Automated Determination of Glacier Ablation Zones

A dissertation submitted in
partial fulfilment for the degree of

Master of Technology

in

Computer Science

by

Anupam Mondal
Roll no. - **CS2205**

under the supervision of

Dr. Sarbani Palit
Compute Vision and Pattern Recognition Unit (CVPRU)

Indian Statistical Institute, Kolkata
June, 2024

CERTIFICATE

This is to certify that the dissertation entitled **"Automated Determination of Glacier Ablation Zones"** submitted by **Anupam Mondal** to the Indian Statistical Institute, Kolkata, in partial fulfillment of the requirements for the degree of Master of Technology in Computer Science, is an authentic and genuine record of the research work carried out by the candidate under my supervision and guidance. I affirm that the dissertation has met all the necessary requirements in accordance with the regulations of this institute.

Sarbani Palit

Dr. Sarbani Palit
CVPR Unit
Indian Statistical Institute
Kolkata - 700108
India

Acknowledgement

I extend my sincere appreciation to Dr. Sarbani Palit, my advisor at the Computer Vision and Pattern Recognition Unit of the Indian Statistical Institute in Kolkata, for her guidance, continuous support, and inspiration. Her profound knowledge and creative suggestions have taught me a great deal in every subject and have shown me how to conduct solid research.

I would like to sincerely thank Ayoti Banerjee, Senior Research Fellow at the Indian Statistical Institute, for her invaluable assistance in gathering the datasets essential for this research. Her consistent provision of ideas and unwavering support have been instrumental to the success of this project.

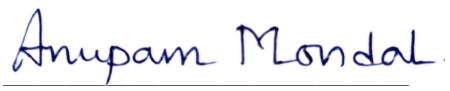
I am deeply grateful to all the teachers at the Indian Statistical Institute for their invaluable advice, insights, and instruction, which provided a crucial perspective to my research. Special thanks go to Subhranil Mustafi and all the other seniors at the SIPL (Signal and Image Processing Lab) for their constant mentoring.

Finally, I want to express my gratitude to my parents and extended family for their unwavering support. I also extend my sincere appreciation to all my friends for their continuous assistance and encouragement. I am thankful to everyone who has contributed to my growth and success, even if I have inadvertently missed mentioning them in the above list.

Declaration

I, **Anupam Mondal**, with Roll No. **CS2205**, hereby declare that the material presented in the dissertation titled **Automated Determination of Glacier Ablation Zones** represents original work carried out by me for the degree of **Master of Technology in Computer Science** at the **Indian Statistical Institute, Kolkata**.

Furthermore, I affirm that no sections of this report have been sourced or copied from external references without proper attribution. I am aware that any instances of plagiarism or the use of unacknowledged materials from third parties will be treated with the utmost seriousness and consequences.



Anupam Mondal
M.Tech (CS), CS2205
Indian Statistical Institute

Abstract

In recent decades, global temperature rises have significantly influenced glacier dynamics [1][2], underscoring the vital need for accurately delineating glacier boundaries to comprehend these shifts and document regional patterns. Despite this urgency, conventional methods struggle to map debris-covered glaciers (DCGs) due to their intricate nature. Climate change exacerbates glacier mass loss and intensifies glacier-related risks, necessitating ongoing monitoring and thorough analysis of climate-glacier interactions. Our research assesses the effectiveness of a convolutional neural network (CNN) in glacier mapping, utilizing Landsat satellite images, digital elevation models (DEMs), and DEM-derived land-surface parameters. Specifically, we seek to enhance the GlacierNet methodology by employing a CNN segmentation model to precisely identify regional DCG ablation zones. By training the models with satellite data from USGS and snow labeling from QGIS, and testing them on glaciers in the Karakoram region, we achieve improved estimations of the ablation zone, yielding high intersection over union scores. This study advances glacier mapping techniques, offering critical insights into climate change impacts on glacier dynamics in the Karakoram region. Furthermore, it marks a significant stride towards automating comprehensive glacier mapping, with potential applications in accurate glacier modeling and mass-balance analysis.

Keywords: Debris-covered glacier (DCG), Pixel wise classification, Digital elevation model (DEM), Landsat, Karakoram, convolutional neural network (CNN), NDSI

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Debris-covered glaciers (DCGs) are defined by a layer of rock and sediment debris enveloping their surfaces, displaying a spectrum of characteristics ranging from pristine ice in the accumulation zone to dense debris cover near the terminus. Nevertheless, the limited spectral variation in the mineralogical composition of supraglacial debris presents difficulties in distinguishing glaciers from their surrounding environment using traditional approaches, resulting in subjective and imprecise manual mapping of glaciers. However, subtle distinctions in debris spectral attributes influenced by factors such as topography, mineral composition, vegetation etc. can be utilized. Artificial intelligence (AI) techniques have the potential to utilize these nuances to enhance the identification of complex DCG boundary zones.

Debris-covered glaciers (DCGs) can be mapped using a variety of techniques that combine digital elevation models (DEMs) and data from satellite sensors. These techniques mainly involve pixel- and object-based analysis using spectral data [3], indices, and geomorphometric factors. Even though object-based approaches are generally preferred, a lot of methods use empirical thresholding of topographic or spectral data [4], which poses problems with morphological variation and anisotropic reflectance. Efficient thresholding becomes complicated due to the spatiotemporal variability of spectral and morphometric features.

In comparison to simple and traditional pattern-recognition algorithms, advanced AI machine learning approaches offer significant advantages, particularly in their ability to characterise visual elements and spatial notions, manage non-linear patterns, and generalise. Artificial neural networks (ANNs) such feed-forward convolutional neural networks (CNN) and deep neural networks (DNN) are used in deep-learning techniques, which are a subset of machine-learning techniques. In this work, we automated the mapping of debris-covered glaciers (DCGs) in the ablation zone using four different architectures, comparing the algorithmic performance of each to identify the best one.

Chapter 2

Related Work

GlacierNet [5] is a state-of-the-art segmentation model, modified from SegNet[6], featuring an axisymmetric architecture that incorporates the initial 13 convolutional layers from the VGG16 network. The model achieves a balanced structure with an equal number of encoders and decoders, mirroring SegNet's design strategy. The encoder component of GlacierNet[5] compresses the input data and extracts significant features using max-pooling layers, thereby reducing the data's dimensionality while highlighting key aspects. In contrast, the decoder component employs a reverse process known as max-unpooling to restore the data to its original size and accurately reposition features using pooling indices from the max-pooling layers.

Both the encoder and decoder are built around convolutional blocks, each consisting of three sequential convolutional layers with 32 kernels of 5x5 size[5]. These layers collaborate to refine the input data through multiple stages of convolution, enhancing feature extraction. GlacierNet's design philosophy aims to relax similarity constraints within the network, allowing it to more effectively identify and mark positive pixels. This approach improves the network's ability to recognize and segment intricate features, which is essential for accurately delineating glacier boundaries.

The sophisticated structure of GlacierNet[5] enhances the precision of glacier mapping, enabled more detailed analysis and understanding of glacier dynamics and their response to environmental changes. Its innovative balance of encoding and decoding processes allowed it to effectively handle the complex nature of debris-covered glaciers (DCGs) and other challenging terrains. GlacierNet thus represented a significant advancement in the field of glaciology and climate change research, providing a valuable tool for researchers and practitioners alike.

Chapter 3

Study Area

We assessed glaciers in Pakistan's central Karakoram, an area of High-Mountain Asia (HMA) distinguished by notable glacial characteristic diversity.[7] The surface and geometric characteristics of the glaciers in the Karakoram vary greatly, showing variances in size, status (advancing, retreating, and surge-type), and rate of change.[8] [9] Special occurrences like the Karakoram anomaly point to a generally marginally positive mass balance in the last few decades. Extreme relief and a greater prevalence of glaciers covered in debris—something that distinguishes this location from many others—are other features. In the Karakoram, westerly-driven systems dominate the patterns of precipitation in the winter and spring, with additional, albeit smaller, monsoon precipitation in the summer. Because of these different circumstances, the Karakoram is a perfect place to test our CNN method of mapping glaciers. The area under investigation encompassed roughly $160 \times 160 = 25600$ square km.

Figure 3.1: Marked box is our study area (Central Karakoram)

Chapter 4

Dataset

Landsat 9, the latest in the Landsat series of satellites, provides critical earth observation data that supports a wide range of scientific and practical applications.[10] Landsat 9 carries two instruments: the Operational Land Imager (OLI) and the Thermal Infrared Sensor (TIRS). These instruments capture high-resolution imagery across various spectral bands, which are essential for monitoring land use, land cover changes, and environmental conditions.

The eight bands of Landsat-9 used in this study include:

- **Band 1 (Coastal/Aerosol)**: Useful for coastal and aerosol studies.
- **Band 2 (Blue)**: Effective for water body mapping and distinguishing soil from vegetation.
- **Band 3 (Green)**: Helpful for assessing plant vigor.
- **Band 4 (Red)**: Important for discriminating vegetation types and soil boundaries.
- **Band 5 (Near Infrared - NIR)**: Critical for biomass studies and water body delineation.
- **Band 6 (Shortwave Infrared 1 - SWIR 1)**: Useful for differentiating moisture content in soil and vegetation.
- **Band 7 (Shortwave Infrared 2 - SWIR 2)**: Helpful in detecting fires and studying soil and geological composition.
- **Band 10 (Thermal Infrared - TIRS 1)**: Used for surface temperature mapping and analysis.

In the context of Landsat data, the wavelength of a band refers to the specific range of electromagnetic radiation (light) that the sensor on the Landsat satellite is designed to detect. Each band captures data within a specific portion of the electromagnetic spectrum, which includes visible light, near-infrared (NIR), short-wave infrared (SWIR), and thermal infrared (TIR) wavelengths. The wavelength determines the type of information that the band can capture and what features or phenomena it can observe.

