Tracing Englacial Layers in Radargram via Semi-supervised Method: A Preliminary Result

Atefeh Jebeli[1,2], Bayu Adhi Tama[2], Sanjay Purushotham[1], Vandana P. Janeja[1,2]

[1] Department of Information Systems, UMBC, MD, USA
[2] Institute for Harnessing Data and Model Revolution in the Polar Regions (iHARP), UMBC, MD, USA
a.jebeli@umbc.edu, bayu@umbc.edu, psanjay@umbc.edu, vjaneja@umbc.edu

Abstract

The melting of ice sheets significantly contributes to sea level rise, highlighting the crucial need to comprehend ice structure for climate benefits. The stratigraphy of ice sheets revealed through ice layer radargrams gives us a window into historical depth-age correlations and accumulation rates. Harnessing this knowledge is not only crucial for interpreting both past and present ice dynamics, especially concerning the Greenland ice sheet, but also for making informed decisions to mitigate the impacts of climate change. Ice layer tracing is prevalently conducted using manual or semi-automatic approaches, requiring significant time and expertise. This study aims to address the need for efficient and precise tracing methods in a two-step process. This is achieved by utilizing an unsupervised annotation method (i.e., ARESELP) to train deep learning models, thereby reducing the need for extensive and time-consuming manual annotations. We benchmarked the popular U-Net segmentation model and its variants, such as U-Net with VGG19, U-Net with Attention mechanism, and U-Net with Inception. Additionally, various thresholding methods such as binary, Otsu, and CLAHE have been explored to achieve optimal enhancement for the true label annotation images. Our preliminary experiments indicate that the combination of attention U-Net with specific processing techniques yields the best performance in terms of the binary IoU metric.

Introduction

Global sea levels are rising due to climate change and ice sheet melting. Researching the causes of melting is a priority for agencies such as the National Science Foundation (NSF). Scientists could calculate historical accumulation rates and ice layer age-depth connections using radargrams. The reconstruction of historical ice sheet dynamics, such as those observed in Greenland, plays a crucial role in elucidating the sea-level fluctuations. It is imperative to conduct a comprehensive analysis of these ice layers and employ automated approaches as a means to mitigate the adverse effects of rising sea levels. The utilization of automated approaches has the potential to expedite this process and offer insights that can enhance our comprehension of global climate dynamics. The elucidation of the correlation between ice sheet *stratigraphy* and the phenomenon of sea level rise holds promise for enhancing our ability to effectively mitigate and anticipate prevailing challenges. The presence of noise in radargrams poses challenges to detecting ice layers. Here, most previous works have relied on manual or semi-automatic approaches. Despite their effectiveness, these approaches are characterized by their labor-intensive nature and demand a high level of expertise (Dong et al. 2021; Liu-Schiaffini et al. 2022). This study examines the potential of artificial intelligence (AI) to automate the recognition of ice layers. The use of AI not only enhances our understanding of ice sheet dynamics but also holds the promise of aiding policymakers and stakeholders in making more informed decisions regarding climate action. We employ semi-supervised learning as an alternative to manual annotations. An unsupervised method, e.g. ARESELP (Xiong, Muller, and Carretero 2017) is employed to generate mask label images, while the original radargram images are obtained from the Center for Remote Sensing of Ice Sheets (CReSIS). Our study encompasses several deep learning techniques for a segmentation task, such as U-Net (Ronneberger, Fischer, and Brox 2015), U-Net+VGG19 (Simonyan and Zisserman 2014), U-Net+Inception (Delibasoglu and Cetin 2020), and Attention U-Net (Oktay et al. 2018). In addition, we evaluate various pixel thresholding techniques, including binary, Otsu, and CLAHE, with a specific emphasis on enhancing the quality of binary mask labels. Our study serves as a fundamental study, highlighting the need to consider the availability and quality of data while choosing approaches. Our primary motivation is not only to demonstrate the effectiveness of our approach but also to overcome the labor-intensive procedure of human annotation.

