Enhancing Sea Ice Segmentation Model: Exploring Neural Network Ensembles and Optimization Strategies for Improved Performance

Bruno Henrique dos Santos Marques
Centro de Informática
UFPE
Recife, Brazil
bhsm@cin.ufpe.br

José Vinicius de Santana Souza
Centro de Informática
UFPE
Recife, Brazil
jvss2@cin.ufpe.br

Victor Gabriel de Carvalho
Centro de Informática
UFPE
Recife, Brazil
vgc3@cin.ufpe.br

***Abstract*—Monitoring sea ice extent holds significant importance in both navigation and environmental science. Utilizing segmentation models like UNet enables us to achieve a nuanced differentiation of sea ice chunks in radar images. This approach facilitates researchers in gaining more precise insights into climate change and water dynamics, enhancing overall efficiency. The objective of this project is to deploy a UNet model for sea ice segmentation, incorporating ensemble methods and a hyperparameter tuning framework to optimize performance. The goal is to demonstrate the substantial impact of proper optimization on the accuracy and effectiveness of the model.**

***Index Terms*—segmentation models, UNet, hyperparameter tuning, ensemble methods, model optimization, image processing, machine learning, deep learning**

I. Introduction

Polar ice covers about 10% of Earth's surface [1] and represents almost 70% of all freshwater [2]. Given this, it is undeniable that the polar regions play a pivotal role in the Earth's climate and socioeconomic system. Monitoring the dynamics of sea ice is essential for understanding temperature variations and their broader implications. Knowing precisely where ice chunks are and their dimensions can provide us with a clear oversight of the consequences of climate change and water dynamics worldwide.

A segmentation task, as the name suggests, is the process of dividing an image into different regions based on the characteristics of adjacent pixels. Consequently, ice segmentation consists of the intricate task of accurately delineating sea-ice boundaries using advanced computer vision techniques. Sea ice segmentation poses a unique set of challenges due to its complex and dynamic nature. Boundaries are often indistinguishable to the human eye.

This project unfolds the journey of modeling and enhancing a sea ice segmentation model using a data set of ice chunk images taken from space. Along with ranking the best approaches and results.

II. Objectives

Within this project, our objective is to construct a UNet-based image segmentation system [4] from the ground up. This system is designed to accurately segment ice chunks observed from space, with a particular focus on distinguishing subtle ice and water interfaces.

Moreover, we aspire to enhance the model's performance by exploring and assessing various optimization techniques in conjunction with ensemble methods. The assessment will employ established segmentation metrics, including Intersection over Union (IoU), Pixel Accuracy, and Dice Coefficient. Additionally, we will provide visualizations of predictions and ground truths to further evaluate and interpret the results.

III. Justification

Due to the important role played by sea ice in both navigation and environmental science questions,' it is utterly necessary to provide precise and reliable information about sea ice's location, size and extent. This information can help prevent future problems and provide valuable data for projects related to sea ice throughout the world.

We believe that our project can both make the climate change a more traceable problem and show the impact of optimization methods in a model's results.

IV. Methodology

This paper will rely on the Leeds SciML Sea Ice Segmentation Kaggle competition dataset [3], in which its main objective is to generate a sea ice segmented image.

A. Dataset Structure

a) Data: The dataset is composed entirely by ".tiff" images. Each data sample is separated into three different image files by its titles ending:

1) Image files with titles ending with "sar.tiff", as shown in Fig. 1, are single-band SAR (synthetic aperture radar) images, providing useful information about the ice-water interface and its characteristics [3], representing the inputs for the models. It is worth saying that those images have low absolute valued pixels, that makes them hard to see as a standard RGB picture.

2) Image files with titles ending with “vis.tiff”, as shown in Fig. 2, are three-band optical images that provide visual data that will also be used as inputs as complements for the “sar.tiff” images [3]. These images will as well be used in data analysis by the team.
3) Image files with titles ending with “ref.tiff”, as shown in Fig. 3, are the ground truth label image files, providing the reference information about the dominant ice-water interface, and will be used to evaluate the accuracy of the trained models [3].

b) Train and Test datasets: The dataset is separated into two different folders:

1) “train” is the folder that contains the data used for training the models, and, as said before, each data sample is separated into three image files ending with: “sar.tiff” (inputs for the model), “vis.tiff” (inputs for the model and data visualization) and “ref.tiff” (ground truth for evalauating the model).
2) “test” is the folder that contains the data used for evaluate the trained models, and each data is separated into two image files ending with: “sar.tiff” (inputs for the model) and “vis.tiff” (data visualization).

