Segmentation of Arctic Sea Ice

Alberto Dorizza[1], **Dario Mameli**[1], **Michele Russo**[1]

[1]Department of Information Engineering, University of Padova, Italy; alberto.dorizza@studenti.unipd.it, dario.mameli@studenti.unipd.it, michele.russo.2@studenti.unipd.it
Department of Information Engineering
University of Padova
35121 Padova, Italy

Abstract

The existence of an automatic and accurate sea ice mapping system capable of performing accurate and efficient sea ice segmentation can play an essential role in analysing and forecasting the effects of climate change on the Arctic region. In this article, after a presentation of the AutoICE challenge, two possible methods of improvement in the automatic sea ice mapping process are presented, namely a new pre-processing of Sentinel-1 SAR input data and a possible variant of the network architecture.

Introduction

Artic sea ice plays a crucial role in our ability to understand the impact of global warming, and the health status of the entire Earth. In fact, even though sea ice is mostly present in the polar regions, it is still able to influence our global climate by acting as a thermoregulator of the entire planet.

The rise of sea temperatures and the dramatic consequences of ice melting influence the path of ocean currents, that are one important factor for a milder climate.

The impact of sea ice mapping

The continuous monitoring and mapping of sea ice is crucial for scientists' understanding of the Arctic's susceptibility to climate change and also plays a key role in understanding the habitat loss of animals living in these areas, such as seals, walruses and polar bears.

It is true that sea ice follows a seasonal pattern, forming in the colder seasons and then melting until it reaches its minimum in spring and summer, usually around September. But the increasing availability of satellite images and the growing number of people accessing remote parts of the Arctic due to the ice thinning, make the interpretation of satellite imagery more complex and time-consuming.

In addition to furthering the understanding of the Earth's climate through the acquisition of climate models and their variations over time (useful for example to understand sea level rise), this activity is also useful for navigation, exploration and management of maritime routes and for the sustainability of our planet by protecting these environments.

Hence, satellite sensors play a key role in tracking the extent and movement of sea ice throughout the year, which would not always be possible with on-site analysis in difficult or inaccessible areas. Indeed, the creation of automated systems for mapping Arctic ice from satellite images not only offers several advantages in terms of speed and efficiency due to a more comprehensive analysis, but sometimes also makes the measurements themselves possible, when the extent of these vast sea ice shelves is larger.

ESA, and in particular its centers of Earth Observation, are studying polar ices, through satellites images, in order to understand possible future scenarios and raise awareness about these changes.

In particular, the technique actually used to study ice is called *Manual ice charting*. Since this technique needs a high number of experts in order to analyze the huge amount of data coming from satellites, they have presented the *AutoIce Challenge* in order to find an automatic approach to analyze ice data.

AutoICE Challenge

The AutoICE challenge (referred to by the acronym AI4EO) is a competition initiated by the ϕ-lab division of the European Space Agency and joined by the Norwegian Computing Center, the Danish Meteorological Institute (DMI), the Technical University of Denmark (DTU), Polar View and Nansen Environmental Remote Sensing Center (NERSC).

It aims to bring together people skilled in Artificial Intelligence (AI) and/or Earth Observation (EO) in order to create an efficient automatic sea ice mapping system from Sentinel-1 SAR data.

The need to create such a system derives from the fact that the manual method of ice mapping is time-consuming and limited in spatial and temporal coverage.

In particular, this challenge seeks to derive more robust and accurate automated sea ice maps and to advance the state of the art of sea ice parameter retrieval from SAR data, by considering more parameters such as sea ice concentration, stage of development and size (shape) of ice floes.

AI4EO took place online from 1 November 2022 to 17 April 2023.

Paper organization

In this paper we are going to present our segmentation model as a solution to the AutoICE challenge, highlighting the differences with the current model implementation.

The paper is organized as follows:

- In **Dataset** we discuss the data used by the model.
- In **Measures** we discuss the evaluation measures for the model.
- In **Current Segmentation** we discuss the current implementation of the segmentation model.
- In **Proposed Segmentation** we discuss the proposed implementation of the segmentation model.
- In **Conclusions and future work** we draw the final conclusions and delineate the future work.

