

## Article

# Segmentation of Glacier Area Using U-Net through Landsat Satellite Imagery for Quantification of Glacier Recession and Its Impact on Marine Systems

Edmund Robbins <sup>1,†</sup> , Robert D. Breininger <sup>1,†</sup>, Maxwell Jiang <sup>2</sup>, Michelle Madera <sup>3</sup>, Ryan T. White <sup>1</sup>  and Nezamoddin N. Kachouie <sup>1,\*</sup>



**Citation:** Robbins, E.; Breininger, R.D.; Jiang, M.; Madera, M.; White, R.T.; Kachouie, N.N. Segmentation of Glacier Area Using U-Net through Landsat Satellite Imagery for Quantification of Glacier Recession and Its Impact on Marine Systems. *J. Mar. Sci. Eng.* **2024**, *12*, 1788. <https://doi.org/10.3390/jmse12101788>

Academic Editor: Mike Meylan

Received: 21 August 2024

Revised: 29 September 2024

Accepted: 2 October 2024

Published: 8 October 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

<sup>1</sup> Department of Mathematics and Systems Engineering, Florida Institute of Technology, Melbourne, FL 32901, USA; [erobbins2017@my.fit.edu](mailto:erobbins2017@my.fit.edu) (E.R.)

<sup>2</sup> Georgia Institute of Technology, Atlanta, GA 30332, USA

<sup>3</sup> The City College of New York, New York, NY 10031, USA

\* Correspondence: [nezamoddin@fit.edu](mailto:nezamoddin@fit.edu) † These authors contributed equally to this work.

**Abstract:** Glaciers have experienced a global trend of recession within the past century. Quantification of glacier variations using satellite imagery has been of great interest due to the importance of glaciers as freshwater resources and as indicators of climate change. Spatiotemporal glacier dynamics must be monitored to quantify glacier variations. The potential methods to quantify spatiotemporal glacier dynamics with increasing complexity levels include detecting the terminus location, measuring the length of the glacier from the accumulation zone to the terminus, quantifying the glacier surface area, and measuring glacier volume. Although some deep learning methods designed purposefully for glacier boundary segmentation have achieved acceptable results, these models are often localized to the region where their training data were acquired and further rely on the training sets that were often curated manually to highlight glacial regions. Due to the very large number of glaciers, it is practically impossible to perform a worldwide study of glacier dynamics using manual methods. As a result, an automated or semi-automated method is highly desirable. The current study has built upon our previous works moving towards identification methods of the 2D glacier profile for glacier area segmentation. In this study, a deep learning method is proposed for segmentation of temporal Landsat images to quantify the glacial region within the Mount Cook/Aoraki massif located in the Southern Alps/Ka Tiritiri o te Moana of New Zealand/Aotearoa. Segmented glacial regions can be further utilized to determine the relationship of their variations due to climate change. This model has demonstrated promising performance while trained on a relatively small dataset. The permanent ice and snow class was accurately segmented at a 92% rate by the proposed model.

**Keywords:** glacier variation; deep learning; U-net; satellite imagery; segmentation; land cover

*J. Mar. Sci. Eng.* **2024**, *12*, 1788. <https://doi.org/10.3390/jmse12101788>

<https://www.mdpi.com/journal/jmse>

ocean and air temperatures that create more frequent and intense coastal storms. Melting glaciers have been shown to impact oceanic currents, which can alter the global climate and lead to more extreme weather events [12,13]. For example, melting glaciers in the Arctic have led to an ocean current becoming faster and more turbulent [12]. However, increased Antarctic ice melt is predicted to cause the cold-water current cell of the global overturning circulation to slow down within a few decades [13].

## 1. Introduction

Glaciers are highly sensitive to climatic conditions, serving as a key indicator of climate change. Global glacial mass and the conditions of specific regional glaciers are often used as a proxy of the climate in global and regional areas as past quantitative data have shown that glacier variation is associated with climate change [1–5]. Over the past century, glaciers have exhibited a general trend of recession [6–8]. Rising CO<sub>2</sub> levels increase global temperatures, which melt glaciers faster, but precipitation patterns also change with the climate, making the exact relationship between climate change and glacial variation unclear.

Melting glaciers have been one of the main contributors to sea level change during the last century [8–10]. The mass loss of Greenland and Antarctic ice sheets and Arctic glaciers has been proven to contribute to global sea level rise [11]. Increasing sea levels can lead to increased coastal erosion and storm surge, which can be exacerbated by warming

Glaciers are home to many animals, both aquatic and terrestrial, and their melting can lead to the extinction of many species [14]. For example, glacier-fed freshwater ecosystems are expected to lose 11–38% of macroinvertebrate species following the disappearance of glaciers. In addition, glaciers are important as a water resource for downstream communities, and their decline with anthropogenic climate change is causing concern for the health and security of communities [15–18]. Therefore, it is a necessity to not only monitor glaciers to better understand global climate but to also study them to understand their changes in a shifting climate.

Understanding glacier variation is a challenging spatiotemporal problem for several reasons. One reason is that there is an abundance of climate and environmental factors that contribute to glacier behavior, such as elevation, temperature, and precipitation [19]. Another reason is that glacier variations are quantified by temporal dynamics of the glacier through the location of the glacier's terminus or the boundary of the glacier's surface area. For example, glacier's variations can be quantified by measuring the location of its terminus or the boundary of its surface area across many years; however, many glaciers are in remote locations, which means that quantifying glacier variation in the terminus point, surface area, volume, or mass is difficult with limited access on the ground. An alternative to ground measurement that has been widely used in the literature is remote sensing. Remote sensing by satellite is often the only feasible way of collecting large amounts of glacier data [20–22].

### 1.1. Glacial Surface Area Segmentation

Remote sensing is the acquiring of information from a distance, most frequently through aircraft and satellites. Remote sensing platforms employ multispectral sensors that collect data such as spectral wavelength bands. These bands can be combined to produce imagery of the data to study different characteristics of the landscape. A significant source of data used in the analysis of glacier change is the satellite imagery of glaciers. There are several satellites that have existing datasets and databases that contain decades of imagery available to download, such as the NASA/USGS Landsat series and the Copernicus Sentinel series. For this project, Landsat data were used because of its public availability and because the Landsat series is the longest continuous space-based record of satellite imagery. Other satellites, such as the Sentinel series, collect imagery at a better resolution and quality than Landsat but they have not collected data for the same duration as the Landsat program.

Collecting data by these satellites are useful but not perfect as clouds and mountain shadows may obscure the glaciers [23]. Various image processing techniques have been developed to detect clouds and remove their shadows to ensure that these imperfect images can be corrected to create a dataset of high-quality images. These techniques often use the physical properties of clouds such as their relatively brighter color and low temperatures, or a reference to a clear sky image, to detect clouds. Shadows can be detected in similar fashion by observing relatively darker colors and lower temperatures. However, as clouds are common over most glaciers, the number of images removed is large, resulting in a significantly diminished dataset [23].

This dataset is then further processed to highlight the glaciers, allowing for their size—and therefore approximate health—to be determined. This process is often complicated by snow and debris cover, making it unclear where the glacier terminates and rock or seasonal snow begins. A general model for the detection of all types of glaciers is sought after but very difficult to achieve [24].

