ULOG3

Image Processing Team

Project Title**:**

**Water Body Segmentation Using NIR Imagery**

***Team Members:***

1. **Ranjitha – Team Leader**
2. **Swedha Varshini C**
3. **Mohammed Riyan H**

***Abstract:***

This project aims to develop an **automated system** for segmenting **water** and **land** **areas** from **Near**-**Infrared** (**NIR**) **satellite** **images**. By utilizing the distinct **reflectance** **properties** **of** **NIR**, where **water** **bodies** **absorb** and **land** **surfaces** **reflect**, the system generates **precise** **segmentation** **masks**. The solution enhances **large**-**scale** **environmental** **monitoring**, enabling **accurate** **water** **resource** **management** and reducing **manual** **effort** through **advanced** **image** **processing** and **machine** **learning** **techniques**.

***Objective:***

The objective of this project is to **design** and **implement** an **automated** **system** that accurately **detects** and **quantifies** **water** and **land** **areas** from **Near**-**Infrared** (**NIR**) **satellite** **images**. By leveraging the unique **reflectance** **properties** of **NIR**, the system will generate **precise** **segmentation** **masks**, enabling **scalable** **environmental** **monitoring**, **efficient** **water** **resource** **management**, and minimizing **manual** **intervention** **in** **geographic** **analysis**.

***Vision:***

To create an **intelligent** and **automated** **system** that enables **accurate** and **large**-**scale** **monitoring** of Earth’s **water** and **land** **resources** using **satellite** **imagery**, providing reliable **geographic** **insights** to support **sustainable** **development**, **environmental** **management**, and **data**-**driven** **decision**-**making** through **advanced** **image** **processing** and **machine** **learning** **technologies**.

***Mission:***

To develop an **automated** **NIR** **image** **segmentation** **system** that **accurately** distinguishes **water** and **land** areas, utilizing **deep** **learning** **techniques** to deliver **precise**, **scalable**, and **efficient** **solutions** for **environmental** **monitoring** and **resource** **management**.

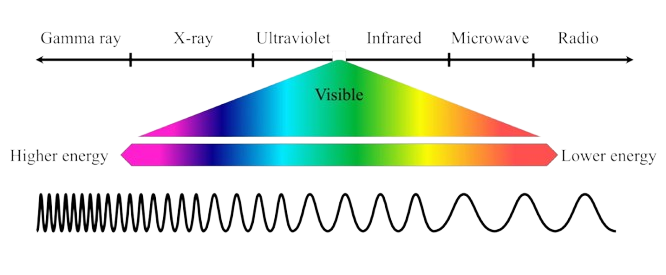
*1. Introduction:*

* **Satellite imagery** has become an essential tool for **observing** and **analysing** **Earth's** **surface**, enabling applications in **environmental** **monitoring**, **urban** **planning**, **disaster** **management**, and **resource** **assessment**. Among various imaging technologies, **Near**-**Infrared** (**NIR**) imaging plays a crucial role in **differentiating** **between** **water** **bodies** and **land** **surfaces** due to its **unique** interaction with **moisture** and **vegetation**.
* The **NIR** **spectrum**, ranging from **700nm** to **1100nm**, is invisible to the **human** **eye** but is **highly** **effective** in **remote** **sensing**. **Water** bodies exhibit **high** **absorption** of **NIR** **wavelengths**, making them appear **dark** in NIR images, whereas **land surfaces**, especially **vegetation, reflect NIR strongly** and **appear** **bright**. These **distinct** **reflectance** **properties** allow for **clear** **separation** between **water** and **land** **regions**, which is often **challenging** **in** **visible** **spectrum** **images** due to **atmospheric** **conditions**, **turbidity**, or **varying surface textures.**
* **Manual interpretation and segmentation** of satellite images are **time-consuming**, **prone to errors**, and **not scalable** for large datasets. The need for an **automated**, **accurate**, and **efficient** system to classify **water** **and** **land** **areas** has become **critical**, especially for applications involving **real**-**time** **environmental monitoring and resource management.**
* This project focuses on leveraging the **physics** of **NIR** **reflectance** in combination with **deep** **learning**-based **image** **processing** **algorithms** to develop a **robust** and **automated** **segmentation** **system**. By generating **precise** **segmentation** **masks** that **differentiate** **between** **water** **bodies** and **land** **surfaces**, the system aims to facilitate large-scale analysis with minimal manual intervention, ensuring high accuracy across diverse terrains and environmental conditions.

*2. Physics Of NIR Reflectance:*

*2.1 What Is NIR:*

**NIR** stands for **Near Infrared**, which refers to a segment of the **electromagnetic spectrum** with wavelengths from approximately **750 nm to 2500 nm**. It is commonly used in **Near-Infrared Spectroscopy (NIRS)**, an **analytical technique** that utilizes near-infrared radiation to analyze the **chemical and physical properties** of samples, both in **solids and liquids**. NIR is **invisible to the human eye** and is positioned just beyond the **visible red**.