Dataset

Sentinel-1 and AMSR-2 data

The data available for this challenge is in a particular format called *Netcdf*. It stores multidimensional scientific data provided by *Sentinel-1* satellites and *AMSR-2*.

Going deeper on the data given by the satellites, we have:

- Sea ice chart parameters
- Geographical information
- HH (horizontal transmit and horizontal received) polarization
- HV(horizontal transmit and vertical received) polarization
- SAR incident angle
- Microwave Radiometer
- Enviromental Variables(temperature, speed, wind, salinity)

The **Sea ice chart** parameters are SIC, SOD and FLOE. In details, **SIC** is the percentage of sea ice to open water for an area and it is discretised into bins of 10%.

SOD is a kind of sea ice percentage, but in this case we distinguish 5 labels. In particular it is focused on the quantity of ice in a certain area and its thickness.
FLOE describes how large or continuous the sea ice pieces/chunks are, and it has 6 labels.

The charts are manually computed and provided by the *Canadian and DMI operational ice services*. Each pixel of every image is associated with a colour indicating the label or bin it represents.

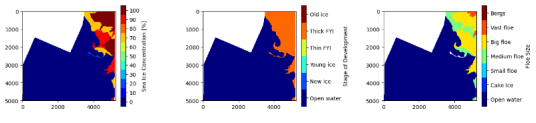

Figure 1: The image shows the plot of each chart

HH polarization, HV polarization and incidence angle are taken from a particular instrument, supplied on *Sentinel-satellites*, called **SAR synthetic-aperture radar**. This instrument is able to collect data in any atmospheric condition, day or night. The *Sentinel-1*'s sensor transmits a radar signal towards the ground, and the portion of the outgoing radar signal that the target on the Earth's surface redirects directly back towards the radar antenna is the backscatter. Sentinel-1 satellites can work in four modes, and for each mode it corresponds different typologies of images. For this challenge it has been used Extra Wide Swath Mode (EW) Level-1 Ground Range Detected products in Medium resolution (GRDM) and in dual polarization, HH and HV.(SAR -)

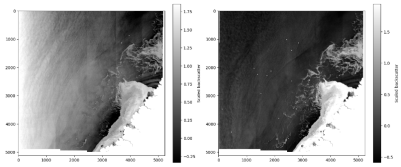

Figure 2: HH and HV images

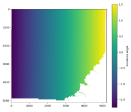

Figure 3: This image indicates the incidence angle of the land w.r.t. the satellites

Microwave Radiometer images are provided by **AMSR-2**, which is a passive instrument that provides radiometer scanning of the land with different microwave frequencies that goes from 6.9Hz to 89Hz. These waves provide crucial information about the state of ice, and in particular about its thickness.(AMS -)(Center 2016)

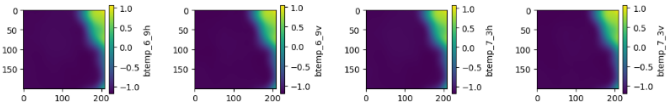

Figure 4: Images taken considering different frequencies and different axis (horizontal and vertical)

Finally, enviromental variables are taken with the use of **ERA5**, which is the fifth generation ECMWF atmospheric reanalysis of the global climate covering the period from January 1940 to the present day. ERA5 provides hourly estimates of a large number of atmospheric, land and oceanic climate variables.
For each Sentinel-1 scene the enviromental variables with the smallest difference in time to the Sentinel-1's acquisitions are retrieved and resampled.(Hersbach -)
Enviromental variables are various. Here are some of them:

- ERA5 eastward component of the 10m wind rotated to account for the Sentinel-1 flight direction.
- ERA5 northward component of the 10m wind rotated to account for the Sentinel-1 flight direction.

- ERA5 2m air temperature.
- ERA5 skin temperature.
- ERA5 total column water vapour.
- ERA5 total column cloud liquid water.

