Journal of *Marine Science and Engineering*

MDPI

Article

A Shape-Aware Network for Arctic Lead Detection from Sentinel-1 SAR Images

Wei Song [1], Min Zhu [1], Mengying Ge [1,2] and Bin Liu [3,*]

[1] College of Information Technology, Shanghai Ocean University, Shanghai 201306, China; wsong@shou.edu.cn (W.S.); m210911533@st.shou.edu.cn (M.Z.); gemengying@shu.edu.cn (M.G.)
[2] Engineering Training Centre, Shanghai University, Shanghai 200444, China
[3] College of Oceanography and Ecological Science, Shanghai Ocean University, Shanghai 201306, China
* Correspondence: bliu@shou.edu.cn

Abstract: Accurate detection of sea ice leads is essential for safe navigation in polar regions. In this paper, a shape-aware (SA) network, SA-DeepLabv3+, is proposed for automatic lead detection from synthetic aperture radar (SAR) images. Considering the fact that training data are limited in the task of lead detection, we construct a dataset fusing dual-polarized (HH, HV) SAR images from the C-band Sentinel-1 satellite. Taking the DeepLabv3+ as the baseline network, we introduce a shape-aware module (SAM) to combine multi-scale semantic features and shape information and, therefore, better capture the shape characteristics of leads. A squeeze-and-excitation channel-position attention module (SECPAM) is designed to enhance lead feature extraction. Segmentation loss generated by the segmentation network and shape loss generated by the shape-aware stream are combined to optimize the network during training. Postprocessing is performed to filter out segmentation errors based on the aspect ratio of leads. Experimental results show that the proposed method outperforms the existing benchmarking deep learning methods, reaching 96.82% for overall accuracy, 93.01% for F1-score, and 91.48% for mIoU. It is also found that the fusion of dual-polarimetric SAR channels as the input could effectively improve the accuracy of sea ice lead detection.

Keywords: lead detection; synthetic aperture radar (SAR) images; shape-aware module (SAM); squeeze-and-excitation channel-position attention module (SECPAM)

check for updates

Citation: Song, W.; Zhu, M.; Ge, M.; Liu, B. A Shape-Aware Network for Arctic Lead Detection from Sentinel-1 SAR Images. *J. Mar. Sci. Eng.* **2024**, *12*, 856. https://doi.org/10.3390/jmse12060856

Academic Editor: Sang Heon Lee

Received: 7 April 2024
Revised: 14 May 2024
Accepted: 20 May 2024
Published: 22 May 2024

Copyright: © 2024 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 leads are linear openings in the sea ice cover, typically ranging in length from a few kilometers to hundreds of kilometers and with a width of at least 50 m [1]. They result from the movement of sea ice and external forces such as wind and ocean currents [2]. Compared to surrounding sea ice, leads have a lower albedo and, therefore, can absorb more solar energy, which accelerates the ice melting and increases the amount of open water. As a result, leads play a significant role in opening up Arctic shipping routes and supporting scientific research missions [3]. Furthermore, sea ice lead detection plays a crucial role in estimating sea ice thickness using radar altimeter data [4]. Obtaining the accurate distribution of leads in polar regions is of great importance.

In recent years, satellite remote sensing has become a crucial technique for monitoring sea ice leads in polar regions due to the limitations of in situ observations [3]. Since the early 1980s, spaceborne optical, thermal infrared, and microwave sensor data have been widely used in the study of sea ice leads [5]. In particular, synthetic aperture radar (SAR) has progressively been considered an essential data source for lead observation due to its all-weather and day-and-night sensing capability, wide coverage, and suitable spatial and temporal resolution [6]. Additionally, a large number of studies have demonstrated that combining HH and HV polarization data of SAR images is helpful for distinguishing between open water and sea ice [7,8]. There are many studies utilizing dual-polarization

SAR data to discriminate between open water and sea ice [9,10]. Zakhvatkina et al. [11] introduced a neural network (NN) classification method for lead detection, using polarization ratio and polarization difference of HH and HV channels in SAR images together with texture features to detect leads. Murashkin et al. [2] used a random forest classifier to detect leads in SAR images based on the dual-polarization features and texture features of Sentinel-1 SAR images. However, accurate lead detection based on SAR imagery is a challenging task. On the one hand, there are many factors that restrict the performance of lead detection based on SAR images, such as speckle noise, system parameters (polarization mode, incidence angle), and environmental factors such as wind and temperature [12]. On the other hand, influenced by the physical characteristics of sea ice, the growth of leads is complex, and the ice–water boundary presents dynamic changes.

Current methods that could be employed for lead detection from SAR images can be categorized into three: threshold-based methods [13], machine learning (ML) methods such as the k-nearest neighbors [12], K-means [4], and neural network [3], and deep learning (DL) methods [14,15]. The aforementioned threshold-based and ML methods all need manual involvement, such as determining thresholds and selecting features. In contrast, DL-based methods with an end-to-end process have shown great potential for lead detection from SAR images due to their ability to provide powerful nonlinear representations and to automatically extract reliable features from large datasets [14,16]. Nonetheless, these studies mostly rely on texture features and backscattering coefficients [17,18]. More recently, Liang et al. [1] developed an entropy-weighted network (EW-Net) to detect leads, which utilizes an entropy-weighted feature fusion block to merge texture and entropy information from SAR images. Liu et al. [15] proposed a lightweight semantic segmentation model based on the U-Net framework for precise and fast extraction of sea ice leads from SAR images, demonstrating improved performance on non-preprocessed data compared to classical methods. They also use texture features as input. It is evident that capturing additional attributes of ice leads will contribute to enhanced precision in lead detection.

We noticed that leads have obvious linear characteristics, and little attention has been paid to utilizing shape information for lead detection in previous studies. Inspired by the previous work [19] on integrating edge information into SAR segmentation, we explore the usability of shape information to improve the performance of DL-based lead detection and propose a network, SA-DeepLabv3+, which is based on DeepLabv3+ [20]. It is worth mentioning that although recently proposed models like the Segment Anything Model [21] perform well on general public datasets, its original design aims to segment various objects rather than specific ones. Since its ViT backbone encoder is trained on large-scale close-range images, its effectiveness in processing spaceborne SAR images is limited [22]. Additionally, training big models specifically for remote sensing images requires a substantial amount of data, which is currently scarce in the field of ice lead segmentation. Therefore, we chose DeepLabv3+ as the benchmark model, as it has proven effective in remote sensing tasks and still has significant room for performance improvement.

In this paper, a shape-aware module (SAM) is employed in this network to combine multi-scale semantic features with shape information of leads, thereby better capturing the shape of leads. To the best of our knowledge, this is the first attempt to use shape information for DL-based lead detection from SAR images. What is more, in order to enhance lead features from both polarization and spatial perspectives, a squeeze-and-excitation channel-position attention module (SECPAM) is designed. Since dual-polarization SAR data fusion can provide more discriminative features for the observed scene [23,24] and the lack of reliable publicly available lead datasets poses significant challenges for lead research, we take the fused dual-polarization SAR image as input and construct a dataset, which is favorable for the detection of sea ice leads. The comparative experiments conducted on the dataset constructed in this paper demonstrate that the proposed network has achieved good segmentation performance and the effectiveness of the proposed modules. In addition, the constructed dataset significantly improved the performance of all segmentation models compared in this paper.

The main contributions of this paper are as follows:

1. We construct a sea ice lead dataset by fusing SAR dual-polarization data to address the scarcity of datasets in lead detection, which can effectively improve the accuracy of lead segmentation and provide innovative ideas for future research on lead datasets.
2. We propose a squeeze-and-excitation channel-position attention module (SECPAM) to enhance the encoder output features, addressing the issue of insufficient extraction of contextual and spatial information from remote sensing images by the model.
3. We propose a shape-aware module (SAM) that can combine multi-scale semantic features with the shape information of leads. Then, a joint loss function that combines segmentation loss and shape loss is designed. Using shape information learned from SAM to train the model can help the model learn the shape of the leads, making the segmented lead shape more complete and clearer.

2. Dataset

2.1. Study Area and Data Source

Considering the potential of lead growth, we chose the Beaufort Sea, located in northwestern Canada, as our study area, which is one of the most active Arctic lead areas [25]. Sentinel-1 dual-polarized, HH and HV, SAR images in extra-wide-swath mode at medium resolution are used as data sources. They are freely accessed from the Copernicus Open Access Center (https://scihub.copernicus.eu (accessed on 20 May 2021)). Since the sea ice movement and fracture during the Arctic winter and spring can promote the formation of leads, we collected a total of 10 SAR scenes of the Beaufort Sea from April to May and November to December from 2020 to 2022. The spatial distribution of these views can be seen in Figure 1, and the image information is shown in Table 1.