Related Work

Previous study has examined the utilization of AI methods for identifying ice layers in radar imagery. The study done by (Varshney et al. 2020) utilized a dataset consisting of 2,621 radar images and corresponding ground truth images. The study limits its analysis to well-annotated layers for ice layers and converts discontinued ice layers into background pixels. The networks were trained on 28 classes of ice layers, and the model's performance was evaluated on the top 10 ice layers of test set images. Subsequently, a different study by (Yari et al. 2019) employs a multi-scale deep learning model (e.g., HED) to identify internal ice layers and uses synthetic data; however, the outcomes are unsatisfactory. A recent study conducted by (Varshney et al. 2021) highlights the importance of utilizing cropped radargram images and their corresponding masks instead of generating predictions for full-size radargram images. This approach is chosen due to the presence of only a restricted number of ice layers in the radargram images, which are accompanied by accurately annotated ground truth data. In a separate, (Dong et al. 2021) uses synthetic radargram images to train a deep learning model (e.g., U-Net). Our current work, however, adopts a unique approach by incorporating an unsupervised method's annotations into a supervised deep learning model, aiming to automate detecting ice layers in radar images. In contrast to previous studies, we utilize the complete radargram images without any cropping or eliminating ice layers, along with their generated ground truth obtained from an unsupervised method (e.g., ARESELP). This work is an extension of our previous study (Jebeli et al. 2023) that developed a two-step annotation method for bedrock and top surface detection.

Methods

In our study, we leverage the U-Net model and its variants, a widely recognized deep learning approach for semantic image segmentation tasks (Ronneberger, Fischer, and Brox 2015). Fundamentally, U-Net employs an encoder to extract features and a decoder for image reconstruction. These two components are connected by unique skip connections, which are distinct and specific to the U-Net model. To amplify segmentation performance, we integrate U-Net architecture with two prominent pre-trained models. Firstly, the U-Net+VGG19 leverages the VGG19 model for high-level feature extraction (Simonyan and Zisserman 2014). Secondly, the U-Net+Inception can be described as a fusion of the U-Net with Inception blocks (Delibasoglu and Cetin 2020). The use of parallel in these blocks enhances feature extraction, resulting in improved computational performance and reduced network size. Furthermore, we also explore the Attention U-Net architecture, which incorporates attention gates to guide the model's focus during segmentation. The Attention U-Net model improves segmentation accuracy by enabling the network to focus on specific regions of interest. This refinement is beneficial in complex scenarios where differentiating between the target and background is difficult. The utilization of this attention-driven improvement has played a crucial role in our work, resulting in a substantial improvement in segmentation accuracy (Oktay et al. 2018).

Material and Experimental Setting

Radargram Data Preprocessing

The CReSIS serves as a prominent provider of radargrams, considered primary sources in the field. This study involves 300 radargram imagery, each processing dimension of 1408×1024 pixels and 3 color channels. Furthermore, due to the unavailability of public ground truth images, we employed an unsupervised model, specifically ARESELP (Xiong, Muller, and Carretero 2017), to generate a collection of ground truth annotations. To measure the quality of an unsupervised model annotation, we use various metrics such as layer breakpoints and local layer density (Tack et al. 2023). Table 1 compares the example of ice layer annotation between ARESELP and manual approaches conducted by a domain expert[1]. In addition, we explore various techniques to produce a binary ground truth tailored to the requirements of binary semantic segmentation deep learning models. This process entails the application of binary thresholding, the Otsu method, and the Contrast Limited Adaptive Histogram Equalization (CLAHE) technique to transform the original ground truth images, which have varying pixel values, into an appropriate binary format. These strategies aim to address the intricacies inherent in this task and further the advancements in binary semantic segmentation. We briefly discuss these three techniques as follows: (i) *Binary thresholding* is done by setting a threshold value. We can consider the value below the threshold as background pixels, whereas above are subject of interest pixels. We test different threshold levels to find the best points for binarizing our unsupervised annotations, and the test shows that 127 is the appropriate threshold for differentiating background and ice layers. (ii) *Otsu's thresholding* automatically identifies the optimal threshold for grayscale images by maximizing inter-class variance, minimizing intra-class variance, and efficiently separating foreground and background pixels. It is computationally economical, versatile, effective for binary semantic segmentation, and requires no user input. (iii) *CLAHE* is an advanced contrast enhancement approach that overcomes traditional histogram equalization. As opposed to global contrast transformation, CLAHE transforms local image regions. This makes the enhancement adaptable, preserving smaller features without over-amplifying noise. CLAHE excels at improving feature differentiation, which can significantly benefit binary thresholding for segmentation tasks.

