VICTORIA UNIVERSITY OF WELLINGTON
Te Whare Wānanga o te Ūpoko o te Ika a Māui

School of Engineering and Computer Science
Te Kura Mātai Pūkaha, Pūrorohiko

PO Box 600
Wellington
New Zealand

Tel: +64 4 463 5341
Fax: +64 4 463 5045
Internet: office@ecs.vuw.ac.nz

Automating Glacier Change Monitoring in the Southern Alps of New Zealand

Jackson Jourdain

Supervisors: Bach Nguyen, Andrew Lensen, Lauren Vargo

Submitted in partial fulfilment of the requirements for
Bachelor of Engineering with Honours.

Abstract

The response of mountain glaciers to annual climate fluctuations is an important indicator of climate change. A subset of 50 glaciers in the Southern Alps of New Zealand is monitored using aerial imagery, to detect and measure the response of glaciers by identifying the glacier outline. The manual identification of the glacier area and subsequent glacier outline is time consuming. This work aims to design an unsupervised system to automate this process accurately. A system was designed that implements and extends upon an existing unsupervised segmentation method with image processing techniques such as contour identification and image morphology. A well-known U-Net segmentation model is extended using image augmentations and transfer learning, then used to detect the glacier boundary. The performance is compared with the designed unsupervised system. The results showed that the designed unsupervised system was able to improve upon the performance of benchmark unsupervised segmentation techniques, and also outperform the best performing U-Net model by 18.2%.

Contents

Chapter 1

Introduction

1.1 Problem statement

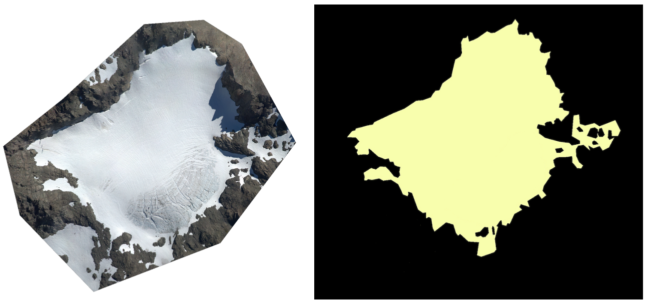
The response of mountain glaciers to annual climate fluctuations is an important indicator of climate change [1]. Mountain glaciers respond to changes in climate within years to decades [2], varying in length and volume in response to changes in temperature and precipitation. Monitoring glacier changes can indicate trends of climate change, and provide insight into the impacts and effects of climate change and glacier melt at global, regional, and local levels [3]. Glacier mass balance measures the net gain or loss of glacial mass over a period of time. Mass balance data can help explain why glaciers are advancing or retreating, and what changes in climate are responsible [4]. More importantly, mass balance shows a direct and undelayed response of glaciers to atmospheric conditions, making annual glacier mass balance a useful signal of glacier and climate changes [5].

1.2 Motivations

In the Southern Alps of New Zealand, there are $\sim$2900 glaciers (as of 2016) [6]. Since 1977, a subset of 50 glaciers has been monitored using aerial imagery to detect and measure the response of glaciers to annual climate fluctuations. At the end of summer (March-April), when snow cover is at a minimum, certain glacial features can be identified in aerial imagery. From these features a glacier mass balance can be estimated [7, 8, 9]. Three useful features are commonly identified. These are the outline of the glacier area, the accumulation and ablation zones, and the end of summer snowline. The snowline is the boundary between the snow-covered accumulation zone, and ice surfaced ablation zone, and when measured at the end of summer, is a proxy for the equilibrium line altitude (ELA) and annual glacier mass balance[10, 11]. Aerial images of New Zealand glaciers are obtained annually, and then manually analysed [12, 13]. The analysis includes identifying the glacier area, and the separation of the accumulation and ablation areas by identifying an ELA. These identifications can provide more insight into annual glacier-climate interactions [14]. Manual processing is usually time-consuming, especially with the large number of images collected for each glacier.

1.3 Goals

The original aim of this project was to implement an unsupervised system that could automate three manual processes done on aerial glacier imagery. Each manual process formed



(a) Rolleston glacier 2017 (b) Segmented glacier area

Figure 1.1: Ortho-rectified image of Rolleston glacier in 2017, and a manually generated mask segmenting the main glacier area.

one objective as part of developing an end-to-end system to be applied to aerial glacier images. These objectives were to: (a) identify the main glacier area, (b) identify accumulation and ablation zones and (c) identify an ELA. While designing a system to complete the first objective, it became apparent that identifying the main glacier area was a much more complicated and challenging task than originally thought. Moving on to the next objective without solving these challenges would have limited the end-to-end system's ability to complete the second and third objectives with accuracy. The challenges were found to be complex and interesting. For this reason, the aim of the project was pivoted to explore potential solutions to these challenges.

The new aim of this project is to design and implement a novel unsupervised image segmentation system that can automatically and accurately identify the area of a glacier in aerial images. The designed system needs to be unsupervised due to the lack of labelled training data. The accuracy of the glacier identification system will be evaluated by comparing predicted segmentation masks with manually generated ground truth masks. The objectives of this project are as follows:

To design an unsupervised segmentation system that can automate the time-consuming manual glacier identification task completed as part of the glacier monitoring process by:

1. Designing and implementing an unsupervised system to identify the main glacier area.

   Identify the region(s) in an image that is defined as glacier area. An example of this identification can be seen in Figure 1.1 (b), where the glacier region is identified as a yellow area, while rock areas and disconnected snow/ice is shown as a black area. This will be achieved by developing a glacier segmentation system that separates the main area(s) with glacial features and areas with non-glacial features. The system must be unsupervised, as there is very little labelled training data available.

2. Implementing and extending an existing supervised image segmentation technique to identify the main glacier area.

   One benchmark, and three extended U-Net image segmentation models will be implemented and applied to the given glacier image dataset to explore the possibility of effective supervised learning on this task. For this task labelled training data must be manually generated. Completion of this objective will give insight into whether supervised learning is a viable option for this task.

3. Comparing the proposed unsupervised system to the supervised image segmentation models, and two other well-known benchmark unsupervised methods.

   The results of the designed unsupervised glacier segmentation system will be compared with the results of the four different variations of supervised U-Net models. Dice score (F1) will be used as a metric to measure the accuracy of the predictions. The comparison of models will clarify which type of model is effective (unsupervised or supervised) and should be further explored in future work.

Chapter 2

Background and Related Work

2.1 Glacier features

Glacier features visible in images such as the main glacier area, accumulation area, ablation area, and ELA can be used to estimate a mass balance [7, 8, 9]. Glacier mass balance measures the net gain or loss of glacial mass over a period of time [4], and can give important indications of changes in climate [5]. The main glacier area is the surface of snow and ice making up the perennial mass of the glacier. In this work, this definition includes significant disjointed areas of snow and ice. More specifically, areas of ice or snow surface that are larger than 2/9 of the size of the largest continuous ice area are defined as part of the glacier. The main glacier area can be divided into two subregions: the accumulation zone and the ablation zone.

The accumulation zone is the area of a glacier where snow that has accumulated during winter does not melt completely during the subsequent summer. The ablation zone is the part of the glacier where summer melting exceeds winter accumulation. The accumulation zone often appears as a smooth white area of snow, while the ablation zone often appears as a more textured light blue-grey area of ice. The end-of-summer snowline, or ELA, is found at the altitude where the net balance of annual snowfall and snow-melt is zero [15]. The ELA is often identified as a visible line separating the accumulation and ablation zones of the glacier. [16]. The position of the ELA can be identified in oblique aerial photos at the end of summer when the glacier snowline has retreated to its maximum altitude [15].

The identification of the main glacier area is a difficult but important challenge. The resulting identifications can be used in the future as part of the process of automatically identifying the accumulation and ablation zones, the ELA, and finally the glacier mass balance. This outlines the importance of the system designed in this work. An accurate system can be used to directly monitor the surface area of the glacier (including outlines), but more importantly, can be implemented as part of a larger system to automatically monitor glacier mass balance.

2.2 Current data collection and manual analysis

In most years since 1977, annual surveys of New Zealand glaciers have been conducted to record annual ELAs. These surveys utilise light aircraft to monitor a selection of 50 glaciers within the South Island of New Zealand. On-board handheld cameras are used to capture oblique photographs from different angles [12]. More recently, sets of oblique aerial photographs have been compiled to produce single high-quality ortho-rectified images. This can be done using many oblique images, geo-referencing data of such images, and featurematching algorithms to generate a 3-D point cloud of the glacier surface. The point cloud is then interpolated to produce digital elevation models and orthophotos of the glacier [17]. The quality of the ortho-rectified image data presents an opportunity for image processing techniques to be applied more accurately on glacier analysis tasks.

The image data available for this project comes from 12 different glaciers. One ortho-rectified image is generated per glacier each year, however the number of images available ranges for each glacier. Some glaciers have images from only two years available, while some have up to seven. For use in this work, 47 total images were available. Ortho-rectified images present glaciers from a bird's eye view, making different images of the same glacier easily comparable. The high resolution of the images allows for small details, textures, and patterns to be more clearly identifiable. Although images are high quality, there is a lack of labelled/masked images because of the time-consuming process needed to produce them. This favours the implementation of an unsupervised system, which can be used on many images in the future. However, to evaluate and compare the unsupervised glacier segmentation system to the U-Net model variations, labelled data was required. Subsequently, labelled data was manually created for 40 images using image labelling sofware. Seven of the 47 images were deemed unfit for use in training a supervised learning model due to defects such as significant discolouration and blank images. This allows for both supervised and unsupervised segmentation models to be applied to this dataset and evaluated.