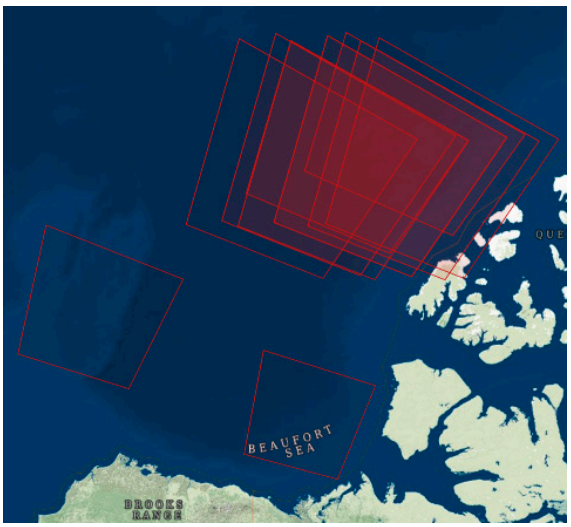


Figure 1. Position distribution of the 10 SAR scenes used in the dataset.

Table 1. Sentinel-1 data information used in our dataset.

Satellite	Acquisition	Wind (m/s)
Sentinel-1	30 April 2020; 16:00	6.1061
Sentinel-1	31 May 2020; 15:50	4.5122
Sentinel-1	30 November 2020; 16:15	5.2155
Sentinel-1	1 April 2021; 16:00	5.0329
Sentinel-1	28 November 2021; 16:40	4.0026
Sentinel-1	30 November 2021; 16:24	4.0610
Sentinel-1	30 December 2021; 15:34	3.7324

Table 1. *Cont.*

Satellite	Acquisition	Wind (m/s)
Sentinel-1	30 April 2022; 16:15	6.7837
Sentinel-1	22 November 2022; 15:59	6.7292
Sentinel-1	30 December 2022; 15:43	3.9218

2.2. Preprocessing and Lead Map Generation

We preprocess the original dual-polarization SAR data using the Sentinel Application Platform 9.0 (SNAP), following the standard procedure for noise removal and calibration [26].

Then, we utilize HH, HV, and HH × HV bands of SAR images as research data, as previous studies have shown the effectiveness of the HH × HV band product in classifying leads and sea ice [2,8]. After normalizing HH and HV data to [0, 1] using min–max normalization, we perform element-wise multiplication to obtain fused HH × HV data. This enhances the contrast between leads and sea ice, facilitating lead detection.

To generate labels for lead identification, we utilize the HH × HV data to create initial binary maps using SAR automated processing techniques [27]. Manual adjustments are made using image editing software to correct misclassifications and add missing pixels, resulting in final binary lead maps where 0 represents sea ice, and 1 represents leads. We crop HH, HV, and HH × HV data into 512 × 512-pixel patches using a non-overlapping sliding window approach. Finally, we synthesize three single-channel grayscale images into a pseudo-color image as input for our network. The dataset consists of 1200 samples, each composed of HH, HV, and HH × HV channels, split into training, validation, and testing sets at the ratio of 7:2:1. Data augmentation techniques, including mirror, vertical and horizontal flips are applied to enhance the network's generalization ability and robustness.

3. Methodology

3.1. Lead Detection Workflow

The workflow of sea ice lead detection in this paper consists of three main steps: dataset construction, image segmentation, and postprocessing. The flowchart of the lead detection workflow is shown in Figure 2. For given raw SAR data, dataset construction includes steps such as data preprocessing, extraction and fusion of dual-polarization data, and label generation. In the segmentation stage, we input the constructed dataset into our proposed segmentation network, SA-DeepLabv3+. Through the SA-DeepLabv3+ model, we obtain preliminary segmentation results for sea ice leads. Finally, we perform post-processing on the preliminary segmentation results. The postprocessing steps involve filtering operations to remove potential erroneous segmentations, fine-tuning, and optimization of the segmentation results to obtain the final ice lead detection results. Through this process, we can effectively and accurately detect ice leads from SAR images, providing crucial support and assistance for polar research and navigation.

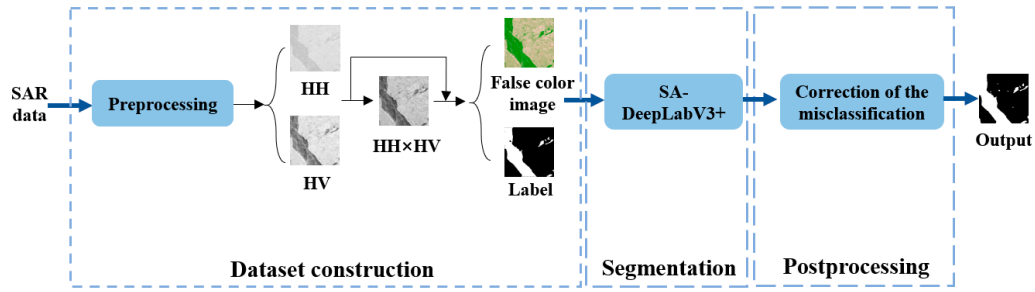


Figure 2. Processing flowchart of the sea ice lead detection.

3.2. Overview of the SA-DeepLabv3+ Structure

The proposed lead detection network is constructed based on the architecture of DeepLabv3+ while focusing on both texture and shape information. As shown in Figure 3,

the SA-DeepLabv3+ includes an encoder, decoder, and SAM. The input is three-channel pseudo-color SAR images, which consist of HH-polarized image, HV-polarized image, and HH × HV. With this input, the ResNet-50 [28] network works as the backbone to extract low-level and high-level features for further use. It then processes high-level features through an atrous spatial pyramid pooling (ASPP) module [29] and SECPAM. The ASPP module processes the high-level features using a 1 × 1 convolution, three 3 × 3 dilated convolutions with different dilation rates {6, 12, 18}, and a global average pooling (GAP) to obtain feature maps with different receptive fields. As a result, the ASPP module generates five texture feature vectors, T_1–T_5, which capture the features of objects at different scales. T_1–T_5 are concatenated on the channel dimension to form the output feature tensor T and T is further processed by the SECPAM, which is designed to emphasize the distinction between ice and leads in both polarimetric and spatial dimensions. Then, a 1 × 1 convolution is applied to comprehensively incorporate feature maps of different scales and generate the encoder output E. The SAM is used to combine multi-scale semantic features (T_1–T_4) with shape features of leads. The decoder uses the low-level features from the backbone and the output features of the encoder to obtain the final prediction result.

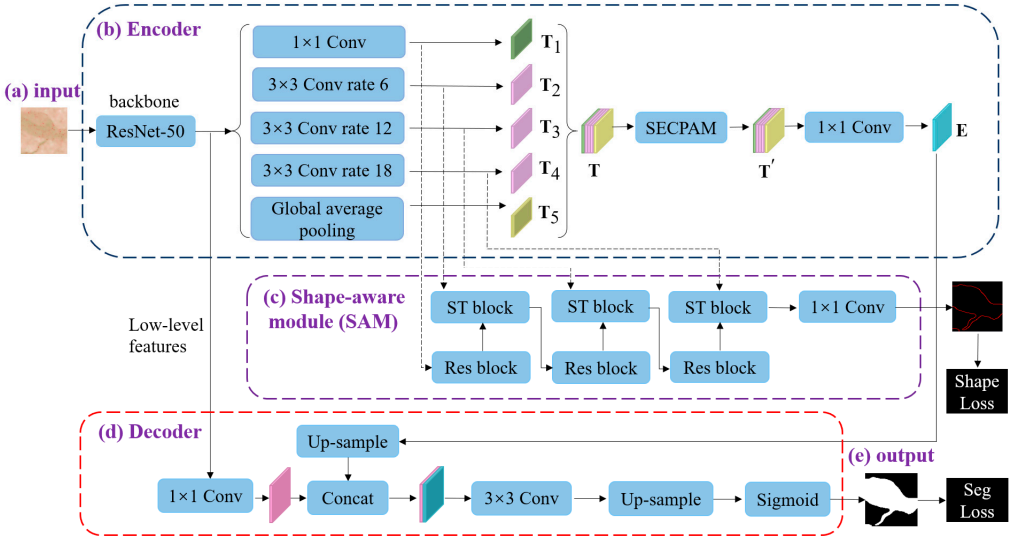


Figure 3. The overall structure of the proposed SA-DeepLabv3+. (**a**) Input: three-channel pseudo-color SAR images; (**b**) Encoder, enhanced by the proposed squeeze-and-excitation channel-position attention module (SECPAM); (**c**) Shape-aware module (SAM), composed of shape and texture integration blocks (ST block) and residual blocks (Res block); (**d**) Decoder; (**e**) Output: binary maps with white for lead and black for sea ice.

3.3. Squeeze-and-Excitation Channel-Position Attention Module

The SECPAM consists of a squeeze-and-excitation channel attention block (SE) [30] and a position attention module (PAM) [14], as shown in Figure 4. It allows the network to focus on more informative channels and better understand the spatial context of leads.