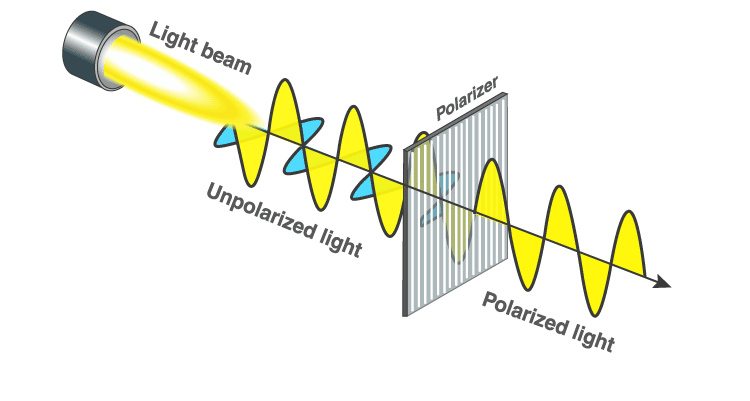
**Fig 1.** Electromagnetic Spectrum

***2.2 Why NIR:***

* **Non-destructive sensing** – Doesn't damage materials.
* **High penetration ability** – Passes through atmospheric particles like haze, smoke, and clouds better than visible light.
* **Thermal independence** – Unaffected by temperature variations during daytime or nighttime.
* **High contrast between materials** – Especially between water and land surfaces due to their differing absorption and reflection rates.

***2.3 How NIR Camera Works:***

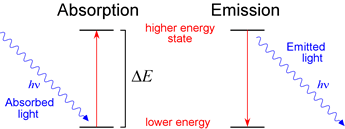
NIR (Near-Infrared) cameras work by capturing the **reflected or emitted radiation** from objects in the **near-infrared spectrum** (typically 700–1100 nm). To ensure only NIR wavelengths are observed from full-spectrum sunlight, the camera uses **gratings** (to split light by wavelength) and **polarizers** (to filter specific orientations of light waves). This helps isolate the desired NIR band and eliminate unwanted visible light or noise.



**Fig 2.** Working of Polarizers

When NIR light interacts with materials on Earth’s surface, four primary behaviours can occur:

1. **Absorption** – The material absorbs the NIR light and converts it to internal energy.
2. **Emission** – The material emits radiation, including NIR, depending on its temperature or properties.
3. **Absorption + Emission** – Some materials first absorb and then re-emit part of the energy.
4. **Reflection** – The surface reflects the incoming NIR radiation without absorption.

****

**Fig 3.** Photoelectric Effect

***2.4 Grayscale Representation***

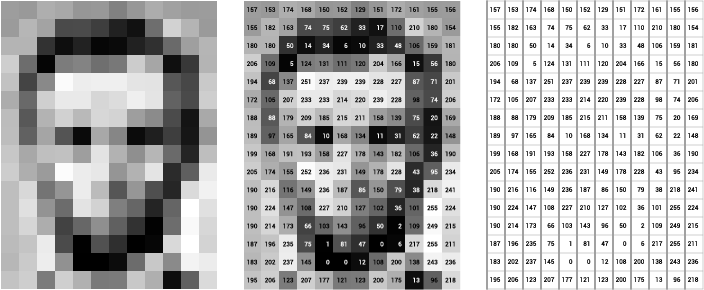
Each pixel in an NIR image corresponds to a brightness value from **0 to 255**:

* **0 (Pure Black)** → Indicates **very low reflection or emission**, meaning most of the NIR was absorbed (e.g., **water**).
* **255 (Pure White)** → Indicates **high reflection or emission**, meaning the surface strongly reflects NIR (e.g., **dry soil, land, vegetation**).

This physical interaction is the **core principle** behind distinguishing materials like water and land in NIR images:

* **Water absorbs NIR heavily** → reflects less → appears **darker**.
* **Land reflects more NIR** → appears **brighter**.

This difference in pixel brightness forms the **basis for image segmentation and classification** in NIR-based Earth observation.



**Fig 4.** Pixel Intensity Representation

***2.5 Mathematical Representation:***

When Near-Infrared (NIR) radiation interacts with a surface, the total incident energy **E**in distributed into three components:

**E**in = **E**abs + **E**ref + **E**emi

Where:

**E**in - Incident NIR energy

**E**ref - Energy Reflected

**E**abs - Energy Absorbed by material

**E**emi - Energy Emitted

For Earth surface materials in **NIR satellite imaging**, we usually assume **no significant emission** (especially during the day), so:

**E**in ≈ **E**abs + **E**ref

Reflectance Coefficient **R**:

Where:

* **R** ∈ [0,1]
* High Ref → High reflection (e.g., land/vegetation)
* Low Ref→ Low reflection (e.g., water)

In image processing, this reflectance **R** is **mapped to pixel intensity** I ∈ [0,255]

**Ipx =** 255x **R**

**Thus:**

**R ≈** 0 **🡪 I ≈** 0 **🡪 Black pixel** (low reflectance, water)

**R ≈** 1 **🡪 I ≈** 255 **🡪 White pixel** (high reflectance, land)

**Case Study:** Interaction of NIR Light with Different Surfaces

|  |  |  |  |
| --- | --- | --- | --- |
| Surface Type | NIR Reflectance (%) | Pixel Value (0–255) | Appearance |
| Water | 5% | ~13 | Very Dark |
| Bare Soil | 30–40% | ~76–102 | Mid-gray |
| Vegetation | 50–60% | ~127–153 | Bright |
| Snow/Ice | 80–90% | ~204–230 | Very Bright |

**a) Case 1:** Water Surface

* **Input NIR Light**: 100%
* **Absorption**: 60%
* **Reflection/Emission**: 40%
* **Behaviour**:
  + Water absorbs a significant portion of the incoming NIR radiation due to its molecular structure (hydrogen bonds absorb longer wavelengths).
  + Very little light is reflected back to the sensor.
* **Pixel Value**:

**I**px = **R** × 255 =0.4×255=102

Appears **mid-Black** in the image.