### 1.2. Previous Work

Quantification of glacier changes using satellite imagery has been of great interest. Spatiotemporal glacier dynamics must be monitored to quantify glacier variations. The potential methods to quantify glacier dynamics with different complexity levels include detecting the terminus location over time, measuring the length of the glacier from the accumulation zone to the terminus over time, quantifying the glacier surface area over time, and measuring glacier volume over time. In our previous work, Kachouie et al., 2013 [25], inflection points were identified along smoothed glacier intensity profiles to locate the terminus points of a glacier and recover the pattern of terminus change over time. A constrained bandwidth selection method in local polynomial regression was developed for locating inflection points, which allows for flexibility in estimating change point locations when the underlying regression function has a single change point. Then, we extended that method in Kachouie et al., 2015 [26] by improving the estimation

of the first derivative of the intensity profile using nonparametric local polynomial regression, including inflection points as potential terminus candidates to track the pool of candidates over time. The technique was applied to the Franz Josef glacier/Ka Roimata o Hine Hukatere of New Zealand/Aotearoa and the Gorner glacier of Switzerland, and then extrapolated to the Viedma glacier in Patagonia. In our work in Onyejekwe et al., 2017 [27], we further developed the methodology of the previous papers by using multivariate regression to predict the glacier terminus location over time based on observed climatic factors. The terminus location was calculated as the distance from a previous reference point, and then converted into a cumulative terminus change time-series dataset. The terminus change was modeled against climate factors such as the global and local average annual temperature, and annual average atmospheric carbon dioxide concentration. Most recently, in Robbins et al., 2023 [28], we continued our previous work to develop a semiautomated method for the identification of the glacier terminus and area [28]. This method used a manually drawn glacier path/area to find the intensity of the pixel values along the paths or the binarized region that is considered the largest connected glacier area. The glacier variation was also modeled against climate factors using parametric and nonparametric regression.

Oerlemans 2005 used glacier length change records to reconstruct temperature histories for different parts of the world using a first-order theory of glacier dynamics [29]. The temperature data and reprocessed glacier data were able to show the magnitude of current global warming, the time that this warming started, and highlighted that the warming appears to be independent of elevation in the lower troposphere. Cullen et al., 2006 is a case study on the glacial recession of Kilimanjaro glaciers [30]. Images of Kilimanjaro glaciers from Quickbird satellite imagery, a georeferenced image, and aerial photography of Kilimanjaro were compared to track glacier change through time. First, transient snow was masked from the images and then the boundary between transient snow and glacial ice was delineated to find the change between glaciers. The results suggest glaciers on Kilimanjaro are merely remnants of past climate conditions rather than indicators of 20th-century climate change. Bhattacharya et al., 2021 found—using a satellite-based time-series of glacier mass balance for seven climatically different regions across High Mountain Asia since the 1960s—that glacier mass loss rates have persistently increased at most sites [31]. This was achieved through calculating glacier mass budgets from satellite imagery and digital elevation models. It was found that an increase in summer temperature explains the long-term trend in mass loss.

The current study has built upon our previous works moving toward identification methods of the 2D glacier profile for glacier area segmentation. Terminus points provide insights into glacial variation, but as a zero-dimensional (0D) representation, do not supply an encompassing view of glacial variations. Quantification of glacier dynamics with higher dimensionality is more informative but requires more complex methods. The 2D representation, i.e., quantification of the glacier surface area, is preferred over 0D and 1D as it quantifies glacier dynamics at a higher dimension. Hence, as a natural extension of our previous work, a 2D quantification method using deep learning is proposed here.

In this study, image segmentation methods will be used to segment glacier regions in temporal satellite imagery. Segmented glacial regions can be further utilized to determine the relationship of their variations due to the climate change.

Deep learning has shown promise in improving image segmentation algorithms [32]. Deep learning has significant applications in various fields of environmental science, such as climate modeling, water monitoring, and glacier area segmentation. Convolutional neural network (CNN) has been broadly used for image segmentation. CNN belongs to the category of feed-forward neural networks that is designed to process grid data, such as image and visual data, using convolutional layers to form kernels and create feature maps of the input data [33]. CNN excels in classification and computer vision tasks, such as image classification and segmentation.

The U-net (an extension of CNN) has demonstrated high performance in the field of image segmentation [34]. The U-net was introduced by Ronneberger et al. based on the fully convolutional network developed by Long et al. [35,36]. Importantly, U-net can be trained on a relatively small dataset and intensive augmentation methods. U-net has outperformed other popular CNNs. Further, CNNs have been used in the literature for glacier segmentation. For example, Xie et al. utilized a CNN called GlacierNet on Landsat imagery combined with a digital

elevation model to identify glaciers with decent accuracy [36]. However, this work was limited in its generalizability and ease of use because of the complex dataset required to train the network.

Although there are existing neural network architectures created and trained purposefully for glaciers—such as GlacierNet—that achieve acceptable results, these models are often localized to where their training data were acquired. Further, these training sets were often curated by hand to highlight glacial features and represent high-quality data. However, there are too many glaciers and glacial regions in the world for similar training sets to be created. Therefore, an effective computer image processing pipeline needs to be established to turn low-quality data into higher quality comparable to labeled datasets.

This paper intends to quantify variations in the glaciers within the Mount Cook/Aoraki massif located in the Southern Alps/Ka Tiritiri o te Moana of New Zealand/Aotearoa using Landsat satellite imagery and image processing methods. In particular, the goal of this paper is to apply image segmentation methods through deep learning to quantify glacier variation using satellite imagery.

## 2. Data

The Mount Cook/Aoraki massif in the South Island/Te Waipounamu of New Zealand/Aotearoa was chosen as the study area as this work is a continuation of our previous work focused on the Franz Josef glacier/Ka Roimata o Hine Hukatere on the northwest flank of Mount Cook/Aoraki. Particularly, the Franz Josef Glacier/Ka Roimata o Hine Hukatere is unique in how it fluctuates between retreat and advancement in contrast with general retreating trends among mountain glaciers [37,38]. Further, the mountain range possesses an abundance of glaciers flowing off its sides in both a continental (eastern) and maritime (western) setting, expanding the variance in the data. In order to quantify the glacier area using image segmentation, satellite imagery and annotated/labeled data are required. The areas of interest in this study are visualized in Figure 1.

### 2.1. Landsat

The Landsat satellite project is the longest continuous space-based record of satellite imagery, run by both NASA and USGS. The first Landsat satellite was launched in 1972, and currently, satellites 7, 8, and 9 remain in orbit. Landsat 7, launched in April of 1999, was drawn on for images prior to the launch of Landsat 8 in February of 2013. Both satellites have a repeat cycle of 16 days and complete an orbit around the Earth in 99 min (USGS, 2023, Landsat 7 Satellite Orbit Facts section) [39]. Landsat 7 has an Enhanced Thematic Mapper Plus (ETM+) instrument that contains eight spectral bands, as seen in Table 1 (USGS, 2023, Landsat 7 Enhanced Thematic Mapper Plus (ETM+) Instrument section). Landsat 8 and 9 hold two sensors—the Operational Land Imager sensor and the Thermal Infrared Sensor—comprising nine and two spectral bands, respectively, as seen in Table 2 (USGS, 2023, Landsat 8 Instruments section). To produce the color image using Landsat 7, bands 3, 2, and 1 are composited together as they are the red, green, and blue bands, respectively; bands 4, 3, and 2 are employed with Landsat 8 and 9. In addition to representing the same information, these bands were also selected for their consistent resolution of 30 m. This ensures each band contains the same amount of information and can be composited together seamlessly.



**Figure 1.** Glacier of interest for this study, with cropped satellite scenes superimposed onto a world map. The Franz Josef glacier/Ka Roimata o Hine Hukatere is located at [ $\sim -43.467, 170.192$ ]. Red arrows point to the zoomed-in region of Franz Josef glacier.

