SWIR2 bands. Here are the results of the analysis NIR, SWIR1,SWIR2 respectively.

1. **NOISE REMOVAL TECHNIQUES**

Noise is always presents in digital images during image acquisition, coding, transmission, and processing steps. Filtering image data is a standard process used in almost every image processing system. Filters are used for this purpose. They remove noise from images by preserving the details of the same.

* 1. **Connected components analysis:**

Connected components labeling scans an image and groups its [pixels](https://homepages.inf.ed.ac.uk/rbf/HIPR2/pixel.htm) into components based on [pixel connectivity](https://homepages.inf.ed.ac.uk/rbf/HIPR2/connect.htm), *i.e.* all pixels in a connected component share similar [pixel intensity values](https://homepages.inf.ed.ac.uk/rbf/HIPR2/value.htm) and are in some way connected with each other. Once all groups have been determined, each pixel is labeled with a gray level or a color according to the component it was assigned to. Connected component labeling works by scanning an image, pixel-by-pixel in order to identify connected pixel regions, regions of adjacent pixels which share the same set of intensity values.

* 1. **Median Filter:**

The median filter is a non-linear [digital filtering](https://en.wikipedia.org/wiki/Digital_filter) technique, often used to remove [noise](https://en.wikipedia.org/wiki/Signal_noise) from an image or signal. Median filtering is very widely used in digital [image processing](https://en.wikipedia.org/wiki/Image_processing) because, under certain conditions, it preserves edges while removing noise. The median filter works by moving through the image pixel by pixel, replacing each value with the median value of neighbouring pixels. The median is calculated by first sorting all the pixel values from the window into numerical order, and then replacing the pixel being considered with the middle (median) pixel value.

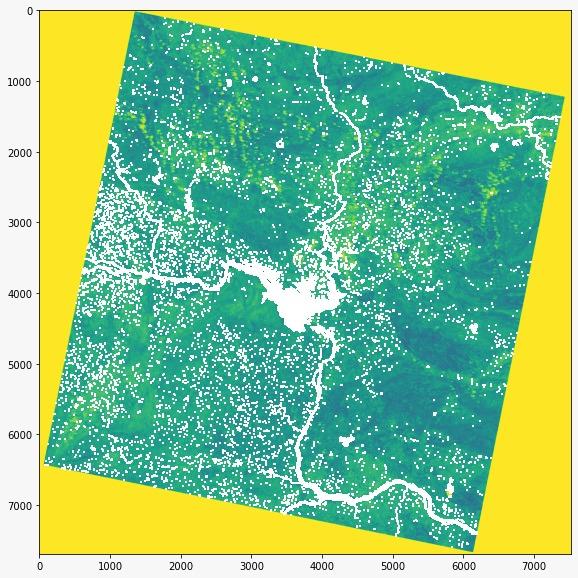
1. **Edge detection**

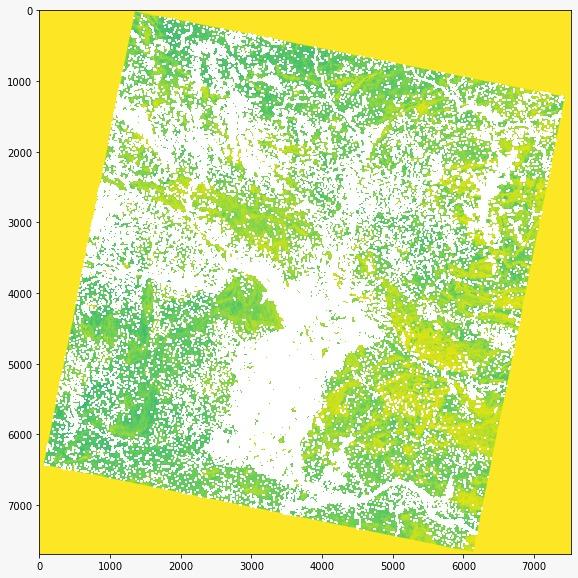
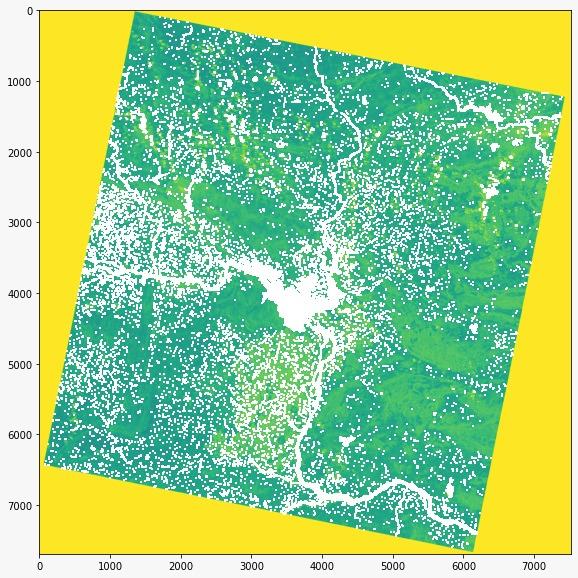
[Edge detection](https://www.mathworks.com/products/image.html) is an image processing technique for finding the boundaries of objects within images. It works by detecting discontinuities in brightness. Edge detection is used for [image segmentation](https://www.mathworks.com/discovery/image-segmentation.html) and data extraction in areas such as image processing, computer vision, and machine vision.