**b) Case 2:** Land Surface

* **Input NIR Light**: 100%
* **Absorption**: 40%
* **Reflection/Emission**: 60%
* **Behaviour**:
  + Soil and vegetation reflect a large portion of NIR.
  + Especially healthy vegetation reflects more NIR due to internal leaf structures (spongy mesophyll).
* **Pixel Value**:

**I**px = **R** x 255 = 0.6 x 255 = 153

Appears **grey** in the image.



**Fig 5.** Comparison of RGB and NIR Views

***2.6 Physics Final Word***

**I**px ∝

***2.7 Impact on CNN-based Water-Land Segmentation***

* This **consistent contrast** is exploited by **Convolutional Neural Networks (CNNs)** and **segmentation models**.
* The model learns from these physical differences during training:
  + **Lower pixel regions** = Water
  + **Higher pixel regions** = Land

*3. Methodology*

*3.1 Dataset Collection*

* The dataset consists of **Near-Infrared (NIR)** satellite images sourced from **Google Cloud Storage**.
* Images are in **TIFF (.tif) format**, preserving **high-resolution spectral data** for accurate segmentation.
* Each image has a spatial resolution of **516 × 513 pixels**.
* The dataset includes **various geographic locations** and **environmental conditions** for better generalization.
* The dataset is split into **70% for training**, **15% for validation**, and **15% for testing**.
* This split ensures **balanced and effective training**, **fine-tuning**, and **evaluation**.

***3.1.1 Summary:***

|  |  |  |
| --- | --- | --- |
| Dataset Partition | Number of Images | Percentage |
| Training Set | *X* images | 70% |
| Validation Set | *Y* images | 15% |
| Test Set | *Z* images | 15% |

***3.2 Data Preprocessing***

To ensure **consistent input** to the segmentation model and enhance **training efficiency**, a comprehensive **preprocessing pipeline** was implemented for the **Near-Infrared (NIR) satellite images**. The key steps include:

**a) Recursive Folder Scanning & File Handling**

* A **custom data generator** was designed to **recursively scan** the dataset directory, identifying all relevant image files in formats such as **TIFF (.tif/.tiff)**, **JPG**, **JPEG**, and **PNG**.
* This **dynamic and flexible approach** ensures compatibility with **varied dataset organizations**.

**b) NIR Band Extraction & Normalization**

* For **multi-band TIFF images**, the **NIR band** (usually the **last channel**) is extracted.
* The extracted data is **normalized to a 0–255-pixel intensity range** using **OpenCV’s normalization** functions.
* This **normalization** ensures **uniform pixel values**, which is **critical for stable and consistent model training**.

**c) Image Resizing**

All images are **resized to 256×256 pixels**, maintaining a balance between:

* **Retaining spatial detail**
* **Reducing computational cost**

This ensures a **uniform input shape** to the CNN model.

**d) On-the-Fly Binary Mask Generation (Otsu Thresholding)**

Instead of relying on **manually labelled masks**, a **dynamic mask generation strategy** using **Otsu’s thresholding** is employed. For each input image:

* **Water regions** (high **NIR absorption**) → **Foreground (1)**
* **Land regions** (high **NIR reflectance**) → **Background (0)**

This **real-time mask generation** provides a **robust and automated labelling mechanism**, ideal for **scaling** the segmentation pipeline.

**e) Data Augmentation (Optional for Future Work)**

Although basic **normalization** and **resizing** are implemented, **data augmentation** (e.g., **flipping**, **rotation**, **scaling**) is identified as a potential enhancement to:

* Improve **model generalization**
* Reduce **overfitting risks**

This can be incorporated in **future iterations** of the pipeline.

***3.3 CNN Model Architecture***

For the task of **water-land segmentation**, a **Convolutional Neural Network (CNN)** based **encoder-decoder architecture** was designed, optimized for **binary segmentation** of **NIR satellite images**.

**a) Encoder (Down sampling Path)**

The **encoder** extracts **hierarchical spatial features** using a series of convolutional and pooling layers:

* **Conv2D layers** with **32, 64, and 128 filters**, using **3×3 kernels**, activated by **ReLU**, to learn spatial patterns.
* **Batch Normalization** is applied after each convolution to **stabilize training** and **speed up convergence**.
* **MaxPooling2D layers** down sample feature maps by a factor of 2, enabling the model to learn **spatial hierarchies** and **contextual features**.

**b) Decoder (Up sampling Path)**

The **decoder** reconstructs the feature maps to match the **original input resolution**, generating the **segmentation mask**:

* **UpSampling2D layers** increase spatial resolution step-by-step.
* Each up sampling is followed by **Conv2D layers** (with **64** and **32 filters**) and **ReLU activation**, to **refine features**.
* **Batch Normalization** ensures **model stability** during reconstruction.

