



## Article

# E-MPSPNet: Ice–Water SAR Scene Segmentation Based on Multi-Scale Semantic Features and Edge Supervision

Wei Song <sup>1</sup>, Hongtao Li <sup>1</sup>, Qi He <sup>1,\*</sup>, Guoping Gao <sup>2</sup> and Antonio Liotta <sup>3</sup><sup>1</sup> College of Information Technology, Shanghai Ocean University, Shanghai 201306, China<sup>2</sup> College of Marine Ecology and Environment, Shanghai Ocean University, Shanghai 201306, China<sup>3</sup> Faculty of Computer Science, Free University of Bozen-Bolzano, 39100 Bolzano, Italy

\* Correspondence: qihe@shou.edu.cn

**Abstract:** Distinguishing sea ice and water is crucial for safe navigation and carrying out offshore activities in ice zones. However, due to the complexity and dynamics of the ice–water boundary, it is difficult for many deep learning-based segmentation algorithms to achieve accurate ice–water segmentation in synthetic aperture radar (SAR) images. In this paper, we propose an ice–water SAR segmentation network, E-MPSPNet, which can provide effective ice–water segmentation by fusing semantic features and edge information. The E-MPSPNet introduces a multi-scale attention mechanism to better fuse the ice–water semantic features and designs an edge supervision module (ESM) to learn ice–water edge features. The ESM not only provides ice–water edge prediction but also imposes constraints on the semantic feature extraction to better express the edge information. We also design a loss function that focuses on both ice–water edges and semantic segmentations of ice and water for overall network optimization. With the AI4Arctic/ASIP Sea Ice Dataset as the benchmark, experimental results show our E-MPSPNet achieves the best performance compared with other commonly used segmentation models, reaching 94.2% for accuracy, 93.0% for F-score, and 89.2% for MIOU. Moreover, our E-MPSPNet shows a relatively smaller model size and faster processing speed. The application of the E-MPSPNet for processing a SAR scene demonstrates its potential for operational use in drawing near real-time navigation charts of sea ice.

**Keywords:** ice–water segmentation; edge prediction; multi-scale attention mechanism; semantic features; ice concentration



**Citation:** Song, W.; Li, H.; He, Q.; Gao, G.; Liotta, A. E-MPSPNet: Ice–Water SAR Scene Segmentation Based on Multi-Scale Semantic Features and Edge Supervision. *Remote Sens.* **2022**, *14*, 5753. <https://doi.org/10.3390/rs14225753>

Academic Editors: Weimin Huang, Deepak R. Mishra and Ana C. Brito

Received: 30 September 2022

Accepted: 11 November 2022

Published: 14 November 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Sea ice is an important part of the Arctic cryosphere. In recent years, the constant rise in Arctic temperatures has led to a record low in sea ice extent and a prolonged thin ice period in the Arctic sea, providing opportunities for countries to use the Arctic shipping routes and develop Arctic resources [1]. However, the thinning of sea ice results in more dynamic ice cover and faster movement of the sea ice edge. In addition, the general retreat of the Arctic sea ice cover is exposing glacier fronts to open water, resulting in the calving of more icebergs. This will bring unpredictable risks to navigation. Near real-time information on sea ice is useful for route planning of ships and icebreakers. Such applications are often supported by sea ice charts or related reports published by national ice monitoring agencies in many countries, such as the United States National Ice Center (USNIC), the Canadian Ice Service, the Danish Meteorological Institute (DMI), etc. The ice charts depict ice condition and distribution in the form of ice eggs based on the experience of sea ice analysts and with reference to remote sensing data, which are time-consuming and labor-intensive. For ship route planning in ice zones, it becomes particularly important to design a fully automated method to achieve automatic segmentation between sea ice and open water based on remote sensing images.

Synthetic Aperture Radar (SAR) has the capability of all-weather continuous observation and high temporal and spatial resolution. In recent years, research based on SAR

images has been widely conducted for automatic object registration, change detection, and measurement [2–4]. Cross-polarization SAR images acquired in wide swath modes show many advantages in sea ice detection. Studies have shown that cross-polarization (horizontal-vertical (HV) and vertical-horizontal (VH)) is helpful in distinguishing between rough thin ice and open water [5,6], is insensitive to changes in sea surface roughness caused by incident angle and wind [7], and sea ice and open water have greater contrast in backscatter. Therefore, many sea ice studies tend to use dual polarization or full polarization data.

Early studies on sea ice segmentation in SAR images rely on texture features [8], which can be generated by algorithms such as the Gray-level Co-occurrence Matrix, Markov Random Fields, and Discrete Wavelet [9–12]. Combining texture features with machine learning algorithms such as Support Vector Machine (SVM) has presented excellent performance in sea ice segmentation [13,14] and has been successfully applied in an automatic sea ice segmentation system [15]. Decision tree and random forest have been used to map fast ice over the Antarctic area with good results [16]. Subsequently, artificial neural networks (NNs) are also frequently used [17,18]. Both SVM and NNs focus on the input of a large number of features to enhance texture information. These features are easily affected by noise, and the algorithm parameter adjustment process requires too much manual intervention and is affected by data and region. Thus, traditional segmentation methods work well on specific problems but have poor sea ice analytical ability in complex situations.

Typical image segmentation only segments pixels that have similar properties into regions, and the class of each region needs to be further processed. However, the region-based segmentation methods are not accurate enough to meet the needs of automated sea ice operations [19]. In contrast, semantic segmentation provides further understanding of the image, directly assigning a semantic label to each pixel in the segments. Segmentation methods based on deep learning still require a large amount of data support. A few public SAR image datasets have been developed in the field of sea ice detection with deep learning [20–23]. Khaleghian et al. [24] proposed a dataset containing six classes of ice types and ice edge analysis. Song et al. [25] established a dataset to facilitate sea-ice classification with spatial and temporal information. However, due to the process of making sea-ice semantic segmentation datasets being complicated and time consuming, only the AI4Arctic/ASIP Sea Ice Dataset [26] is applied to scene segmentation. In ice-water segmentation in this paper, the semantics mainly refer to sea ice and open water. Convolutional neural networks (CNN) have become a mature method for solving semantic segmentation, because they are less sensitive to pixel-level details and produce ice concentration that is less noisy and in closer agreement with that from image analysis charts [27]. Despite growing interest in deep learning techniques for sea ice segmentation, several recent studies on ice-water segmentation [28,29] and ice type classification [30–33] using dual-polarized C-band SAR have suggested a high difficulty in accurately predicting the ice types with low ice concentrations, especially near the edges of sea ice and open water.

To address the problem of edge detail loss, researchers have considered training a multi-task network or designing a special network module to obtain more accurate semantic segmentation results. Chen et al. [34] proposed to optimize the segmentation results obtained by a full convolutional network in detail through conditional random fields (CRF), but this processing method only uses low-level features such as texture information to correct the segmentation results. On this basis, they further proposed an edge-preserving filtering method with domain transform instead of CRF to improve the accuracy of object localization in semantic segmentation [35]. Takikawa et al. [36] constructed a semantic segmentation model with parallel CNN that treats edge information as a separate processing branch and processes the information in parallel with the segmentation branch, which can produce clearer object boundaries. In summary, edge information can indeed help the segmentation network to locate the objects precisely.

These common image processing methods are not completely applicable to remote sensing images, because in the remote sensing images, the actual edges are always un-

certain to some extent due to the sensor errors and atmospheric distortions. Most sea ice images do not have strictly ideal edges, and their boundaries are not always lines in the geometric sense, but transitions with a certain width, and the actual edges basically present uncertainty and fuzziness. Uncertainty has caused slow processes and low accuracy in edge detection tasks, and efforts have previously focused on how to reduce the uncertainty of edge information in edge detection tasks, and in recent years fuzziness edges have attracted the attention of experts [37,38].

In order to improve the accuracy of ice–water segmentation of SAR images, this paper designs an ice–water scene segmentation network combining multi-scale semantic features and edge supervision, with the assumption that the inherent correlation exists between edge detection and semantic segmentation.

A concise summary of our contributions are as follows

1. We propose an ice–water scene segmentation network, E-MPSPNet. It fuses the multi-scale features with scale-wise attention to produce an ice–water segmentation feature map and combines the segmentation feature map with an edge feature map to achieve better segmentation accuracy. The proposed E-MPSPNet performs well with a relatively higher efficiency compared to mainstream segmentation networks, U-Net, PSPNet, DeepLabV3, and HED-UNet.
2. To eliminate the uncertainty of ice–water segmentation edges, we design an edge supervision module based on the idea of deep supervision. It plays a two-fold role: directly predicting the ice–water edge feature map and providing additional edge constraints to feature extraction. This module helps capture the edge characteristics of ice and water more effectively.
3. We design a joint loss function that combines the edge loss and the semantic loss for the network optimization and take into account the problem of class imbalance between edge pixels and non-edge pixels.

This paper is organized as follows. Section 2 presents the secondary process of the dataset. Section 3 presents the E-MPSPNet network architecture. Section 4 presents the baseline experimental results of our method and other commonly used segmentation models on the constructed dataset. Section 5 discusses the applications and limitations of our model in coastal environments. Section 6 is the conclusion.

## 2. Study Area and Data

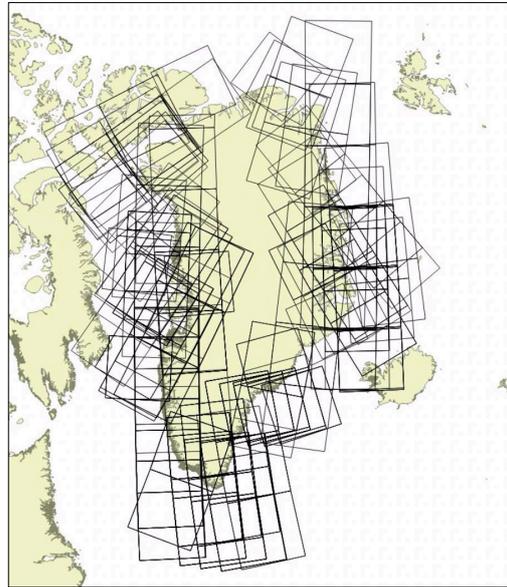
### 2.1. Data Source

In this study, we employed a public dataset AI4Arctic/ASIP Sea Ice Dataset—version 2 as the data source and re-produced training and testing datasets on the basis of it.