* 1. **Sobel Edge Detection:**

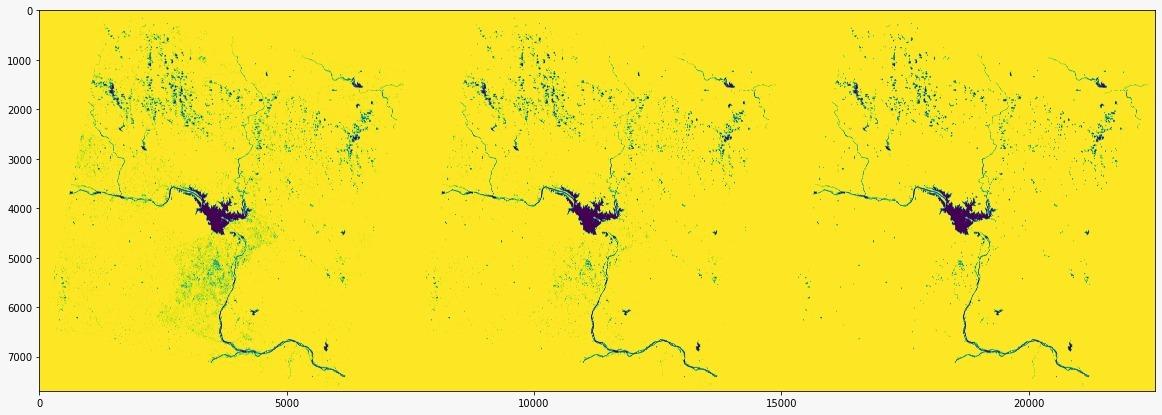
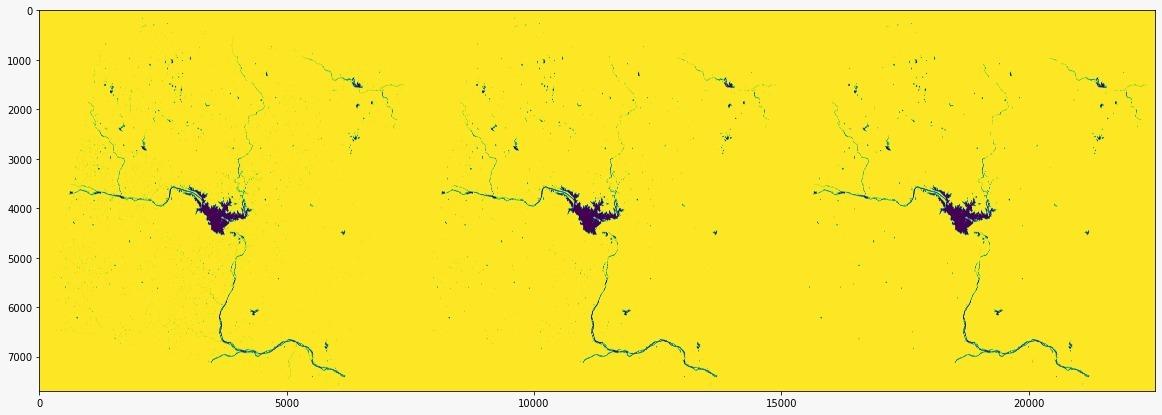
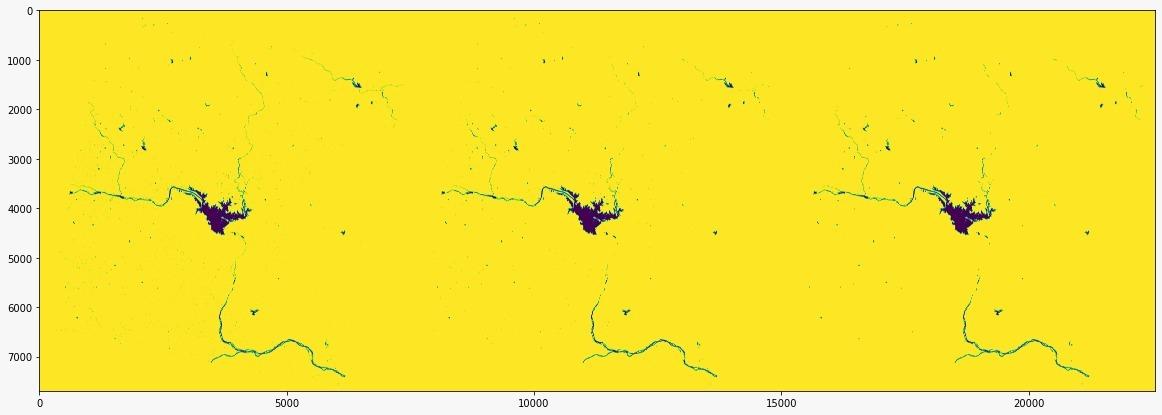
The Sobel operator performs a 2-D spatial gradient measurement on an image and so emphasizes regions of [high spatial frequency](https://homepages.inf.ed.ac.uk/rbf/HIPR2/freqdom.htm) that correspond to edges. It is used to find the approximate absolute gradient magnitude at each point in an input grayscale image. The operator consists of a pair of 3×3 [convolution kernels](https://homepages.inf.ed.ac.uk/rbf/HIPR2/convolve.htm). These kernels are designed to respond maximally to edges running vertically and horizontally relative to the pixel grid, one kernel for each of the two perpendicular orientations. The kernels can be applied separately to the input image, to produce separate measurements of the gradient component in each orientation (call these *Gx* and *Gy*).