2.3 Image segmentation overview

Semantic image segmentation is the process of partitioning an image into multiple subregions, where each subregion belongs to a class and shares similar features [18]. In an image segmentation task, a model is used to assign each pixel in an image to a predicted class. The resulting predicted image mask shows the segmented image. The system designed in this work solves a binary pixel classification problem. The first of the classes is glacial surface, and the second is non-glacial surface. In this work, resulting prediction masks consist of 0's for non-glacial surfaces, and 1's for glacial surfaces. This type of mask can be visualised as a white and black image where white sections are glacial surfaces (as seen in further sections). Segmentation techniques can be divided into two main categories: unsupervised and supervised techniques.

2.3.1 Unsupervised image segmentation techniques

In the absence of labelled training data, only unsupervised segmentation techniques can be implemented. Depending on the problem, it can be a challenge to yield accurate results from the use of unsupervised segmentation techniques, as the system has no knowledge of the problem. Most unsupervised image segmentation techniques use colour, brightness, and local patterns to make pixel-level clusters based on these features [19]. Several sub-categories of unsupervised image segmentation exist, such as thresholding, clustering, edge detection, and unsupervised deep learning [20, 19]. Felzenszwalab and Huttenlocher [21] use a simple graph-based automatic thresholding technique considering local neighbourhoods. Shi and Malik [22] use the dissimilarity between groups and similarity within groups as a global criterion for segmenting an image. Dhanachandra [23] uses simple K-means clustering in combination with medial filtering to produce segmentation results. Further categories include unsupervised deep learning, and adaptive thresholding.

One category of unsupervised segmentation techniques is unsupervised deep learning models. These techniques often use an encoder-decoder model to learn to extract a segmentation mask from an image. One example of this is Xia's W-Net model [19]. Alternative deep learning methods have been proposed such as Kanezaki's unsupervised backpropagation model [24]. These techniques are relatively new, so there is little available literature surrounding the performance of these models when applied to different data. In addition, most of these techniques are complex, and would take a long time to implement. Due to this, they are not a viable option technique for this work, considering that it is unknown if these techniques will be able to produce good results.

Adaptive thresholding techniques are commonly used for unsupervised segmentation tasks. Adaptive thresholding techniques are easy to implement, and often have very few parameters, a reason for their use in this work. Two main categories of adaptive thresholding can be derived. The first is local adaptive thresholding, in which unique thresholds are generated based on pixel information from local neighbouring pixels. The second is global adaptive thresholding, in which pixel information from the whole image is considered when generating a threshold for the whole image. Some local adaptive thresholding techniques include Liu and Jezek's [25], Yu's integration of canny edge detection with local adaptive thresholding [26], and Yan's multi-stage local adaptive thresholding [27]. Both Liu and Jezek's, and Yu's methods are applied to antarctic coastline segmentation problems – relevant to the glacier domain this work is concerned with. Two relevant global adaptive thresholding techniques include Otsu's method [28], and Sezan's peak detection algorithm [29]. The adaptive thresholding technique most important to this work is Otsu's method.

Otsu's method is a global adaptive thresholding technique that finds an optimal pixel threshold that maximises between class variance. The method is motivated by the conjecture that well-thresholded classes would be separated in grey levels, and therefore, a threshold resulting in the best separation of classes in grey levels would be the best threshold. The method adopts a criterion of the between class variance, which measures the 'goodness' of the threshold. A sequential search is conducted to find the threshold setting that maximises this criterion. The method is non-parametric, can be extended for multi-threshold problems, and automatically and stably selects threshold settings. These advantages make it useful for a wide range of applications, including as part of a solution to the objectives of this work.

2.3.2 Supervised image segmentation techniques

A well-known supervised segmentation technique called UNet [30] is a popular deep learning method that has sparked major developments in the field of supervised image segmentation since its design in 2015. The UNet model, as well as state-of-the-art extensions of itself such as Res-UNet [31] has been shown to achieve extremely accurate segmentation results, even with small datasets [32]. However, these models are often prone to overfitting when applied to such datasets [33, 34]. Originally designed for biomedical imaging, the UNet and Res-UNet models have been successfully applied to a wide variety of domains, including ones with relevance to the domain of this work. Recent research has been done applying UNet [35] and Res-UNet [36] models to relevant glacial segmentation problems, showing that such a model could be effective when used for the objectives of this work. The work of [35] analyses calving fronts of glacial sheets in Antarctica using UNet models. This is different from a mountain glacier domain, however, there are some shared features between domains such as the colours and textures of ice and snow. The work of [36] applies a Res-UNet to identify calving fronts in Greenland mountain glaciers. Calving is the process of ice breaking off the end of the glacier.

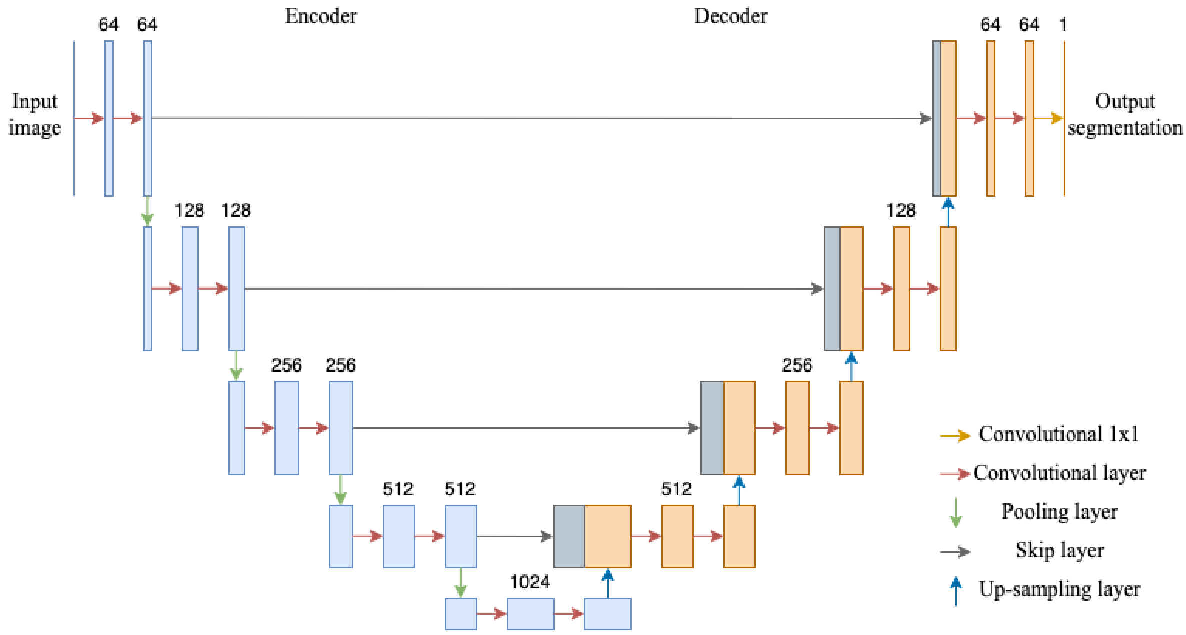


Figure 2.1: U-Net architecture.

UNet model

The UNet model is a deep learning model where the only trainable parameters are within convolutional layers. U-Net consists of an encoder (that extracts features from the input images) and a decoder (that concatenates past feature maps from the skip layer). The encoder contains the first half of the architecture, consisting of only convolutional and pooling layers. The decoder contains the second half of the architecture, which consists of convolutional and up-sampling layers. Skip-layers connect different parts of the encoder and decoder. The encoder extracts feature maps from the input images, and uses pooling layers to downsize these feature maps to reduce spatial information. The decoder enlarges the feature maps and then concatenates past feature maps from 'skip' layers to restore some spatial information. The output of the UNet can be used to create a predicted segmentation mask. Figure 2.1 illustrates the structure of a standard UNet model.

One commonly used method that can improve the accuracy of UNet models on small datasets is image augmentation [30]. This is the process of enriching the training data by manipulating images differently in every training epoch. This has the effect of artificially enlarging the dataset, to provide a model with more training data. A model can be made more generalisable, and less prone to overfitting. Due to the small size of the glacier image dataset available in this project, image augmentation is likely to be important if a U-Net model is going to perform well.

Another method that can improve the accuracy of UNet models on small datasets is transfer learning. A technique designed by Iglovikov and Shvets called TernausNet [37], uses transfer learning to improve a U-Net model. A U-Net model was trained on the ImageNet [38] data set. This network was kept and then applied to a new data set where only the decoder portion of the network was re-trained. The advantage of this is that the network has already learned what features to extract from the images, and only needs to learn how to use these features to achieve accurate segmentations on the new data.