The AI4Arctic/ASIP Sea Ice Dataset has 461 views covering different waters around Greenland from March 2018 to May 2019 at different sea and ice conditions in different seasons. The spatial distribution of these views can be seen in Figure 1. Each file in network Common Data Form (netCDF) format contains mainly dual-polarized (HH and HV) Extra Wide swath mode (EW) SAR images and low-resolution AMSR2 microwave radiometer data, as well as the corresponding sea ice charts. The SAR images were from two C-band satellites, Sentinel-1 A and B, with a standard strip width of 400 km, a resolution of about 90 m, a pixel spacing of 40 m, and a range of incidence angle of  $18.9^{\circ}$ – $47.0^{\circ}$ . Additionally, the file provides auxiliary information such as the distance between each pixel and the land. We only used the dual-polarized SAR images and the corresponding sea ice charts for ice–water segmentation.

The ice charts in the AI4Arctic/ASIP Sea Ice Dataset are from the ice chart archives produced by DMI Greenland Ice Service. They are provided in the Sea Ice Georeferenced Information and Data (SIGRID3) ice code (JCOMM Expert Team on Sea Ice 2014), following the World Meteorological Organization (WMO) standard. Together with the dataset, two noise correction schemes are provided, which are from the European Space Agency (ESA) and the Nansen Environmental and Remote Sensing Center (NERSC), respectively.

We chose to go through the NERSC noise correction scheme [39], because it is more effective for HV polarized noise floor stripes removal [28].



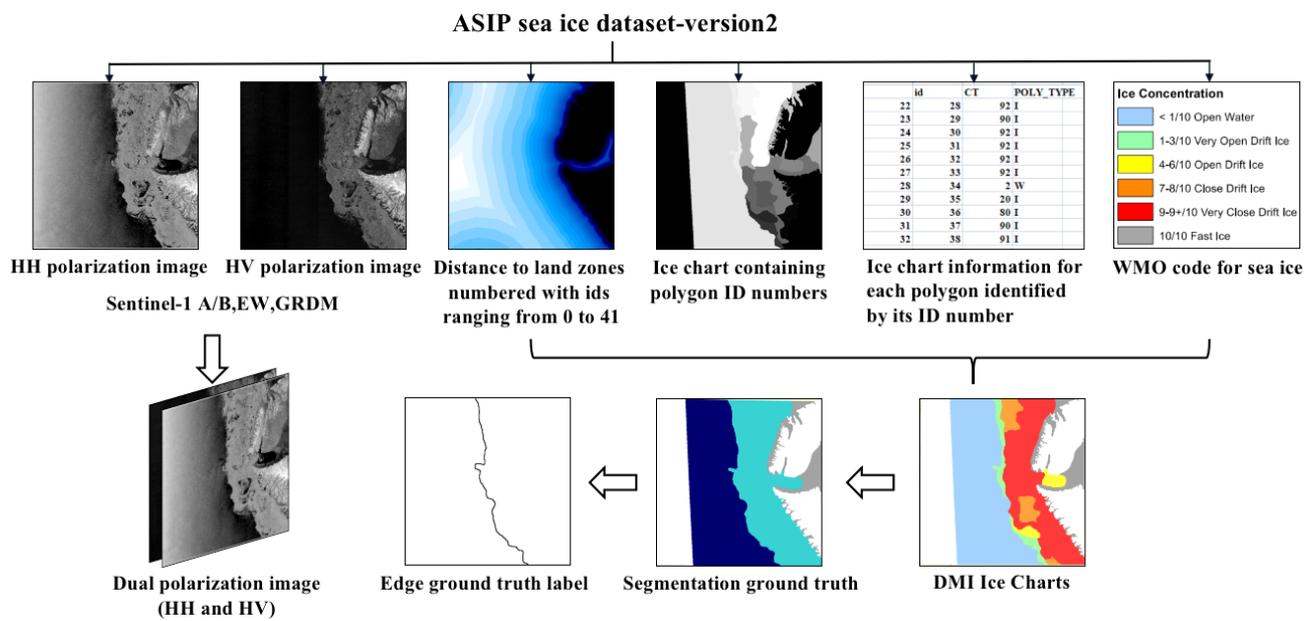
**Figure 1.** Spatial distribution of the scenes in the AI4Arctic/ASIP Sea Ice Dataset.

## 2.2. Dataset Processing

The characteristics and extent of the marginal area are very important for navigation or exploration in or near ice areas. Although icebreakers are able to pass through low-concentration ice areas, sea ice changes rapidly. In order to depict more accurate ice–water edges in ice maps, it is beneficial to treat ice of different concentrations as “sea ice” for safe navigation. So, we secondary processed the AI4Arctic/ASIP sea ice dataset and created an ice–water segmentation dataset using SAR images and DMI ice maps as auxiliary information.

The steps of processing ground truth labels for ice and water segmentation are as follows and are demonstrated in Figure 2.

1. Identify polygons for ice–water segments. The AI4Arctic/ASIP Sea Ice Dataset contains a DMI ice chart for the area corresponding to each SAR image. Each polygon in the ice chart is recorded in a table with its unique ID and the code of ice concentration in SIGRID3. To generate ice and water polygons for this study, we simplify the ice concentration SIGRID3 codes into two categories, as shown in Table 1. Label “0” defines pixels with ice concentrations less than 1/10 as sea water (according to the WMO’s definition), and label “1” defines pixels with ice concentrations in the range 1–10/10 as sea ice.
2. Identify land masks. According to the distance information between pixels and the land zones provided in the netCDF files, the pixels containing land are used as masks. The parts of the SAR image outside the ice chart area are also considered as masks. The pixels being masked are not used for the training of the model.
3. Generate ground truth labels. After completing the above two steps, the ground truth maps for ice–water segmentation can be generated. Then, a Sobel operator is run on the ice–water segmentation maps to produce ice–water boundaries. The produced edge ground truth map has the value “zero” for the ice–water boundaries, and it will be used for edge supervision in this paper.



**Figure 2.** Preprocessing for the dataset example case: 20180421T080346\_S1B\_AMSR2\_IcechartGreenlandCentralEast.

**Table 1.** Simplified labeling scheme reducing DMI codes to two categories.

Definition, Concentration	Sigrid3 Code (CT, CA, CB, and CC)	Category	Label
Ice Free	00	Sea corresponds to the concentration of codes < 1/10	0
Less than 1/10	01		
Bergy water	02 *		
1/10	10	Ice corresponds to the concentration of codes 1–10/10	1
2/10	20		
3/10	30		
4/10	40		
5/10	50		
6/10	60		
7/10	70		
8/10	80		
9/10	90		
9+/10 (95%)	91 **		
10/10	92		

\* The category “Bergy water” is used for open sea (water category) in the DMI ice charts. The category “Ice Free” is not used in the DMI ice charts, since icebergs can appear everywhere in Greenland waters. \*\* The category “9+/10” is used in the DMI ice charts for sea ice that is fully compacted, but not fast ice (100% ice).

Since each SAR image is too large to be directly fed into the network for training and too much detail information will be lost if it is downscaled, this paper crops the SAR image into patches of 800 × 800 pixels with a non-overlapping sliding window. The set of sub-regions is filtered to exclude the sub-regions containing mask pixels. To increase the generalization ability and robustness of the network model, the diversity of the dataset can be increased by data augmentation. We expand the data by simultaneously flipping horizontally, flipping vertically, and mirroring diagonally for each SAR image and its corresponding label image. The dataset created per the methodology in this section generated 8244 samples, which were then normalized and divided into three parts according to 7:2:1 for the training set, test set, and validation set.

### 3. Methodology

#### 3.1. Overview of the Network Structure

This paper builds an ice and water segmentation network, E-MPSPNet, that integrates the sea ice edge information and semantic features, and the overall structure of the network is shown in Figure 3. E-MPSPNet consists of a backbone network, a multi-scale feature fusion module (MFFM) and an edge supervision module (ESM). The backbone network is mainly used to obtain shallow and deep semantic feature maps for subsequent feature fusion. The MFFM introduces a multi-scale attention mechanism to fuse the features at different levels of details to semantic segmentation feature maps, making better use of global and local information. The ESM imposes supervision on each of the feature extraction layers in the backbone network by backpropagating edge prediction errors. Meanwhile, it generates predicted edge feature maps. Finally, the fused semantic segmentation features and the predicted edge feature maps are combined to improve the network segmentation effect.

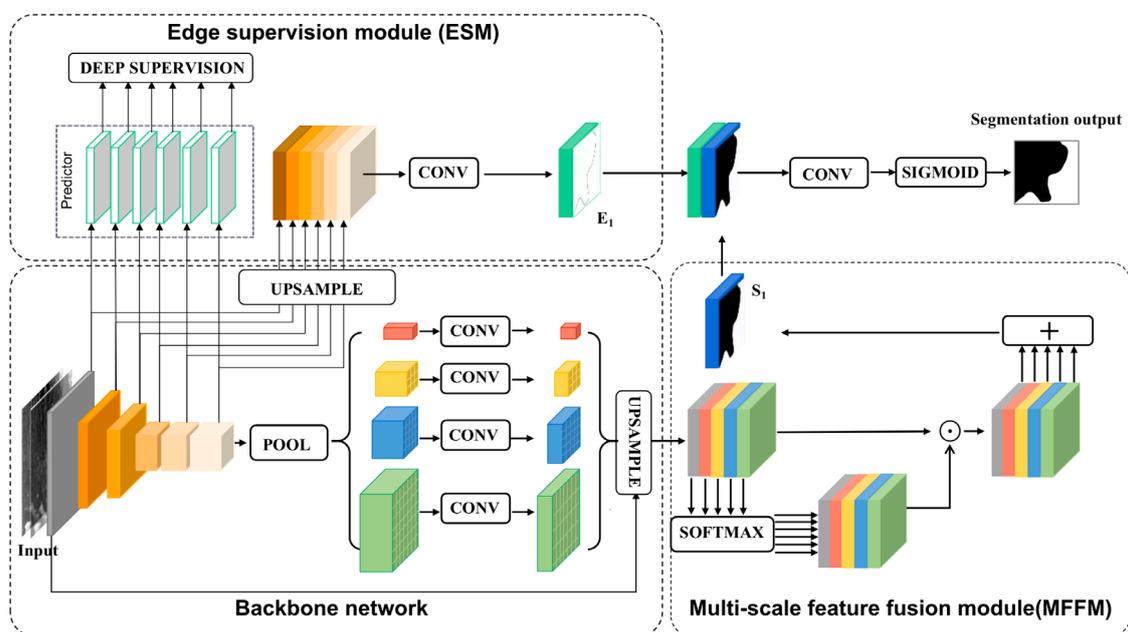


Figure 3. The E-MPSPNet architecture.