The SE channel attention module focuses on the channel information, obtains the importance of each channel by modeling the dependency relationship between channels, and readjusts it according to the importance of each channel, thus enhancing the presentation ability. The SE module includes a global average pooling, two fully connected layers, and a sigmoid activation function used to convert channel values to [0, 1] to get the output T_1 of the SE. The calculation formula is as follows:

$$T_1 = \sigma(FC(FC(AvgPool(T)))) \quad (1)$$

where σ is the sigmoid function, $AvgPool$ is the global average pooling operation and FC represents the fully connected layer with ReLU activation function.

By capturing the spatial dependency between any two positions of the feature map, PAM can learn the feature similarity between any two positions and obtain the corresponding weight map. The T_1 is first multiplied with the learned weight map β to get the weighted feature D_β. Finally, D_β is added with T_1 to get the final output feature map $T\prime$. This process can be expressed as:

$$T\prime = \varepsilon D_\beta + T_1 \tag{2}$$

where ε is a trainable parameter, which is gradually learned from 0 and assigned with weights through backpropagation. The final output features integrate local features and global features, so the feature map is optimized through the PAM module, which can help the model to segment more accurately.

In general, the SE strengthens the input features by integrating all channels, while PAM selectively aggregates the features of each position through the weighted sum of features at all positions. The combination of the two can model the original features from both channel dimensions and spatial dimensions so as to extract more useful features more effectively.

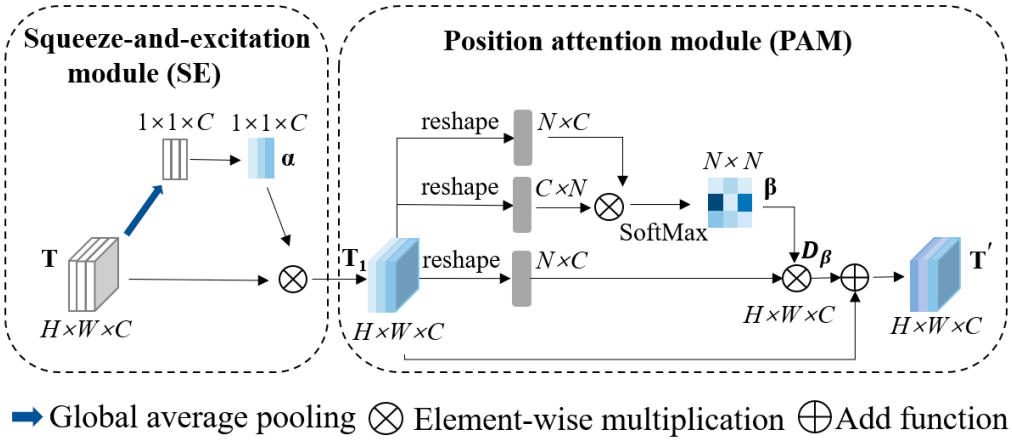


Figure 4. The structure of the squeeze-and-excitation channel-position attention module (SECPAM). *H*, *W*, and *C* are the rows, columns, and channels of the feature map, respectively ($N = H \times W$).

3.4. Shape-Aware Module

The SAM is composed of three shape and texture integration blocks (ST block), three residual blocks (Res block), and a 1×1 convolution layer; the network structure is shown in Figure 5. The feature map of the lth layer in the SAM is calculated as:

$$\begin{cases} \boldsymbol{S}_l = R(\boldsymbol{T}_l), & l = 1 \\ \boldsymbol{S}_l = R(\boldsymbol{S}'_{l-1}), & l > 1 \end{cases} \tag{3}$$

where $R(\cdot)$ denotes the Res block; $\boldsymbol{S}_l{}'$ denotes the output of the ST block; $\boldsymbol{S}l$ denotes the lth SAM feature maps; and $\boldsymbol{T}l$ represents the different scale semantic features from the encoder. The ST block, shown in Figure 6, fuses the semantic and shape feature maps, which can be expressed as follows:

$$\boldsymbol{S}_l{}' = C_{3\times3}(\boldsymbol{S}_l \oplus (\boldsymbol{S}_l \otimes \boldsymbol{\alpha}_l)) \tag{4}$$

$$\boldsymbol{\alpha}_l = \sigma(C_{1\times1}(\boldsymbol{S}_l \parallel \boldsymbol{T}_{l+1})) \tag{5}$$

where $\otimes$ and $\oplus$ denote the element-wise multiplication and addition, respectively; $\parallel$ denotes the operation of concatenating feature maps along the channel dimension; $\boldsymbol{\alpha}_l$ is computed attention map that emphasizes the lead shapes, and σ is the sigmoid function. Finally, the output feature map of the last ST block is processed by 1×1 convolution to predict a boundary map.

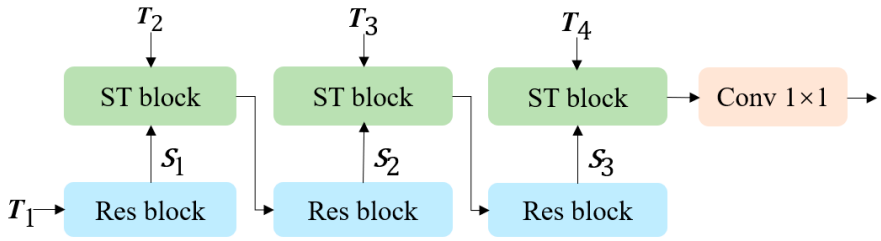


Figure 5. The structure of the shape-aware module.

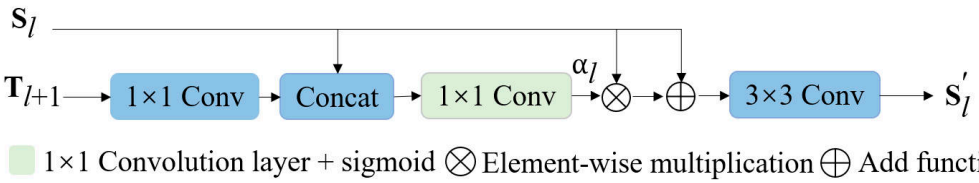


Figure 6. The structure of shape and texture integration (ST) block.

3.5. Postprocessing

Due to the fact that leads are a natural phenomenon that occurs in polar regions, they not only have their own unique formation mechanisms but are also related to environmental factors in the region, such as temperature, wind, etc. To increase the accuracy of the segmentation results and minimize the impact of regional differences, we designed a postprocessing operation as an auxiliary module in the lead detection process. Specifically, we applied two correction steps to the detection results due to the obvious linear shape and regional connectivity of the leads. Firstly, we derived the aspect ratio (AR) of the lead using the minimal bounding rectangle, and areas with AR lower than 2.2 were filtered from the results in order to preserve the long band shape of the ice lead. The setting of the AR threshold will be discussed in Section 5.3. Additionally, small holes that are misclassified inside the lead area's contour were filled.

$$AR = \frac{\max(w, h)}{\min(w, h)} \tag{6}$$

where w and h denote the width and height of the minimal bounding rectangle of the detected lead, respectively.

3.6. Jointed Loss Function

Our loss function consists of two parts: the segmentation loss term and the shape loss term, which concentrates on the texture information and shape information of the region to be segmented, respectively. Due to the fact that the number of leads in the image samples of the dataset in this study is relatively small, the traditional cross-entropy loss function will bias the model to a large number of background pixels, resulting in poor training results. Dice loss [31,32] can alleviate the negative impact of sample imbalance, but there will be loss saturation. Focal loss [33] is an upgraded version of cross-entropy loss, which can better solve the imbalance problem. Therefore, Focal loss and Dice loss are combined to measure the overlap between the ground truth and the segmentation result of the network to train the segmentation model. The form of the segmentation loss function is shown in Equation (7).

$$L_{\text{seg}} = \lambda_{\text{Dice}} * L_{\text{DiceLoss}} + \lambda_{\text{Focal}} * L_{\text{FocalLoss}} \tag{7}$$

where λ_{Dice} and λ_{Focal} are the weighting factors to balance the two losses, which are set as 1 and 2, respectively, in the experiment.

Due to the obvious linear shape of the leads, focusing on the shape information of the target can help to detect the leads. The predicted class boundary from the SAM is deeply supervised to generate the shape loss L_{shape}, which is defined as a binary cross-entropy loss

between the canny [34] edge of the ground truth and the predicted shape map, which is defined as:

$$L_{shape} = \frac{1}{N}\sum_{i=1}^{N} l_{bce}(f_{SAM}(S_i;\theta), Z_i) \quad (8)$$

where $f_{SAM}(S_i;\theta)$ represents the output shape predicted by the network on the input feature map, θ is the parameter of the SAM, and Z_i represents the real boundary obtained from the target. Minimizing shape loss enables the network to learn effective shape features, thus learning more boundary details, which is conducive to the final segmentation.