2.4 Summary of existing work

While image segmentation techniques have been applied to many types of problems, there is little work involving the identification of glacier areas. Although there is existing research applying supervised segmentation models to different kinds of glacier related problems, none were intended for the identification of area. Considering the lack of labelled training data available in this project, existing unsupervised segmentation techniques are an important area of work informing the approaches taken to design an unsupervised system. Experimenting with existing unsupervised techniques such as adaptive thresholding formed a large part of the design process, and such techniques contribute heavily to the final designed system. Past work involving supervised techniques such as UNet is important to the objective of this project. The implementation and extension of existing supervised models are important to compare with the designed unsupervised system. The results of this comparison will likely reinforce the importance of a system being unsupervised, or show the potential of supervised learning to be used in this application. Overall, the use of existing work to inform the design of a new system that is specific to the domain of mountain glacier identification forms the main contribution of this project.

Chapter 3

Design

This chapter outlines the process which was taken to narrow down potential solutions and come up with a final design to address the objective of identifying the main glacier area. Several techniques were tested, however, each had its advantages and disadvantages when solving the problem at hand. From this process, the difficulties of the problem were identified. The result of this process was the designed glacier identification system, which was better at identifying the main glacier area than other techniques.

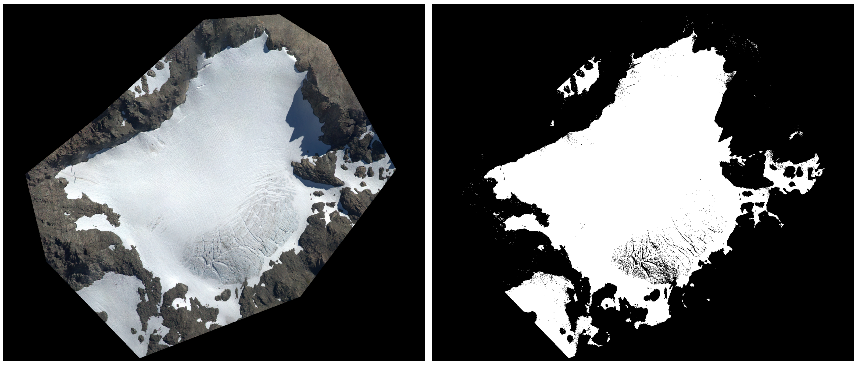
3.1 Design Criterion

The criterion for successful identification of the main glacier area is to automatically and accurately segment the main surface area of a glacier mass from a background in a variety of aerial orthophotos. The system should take an input image and produce a binary segmentation mask, clearly presenting the outlines of the main glacier surface. As well as the accurate outline of the glacier, significant non-glacier features within the glacier outline must be visible in the mask. The system must be able to repeatedly segment an image, and reproduce the same results. The system must be robust enough to produce accurate results on a variety of comparable glacier images, including different lighting conditions, image quality, image size, glacier shape, glacier features, glacier colour and rock colour.

3.2 Unsupervised System Design Process

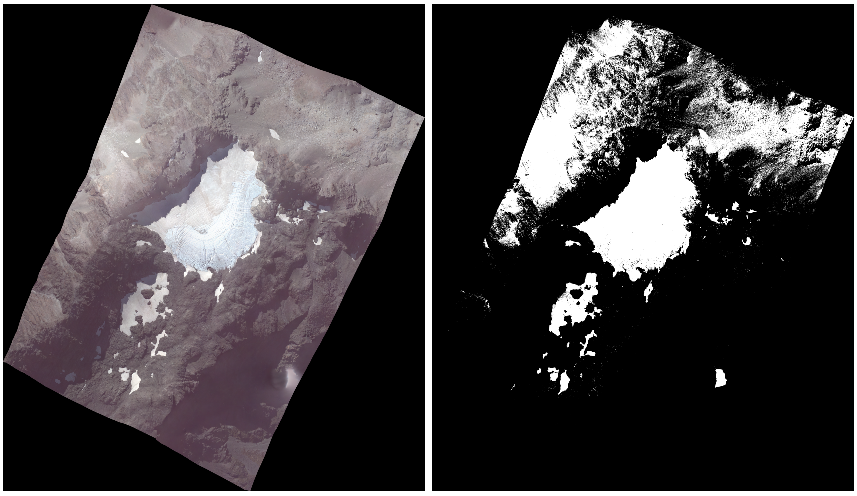
3.2.1 Manually set global thresholding

Firstly, a technique of manually set global thresholding was used. Being the simplest and least dynamic of unsupervised segmentation techniques, this would provide a valid baseline to compare against when designing and testing more advanced techniques. This basic technique was able to accurately segment many glacier areas across a range of images, however, was very limited. The first limitation was that not one threshold setting fits all images. Different images have different brightness levels, so a threshold that works well on one image may not work well on another. The second limitation identified is that only one threshold is used for segmenting all regions in an image. Fluctuations in the brightness of rock and snow areas across areas of an image led to inconsistent segmentation results. The third limitation identified was that no image features aside from each pixel's grey level are used. Figure 3.1 visualises both a good and poor segmentation result from using manually set global thresholding. Other image features such as colour and patterns have the potential to be used to improve segmentation results. Different unsupervised segmentation techniques such as global adaptive thresholding, and local adaptive thresholding, were implemented as a more dynamic form of thresholding.



(a) Rolleston Glacier 2017

(b) Rolleston Glacier 2017 segmented mask



(c) Rolleston Glacier 2020

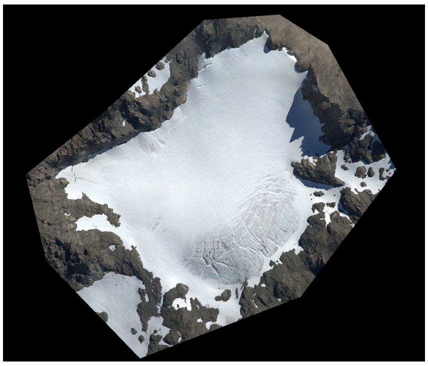
(d) Rolleston Glacier 2020 segmented mask

Figure 3.1: Examples of segmentation results using manually set global thresholding. (b) is an example of a good segmentation result, where only the shadow-covered glacier area was incorrectly segmented. (d) is an example of a poor segmentation result where rock areas were bright enough to be segmented as glacier area.

3.2.2 Global adaptive thresholding

Histogram based

The first global adaptive thresholding technique implemented was a designed histogram based pixel thresholding system. This system used a technique inspired by Sezan's peak detection algorithm [29], because it shows that analysing a pixel intensity histogram can be effective for choosing a threshold value. A grey level frequency histogram was generated for each image and analysed to select an optimal thresholding setting. Figure 3.2 visualises a pixel frequency histogram and selected threshold for Rolleston glacier 2017. A one dimensional Guassian blur was applied to the frequency values of each pixel intensity in the histogram. This would reduce the number of small peaks and troughs in the histogram, allowing for the identification of the main peaks and troughs. A threshold was selected which



(a) Rolleston Glacier 2017

(b) Pixel frequency histogram with threshold

Figure 3.2: Rolleston glacier 2017, and a corresponding pixel frequency histogram with a selected threshold value.

is located in the trough between the two largest peaks. This system showed an improvement over manually set global thresholding, however, was not robust enough to work flawlessly on enough images in the dataset. For this reason, Otsu's method, another global adaptive thresholding method was implemented. This is because it has been shown to produce good results on a variety of problems, depending on certain image parameters such as intensity level between classes and relative class sizes [39].

Otsu based

Otsu's global adaptive thresholding technique was implemented and applied to greyscale glacier imagery. The results of this technique showed significant improvements over manually set global thresholding. Improvements were primarily due to Otsu's method being able to select a more optimal threshold setting for each individual image. This makes the technique more dynamic across a whole dataset, giving more accurate segmentations on more images. This is unlike manually set thresholding, which can work well on many images, however, does not work well on images where grey levels in one or both classes stray too far from the mean levels. In these cases, manually set global thresholding will incorrectly segment one of the classes, resulting in extremely poor segmentations. This outcome is reduced significantly with the use of Otsu's method. Two limitations of Otsu's thresholding were identified. The first is that only a singular greyscale channel is used, potentially missing out on valuable information available in a three colour channel image. The second is that Otsu's method selects a threshold value by which the whole image will be thresholded, using an optimised between class variance value based on information from the whole image. This is a limitation, as there is often variation across different regions of an image. The threshold value might work for one region of the image, but produce inaccurate segmentation results across a whole image. Two further techniques were explored to try to improve Otsu's method by mitigating these limitations. These were local adaptive thresholding and multi-channel thresholding.

3.2.3 Local adaptive thresholding

Local adaptive thresholding techniques were implemented to account for regional variation within aerial glacier images. In local adaptive thresholding, unique thresholds are generated based on pixel information from local neighbouring pixels. Generally, a threshold is calculated for each pixel based on its surroundings, and that pixel is classified using the unique threshold. A basic local adaptive thresholding algorithm was implemented. Otsu's thresholding is applied to the pixels in a local neighbourhood to select a threshold value and classify each pixel in the image. This way, more optimal thresholds can be selected for every varying subregion of the image. Different neighbourhood sizes were experimented with, which produced interesting results. Large neighbourhoods (1/4 the area of the image) were not found to greatly improve the resulting segmentations over Otsu's method. Small neighbourhoods (areas between 5x5, and 50x50 pixels) performed significantly worse than Otsu's method. This was because for each local neighbourhood the algorithm will always find a threshold. This is fine in most circumstances, as there is a mix of pixels from each class. However, when the local neighbourhood contains all or almost all pixels from a single class, a detrimental effect on the segmentation result is seen. In these situations a threshold is still selected, usually resulting in many pixels in the region being incorrectly classified.