As stated by the Kaggle competition description, the data has limitations such as coarse resolution and cloud cover [3] that may become a challenge for the models to extract the right features from the data.

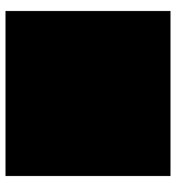

Fig. 1. Example of “sar.tiff” image

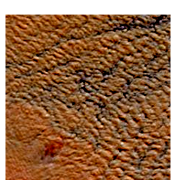

Fig. 2. Example of “vis.tiff” image

Fig. 3. Example of “ref.tiff” image

B. Model

In order to learn from data, a UNet [4] model will be used. It is one of the most popular approaches for image segmentation, combining the efficiency of convolutional neural networks and residual links. The encoder pathway captures features at multiple scales, while the decoder pathway reconstructs the segmented output with remarkable spatial resolution. Fig. 4 shows a simplified architecture for UNet, it is clear to see some main characteristics like the skip connections and its U shape.

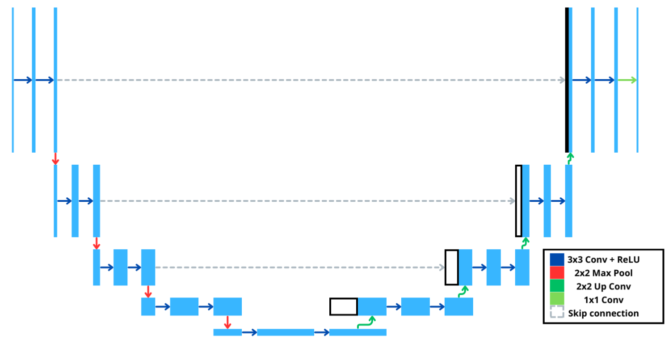

Fig. 4. Simplified diagram of UNet.

C. Optimization Techniques

To obtain the best possible results from the UNet [4] model, a series of optimization techniques will be used. We will focus on using different types of ensemble methods, and also, search for the best hyperparameters combination automatically with an hyperparameter optimization framework.

It is worth mentioning that these are state-of-the-art or already consolidated techniques, which contributed to our preference.

a) Hyperparameter Tuning: An important step in obtaining the best possible model, is the hyperparameter tuning phase, in which we will search for the best possible hyperparameter combination so that the model can perform with higher efficiency maintaining the best possible results. For an automatic hyperparameter tuning, the Optuna Framework [8] that employs a series of state of art optimization algorithms will be used.

b) Bagging [6]: Is the ensemble learning method that consists in training multiple instances of the same learning algorithm on different subsets of the training data. At the end, all outputs are averaged or voted for a final result. This technique defends that the combination of diverse models helps reduce variance and improve overall performance. Figure 5 shows Bagging workflow.

c) XGBoosting [5]: Belongs to the family of gradient boosting algorithms, designed for tree-based ensemble learning. It is particularly popular for its speed and performance in competitions on platforms like Kaggle. Its simplified diagram is shown in figure 6.

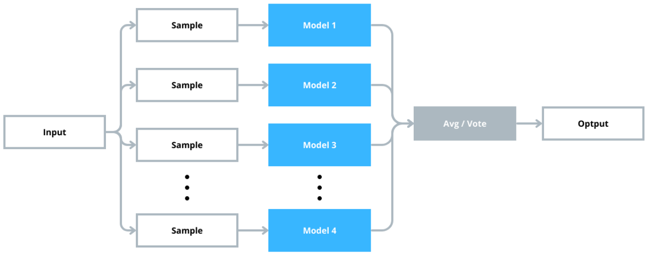

Fig. 5. Simplified diagram of Bagging technique.

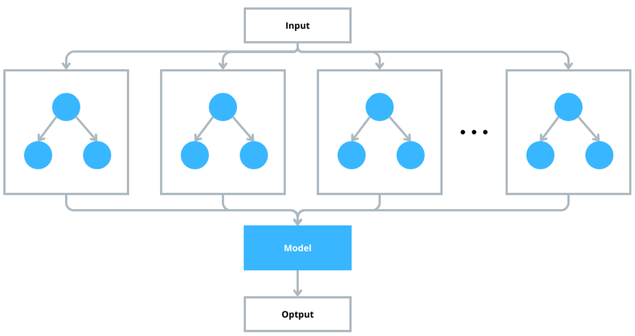

Fig. 6. Simplified diagram of XGBoosting technique.

d) Stacking [7]: Is an ensemble learning technique that involves combining predictions from multiple base models using V -fold cross-validation to build the optimal weighted combination of predictions from those base models. It re-assembles Bagging in some aspects, bur figure 7 can make its differences more intuitive.