In this study, a jointed loss function is obtained by combining the above two losses, and the final total loss is expressed as:

$$L = L_{seg} + \lambda_{shape} * L_{shape} \quad (9)$$

where λ_{shape} is the weight coefficient of the balance segmentation loss and the shape loss, which is set to 1 in the experiments.

4. Experiments

4.1. Experimental Setup

All of the experiments are implemented with PyTorch 1.7.1 on a GeForce RTX TITAN GPU. The model is optimized using a stochastic gradient descent (SGD) optimizer with an initial learning rate set to 1×10^{-2}. Weight decay and momentum are set to 5×10^{-5} and 0.99, respectively.

The overall accuracy (OA), F1-score (F1), and mean Intersection over Union (mIoU) are used as evaluation indicators. The metrics mentioned can be derived from the confusion matrix using four statistical measures: True Positives (*TP*), False Positives (*FP*), True Negatives (*TN*), and False Negatives (*FN*). OA represents the ratio of correctly predicted pixels to the total number of pixels:

$$OA = \frac{TP + TN}{TP + TN + FP + FN} \quad (10)$$

F1, computed as the harmonic mean of recall (R) and precision (P), evaluates the similarity between the segmentation results predicted by the network and the ground truth.

$$F1 = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \quad (11)$$

where the recall and precision are represented as:

$$\text{Precision} = \frac{TP}{TP + FP} \quad (12)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (13)$$

mIoU measures the average overlap between the ground truth and predicted values for all classes.

$$mIoU = \frac{1}{N_{class}} \sum_{i=1}^{N_{class}} \frac{TP(i)}{TP(i) + TP(i) + FN(i)} \quad (14)$$

where $N_{class} = 2$ indicates two classes of objects, sea ice and leads.

4.2. Comparative Experiments

To verify the performance of the proposed method, comparative experiments are conducted to compare SA-DeepLabv3+ with the baseline network DeepLabv3+ [20], U-Net [35], EW-Net [1], TransUNet [36], and CL-Net [15]. U-Net and DeepLabv3+ are benchmarking DL models for remote sensing image segmentation. EW-Net was recently proposed for lead detection by taking into account the entropy features of SAR images.

Recently, Transformer has become a research hotspot in the DL community. TransUNet combines U-Net with a Transformer [37] encoder, which performs well in medical image segmentation. CL-Net is a newly proposed lightweight segmentation network for leads that is based on U-Net.

4.2.1. Quantitative Results

Table 2 shows the quantitative comparison results of six segmentation methods on the testing set. It can be seen that the three indicators of TransUNet have the lowest values. The possible reason is that TransUNet is based on the Transformer model. Because the model is complex, it often needs a lot of training data. The lack of datasets in the remote sensing field is a common problem. The dataset produced in this study is also limited, resulting in insufficient training of TransUNet and low precision of the model. The segmentation accuracy of U-Net and DeepLabv3+ is slightly higher than that of TransUNet. DeepLabv3+ expands the receptive field through the ASPP module to obtain more context information, and its performance is significantly better than U-Net in terms of indicators. EW-Net further improves network performance by effectively fusing entropy features through an entropy-weighted module. CL-Net is a lightweight lead detection model with a 0.37% higher F1, 0.62% higher mIoU, and 0.35% higher OA than EW-Net. It can be seen that the proposed method SA-DeepLabv3+ achieved the highest results in all three indicators. Specifically, compared to TransUNet, U-Net, DeepLabv3+, EW-Net, and CL-Net, the proposed method shows an improvement of 4%, 3.6%, 2.51%, 2.13%, and 1.51% in mIoU, respectively. This proves that the addition of shape information can improve the segmentation accuracy of the lead segmentation task.

Table 2. Comparison of precision between different segmentation models.

Model	F1/%	mIoU/%	OA/%
TransUNet [36]	89.24	87.26	95.11
U-Net [35]	89.50	87.66	95.03
DeepLabv3+ [20]	90.60	88.75	95.68
EW-Net [1]	91.45	89.13	95.88
CL-Net [15]	91.82	89.75	96.23
SA-DeepLabv3+	93.01	91.48	96.82

4.2.2. Qualitative Results

Figure 7 shows the visual comparison results of the six methods on three SAR image patches from the testing set. The TransUNet segmentation results in Figure 7b show obvious misclassifications and incomplete lead contours. The segmentation results of U-Net and DeepLabv3+ are superior to those of TransUNet, as shown in Figure 7c,d. However, some lead pixels are still misclassified as sea ice pixels, presenting small holes in the detected lead areas and a few fractured leads. Compared to Figure 7c,d, the results of the EW-Net and CL-Net in Figure 7e,f are obviously improved with less noise and misclassifications in the middle of the leads. However, there are still lead disconnections and visible misclassifications at the edges of the leads. It should be noted that our implementation of the EW-Net and CL-Net followed the network architecture description in [1,15], but its input differs from the original model. The proposed model achieves the best performance. As shown in Figure 7g, the linear shape and contour of the leads identified by the proposed model are more complete and more distinct. This is attributed to both the shape information processed by the SAM and the improved encoder feature processed by the SECPAM. In addition, some holes are removed by the postprocessing. Therefore, we can conclude that the proposed improvement on DeepLabv3+ is effective and the proposed method outperforms the compared existing models.

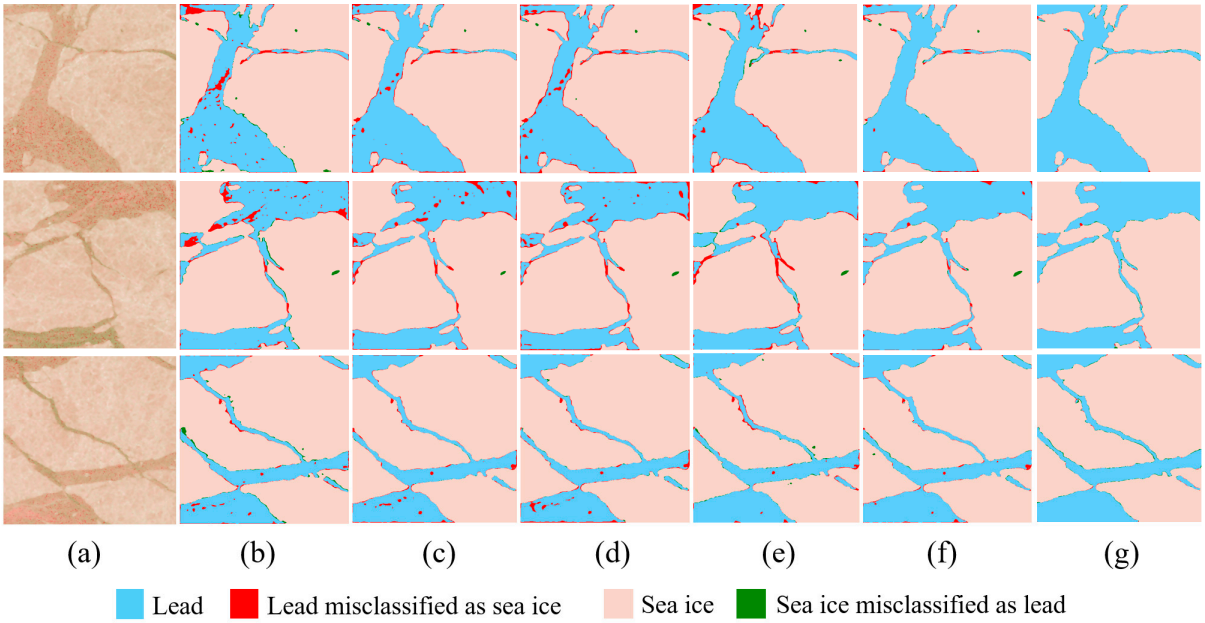


Figure 7. Visualized results of different lead detection models on the three subregion SAR images from the test set. (**a**) the pseudo-color SAR images; (**b**) results of TransUNet; (**c**) results of U-Net; (**d**) results of DeepLabv3+; (**e**) results of EW-Net; (**f**) results of CL-Net; (**g**) results of SA-DeepLabv3+.

4.3. Ablation Experiments

To study the effect of the SECPAM and SAM, ablation experiments were carried out. Table 3 shows the comparison results of DeepLabv3+ as the baseline network with different combinations of the SECPAM and SAM. Figure 8 demonstrates the visualized segmentation results of different situations, for example, the sub-region of SAR images selected from the test set.

Table 3. Comparison results of DeepLabv3+ with different combinations of the SECPAM and SAM.