- **Visible Light Bands (Blue, Green, Red)**
  - **Wavelength Range:** 0.4 to 0.7 µm
  - **Uses:** Natural color imagery, studying vegetation, water bodies, soil, urban areas.
- **Near-Infrared (NIR)**
  - **Wavelength Range:** 0.7 to 1.0 µm
  - **Uses:** Assessing plant health, biomass, soil moisture, delineating water bodies.
- **Shortwave Infrared (SWIR)**
  - **Wavelength Range:** 1.0 to 3.0 µm
  - **Uses:** Monitoring soil and vegetation moisture, identifying geological formations, detecting fire scars.
- **Thermal Infrared (TIR)**
  - **Wavelength Range:** 10 to 12 µm
  - **Uses:** Measuring surface temperatures, studying urban heat islands, soil moisture, volcanic activity.
- **Specific Bands and Their Applications**
  - **Coastal/Aerosol Band (Blue, 0.433-0.453 µm):** Used for coastal water studies and aerosol detection.
  - **Cirrus Band (SWIR, 1.360-1.390 µm):** Designed to detect cirrus clouds, aiding in cloud masking.

For the training and testing of the four different architectures, we utilized ten data layers as inputs as in Figure: 4.4. These included the eight aforementioned bands from Landsat-9, a Digital Elevation Model (DEM), and one geomorphometric parameter derived from the DEM. The Landsat-9 bands were sourced from the USGS Earth Explorer website, specifically from the Landsat Collection 2 Level-2 dataset (Landsat 8-9 OLI/TIRS CS L2).

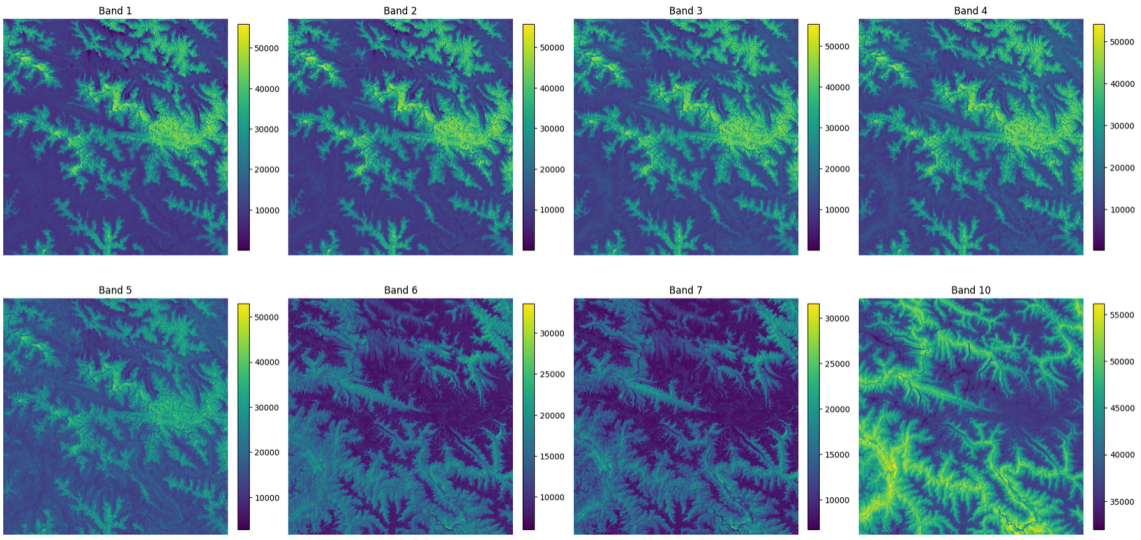


Figure 4.1: Input Bands (Band 1-7 and Band 8)

The DEM data was obtained from the Shuttle Radar Topography Mission (SRTM) 1-Arc Second Global dataset. Additionally, the slope parameter was derived from the DEM using QGIS software. Each of these layers had a resolution of 30 meters per pixel.

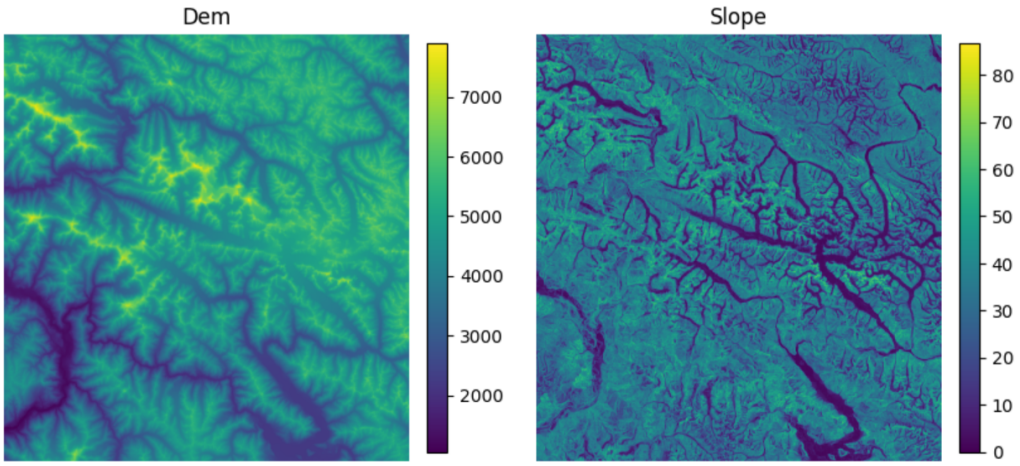


Figure 4.2: Dem and Slope of the corresponding study area

Each of the ten layers was selected for their complementary advantages, with none considered indispensable. The Landsat-9 bands provide a wide range of spectral information, crucial for various types of land cover and land use analysis. The DEM offers vital topographic information, while the geomorphometric parameter adds further granularity to the terrain analysis. To ensure consistency in magnitude, data values were normalized based on their minimum and maximum values. This normalization process helps in maintaining uniformity across different data layers, which is essential for accurate model training and testing.

Cloud cover, particularly in high-altitude glacierized basins, presents significant challenges in optical remote sensing image analysis. Cloud cover can obscure critical features of the landscape, leading to incomplete or inaccurate data. Therefore, it was essential to select images with minimal cloud cover, specifically less than 5%. The training of the models utilized images acquired on June 23, 2023 (LC09L2SP14903520230623), while images from different rows and paths were employed to assess model accuracy. Using images from various paths and rows ensures that the model is robust and can generalize well to different spatial contexts.

To obtain the desired output, the Normalized Difference Snow Index (NDSI)[11] was computed using Band-3 and Band-6 of Landsat-9. The NDSI is calculated using the following formula:

$$\text{NDSI} = \frac{\text{Band 3} - \text{Band 6}}{\text{Band 3} + \text{Band 6}}$$

This index helps in distinguishing snow from non-snow surfaces. Snow was then identified by applying a threshold value exceeding 0.4, resulting in a binary snow-labeled image. The thresholding process is crucial as it converts the continuous NDSI values into a binary classification, where pixels with NDSI values greater than 0.4 are labeled as snow, which has been shown in Figure:4.3

Band 3 (green) and Band 6 (shortwave infrared, SWIR) are particularly useful for NDSI calculation because they leverage the distinct reflectance properties of snow. Snow has high reflectance in the green band and low reflectance in the SWIR band, creating a strong contrast that enhances snow detection. The green band helps distinguish snow from vegetation, which has lower reflectance in this band, while the SWIR band aids in differentiating snow from clouds and liquid water, both of which have different reflectance characteristics in this spectrum. This combination allows for accurate identification and mapping of snow cover, making the NDSI a reliable tool in remote sensing.

Figure 4.3: Snow label image in the month of June, 2023

Consequently, the ground truth was effectively generated, approximating the ablation zone through these outlined procedures. The ablation zone, which is the area where snow and ice melt faster than they accumulate, is critical for understanding glacier dynamics. Due to satellite errors, some areas may not be covered, resulting in no-data values. These no-data values were filled using QGIS with k-nearest interpolation. This interpolation method estimates the missing values based on the values of the nearest neighbors, ensuring a complete and continuous dataset for analysis.

Using Landsat-9 instead of Landsat-8 for determining glacier ablation zones offers several advantages. Firstly, Landsat-9 provides improved data continuity and reliability, ensuring a seamless transition from Landsat-8 data. This continuity is crucial for long-term monitoring and analysis of glacier dynamics. Additionally, Landsat-9 carries upgraded sensors, such as the Operational Land Imager (OLI), which offer enhanced spectral and radiometric capabilities compared to Landsat-8. These improved sensors enable more accurate detection and characterization of

glacier features, including ablation zones.

Furthermore, Landsat-9's thermal infrared sensor (TIRS) provides valuable thermal data that can aid in assessing surface temperatures and thermal gradients, which are essential for identifying and delineating ablation zones. The combination of multi-spectral and thermal data from Landsat-9 allows for a comprehensive analysis of glacier ablation processes, including the melting of snow and ice.

Moreover, Landsat-9's extended lifespan and continued operation ensure a consistent supply of high-quality satellite imagery for ongoing monitoring of glacier ablation zones. This long-term dataset is invaluable for studying trends and changes in glacier dynamics over time, providing valuable insights into the impacts of climate change on glacier systems.[2]

In summary, the use of Landsat-9 offers improved data quality, enhanced spectral and thermal capabilities, and long-term continuity, making it a preferred choice for determining glacier ablation zones compared to Landsat-8.