**c) Output Layer**

The final layer is a **1×1 Conv2D layer** with a **sigmoid activation function**, which outputs a **binary mask**:

* Each pixel represents the **probability of being Water (foreground = 1)** or **Land (background = 0)**.

|  |  |
| --- | --- |
| Layer Type | Details |
| Input Layer | **256x256x1 grayscale NIR image** |
| Conv2D + Batch Norm | **32 filters, 3x3 kernel** |
| MaxPooling2D | **2x2 pool size** |
| Conv2D + Batch Norm | **64 filters, 3x3 kernel** |
| MaxPooling2D | **2x2 pool size** |
| Conv2D + Batch Norm | **128 filters, 3x3 kernel** |
| UpSampling2D | **2x2 up sample** |
| Conv2D + Batch Norm | **64 filters, 3x3 kernel** |
| UpSampling2D | **2x2 up sample** |
| Conv2D + Batch Norm | **32 filters, 3x3 kernel** |
| Output Layer | **1 filter, 1x1 kernel, Sigmoid Activation** |

***3.4 Training Strategy***

To ensure **efficient** **learning** and **accurate segmentation**, a **well-structured training strategy** was implemented. This strategy involved:

* Selecting an appropriate **loss function** — **Binary Cross-Entropy (BCE),** ideal for pixel-wise binary classification
* Using the **Adam optimizer**, known for **adaptive learning rates** and **fast convergence**
* Integrating **callbacks** such as **Early Stopping** and **ReduceLROnPlateau** to:
  + **Prevent overfitting**
  + **Stabilize training**
  + **Optimize learning dynamics**

This combined setup enabled the model to **generalize effectively, minimize training loss**, and maintain **robust performance** across unseen satellite images.

**a) Loss Function**

* The task of **water-land segmentation** is formulated as a **binary classification** **problem at the pixel level**. Therefore**, Binary Cross-Entropy (BCE) Loss** was used as the **loss function**.
* BCE is well-suited for **segmentation tasks** where the output mask consists of two classes:
  + **Water → Foreground = 1**
  + **Land → Background = 0**
* This loss function **penalizes incorrect pixel classifications**, effectively **guiding the model** toward **accurate mask generation** by minimizing the difference between predicted and true class probabilities at each pixel.

**b) Optimizer**

* The model was trained using the **Adam Optimizer**, an **adaptive learning rate optimization algorithm** known for its **efficiency and faster convergence** in deep learning tasks.
* Adam combines the advantages of both **AdaGrad** (adaptive learning for sparse data) and **RMSProp** (adaptive learning for non-stationary objectives), making it particularly effective for handling **sparse gradients** and **non-stationary loss** **landscapes** — common in segmentation problems.

**c) Training Parameters**

|  |  |
| --- | --- |
| Parameter | Value |
| Batch Size | 8 |
| Epochs | 10 |
| Learning Rate | Adaptive (Reduced on Plateau) |
| Input Image Size | 256 x 256 x 1 (Grayscale NIR) |

**d) Callbacks to Enhance Training**

To **prevent overfitting** and ensure **optimal model convergence**, the following **callbacks** were integrated into the **training loop**:

* **Early Stopping**:  
  Monitors the **loss** and **halts training** if it does not improve after **10 consecutive epochs**. Automatically **restores the best model weights**, preventing performance degradation.
* **ReduceLROnPlateau**:  
  Monitors training stagnation and **reduces the learning rate by a factor of 0.5** if the loss **plateaus for 5 epochs**. This enables **finer weight updates** during flat regions of the training curve.

**e) Training Execution**

The model was trained on the **pre-processed dataset** using the **custom data generator**, ensuring **efficient batch load**ing and **on-the-fly mask generation.** The training was performed for **10 epochs**, with **callbacks** actively managing **learning dynamics.**

***3.5 Post-processing & Output Generation***

* After the **CNN** **model** generates the **initial** **segmentation** **masks**, a series of **post**-**processing** **steps** are applied to **refine** the **results** and **extract** **meaningful** **insights**, such as **area** **percentages** of **water** and **land** **regions**.
* These steps ensure that the outputs are **not** **just** **binary** **masks**, but transformed into **actionable** **geographic** **intelligence** — suitable for applications in **environmental** **analysis**, GIS **systems, and real-time decision-making.**

**a) Binary Mask Thresholding**

The model outputs **probability** **maps** (values between **0** and **1**) for each pixel. A **fixed** **threshold** (commonly **0.5**) is applied to **convert** these probability maps into **binary segmentation masks**, where:

* **Pixel value ≥ 0.5** → Classified as **Water** **(Foreground = 1)**
* **Pixel value < 0.5** → Classified as **Land (Background = 0)**

**b) Morphological Operations (Optional Enhancement)**

To remove small noise and refine segmentation boundaries, morphological operations such as:

* **Opening** (Erosion followed by Dilation): Removes **small false positives**.
* **Closing** (Dilation followed by erosion): Fills **small** **holes** in water regions.  
  These steps **enhance mask clarity**, especially in **noisy or complex terrains.**

**c) Area Percentage Calculation**

* Post **segmentation**, the total **pixel** **count** for **water** and **land** regions is computed to determine the **percentage** **area** **coverage**:
* This **quantification** enables **objective** **assessment** of **water** **body** **coverage** in each satellite image.