#### 3.2. Backbone Network

The E-MPSPNet backbone network is based on PSPNet [40] with different structure parameters in order to adapt to our SAR image setting. In the backbone, ResNet50 [41] is adopted as a feature extractor. The extracted features are processed through a pyramid parsing module (PPM), which consists of different sizes of pooling kernels to output different scales of feature maps.

The feature extractor is composed of five residual blocks (RESblocks), whose network structure is shown in Table 2. The feature extractor reduces the resolution of the input image ( $800 \times 800$ ) to  $1/2$ ,  $1/4$ , and  $1/8$  via three-time pooling. Together with the input image, the feature maps output by the five RESblocks provide six side-layer inputs for the ESM module.

**Table 2.** Convolutional layer structure of the backbone network.

Layer Type	Kernel Size	Channels	Stride	Output Size
input	$3 \times 3$	64	1	$800 \times 800 \times 64$
conv1	$7 \times 7$	128	2	$400 \times 400 \times 128$
conv2	Max pooling	128	2	$200 \times 200 \times 128$
	$\begin{bmatrix} 1 \times 1 \\ 3 \times 3 \\ 1 \times 1 \end{bmatrix} \times 3$	256	1	$200 \times 200 \times 256$
conv3	$\begin{bmatrix} 1 \times 1 \\ 3 \times 3 \\ 1 \times 1 \end{bmatrix} \times 4$	512	1	$100 \times 100 \times 512$
conv4	$\begin{bmatrix} 1 \times 1 \\ 3 \times 3 \\ 1 \times 1 \end{bmatrix} \times 6$	1024	1	$100 \times 100 \times 1024$
conv5	$\begin{bmatrix} 1 \times 1 \\ 3 \times 3 \\ 1 \times 1 \end{bmatrix} \times 3$	2048	1	$100 \times 100 \times 2048$

### 3.3. Edge Supervision Module

The ESM is constructed based on depth supervision [42]. It has two objectives: The first is to directly provide edge prediction information for ice–water segmentation. The second is to provide additional constraints to each feature extraction layer of the backbone network so that the extracted features can capture the edge characteristics of ice and water more effectively.

We denote our input training data set by  $S = \{(x_i, y_i), i = 1 \dots N\}$ , where  $x_i \in \mathbb{R}^{h \times w \times c}$  denotes the input image, in which  $h, w$  are the length and width of the feature map, respectively, and  $c = 2$  indicates the channels of HH and HV polarizations and  $y_i \in \mathbb{R}^{h \times w}$  denotes the corresponding ground truth binary edge map for image  $x_i$ . We consider each sample from a holistic perspective, so the subscript  $i$  is dropped for notational simplicity.

$x$  leads to  $M$  side output layers from the feature extraction module of the backbone network, and for each layer  $m = 1 \dots M$ , there is:

$$F^{(m)} = f(W^{(m)} \times F^{(m-1)}), F^{(0)} \equiv x \quad (1)$$

where  $W^{(m)}$  denotes the weight of the feature extraction layer, and  $F^{(m-1)}$  denotes the feature map extracted by the  $m - 1$  layer.

In ESM, each side output layer is connected to a  $1 \times 1$  convolution layer, whose corresponding weight is defined as  $w = (w^1 \dots w^M)$ . These side layers produce ice–water edge prediction results by supervised learning with the edge ground truth label (scaled by  $1/2, 1/4,$  and  $1/8$  scales correspondingly) and exert external constraints on the feature extraction intermediate layers in the backbone network via error backpropagation, which enhances the expression of the ice–water edge in the feature maps. Meanwhile, the feature maps output from all the side layers are upsampled to the original image size and stacked, and then the edge prediction map  $E_1$  is generated by  $1 \times 1$  convolution.

The objective function of the whole ESM module consists of two parts, denoted as:

$$J(W) = J_{side}(W) + J_{out}(W) \quad (2)$$

$$J_{side}(W) = \sum_{m=1}^M \alpha_m \left[ \|w^{(m)}\|^2 + \ell_{side}(W, w^{(m)}) \right] \quad (3)$$

$$J_{out}(W) = \|w^{(out)}\|^2 + \ell_{fuse}(W, w^{(out)}) \quad (4)$$

where  $w^{(out)}$  denotes the weight of the output layer of the edge prediction map  $E_1$ ,  $\ell_{fuse}(W, w^{(out)})$  denotes the loss function of multilayer fusion, and  $\ell_{side}(W, w^{(m)})$  de-

notes the pixel-level loss function of the image output from the side layer. The specific definition of the loss function is given in Section 3.5.

### 3.4. Multi-Scale Feature Fusion Module

When distinguishing the boundary between ice and water, the model is expected to obtain a more accurate description of the ice segmentation line by learning from a high-resolution image. Learning lower-resolution-level image information farther away from the boundary can provide a more comprehensive evaluation of the scene, leading to better segmentation in these regions. Therefore, in this paper, attention weight is assigned to the multi-scale feature images to fully emphasize the details of the ice and water classes at different scales. Then, simple summation processing is used to integrate the multi-scale feature images for ice–water segmentation prediction.

The upsampled multi-scale pooling feature from the backbone network is denoted as  $P = \{p'_j\}$ ,  $p'_j \in \mathbb{R}^{h \times w \times c}$ , and  $j = 1 \dots 5$  represents the serial number of different pooling feature maps. The MFFM module first performs a channel downscaling operation  $r(\cdot)$  on each feature map, which is implemented by a  $1 \times 1$  convolution operation. The reduced dimensional feature is denoted as  $p_j = r(p'_j) \in \mathbb{R}^{h \times w \times 2}$ . Then, softmax function is applied to the feature map  $p_j$  to transform it into an attention weight matrix, and the dot product of the feature map  $p_j$  and the attention weight matrix are calculated. Finally, semantic feature  $S_1$  is obtained by summing the multiple feature maps. The process is formulated as follows:

$$h_j = p_j \odot \text{softmax}(p_j) \quad (5)$$

$$S_1 = \sum_j h_j \quad (6)$$

In the end, the E-PSPNet network fuses the semantic features  $S_1$  and the edge features  $E_1$  by stacking and  $1 \times 1$  convolution and outputs the final ice–water segmentation prediction result by pixel-level classification with the sigmoid function.

### 3.5. The Joint Loss Function

The loss function is defined as the joint loss  $L$  of the MFFM and ESM, and the expression is:

$$L = l_{seg} + l_{edg} \quad (7)$$

where  $l_{seg}$  represents the semantic loss, the loss value predicted by MFFM, and  $l_{edg}$  is the loss value predicted by EEM.

The semantic loss  $l_{seg}$  is calculated by mixing a focal loss [43] and a MIOU loss function, which is defined as:

$$l_{seg} = l_{Focal} + l_{MIOU} \quad (8)$$

In the sea ice segmentation task, it is difficult to distinguish low-concentration ice from water, and the features of “edge” and “non-edge” are highly unbalanced. Focal loss reduces the weight of the easy sample (i.e., water), so that the model can focus more on a difficult sample (i.e., ice) during training. The expression  $l_{Focal}$  is in (9), where  $\gamma = 2$  and  $(i, j)$  are the coordinates of an image position,  $(i, j) \in (W, H)$ ;  $W, H$  are the width and height of the image, respectively;  $\hat{y}_{i,j}$  is the pixel value at the prediction map  $(i, j)$ ,  $\hat{y}_{i,j} \in (0, 1)$ ; and  $y_{i,j}$  is the pixel value at label map  $(i, j)$ ,  $y_{i,j} \in \{0, 1\}$ .

$$l_{Focal} = - \sum_{i,j} \left( y_{i,j} (1 - \hat{y}_{i,j})^\gamma \log(\hat{y}_{i,j}) + (1 - y_{i,j}) (\hat{y}_{i,j})^\gamma \log(1 - \hat{y}_{i,j}) \right) \quad (9)$$

The MIOU loss supervises the network learning by measuring the mean intersection over union between the predicted segmentation and the ground truth for multiple classes. In the SAR image, the proportion of sea ice and sea water is often unbalanced. The MIOU

loss function can give equal treatment to ice and water, and thus effectively improve the training difficulties caused by the class imbalance.

The loss function of MIOU is defined as:

$$l_{MIOU} = 1 - \sum_{i,j} \frac{y_{i,j}\hat{y}_{i,j}}{y_{i,j} + \hat{y}_{i,j} - y_{i,j}\hat{y}_{i,j}} \quad (10)$$

The loss for the EEM module,  $l_{edg}$ , is defined as the average loss of the six side outputs and the final edge prediction loss, as shown in (11)

$$l_{edg} = \frac{1}{M-1} \sum_{M=1}^{M-1} l_{side}^{M-1} + l_{fuse} \quad (11)$$

where  $M = 6$  is the total number of output layers. Both  $l_{side}$  and  $l_{fuse}$  are in the form of focal loss, as (9). The focal loss has a modulating factor  $(1 - \hat{y}_{i,j})^\gamma$ . For easily classified samples, the factor tends to 0, which has little influence on the weight of the loss function; but, for hard samples, the weight in the loss function increases, allowing the network to focus on “ambiguous” samples at the edge of the sea ice. Thus, the  $l_{edg}$  based on focal loss is not only suitable for extracting uncertain edge features of sea ice, but also improves the generalization ability of the model and helps the network training to focus more on the optimization of the edges.