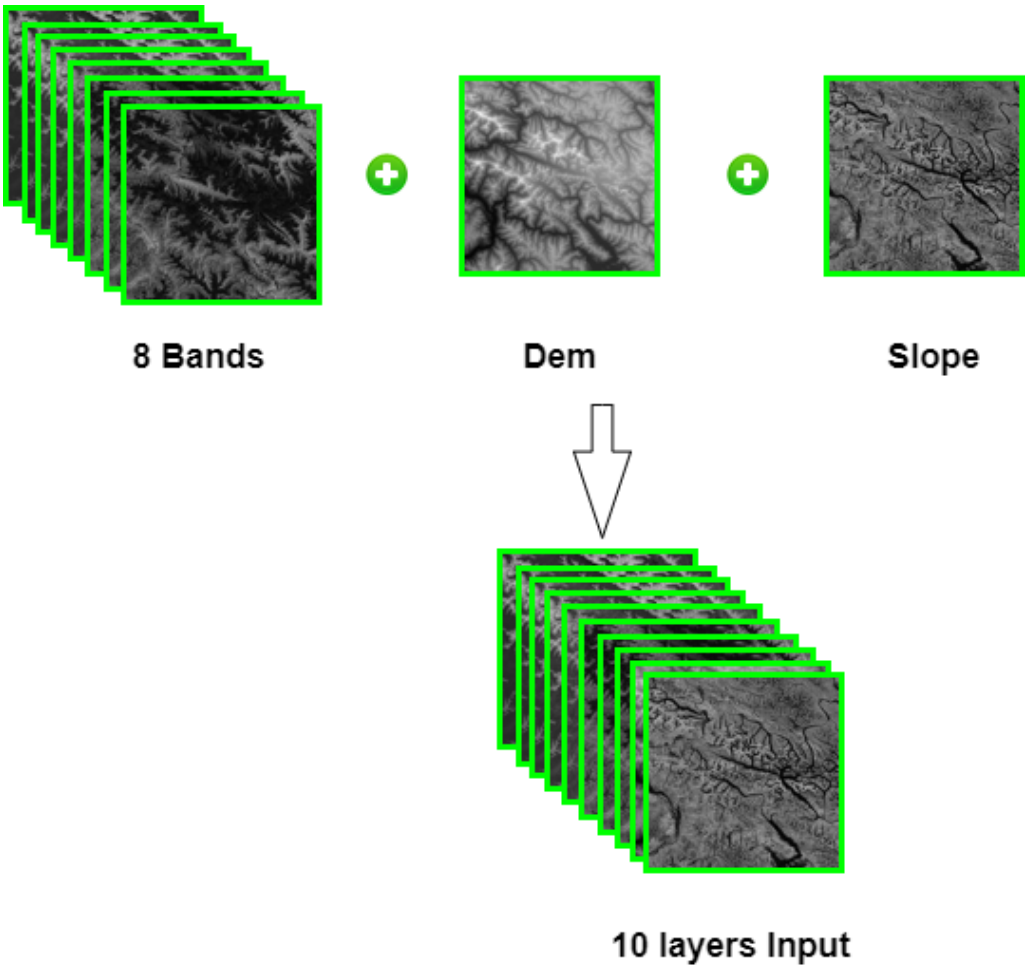


Figure 4.4: 10 layer input image of shape (5472, 5203, 10)

Chapter 5

Methodology

5.1 Data Preprocessing

To prepare our dataset for training and testing the CNN architectures, we followed several key steps:

1. **Normalization of Input Data Layers:**

- We normalized all ten input data layers, which included eight Landsat-9 bands (bands 1 to 7 and band 10), a Digital Elevation Model (DEM), and a geomorphometric parameter derived from the DEM, as shown in Figure: 4.4.

- Normalization was performed using the formula:

$$\text{Normalized Data} = \frac{\text{Original Pixel Value} - \text{Minimum Pixel Value}}{\text{Maximum Pixel Value} - \text{Minimum Pixel Value}}$$

  Normalization ensures that all input features are on a similar scale, which accelerates the training process, improves convergence, and enhances overall model performance. It prevents features with larger ranges from dominating the learning process, leading to better-balanced training.

2. **Concatenation of Data Layers:**

- Once normalized, the ten layers were concatenated to create a three-dimensional image.

  This step integrates spectral, topographic, and geomorphometric information into a single dataset, providing a comprehensive view of the data. This allows the CNN model to consider various aspects of the terrain and land cover, improving its ability to learn relevant features.

3. **Processing Snow Label Image:**

- The snow label image initially contained both ablation and accumulation zones. We removed accumulation zones by applying a threshold to the DEM, excluding areas with pixel values greater than 5000.

  Focusing solely on ablation zones ensures that the model is trained specifically to identify areas of interest. This is crucial for understanding glacier melt dynamics.

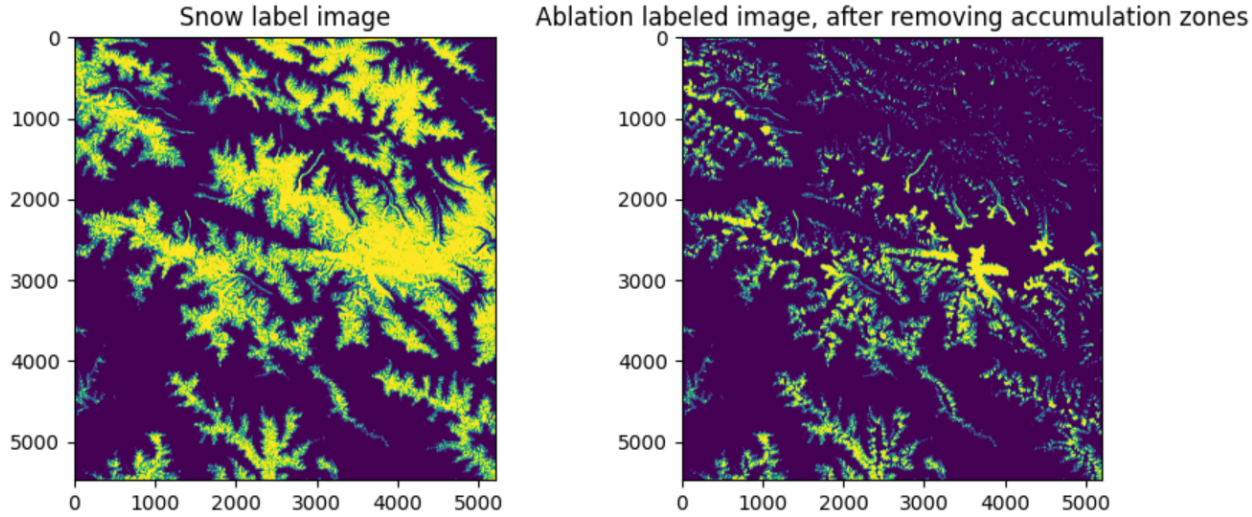


Figure 5.1: Accumulation zone removing (Before and After image)

4. **Connected Component Analysis:**

- We used connected component analysis to remove small components with fewer than 1000 pixels.

  Removing small, insignificant areas reduces noise in the dataset, allowing the model to focus on more relevant features. This step helps in improving the accuracy and reliability of the model predictions.

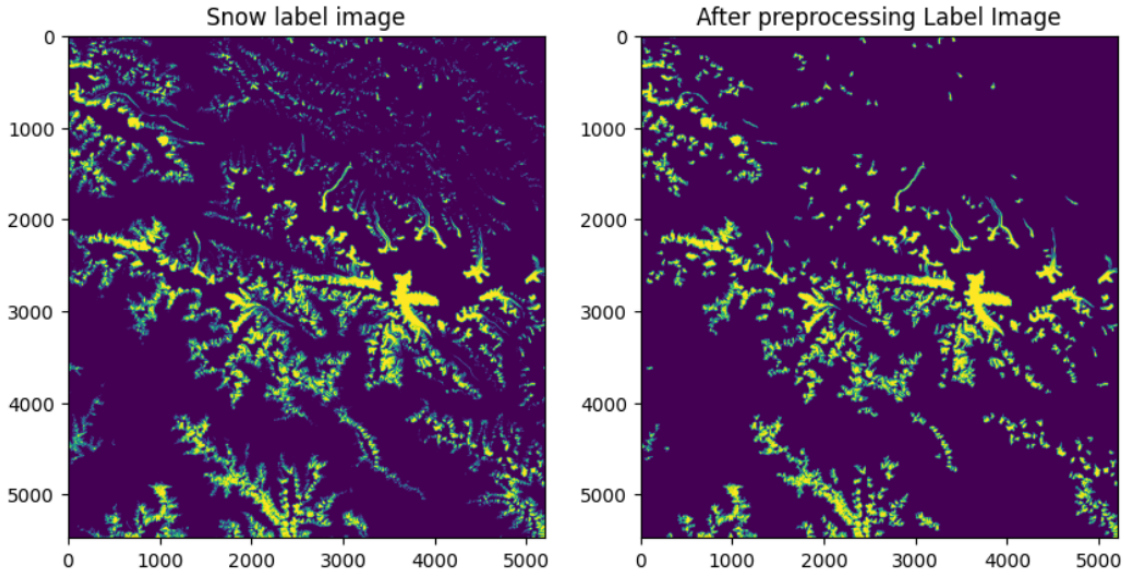


Figure 5.2: Removing small connected components

5. **Filling Holes Using K-Nearest Neighbor Interpolation[12]:**

- We filled any holes in the label image using k-nearest neighbor interpolation.

  Filling gaps in the data ensures a complete and continuous dataset, which is necessary to prevent inaccuracies during model training and predictions. The k-nearest neighbor method estimates missing values based on the nearest neighbors, maintaining data integrity.

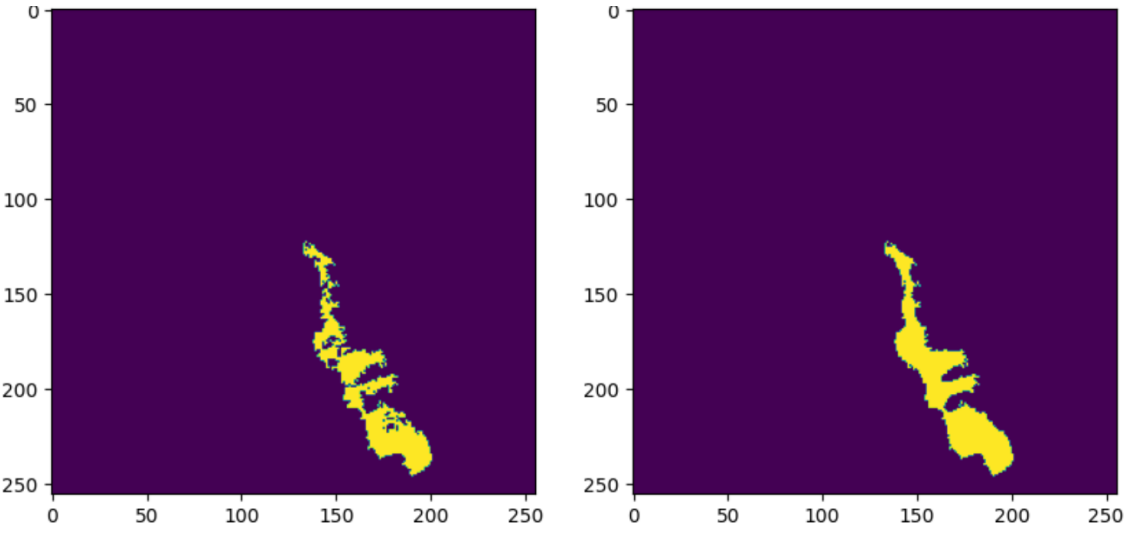


Figure 5.3: Label image before and after hole filling

6. **One-Hot Encoding:**

- We performed one-hot encoding on the label data, converting categorical labels into a binary matrix representation.