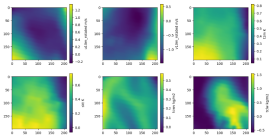

Figure 5: Enviromentals variables

Measures

The submitted models are evaluated according to the following three sea ice parameters:

- sea ice concentration (SIC)
- stage-of-development (SOD)
- floe size

In particular, there are used the following individual evaluation statistical metrics:

- SIC will be evaluated by calculating the R^2 coefficient formulated as:

$$R^2 = 1 - \frac{\sum_{i=1}^{N_{\text{pixel}}} \left(y_i^{\text{true}} - y_i^{\text{pred}} \right)^2}{\sum_{i=1}^{N_{\text{pixel}}} \left(y_i^{\text{true}} - \hat{y}_i^{\text{true}} \right)^2} \tag{1}$$

  where y_i^{true} is the true i^{th}, $\hat{y}_i^{\text{true}}$ is the mean true pixel value, and y_i^{pred} is the predicted class of the i^{th} pixel.
- SOD and Floe size will both be evaluated using the F1 score:

$$F1 = 2\frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}} \tag{2}$$

  where precision (P) and recall (R) are defined as:

$$\begin{aligned} P &= \frac{T_P}{T_P + F_P} \\ R &= \frac{T_P}{T_P + F_N} \end{aligned} \tag{3}$$

  where T_P is the number of predicted true positives, F_P the number of false positive and F_N the number of false negatives.

Subsequently, the three sea ice parameter scores will be combined into one single final score as:

$$\text{Score} = \frac{2}{5} \cdot R^2(\text{SIC}) + \frac{2}{5} \cdot F1(\text{SOD}) + \frac{1}{5} \cdot F1(\text{Floe size}) \tag{4}$$

Notice that when deriving the final score, SIC and SOD will be weighted over floe size.

Current Segmentation

Preprocessing

All the images from *Sentinel-1* were preprocessed by applying SAR noise correction, performed by NERSC (Korosov et al. 2022).

U-Net

In order to process the data, the organizers of the challenge have fine-tuned an implementation of *U-Net* using *PyTorch*. **U-Net** is a *fully convolutional neural network* (**FCN**) that was first presented in (Ronneberger, Fischer, and Brox 2015) for the segmentation of biomedical images.

Architecture The implemented structure mostly follows the same pipeline as in **Figure 6**. It is a U-shaped **encoder-decoder** network, where the encoder is also known as *contracting path*, while the decoder is often referred as *expanding path*.

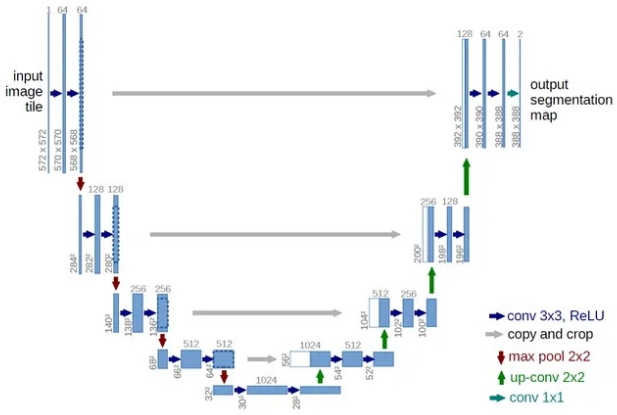

Figure 6: U-Net original architecture, as described in (Ronneberger, Fischer, and Brox 2015)

- **Contracting path**: it consists of 4 consecutive *contracting blocks*. These blocks are composed by two convolutional layers made of a number of 3x3 convolutional filters with *zero padding* (to preserve the dimension of the input), each followed by a rectified linear unit (ReLU) and a 2x2 max pooling operation with stride 2, for downsampling. This last operation effectively halves the size of the image. Each block doubles the number of convolutional filters of the previous one, resulting in double the amount of feature maps, with the exception of the last contracting block. This last block, as opposed to the original paper (Ronneberger, Fischer, and Brox 2015), keeps the number of filters equal. Consequently, in the last block the number of feature maps is also kept unchanged. The number of convolutional filters used is [16, 32, 64, 64] in this order, against [64, 128, 256, 512] of (Ronneberger, Fischer, and Brox 2015). Batch normalization is also performed in between the convolutional layers and ReLU, to reduces internal covariate shift and making the network more stable while training.
- **Expanding path**: it consists of 4 consecutive *expanding blocks*. Differently from the other blocks, these ones operate by first applying upsampling via bilinear interpolation with factor 2, as opposed to the transpose convolution (or "up-convolution") used in (Ronneberger, Fischer, and Brox 2015), with the same effect of doubling