To mitigate this limitation, an extension to this local thresholding algorithm was designed. The extension used a histogram analysis to first identify whether a threshold should be created or not. If so, Otsu's thresholding was used. If not, the pixel is classified based on a pre-determined global threshold found through Otsu's method. This algorithm showed promise, but unfortunately was not robust and was unable to outperform either the normal local thresholding algorithm or Otsu's global adaptive thresholding. This was because the algorithm was unable to decide whether a local threshold should be selected or not correctly and consistently.

3.2.4 Proposed multi-channel unsupervised segmentation

A multi-channel thresholding system was designed to take advantage of information across each of the colour channels. The system splits an image into its three colour channels and computes a threshold for each channel. The three resulting segmentation masks would then be combined to form a final segmentation mask. The goal of this system was to take advantage of any extra information stored in individual colour channels, as opposed to a singular grey channel. Through basic analysis of the dataset, ice areas were often found to possess a proportionally high blue value, and rock areas a proportionally high red value. Subsequently, the idea of this system aimed to exploit this pattern to obtain better segmentation results.

Histogram based

The system first applied the previously designed histogram based pixel thresholding algorithm to each channel. The full system worked well on many images but came with certain limitations which were visible in the results on some images. One limitation was that the system could still not correctly segment shadow-covered glacier areas and light-reflective rock areas. Another limitation was that the system was not robust. Certain images would break the histogram based automatic threshold setting system, resulting in extremely poor choices of threshold values. To fix this issue, it was hypothesised that Otsu's thresholding method could be used in place of the histogram based automatic pixel thresholding method. This change would potentially increase the robustness of the system across more images increasing overall performance.

Otsu based

Otsu's thresholding was implemented on each colour channel of an image, instead of the previously used histogram based method. This made the system more robust. To further enhance this design, and fully take advantage of multi-channel thresholding, bias values were added to the threshold values for each colour channel. It was found that the blue channel was able to achieve more accurate segmentation results. This result is likely due to the fact that ice areas were found to often possess a proportionally higher blue value than other areas. Although common, this trend was not found in every image, and in some cases was found to be the opposite. The system was able to mitigate some limitations of other systems. The system was able to consider colour information through the use of individual colour channel thresholding. The system could also work well across a wide range of images because different threshold values are automatically found for each image. However, one main limitation still existed. The limitation was that shadow-covered glacier and light-reflective rock areas were often incorrectly segmented. In an attempt to mitigate this limitation, further techniques were designed. Further techniques were experimented with to try to enable the correct identification of shadow-covered glacier areas and light-reflective rock areas. The experiments included using canny edge detection to try and detect the boundary of the glacier area to assist the thresholding system with identification. However, a lack of progress with unsupervised techniques led to focusing on the refinement of the multi-channel Otsu's thresholding system.

3.2.5 System enhancements

Multiple techniques were explored to enhance the multi-channel Otsu's thresholding system. Some techniques were implemented within the system, while others were applied to the system output. By experimenting with noise reduction techniques, contour finding, and image morphology, a final system was composed. This system was able to produce accurate segmentations, with smooth boundaries. The system could also remove small disconnected areas classified as glacier area from being included in the final identified main glacier area. Unfortunately, the system was still unable to correctly segment shadow-covered glacier areas and light-reflective rock areas consistently. These system enhancements are covered in detail in the system implementation.

3.3 Supervised System Design Process

As part of the objectives of this project, a supervised learning model was to be implemented to compare against the designed unsupervised system. To use such a model, labelled training data would be required. For this reason, ground truth masks were generated for the 40 image dataset. Because it is known to perform well, especially on smaller datasets, a UNet model was implemented and applied to the dataset. The UNet model was tested with different architectures, different image preprocessing strategies, different optimisers, and different parameters. Through different experiments, it was found that the results of the UNet model were not very accurate. In most test cases, the glacier area could definitely be seen, but the segmentations were messy and sections of the glacier would often be missing. It was hypothesised that the small dataset was the factor limiting the performance of the UNet model. For this reason, extensions to the UNet model were explored.

Image augmentation was applied to the baseline UNet model to enrich the training data to improve accuracy and reduce the potential of overfitting. Different types of augmentation were experimented with, ranging from simple augmentations to more extreme augmentations which would change the image more drastically. Interestingly, little success was found in applying image augmentations to the model. This led to the exploration of a transfer learning technique to apply to the UNet model. A technique designed by Iglovikov and

Shvets called TernausNet [37] was used to extend the baseline model. This technique was also used in combination with image augmentation.

Chapter 4

Implementation

In this section, the proposed system for identifying the main glacier area is outlined in detail. The system aims to segment the main area with glacial features from areas with non-glacial features. In other words, it aims to separate the regions of snow and ice which form the main glacier surface from the areas of rock, water, and other non-glacial features. Non-glacial features can also include small sections of snow and ice which are deemed not large enough to be not part of the main glacier mass. The glacier segmentation system implements and extends Otsu's automatic pixel thresholding through the design of a multi-channel thresholding system. A contour finding technique is implemented within the multi-channel thresholding system to narrow down the identification of snow and ice areas to the main glacier area. The system implements a technique that reduces noise and manipulates the output mask to further enhance the results of multi-channel thresholding. This section will also detail the implementation of four supervised U-Net model variations, to be used to compare against the proposed system.

4.1 Unsupervised thresholding system

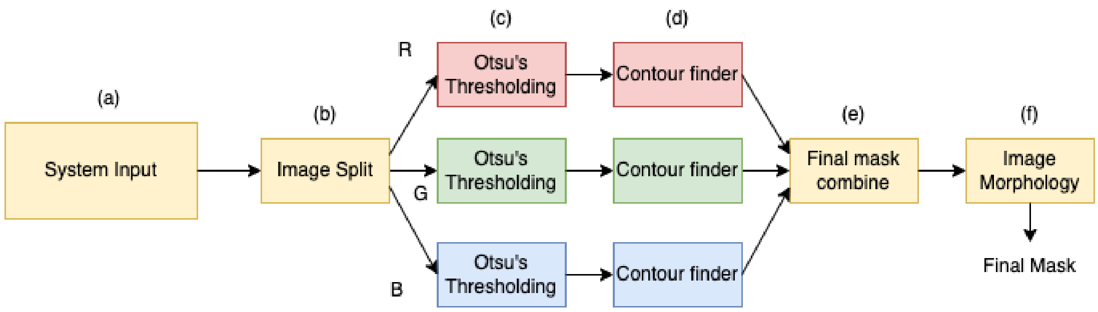


Figure 4.1: High level system overview.

The proposed unsupervised system has four main phases. A high level overview of the system architecture is shown in Figure 4.1. These are individual channel pixel thresholding, contour identification, final mask combining, and image morphology. The individual channel pixel thresholding phase can be further broken down into input, splitting, and thresholding steps. The system works by taking an input aerial glacier image from a user, processing the image through each phase of the system, and outputting a binary segmentation mask. The following sections will detail each phase of the glacier segmentation system.