  This transformation is essential for training CNN models, enabling them to effectively learn from categorical data. It allows the model to distinguish between different classes of pixels, such as snow and non-snow areas.

  Here is an example of one hot encoding:

$$\text{Input Matrix} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\text{Class 0 Matrix} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix}$$

$$\text{Class 1 Matrix} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

7. **Subsampling the Image:**

- Given the extensive swath of the original multi-channel image, we subsampled the image using a sliding window approach with a pixel size of 256 $\times$ 256 and a shifting stride of 128 pixels.[5]

  Subsampling makes the data more manageable and reduces computational load, facilitating efficient training and validation of the CNN models. It ensures that the window size is sufficiently large to maintain performance while determining the size of the intermediate feature map generated by the convolutional layers.

8. **Selection of Subsampled Images:**

- We selected subsampled images containing at least 10% non-snow (label 0) and 14% snow (label 1), ensuring a balanced dataset.

  This selection criterion ensured that the dataset had sufficient representation of both classes, which is crucial for balanced training. We divided these subsampled images into training and validation sets, preparing the data for effective model training and validation.

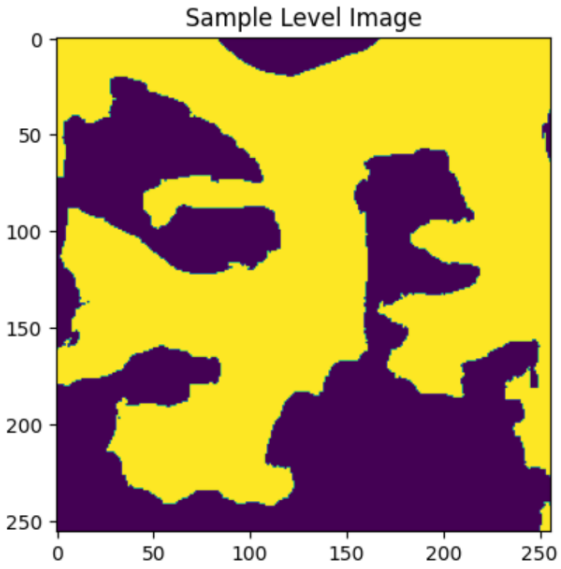


Figure 5.4: One subsampled label image of shape (256 ×256)

5.2 CNN Models

In this section, we explore four different Convolutional Neural Network (CNN) models to compare their performance in detecting glacier ablation zones. The models used are SegNet[6], U-Net[13], Residual U-Net[14], and Attention U-Net[15]. Each model's architecture and its suitability for our task are described in detail below.

5.2.1 SegNet

SegNet is a deep learning model specifically designed for semantic segmentation. It features an encoder network, a matching decoder network, and concludes with a pixel-wise classification layer using Softmax. The encoder network is typically a 13-layer VGG16 architecture, which downsamples the input image to extract features at different levels. The decoder network upsamples these features to reconstruct the original resolution, providing a dense pixel-wise classification map.

The key feature of SegNet is its ability to retain spatial information during the upsampling process by using pooling indices from the encoder network. This helps in accurately delineating the boundaries of glacier ablation zones. SegNet's structured architecture and ability to manage spatial hierarchies make it suitable for our glacier ablation zone detection task, where precise boundary detection is critical. (See Figure: 5.5)

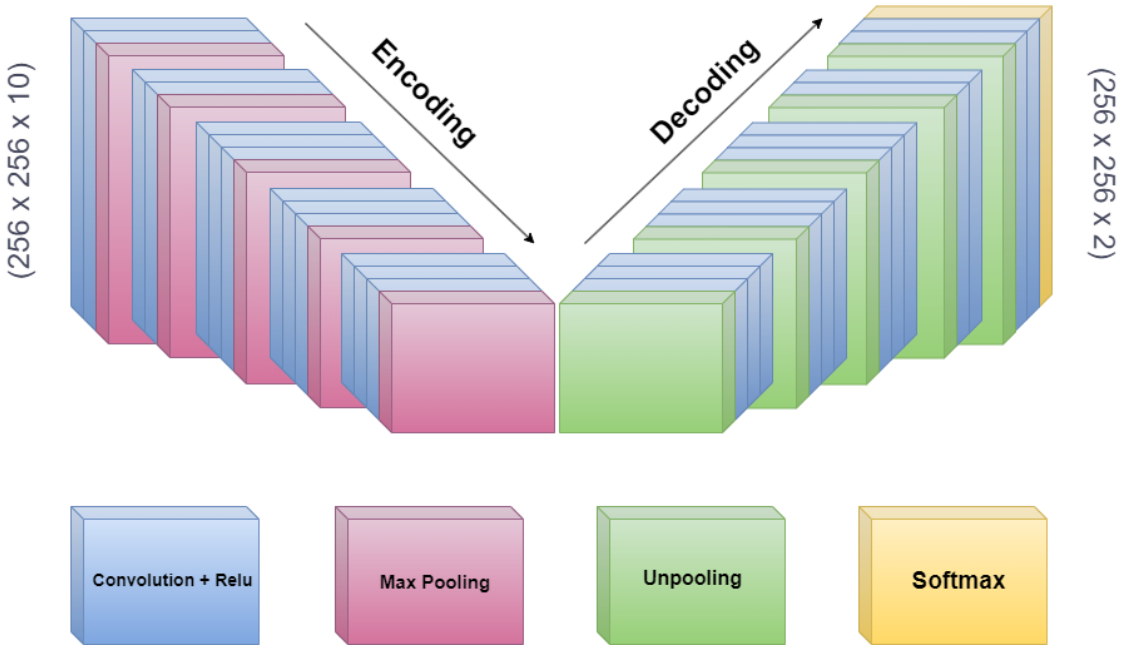


Figure 5.5: Segnet Architecture

5.2.2 U-Net

U-Net is another popular CNN architecture used for image segmentation. It comprises a contracting path (encoder) and an expansive path (decoder), creating a U-shaped architecture. The encoder captures context by progressively downsampling the input image, while the decoder reconstructs the image by upsampling and combining feature maps from corresponding encoder layers.

The main difference between U-Net and SegNet is the skip connections that directly connect corresponding layers in the encoder and decoder paths. These connections help retain fine-grained details, which are crucial for segmenting glacier ablation zones accurately. U-Net's ability to leverage these skip connections improves its performance in scenarios requiring precise segmentation, such as glacier ablation zone detection.(See Figure: 5.6)

5.2.3 Residual U-Net

Residual U-Net is an enhanced version of the U-Net architecture that incorporates residual connections within the U-Net framework. Residual connections, introduced by ResNet, help in training deeper networks by allowing the gradient to bypass certain layers, thus mitigating the vanishing gradient problem.

In Residual U-Net, each convolutional block in the encoder and decoder paths includes a residual connection that adds the input of the block to its output. This