Model	F1/%	mIoU/%	OA/%
baseline	90.60	88.75	95.68
baseline + SECPAM	91.04	89.21	95.88
baseline + SAM	91.53	89.77	96.10
baseline + SECPAM + SAM	92.78	91.26	96.75

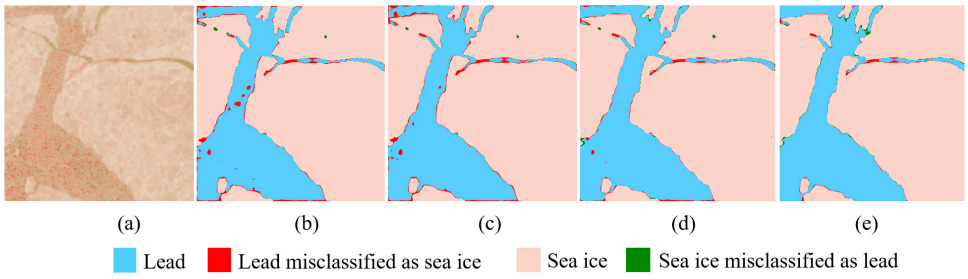


Figure 8. Visualized results of different models on the one example sub-region of SAR image from the test set. (**a**) the pseudo-color SAR image; (**b**) result of the baseline network; (**c**) result of the baseline and SECPAM; (**d**) result of the baseline and SAM; (**e**) result of the baseline and SECPAM and SAM.

It can be seen from Table 3 that the addition of SECPAM has improved all the evaluation indicators compared with the baseline network. Especially, the SECPAM module increases F1 by 0.44%, mIoU by 0.46% and OA by 0.20%. The false detection rate in segmentation results is significantly reduced, as shown in Figure 8c. The SECPAM module

can learn the importance of different channels and positions and adaptively adjust features. Therefore, DeepLabv3+ with SECPAM can reduce error detection when compared to the DeepLabv3+. The SAM increases F1 by 0.93%, mIoU by 0.95%, and OA by 0.42%, bringing more improvement than the SECPAM. This proves that shape features of leads fused into the segmentation network can produce positive feedback to the segmentation results and improve the details of the leads (Figure 8d), demonstrating the effectiveness of the proposed SAM. Moreover, the addition of both the SECPAM and SAM results in a considerable increase in F1, mIoU, and OA of 2.18%, 2.51%, and 1.07%, respectively. The network with the proposed two modules (Figure 8e) further improved the segmentation results with more complete shape details, yielding a segmentation result that is closer to the ground truth compared with other results. This shows that the SECPAM can capture complex channel and spatial relationships in the images, improving the model's ability to achieve accurate segmentation; on the other hand, the SAM allows the network to focus on shape information.

The ST block in the SAM computes an attention map, α_l, of boundaries by using information from the SAM and multi-scale semantic features, where l denotes the layer number in our SAM. As there are three ST blocks in the SAM, three learned shape attention maps can be extracted from the outputs of these ST blocks, and they are able to demonstrate how the SAM pays more attention to the lead shapes.

Figure 9 contains three samples from the test set, their ground truth (GT) maps, and the intermediate shape stream attentions α_l. Figure 8 clearly illustrates the correlation between the attention maps in the SAM and the original input images. It can be seen that as the layer increases, the shape information learned by the attention map becomes more refined and specific, demonstrating its ability to learn the correct shape characteristics.

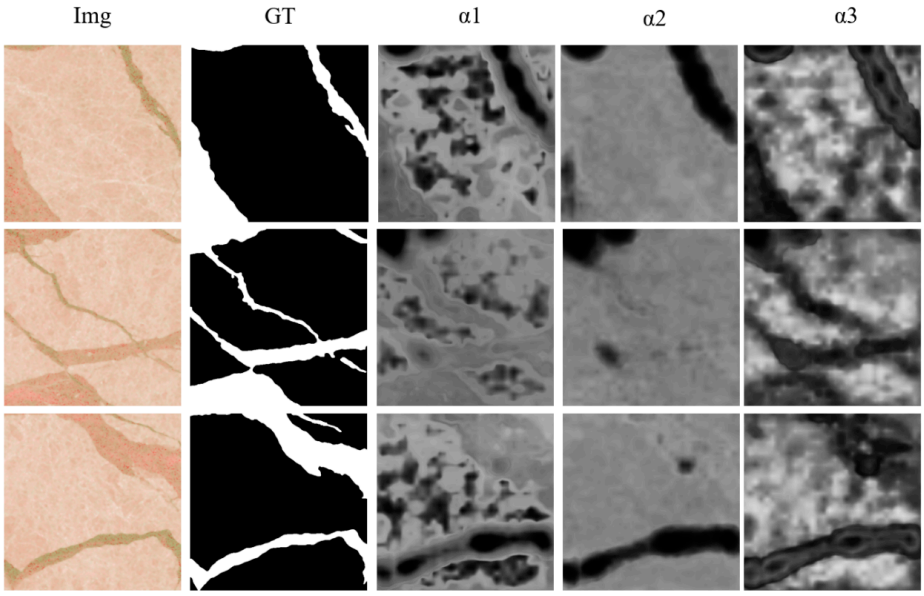


Figure 9. Visual comparison of the SAM's attentions. α_l is l_{th} shape attention map.

5. Discussion

5.1. Different Attention Modules

We use different attention modules in the network to compare their performance in lead detection. Mix attention (MA) [38] is one of the latest attention modules based on lightweight CNN and ViT; it combines channel, spatial, and global context attention while enhancing both the feature representation of the target itself and the correlation between targets. Furthermore, the convolutional block attention module (CBAM) [39] is a highly cited module that combines channel attention and spatial attention.

Table 4 shows the segmentation performance of the proposed model using different attention modules. It is evident that the segmentation results achieved by utilizing a single type of attention mechanism (such as SE, PAM) are inferior to those obtained by modules that combine multiple types of attention (such as MA, CBAM, and SECPAM). Furthermore, the segmentation accuracy of the proposed attention mechanism SECPAM module is higher than that of MA and CBAM. Therefore, the SECPAM effectively exploits the relationships between different channels and positions in SAR images, providing valuable guidance for lead detection. Unlike the general image field, the SECPAM module proposed in this paper may be more competitive in remote sensing image segmentation.

Table 4. Segmentation performance of SA-DeepLabv3+ with different attention modules.

Attention Module	F1/%	mIoU/%	OA/%
SE [30]	88.70	86.70	94.82
PAM [14]	89.57	87.64	95.20
MA [38]	89.87	87.93	95.28
CBAM [39]	90.64	88.76	95.64
SECPAM	92.78	91.26	96.75

Figure 10 illustrates the segmentation results of the proposed model using different attention modules. It is observed that when employing the SE attention module and PAM attention module, the model produces the highest number of misclassified pixels, resulting in fragmented lead shapes and numerous holes, as depicted in Figure 10b,c. However, with the adoption of the MA and CBAM attention modules, the segmentation accuracy of the model significantly improves, evidenced by the reduction in misclassified areas, as shown in Figure 10d,e. Moreover, when utilizing the SECPAM attention module proposed in this paper, the model generates more complete lead shapes with a greatly reduced misclassification rate, as shown in Figure 10f. This highlights the significant advantage of the proposed SECPAM attention module in the domain of lead segmentation.

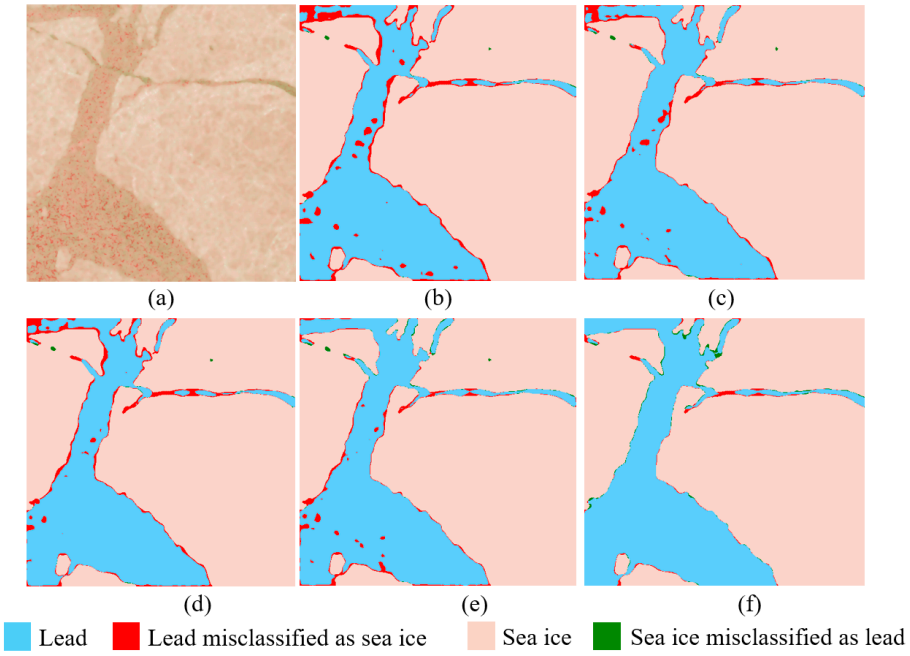


Figure 10. Segmentation results of the proposed model using different attention modules. (**a**) the pseudo-color SAR image; (**b**) result of the proposed model with SE; (**c**) result of the proposed model with PAM; (**d**) result of the proposed model with MA; (**e**) result of the proposed model with CBAM; (**f**) result of the proposed model with SECPAM.