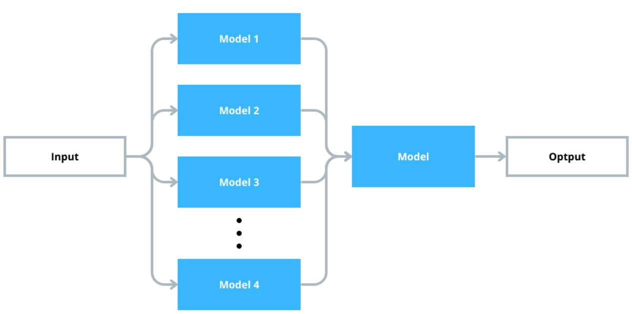

Fig. 7. Simplified diagram of Stacking technique.

D. Evaluation Metrics

a) Intersection over Union (IoU): One of the most used image segmentation metrics. It consists in quantifying how well the model can distinguish objects from their backgrounds in an image. It consists in determining how closely two areas (in our case, masks) overlap and can be defined as follows:

$$IoU = \frac{|A \cap B|}{|A| \cup |B|}$$

Note that A and B represent the conjunct of points on each related mask.

b) Dice Coefficient: It is a similarity metric that acts between two sets. In our case, it will check the similarity between our reference and our predicted image. Dice Coefficient is calculated from the precision and recall of a prediction, essentially being their harmonic mean.

$$DC = \frac{2 \times |A \cap B|}{|A| + |B|}$$

The main difference between DC and IoU is that Dice can be more robust to class imbalance. While IoU penalizes more under and over-segmentation.

c) Pixel Accuracy: Measures the overall accuracy of pixel-wise predictions. It is a really simple and straightforward approach, it is given by:

$$PA = \frac{Number\ of\ correctly\ classified\ pixels}{Total\ number\ of\ pixels}$$

d) Visualization: Model's outputs will also be visually analysed.

E. Coding Environment

During our implementation and evaluation, we will use entirely Google Colab, an online Python coding platform that provides us T4 GPUs, essential for training convolutional neural networks more rapidly. Additionally, Colab has the benefit of working remotely and without the need of a local powerful computer, code can also be shared and version tracked between our team.

V. Schedule

TABLE I
PROJECT SCHEDULE

Week	Activity
27/11	Project assignment
01/12 - 20/12	Choosing the project theme and preparing the proposal
28/01 - 03/02	Study and in-depth analysis of the dataset
04/02 - 10/02	Development of a segmentation model
11/02 - 17/02	Preparation of multiple models for ensemble
18/02 - 24/02	Application and evaluation of optimization techniques
25/02 - 02/03	Analysis of metrics and preparation of results
03/03 - 05/03	Finalization of the article and presentation
06/03	Project deadline

References

[1] NASA, “Global Ice Viewer” climate.nasa.gov. https://climate.nasa.gov/interactives/global-ice-viewer/#/ (accessed Dec. 15, 2023)

[2] T. Bralower, D. Bice, “Distribution of Water on the Earth’s Surface,” e-education.psu.edu. https://www.e-education.psu.edu/earth103/node/701 (accessed Dec. 15, 2023)

[3] Eilish O’Grady, Naomi Shakespeare-Rees, spiruel. (2023). ”Leeds SciML Sea Ice Segmentation,” Kaggle. https://kaggle.com/competitions/leeds-sciml-sea-ice-segmentation (accessed Dec. 18, 2023)

[4] Ronneberger, O., Fischer, P., Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab, N., Hornegger, J., Wells, W., Frangi, A. (eds) Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. MICCAI 2015. Lecture Notes in Computer Science(), vol 9351. Springer, Cham. https://doi.org/10.1007/978-3-319-24574-4_28

[5] Tianqi Chen and Carlos Guestrin. 2016. XGBoost: A Scalable Tree Boosting System. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD ’16). Association for Computing Machinery, New York, NY, USA, 785–794. https://doi.org/10.1145/2939672.2939785

[6] Breiman, L. Bagging predictors. Mach Learn 24, 123–140 (1996). https://doi.org/10.1007/BF00058655

[7] David H. Wolpert, Stacked generalization, Neural Networks, Volume 5, Issue 2, 1992, Pages 241-259, ISSN 0893-6080, https://doi.org/10.1016/S0893-6080(05)80023-1.

[8] Takuya Akiba, Shotaro Sano, Toshihiko Yanase, Takeru Ohta, and Masanori Koyama. 2019. Optuna: A Next-generation Hyperparameter Optimization Framework. In KDD.