## 4. Experiment and Analysis

### 4.1. Experimental Environment and Settings

The experiments in this paper were all performed on a GeForce RTX 2080 Ti GPU. The PyTorch deep learning framework was used to implement the algorithm and the Adam optimizer [44] was used, with an initial learning rate set to  $1 \times 10^{-2}$  and an exponential learning rate decay of 0.9 applied after each epoch.

### 4.2. Evaluation Indicators

Accuracy of pixel-level ice–water classification is used for evaluating the overall performance of the models. At the same time, in order to avoid the limitation of accuracy in class-imbalanced classification, this paper also uses F-score and mean intersection over union (MIoU) as the evaluation indicators for pixel-level semantic segmentation. All these metrics can be obtained from the confusion matrix using the following four statistical measures: True Positives (TP), False Positives (FP), True Negatives (TN), and False Negatives (FN).

Accuracy is denoted as the proportion of correctly predicted pixels to the total number of pixels:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (12)$$

F-score is calculated by the harmonic mean of recall and precision, as in (13)–(15). The precision and recall can indicate the correctness and completeness of the segmented regions, respectively.

$$Precision = \frac{TP}{TP + FP} \quad (13)$$

$$Recall = \frac{TP}{TP + FN} \quad (14)$$

$$F - Score = 2 \times \frac{precision \times recall}{precision + recall} \quad (15)$$

MIoU is a commonly used metric in remote sensing image segmentation, as in (16), and  $N_{class} = 2$  indicates the ice and water. It measures the average of the overlapping ratio between the ground truth and the prediction over all classes.

$$MIoU = \frac{1}{N_{class}} \sum_{i=1}^{N_{class}} \frac{TP(i)}{TP(i) + FP(i) + FN(i)} \quad (16)$$

All the three evaluation indicators are between 0 and 1 in value, and the higher the value, the more accurate the segmentation result.

#### 4.3. Edge Supervision Module

##### 4.3.1. Comparison with Different Segmentation Models

For comparative analysis, we chose four deep learning segmentation models: U-Net, PSPNet, DeepLabV3, and HED-UNet [23]. The U-Net, PSPNet, and DeepLabV3 are well-known segmentation models, which have also been used for remote sensing image segmentation [45–47]. In particular, our model uses the PSPNet as the backbone network. Thus, it is reasonable to compare with it. The HED-UNet combines semantic segmentation and edge detection and was applied for Antarctic coastline detection. Our model has a similar idea to HED-UNet.

We evaluate these models' performance from both quantitative and qualitative aspects. Quantitative analysis is based on accuracy, F-score, and MIoU, and qualitative analysis is based on the observation of the segmentation effect in the predicted maps.

As shown in Table 3, the ice and water segmentation network designed in this paper has better performance on the test set than U-Net, PSPNet, DeepLabV3, and HED-UNet. The U-Net model fuses the deep and shallow ice–water feature maps and performs upsampling to solve the loss of feature information caused by the deepening of the network, but the prediction results are still misclassified and incomplete, and the prediction accuracy has room for improvement. DeepLabV3 expands the receptive field through expansive convolution to obtain more context information, and its performance is significantly better than U-Net in terms of indicators. PSPNet further improves network performance by effectively extracting multi-scale features through its pyramid pooling module. HED-UNet is built on U-Net, with a 4.8% higher MIoU, 2.9% higher F-score, and 2.9% higher accuracy than U-Net. This proves that the addition of edge information can improve the segmentation accuracy of the semantic segmentation task.

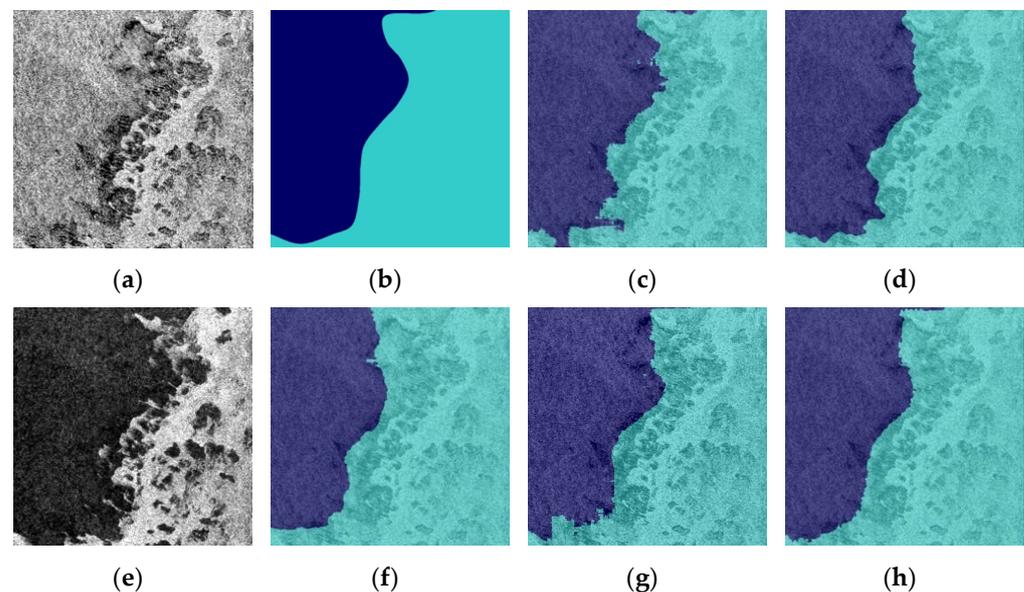
**Table 3.** Comparison results of different model segmentation.

Methodology	Accuracy	F-Score	MIoU
UNet	0.903	0.881	0.822
DeepLabV3	0.928	0.919	0.864
HED-UNet	0.932	0.910	0.870
PSPNet	0.932	0.921	0.873
E-MPSPNet	<b>0.942</b>	<b>0.930</b>	<b>0.892</b>

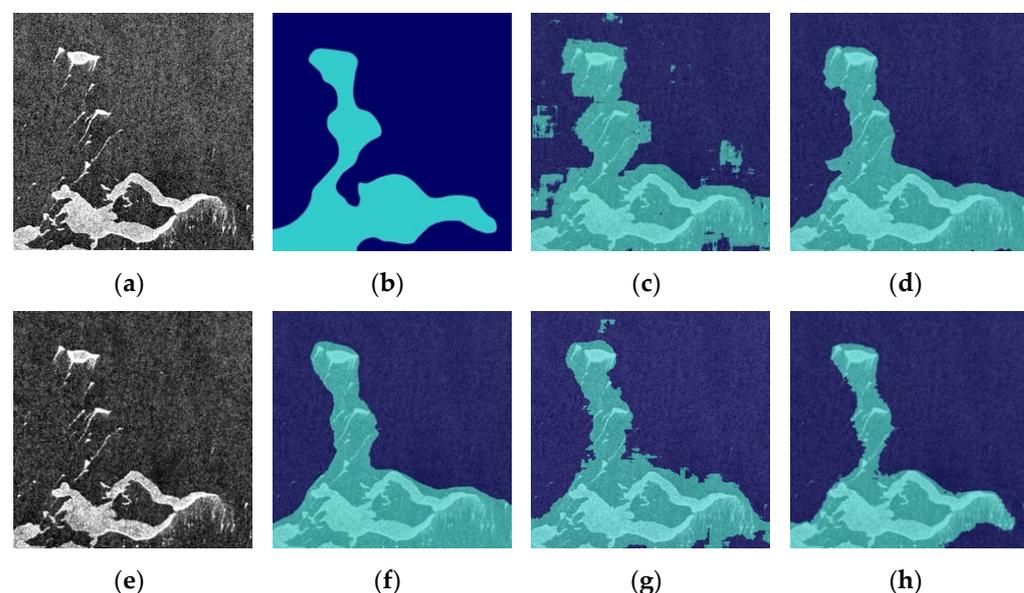
Our E-MPSPNet reaches the highest for all the evaluation metrics: accuracy of 0.942, F-score of 0.930, and MIoU of 0.892. Compared with the better-performed PSPNet network, the design of the edge supervision module and multi-scale feature fusion module in E-MPSPNet can prompt the network to pay more attention to edge details during the training process, so as to obtain better segmentation performance.

Qualitatively, Figures 4 and 5 demonstrate the segmentation results of different models for two sub-regions of SAR images selected from the test set. Figure 4b shows the ground truth segmentation maps for the two SAR images, where dark blue is sea water and light blue is sea ice. The predicted segmentation map by each model is overlaid with a transparent ratio of 0.7 on the original sub-region SAR image, in order to show the

segmentation effect visually. The rest of the demonstration figures in this paper apply the same overlaying method.



**Figure 4.** Example 1 of the segmentation results of a sub-region SAR image, the edge area of ice water is mostly low concentration sea ice with open water gaps. (a) The original SAR image (HH), (b) the ground truth of semantic segmentation (dark blue is sea water and light blue is sea ice), (c) the prediction result of U-Net, (d) the prediction result of DeepLabV3, (e) the original SAR image (HV), (f) the prediction result of PSPNet, (g) the prediction result of HED-UNet, (h) the prediction result of E-MPSPNet.



**Figure 5.** Example 2 of the segmentation results of a sub-region SAR image with open water inside the sea ice. (a) The original SAR image (HH), (b) the ground truth of semantic segmentation (dark blue is sea water and light blue is sea ice), (c) the prediction result of U-Net, (d) the prediction result of DeepLabV3, (e) the original SAR image (HV), (f) the prediction result of PSPNet, (g) the prediction result of HED-UNet, (h) the prediction result of E-MPSPNet.