5.2. Different Inputs

Leads with open water or thin ice, in most cases, have low surface roughness and, therefore, have low backscatter values on both SAR polarization modes HH and HV, appearing dark. The reason for using the band product HH × HV is that if both HH and HV have low backscatter intensities, HH x HV can amplify the difference between leads and open water. Leads with open water, however, can appear bright (high intensity) on the HH polarization if wind roughens the water surface and appear dark (low intensity) on the HV polarization under the same conditions [40,41]. In the band ratio, HH/HV image leads appear bright if HH is high and HV is low.

To investigate the impact of different polarization data on lead detection, we evaluated the model performance using combinations of HH and HV polarizations as input. The quantitative experimental results are presented in Table 5, while the visual segmentation results are depicted in Figure 11. From the comparisons, it can be concluded that the combined use of dual-polarization data can compensate for information deficiency when single-polarization data are used alone, enabling the model to learn more and improve segmentation accuracy. This is evident in the segmentation results, where compared to Figure 11a,b, fewer pixels of leads are misclassified as sea ice pixels in Figure 11c. Additionally, adding HH × HV as input enhances segmentation performance, as shown in Figure 11d. The model's segmentation performance is optimal when the fused data are incorporated, with leads exhibiting complete contours and significantly reduced internal hole phenomena. Conversely, adding HH/HV results in a decrease in all indicators, as depicted in Figure 11e, where numerous misclassified pixels appear at the edges of leads. This could be attributed to the exclusion of high wind speed effects during dataset construction, selecting only SAR data under low wind speed conditions, thus favoring HH × HV for lead detection and potentially introducing interference and misclassification with HH/HV. Furthermore, incorporating data from all channels leads to the lowest segmentation accuracy, as reflected in Figure 11f, where leads lose their original shapes, and numerous pixels of the lead are misclassified as sea ice pixels. It is evident that more information does not necessarily lead to better segmentation results; only the addition of data conducive to lead detection can enhance segmentation accuracy.

Table 5. Segmentation results of SA-DeepLabv3+ with different inputs.

Input	F1/%	mIoU/%	OA/%
HH	89.58	87.65	95.20
HV	89.26	87.42	95.08
HH, HV	91.35	89.60	96.04
HH, HV, HH × HV	92.78	91.26	96.75
HH, HV, HH/HV	90.98	89.21	95.87
HH, HV, HH/HV, HH × HV	83.49	81.86	93.03

To verify the impact of HH × HV data in different models, we use SAR dual polarization data (HH, HV) as inputs to train different segmentation networks and compare them with our dataset. The experimental results are illustrated in Table 6. It can be seen that, compared to using HH and HV channels as inputs, the addition of HH × HV improves the segmentation performance of all networks, which proves the effectiveness of our dataset.

Table 6. Comparison of different inputs on different models.

Model	Input Channel	F1/%	mIoU/%	OA/%
TransUNet [36]	2	88.16	86.34	94.21
	3	89.24	87.26	95.11
U-Net [35]	2	88.45	86.61	94.63
	3	89.50	87.66	95.03

Table 6. *Cont.*

Model	Input Channel	F1/%	mIoU/%	OA/%
DeepLabv3+ [20]	2	90.03	88.11	95.37
	3	90.60	88.75	95.68
EW-Net [1]	2	90.87	88.67	95.56
	3	91.45	89.13	95.88
CL-Net [15]	2	91.23	89.27	95.91
	3	91.82	89.75	96.23
SA-DeepLabv3+	2	91.85	90.27	96.31
	3	92.78	91.26	96.75

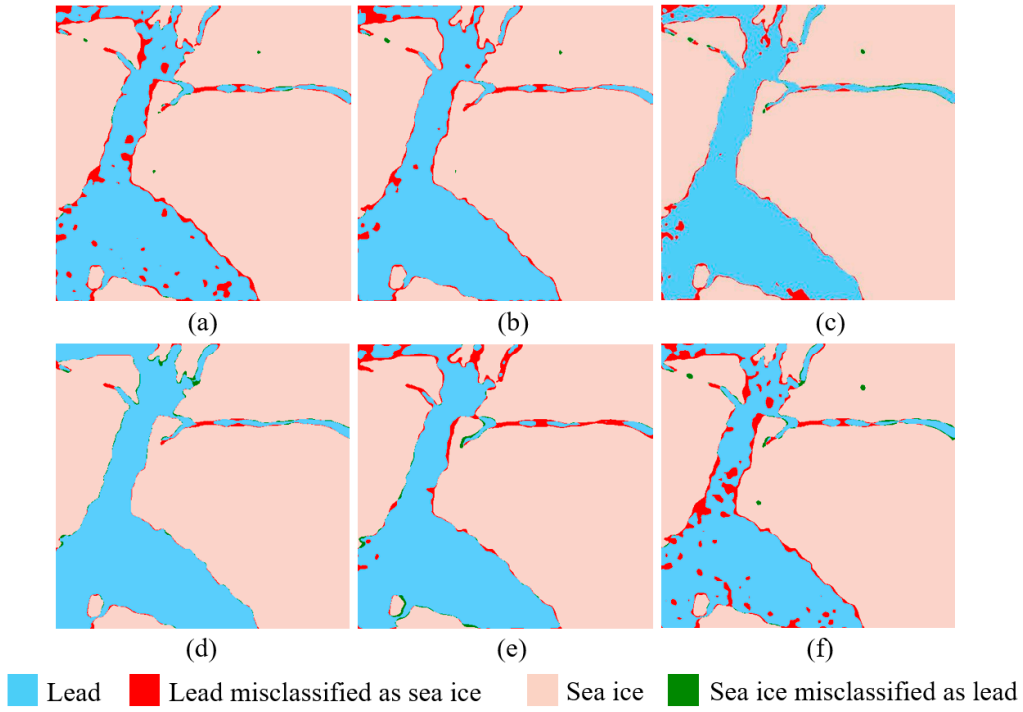


Figure 11. Segmentation results of the proposed model using different inputs. (**a**) HH; (**b**) HV; (**c**) HH, HV; (**d**) HH, HV, HH × HV; (**e**) HH, HV, HH/HV; (**f**) HH, HV, HH/HV, HH × HV.

5.3. The Impact of Postprocessing

In order to ensure that the detected areas correspond to leads that meet certain length and width requirements, postprocessing operations are necessary. We calculate the aspect ratio of the detected lead using the minimal bounding rectangle, and areas with an aspect ratio below a certain threshold (AR) are filtered out from the results. We explore the effect of different AR values ranging from 1.0 to 4.0 on the algorithm. As shown in Figure 12, the segmentation accuracy initially increases and then decreases as the AR value increases. After conducting a thorough analysis, we have determined that a threshold AR value of 2.2 yields the best results. This value strikes a balance between capturing leads with the desired length and width requirements while maintaining high segmentation accuracy.

We also compare the segmentation effects of different models with and without postprocessing. After adding postprocessing, all three indicators of the six networks have been improved (as shown in Table 7), so this operation is necessary.

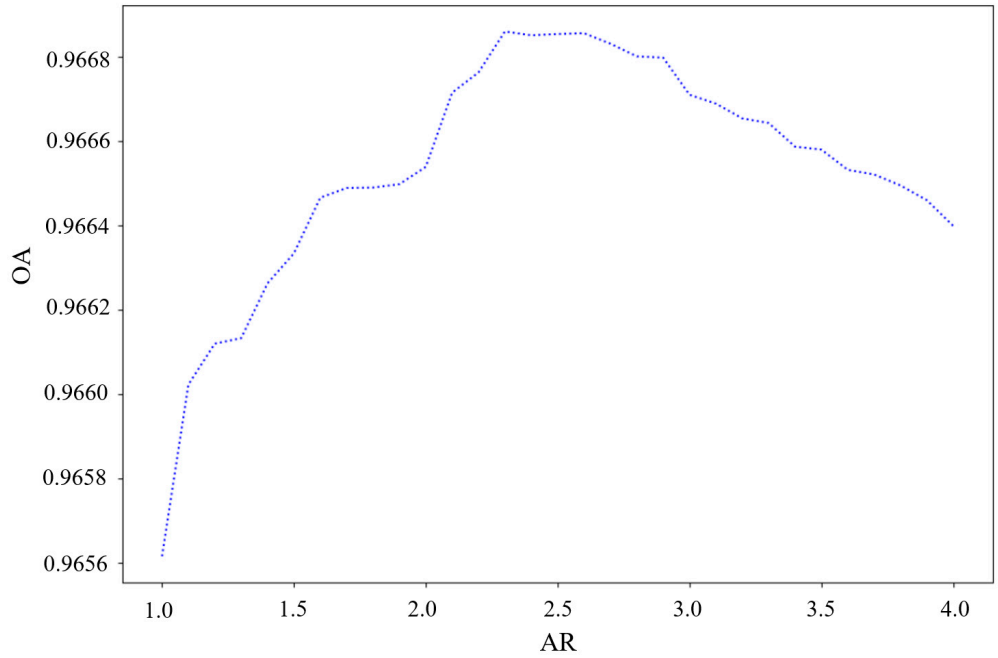


Figure 12. Segmentation results of SA-DeepLabv3+ after using postprocessing operations with different threshold AR.