1. **Experimental Results**

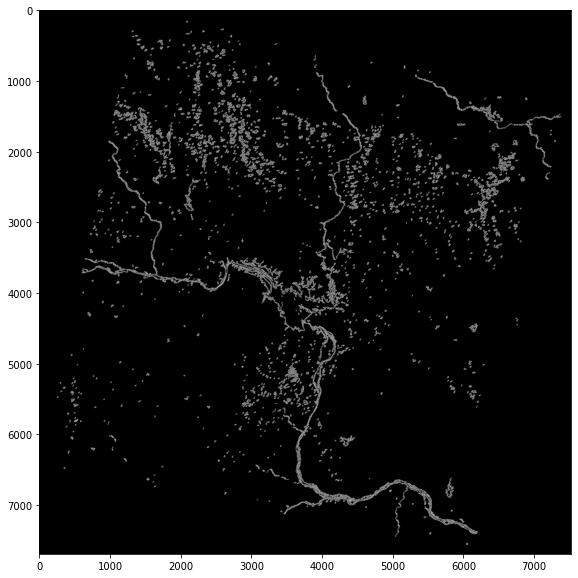
* **Applying NDWI on NIR, SWIR1, SWIR2**



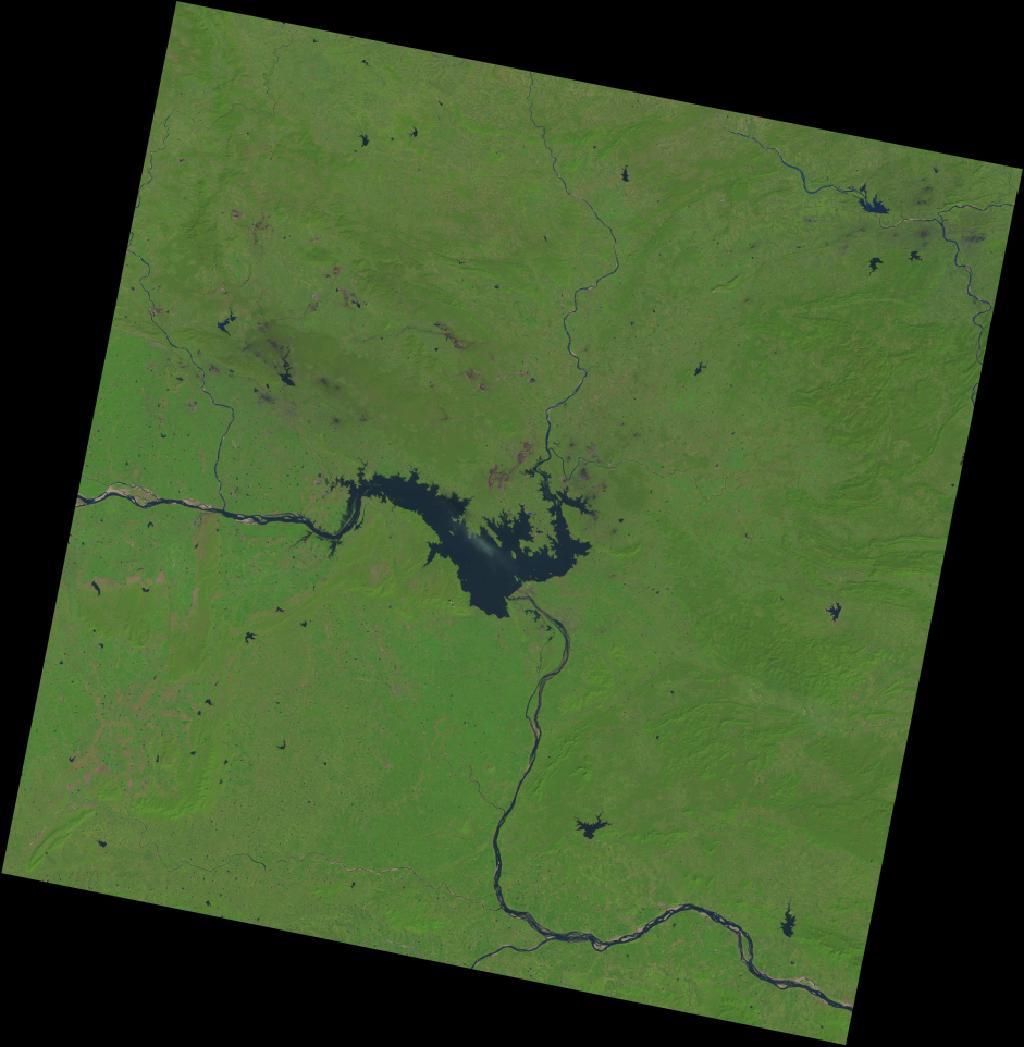
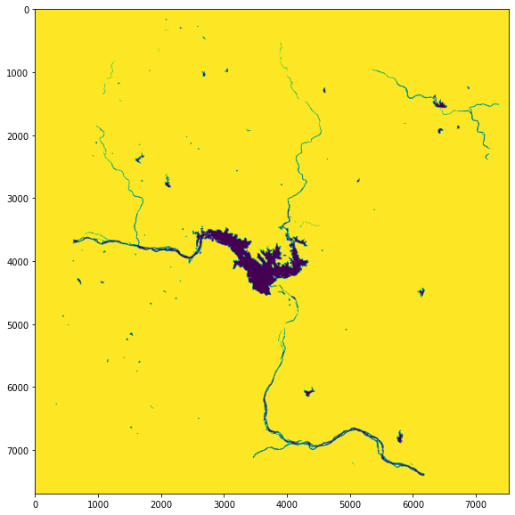
* From the meta data provided by the Earth Explorer i.e. the Latitudes and Longitudes Area of the tile is calculated which helps in the calculation of the area covered by the waterbody.
* Area of the Tile = 323
* And Calculated the Area covered by Water bodies in the tile. (i.e. the area of the white pixels in the tile.)
* = 2.6
* = 3.4
* = 6.3
* It can be observed that there are a lot of noise in SWIR2 band images.
* **Noise Removal Techniques**



* NIR
  1. Base : 3.3909186207713016 KM2
  2. Median Filter : 3.1232392062161805 KM2
  3. Connected Components : 3.083140621074842 KM2
* SWIR1
  1. Base : 4.09051357105924 KM2
  2. Median Filter : 3.6689060687717525 KM2
  3. Connected Components : 3.583258930643658 KM2
* SWIR2
  1. Base : 7.941295564721472 KM2
  2. Median Filter : 6.313888260618718 KM2
  3. Connected Components : 5.690112637590087 KM2
* **Edge Detection**

1. **Discussion**

Now we have performed analysis on the same region which was collected on 03/11/2021. (The Analysis was performed on SWIR1 band)   

* We calculated the area of the water bodies.
  + = 5.36
* Previously calculated Area was 3.58
* So the Area of the water body have increased by 2 .
* The amount of water in rivers and lakes is always changing due to inflows and outflows. Inflows to these water bodies will be from precipitation overland runoff, groundwater seepage, and tributary inflows. Outflows from lakes and rivers include evaporation.
* Moreover May is the peak of Summer season which resulted in the decrease of Water level from the Hirakud lake region.
* So this might be the reason for the increase in the area of the Water region in the November

1. **Future Scope**

* We would perform analysis on the previous satellite images of this region and collect information such as Area and other small connected water bodies near it.
* We further would develop a machine learning model which would predict the future changes that can be occurred in that region.
* Generalize the model for any Water body region**.**

1. **References**

* NDWI - [Link](https://en.m.wikipedia.org/wiki/Normalized_difference_water_index)
* Sobel operator - [Link](https://en.m.wikipedia.org/wiki/Sobel_operator)
* Improved Sobel Edge Detection - [Link](https://ieeexplore.ieee.org/abstract/document/5563693)
* Earth Explorer - [Link](https://earthexplorer.usgs.gov/)