4.1.1 Individual channel pixel thresholding

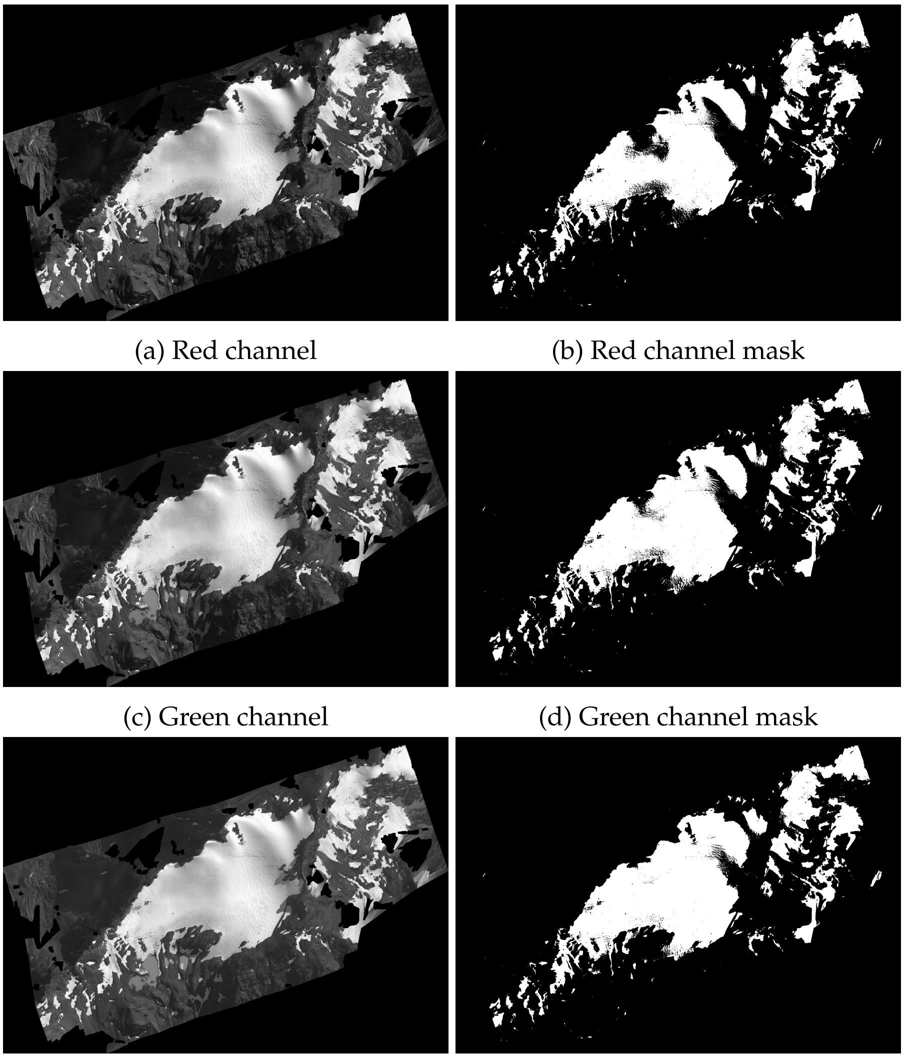
The channel thresholding system implements and extends Otsu's automatic thresholding to segment each of the three colour channels of an image. Before thresholding, an input image is read, and processed by splitting up each colour channel. Otsu's thresholding identifies a threshold that maximises between class variance usually on a single channel greyscale image. Otsu's thresholding is altered to directly deal with RGB images by treating each individual colour channel as a greyscale channel. A threshold is found for each channel while ignoring pixels with a value of 0. This is because the glacier orthophotos often contain a large black border area, and this area does not need to be considered when selecting a threshold. The three threshold's are then used on their corresponding colour channel to create a segmentation mask of the glacier area. The results from this system phase for each colour channel are visualised in Figure 4.2. The raw segmentation masks found for each colour channel are displayed. It should be noted that the blue channel mask seems to produce the most accurate segmented glacier area. The green and red channels still produce a clear mask of the glacier, however, are less accurate in some areas, especially where there are shadows on the glacier.

4.1.2 Threshold bias values

One of the original reasons for implementing a multi-channel thresholding system was to take advantage of the information accessible in multiple colour channels, instead of a singular grey channel. With this system set up, the influence each colour channel has on the final segmentation can be manually tuned. While experimenting with different configurations of the system, it is found that heavily biasing the blue colour channel is yielding better segmentation mask results. This is likely because glacier areas, particularly areas of ice, are commonly found to have a proportionally higher blue value than other areas of an image. Due to this finding, bias values are added to the threshold values found using Otsu's method in the first phase of the system. The bias values raised the threshold values of the red and green colour channels significantly, causing them to be less sensitive. The bias value for the blue channel lowered the threshold, causing it to be more sensitive to darker areas. This resulted in the blue colour channel being emphasised. The final bias values used were 20, 20, and -10, for the red, green, and blue colour channels, respectively. Essentially, if a blue channel threshold was set at a pixel intensity of 150, any pixel with a value below that is classified as non-glacial surface, while every pixel with a value above that is classified as glacial surface. When a bias of -10 is added to the threshold value, the threshold becomes more lenient on slightly darker pixels. This results in more pixels being classified as glacial surface in the blue channel. When bias values of 20 are added to the green and red channel thresholds, they become far stricter. This results in significantly fewer pixels being classified as glacial surface. Although using stricter thresholds, the green and red channels can often correctly identify areas that the blue channel can't. This result is mainly seen in bright rock areas, where pixels are sometimes found to have a proportionally higher red value. Ultimately, these bias settings result in the blue channel being the main contributor to the final segmentation results, while the green and red channels occasionally pick up areas that the blue channel doesn't.

4.1.3 Contour identification

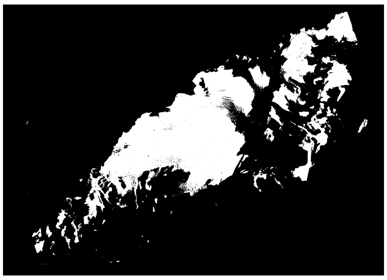
The contour finding system aims to find the largest joined area(s) of a glacier, as well as any significant non-glacier regions enclosed within the identified area(s). Contours are defined as a curve connecting all the continuous points along a boundary having the same



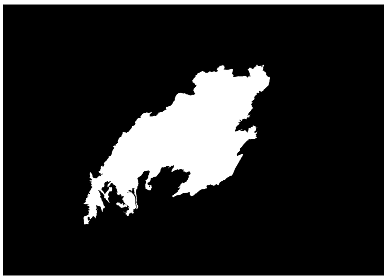
(a) Red channel
(b) Red channel mask
(c) Green channel
(d) Green channel mask
(e) Blue channel
(f) Blue channel mask

Figure 4.2: Before and after Otsu's thresholding is applied on each colour channel (Ridge 2017).

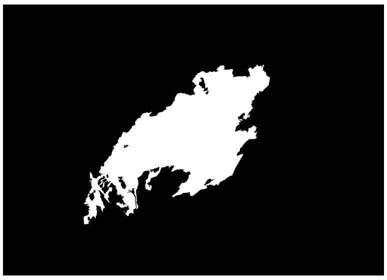
colour/intensity. This continuous curve encloses a section of glacial surface or non-glacial surface. For each colour channel mask generated in the previous system phase, all the contours in the image are found. The total area enclosed by each contour can be calculated so that the largest two contours enclosing glacial surface can be found. In the designed system, a threshold was used to define if a secondary glacier area is large enough to be considered a main part of the glacier area. This threshold was defined as: the second-largest contour needs to have a larger area than 2/9 of the area of the largest contour. This value was chosen by analysing the imagery to identify a common trend and was able to produce good results. If the second-largest contour does not meet the criteria, it is not considered part of the main glacier area. Once the main glacier contours are identified, any second order contours are identified. These are the contours that are fully enclosed within the main glacier area. These contours are identified as non-glacier areas. Only second order contours are identified, so that contours within these (greater than second order) are ignored. This technique is able to narrow down the identified snow and ice areas to one or two regions of the main glacier surface area. The output of this system phase is visualised in Figure 4.3.



(a) Blue channel mask



(b) Largest contour in blue channel mask



(c) Largest contour with identified sub-contours in blue channel mask

Figure 4.3: Contour identification process of blue channel mask (Ridge 2017).

4.1.4 Final mask combining

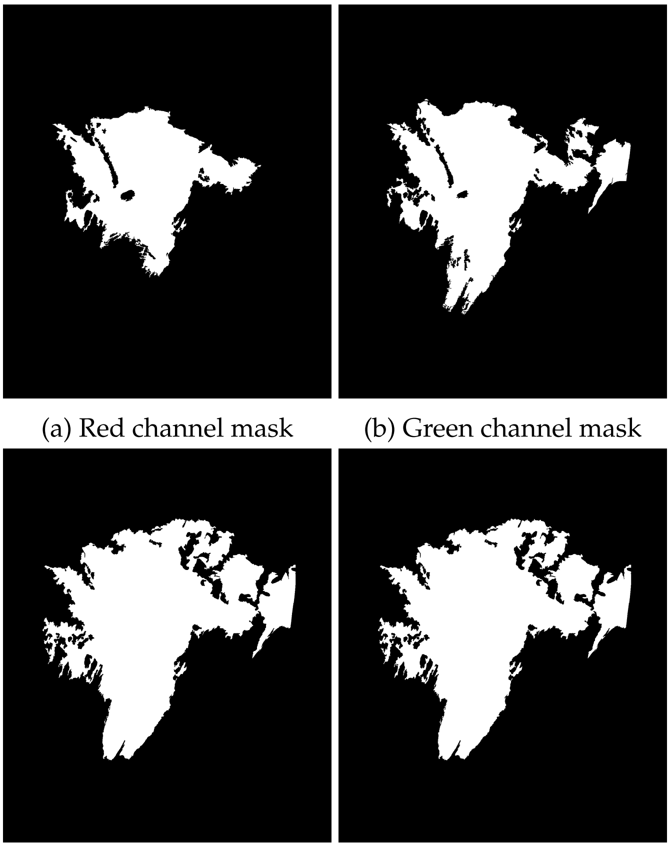
The final mask combining phase used a bitwise-or function to combine the three mask layers into one. Wherever there is a pixel value of 1 in a sub-channel mask, there will be a pixel value of 1 in the final mask. Therefore, glacier areas found in each colour channel are never reduced. It should be noted that the blue mask provides the majority of the area in the final mask due to the thresholding bias values applied in the thresholding phase. Figure 4.4 shows the segmentation mask from each colour channel, and the combined segmentation mask.

4.1.5 Image morphology

An image morphology technique is used on the final combined mask to increase the smoothness of edges, giving a more refined segmentation mask. The technique used is dilation. Dilation works by convolving the glacier mask with a kernel of a defined shape and size. A local maximum value is computed for each placement of the kernel, and the pixel corresponding to the centre of the kernel is set to this value. Therefore, every time the kernel is placed and a pixel of value 1 is present in the bounds of the kernel, the centre pixel will be set to 1. This process results in the white areas of the image mask expanding with each iteration of dilation. Small areas of the non-glacier area are erased, and the edges between regions are smoothed. The output of this system phase is visualised in Figure 4.5.