**Table 1.** Landsat 7 band wavelengths and resolutions. Note: adapted from USGS (USGS, 2023, Landsat 7 Enhanced Thematic Mapper Plus (ETM+) section).

| Bands                                | Wavelength ( $\mu\text{m}$ ) | Resolution (m) |
|--------------------------------------|------------------------------|----------------|
| Band 1—Blue                          | 0.45–0.52                    | 30             |
| Band 2—Green                         | 0.52–0.60                    | 30             |
| Band 3—Red                           | 0.63–0.69                    | 30             |
| Band 4—Near Infrared                 | 0.77–0.90                    | 30             |
| Band 5—Shortwave Infrared 1          | 1.55–1.75                    | 30             |
| Band 6—Thermal                       | 10.40–12.50                  | 60             |
| Band 7—Shortwave Infrared 2          | 2.09–2.35                    | 30             |
| Band 8—Panchromatic (entire visible) | 0.52–0.90                    | 15             |

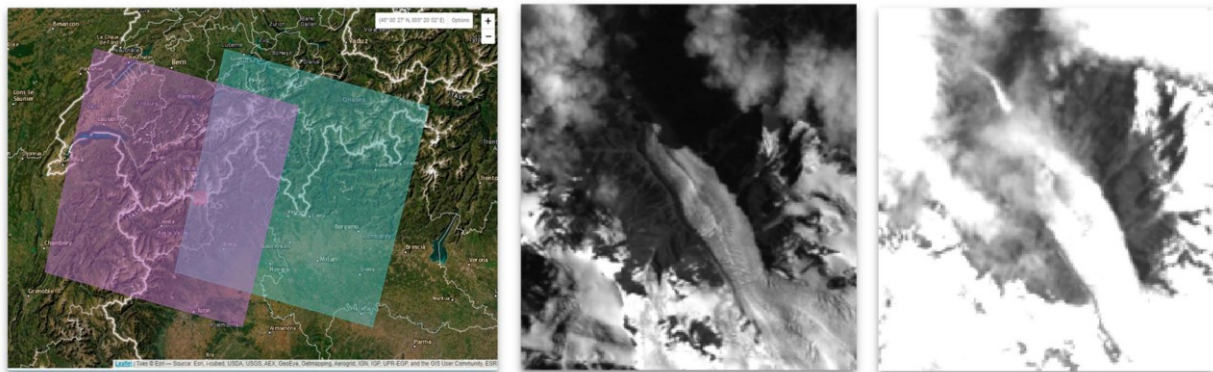
In May of 2003, Landsat 7's scan line corrector failed, creating gaps in its images with only 78 percent of the total pixels remaining (USGS, 2023, Landsat 7 Scan Line Corrector (SLC) Failure section). Therefore, the Landsat 7 images utilized in this study were processed by interpolation of the missing data to create complete images for the model input.

To download a package of scenes from the USGS Earth Explorer website, the Mount Cook/Aoraki massif is first selected using the map coordinates feature. The date range and cloud cover ramifications are then specified. Scenes from 1 January to 30 May of the relevant years were collected, with a preferred cloud cover range under 30 percent. Satellites of interest were then selected as datasets, with Landsat 7 Collection 2 Levels 1 and 2 for scenes prior to 2013. For the remaining years up to 2023, the same Landsat 7 selection was performed, with the addition of Landsat 8 and 9 Collection 2 Level-2. Landsat 8 and 9 Collection 2 Level-2 scenes were preprocessed and thus contain regions where imperfect images were removed, leaving areas of no data. As the labels are built from the Landsat images, these areas are correspondingly marked as containing no data and thus do not disrupt the model.

**Table 2.** Landsat 8–9 band wavelengths and resolutions. Note: adapted from USGS (USGS, 2023, Landsat 8–9 Operational Land Imager (OLI) and Thermal Infrared Sensor (TIRS) section).

| Bands                                | Wavelength ( $\mu\text{m}$ ) | Resolution (m) |
|--------------------------------------|------------------------------|----------------|
| Band 1—Visible Coastal Aerosol       | 0.43–0.45                    | 30             |
| Band 2—Visible Blue                  | 0.45–0.51                    | 30             |
| Band 3—Visible Green                 | 0.53–0.59                    | 30             |
| Band 4—Red                           | 0.64–0.67                    | 30             |
| Band 5—Near Infrared                 | 0.85–0.88                    | 30             |
| Band 6—Shortwave Infrared 1          | 1.57–1.65                    | 60             |
| Band 7—Shortwave Infrared 2          | 2.11–2.29                    | 30             |
| Band 8—Panchromatic (entire visible) | 0.50–0.68                    | 15             |
| Band 9—Cirrus                        | 1.36–1.38                    | 30             |
| Band 10—Thermal Infrared Sensor 1    | 10.60–11.19                  | 100            |
| Band 11—Thermal Infrared Sensor 2    | 11.50–12.51                  | 100            |

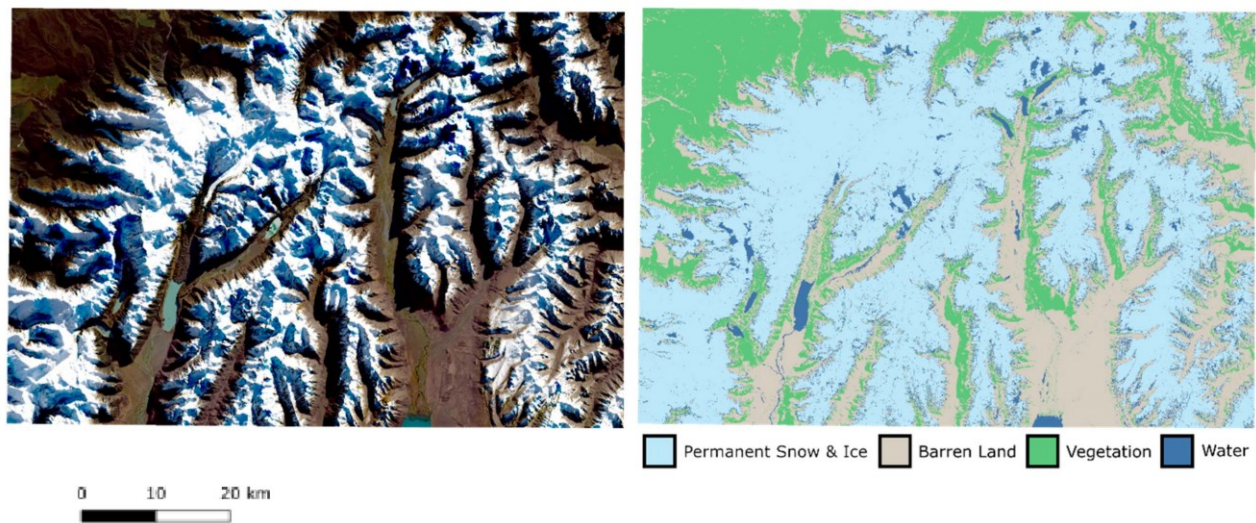
Landsat satellites have different paths and alignments, which can make collecting scenes of a particular location difficult, visualized in Figure 2 (left). There are many types of issues with collecting Landsat scenes of an exact location. A glacier can be obscured by mountain shadow due to the time of day or year the image was taken, as seen in Figure 2 (middle). The terminus of the Franz Josef glacier/Ka Roimata o Hine Hukatere, as seen in Figure 2 (middle), is contained within a mountain valley and is often obscured by surrounding mountains. Also, glaciers can be covered by clouds in satellite images, as seen in Figure 2 (left).



**Figure 2.** (left) Two overlapping Landsat tiles (purple and green) with different Landsat paths and alignments; (middle) Cropped Landsat scene of Franz Josef glacier/Ka Roimata o Hine Hukatere with shadows; (right) Cropped Landsat scene of Franz Josef glacier/Ka Roimata o Hine Hukatere with cloud cover. The snout of Franz Josef glacier/Ka Roimata o Hine Hukatere is located at  $[-43.458, 170.185]$ .