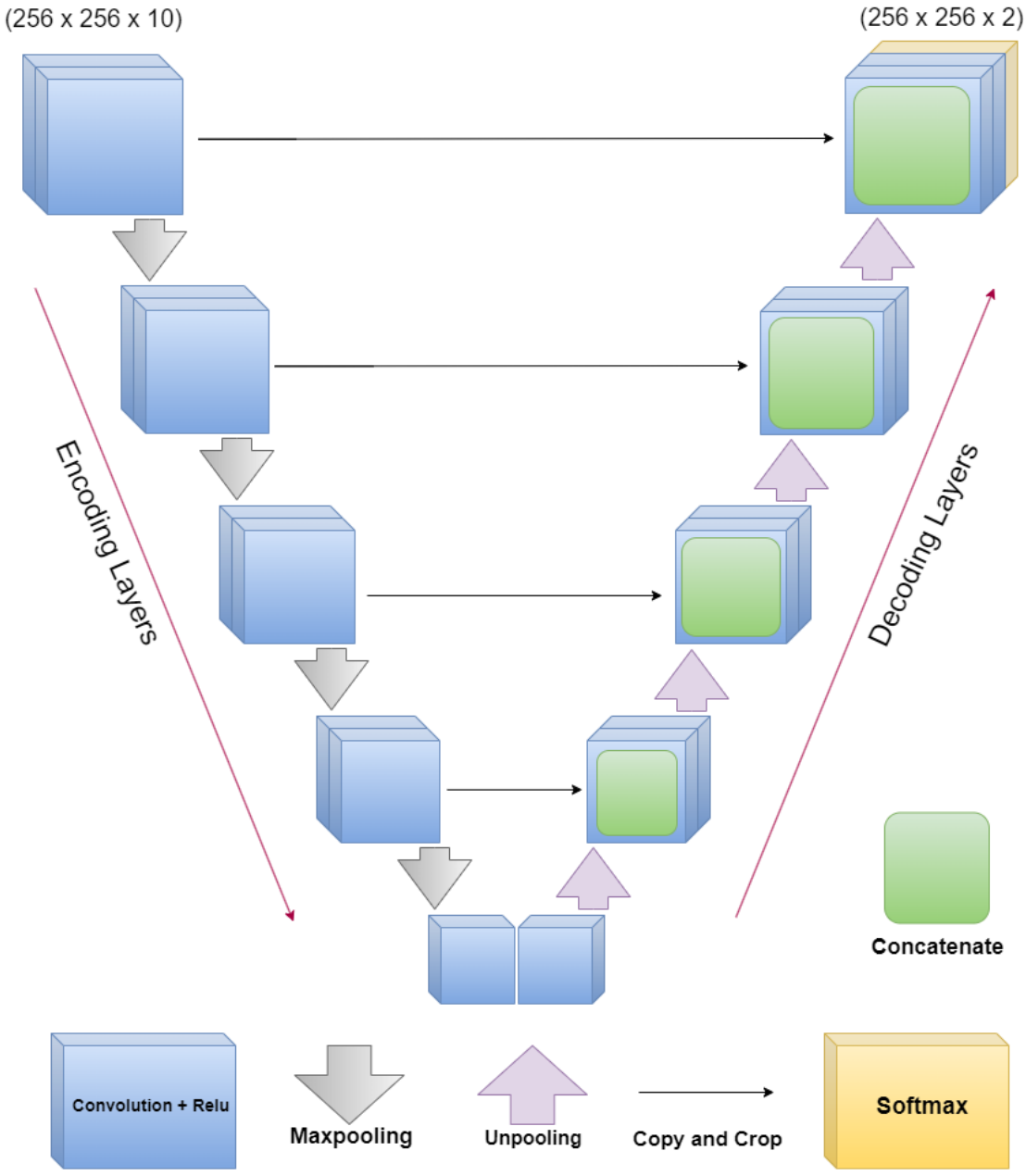


Figure 5.6: U-Net Architecture

modification enables the model to learn more complex features without degradation in performance. For our task, Residual U-Net's improved learning capability is expected to enhance the detection of glacier ablation zones, particularly in complex and varied terrains. (See Figure: 5.7)

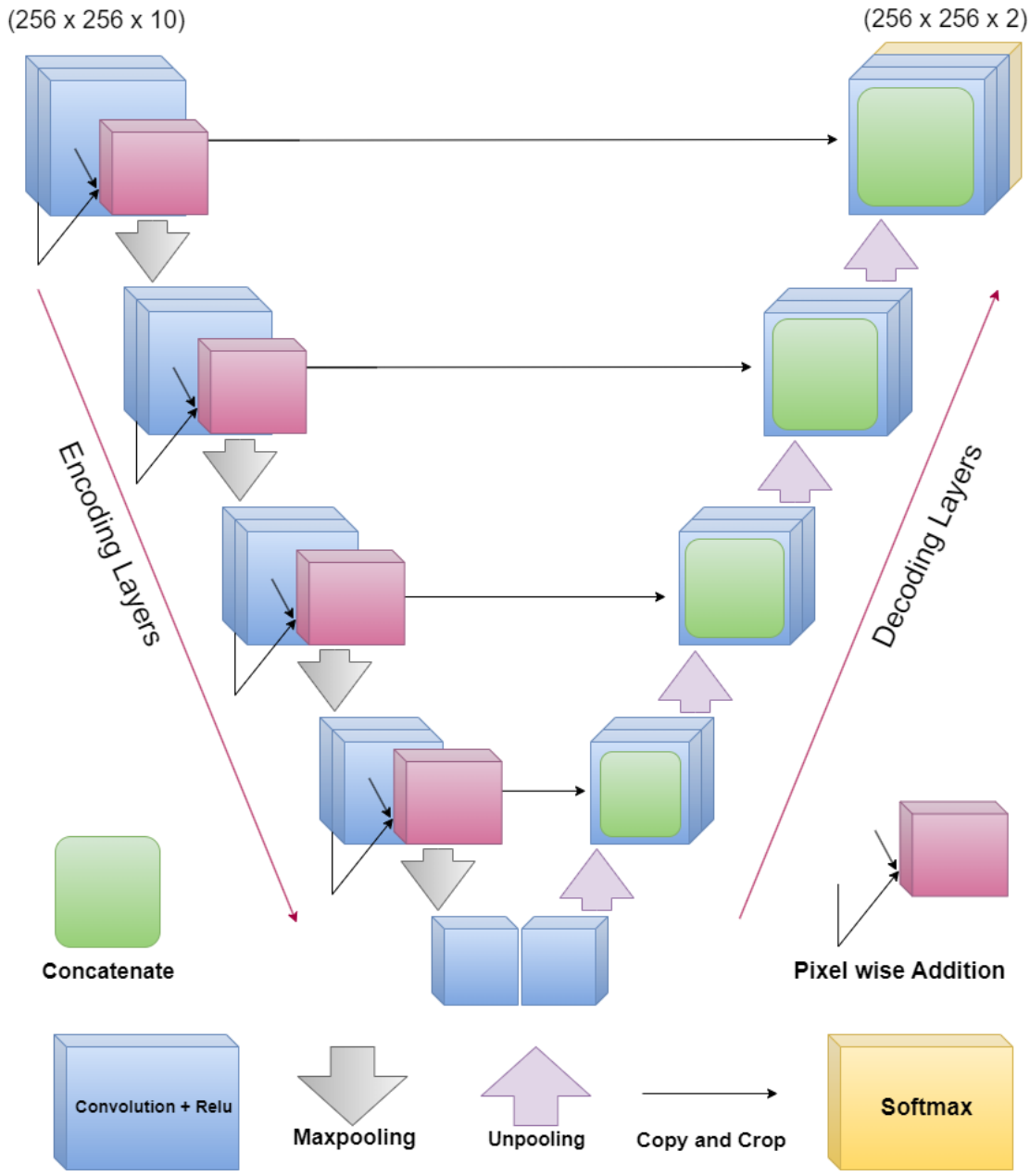


Figure 5.7: Residual U-Net Architecture

5.2.4 Attention U-Net

Attention U-Net extends the U-Net architecture by integrating attention mechanisms. Attention mechanisms allow the network to focus on relevant parts of the input image while suppressing irrelevant regions. In Attention U-Net, attention gates are added to the skip connections between the encoder and decoder paths.

These attention gates dynamically highlight salient features and suppress background noise, leading to more precise segmentation outputs. For glacier ablation zone detection, the attention mechanism helps in accurately identifying and delineating ablation zones by focusing on the most relevant features in the input data. This model is particularly useful in scenarios where the target regions are small or exhibit significant variability.

By employing these four CNN models, we aim to identify which architecture best suits our glacier ablation zone detection task. Each model brings unique advantages, from retaining spatial information in SegNet to leveraging attention mechanisms in Attention U-Net, thus providing a comprehensive comparison of their capabilities. (See Figure: 5.8)

5.2.5 Internal Components of CNN Models

To understand the internal workings of the CNN models, we detail the following components and techniques used in our implementations:

1. **Loss Function:** The loss function measures the difference between the predicted output and the actual target values. For our segmentation task, we used binary cross-entropy loss [16], which is well-suited for multi-class classification problems. The formula for binary cross-entropy loss is:

$$L = -\frac{1}{N}\sum_{i=1}^{N}\left[y_i \log(p_i) + (1 - y_i)\log(1 - p_i)\right]$$

where y_i is the actual binary class label for observation i (0 or 1), and p_i is the predicted probability that observation i belongs to the positive class.

2. **Optimizer:** We used the Adam [17] optimizer, which combines the advantages of both Adaptive Gradient Algorithm (AdaGrad) and Root Mean Square Propagation (RMSProp). Adam is computationally efficient, requires little memory, and is well-suited for problems with large datasets and high-dimensional parameter

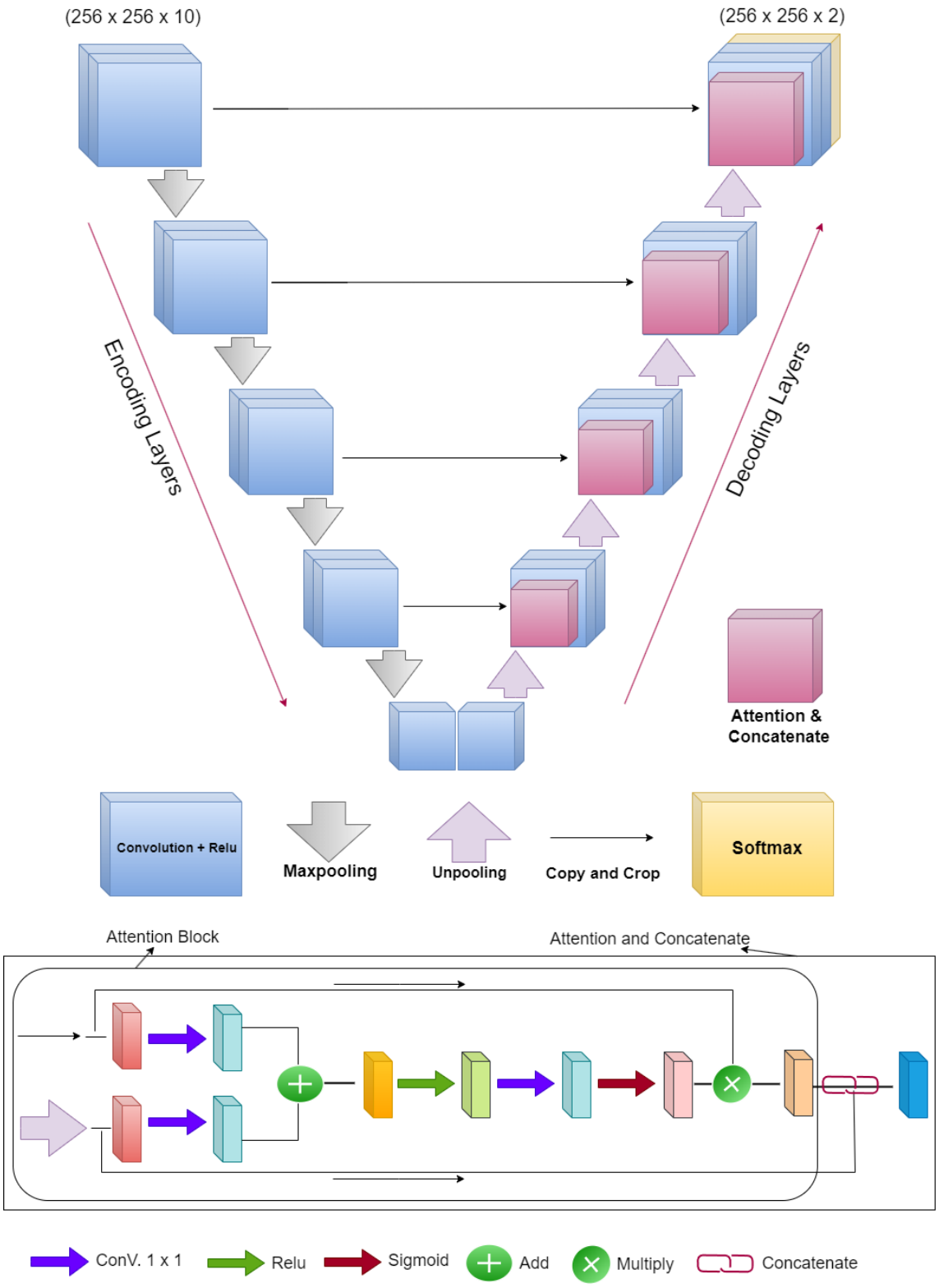


Figure 5.8: Attention U-Net Architecture

spaces.