Table 7. Comparison of with (w/) and without (w/o) postprocessing on different models.

Model	Postprocessing	F1/%	mIoU/%	OA/%
TransUNet [36]	w/o	89.24	87.26	95.11
	w/	89.83	87.91	95.39
U-Net [35]	w/o	89.50	87.66	95.03
	w/	90.46	88.66	95.51
DeepLabv3+ [20]	w/o	90.60	88.75	95.68
	w/	91.43	89.69	96.07
EW-Net [1]	w/o	91.45	89.13	95.88
	w/	91.83	90.03	96.12
CL-Net [15]	w/o	91.82	89.75	96.23
	w/	92.56	90.32	96.57
SA-DeepLabv3+	w/o	92.78	91.26	96.75
	w/	93.01	91.48	96.82

5.4. Training Details

To evaluate the training effort of the compared models in this paper, we recorded the training and validation losses of each model with the number of training epochs, as well as the time required to train each model. Figure 13 illustrates the variations of training–validation losses within 30 epochs, and Table 8 presents the average training time required for each epoch of different models.

Table 8. Training time per epoch for different segmentation models.

	TransUNet	U-Net	DeepLabv3+	EW-Net	CL-Net	SA-DeepLabv3+
Training time (min/epoch)	60.2	54.3	53.2	44.5	42.7	43.1

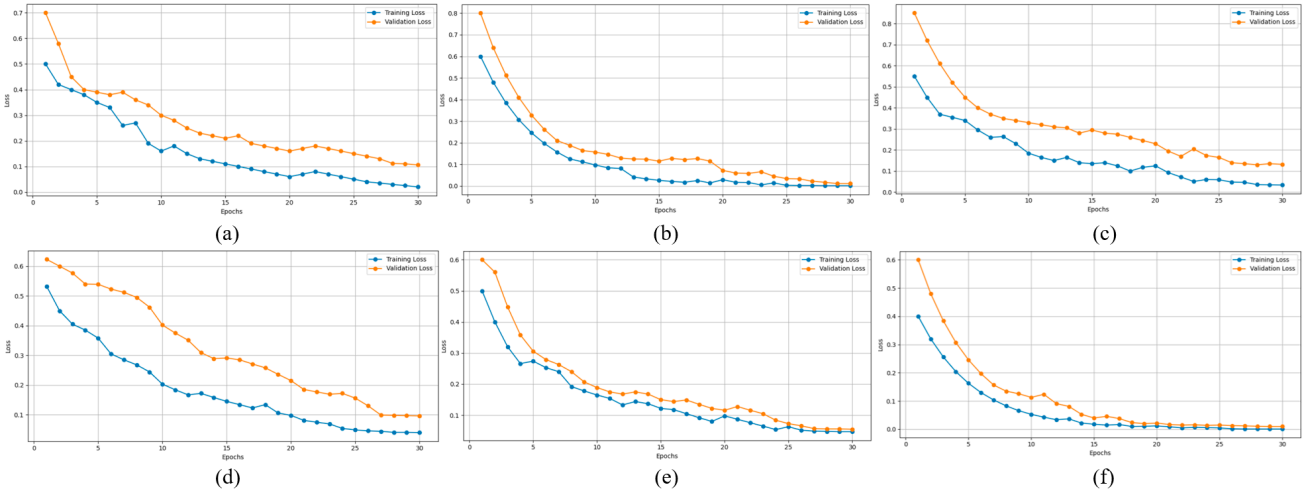


Figure 13. The variation of training–validation losses for different segmentation models. (**a**) TransUNet; (**b**) U-Net; (**c**) DeepLabv3+; (**d**) EW-Net; (**e**) CL-Net; (**f**) SA-DeepLabv3+.

From Figure 13, it can be observed that the TransUNet (Figure 13a), DeepLabv3+ (Figure 13c), and EW-Net (Figure 13d) converge slower than other models. U-Net and CL-Net (which is based on U-Net) exhibit faster convergency, which may benefit from their simple network structure. Our model, SA-DeepLabv3+, demonstrates the fastest convergence, even though it is based on DeepLabv3+. This can be largely attributed to the use of the shape-aware module. The faster convergence rates suggest higher training efficiency.

Table 8 reveals that the TransUNet model requires the longest training time per epoch, followed by U-Net and DeepLabv3+, while EW-Net and CL-Net require relatively shorter training times. The proposed model's training time is slightly longer than that of CL-Net. This is expected because CL-Net is a lightweight model.

Taking into account the convergence of training–validation losses and training time of the models, the proposed model demonstrates superior performance in this study, with lower training–validation losses and shorter training time.

6. Conclusions

In this paper, we propose a model for lead segmentation built upon the DeepLabv3+. To address the problem of the lack of lead datasets, a sea ice lead dataset fusing dual-polarization SAR data is constructed, and the effectiveness of the dataset in lead segmentation is demonstrated through experiments. Secondly, since shape information has not been considered in existing research on lead segmentation, a shape-aware module is proposed that can combine multi-scale semantic features and shape information and capture linear shapes of leads. We also design a squeeze-and-excitation channel-position attention module to enhance the encoder features. The experiments show that the proposed method outperforms the DeepLabv3+ and other existing benchmarking methods. These experiments confirm the effectiveness of the proposed designs, which are derived from the physical mechanisms of radar remote sensing and guided by the lead characteristics in SAR images.

However, some limitations of this work should be mentioned. Due to the significant impact of wind speed on the growth and movement status of leads, the SAR data used in this study were all collected in low wind speed scenarios, thus excluding the influence of wind speed. Therefore, future research can focus on the segmentation of leads under high wind speed conditions and incorporate more varied data to enhance the robustness of the model in different environments and regions. Additionally, due to the high real-time requirements of deep learning models when facing applications, our future research will design a lightweight model based on the model proposed in this paper.

To facilitate further sea ice lead research, we have made all the code utilized in this paper openly accessible and available for public use. You can find the code at the following link: https://github.com/0814zm/Lead-detection (accessed on 10 May 2024).

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

Funding: This research was funded by the National Natural Science Foundation of China (Grant No. 42006159, 61972240).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The Sentinel-1 SAR data used in this study are available at https://scihub.copernicus.eu (accessed on 20 May 2021). The code is available at https://github.com/0814zm/Lead-detection (accessed on 10 May 2024).