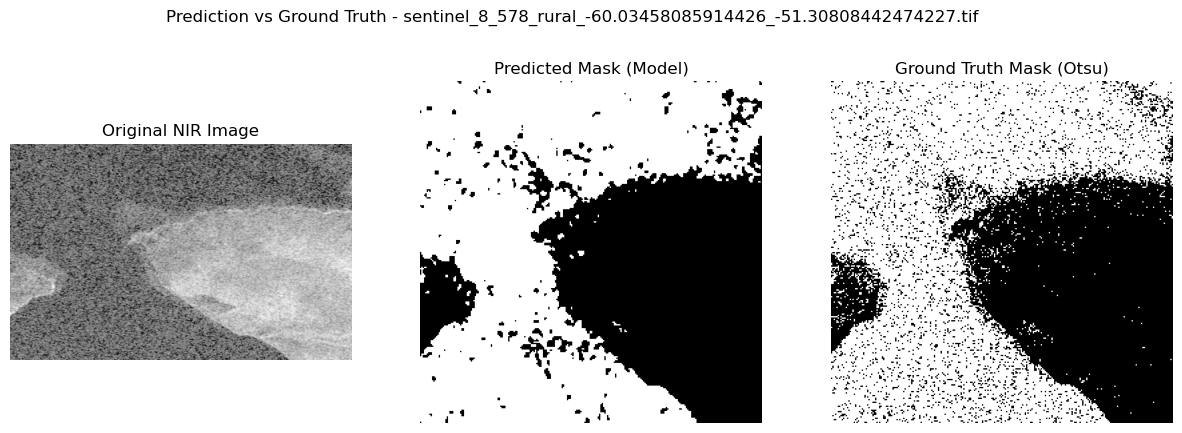
**d) Visualization of Results**

The processed outputs are visualized through:

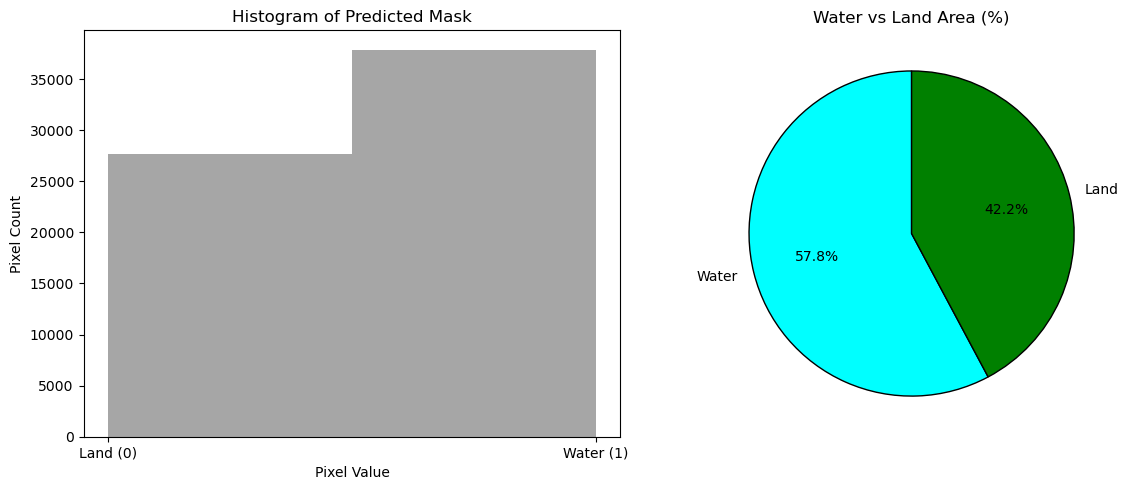
* **Overlayed Segmentation Masks** on **Original** **NIR** **Images** (for **qualitative** inspection)
* **Histograms** of **Pixel** **Distributions** (showing **intensity** **spread**)
* **Pie** **Charts** depicting the **calculated** **Water** vs **Land** **area** **percentages** for **clear** **representation** of segmentation results.