In the SAR image (HH) shown in Figure 4a, the SAR image (HV) shown in Figure 4e, the edge area of ice and water is mostly low concentration sea ice with open water gaps. As seen in Figure 4c, U-Net can only vaguely predict the extent of sea ice, and the part of

sea ice at the edge is incorrectly classified as seawater. According to the prediction results of DeepLabV3 and PSPNet in Figure 4d,f, respectively, although a relatively complete region can be segmented, it is easy to incorrectly classify open-water pixels as sea ice pixels, which is not conducive to subsequent edge localization. The misclassification effect of HED-UNet (Figure 4g) at the edge of sea ice has been improved to some extent, but there are still a lot of inaccurate classifications at the edge. In comparison, the detection effect of sea ice by E-MPSPNet (Figure 4h) is closer to the true value, and the edge information can be retained better.

Figure 5a depicts sea ice with no clear boundary and open water within the sea ice. The U-Net segmentation result (Figure 5c) reveals obvious mis-segmentations and uneven sea ice contours, as well as a severe lack of edge information. The result of DeepLabV3 (Figure 5d) has a much better segmentation effect, but the edges are not smooth enough. The prediction results of PSPNet and HED-UNet in Figure 5f,g have fewer mis-segmented sea ice pixels, but like the previous two models, there is no effective distinction between ice and water in the lower left corner of the region, resulting in an incomplete presentation of the sea ice contours. This indicates that HED-UNet does not have a good segmentation effect on sea ice images with complex edges. E-MPSPNet (Figure 5h) shows clear segmentation regions of ice and water. Although there is a certain deviation from the ground true map (Figure 5b), the edge details are better grasped.

Our network mainly targets the edges between ice and water. Because the number of edge pixels is much smaller than the number of total image pixels, the advantage of E-MPSPNet over other models is small based on the evaluation metrics. However, in observing the segmentation effect of the prediction results, it can be seen that our proposed method has obvious improvements at the ice and water boundaries as well as for the overall segmentation of complex ice–water scenes. It needs to be noted that E-MPSPNet has a rough segmentation problem in the transition regions between open water and low-concentration of sea ice for the scene like Figure 5. More discussion about this problem is introduced in Section 5.3.

#### 4.3.2. Influence of MFFM and EEM on Network Segmentation Performance

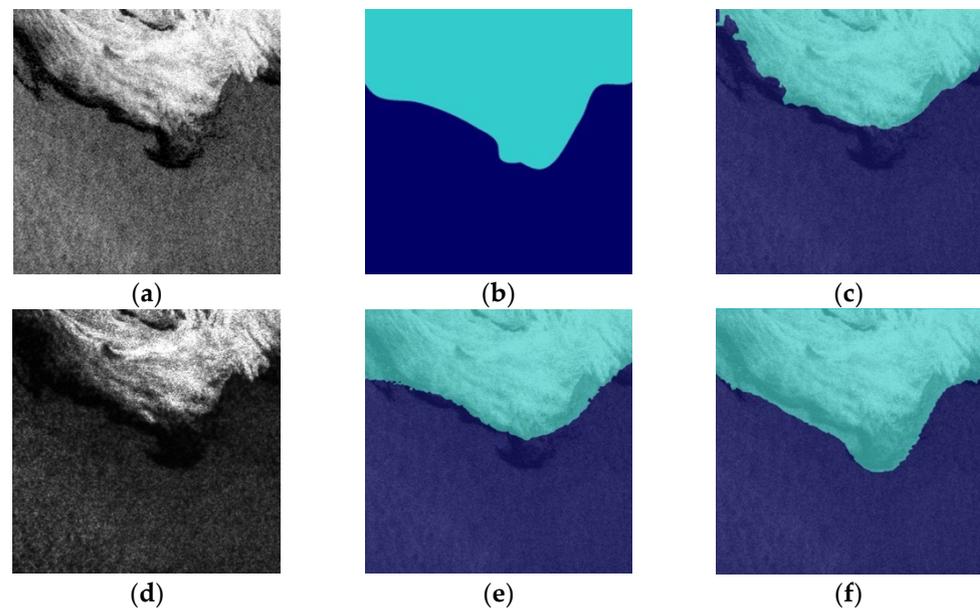
In order to verify the effectiveness of the MFFM and ESM in E-MPSPNet, three structures, backbone network, backbone and MFFM, and backbone and MFFM and ESM are compared based on the test set. It can be seen from Table 4 that the addition of MFFM and ESM has improved all the evaluation indicators compared with the backbone network. Specially, the MFFM module increases MIoU by 0.7% and F-score by 0.4%. This shows that MFFM can improve the expression of semantic features of sea ice and water by fusing multi-scale features with attention mechanisms. The ESM module increases MIoU by 1% and F-score by 0.7%, bringing more improvement than the MFFM. This proves that edge features added to the segmentation network can generate positive feedback and can solve the problem of the loss of sea ice edge details.

**Table 4.** The effects of MFFM and ESM on segmentation performance.

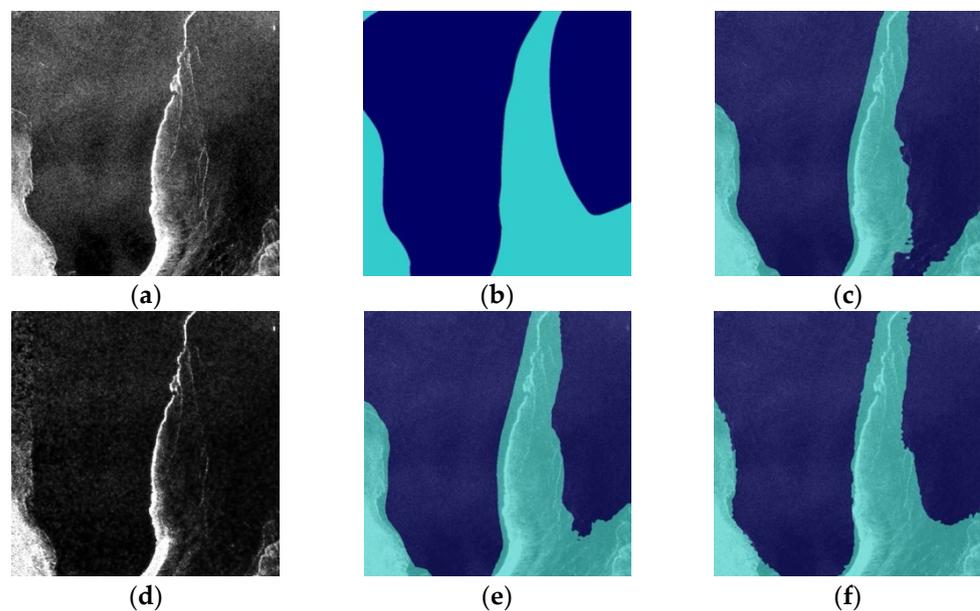
Network Structure	Accuracy	F-Score	MIoU
Backbone network	0.934	0.919	0.875
Backbone + MFFM	0.936	0.923	0.882
Backbone + MFFM + ESM	<b>0.942</b>	<b>0.930</b>	<b>0.892</b>

The results of the subjective evaluation are shown in Figures 6 and 7. Figure 6a depicts a sea ice area with a clear ice–water boundary. When the prediction results are compared, it is discovered that the backbone network (Figure 6c) is better for segmenting large areas of sea ice and water but loses a significant amount of edge details. The same problem of inadequate and incomplete segmentation exists after adding MFFM (Figure 6e), but the effect is improved because effective multi-scale contextual information is extracted after adding the multi-scale attention mechanism, so that the mis-segmentation is greatly

reduced. The addition of the ESM module (Figure 6f) alleviates the problems of incorrect and missed segmentation, and the edge contour can be better preserved.



**Figure 6.** Segmentation results for an example sub-region SAR image with a clear ice–water boundary. (a) The original SAR image (HH), (b) the ground truth semantic segmentation map, (c) the prediction result of the backbone network, (d) the original SAR image (HV), (e) the prediction result of the backbone and MFFM structure, and (f) the prediction result of the backbone and MFFM and ESM structure.



**Figure 7.** Segmentation results for an example sub-region SAR image with low-concentration sea ice filled at the edges. (a) The original SAR image (HH), (b) the ground truth semantic segmentation map, (c) the prediction result of the backbone network, (d) the original SAR image (HV), (e) the prediction result of the backbone and MFFM structure, and (f) the prediction result of the backbone and MFFM and ESM structure.

In Figure 7a, low-concentration sea ice is filled at the edges, and similar texture characteristics are observed between low-concentration sea ice and open water. According

to the prediction results, the segmentation in the backbone network (Figure 7c) performs the worst, misclassifying the low-concentration sea ice as water, and the reason for the unsatisfactory segmentation boundary could be a lack of dependency information among pixels. The addition of the MFFM module (Figure 7e) to the backbone network improves the edge segmentation effect by obtaining more contextual information. The backbone and MFFM and ESM structure (Figure 7f) has further improved the edge details, yielding a segmentation result that is closer to the ground truth compared with other results. In summary, MFFM provides more effective global information for the model, and the shallow feature maps in ESM can supplement the missing edge information, showing good segmentation performance for the case of low-concentration sea ice edges.

#### 4.3.3. Influence of Different Loss Functions on Network Segmentation Performance

For a deep learning model, the neural network weights in the model are trained through loss backpropagation, so the loss function determines the training effect of the deep learning model. In order to verify the effectiveness of the loss function designed in this paper, ablation experiments were performed on the loss function of E-MPSPNet. To simplify the comparison, the joint loss function is defined as  $l_1 = l_{Focal}$ ,  $l_2 = l_{Focal} + l_{MIoU}$ ,  $l_3 = l_{Focal} + l_{MIoU} + l_{edg}$ , where  $l_1$  and  $l_2$  are the loss functions for the semantic segmentation task and  $l_3$  is the loss function after adding edge features. The segmentation performance with different loss functions is shown in Table 5.