4.2 Extended supervised U-Net models

A supervised learning method is implemented to compare with the performance of the proposed glacier identification system. This is done to evaluate the possibility that supervised learning is a viable option to identify glacier areas, and potentially outperform the designed unsupervised system. The method implemented was a fully convolutional neural network for image segmentation called a UNet. Four different variations of UNet model were implemented in an attempt to mitigate the weaknesses of the UNet model and obtain the best performance possible. Each variation of UNet model uses the same standard network architecture outlined in Section 2.4. Firstly, a UNet model was implemented using PyTorch, based on the implementation in [40]. Secondly, image augmentations are added to the data loader for the baseline UNet model to try to enrich the small dataset and improve accuracy. This setup means that image augmentations are applied to each batch of training data during training. This ensures that there is always variety in the training data even when



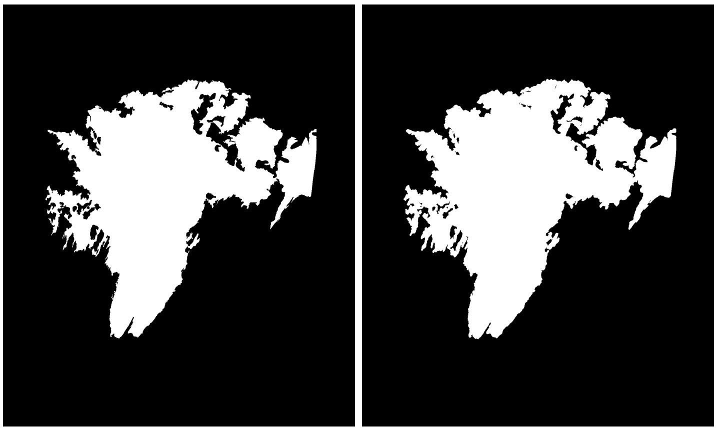
(a) Red channel mask (b) Green channel mask

(c) Blue channel mask (d) Combined mask

Figure 4.4: Each colour channel mask before combining phase, and combined mask (Brewster 2017).

training for a larger number of epochs. Image augmentations were set up sequentially so that images have the possibility of being augmented with multiple techniques. The image augmentation details can be found in Table 4.1. Thirdly, a model of UNet using transfer learning is implemented. An implementation by Iglovikov and Shvets used in [37] called TernausNet is adapted to the framework already built in the baseline model. Lastly, the TernausNet transfer learning model is extended by adding image augmentations to the data loader. This resulted in four different UNet models: a baseline UNet, a baseline UNet with augmentations, TernausNet, and TernausNet with augmentations.

The data preprocessing and parameter settings are detailed in Table 4.2. For the UNet models to train properly, each training image needs to have identical dimensions. The orthophotos in the dataset were a wide range of sizes. Each image was resized to dimensions of 512x512 pixels using bilinear interpolation. Each training run used a training set of 32 images, and a validation set of 8 images was used. A batch size of 8 was used to ensure whole batches can be used for both the training and validation sets. A loss function of cross entropy loss + dice loss was used. This loss function can take into account the loss of the raw network output, as well as the loss of the one-hot encoded (actual predicted) output. A stopping criterion was implemented to prevent the model from over training, since that is likely to be a possibility due to the small dataset. If the validation accuracy has not improved for 5 epochs, or training has reached 50 epochs, training is terminated and the model with the highest validation accuracy is saved.



(a) Before image morphology (b) Final mask after image morphology

(c) Before image morphology (close up area) (d) Final mask after image morphology (close up area)

Figure 4.5: Before and after image morphology with close up examples (Brewster 2017).

4.3 Experiment design

The manually generated ground truth masks were used to evaluate the accuracy of the unsupervised glacier identification system, and to train and evaluate the U-Net models. A baseline Otsu's grayscale thresholding method and a global grayscale thresholding method were tested to compare against the proposed unsupervised method. In the global thresholding method, a threshold of 150 was permanently set. Dice score (F1) was used as a metric to measure the accuracy of predicted segmentation masks against the ground truth masks. 1 indicates a perfect segmentation result, while 0 indicates the worst possible segmentation result.

$$Dicescore = \frac{2 \cdot TP}{2 \cdot TP + FP + FN}$$

To train and evaluate the U-Net models, 10-fold cross validation was implemented. The training was halted once validation accuracy reached a plateau. The peak validation accuracy was taken from each of the 10 runs, and an average validation accuracy was calculated.

Augmentation	Parameters
Vertical flip	Probability = 0.5
Horizontal flip	Probability = 0.5
Rotate	Probability = 0.5 Rotation angle 0-90 degrees Padding value = (0, 0, 0)
Random resized crop	Probability = 0.3 Scale = 0.7-0.9 Ratio = 0.6-1.4
Adjust brightness	Change between -0.5 and -0.3, or 0.3 and 0.5. -0.5 causes a decrease in brightness by 50%, 0.5 causes an increase in brightness by 50%. These settings were chosen to apply a mild to moderate change in brightness which would be able to effectively emulate the lighting conditions found across images in the dataset.

Table 4.1: Image augmentation settings.

Parameter	Setting
Image scale	512x512
Training percentage	0.8 (32 images)
Validation percentage	0.2 (8 images)
Batch size	8
Optimiser	Adam
Loss function	Cross entropy loss + dice loss (F1 loss)
Stopping criterion	Plateau in validation accuracy for 5 epochs, or reached 50 epoch limit

Table 4.2: U-Net parameter settings.

Chapter 5

Results and Evaluation

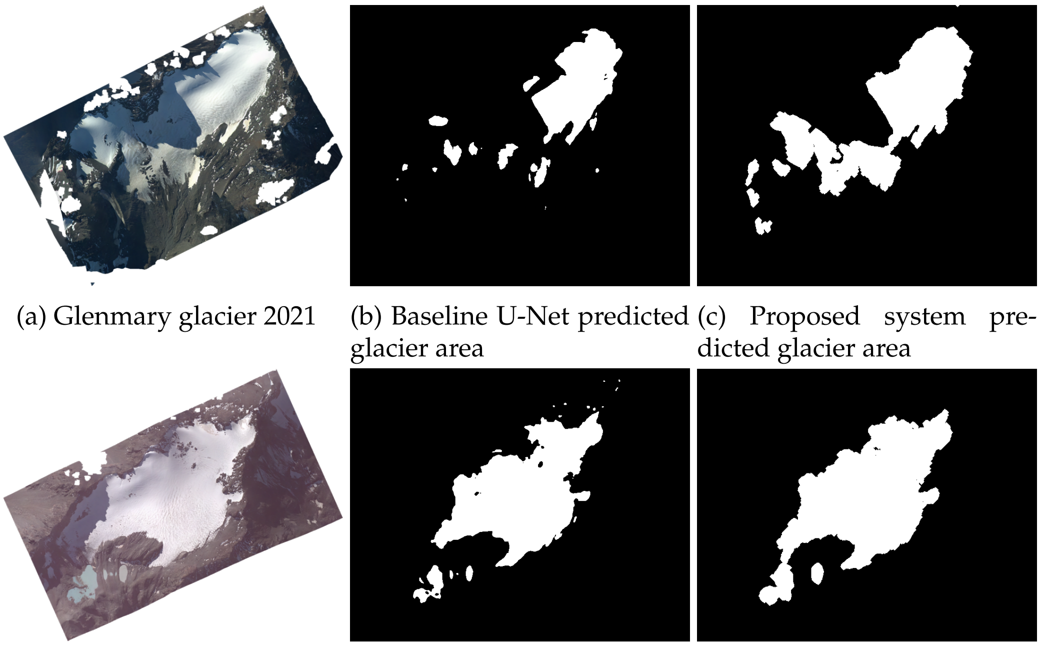
This section will present, analyse, and discuss the results obtained from the proposed system, as well the baseline techniques, and the supervised UNet models. Limitations of the proposed system are discussed, and future avenues of improvement are considered. An evaluation will be made as to how well the project objectives were completed.

5.1 Summary of model results

Model	**Dice score (F1)**
Glacier identification system (proposed method)	0.849 (test)
Benchmark Otsu's thresholding	0.802 (test)
Benchmark manually set global thresholding	0.772 (test)
Baseline UNet	0.718 (validation)
Baseline UNet with augmentation	0.694 (validation)
TernausNet	0.714 (validation)
TernausNet with augmentation	0.686 (validation)

Table 5.1: Test results of glacier identification system, benchmark unsupervised methods, average validation accuracies of U-Net model variations.

5.2 Visual results



(a) Glenmary glacier 2021

(b) Baseline U-Net predicted glacier area

(c) Proposed system predicted glacier area

(d) Ridge glacier 2021

(e) Baseline U-Net predicted glacier area

(f) Proposed system predicted glacier area

Figure 5.1: Visualised prediction results from baseline supervised U-Net model and proposed unsupervised system on Glenmary and Ridge glaciers.

5.3 Designed system results

The designed system was able to achieve a Dice score of 0.849 when tested on the entire available dataset. Considering the nature and challenges of the problem, this is an unexpectedly high performance for an unsupervised model. Areas of glacier with no shadow coverage are easily segmented correctly due to the stark difference between white ice/snow and rock areas. By analysing some segmentation masks produced by the system, it becomes obvious that the system struggles to correctly segment certain areas. This is what limits the system from achieving more accurate results. In the case of some images, rock areas are lit up intensely by the sun. This results in these areas having pixel values greater than the selected thresholds, and being incorrectly classified as glacial surfaces.