3. **Learning Rate:** The learning rate used was $\alpha = 0.0001$. This parameter controls the size of the steps the optimizer takes to reach the minimum of the loss function. A smaller learning rate ensures more precise convergence but requires more epochs to train.

4. **Convolutional Layers (Conv2D):** Convolutional layers are the core building blocks of a CNN. They apply a convolution operation to the input, passing the result to the next layer. We used 2D convolutions with padding set to 'same' to ensure that the output dimensions match the input dimensions, preserving the spatial resolution.

Here's an example of a convolution operation using a 3×3 kernel and a stride of 1 with 'same' padding:

$$\text{Input Matrix} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{bmatrix} \quad \text{Kernel} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

$$\text{Output Matrix} = \begin{bmatrix} 7 & 14 & 17 & 11 \\ 17 & 30 & 35 & 22 \\ 29 & 50 & 55 & 34 \\ 23 & 34 & 37 & 27 \end{bmatrix}$$

5. **Activation Function:** We used the ReLU (Rectified Linear Unit) [18] activation function, defined as:

$$\text{ReLU}(x) = \max(0, x)$$

ReLU introduces non-linearity to the model and helps in addressing the vanishing gradient problem, leading to faster and more effective training.

6. **MaxPooling:** MaxPooling layers are used to reduce the spatial dimensions of the input, decreasing the computational load and helping the model to become invariant to small translations. The operation selects the maximum value from each patch of the feature map.

$$\text{Input Matrix} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 14 & 13 & 15 & 16 \end{bmatrix}$$

$$\text{MaxPooling with Stride 2} = \begin{bmatrix} \max(1,2,5,6) & \max(3,4,7,8) \\ \max(9,10,14,13) & \max(11,12,15,16) \end{bmatrix} = \begin{bmatrix} 6 & 8 \\ 14 & 16 \end{bmatrix}$$

7. **Unpooling:** Unpooling layers are used in the decoder part of architectures like SegNet to upsample the feature maps. They reverse the max pooling operation by placing the pooled values back to their original positions, helping to reconstruct the spatial resolution of the input image.

$$\text{Input Matrix} = \begin{bmatrix} 6 & 8 \\ 14 & 16 \end{bmatrix}$$

$$\text{Max Unpooling / Upsampling} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 6 & 0 & 8 \\ 0 & 0 & 0 & 0 \\ 14 & 0 & 0 & 16 \end{bmatrix}$$

By utilizing these components, we ensured that our CNN models were optimized for the task of glacier ablation zone detection. Each element, from the choice of activation function to the use of advanced optimizers like Adam, was selected to enhance the model's ability to learn complex features and make accurate predictions.

5.3 Post Processing

In this section, we describe the post-processing steps applied to the binary images predicted by the trained CNN models:

1. Merging Predicted Images:

After the CNN model training is done, we obtain small predicted images by predicting many small size $(256 \times 256 \times 10)$ subsampled images. These small predictions are merged to reconstruct the output for the full original image. [5]

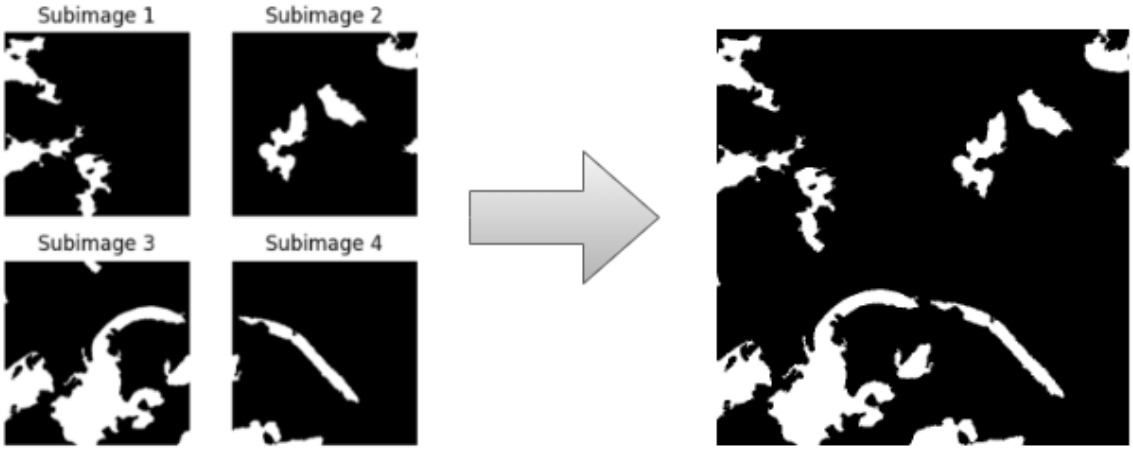


Figure 5.9: Concatenation of small binary images

2. Conversion to Binary Class Labels:

After the model outputs the probability of each pixel belonging to class 0 or class 1 using the Softmax function, we convert these probabilities into binary class labels. A pixel is assigned to class 1 if its probability is greater than 0.9, and to class 0 if its probability is less than 0.1. For pixels with probabilities between 0.1 and 0.9, we utilize the k-Nearest Neighbors (k-NN)[12] algorithm to determine their class based on the labeled pixels.

k-Nearest Neighbors (k-NN) Algorithm:

k-NN is a simple, non-parametric method used for classification. It assigns a class to an unlabeled pixel by considering the classes of its nearest neighbors in the feature space. The class that is most common among the neighbors is assigned to the pixel. This approach helps in making a more informed decision for the pixels with ambiguous probabilities, enhancing the overall accuracy of the classification.

3. Connected Component Analysis and Removing Small Components:

After converting the probabilistic outputs to binary labels and refining them with k-NN, we apply connected component analysis to the binary image. This step involves identifying and labeling all the connected regions (components) in the image. Each connected component represents a contiguous region of pixels that share the same label.

Additionally, to improve the quality of the final binary image, we remove small components that are likely to be noise or insignificant regions. Specifically, we eliminate components that have fewer than 1100 pixels. This threshold value was determined empirically to provide the best results for our application. By discarding these small components, we reduce noise and enhance the clarity of the detected glacier ablation zones.

These post-processing steps are crucial for refining the raw outputs of the CNN models, ensuring that the final binary images are accurate and meaningful for identifying glacier ablation zones.

5.4 A New Approach

The predefined models perform well in predicting ablation zones for the same timestamp data. However, when attempting to predict different timestamp data, particularly from the end of ablation season, the predictions are not accurate. These models are trained on data from June 2023, which marks the beginning of ablation, but we are trying to predict data from September 2023, which signifies the end of ablation (glacier melting).

To address this issue, we propose a new approach. Instead of directly predicting ablation zones, we predict the presence of snow.[19] By focusing on predicting snow, we can overcome the challenges associated with predicting ablation zones accurately. After obtaining the predicted snow areas, we apply a digital elevation model (DEM) threshold of 5000 on these predictions to identify the ablation zones. This approach yields better Intersection over Union (IoU) scores for predicting the data, even for different timestamp data.

The Normalized Difference Snow Index (NDSI) is an index related to the presence of snow in a pixel, which uses bands 3 and 6 to determine the snow. To determine snow using CNN models, we experimented with two approaches. First, using 10 layers (including bands 3 and 6), and secondly using 8 layers (excluding bands 3 and 6) for each of four CNN models: SegNet, U-Net, Residual U-Net, and Attention U-Net.

This new approach allows us to leverage the consistency of snow presence across different timestamps, making the prediction task more robust. The preprocessing and post-processing steps remain the same as described earlier, ensuring the reliability and consistency of the predictions.

By adopting this strategy, we enhance the accuracy of our predictions, particularly when dealing with data from different timestamps, ultimately improving the effectiveness of our models in identifying glacier ablation zones.

Chapter 6

Experiments and Results

In this section, we outline the experimental setup and report the results of our investigation, which focuses on assessing the performance of a Convolutional Neural Network (CNN) model for the identification of glacier ablation zones. We provide a detailed analysis of key performance metrics, including Intersection over Union (IoU), Accuracy, Precision, Recall, and F1 Score. Our evaluation includes a comparison with baseline models to gauge the efficacy of our approach.