## 2.2. Land Cover Database

To provide a truth for the model to reference as it trained, land cover masks of the study area were created from the New Zealand Landcover Database (LCDB) from the Manaaki Whenua Landcare Research (2021) [40]. This database was developed for the nation of New Zealand/Aotearoa from classified SPOT satellite imagery, Landsat 7 imagery, ETM+ satellite imagery, aerial photography, and ancillary data such as digital topographical data and published topographical maps. The land classifications developed using imagery were also hand corrected. The database contained land classifications for the years 2001, 2008, 2012, and 2018 in the form of shapefiles containing polygons demarcating various classifications. Figure 3 gives an example of one of the rasterized masks created using the land cover database alongside the corresponding Landsat scene.

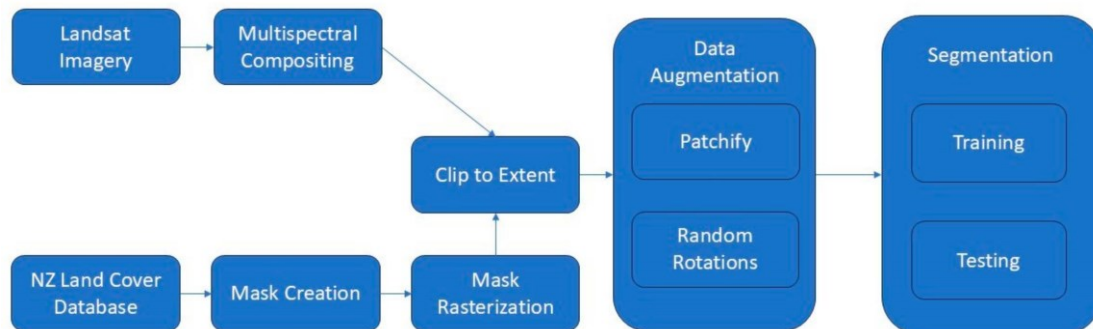


**Figure 3.** Comparison of Landsat scene of the Mount Cook/Aoraki massif (**left**) and New Zealand land cover database rasterized mask (**right**). The center of the scenes of the Mount Cook/Aoraki massif is located at  $[-43.611, 170.367]$ .

The land cover database's shapefiles contained 47 Vegetative Cover classes and we combined them to a single vegetation class. The total number of classes was reduced to four, including vegetation, snow, water, and barren land/ground classes. This new shape file was then employed for the land cover masks for training.

### 3. Methods

The methodology used in this study is visualized in Figure 4. The satellite imagery and the labeled land cover masks are preprocessed and then augmented before training the deep learning segmentation model.



**Figure 4.** Flow diagram of the proposed methods in this study.

Preprocessing methods were applied to the collected Landsat scenes in QGIS for the correction of post-2003 Landsat 7 images and the production of composite imagery. Land cover masks were generated from the New Zealand land cover database. The land cover shapefiles were clipped to the extent of their relevant band images, and the clipped mask was rasterized with a resolution of 10,000 by 10,000 pixels. Symbology properties were then applied to provide each class with a distinct color, and the mask was exported as a rendered image. The rasterized masks were then fine-tuned using an ensemble of four machine learning and statistical classification algorithms—random forest, maximum likelihood estimation (MLE), k-nearest neighbor, and support vector machine (SVM). The results of those methods were combined with maximum voting. These rasterized land cover masks are considered the “ground truth” for this study. However, it is important to note that these land cover masks are generated from derived data, not direct in situ measurements from either ground level or a remote sensing platform. The derived data were validated and corrected by hand. The derived data are considered “ground truth” because in situ labeled data for glaciers are sparse.

For each of the years that the land cover database had data, Landsat scenes were acquired for the year of interest and years directly before and after. For example, the land cover data for

2012 was applied to scenes from 2011, 2012, and 2013. To generate a sufficient number of readable inputs for the model, various preprocessing techniques were performed. Landsat 7 scenes were adjusted to account for missing data points and cropped to focus on more relevant (glacial) areas. The land cover of these scenes was then estimated utilizing an ensemble of various other models. Composite images were procured for color. These composites along with the land cover masks were then divided into segments of 256 by 256 pixels.

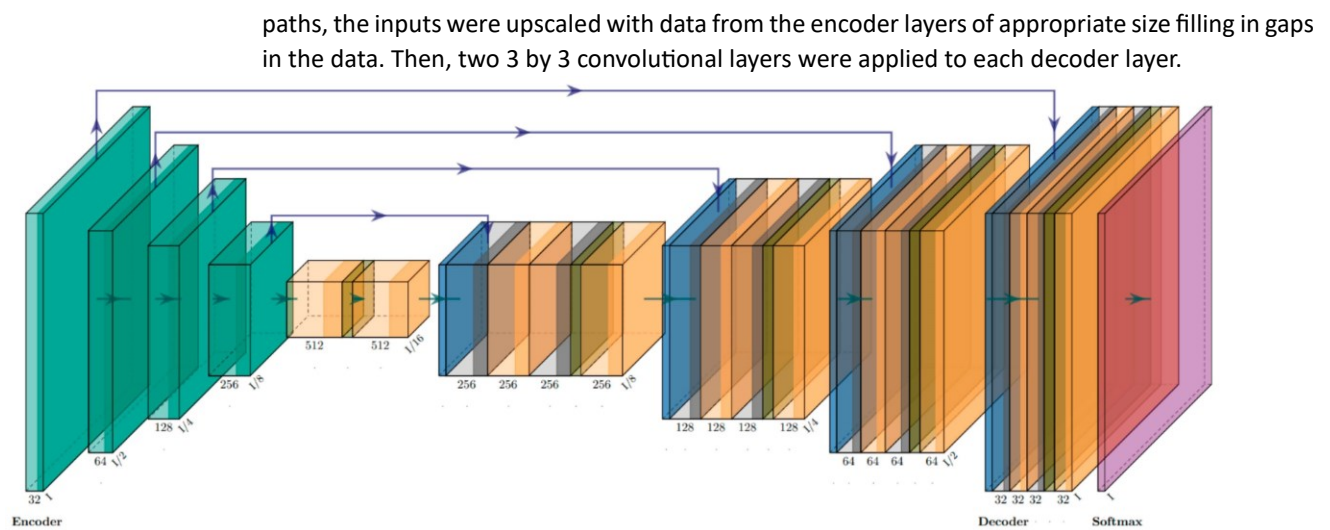
To adjust for the malfunction of Landsat 7's scan line corrector, the USGS provides gap mask information with the download of their satellite packages. Each band has a corresponding gap mask that relays the location of pixels affected by the scan line corrector that in turn contains no data. These gap mask files were employed through QGIS's "Fill nodata" function as validity masks for the corresponding input band layer to replace the sections of the scene without data (and visible striations).

Complete Landsat 7 and 8/9 RGB bands were composed together to create a singular input for the model. The input size was limited to three bands to reduce the necessary number of computational resources needed for model training. The three optical bands were chosen because they possessed the highest resolution with multiple bands (30 m) and the alternative, the infrared bands, do not offer any significant advantages over optical. Both band types face problems with shadows; therefore, to make the model more familiar to the human eye, optical was chosen. Then, as Landsat scenes are significantly larger than required and cover significantly more non-ice areas than permanent ice, the images were cropped to the mountains around Mount Cook/Aoraki. A total of 16 Landsat scenes were used for training the neural network.

