Region-level Ground Truth can produce Pixel-level Segmentation for Sea Ice Concentration

Muhammed Patel[1] **Xinwei Chen**[1] **Linlin Xu**[1] **Yuhao Chen**[1]
K. Andrea Scott[2] **David A. Clausi**[1]
[1]Vision and Image Processing Group, Systems Design Engineering, University of Waterloo
[2]Department of Mechanical and Mechatronics Engineering, University of Waterloo

1 Extended Abstract

Sea ice concentration (SIC) refers to the percentage of sea ice in a given ocean area. Accurate high-detail SIC maps in polar regions are vital for a range of human activities (e.g., Arctic shipping) and the precise monitoring of polar sea ice changes [1]. However, the current method of creating SIC maps through manual expert annotations on synthetic aperture radar (SAR) imagery and other satellite data is both imprecise in terms of spatial resolution and time-intensive to prepare [2]. Although various convolutional neural network (CNN)-based techniques have emerged for automating the process of generating sea ice maps from SAR images, a major challenge lies in the absence of finely-grained pixel-level labels for training these models. This lack of detailed training data limits the ability of these methods to yield reliable high-resolution mapping outcomes.

To address this issue, this research generate predictions for pixel-level SIC by training a model to learn from region-level SIC ground truth data. Specifically, a novel regional loss function is developed described in Algorithm 1, which allows direct integration of the regional SIC values in ice charts to train a U-Net-based SIC estimation model. The network architecture of the proposed pipeline, built upon the U-Net framework [3], is illustrated in Fig. 1. This architecture is composed of four encoder blocks and four decoder blocks. The number of filters utilized in each convolutional or deconvolutional layer is represented by the corresponding digit beneath them in Fig. 1. To calculate the likelihood of each pixel belonging to the ice class, a linear layer followed by the sigmoid function is incorporated after the outputs of the final deconvolutional layer. This process also corresponds to pixel-level SIC estimation.

In order to circumvent the potential inaccuracies stemming from the translation of region-level SIC values to ground-truth pixel-level SIC values, we adopt a novel regional loss function during model training, which can be regarded as weakly supervised learning approach that empowers the CNN-based model to directly learn from the ground truth SIC values at the polygon level. Specifically, during each iteration of the training phase, the sigmoid function outputs for each input patch are transformed into predictions for polygon-based SIC, as depicted on the right side of Fig. 1. The boundaries of these polygons are determined based on the corresponding ground truth SIC at the polygon level. The predicted SIC for a given polygon is computed by summing the logits following the sigmoid function (which represents pixel-based SIC estimates) and then dividing by the total number of pixels within the polygon. Then, a polygon-based L2 loss function is introduced. This loss function operates on each patch and quantifies the squared error between the predicted polygon-based SIC and the actual polygon-level ground truth SIC from ice charts. The loss calculation is expressed as:

$$L_2 = \frac{1}{N}\sum_{i=1}^{N}(\hat{C}_i - C_i)^2, \tag{1}$$

Here, N represents the number of polygons within the input patch, $\hat{C}_i$ denotes the prediction for polygon-based SIC of the i^{th} polygon, and C_i signifies the actual SIC value at the i^{th} polygon level in the ice charts. This approach allows for learning from the polygon-level SIC labels without requiring direct pixel-level labels, mitigating the potential inaccuracies associated with the traditional pixel-based approach.

The proposed method and the U-Net-based model is trained and evaluated using the recently released AI4Arctic Sea Ice Challenge Dataset, consisting of 533 Sentinel-1 SAR scenes and ancillary data along with their corresponding ice charts. A comprehensive explanation of the dataset can be found in the user manual [4]. The testing results showcase the capability of the proposed model in generating high-resolution SIC maps at the pixel level that align well with existing ice charts and visual interpretations with one example depicted in Fig. 2. Compared to the results obtained from a benchmark U-Net trained with pixel-based ice chart labels and cross-entropy loss (Fig. 2(h)), the proposed model improves the mapping resolution significantly, as shown in Fig. 2(e).