**Table 5.** Results of the effect of different loss on network segmentation performance.

Methodology	Accuracy	F-Score	MIoU
$l_1$	0.928	0.915	0.866
$l_2$	0.936	0.924	0.879
$l_3$	<b>0.942</b>	<b>0.930</b>	<b>0.892</b>

As can be seen from Table 5, using only focal loss can achieve relatively good performance. Comparing  $l_2$  to  $l_1$ , it can be concluded that the MIoU loss function can greatly lift the segmentation performance, which makes the model focus more on difficult and easily misclassified samples during training, and eliminates the problems caused by the unbalanced categories of samples. Compared with  $l_1$ , accuracy, F-score, and MIoU using  $l_3$  loss increased by 1.4%, 1.5%, and 2.6%, respectively. The ablation experiments show that predicting uncertain edges can help to improve the overall segmentation effect of the network, and the joint loss function designed in this paper has obvious advantages to make up for the shortage of segmentation results of single loss models, which can reduce the gap between the model prediction results and the true values and improve the model ability of detail detection.

## 5. Discussion

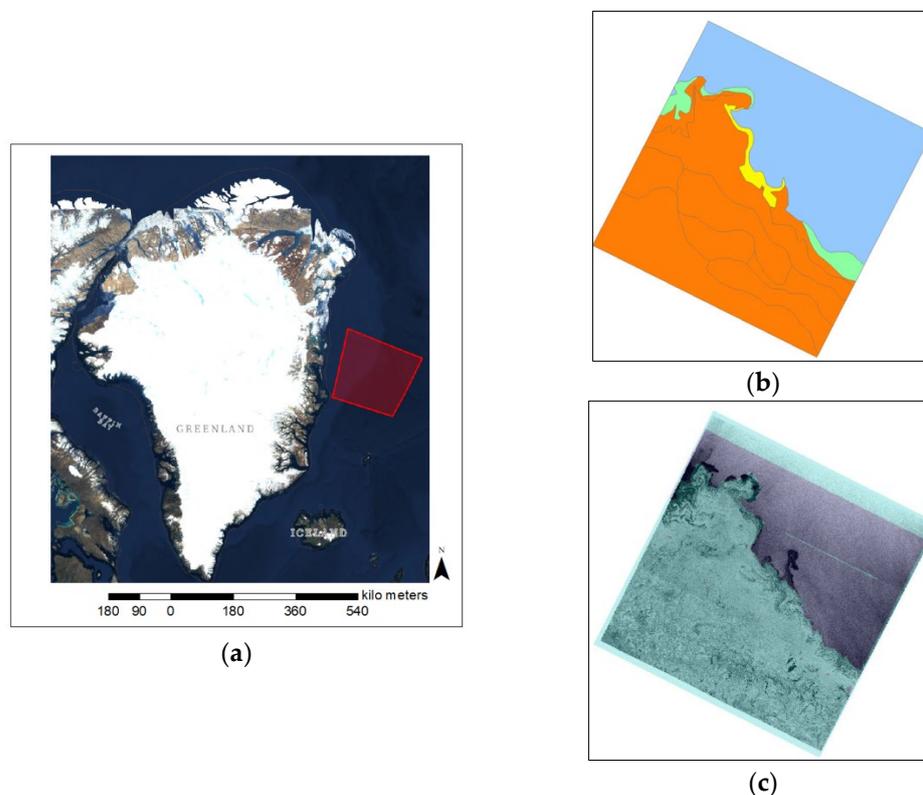
### 5.1. The Application of Ice–Water Segmentation in a SAR Scene

A fully automated sea ice segmentation system can not only provide a near real-time service for sea ice mapping, but also minimizes the impact of human bias when compared to manual ice map drawing by different analysts. Thus, it is required to test the performance of an ice–water segmentation algorithm in practice.

To demonstrate the utility of our proposed model in automated ice–water segmentation applications, we applied the trained E-MPSPNet to the ice–water scene segmentation for the whole SAR image. The process includes three steps: (1) cropping the entire SAR image into sub-regions with a sliding window of  $800 \times 800$  pixels; (2) performing segmentation for each sub-region of the SAR image by the E-MPSPNet; (3) reassembling the segmentation outputs of these sub-regions in their original order to form a complete segmentation map.

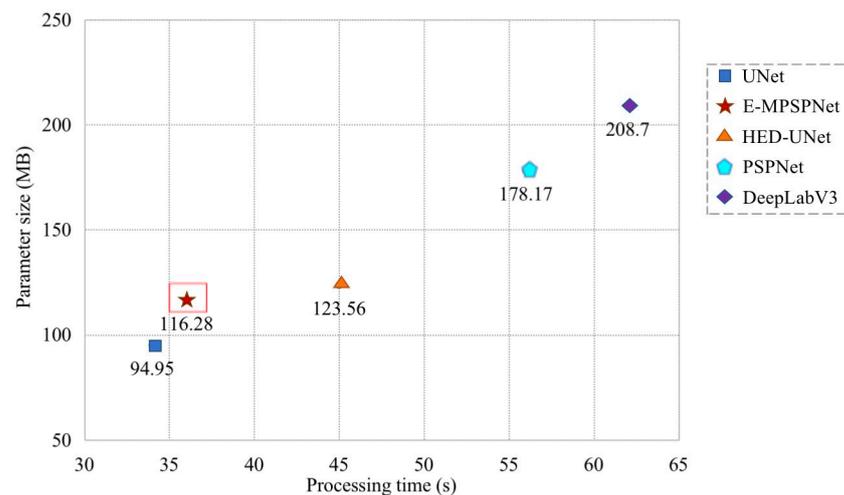
We took the SAR image from the Greenland Sea on 22 December 2021 (Figure 8) as an example, which contained data that had not been used to train the E-MPSPNet

network. Figure 8a shows the USNIC ice chart corresponding to the selected SAR image, and Figure 8b shows the segmentation result overlaid on the original SAR image. It can be seen that the overall edge between sea ice and open water is quite close to the true ice chart, showing the good generalization capability of our model. There is a small line of mis-segmentation in the middle of open water, which is often caused by the swath band effect of SAR imagery. This kind of mis-segmentation can be eliminated by simple filtering.



**Figure 8.** Applying E-MPSPNet for ice–water segmentation of the SAR image. (a) The SAR image located around the Greenland east coast on December 22, 2021. (b) USNIC ice chart, where blue represents the ice concentration  $< 1/10$  Open Water, green represents the ice concentration  $1\text{--}3/10$  Very Open Drift Ice, yellow represents the ice concentration  $4\text{--}6/10$  Open Drift Ice, and orange represents the ice concentration  $7\text{--}8/10$  Close Drift Ice, (c) the predicted segmentation overlapped on the SAR image, where dark blue is sea water and light blue is sea ice.

Moreover, we test the efficiency of our model for the above example. Figure 9 illustrates a scatter diagram of parameter sizes versus processing times for different models. Without any parallel optimization, the time spent to completely process a SAR image of  $10,000 \times 10,000$  pixels is only 38 s on a single GeForce RTX 2080 Ti GPU, which indicates the effectiveness of the proposed model. Additionally, the parameter sizes for various models are examined. E-MPSPNet with a parameter size of 116.28 MB is smaller than other models except UNet. Although UNet has fewer parameters, its accuracy performance is poorer than our model. In real applications, limited by the memory space of a fully automated sea ice segmentation system, it is critical to design a model with a small number of parameters while ensuring high performance. In conclusion, the E-MPSPNet model provides technical support for drawing a navigation map that accurately distinguishes ice and water, and it has a broad prospect for designing an automatic segmentation system for timely analysis of SAR images using E-MPSPNet.



**Figure 9.** The parameter size–time scatter diagram. Comparison of processing time(s) and parameter size (MB) of E-MPSPNet and four segmentation models: U-Net, PSPNet, DeepLabV3, and HED-UNet for processing a SAR image of  $10,000 \times 10,000$  pixels.

### 5.2. The Impact of the Incident Angle

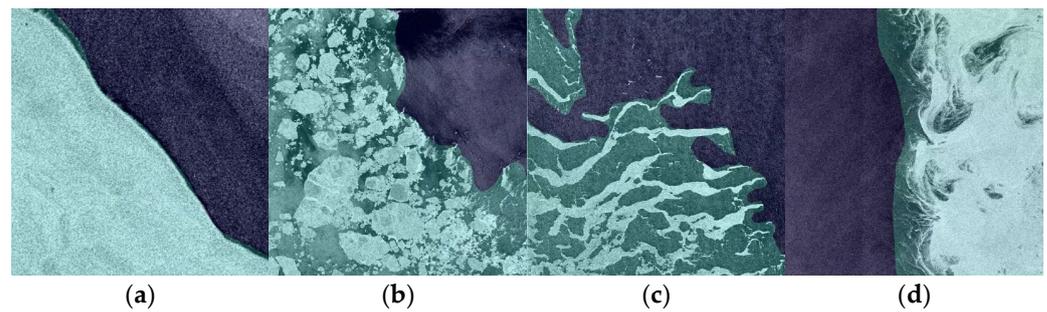
Sea ice conditions are complex and it is still a challenge to determine a common threshold for ice and water segmentation using only dual-polarization radar intensity. The correlation between image backscatter and incidence angle has been studied for the ocean and sea ice in the past [48,49]; in the observation of satellite-based radar data, the incidence angle is an important factor affecting the scattering characteristics of the features, and there is a direct relationship between the backscattering intensity of SAR images and the incidence angle. Thus, this paper also investigates the effect of incidence angle on the segmentation performance of the model.

We experimented with two ways of introducing the incident angle information into the ice–water segmentation model. Firstly, the incident angle information of Sentinel-1 images provided in the ASIP dataset was used as the third channel of our model input, together with the channels of HH and HV polarization, to train the E-MPSPNet. The experimental results were not satisfactory, and all evaluation indicators decreased after adding the third channel. Considering that the incident angle has the same value on each column in the image numerical matrix, the performance drop might be because highly redundant information affects the model’s feature learning. In the second experiment, the incident angle information was used as a separate feature vector and fused with other extracted features in the MFFM module, and the experimental results show that the segmentation effect does not bring much improvement.