the size of the image. Secondly, the result is concatenated with the corresponding cropped feature map from the contracting path. Lastly, this output is convolved similarly to how it is done in the *contracting blocks*, though now the number of convolutional filters halves at each step of the cascade. Also, as opposed to the original paper (Ronneberger, Fischer, and Brox 2015), the second *expanding block* maintains the same number of convolutional filters as the first. The number of convolutional filters used is [64, 64, 32, 16] in this order, against [512, 256, 128, 64] of (Ronneberger, Fischer, and Brox 2015). Again, batch normalization is actuated.

- **Bridge**: it is a block which is used to connect the two paths. In this implementation it is composed by a single block consisting of just the two convolutional layers as described before. Unlike in (Ronneberger, Fischer, and Brox 2015), the number of convolutional filters is kept equal, resulting in the number of feature maps being left unchanged, while originally the number would double.
- **Final layers**: after the last *expanding block*, 3 convolutional layers with a different number of convolutional filters of size (1x1) are applied in parallel on the previous output, resulting in 3 different outputs, 1 for each of the 3 features SIC, SOD, FLOE:
  - for SIC, we obtain an image with 10 feature maps, one for each SIC class
  - for SOD, we obtain an image with 5 feature maps, one for each SOD class
  - for FLOE, we obtain an image with 6 feature maps, one for each FLOE class

Training The training pipeline follows a general train-validation-test scheme, where ***cross-entropy*** loss function is used to calculate the error on each of the three predictions for the SIC, SOD and FLOE variables.

Algorithm 1: Training & Validation

$T, V \leftarrow \text{SPLIT}(dataset)$
$net \leftarrow UNet$
$vars \leftarrow SIC, SOD, FLOE$
${error^0}_{\text{train}} \leftarrow 0$
${error^0}_{\text{val}} \leftarrow 0$
$epoch \leftarrow 1$
while $epoch < numEpochs$ **do**
  ${loss}_{\text{sum}} \leftarrow 0$
  ${error^{epoch}}_{\text{train}} \leftarrow \text{TRAINING}(net, vars, T)$
  $\text{DISPLAY}({error^{epoch}}_{\text{train}})$
  ${error^{epoch}}_{\text{val}} \leftarrow \text{VALIDATION}(net, vars, V)$
  $\text{DISPLAY}({error^{epoch}}_{\text{val}})$
  if ${error^{epoch}}_{\text{val}} < {error^{epoch-1}}_{\text{val}}$ **then**
    $\text{SAVE}(net)$
  end if
  $epoch \leftarrow epoch + 1$
end while

Results Using this implementation of U-Net, we report the obtained results.

SCORE	SIC	SOD	FLOE
74.85	75.34	76.43	70.74

Table 1: Results obtained on the test set, following the evaluation measures as described in **Measures**

Proposed Segmentation

Alternative pre-processing

The current preprocessing suffers from problems in areas without features (e.g open water). We intend to find these areas by using a threshold and then substituting the pixel values with the following formula:

$$\sigma_{0i} = \frac{DN_i^2 - n_i}{A_i} \tag{5}$$

where DN_i is the image pixel value at location i, n_i is the provided noise data at location i, and A_i is the provided calibration coefficient at location i. The information about A_i and n_i are specific to the image and are provided directly by ESA.(Karvonenr 2017)(Guillaume Hajduch 2021)

Another problem that we found while analyzing images is that many of them are cut in half by the satellite's sensor due to quantization errors and noise, as shown on the image:

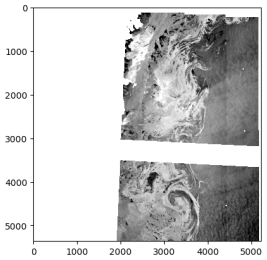

Figure 7: Example of a cut image

This artifact can be detected by using a sliding window, calculating the average of pixel intensities inside the window, checking whether the difference between the average and the pixel is more than 200 with the pixel having intensity equal to 255, and finally substituting the pixel value with the average value inside the window.