Algorithm 1: Weak label loss

```
Data: output_logits, land_masks, polygon_icecharts,
polygon_codes_batch
Result: total_loss
total_loss = 0
batch_size = output.shape[0]
output_sigmoid = σ(output_logits)
for img_id in range(batch_size) do
    polygon_ids = np.unique(polygon_icecharts[img_id]);
    for polygon_id in polygon_ids do
        // create a mask of current polygon
        poly_mask = polygon_icecharts[img_id] == poly_id
        // create final mask by removing land
           area from the poly_mask
        final_mask = poly_mask ∩ land_masks[img_id]
        // Get the Ground truth ct (total ice
           concentration) for polygon_id using
           Lookup table
        c_t = LUT(polygon_codes[img_id], polygon_id)
        // AOI represents the area of interest
        AOI = output_sigmoid[img_id][final_mask]
        ĉ_t = AOI.sum()/AOI.numel()
        loss = ||c_t − ĉ_t||
        total_loss +=loss
return total_loss
```

In addition to visual interpretation, we also assess the model's performance quantitatively. Given the absence of actual pixel-based labels for Sea Ice Concentration (SIC), we employ rasterized polygon-based SIC maps for the purpose of comparison, as illustrated in Fig. 2(c). Similar to previous works [2], the accuracy of SIC estimation is gauged using the R^2 score, which measures how effectively the SIC predictions approximate the corresponding SIC labels. Given the significant visual dissimilarity between the polygon-based labels with their coarse resolution and the pixel-level SIC predictions, an R^2 score below 0.85 (0.837 for validation scenes and 0.737 for testing scenes) is in accordance with our expectations, as shown in Table 1. Nevertheless, the agreement between polygon-based labels and predictions (e.g., Fig. 2(g)) is evidenced by remarkably high R^2 scores of 0.978 and 0.961 for the validation and testing scenes, respectively.

To sum up, our proposed approach not only yields pixel-level SIC estimations but also generates pixel-level ice-water segmentation maps and enhanced polygon-based SIC maps. All these outputs can be transformed into valuable sea ice mapping products that greatly enhance Arctic-related activities, including ship navigation, and contribute significantly to climate-related research efforts.

Table 1: Numerical results obtained from validation and testing sets.

Comparison	Validation set		Testing set	
	RMSE	R^2 score	RMSE	R^2 score
Pixel-level SIC estimates VS Ground truth SIC	0.163	0.840	0.211	0.737
Polygon-level SIC estimates VS Ground truth SIC	0.140	0.978	0.146	0.961

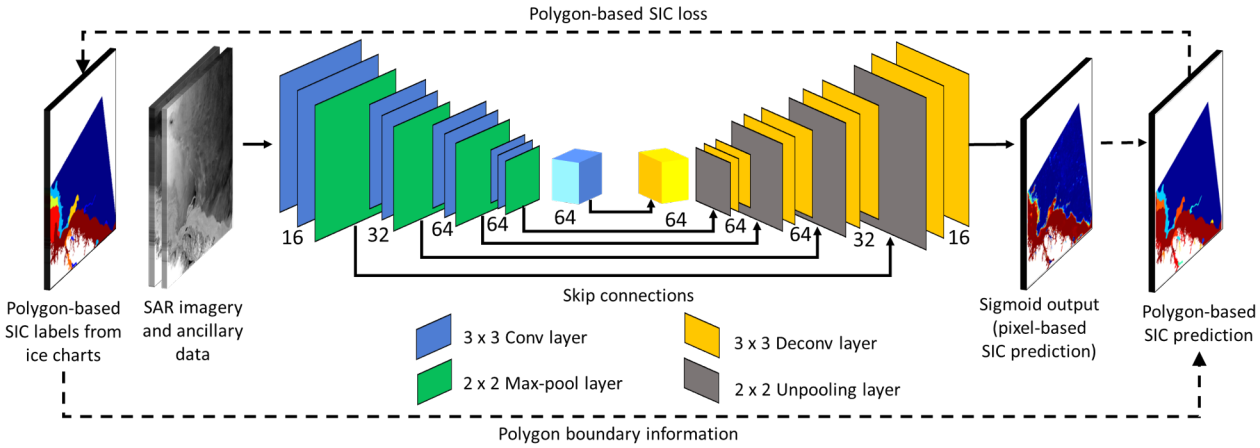

Fig. 1: The design of the suggested model for mapping SIC at the pixel level, which utilizes a U-Net architecture alongside the regional loss function (represented by dashed lines).