Performance Metrics:

- **Intersection over Union (IoU)**: Measures the overlap between predicted and ground truth regions. It is calculated as the intersection of the predicted and ground truth regions divided by their union:

$$IoU = \frac{\text{Area of Intersection}}{\text{Area of Union}} = \frac{TP}{TP + FP + FN}$$

- **Accuracy**: Measures the overall correctness of predictions. It is calculated as the ratio of correctly predicted samples to the total number of samples:

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

- **Precision**: Measures the ratio of correctly predicted positive observations to the total predicted positives. It is calculated as:

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

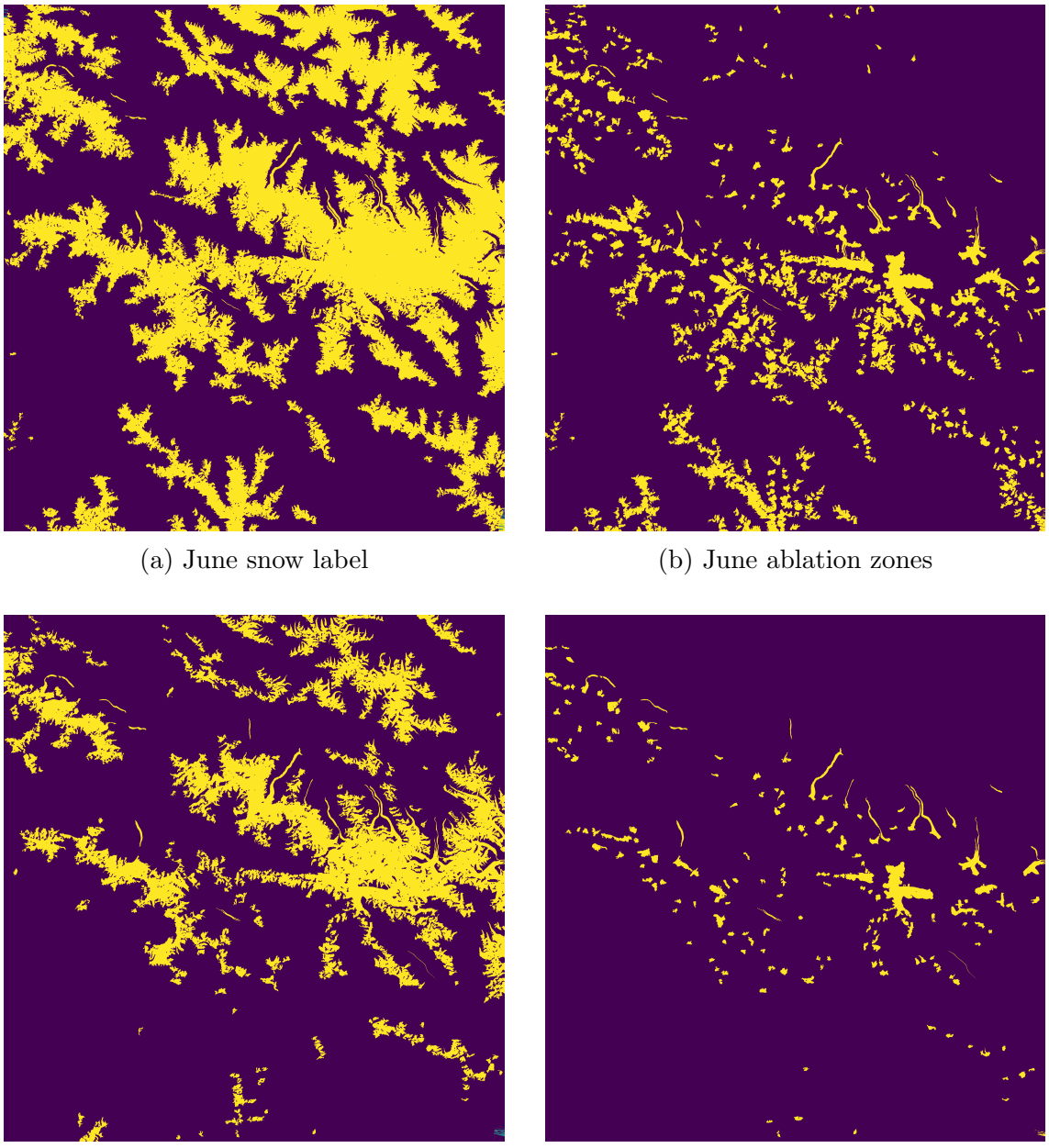
- **Recall (Sensitivity)**: Measures the ratio of correctly predicted positive observations to the total actual positives. It is calculated as:

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

- **F1 Score**: Harmonic mean of precision and recall, provides a balance between precision and recall. It is calculated as:

$$F1Score = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

Where, TP - True Positive, TN - True Negative, FP - False Positive, FN - False Negative



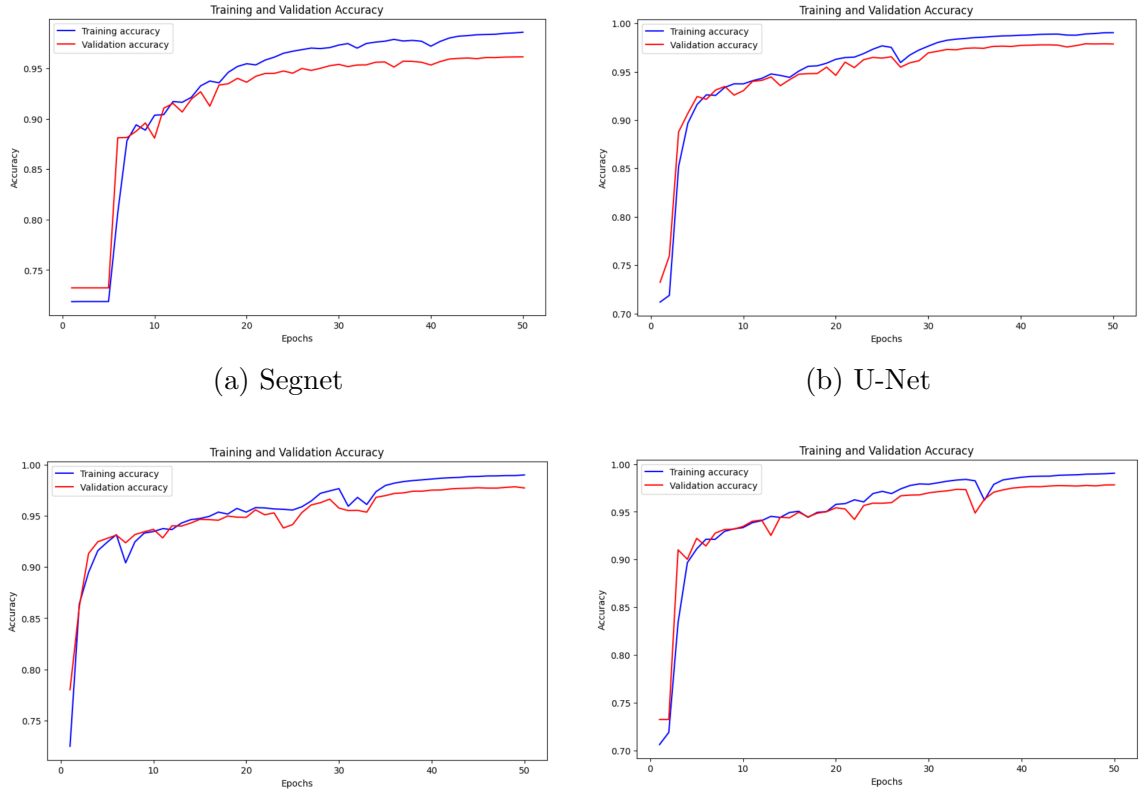
(a) June snow label

(b) June ablation zones

(c) September snow label

(d) September ablation zones

Figure 6.1: Ground Truth Labels of Snow and Ablation Zones for June and September



(a) Segnet

(b) U-Net

(c) Residual U-Net

(d) Attention U-Net

Figure 6.2: Train vs Validation accuracy for four models

6.1 Ablation Zone Prediction:

- Comparison of predicting ablation zones using various CNN models:

Model	IoU	Accuracy	Precision	Recall	F1 Score
SegNet	0.7609	0.9782	0.7993	0.9407	0.8642
U-Net	0.8269	0.9854	0.8660	0.9482	0.9053
Res U-Net	0.8385	0.9865	0.8769	0.9503	0.9122
Attention U-Net	0.8462	0.9873	0.8884	0.9468	0.9167

Table 6.1: Performance metrics of different models

- After the initial prediction by the CNN model, further processing steps were performed to enhance the prediction accuracy. Firstly, pixels with probabilities falling between 0.1 and 0.9 (Thresholds were chosen without any assumptions, leveraging the fact that for lower probabilities, it can be reasonably concluded that the pixel belongs to class 0, while for higher probabilities, it can be inferred that the pixel belongs to class 1. This approach was adopted to ensure a more

robust classification. The results obtained using this approach are presented in the table below.) were identified as 'holes' using K-nearest neighbors (KNN) algorithm. Subsequently, connected component analysis (CCA) was applied to remove small connected pixels or noises for various thresholds. Performance metrics were then calculated after each step of processing to evaluate the improvement in prediction accuracy.

Steps	IoU	Accuracy	Precision	Recall	F1 Score
1.	0.8096	0.9833	0.8337	0.9655	0.8948
2.	0.8162	0.9841	0.8462	0.9584	0.8988
3A.	0.8445	0.9870	0.8760	0.9592	0.9157
3B.	0.8462	0.9873	0.8884	0.9468	0.9167

Table 6.2: Comparison of various post processing steps for Attention U-Net

- Step 1. represents the performance metrics after CNN prediction.
- Step 2. represents the performance metrics after post-processing steps (hole filling using kNN).
- Step 3A. and Step 3B. represent the performance metrics after removing small connected components using thresholds 1000 and 1100, respectively.

6.2 Snow Prediction:

In this study, we first predict snow cover using CNN models. Subsequently, we determine the ablation zones by applying a DEM threshold to the predicted snow cover.

- This experiment involves predicting snow cover without performing any preprocessing or postprocessing on the snow labels.

Configuration	Test	IoU	Accuracy	Precision	Recall	F1 Score
Without B3, B6	Same timestamp	0.8453	0.9446	0.9251	0.9074	0.9162
	Different timestamp	0.7090	0.9420	0.8740	0.7897	0.8297
With B3, B6	Same timestamp	0.8461	0.9436	0.9037	0.9299	0.9166
	Different timestamp	0.7333	0.9453	0.8513	0.8410	0.8461

Table 6.3: without preprocessing and postprocessing snow prediction of segnet model

- All subsequent experiments will incorporate preprocessing and postprocessing steps to improve snow prediction accuracy. The models will be tested on both the same timestamp data and different timestamp data, but on the same location.