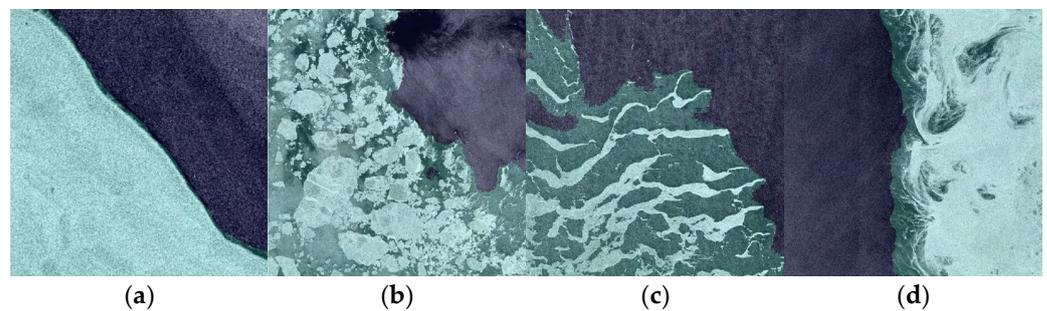
We conclude that the proposed network in this paper is insensitive to the effect of incidence angle. However, to determine whether the incidence angle information of SAR images can provide useful information in the research of ice and water segmentation, we need to investigate more types of SAR images.

### 5.3. The Impact of Characteristics of Ice–Water Boundaries

As shown in Section 4.3, the problem of inaccurate partitioning is mainly in the transition region between open water and low-concentration sea ice. To further discuss the impact of the characteristics of ice–water boundaries on the segmentation, we illustrate four typical scenarios of different ice–water boundaries in Figure 10 (overlapping the ground truth segmentation on the original SAR image (HH)). High-concentration sea ice tends to have clear boundaries (Figure 10a), while brash ice leads to complex ice–water mixing boundaries (Figure 10b), ice and water intersecting with each other has unclear boundaries (Figure 10c), and grease ice gives transitional boundaries (Figure 10d). Figure 11a–d show the predicted results by our E-MPSPNet corresponding to the figures of Figure 10a–d.



**Figure 10.** The labeled segmentations of four scenarios with different characteristics of ice–water boundaries. (a) High-concentration sea ice has clear boundaries, (b) brash ice has complex ice–water mixing boundaries, (c) ice–water intersection with each other has unclear boundaries, (d) grease ice has transitional boundaries.



**Figure 11.** The predicted segmentation results. (a) The result corresponding to Figure 10a, (b) the result corresponding to Figure 10b, (c) the result corresponding to Figure 10c, (d) the result corresponding to Figure 10d.

Although our model is effective in edge detail detection and outperforms other models in all evaluation indicators, there are still imperfections to match the ground truth (comparison between Figures 10 and 11).

We consider that part of the reason for the mismatch is the limitation of ice chart production. For high-dynamic sea ice regions, ice chart information has limited value due to the time delay from the acquisition of satellite data to the production of the chart. Furthermore, C-band SAR satellites have difficulty to capture very thin ice, so the true ice–water boundary in this study may actually be limited by the threshold at which sensors recognize ice. In addition, for navigation safety, DMI ice chart analysts sometimes expand the coverage of ice to generate ice charts, but actual sea ice conditions are unlikely to have smooth boundary lines, as demonstrated in Figure 10d.

## 6. Conclusions

In this paper, after investigating the development of sea ice research based on SAR images from the perspective of realizing accurate and automatic segmentation of ice and water, we propose an ice–water SAR segmentation network, E-MPSPNet, to enable high precision in ice edges that are hard to segment.

The proposed E-MPSPNet is based on the improved PSPNet framework, followed by a multi-scale attention mechanism to add weights to the different scales of semantic feature maps collected by the spatial pyramid pooling module, which can enhance the attention to the most relevant dimensions of the current pixel and improve the segmentation effect. Meanwhile, the edge supervision module is constructed by introducing lateral output layers on the residual blocks of the PSPNet backbone network for edge prediction. It facilitates edge expression on the semantic feature maps and combines the edge features from shallow to deep layers to achieve multi-scale edge map prediction. Finally, the edge map and the semantic segmentation map are fused by a convolution operation to extract

rich feature details of the image, providing more accurate results for the ice–water semantic segmentation task. The E-MPSPNet is trained with a joint loss function that minimizes both the semantic segmentation loss and edge loss. Experiments were conducted on a public AI4Arctic/ASIP sea ice dataset that has been modified for ice–water segmentation. The proposed model in this paper achieves an accuracy of 0.942, F-score of 0.930, and MIoU of 0.892. Compared with U-Net, PSPNet, DeepLabV3, and HED-UNet, our model has better performance when segmenting open water and low-concentration sea ice. Results of ablation experiments have shown the modules of multi-scale feature fusion and edge supervision play important roles in distinguishing sea ice and water and drawing clear boundaries between them. The applicability of E-MPSPNet has been verified on a SAR scene segmentation, achieving a good match with the USNIC ice chart in a relatively quick process. We have also discussed the potential impacts of the incident angle of SAR imagery and the segmentation performance of our model on ice–water boundaries with different characteristics.

Some limitations of this work should be mentioned. The effect of segmentation is restricted by the ice charts, given that the evaluation is based on the available ice charts, and the model performance may also vary from the dataset. At present, weak supervision and semi-supervised learning semantic segmentation algorithms can reduce the dependence of algorithm models on finely labeled data and the cost of labeling. Our further work will consider establishing a dataset with more detailed edge mapping and examine our model's performance with other SAR images, such as those from RADARSAT satellites. We will also try to migrate semi-supervised algorithms to sea ice segmentation to improve the real-time performance of automatic sea ice segmentation.

**Author Contributions:** Conceptualization, W.S. and Q.H.; data processing, H.L.; methodology, W.S. and H.L.; writing—original draft preparation, H.L.; writing—review and editing, W.S., A.L. and Q.H.; investigation and supervision, G.G. and A.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by National Natural Science Foundation of China (NSFC), grant number 61972240 and the Program for the Capacity Development of Shanghai Local Colleges, funded by Shanghai Science and Technology Commission, grant number 20050501900.

**Data Availability Statement:** The AI4Arctic/ASIP Sea Ice Dataset—version2 used in this study is available at [https://data.dtu.dk/articles/dataset/AI4Arctic\\_ASIP\\_Sea\\_Ice\\_Dataset\\_-\\_version\\_2/1301134/3](https://data.dtu.dk/articles/dataset/AI4Arctic_ASIP_Sea_Ice_Dataset_-_version_2/1301134/3) (accessed on 12 July 2021).

**Acknowledgments:** We would like to express our gratitude to the Danish Meteorological Institute (DMI), the Technical University of Denmark (DTU), and Nansen Environmental Remote Sensing Center (NERSC) for providing the AI4Arctic/ASIP Sea Ice Dataset—version2. We gratefully thank the National Natural Science Foundation of China, China (grant no.61972240) and the Program for the Capacity Development of Shanghai Local Colleges (grant no.20050501900) for supporting this study.

**Conflicts of Interest:** No potential conflicts of interest are reported by the authors.

## References

1. Carter, N.; Dawson, J.; Joyce, J.; Ogilvie, A. *Arctic Corridors and Northern Voices: Governing Marine Transportation in the Canadian Arctic (Arviat, Nunavut Community Report)*; Arctic Corridors: Ottawa, ON, Canada, 2017.
2. Kang, M.-S.; Kim, K.-T. Automatic SAR Image Registration via Tsallis Entropy and Iterative Search Process. *IEEE Sens. J.* **2020**, *20*, 7711–7720. [[CrossRef](#)]
3. Gong, M.; Cao, Y.; Wu, Q. A Neighborhood-Based Ratio Approach for Change Detection in SAR Images. *IEEE Geosci. Remote Sens. Lett.* **2012**, *9*, 307–311. [[CrossRef](#)]
4. Hakim, W.L.; Achmad, A.R.; Eom, J.; Lee, C.-W. Land Subsidence Measurement of Jakarta Coastal Area Using Time Series Interferometry with Sentinel-1 SAR Data. *J. Coast. Res.* **2020**, *102*, 75–81. [[CrossRef](#)]
5. Partington, K.C.; Flach, J.D.; Barber, D.; Isleifson, D.; Meadows, P.J.; Verlaan, P. Dual-Polarization C-Band Radar Observations of Sea Ice in the Amundsen Gulf. *IEEE Trans. Geosci. Remote Sens.* **2010**, *48*, 2685–2691. [[CrossRef](#)]
6. Makynen, M.; Manninen, A.; Simila, M.; Karvonen, J.; Hallikainen, M. Incidence angle dependence of the statistical properties of C-band HH-polarization backscattering signatures of the Baltic Sea ice. *IEEE Trans. Geosci. Remote Sens.* **2002**, *40*, 2593–2605. [[CrossRef](#)]