***Test 01***

**Mask Prediction**

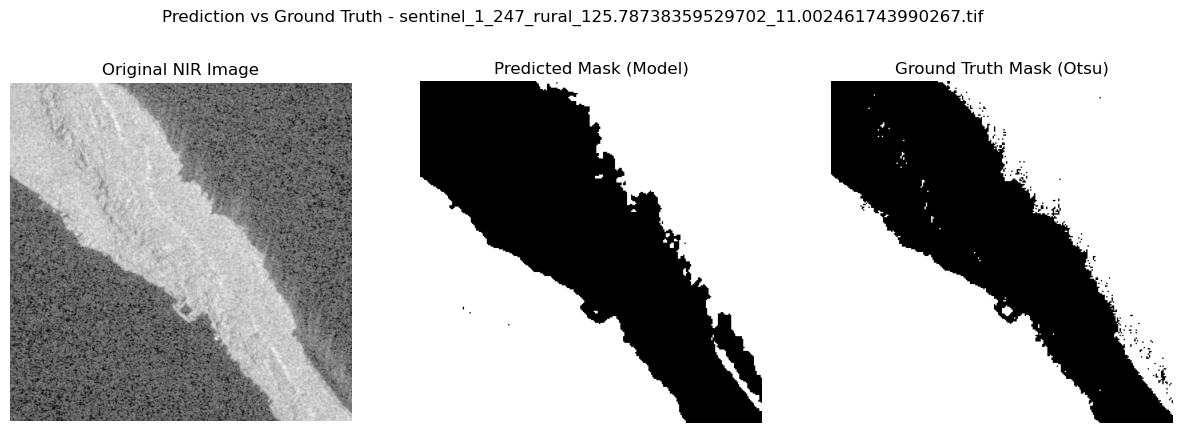
****

**Information From NIR Image**

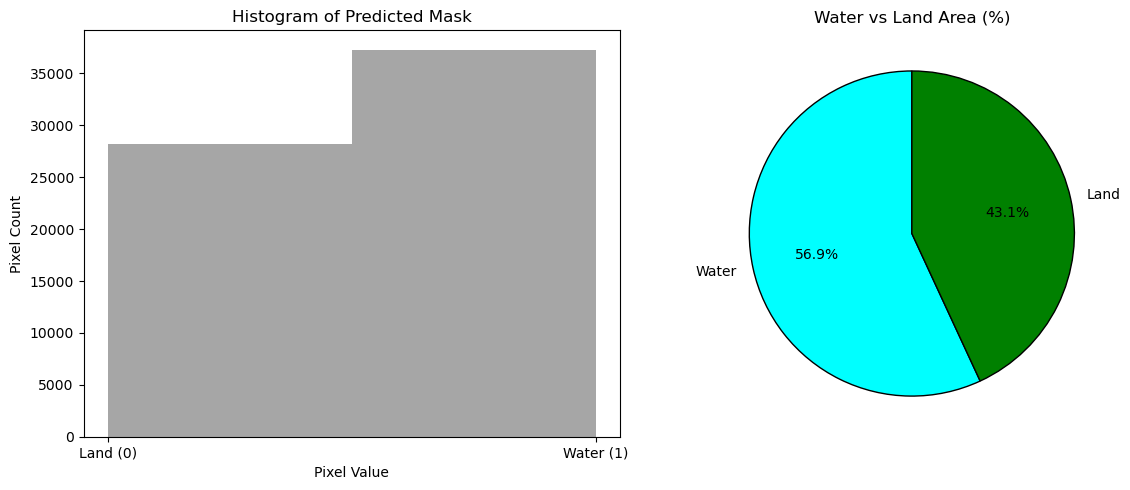
****

***Test 02***

**Mask Prediction**

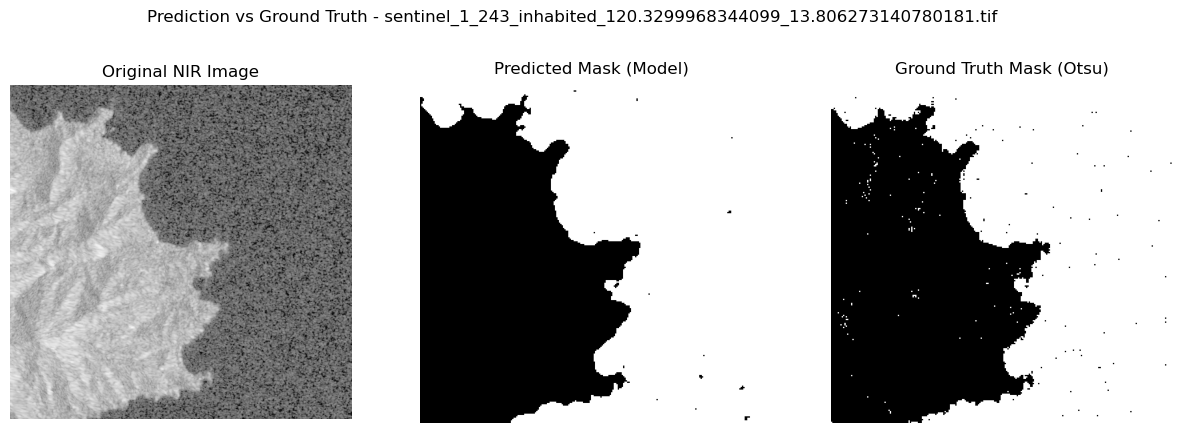
****

**Information From NIR Image**

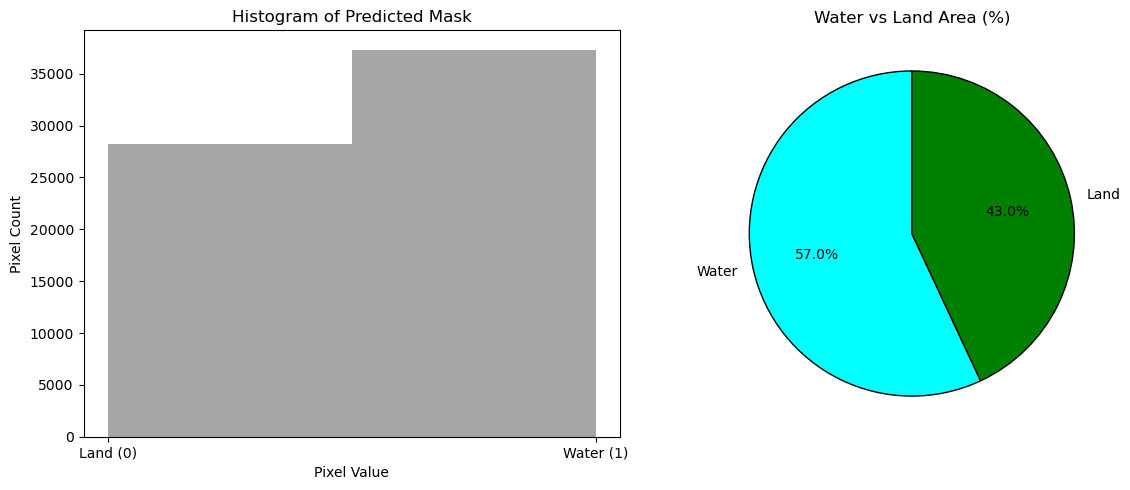
****

***Test 03***

**Mask Prediction**

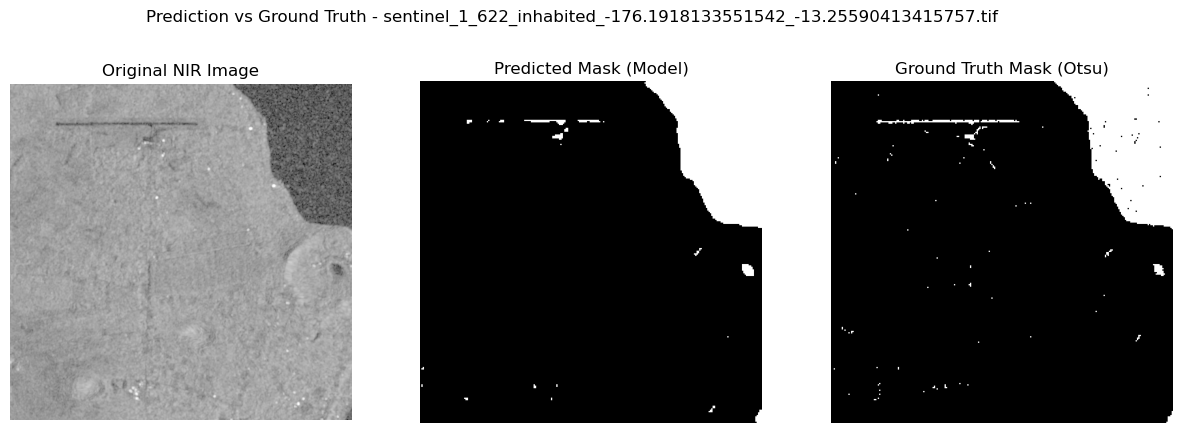
****

**Information From NIR Image**

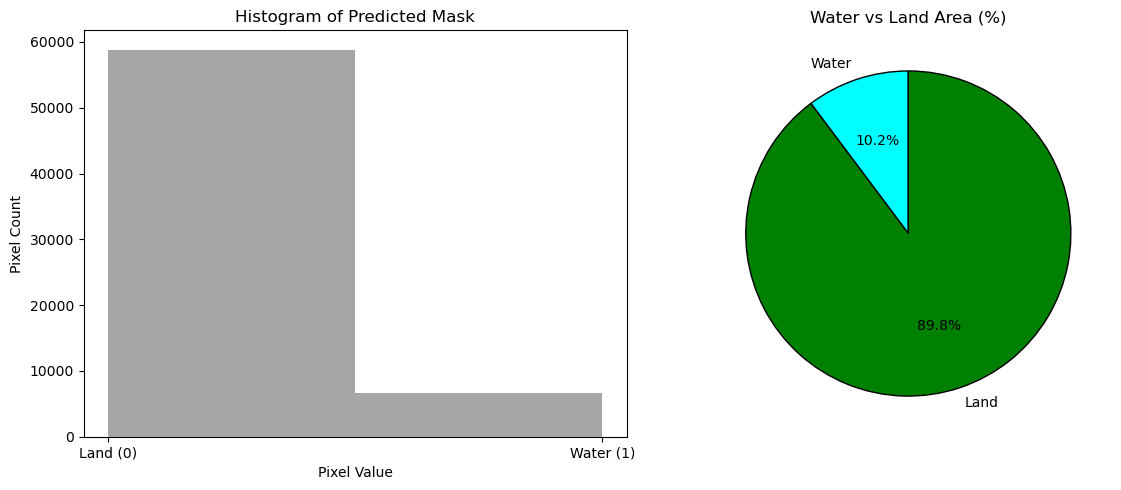
****

**Test 04**

**Mask Prediction**

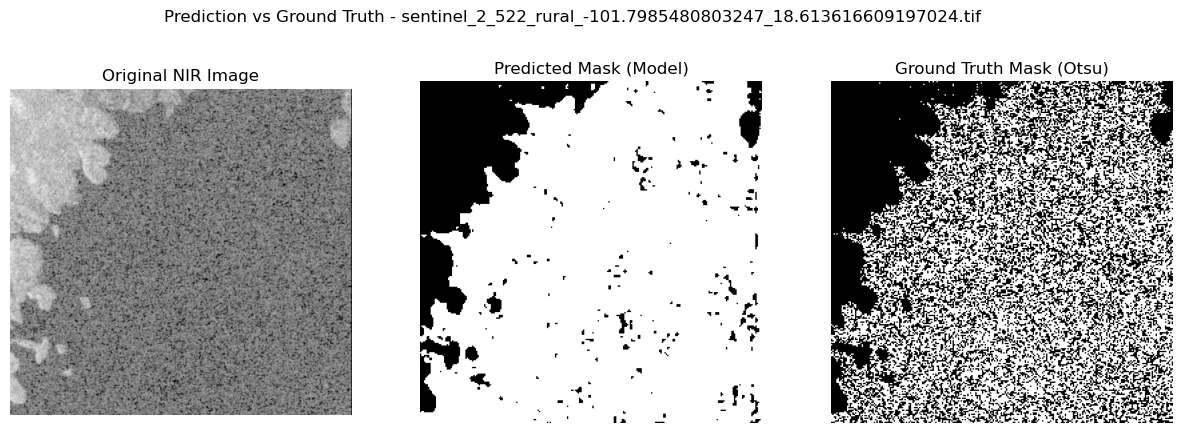
****

**Information From NIR Image**

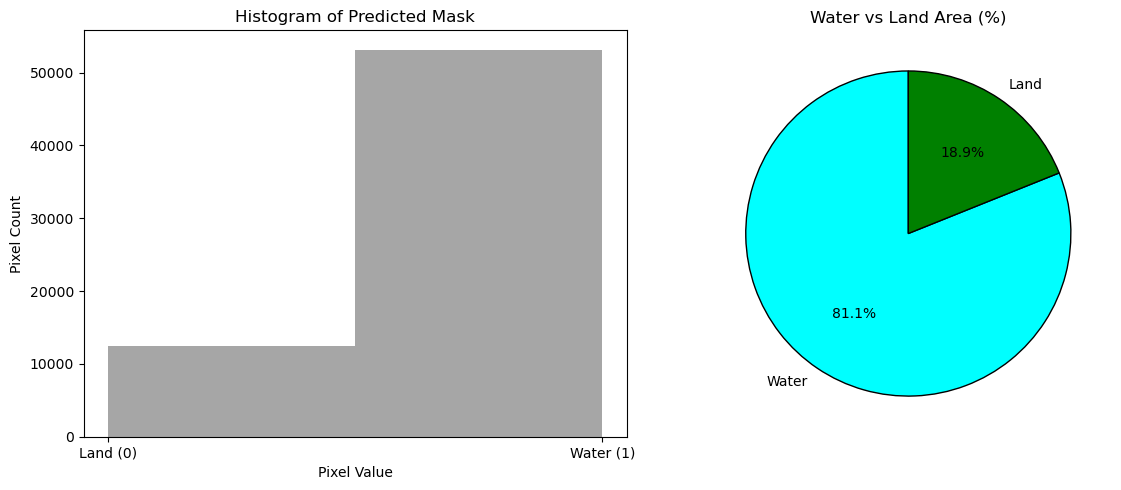
****

**Test 05**

**Mask Prediction**

****

**Information From NIR Image**

****

**e) Saving Segmented Images**

The final **segmented** **masks**, along with **computed** **area** **percentages**, are **saved** for **each** **image**. This output can be used for **downstream** **applications** such as **environmental** **monitoring** **reports**, **GIS** **analysis**, and **real**-**time** **geographic** **intelligence** **systems**.

*4. Challenges Faced:*