The model used to segment the glacier images was derived from the U-net convolutional neural network [34]. The U-net is a neural network architecture that was designed for image segmentation, originally being applied to the segmentation of human cells. The U-net architecture is known for its U-shaped structure, which consists of a contracting path (left side) and an expansive path (right side). The contracting path captures context and extracts features, while the expansive path allows precise localization. The contracting path, also known as the encoder, starts with multiple convolutional layers followed by rectified linear unit (ReLU) activations. Then, max-pooling layers are used to down-sample the spatial dimensions, reducing the size of feature maps while increasing the receptive field. Skip connections (or shortcuts) are established between the contracting and expansive paths to facilitate the transfer of low-level information and gradients. Skip connections are a crucial component of the U-Net architecture. They help preserve fine-grained details by allowing low-level features from the contracting path to be directly combined with high-level features in the expansive path. The expansive path, also known as the decoder, gradually up-samples the feature maps to the original input size. Each up-sampling step involves a combination of up-sampling layers and concatenation with feature maps from the contracting path. These concatenated feature maps are then processed with convolutional layers. The expansive path ends with a final convolutional layer to produce the segmentation mask.

The U-Net architecture has proven to be effective for a wide range of image segmentation tasks, and its adaptability and ability to preserve spatial information through skip connections make it a popular choice in the computer vision and medical imaging communities. The U-net architecture was used in this study as it has been proven to provide detailed segmentations on a few training images.

The model generally followed a U-net architecture in that it had five encoder layers, one bridge layer, and five decoder layers, as shown in Figure 5; however, the convolutional layers of the encoder were replaced with the convolutional layers from the VGG16 model trained on ImageNet. VGG16 is another CNN designed for image classification, while ImageNet is an image dataset purposefully created for image classification training [41]. The pre-trained VGG16 convolutional layers used were already tuned to break down an image into structure-identifying features. By using convolutional layers from the already trained VGG16 network, the performance of the U-net used for this study can be improved through transfer learning. Transfer learning is a machine learning method where a model for one domain/task can be improved by transferring information from a related domain/task [42,43]. This would allow for the model to improve its accuracy or reduce the time of training, as the model would no longer need to learn such techniques independently. In the bridge layer, there was simply a 3 by 3 convolutional layer. In the decoder



**Figure 5.** U-net architecture (created using LaTeX package TikZ v2). The blue-green boxes in the encoder (**left half**) are the VGG16 convolutional layers. The light orange banded boxes in the bottleneck and the decoder (**right half**) are 3 (standard) convolutional layers. The yellow-green boxes between the convolutions are dropout layers. The dark blue boxes in the decoder are the up-convolutional layers. The final purple layer is the softmax layer. The arrows connecting the encoder layers to the decoder layers (above the boxes) are the skip connections.

To reduce model complexity, and therefore required training time and resources, each input image and its associated labels were split into 256-by-256 patches for the model to read. Each patch was initially 256 pixels horizontally and vertically apart from the surrounding patches. Intermediary patches starting and ending in-between the initial patches were then added to increase the diversity of patches from the original images to increase the size of the dataset and reduce overfitting. Then, to further help reduce overfitting, the patches were supplemented with random rotations and flips. Each patch had a one-third chance for either its horizontal flip, vertical flip, or rotation to be added to the inputs. If a patch was being rotated, it had a one-third chance of either being rotated 90°, 180°, or 270°. After patchification and data augmentation, 20% of the data—1494 patches—was allocated randomly to a testing/validation dataset and the rest to a training dataset comprising 5976 patches.

All optimizers and loss functions available in Keras were tested on a diminished dataset to evaluate their impact on and performance with the established model. The Adam optimizer was found to be the most effective, and the accessible loss functions were run on a version of the model with this optimizer. After that round of testing, a custom total loss function was found to be the optimal loss function. This loss function was defined as the combination of dice loss and focal loss and was implemented with the Adam optimizer for all future explorations.

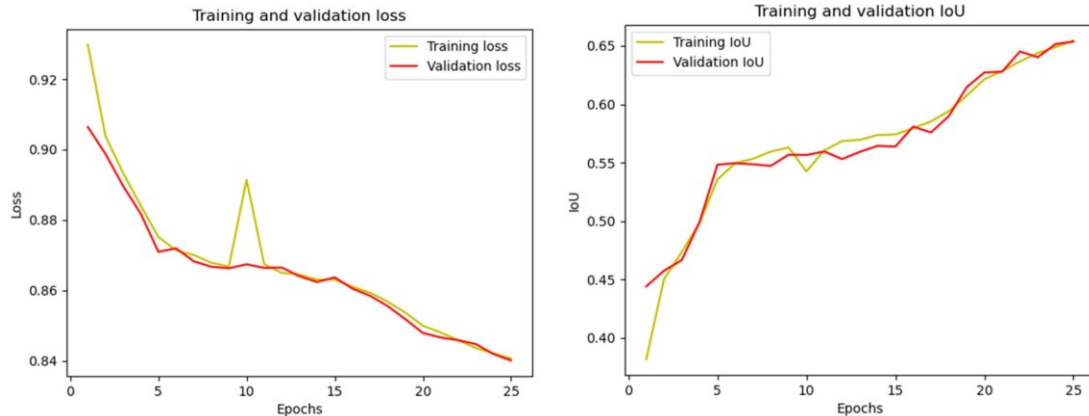
Certain parameters like epochs, dropouts, class weights, and learning rate were also tuned. The final model was trained for 400 epochs. A dropout of 10 percent was found to be ideal, with higher dropouts causing the model to learn too little while lower dropouts saw the model's training and validation results diverge. The study area saw a roughly even mix of all classes, thus, each class was given the same weight. Initially, the learning rate was set high, but as the model plateaued, the learning rate was lowered to help the model make fine adjustments.

The model was evaluated based on the Jaccard coefficient, otherwise known as Intersection over Union (IoU). Predicted pixels that matched the truth were counted and divided by the sum of pixels in the input image and predicted pixels that did not match the truth. As the model was tasked to classify 4 classes to balance the distribution of pixels but only the classification ability of ice and snow was of interest, an imperfect IoU was acceptable if the bulk of the error was from non-snow/ice pixels.

#### 4. Results

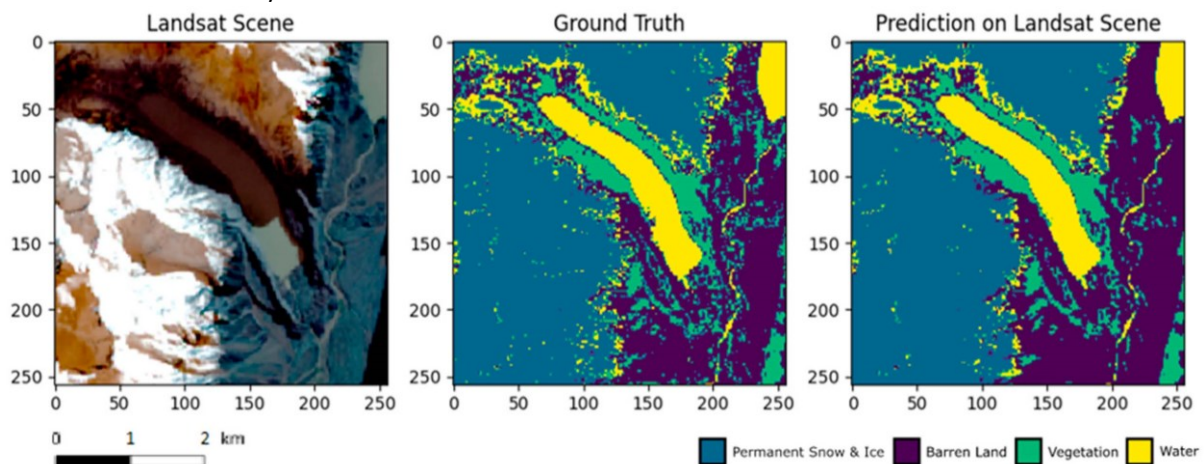
In this study, the performance of the neural network was assessed using IoU during training and validation and a confusion matrix with all of the classes. As seen in Figure 6, the training and validation loss and IoU reached a value of 0.65 at 25 epochs. As the number of epochs increased,

the loss function dropped and reached to 0.84 and the IoU value increased to 0.65, while the validation curve (in red) closely followed the training curve (in yellow). This implies that the model did not overfit or underfit to the data and was generalized. Loss and IoU were computed to quantify the performance of the model. Next, the prediction performance of the model is computed and compared to the land-cover ground truth and satellite scenes.