Experimental Setting

Given the limited size of the available dataset, we apply two augmentation techniques—horizontal flip and noise addition—to boost model performance, yielding a total of 900 images. The dataset is split 80/20 for the training/testing sets. A 5-fold cross-validation ($5CV$) strategy and early stopping are used on the training subset to ensure model generalizability and avoid overfitting. The test subset is used to evaluate the model on unseen data. Each model iteration has 300 epochs with a batch size of 8. We resize the input size image to 512×512, while two metrics, such as Binary Intersection over Union (binary IoU) and dice coefficient, are used as evaluation metrics. Binary IoU is the ratio of the area overlap between the predicted segmentation and the true label to the area of union between the predicted segmentation and the true label. On the other hand, the dice coefficient is computed as twice the intersection of the predicted segmentation and the ground truth, divided by the sum of the pixel counts of both the predicted and ground truth labels.

[1]We acknowledge that expert label annotations are provided by Nicholas D. Holschuh at Amherst College, MA, United States

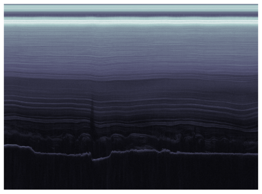
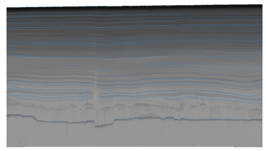
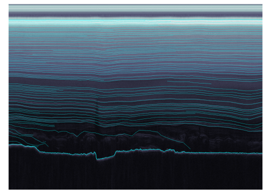
Radargram	Manual	ARESELP
	20120330_01_021	
		

Table 1: A comparison between manual and automatic (e.g., ARESELP (Xiong, Muller, and Carretero 2017)) annotations for a radargram sample.

We have open-sourced our work at github[2].

Result and Discussion

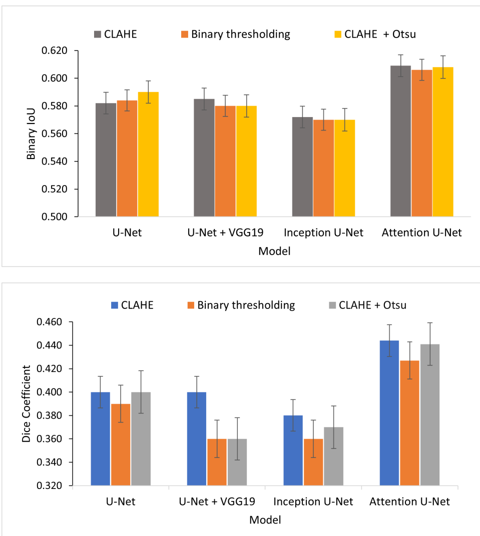


Figure 1: Performance average the test subset of various segmentation models over $5CV$ w.r.t. binary IoU (top) and dice coefficient (bottom).