5.4 Designed system vs baseline unsupervised methods

As previously mentioned, the designed system was able to achieve a Dice score of 0.849 when tested on the dataset. This is a large increase over the two benchmark unsupervised models which were tested. Otsu's method was able to achieve a Dice score of 0.802 which is 0.047 less than that of the designed system. This equates to a 5.9% increase in Dice score when using the designed system over Otsu's method. The manually set global thresholding method was able to achieve a Dice score of 0.772 which is 0.077 less than that of the designed system. This equates to a 10% increase in Dice score when using the designed system over manually set global thresholding. The results show that the designed system significantly outperforms the two existing baseline unsupervised segmentation techniques. This is due to the designed system being specifically designed to identify glacier regions. The use of domain knowledge to apply multi-channel thresholding, thresholding bias values, and implement contour finding techniques effectively gives the designed system advantages over existing methods.

5.5 Designed system vs UNet models

The results show that the accuracies of the four U-Net variations were poor when compared to the unsupervised glacier segmentation system. This indicates that the dataset is most likely too small to be used effectively for supervised learning with U-Net models. This was a factor that had already been considered, hence the implementation of image augmentation, transfer learning, and the combination of both. These techniques were implemented with the goal of mitigating the undesired effects of training on a small dataset such as overfitting. However, mitigation techniques did not prove successful in this application. The use of transfer learning (TernausNet) decreased the accuracy of the baseline U-Net model from a Dice score of 0.718 to 0.714, a 0.5% decrease. Although only a small decrease, this is a surprising result as the TernausNet model has been shown to outperform regular U-Net models in many cases [37]. Image augmentation also had a negative effect on the U-Net models. This effect is commonly seen in many other applications [41, 42]. Models applied to different datasets respond with varying results when image augmentations are used. Those using image augmentation showed a reduction in accuracy when compared with their corresponding models which did not use augmentation. The baseline U-Net model with image augmentations achieved a Dice score of 0.694, a 0.024 (3.3%) decrease from the baseline U-net model which scored 0.718. The TernausNet model with image augmentations achieved a Dice score of 0.686, a decrease of 0.028 (3.9%) from the TernausNet model which scored 0.714.

The designed system outperformed the best performing U-Net model, the baseline U-Net, by 0.131 Dice score. This is an increase in Dice score by 18.2%. Both of the baseline thresholding methods implemented to compare with the designed system also outperformed the best U-Net model. This conveys that unsupervised methods such as the designed system are much more suited to this problem. Achieving high performance with no training data is important for this system to be implemented within the glacier analysis process in future years. With such a small dataset, supervised learning proved ineffective. The accuracy obtained by the U-Net models would not suffice for use in glacier analysis. More training data would likely improve the performance of a U-Net. Figure 5.1 visually compares the predictions of the best performing U-Net model and the designed system.

5.6 Limitations of the designed system

An analysis of the masks obtained by the designed system revealed that certain glacial features were difficult to segment correctly, due to a limitation of the system. These glacier features were shadow-covered sections and light-reflective rock areas of a glacier. Some shadow-covered areas appear very dark and are classified as non-glacier. Some light-reflective rock areas appear very bright and are classified as glacier. The effects of shadow-covered glacier areas on predicted results can be seen in Figure 5.1. Incorrectly classification of light-reflective rock areas only significantly affected three of the 40 test images, obtaining a Dice score of below 0.67. If this limitation can be mitigated, the system could be improved to have an average Dice score of $\sim$0.87. Incorrect classification of shadow-covered glacier areas affects the majority of images in some way. In most cases, these effects are small and do not have a large impact on the accuracy of the final segmentation. However, in some cases, the effect of this limitation is significant, such as the effect observed when segmenting Glenmary glacier (2021) in Figure 5.1. One reason for the limitations of the system is that it only takes into account information from the pixel of interest, and does not use any image patterns or local features. For most images, the proposed system can obtain a segmentation mask with high enough accuracy (>0.80 Dice score) to be used to automate the identification of glacier areas during a glacier analysis process. With further improvements to mitigate limitations and increase segmentation accuracy, the system could be robust and accurate enough to be implemented as a glacier area identification tool for use in a glacier analysis process. It could replace the manual identification of glacier areas and could be used to calculate surface area measurements.

Chapter 6

Conclusions and Future Work

6.1 Major contributions

The aim of this work is to design and implement an unsupervised image segmentation system that can automatically and accurately identify the area of a glacier in an aerial image. The proposed system implemented and extended upon Otsu's thresholding method through the use of a multi-channel thresholding system with the addition of threshold bias values, and contour identification. A mask combination process was used to combine glacier areas identified in each colour channel. Image morphology was used to refine glacier areas and smooth edges in the predicted segmentation masks. Four different variations of U-Net model were trained and evaluated using a small set of manually generated training data. Two baseline unsupervised segmentation methods, Otsu's thresholding, and global thresholding were used to obtain benchmark results to compare against the proposed method. The results of the proposed method were compared with the results of the U-Net models. The proposed method was able to improve significantly on the baseline Otsu's method, and baseline global thresholding method. More importantly, the proposed method greatly outperformed the U-Net models. It was found the dataset was too small for the U-Net models to perform accurately, and techniques used to mitigate the effects of this were unsuccessful.

Limitations of the proposed method were identified, as the system's accuracy greatly decreased when applied to images containing shadow-covered glacier areas and light-reflective rock areas. With some improvements to mitigate the limitations of the proposed method, accuracy and robustness could be increased. An improved system could be implemented as an automatic glacier identification tool to be used in the current analysis process of aerial glacier imagery from the Southern Alps of New Zealand.

The objectives of this project were to (a) design an unsupervised system to identify the main glacier area; (b) implement and extend an existing supervised image segmentation technique to identify the main glacier area, and (c) compare the designed unsupervised system to the supervised image segmentation model. These three objectives were completed through the design of an unsupervised segmentation system and comparing it with supervised learning models. The completion of these objectives is a crucial step towards being able to automate the monitoring of glacier change in the Southern Alps of New Zealand, and potentially in other regions. The problem tackled in this project is not as simple as was first thought, and had many challenges. Despite this, the results obtained from the designed system were impressive, and make an important contribution to progress in this domain. In future work, the designed system can be extended to achieve the original goals of identifying the accumulation zone, ablation zone, and ELA.

6.2 Future work

After an analysis of the limitations of the designed system, avenues for improving the system and obtaining better results have been identified. A system that can use image patterns and local features to assist with pixel classification would likely improve upon the existing system. Similarly to how a UNet model extracts features from an image, local features such as local region variance and patterns could be used to help correctly segment areas. An example of this could be using a local region variance value to assist in deciding what class a pixel belongs to. In shadow-covered glacier areas variance across the area is very small. The region is often smoother with less fluctuation in pixel intensity than rock or sun-lit glacier areas. Extending the designed system to implement techniques such as this could provide an opportunity for more information to be utilised.

Another avenue of improvement could be to revisit an idea explored in Section 3.2.3 and develop a more robust and accurate system. The idea was to use local adaptive thresholding but to use a technique to decide whether to select a threshold or not. A system that can do this reliably could completely mitigate the problem encountered when applying local adaptive threshing to the glacier identification problem. The main problem encountered was that when using a local adaptive thresholding algorithm, a threshold was always selected for every local region. This would result in poorly chosen thresholds for regions containing a majority of pixels from a single class, such as a region of only white snow. In these regions, a global threshold produces more accurate segmentations, as the threshold is selected based on the whole image. Local adaptive thresholding selected good threshold values for regions containing a mixture of pixels from both classes, as a threshold can be found to split the pixels into classes accurately. A point of exploration could be to develop a technique that can decide whether a local threshold is needed, or a global threshold is needed.

One more avenue to explore could be to use the digital elevation models (DEMs) generated when creating orthophotos to identify where areas of shadow are likely to appear over a glacier. Depending on the time of day at which the photos were taken, the DEM of a glacier can be used to figure out where the mountain casts shadows over the glacier. This information could be used to improve the identification of shadow-covered glacier areas.

A more interesting avenue to explore would be to obtain a larger training dataset to see if the supervised models can be significantly improved. Although the UNet model is known to be able to perform well on small datasets, [30] the dataset of 40 images was probably too small. With enough data, the UNet models may be able to learn image features and patterns more consistently. It would be interesting to see if the UNet model can accurately classify shadow-covered glacier areas and light-reflective rock areas, and outperform the designed unsupervised system. Another route to explore with more training data would be to train separate UNet models for each glacier. Through analysis of the current training images, it is obvious that each glacier has unique features and patterns. Training on a dataset of images from many glaciers could be preventing optimal learning of the model. Training a unique UNet model on each glacier could enhance the learning of the models and produce overall better results.

Bibliography

[1] R. Hock, G. Rasul *et al.*, "Chapter 2: High Mountain Areas," in *IPCC Special Report on the Ocean and Cryosphere in a Changing Climate*, H. Portner, D. Roberts, V. Masson-Delmotte, P. Zhai, M. Tignor, E. Poloczanska, K. Mintenbeck, A. Alegria, M. Nicolai, A. Okem, J. Petzold, B. Rama, and N. Weyer, Eds. Cambridge University Press, 2019, ch. 2.