7. Nghiem, S.; Bertoia, C. Study of Multi-Polarization C-Band Backscatter Signatures for Arctic Sea Ice Mapping with Future Satellite SAR. *Can. J. Remote Sens.* **2014**, *27*, 387–402. [[CrossRef](#)]
8. Holmes, Q.A.; Nuesch, D.R.; Shuchman, R.A. Textural Analysis and Real-Time Classification of Sea-Ice Types Using Digital SAR Data. *IEEE Trans. Geosci. Remote Sens.* **2007**, *2*, 113–120. [[CrossRef](#)]
9. Soh, L.-K.; Tsatsoulis, C. Texture analysis of SAR sea ice imagery using gray level co-occurrence matrices. *IEEE Trans. Geosci. Remote Sens.* **1999**, *37*, 780–795. [[CrossRef](#)]
10. Clausi, D.A. An analysis of co-occurrence texture statistics as a function of grey level quantization. *Can. J. Remote Sens.* **2002**, *28*, 45–62. [[CrossRef](#)]
11. Clausi, D.A.; Yue, B. Comparing Cooccurrence Probabilities and Markov Random Fields for Texture Analysis of SAR Sea Ice Imagery. *IEEE Trans. Geosci. Remote Sens.* **2004**, *42*, 215–228. [[CrossRef](#)]
12. Ochilov, S.; Clausi, D.A. Operational SAR Sea-Ice Image Classification. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 4397–4408. [[CrossRef](#)]
13. Korosov, A.; Zakhvatkina, N.; Muckenhuber, S. Ice/Water Classification of Sentinel-1 Images. In *EGU General Assembly Conference Abstracts*; EGU: Munich, Germany, 2015.
14. Liu, H.; Guo, H.; Zhang, L. SVM-Based Sea Ice Classification Using Textural Features and Concentration From RADARSAT-2 Dual-Pol ScanSAR Data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *8*, 1601–1613. [[CrossRef](#)]
15. Zakhvatkina, N.; Korosov, A.; Muckenhuber, S.; Sandven, S.; Babiker, M. Operational algorithm for ice–water classification on dual-polarized RADARSAT-2 images. *Cryosphere* **2017**, *11*, 33–46. [[CrossRef](#)]
16. Kim, M.; Im, J.; Han, H.; Kim, J.; Lee, S.; Shin, M.; Kim, H.-C. Landfast sea ice monitoring using multisensor fusion in the Antarctic. *Mapp. Sci. Remote Sens.* **2015**, *52*, 239–256. [[CrossRef](#)]
17. Wiebke, A.; Céline, H.; Eriksson, L. Comparison of ice/water classification in Fram Strait from C- and L-band SAR imagery. *Ann. Glaciol.* **2018**, *59*, 112–123.
18. Karvonen, J.; Simila, M.; Mäkynen, M. Open Water Detection from Baltic Sea Ice Radarsat-1 SAR Imagery. *IEEE Geosci. Remote Sens. Lett.* **2005**, *2*, 275–279. [[CrossRef](#)]
19. Mohammadimanesh, F.; Salehi, B.; Mahdianpari, M.; Gill, E.; Molinier, M. A new fully convolutional neural network for semantic segmentation of polarimetric SAR imagery in complex land cover ecosystem. *ISPRS J. Photogramm. Remote Sens.* **2019**, *151*, 223–236. [[CrossRef](#)]
20. Huang, L.; Liu, B.; Li, B.; Guo, W.; Yu, W.; Zhang, Z.; Yu, W. OpenSARShip: A dataset dedicated to Sentinel-1 ship interpretation. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, *11*, 195–208. [[CrossRef](#)]
21. Xia, G.-S.; Bai, X.; Ding, J.; Zhu, Z.; Belongie, S.; Luo, J.; Datcu, M.; Pelillo, M.; Zhang, L. DOTA: A Large-Scale Dataset for Object Detection in Aerial Images. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, USA, 18–22 June 2018.
22. Heidler, K.; Mou, L.; Baumhoer, C.; Dietz, A.; Zhu, X.X. HED-UNet: Combined segmentation and edge detection for monitoring the Antarctic coastline. *IEEE Trans. Geosci. Remote Sens.* **2021**, *60*, 4300514. [[CrossRef](#)]
23. Wang, Z.; Bai, L.; Song, G.; Zhang, J.; Tao, J.; Mulvenna, M.; Bond, R.; Chen, L. An Oil Well Dataset Derived from Satellite-Based Remote Sensing. *Remote Sens.* **2021**, *13*, 1132. [[CrossRef](#)]
24. Khaleghian, S.; Ullah, H.; Krämer, T.; Hughes, N.; Eltoft, T.; Marinoni, A. Sea Ice Classification of SAR Imagery Based on Convolution Neural Networks. *Remote Sens.* **2021**, *13*, 1734. [[CrossRef](#)]
25. Song, W.; Gao, W.; He, Q.; Liotta, A.; Guo, W. SI-STARSAR-7: A Large SAR Images Dataset with Spatial and Temporal Information for Classification of Winter Sea Ice in Hudson Bay. *Remote Sens.* **2022**, *14*, 168. [[CrossRef](#)]
26. Wulf, T.; Kreiner, M.B.; Buus-Hinkler, J.; Tonboe, R.T.; Hoyer, J.L.; Saldo, R.; Pedersen, L.T.; Nielsen, A.A.; Skriver, H.; Malmgren-Hansen, D. Fusion of Satellite SAR and Passive Microwave Radiometer Data for Automated Sea Ice Mapping and the Expected Impact of CIMR Observations. In *Proceedings of the From Science to Operations for the Copernicus Imaging Microwave Radiometer (CIMR) Mission*, Noordwijk, The Netherlands, 15–17 April 2020; ESA: Paris, France, 2021.
27. Wang, L.; Scott, K.A.; Clausi, D.A. Sea ice concentration estimation during freeze-up from SAR imagery using a convolutional neural network. *Remote Sens.* **2017**, *9*, 408. [[CrossRef](#)]
28. Malmgren-Hansen, D.; Pedersen, L.T.; Nielsen, A.A.; Kreiner, M.B.; Saldo, R.; Skriver, H.; Lavelle, J.; Buus-Hinkler, J.; Krane, K.H. A Convolutional Neural Network Architecture for Sentinel-1 and AMSR2 Data Fusion. *IEEE Trans. Geosci. Remote Sens.* **2021**, *59*, 1890–1902. [[CrossRef](#)]
29. Dirscherl, M.; Dietz, A.; Kneisel, C.; Kuenzer, C. A Novel Method for Automated Supraglacial Lake Mapping in Antarctica Using Sentinel-1 SAR Imagery and Deep Learning. *Remote Sens.* **2021**, *13*, 197. [[CrossRef](#)]
30. Kruk, R.; Fuller, M.; Komarov, A.; Isleifson, D.; Jeffrey, I. Proof of Concept for Sea Ice Stage of Development Classification Using Deep Learning. *Remote Sens.* **2020**, *12*, 2486. [[CrossRef](#)]
31. Han, Y.; Liu, Y.; Hong, Z.; Zhang, Y.; Yang, S.; Wang, J. Sea Ice Image Classification Based on Heterogeneous Data Fusion and Deep Learning. *Remote Sens.* **2021**, *13*, 592. [[CrossRef](#)]
32. Zhang, T.; Yang, Y.; Shokr, M.; Mi, C.; Li, X.-M.; Cheng, X.; Hui, F. Deep Learning Based Sea Ice Classification with Gaofen-3 Fully Polarimetric SAR Data. *Remote Sens.* **2021**, *13*, 1452. [[CrossRef](#)]
33. Boulze, H.; Korosov, A.A.; Brajard, J. Classification of Sea Ice Types in Sentinel-1 SAR Data Using Convolutional Neural Networks. *Remote Sens.* **2020**, *12*, 2165. [[CrossRef](#)]

34. Chen, L.C.; Papandreou, G.; Kokkinos, I.; Murphy, K.; Yuille, A.L. Semantic Image Segmentation with Deep Convolutional Nets and Fully Connected CRFs. *Comput. Sci.* **2014**, *4*, 357–361.
35. Chen, L.-C.; Barron, J.T.; Papandreou, G.; Murphy, K.; Yuille, A.L. Semantic Image Segmentation with Task-Specific Edge Detection Using CNNs and a Discriminatively Trained Domain Transform. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 26 June–1 July 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 4545–4554.
36. Takikawa, T.; Acuna, D.; Jampani, V.; Fidler, S. Gated-SCNN: Gated Shape CNNs for Semantic Segmentation. In Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Korea, 27 October–2 November 2019; IEEE: Piscataway, NJ, USA, 2020.
37. Sun, G.Y.; Huang, B.H.; Zhao, X.J. The Analysis of Edge Detection Uncertainty of Remote Sensing Images and its Processing Method. *Remote Sens. Inf.* **2010**, *32*, 110–114.
38. Wu, T.; Jin, Y.-F.; Hou, R.; Yang, J.-J. Cognitive physics-based method for image edge representation and extraction with uncertainty. *Acta Phys. Sin.* **2013**, *62*, 675.
39. Park, J.-W.; Korosov, A.A.; Babiker, M.; Sandven, S.; Won, J.-S. Efficient Thermal Noise Removal for Sentinel-1 TOPSAR Cross-Polarization Channel. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 1555–1565. [[CrossRef](#)]
40. Zhao, H.; Shi, J.; Qi, X.; Wang, X.; Jia, J. Pyramid Scene Parsing Network. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 6230–6239.
41. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 26 June–1 July 2016.
42. Lee, C.-Y.; Xie, S.; Gallagher, P.; Zhang, Z.; Tu, Z. Deeply-Supervised Nets. In Proceedings of the Eighteenth International Conference on Artificial Intelligence and Statistics, San Diego, CA, USA, 9–12 May 2015.
43. Lin, T.; Goyal, P.; Girshick, R.; He, K.; Dollár, P. Focal Loss for Dense Object Detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *99*, 2999–3007.
44. Kingma, D.P.; Adam, J.B. A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980.
45. Yang, H.; Yu, B.; Luo, J.; Chen, F. Semantic segmentation of high spatial resolution images with deep neural networks. *GISci. Remote Sens.* **2019**, *56*, 749–768. [[CrossRef](#)]
46. Stokholm, A.; Wulf, T.; Kucik, A.; Saldo, R.; Buus-Hinkler, J.; Hvidegaard, S.M. AI4SeaIce: Toward Solving Ambiguous SAR Textures in Convolutional Neural Networks for Automatic Sea Ice Concentration Charting. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 4304013. [[CrossRef](#)]
47. Wang, Z.; Wang, J.; Yang, K.; Wang, L.; Su, F.; Chen, X. Semantic segmentation of high-resolution remote sensing images based on a class feature attention mechanism fused with Deeplabv3+. *Comput. Geosci.* **2022**, *158*, 104969. [[CrossRef](#)]
48. Singha, S.; Johansson, A.M.; Doulgeris, A.P. Robustness of SAR Sea Ice Type Classification Across Incidence Angles and Seasons at L-Band. *IEEE Trans. Geosci. Remote Sens.* **2020**, *59*, 9941–9952. [[CrossRef](#)]
49. Singha, S.; Johansson, A.M.; Doulgeris, A.P. Incident Angle Dependence of Sentinel-1 Texture Features for Sea Ice Classification. *Remote Sens.* **2021**, *13*, 552.