In Figure 1, we present the mean performance of each segmentation model on the test subset with various binarizing ground truth strategies with regard to binary IoU and dice coefficient, respectively. Our results highlight the superior performance of the Attention U-Net, particularly when employing the CLAHE approach. Note that Attention U-Net has not been considered in the previous studies, and the superiority of Attention U-Net lies in its attention mechanism, which can effectively capture the fine-grained details (e.g., ice layers) information from a broader area around each pixel in the input radargram images. In the second place, the original U-Net still performs better than other variants. Previous work (Dong et al. 2021) also highlights the superiority of the U-Net for ice layer detection on synthetic datasets. However, the lack of accessibility to the dataset and the absence of the publicly available code pose challenges when attempting to do an in-depth evaluation of this aspect. Furthermore, it is important to highlight that the performance variance of all segmentation models across the three thresholding methods is not significant. Therefore, this indicates that all segmentation models are robust to thresholding strategies. All models can generalize well across different thresholds and still produce reasonably good segmentations.

In Table 2 we assess the quality prediction of four segmentation approaches. Here, we include three different radargram frames. The results clearly demonstrate that Attention U-Net produces remarkably more englacial layers than other baselines. In addition, the quality of ice layer annotation generated by Attention U-Net is more consistent and continuous, with fewer discontinuous layers. Nevertheless, it is imperative to acknowledge that our findings remain preliminary and necessitate more verification from experts in the field such as glaciologists.

Conclusion and Future Work

This study introduces a two-step approach for automating the annotation of ice layers in radargrams. Initially, we utilized the ARESELP unsupervised method to generate ice layer annotations. However, due to the method's limited applicability — constrained to specific years of radargrams — we transitioned to a deep learning-based segmentation technique. This approach allows the model, once trained on the generated annotations, to swiftly predict ice layers on radargrams from a wider range of years. Our preliminary investigation assessed four well-known deep learning models, such as U-Net, U-Net+VGG19, U-Net+Inception, and Attention U-Net. In addition to this, we explored various pixel thresholding techniques to enhance the quality of our ground truth images. Given the preliminary stage of this research, our immediate focus lies in refining the unsupervised annotations, devising automated metrics for their evaluation, and working closely with domain experts to get better evaluations. Future studies may delve deeper, potentially introducing advanced segmentation models, such as transformer architectures.

Acknowledgments

This work is funded by the National Science Foundation (iHARP, Award #2118285).

[2]https://github.com/iharp-institute/Tracing-Englacial-Layers-in-Radargram-via-Semi-Supervised-Method

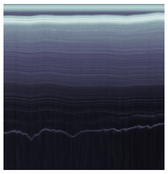
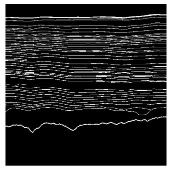
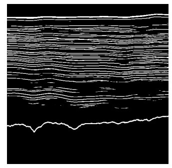
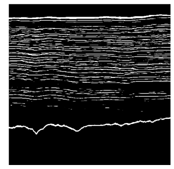
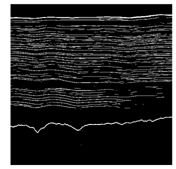
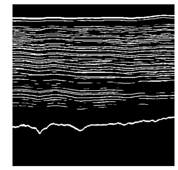
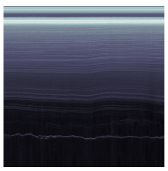
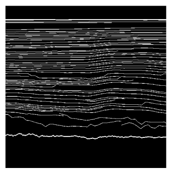
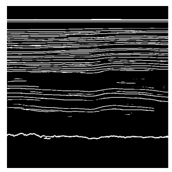
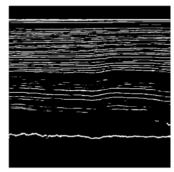
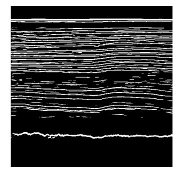
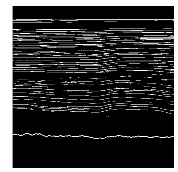
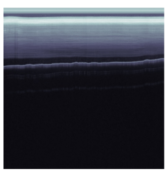
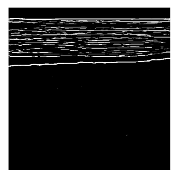
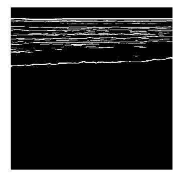
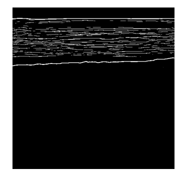
Original radargram	Mask label	Prediction			
		U-Net	U-Net + VGG19	U-Net + Inception	U-Net + Attention
					