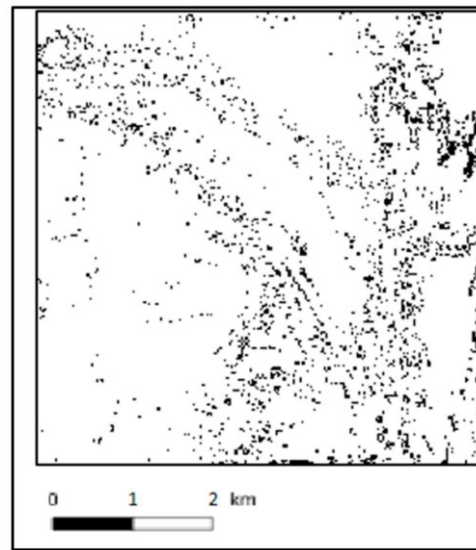


**Figure 6.** Loss and IoU curve for the U-net over 25 epochs. The red curve represents the validation values, and the yellow curve represents the training values.

A visual inspection of Figure 7 reveals that the trained model makes predictions that are very close to the ensemble-generated ground truth land cover masks. The ice–snow region representing the glacier is preserved by the model, as seen in Figure 7 (left). However, there are noticeable differences between the ground truth land cover masks and the prediction on the Landsat scene. This difference is visualized in Figure 8, which shows the pixels that are different by comparing the predicted scene with the ground truth land cover mask. While the primary regions are preserved by the model, there are some small groups of pixels that were not classified correctly by the model.

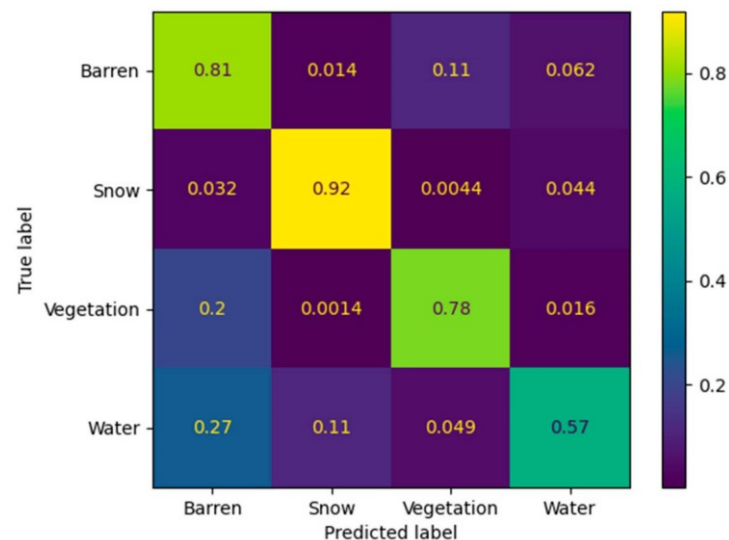


**Figure 7.** Comparison of input patch (left), rasterized land cover mask (middle), and prediction from the U-net (right) using the rasterized land cover mask. The center of these patches is located at  $[-43.508, 170.478]$ .



**Figure 8.** Difference between ground truth land cover mask and prediction from the U-net.

As shown by Figure 9, the model correctly identifies ice and snow with a classification rate of 92% of the pixels that truly were ice and snow. The greatest error of ice and snow was the misclassification of it as water, which occurred at a rate of 4.4%. Water naturally exists around ice and snow and, during the summer, within the ice and snow regions as they melt, providing long borders between the two classes. The barren land class was classified correctly at a rate of 81%; barren land pixels were misclassified most frequently as vegetation pixels at a rate of 11%. This could be due to having vegetation regions adjacent to deforested areas that are classified as barren land. The vegetation class was classified correctly at a rate of 78%; vegetation pixels were misclassified primarily as barren land pixels at a rate of 20%. This could be due to vegetation color being similar to their surrounding environment. The time of year also could have an impact, as all the Landsat scenes used were taken during summer months. The vegetation class is wide, as it includes many different types of vegetated areas, such as agricultural areas and wild forests. Some of these vegetation subclasses could be mistaken for barren land, for example, an agricultural region that was recently harvested would appear to be similar to a barren area. The water class had the lowest correct classification rate at 57%, which means that water pixels were misclassified the most out of all classes. This is likely because water bodies are often small and thus have a relatively large perimeter bounding other regions relative to the total area. These boundaries are where misclassifications occur in general. Water pixels were misclassified as barren land pixels at a rate of 27%. A reason why water is misclassified the most often, especially as barren land, is that all Landsat scenes were taken during the summer months when erosion is the strongest and the runoff discolors the water bodies to be more similar to barren land.



**Figure 9.** Confusion matrix of the U-net. True vs. predicted counts are scaled between zero and one.

## 5. Discussion and Conclusions

Automated or semi-automated methods for the quantification of glacier dynamics at higher dimensions are desirable and essential to further study the impact of climate change on glacier dynamics. A 2D quantification method using deep learning was proposed here. In this paper, the mountain glacial region within the Mount Cook/Aoraki massif located in the Southern Alps/Ka Tiritiri o te Moana of New Zealand/Aotearoa was segmented through Landsat satellite imagery. A deep learning neural network was implemented for glacier segmentation to quantify glacier variation. Segmented glacial regions can be further utilized to determine the relationship of glacier variations with climate change. The contributions of this study are summarized below:

- Proposing an automated method: there is an immense number of glaciers over the globe and an automated or semi-automated method for the quantification of glacier dynamics is crucial;
- A multiclass land cover classification: The proposed model demonstrated great capability in the classification of ice and snow from its surrounding features. The multiclass approach broadens the applicability of the methodology to other related problems, such as the classification of water quality of different water bodies, classification of different landscapes, and land surveying;
- This method was trained on several glaciers contained within the Mount Cook/Aoraki massif in New Zealand/Aotearoa. By training the model on a glacial region, rather than just a single glacier, the model is better trained to generalize and avoid overfitting. Including data from various mountain glacial regions will allow the model to generalize even further during the training phase. The model in this study can be potentially applied to the other glacial regions across the world.

The model in this study has the potential to be applied to any glacial region and is not limited by the availability of external ground truths. Many glaciers have very limited ground truth, no ground truth, or the available ground truth does not have sufficient quality. Detecting snow at a rate of 92% demonstrates promising results and encourages further application of the proposed method to other glacial regions. Most notably, the model ran on a relatively small dataset of patches generated from 16 cropped Landsat scenes and performed up to par with other image segmentation CNNs.