|  |  |
| --- | --- |
| Challenge | Explanation |
| Limited Availability of NIR Datasets | Publicly available **NIR** **datasets** specific to **water** **body** **segmentation** are scarce, **requiring** **extensive** **searching** and **filtering** to find suitable imagery. |
| Dataset Sourcing & Band Selection | The final dataset had to be **carefully** **sourced** from **Google** **Cloud**, ensuring **NIR** **bands** were present and usable for **segmentation** **tasks**. |
| Environment & Library Compatibility | Setting up the **development** **environment** with **TensorFlow**, **OpenCV**, **scikit**-**image**, and **tifffile** caused **compatibility** **issues** across different systems. |
| Dependency Troubleshooting | Resolving **library** **dependencies** and configurations consumed **significant** **time**, requiring **manual** **interventions** and **troubleshooting.** |
| Corrupted Image Preprocessing | Several images became **corrupted** (fully **black**) during **preprocessing** due to **format** **inconsistencies**, necessitating **dataset** **reloading** and careful file handling. |
| Data Splitting Complexity | The dataset was **originally** **unstructured** (single folder), requiring **custom** **scripts** to automate the split into **Train (70%),** **Validation (15%),** and **Test** **(15%).** |
| Computational Limitations | Processing **high-resolution NIR images** and training **segmentation models** demanded **substantial** **computational** **resources**, leading to **memory** **optimization** challenges. |

*5. Conclusion*

This project presents an **automated** **system** for **segmenting** and **quantifying** **water** and **land** **areas** from **NIR** **satellite** **images** using a **CNN**-**based** **approach**. By harnessing the **physics** of **NIR** **reflectance** and **dynamic** **mask** **generation**, the model achieves **accurate**, **scalable**, and **efficient** **segmentation**. The solution supports **real**-**time** **environmental** **monitoring**, **reduces** **manual** **efforts**, and lays the foundation for **intelligent** **geographic** **analysis** in **sustainable** **resource** **management**.