					
					

Table 2: Example of predicted image for four segmentation models with CLAHE + Otsu's thresholding strategy.

References

Delibasoglu, I.; and Cetin, M. 2020. Improved U-Nets with inception blocks for building detection. *Journal of Applied Remote Sensing*, 14(4): 044512–044512.

Dong, S.; Tang, X.; Guo, J.; Fu, L.; Chen, X.; and Sun, B. 2021. EisNet: Extracting Bedrock and Internal Layers from Radiostratigraphy of Ice Sheets with Machine Learning. *IEEE Transactions on Geoscience and Remote Sensing*, 60: 1–12.

Jebeli, A.; Tama, B. A.; Janeja, V. P.; Holschuh, N.; Jensen, C.; Morlighem, M.; MacGregor, J. A.; and Fahnestock, M. A. 2023. TSSA: Two-Step Semi-Supervised Annotation for Radargrams on the Greenland Ice Sheet. In *IGARSS 2023 - 2023 IEEE International Geoscience and Remote Sensing Symposium*, 56–59.

Liu-Schiaffini, M.; Ng, G.; Grima, C.; and Young, D. 2022. Ice Thickness From Deep Learning and Conditional Random Fields: Application to Ice-Penetrating Radar Data With Radiometric Validation. *IEEE Transactions on Geoscience and Remote Sensing*, 60: 1–14.

Oktay, O.; Schlemper, J.; Folgoc, L. L.; Lee, M.; Heinrich, M.; Misawa, K.; Mori, K.; McDonagh, S.; Hammerla, N. Y.; Kainz, B.; et al. 2018. Attention u-net: Learning where to look for the pancreas. *arXiv preprint arXiv:1804.03999*.

Ronneberger, O.; Fischer, P.; and Brox, T. 2015. U-Net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, 234–241. Springer.

Simonyan, K.; and Zisserman, A. 2014. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

Tack, N.; Tama, B. A.; Jebeli, A.; Janeja, V. P.; Engel, D.; and Williams, R. 2023. Metrics for the Quality and Consistency of Ice Layer Annotations. In *IGARSS 2023 - 2023 IEEE International Geoscience and Remote Sensing Symposium*, 4935–4938.

Varshney, D.; Rahnemoonfar, M.; Yari, M.; and Paden, J. 2020. Deep Ice Layer Tracking and Thickness Estimation using Fully Convolutional Networks. In *2020 IEEE International Conference on Big Data (Big Data)*, 3943–3952.

Varshney, D.; Rahnemoonfar, M.; Yari, M.; Paden, J.; Ibikunle, O.; and Li, J. 2021. Deep learning on airborne radar echograms for tracing snow accumulation layers of the Greenland ice sheet. *Remote Sensing*, 13(14): 2707.

Xiong, S.; Muller, J.-P.; and Carretero, R. C. 2017. A new method for automatically tracing englacial layers from MCoRDS data in NW Greenland. *Remote Sensing*, 10(1): 43.

Yari, M.; Rahnemoonfar, M.; Paden, J.; Oluwanisola, I.; Koenig, L.; and Montgomery, L. 2019. Smart tracking of internal layers of ice in radar data via multi-scale learning. In *2019 IEEE International Conference on Big Data (Big Data)*, 5462–5468. IEEE.