Our future work is focusing on addressing the challenges that yet to be addressed to further improve the model and listed below:

- The model was able to classify the snow/ice class with satisfactory accuracy. The water class has the highest misclassification rate compared to other classes, as seen in Figure 9. The model was able to correctly identify certain bodies of water, depicted by the dark blue regions, while some other water bodies were misclassified. Further exploration of the cropped input images displayed that patch boundaries may impact the misclassification rate;
- The amount of data available for training and validation is limited by the New Zealand/Aotearoa land cover database. The model ran on a relatively small dataset of patches generated from 16 cropped Landsat scenes and performed up to par with other image segmentation CNNs. Increasing the number of Landsat scenes used for training and validation will improve the performance of the model. The model can be trained on multiple glacial regions, such as extending the study area beyond the Mount Cook/Aoraki massif in the South Island/Te Waipounamu of New Zealand/Aotearoa to the rest of the country. For example, previous work studied glaciers such as the Gorner and Rhone glaciers in Switzerland, the Viedma glacier in Patagonia, the White Glacier in the northwest USA, and the Zemu Glacier in the Himalayas [25–27];
- The collected annotated data were limited to the temporal availability of the land cover database. For a broader multi-decadal study of mountain glacial dynamics, annotated data over an extended time are required. Glaciers do not have a substantial daily/weekly/monthly variation, rather noticeable changes happen over longer time frames (annual/decadal). Hence, to monitor and quantify glacial dynamics, multidecadal annotated data are required;
- The model performance can be potentially improved by changing the model architecture by supplementing it with additional layers or by changing to a more powerful architecture, such as a transformer. However, one issue is that these architectures require a large amount of

training data. This can be a potential issue, as many glaciers lack sufficient ground truth data. To use the more complex neural network architectures, additional labeled data will be crucial. A particularly well-suited dataset for investigating glacier change is the Global Land Ice Measurements from Space (GLIMS) [44]. GLIMS is an international initiative and a database system used for monitoring and studying changes in Earth's glaciers and ice sheets. The GLIMS Glacier Database currently has 604,986 entries for individual glaciers. However, it does not provide multi-class land cover annotation and can only be used for binary classification;

- An issue caused by the patchification of the original image is imperfections and misclassifications along the stitching lines when segmented patches will be stitched to form the segmented scene. Future research will be conducted to address the segmented inconsistencies along the stitched patches to reduce the occurrence of misclassification due to sharp cut-offs and edges.

**Author Contributions:** Conceptualization, N.N.K.; Methodology, E.R., R.D.B. and N.N.K.; Software, E.R., R.D.B., M.J. and M.M.; Validation, E.R. and R.D.B.; Formal analysis, E.R., R.D.B., M.J., M.M., R.T.W. and N.N.K.; Investigation, E.R., R.D.B., R.T.W. and N.N.K.; Resources, N.N.K.; Data curation, E.R., M.J. and M.M.; Writing—original draft, E.R., R.D.B., M.J., M.M. and N.N.K.; Writing—review & editing, N.N.K. and R.D.B.; Visualization, E.R., R.D.B., M.J. and M.M.; Supervision, R.T.W. and N.N.K.; Project administration, N.N.K.; Funding acquisition, N.N.K. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was performed as part of an REU (Research Experiences for Undergraduates) in statistical models with applications to geoscience awarded by the NSF (Grant # 1950768).

**Data Availability Statement:** Data is publicly available through [USGS.gov](https://www.usgs.gov/).

**Conflicts of Interest:** The authors declare no conflict of interests.

## References

1. Oerlemans, J.; Fortuin, J.P.F. Sensitivity of glaciers and small ice caps to greenhouse warming. *Science* **1992**, *258*, 115–117. [\[CrossRef\]](#) [\[PubMed\]](#)
2. Lüthi, M.P.; Bauder, A.; Funk, M. Volume change reconstruction of Swiss glaciers from length change data. *J. Geophys. Res.* **2010**, *115*. [\[CrossRef\]](#)
3. Glasser, N.F.; Harrison, S.; Jansson, K.N.; Anderson, K.; Cowley, A. Global sea-level contribution from the Patagonian icefields since the little ice age maximum. *Nat. Geosci.* **2011**, *4*, 303–307. [\[CrossRef\]](#)
4. Leclercq, P.W.; Oerlemans, J.; Cogley, J.G. Estimating the glacier contribution to sea-level rise for the period 1800–2005. *Surv. Geophys.* **2011**, *32*, 519–535. [\[CrossRef\]](#)
5. Leclercq, P.W.; Oerlemans, J. Global and hemispheric temperature reconstruction from glacier length fluctuations. *Clim. Dyn.* **2011**, *38*, 1065–1079. [\[CrossRef\]](#)
6. Leclercq, P.W.; Oerlemans, J.; Basagic, H.J.; Bushueva, I.; Cook, A.J.; Le Bris, R. A data set of worldwide glacier length fluctuations. *Cryosphere* **2014**, *8*, 659–672. [\[CrossRef\]](#)
7. Barry, R.G. The status of research on glaciers and global glacier recession: A review. *Prog. Phys. Geogr.* **2006**, *30*, 285–306. [\[CrossRef\]](#)
8. Meier, M.F.; Dyurgerov, M.B.; Rick, U.K.; O'Neel, S.; Pfeffer, W.T.; Anderson, R.S.; Anderson, S.P.; Glazovsky, A.F. Glaciers dominate eustatic sea-level rise in the 21st century. *Science* **2007**, *317*, 1064–1067. [\[CrossRef\]](#)
9. Gregory, J.M.; White, N.J.; Church, J.A.; Bierkins, M.F.P.; Box, J.E.; van den Broeke, M.R.; Cogley, J.G.; Fettweis, X.; Hanna, E.; Huybrechts, P.; et al. Twentieth-century global-mean sea level rise: Is the whole greater than the sum of the parts? *J. Clim.* **2013**, *26*, 4476–4499. [\[CrossRef\]](#)
10. Kaplan, M.R.; Schaefer, J.M.; Strelin, J.A.; Denton, G.H.; Anderson, R.F.; Vandergoes, M.J.; Finkel, R.C.; Schwartz, R.; Travis, S.G.; Garcia, J.L.; et al. Patagonian and southern South Atlantic view of Holocene climate. *Quat. Sci. Rev.* **2016**, *141*, 112–125. [\[CrossRef\]](#)
11. Pörtner, H.-O.; Roberts, D.C.; Masson-Delmotte, V.; Zhai, P.; Tignor, M.; Poloczanska, E.; Mintenbeck, K.; Alegria, A.; Nicolai, M.; Okem, A.; et al. (Eds.) IPCC, 2019: Technical Summary. In *IPCC Special Report on the Ocean and Cryosphere in a Changing Climate*; Cambridge University Press: Cambridge, UK, 2019; pp. 39–69. [\[CrossRef\]](#)
12. Armitage, T.W.K.; Manucharyan, G.E.; Petty, A.A.; Kwok, R.; Thompson, A.F. Enhanced eddy activity in the Beaufort Gyre in response to sea ice loss. *Nat. Commun.* **2020**, *11*, 761. [\[CrossRef\]](#) [\[PubMed\]](#)
13. Li, Q.; England, M.H.; Hogg, A.M.; Rintoul, S.R.; Morrison, A.K. Abyssal ocean overturning slowdown and warming driven by Antarctic meltwater. *Nature* **2023**, *615*, 841–847. [\[CrossRef\]](#) [\[PubMed\]](#)
14. Jacobsen, D.; Milner, A.M.; Brown, L.E.; Dangles, O. Biodiversity under threat in glacier-fed river systems. *Nat. Clim. Chang.* **2012**, *2*, 361–364. [\[CrossRef\]](#)
15. Hall, M.H.P.; Fagre, D.B. Modeled Climate-Induced Glacier Change in Glacier National Park, 1850–2100. *BioScience* **2003**, *53*, 131–140. [\[CrossRef\]](#)