Acknowledgments: We would like to acknowledge the European Commission and European Space Agency (ESA) for supplying the Sentinel-1 SAR data. We gratefully thank the National Natural Science Foundation of China (Grant No. 42006159, 61972240) for supporting this study.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Liang, Z.; Pang, X.; Ji, Q.; Zhao, X.; Li, G.; Chen, Y. An Entropy-Weighted Network for Polar Sea Ice Open Lead Detection from Sentinel-1 SAR Images. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 4304714. [CrossRef]
2. Murashkin, D.; Spreen, G.; Huntemann, M.; Dierking, W. Method for Detection of Leads from Sentinel-1 SAR Images. *Ann. Glaciol.* **2018**, *59*, 124–136. [CrossRef]
3. Longepe, N.; Thibaut, P.; Vadaine, R.; Poisson, J.-C.; Guillot, A.; Boy, F.; Picot, N.; Borde, F. Comparative Evaluation of Sea Ice Lead Detection Based on SAR Imagery and Altimeter Data. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 4050–4061. [CrossRef]
4. Zhong, W.; Jiang, M.; Xu, K.; Jia, Y. Arctic Sea Ice Lead Detection from Chinese HY-2B Radar Altimeter Data. *Remote Sens.* **2023**, *15*, 516. [CrossRef]
5. Qu, M.; Pang, X.; Zhao, X.; Lei, R.; Ji, Q.; Liu, Y.; Chen, Y. Spring Leads in the Beaufort Sea and Its Interannual Trend Using Terra/MODIS Thermal Imagery. *Remote Sens. Environ.* **2021**, *256*, 112342. [CrossRef]
6. Gao, Y.; Gao, F.; Dong, J.; Wang, S. Transferred Deep Learning for Sea Ice Change Detection from Synthetic-Aperture Radar Images. *IEEE Geosci. Remote Sens. Lett.* **2019**, *16*, 1655–1659. [CrossRef]
7. Hong, D.-B.; Yang, C.-S. Automatic Discrimination Approach of Sea Ice in the Arctic Ocean Using Sentinel-1 Extra Wide Swath Dual-Polarized SAR Data. *Int. J. Remote Sens.* **2018**, *39*, 4469–4483. [CrossRef]
8. Jiang, M.; Xu, L.; Clausi, D.A. Sea Ice–Water Classification of RADARSAT-2 Imagery Based on Residual Neural Networks (ResNet) with Regional Pooling. *Remote Sens.* **2022**, *14*, 3025. [CrossRef]
9. Komarov, A.S.; Buehner, M. Detection of First-Year and Multi-Year Sea Ice from Dual-Polarization SAR Images Under Cold Conditions. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 9109–9123. [CrossRef]
10. Lu, Y.; Zhang, B.; Perrie, W. Arctic Sea Ice and Open Water Classification from Spaceborne Fully Polarimetric Synthetic Aperture Radar. *IEEE Trans. Geosci. Remote Sens.* **2023**, *61*, 4203713. [CrossRef]
11. Zakhvatkina, N.; Smirnov, V.; Bychkova, I.; Stepanov, V. Detection of the Leads in the Arctic Drifting Sea Ice on SAR Images. In Proceedings of the 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, Brussels, Belgium, 11–16 July 2021; pp. 4276–4279.
12. Khaleghian, S.; Ullah, H.; Kræmer, T.; Hughes, N.; Eltoft, T.; Marinoni, A. Sea Ice Classification of SAR Imagery Based on Convolution Neural Networks. *Remote Sens.* **2021**, *13*, 1734. [CrossRef]
13. Komarov, A.S.; Buehner, M. Adaptive Probability Thresholding in Automated Ice and Open Water Detection From RADARSAT-2 Images. *IEEE Geosci. Remote Sens. Lett.* **2018**, *15*, 552–556. [CrossRef]
14. Ren, Y.; Li, X.; Yang, X.; Xu, H. Development of a Dual-Attention U-Net Model for Sea Ice and Open Water Classification on SAR Images. *IEEE Geosci. Remote Sens. Lett.* **2022**, *19*, 4010205. [CrossRef]
15. Liu, S.; Li, M.; Xu, M.; Zeng, Z. An Improved Lightweight U-Net for Sea Ice Lead Extraction from Multi-Polarization SAR Images. *IEEE Geosci. Remote Sens. Lett.* **2023**, *20*, 2000705. [CrossRef]
16. Gao, F.; Wang, X.; Gao, Y.; Dong, J.; Wang, S. Sea Ice Change Detection in SAR Images Based on Convolutional-Wavelet Neural Networks. *IEEE Geosci. Remote Sens. Lett.* **2019**, *16*, 1240–1244. [CrossRef]

17. Zakhvatkina, N.Y.; Alexandrov, V.Y.; Johannessen, O.M.; Sandven, S.; Frolov, I.Y. Classification of Sea Ice Types in ENVISAT Synthetic Aperture Radar Images. *IEEE Trans. Geosci. Remote Sens.* **2013**, *51*, 2587–2600. [CrossRef]
18. Karvonen, J.; Simila, M.; Makynen, M. Open Water Detection from Baltic Sea Ice Radarsat-1 SAR Imagery. *IEEE Geosci. Remote Sens. Lett.* **2005**, *2*, 275–279. [CrossRef]
19. 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. [CrossRef]
20. Chen, L.-C.; Zhu, Y.; Papandreou, G.; Schroff, F.; Adam, H. Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Hermany, 8–14 September 2018; pp. 801–818.
21. Kirillov, A.; Mintun, E.; Ravi, N.; Mao, H.; Rolland, C.; Gustafson, L.; Xiao, T.; Whitehead, S.; Berg, A.C.; Lo, W.-Y.; et al. Segment Anything. In Proceedings of the 2023 IEEE/CVF International Conference on Computer Vision (ICCV), Paris, France, 2–6 October 2023; IEEE: Piscataway, NJ, USA, 2023; pp. 3992–4003.
22. Moghimi, A.; Welzel, M.; Celik, T.; Schlurmann, T. A Comparative Performance Analysis of Popular Deep Learning Models and Segment Anything Model (SAM) for River Water Segmentation in Close-Range Remote Sensing Imagery. *IEEE Access* **2024**, *12*, 52067–52085. [CrossRef]
23. Liu, B.; Li, X.; Zheng, G. Coastal Inundation Mapping from Bitemporal and Dual-Polarization SAR Imagery Based on Deep Convolutional Neural Networks. *JGR Oceans* **2019**, *124*, 9101–9113. [CrossRef]
24. Liu, G.; Liu, B.; Zheng, G.; Li, X. Environment Monitoring of Shanghai Nanhui Intertidal Zone with Dual-Polarimetric SAR Data Based on Deep Learning. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 4208918. [CrossRef]
25. Murashkin, D.; Spreen, G. Sea Ice Leads Detected from Sentinel-1 SAR Images. In Proceedings of the IGARSS 2019—2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 28 July–2 August 2019; pp. 174–177.
26. Filipponi, F. Sentinel-1 GRD Preprocessing Workflow. In Proceedings of the 3rd International Electronic Conference on Remote Sensing, Virtual, 22 May–5 June 2019; p. 11.
27. Passaro, M.; Müller, F.L.; Dettmering, D. Lead Detection Using Cryosat-2 Delay-Doppler Processing and Sentinel-1 SAR Images. *Adv. Space Res.* **2018**, *62*, 1610–1625. [CrossRef]
28. 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, 27–30 June 2016; pp. 770–778.
29. Chen, L.-C.; Papandreou, G.; Schroff, F.; Adam, H. Rethinking Atrous Convolution for Semantic Image Segmentation. *arXiv* **2017**, arXiv:1706.05587.
30. Hu, J.; Shen, L.; Albanie, S.; Sun, G.; Wu, E. Squeeze-and-Excitation Networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), San Francisco, CA, USA, 18–20 June 2018; pp. 7132–7141.
31. Milletari, F.; Navab, N.; Ahmadi, S.-A. V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation. In Proceedings of the 2016 Fourth International Conference on 3D Vision (3DV), Stanford, CA, USA, 25–28 October 2016; pp. 565–571.
32. Li, X.; Sun, X.; Meng, Y.; Liang, J.; Wu, F.; Li, J. Dice Loss for Data-Imbalanced NLP Tasks. *arXiv* **2020**, arXiv:1911.02855.
33. Lin, T.-Y.; Goyal, P.; Girshick, R.; He, K.; Dollár, P. Focal Loss for Dense Object Detection. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017.
34. Canny, J. A Computational Approach to Edge Detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **1986**, *PAMI-8*, 679–698. [CrossRef]
35. Ronneberger, O.; Fischer, P.; Brox, T. *U-Net: Convolutional Networks for Biomedical Image Segmentation*; Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F., Eds.; Springer International Publishing: Munich, Germany, 2015; pp. 234–241.
36. Chen, J.; Lu, Y.; Yu, Q.; Luo, X.; Adeli, E.; Wang, Y.; Lu, L.; Yuille, A.L.; Zhou, Y. TransUNet: Transformers Make Strong Encoders for Medical Image Segmentation. *arXiv* **2021**, arXiv:2102.04306.
37. Han, K.; Xiao, A.; Wu, E.; Guo, J.; Xu, C.; Wang, Y. Transformer in Transformer. *arXiv* **2021**, arXiv:2103.00112.
38. Guan, R.; Man, K.L.; Zhao, H.; Zhang, R.; Yao, S.; Smith, J.; Lim, E.G.; Yue, Y. MAN and CAT: Mix Attention to Nn and Concatenate Attention to YOLO. *J. Supercomput.* **2023**, *79*, 2108–2136. [CrossRef]
39. Woo, S.; Park, J.; Lee, J.-Y.; Kweon, I.S. CBAM: Convolutional Block Attention Module. In *Computer Vision—ECCV 2018*; Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y., Eds.; Lecture Notes in Computer Science; Springer International Publishing: Cham, Switzerland, 2018; Volume 11211, pp. 3–19.
40. Dierking, W. Sea Ice Monitoring by Synthetic Aperture Radar. *Oceanography* **2013**, *26*, 100–111. [CrossRef]
41. Scharien, R.K.; Yackel, J.J. Analysis of Surface Roughness and Morphology of First-Year Sea Ice Melt Ponds: Implications for Microwave Scattering. *IEEE Trans. Geosci. Remote Sens.* **2005**, *43*, 2927–2939. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.