Model	Test	IoU	Accuracy	Precision	Recall	F1 Score
Segnet	Same timestamp	0.9380	0.9772	0.9616	0.9746	0.9680
	Different timestamp	0.7212	0.9427	0.9875	0.8288	0.8381
U-Net	Same timestamp	0.9677	0.9885	0.9917	0.9756	0.9836
	Different timestamp	0.8743	0.9769	0.9592	0.9081	0.9329
Res U-Net	Same timestamp	0.9738	0.9906	0.9836	0.9899	0.9868
	Different timestamp	0.8701	0.9761	0.9571	0.9054	0.9305
Attention U-Net	Same timestamp	0.9702	0.9893	0.9902	0.9796	0.9848
	Different timestamp	0.8805	0.9780	0.9574	0.9164	0.9365

Table 6.4: Snow prediction without Band3 and Band6

Model	Test	IoU	Accuracy	Precision	Recall	F1 Score
Segnet	Same timestamp	0.9457	0.9802	0.9717	0.9724	0.9720
	Different timestamp	0.7997	0.9615	0.9076	0.8706	0.8887
U-Net	Same timestamp	0.9746	0.9908	0.9846	0.9897	0.9872
	Different timestamp	0.8839	0.9785	0.9535	0.9237	0.9384
Res U-Net	Same timestamp	0.9712	0.9895	0.9786	0.9923	0.9854
	Different timestamp	0.8837	0.9782	0.9409	0.9356	0.9383
Attention U-Net	Same timestamp	0.9725	0.9907	0.9807	0.9916	0.9861
	Different timestamp	0.8936	0.9802	0.9485	0.9392	0.9438

Table 6.5: Snow prediction with Band3 and Band6

Model	Test	IoU	Accuracy	Precision	Recall	F1 Score
Segnet	Same timestamp	0.8528	0.9837	0.9039	0.9379	0.9206
	Different timestamp	0.6663	0.9910	0.9046	0.7166	0.7997
U-Net	Same timestamp	0.9041	0.9898	0.9472	0.9522	0.9497
	Different timestamp	0.7227	0.9929	0.9658	0.7417	0.8390
Res U-Net	Same timestamp	0.9039	0.9896	0.9273	0.9729	0.9496
	Different timestamp	0.7487	0.9935	0.9534	0.7771	0.8563
Attention U-Net	Same timestamp	0.9037	0.9897	0.9418	0.9571	0.9494
	Different timestamp	0.7569	0.9938	0.9701	0.7749	0.8616

Table 6.6: Ablation zone prediction by predicting snow without Band3 and Band6

Model	Test	IoU	Accuracy	Precision	Recall	F1 Score
Segnet	Same timestamp	0.8469	0.9826	0.8852	0.9513	0.9171
	Different timestamp	0.7114	0.9921	0.8950	0.7761	0.8313
U-Net	Same timestamp	0.9062	0.9898	0.9288	0.9738	0.9507
	Different timestamp	0.7378	0.9932	0.9619	0.7599	0.8491
Res U-Net	Same timestamp	0.8995	0.9890	0.9203	0.9755	0.9471
	Different timestamp	0.7568	0.9937	0.9580	0.7828	0.8616
Attention U-Net	Same timestamp	0.8995	0.9890	0.9206	0.9752	0.9471
	Different timestamp	0.7779	0.9942	0.9552	0.8073	0.8751

Table 6.7: Ablation zone prediction by predicting snow with Band3 and Band6

Chapter 7

Conclusion and Future Work

Our work focused on leveraging Landsat 9 data to predict snow cover and identify glacier ablation zones. We successfully collected and utilized Landsat 9 satellite data along with recent observations to train and test our models. Through the application of the Normalized Difference Snow Index (NDSI) and Digital Elevation Model (DEM), we accurately identified snow cover and labeled glacier ablation zones. Preprocessing and postprocessing steps, including connected component analysis and hole filling, were applied to improve the quality of label images. Our approach effectively addressed initial challenges in ablation zone prediction, yielding higher Intersection over Union (IoU) even for different timestamp data.

In our experiments, we compared four CNN models for snow and ablation zone prediction. Results showed that our models for snow prediction achieved promising performance metrics, including IoU, accuracy, precision, recall, and F1 score, demonstrating the effectiveness of our approach.

Our major contributions are:

- We collected the latest Landsat 9 satellite data and recent observations, training the model on June 23 and testing it on September 23.
- For preprocessing label images (snow label and ablation label), we employed connected component analysis to remove noise and performed hole filling to obtain connected components.
- We subsampled the original image and selected subimages that contained a minimum percentage of both labels to avoid imbalanced datasets.
- In postprocessing, we predicted pixels using k-nearest neighbors (KNN) where the probability was between 0.1 to 0.9, and applied connected component analysis to remove noise.
- Ablation zone prediction using CNN models initially yielded low IoU. However,

our approach mitigated these issues, enabling effective prediction of ablation zones even for different timestamp data.

- We predicted snow initially and then applied DEM thresholding to obtain glacier ablation zones, resulting in higher IoU.

A limitation arises from using the NDSI index to predict snow cover. Obtaining ground truth label data through field visits or manual mapping can significantly improve our predictions. Additionally, incorporating more geomorphometric parameters such as slope, aspect, and terrain roughness, along with geological factors like temperature and humidity, could enhance the accuracy of snow prediction models.

Future research could explore utilizing more complex or advanced CNN models, including attention mechanisms or graph-based approaches, to determine glacier ablation zones more effectively. Integrating multi-source remote sensing data, including SAR and optical imagery, could provide richer information for snow and glacier mapping. Incorporating climate data to analyze the impact of climate change on glacier dynamics would also be beneficial. Developing automated techniques for change detection and monitoring of glacier extents over time could provide valuable insights for understanding long-term glacier behavior and their contribution to water resources and sea-level rise. Investigating the use of high-resolution data, such as Sentinel data, could achieve more accurate results.

By addressing these areas, future work can build upon our findings to improve the prediction and monitoring of snow cover and glacier ablation zones, ultimately contributing to a better understanding of glacier dynamics and their environmental impacts.

Bibliography

[1] B. R. Singh and O. Singh, "A study about realities of climate change: Glacier melting and growing crises," in *Climate Change* (B. R. Singh, ed.), ch. 2, Rijeka: IntechOpen, 2013.

[2] S.-J. WANG and L.-Y. ZHOU, "Integrated impacts of climate change on glacier tourism," *Advances in Climate Change Research*, vol. 10, no. 2, pp. 71–79, 2019. Special issue on cryospheric functions and services.

[3] P. Rastner, T. Bolch, C. Notarnicola, and F. Paul, "A comparison of pixel- and object-based glacier classification with optical satellite images," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 3, pp. 853–862, 2014.

[4] B. A. Robson, C. Nuth, S. O. Dahl, D. Hölbling, T. Strozzi, and P. R. Nielsen, "Automated classification of debris-covered glaciers combining optical, sar and topographic data in an object-based environment," *Remote Sensing of Environment*, vol. 170, pp. 372–387, 2015.

[5] Z. Xie, U. K. Haritashya, V. K. Asari, B. W. Young, M. P. Bishop, and J. S. Kargel, "Glaciernet: A deep-learning approach for debris-covered glacier mapping," *IEEE Access*, vol. 8, pp. 83495–83510, 2020.

[6] V. Badrinarayanan, A. Kendall, and R. Cipolla, "Segnet: A deep convolutional encoder-decoder architecture for image segmentation," 2016.

[7] Z. Xie, U. Haritashya, V. Asari, M. Bishop, J. Kargel, and T. Aspiras, "Glaciernet2: A hybrid multi-model learning architecture for alpine glacier mapping," 04 2022.

[8] U. Minora, D. Bocchiola, C. D'Agata, D. Maragno, C. Mayer, A. Lambrecht, B. Mosconi, E. Vuillermoz, A. Senese, C. Compostella, C. Smiraglia, and G. Diolaiuti, "2001–2010 glacier changes in the central karakoram national park: a contribution to evaluate the magnitude and rate of the "karakoram anomaly"," *The Cryosphere Discussions*, vol. 7, pp. 7538–, 04 2013.

[9] A. Ali, P. Dunlop, S. Coleman, D. Kerr, R. W. McNabb, and R. Noormets, "Glacier area changes in novaya zemlya from 1986–89 to 2019–21 using object-based image analysis in google earth engine," *Journal of Glaciology*, vol. 69, no. 277, p. 1305–1316, 2023.

[10] J. G. Masek, M. A. Wulder, B. Markham, J. McCorkel, C. J. Crawford, J. Storey, and D. T. Jenstrom, "Landsat 9: Empowering open science and applications through continuity," *Remote Sensing of Environment*, vol. 248, p. 111968, 2020.

[11] https://space4water.org/taxonomy/term/1247.

[12] G. Guo, H. Wang, D. Bell, and Y. Bi, "Knn model-based approach in classification," 08 2004.

[13] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," 2015.

[14] F. I. Diakogiannis, F. Waldner, P. Caccetta, and C. Wu, "Resunet-a: A deep learning framework for semantic segmentation of remotely sensed data," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 162, p. 94–114, Apr. 2020.

[15] O. Oktay, J. Schlemper, L. L. Folgoc, M. Lee, M. Heinrich, K. Misawa, K. Mori, S. McDonagh, N. Y. Hammerla, B. Kainz, B. Glocker, and D. Rueckert, "Attention u-net: Learning where to look for the pancreas," 2018.

[16] A. Mao, M. Mohri, and Y. Zhong, "Cross-entropy loss functions: Theoretical analysis and applications," 2023.

[17] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2017.

[18] A. F. Agarap, "Deep learning using rectified linear units (relu)," 2019.

[19] L. Kramareva, A. Andreev, E. Simonenko, and A. Sorokin, "The use of a convolutional neural network for detecting snow according to the data of the multichannel satellite device of meteor-m no.2 spacecraft," *Procedia Computer Science*, vol. 150, pp. 368–375, 2019. Proceedings of the 13th International Symposium "Intelligent Systems 2018" (INTELS'18), 22-24 October, 2018, St. Petersburg, Russia.