16. Carey, M.; Molden, O.C.; Rasmussen, M.B.; Jackson, M.; Nolin, A.W.; Mark, B.G. Impacts of glacier recession and declining meltwater on mountain societies. *Ann. Assoc. Am. Geogr.* **2016**, *107*, 350–359.
17. Dehecq, A.; Gourmelen, N.; Gardner, A.S.; Brun, F.; Goldberg, D.; Nienow, P.W.; Berthier, E.; Vincent, C.; Wagnon, P.; Trouvé, E. Twenty-first century glacier slowdown driven by mass loss in High Mountain Asia. *Nat. Geosci.* **2019**, *12*, 22–27. [\[CrossRef\]](#)
18. Bradley, R.S.; Vuille, M.; Diaz, H.F.; Vergara, W. Threats to water supplies in the tropical Andes. *Science* **2006**, *312*, 1755–1756. [\[CrossRef\]](#)
19. Cullen, N.J.; Mölg, T.; Kaser, G.; Hussein, K.; Steffen, K.; Hardy, D.R. Kilimanjaro glaciers: Recent areal extent from satellite data and new interpretation of observed 20th century retreat rates. *Geophys. Res. Lett.* **2006**, *33*, 675–677. [\[CrossRef\]](#)
20. Bahr, D.B.; Meier, M.F.; Peckham, S.D. The physical basis of glacier volume-area scaling. *J. Geophys. Res.* **1997**, *102*, 20355–20362. [\[CrossRef\]](#)
21. Ke, L.; Ding, X.; Song, C. Heterogeneous changes of glaciers over the western Kunlun Mountains based on ICESat and Landsat-8 derived glacier inventory. *Remote Sens. Environ.* **2015**, *168*, 13–23. [\[CrossRef\]](#)
22. Burns, P.; Nolin, A. Using atmospherically-corrected Landsat imagery to measure glacier area change in the Cordillera Blanca, Peru from 1987 to 2010. *Remote Sens. Environ.* **2014**, *140*, 165–178. [\[CrossRef\]](#)
23. Li, Z.; Shen, H.; Weng, Q.; Zhang, Y.; Dou, P.; Zhang, L. Cloud and cloud shadow detection for optical satellite imagery: Features, algorithms, validation, and prospects. *ISPRS J. Photogramm. Remote Sens.* **2022**, *188*, 89–108. [\[CrossRef\]](#)
24. Paul, F.; Bolch, T.; Kääb, A.; Nagler, T.; Nuth, C.; Scharrer, K.; Shepherd, A.; Strozzi, T.; Ticconi, F.; Bhambri, R.; et al. The glaciers climate change initiative: Methods for creating glacier area, elevation change and velocity products. *Remote Sens. Environ.* **2015**, *162*, 408–426. [\[CrossRef\]](#)
25. Kachouie, N.N.; Huybers, P.; Schwartzman, A. Localization of mountain glacier termini in Landsat multi-spectral images. *Pattern Recognit. Lett.* **2012**, *34*, 94–106. [\[CrossRef\]](#)
26. Kachouie, N.N.; Gerke, T.; Huybers, P.; Schwartzman, A. Nonparametric Regression for Estimation of Spatiotemporal Mountain Glacier Retreat from Satellite Images. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 1135–1149. [\[CrossRef\]](#)
27. Onyejekwe, O.; Holman, B.; Kachouie, N.N. Multivariate models for predicting glacier termini. *Environ. Earth Sci.* **2017**, *76*, 807. [\[CrossRef\]](#)
28. Robbins, E.; Hlaing, T.T.; Webb, J.; Kachouie, N.N. Supervised methods for modeling spatiotemporal glacier variations by quantification of the area and terminus of mountain glaciers using remote sensing. *Algorithms* **2023**, *16*, 486. [\[CrossRef\]](#)
29. Oerlemans, J. Extracting a Climate Signal from 169 Glacier Records. *Science* **2005**, *308*, 675–677. [\[CrossRef\]](#)
30. Cullen, N.J.; Sirguey, P.; Mölg, T.; Kaser, G.; Winkler, M.; Fitzsimons, S.J. A century of ice retreat on Kilimanjaro: The mapping reloaded. *Cryosphere* **2013**, *7*, 419–431. [\[CrossRef\]](#)
31. Bhattacharya, A.; Bolch, T.; Mukherjee, K.; King, O.; Menounos, B.; Kapitsa, V.; Neckel, N.; Yang, W.; Yao, T. High Mountain Asian glacier response to climate revealed by multi-temporal satellite observations since the 1960s. *Nat. Commun.* **2021**, *12*, 4133. [\[CrossRef\]](#)
32. Prieur, C.; Rabatel, A.; Thomas, J.-B.; Farup, I.; Chanussot, J. Machine Learning Approaches to Automatically Detect Glacier Snow Lines on Multi-Spectral Satellite Images. *Remote Sens.* **2022**, *14*, 3868. [\[CrossRef\]](#)
33. Hubel, D.H.; Wiesel, T.N. Receptive fields of single neurones in the cat's striate cortex. *J. Physiol.* **1959**, *148*, 574–591. [\[CrossRef\]](#) [\[PubMed\]](#)
34. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015, Proceedings of the 18th International Conference, Munich, Germany, 5–9 October 2015*; Springer: Cham, Switzerland, 2015. [\[CrossRef\]](#)
35. Shelhamer, E.; Long, J.; Darrell, T. Fully convolutional networks for semantic segmentation. *arXiv* **2014**, arXiv:1411.4038. [\[CrossRef\]](#) [\[PubMed\]](#)
36. Xie, Z.; Haritashya, U.K.; Asari, V.K.; Young, B.W.; Bishop, M.P.; Kargel, J.S. GlacierNet: A deep-learning approach for debris-covered glacier mapping. *IEEE Access* **2020**, *8*, 83495–83510. [\[CrossRef\]](#)
37. Purdie, H.; Anderson, B.; Chinn, T.; Owens, I.; Mackintosh, A.; Lawson, W. Franz Josef and Fox Glaciers, New Zealand: Historic length records. *Glob. Planet. Chang.* **2014**, *121*, 41–52. [\[CrossRef\]](#)
38. Mackintosh, A.N.; Anderson, B.M.; Lorrey, A.M.; Renwick, J.A.; Frei, P.; Dean, S.M. Regional cooling caused recent New Zealand glacier advances in a period of global warming. *Nat. Commun.* **2017**, *8*, 14202. [\[CrossRef\]](#)
39. United States Geological Survey (USGS). *Landsat—Earth Observation Satellites Fact Sheet 2015–3081*; U.S. Geological Survey (USGS): Reston, VA, USA, 2016. [\[CrossRef\]](#)
40. Manaaki Whenua|Landcare Research. LCDB v5.0—Land Cover Database Version 5.0, Mainland, New Zealand. Coordinates. 10 November 2021. Available online: <https://iris.scinfo.org.nz/layer/104400-lcdb-v50-land-cover-database-version-50-mainlandnew-zealand/> (accessed on 21 October 2023).
41. Simonyan, K.; Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. International Conference on Learning Representations. *arXiv* **2015**, arXiv:1409.1556. [\[CrossRef\]](#)
42. Weiss, K.; Khoshgoftar, T.M.; Wang, D.D. A survey of transfer learning. *J. Big Data* **2016**, *3*, 1345–1459. [\[CrossRef\]](#)
43. Tan, C.; Sun, F.; Kong, T.; Zhang, W.; Yang, C.; Liu, C. A Survey on Deep Transfer Learning. In *Proceedings of the 27th International Conference on Artificial Neural Networks, Rhodes, Greece, 4–7 October 2018*; pp. 270–279. [\[CrossRef\]](#)
44. Raup, B.; Racoviteanu, A.; Khalsa, S.J.S.; Helm, C.; Armstrong, R.; Arnaud, Y. The GLIMS geospatial glacier database: A new tool for studying glacier change. *Glob. Planet. Chang.* **2006**, *56*, 101–110. [\[CrossRef\]](#)

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.