Alternative U-Net model

We are now going to present an alternative model to the current U-Net implementation that might allow to produce more satisfactory results. The network is called **DAU-Net**, and it was presented in (Ren et al. 2022).

Rationale As reported in **Figure 8**, DAU-Net has been shown outperforming U-Net on the task of semantic segmentation of sea ice.

Though the data provided and the classification tasks are ultimately different (camera vs SAR images, binary vs multi-class classification) we nevertheless believe that the architecture of the network is capable of yelding superior segmentation results to **Table 1**.

Data	Model	IoU	Accuracy	Precision	Recall
$Data_1$	$DenseNet_{FCN}$	91.62%	95.55%	94.28%	97.01%
	U-Net	87.18%	93.17%	93.73%	92.58%
	DAU-Net	**94.66%**	**97.22%**	**96.14%**	**98.40%**
$Data_2$	$DenseNet_{FCN}$	87.07%	94.20%	93.85%	92.33%
	U-Net	88.64%	95.04%	96.50%	91.58%
	DAU-Net	**89.60%**	**95.49%**	**97.24%**	**91.93%**
$Data_3$	$DenseNet_{FCN}$	89.35%	92.19%	93.07%	94.24%
	U-Net	90.78%	93.27%	94.21%	96.15%
	DAU-Net	**91.61%**	**93.95%**	**95.26%**	**95.17%**

Figure 8: Comparisons of results (Ren et al. 2022)

Architecture DAU-Net is an **encoder-decoder** network, which is heavily based on the U-Net implementation as in (Ronneberger, Fischer, and Brox 2015), but integrating the so called *dual-attention* mechanism, with the goal of *"enchancing the feature characteristics to improve the classification accuracy"*. We report the schema in **Figure 9**.

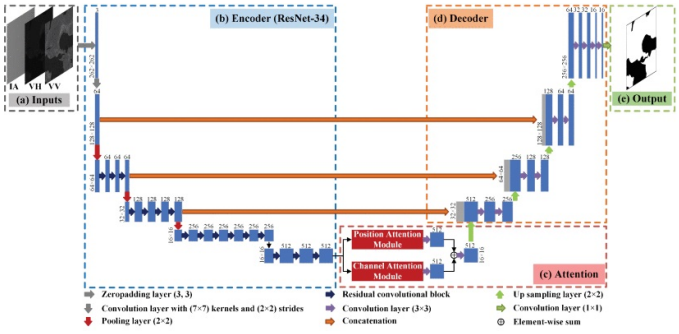

Figure 9: Architecture of DAU-Net (Ren et al. 2022)

The overall structure includes 5 parts:

- **Input**: Unlike the original network reported in (Ren et al. 2022) the input shall be a single scene, i.e. an image with as many channels as the scene variables chosen for training (if all are used, that will be 24).
- **Encoder**: as in (Ren et al. 2022), **ResNet-34** (He et al. 2016) shall be used. It is composed of
  - zero-padding layer with (3x3) kernel
  - convolutional layer with 64 filters with kernel size (7x7) and stride 2
  - 4 ResNet stages with respectively 3, 4, 6 and 3 *Residual convolutional blocks*. Thanks to the *skip connections* mechanisms introduced in these blocks, the network is able to learn the residual between the input and desired output, which can help overcoming issues like vanishing gradients. The number of convolutional filters in these stages is [64, 139, 256, 512] in this order.
- **Attention**: as in (Ren et al. 2022), we shall also add this new module as a *bridge* between encoder and decoder, *"to improve the feature representations of sea ice"*. It is composed by two algorithms:
  - **PAM**: it focuses on capturing spatial dependencies and optimizing local features. It captures global spatial dependencies by aggregating feature values across positions via weighted summation. It therefore produces an output feature map by combining local and global features. We report the schema in **Figure 10**.