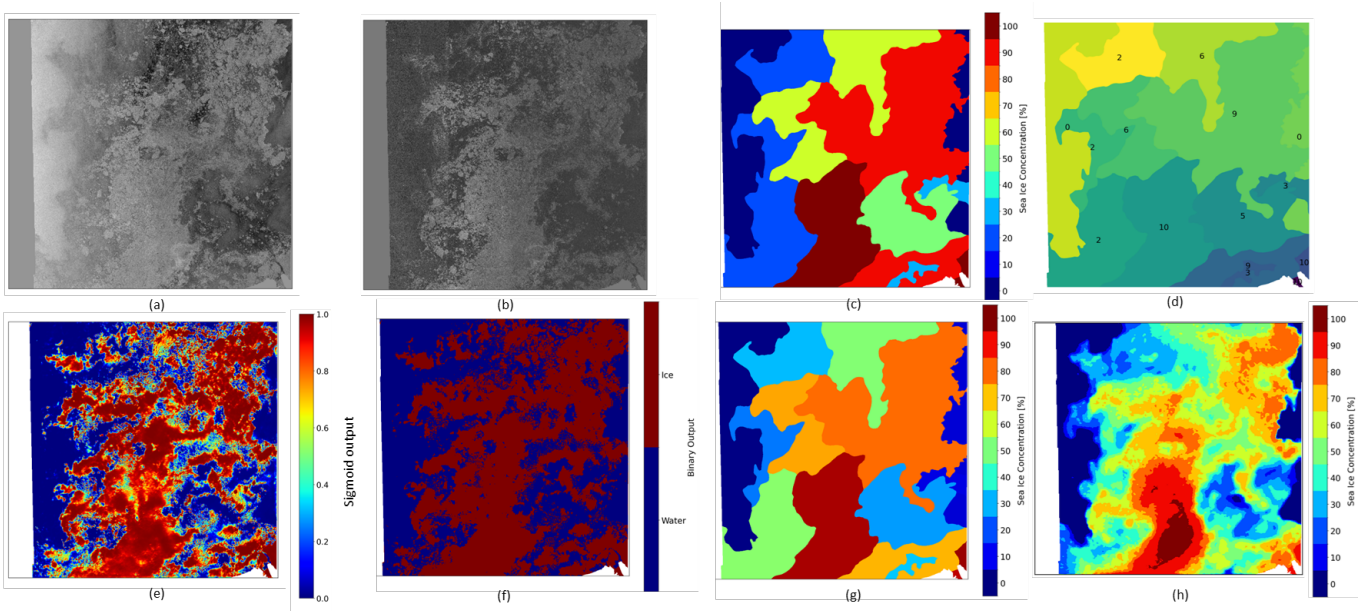

Fig. 2: A Sentinel-1 SAR scene with the identifier 20180716T110418_cis from the testing dataset is depicted in both HH (a) and HV (b) polarizations. Additionally, (c) illustrates the SIC map obtained from the ice chart, with each polygon indicated by a particular color (d). The sigmoid output of the proposed model, serving as a representation of the pixel-based SIC map, is presented in (e). The outcome of ice-water segmentation from (e) is shown in (f). (g) showcases the predicted polygon-based SIC map that combines (e) with the polygon boundary details from (d). It is apparent that (g) correlates well with (c). (h) is the prediction from a benchmark U-Net for comparison with (e), which demonstrates the significant improvement of spatial resolution using the proposed approach.

References

[1] A. Stokholm, T. Wulf, A. Kucik, R. Saldo, J. Buus-Hinkler, and S. M. Hvidegaard, "AI4SeaIce: Toward solving ambiguous SAR textures in convolutional neural networks for automatic sea ice concentration charting," vol. 60, pp. 1–13, 2022.

[2] A. Kucik and A. Stokholm, "AI4SeaIce: selecting loss functions for automated SAR sea ice concentration charting," *Sci. Rep.*, vol. 13, no. 1, p. 5962, 2023.

[3] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Intl. Conf. Med. Image Comput. Comput. Assist. Interv.* Springer, 2015, pp. 234–241.

[4] J. Buus-Hinkler, T. Wulf, A. R. Stokholm, A. Korosov, R. Saldo, L. T. Pedersen, and et al., "AI4Arctic Sea Ice Challenge Dataset," Technical University of Denmark. Collection, 2022. [Online]. Available: https://doi.org/10.11583/DTU.c.6244065.v2