[2] A. Anthwal, V. Joshi, A. Sharma, and S. Anthwal, "Retreat of himalayan glaciers–indicator of climate change," *Nature and Science*, vol. 4, no. 4, pp. 53–59, 2006.

[3] W. Haeberli, J. Cihlar, and R. G. Barry, "Glacier monitoring within the global climate observing system," *Annals of Glaciology*, vol. 31, pp. 241–246, 2000.

[4] J. G. Cogley, A. Arendt, A. Bauder, R. Braithwaite, R. Hock, P. Jansson, G. Kaser, M. Moller, L. Nicholson, L. Rasmussen *et al.*, "Glossary of glacier mass balance and related terms," 2010.

[5] M. Zemp, M. Hoelzle, and W. Haeberli, "Six decades of glacier mass-balance observations: a review of the worldwide monitoring network," *Annals of Glaciology*, vol. 50, no. 50, pp. 101–111, 2009.

[6] S. Baumann, B. Anderson, T. Chinn, A. Mackintosh, C. Collier, A. M. Lorrey, W. Rack, H. Purdie, and S. Eaves, "Updated inventory of glacier ice in New Zealand based on 2016 satellite imagery," *Journal of Glaciology*, vol. 67, no. 261, pp. 13–26, 2021.

[7] S. H. Mernild, M. Pelto, J. K. Malmros, J. C. Yde, N. T. Knudsen, and E. Hanna, "Identification of snow ablation rate, ela, aar and net mass balance using transient snowline variations on two arctic glaciers," *Journal of Glaciology*, vol. 59, no. 216, pp. 649–659, 2013.

[8] A. E. Racoviteanu, M. W. Williams, and R. G. Barry, "Optical remote sensing of glacier characteristics: a review with focus on the himalaya," *Sensors*, vol. 8, no. 5, pp. 3355–3383, 2008.

[9] D. I. Benn and F. Lehmkuhl, "Mass balance and equilibrium-line altitudes of glaciers in high-mountain environments," *Quaternary International*, vol. 65, pp. 15–29, 2000.

[10] E. LaChapelle, "Assessing glacier mass budgets by reconnaissance aerial photography," *journal of Glaciology*, vol. 4, no. 33, pp. 290–297, 1962.

[11] T. Chinn, B. Fitzharris, A. Willsman, and M. Salinger, "Annual ice volume changes 1976–2008 for the New Zealand Southern Alps," *Global and Planetary Change*, vol. 92, pp. 105–118, 2012.

[12] A. Willsman and G. Macara, "New zealand glacier monitoring: End of summer snowline survey 2019," 2019.

[13] A. M. Lorrey, L. Vargo, H. Purdie, B. Anderson, N. J. Cullen, P. Sirguey, A. Mackintosh, A. Willsman, G. Macara, and W. Chinn, "Southern Alps equilibrium line altitudes: four decades of observations show coherent glacier–climate responses and a rising snowline trend," *Journal of Glaciology*, pp. 1–14, 2022.

[14] B. K. Veettil, S. Wang, S. F. de Souza, U. F. Bremer, and J. C. Simoes, "Glacier monitoring and glacier-climate interactions in the tropical andes: A review," *Journal of South American Earth Sciences*, vol. 77, pp. 218–246, 2017.

[15] T. Chinn, "Glacier fluctuations in the southern alps of new zealand determined from snowline elevations," *Arctic and alpine research*, vol. 27, no. 2, pp. 187–198, 1995.

[16] S. A. Elias, *Encyclopedia of Quaternary science*. Elsevier, 2006.

[17] L. J. Vargo, B. M. Anderson, H. J. Horgan, A. N. Mackintosh, A. M. Lorrey, and M. Thornton, "Using structure from motion photogrammetry to measure past glacier changes from historic aerial photographs," *Journal of Glaciology*, vol. 63, no. 242, pp. 1105–1118, 2017.

[18] D. Baswaraj, A. Govardhan, and P. Premchand, "Active contours and image segmentation: The current state of the art," *Global Journal of Computer Science and Technology*, 2012.

[19] X. Xia and B. Kulis, "W-net: A deep model for fully unsupervised image segmentation," *arXiv preprint arXiv:1711.08506*, 2017.

[20] J. S. Sevak, A. D. Kapadia, J. B. Chavda, A. Shah, and M. Rahevar, "Survey on semantic image segmentation techniques," in *2017 International Conference on Intelligent Sustainable Systems (ICISS)*. IEEE, 2017, pp. 306–313.

[21] P. F. Felzenszwalb and D. P. Huttenlocher, "Efficient graph-based image segmentation," *International journal of computer vision*, vol. 59, no. 2, pp. 167–181, 2004.

[22] J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 22, no. 8, pp. 888–905, 2000.

[23] N. Dhanachandra, K. Manglem, and Y. J. Chanu, "Image segmentation using k-means clustering algorithm and subtractive clustering algorithm," *Procedia Computer Science*, vol. 54, pp. 764–771, 2015.

[24] A. Kanezaki, "Unsupervised image segmentation by backpropagation," in *2018 IEEE international conference on acoustics, speech and signal processing (ICASSP)*. IEEE, 2018, pp. 1543–1547.

[25] H. Liu and K. C. Jezek, "A complete high-resolution coastline of antarctica extracted from orthorectified radarsat sar imagery," *Photogrammetric Engineering & Remote Sensing*, vol. 70, no. 5, pp. 605–616, 2004.

[26] Y. Yu, Z. Zhang, M. Shokr, F. Hui, X. Cheng, Z. Chi, P. Heil, and Z. Chen, "Automatically extracted antarctic coastline using remotely-sensed data: An update," *Remote Sensing*, vol. 11, no. 16, p. 1844, 2019.

[27] F. Yan, H. Zhang, and C. R. Kube, "A multistage adaptive thresholding method," *Pattern recognition letters*, vol. 26, no. 8, pp. 1183–1191, 2005.

[28] N. Otsu, “A threshold selection method from gray-level histograms,” *IEEE transactions on systems, man, and cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.

[29] M. I. Sezan, “A peak detection algorithm and its application to histogram-based image data reduction,” *Computer vision, graphics, and image processing*, vol. 49, no. 1, pp. 36–51, 1990.

[30] O. Ronneberger, P. Fischer, and T. Brox, “U-net: Convolutional networks for biomedical image segmentation,” in *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015, pp. 234–241.

[31] X. Xiao, S. Lian, Z. Luo, and S. Li, “Weighted res-unet for high-quality retina vessel segmentation,” in *2018 9th international conference on information technology in medicine and education (ITME)*. IEEE, 2018, pp. 327–331.

[32] J. Jing, Z. Wang, M. Rätsch, and H. Zhang, “Mobile-unet: An efficient convolutional neural network for fabric defect detection,” *Textile Research Journal*, vol. 92, no. 1-2, pp. 30–42, 2022.

[33] B. Baheti, S. Innani, S. Gajre, and S. Talbar, “Eff-unet: A novel architecture for semantic segmentation in unstructured environment,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, 2020, pp. 358–359.

[34] L. Zhang, J. Shen, and B. Zhu, “A research on an improved unet-based concrete crack detection algorithm,” *Structural Health Monitoring*, vol. 20, no. 4, pp. 1864–1879, 2021.

[35] M. Periyasamy, A. Davari, T. Seehaus, M. Braun, A. Maier, and V. Christlein, “How to get the most out of u-net for glacier calving front segmentation,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 15, pp. 1712–1723, 2022.

[36] E. Zhang, L. Liu, L. Huang, and K. S. Ng, “An automated, generalized, deep-learning-based method for delineating the calving fronts of greenland glaciers from multi-sensor remote sensing imagery,” *Remote Sensing of Environment*, vol. 254, p. 112265, 2021.

[37] V. Iglovikov and A. Shvets, “Ternausnet: U-net with vgg11 encoder pre-trained on imagenet for image segmentation,” *ArXiv e-prints*, 2018.

[38] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein *et al.*, “Imagenet large scale visual recognition challenge,” *International journal of computer vision*, vol. 115, no. 3, pp. 211–252, 2015.

[39] T. Y. Goh, S. N. Basah, H. Yazid, M. J. A. Safar, and F. S. A. Saad, “Performance analysis of image thresholding: Otsu technique,” *Measurement*, vol. 114, pp. 298–307, 2018.

[40] milesial, “Pytorch-unet,” https://github.com/milesial/Pytorch-UNet, 2022.

[41] F. Perez, C. Vasconcelos, S. Avila, and E. Valle, “Data augmentation for skin lesion analysis,” in *OR 2.0 Context-Aware Operating Theaters, Computer Assisted Robotic Endoscopy, Clinical Image-Based Procedures, and Skin Image Analysis*. Springer, 2018, pp. 303–311.

[42] M. Sato, K. Horie, A. Hara, Y. Miyamoto, K. Kurihara, K. Tomio, and H. Yokota, “Application of deep learning to the classification of images from colposcopy,” *Oncology letters*, vol. 15, no. 3, pp. 3518–3523, 2018.