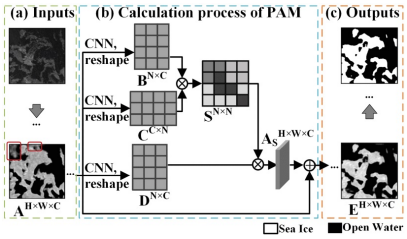

Figure 10: PAM processing (Ren et al. 2022)

  - **CAM**: it focuses on capturing channel dependencies and enhancing feature discriminability. It enhances feature discriminability by emphasizing channels with similar class responses via a weighting scheme. It therefore produces in output feature maps by weighting the original feature maps based on channel dependencies. We report the schema in **Figure 11**.

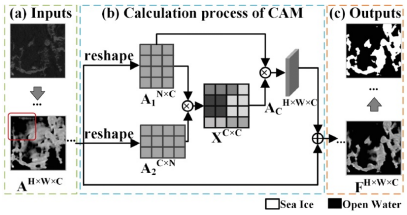

Figure 11: CAM processing (Ren et al. 2022)

The two outputs are then fused by summation using (3x3) filters in a convolutional layer.

- **Decoder**: as in (Ren et al. 2022) it shall consist of five blocks, each one containing one up-sampling layer with (2x2) filters and two convolutional layers with (3x3) filters. The number these filters is [256, 128, 64, 32, 16] in this order. The feature maps in the respective encoding block are cropped and concatenated to the outputs of the up-samplings in the corresponding decoding blocks, as in U-Net.
- **Output**: since our problem is a multi-classification problem (differently from (Ren et al. 2022)), the network shall be fine-tuned: we shall use the final layers as described in **U-Net** as the classification head.

Conclusions and future work

In this article, we have discussed the importance of developing an automatic and accurate sea ice mapping system, using Sentinel-1 SAR and AMSR-2 data, to analyze and forecast the effects of climate change on the Arctic region, in order to understand the impact of global warming and the health status of the Earth.

The manual method of ice mapping is time-consuming and limited in spatial and temporal coverage. Therefore, in

this paper we have reported the current segmentation approach used in the AutoICE challenge, involving the preprocessing of the data through SAR noise correction and the use of U-Net for the the actual segmentation.

Subsequently, we have presented a possible improvement to the preprocessing for a better quality image to highlight its main features and we also proposed an alternative network arhitecture, based on DAU-Net, which should perform better, as we can see in **Figure 8**.

Unfortunately, due to hardware limitations, we were not able to experiment on this methodology ourselves. However, we hope to be able to overcome these constraints in the near future in order to explore the potential of this methodology for the described problem, and others as well.

References

-. AMSR-2 overview. https://www.earthdata.nasa.gov/learn/find-data/near-real-time/amsr2.

-. *Sentinel-1 overview*. https://sentinels.copernicus.eu/web/sentinel/user-guides/sentinel-1-sar/overview.

Center, J. A. E. A. E. O. R. 2016. Data Users' Manual for the Advanced Microwave Scanning Radiometer 2 (AMSR2) onboard the Global Change Observation Mission 1st - Water "SHIZUKU" (GCOM-W1).

Guillaume Hajduch, M. B. 2021. Sentinel-1 Product Specification.

He, K.; Zhang, X.; Ren, S.; and Sun, J. 2016. Deep Residual Learning for Image Recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770–778.

Hersbach. -. ERA5 hourly data on single levels from 1940 to present. https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels?tab=overview.

Karvonenr, J. 2017. Baltic Sea Ice Concentration Estimation Using SENTINEL-1 SAR and AMSR2 Microwave Radiometer Data.

Korosov, A.; Demchev, D.; Miranda, N.; Franceschi, N.; and Park, J.-W. 2022. Thermal Denoising of Cross-Polarized Sentinel-1 Data in Interferometric and Extra Wide Swath Modes. *IEEE Transactions on Geoscience and Remote Sensing*, 60: 1–11.

Ren, Y.; Li, X.; Yang, X.; and Xu, H. 2022. Development of a Dual-Attention U-Net Model for Sea Ice and Open Water Classification on SAR Images. *IEEE Geoscience and Remote Sensing Letters*, 19: 1–5.

Ronneberger, O.; Fischer, P.; and Brox, T. 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation. arXiv:1505.04597.