

Article

Not peer-reviewed version

Integrating Remote Sensing Image and Sea Ice Recognition for Optimal Polar Ship Path Planning

[Jiayi Bian](#) , [Li Zhou](#) ^{*} , [Shifeng Ding](#) , [Pentti Kujala](#) , Sen Han , [Dinghan Zeng](#) , [Roger Skjetne](#)

Posted Date: 29 September 2024

doi: 10.20944/preprints202409.2220.v1

Keywords: Computer vision; You Only Look Once (YOLOv5); remote sensing image; sea ice; ship path planning



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

Integrating Remote Sensing Image and Sea Ice Recognition for Optimal Polar Ship Path Planning

Jiayi Bian ¹, Li Zhou ^{2,*}, Shifeng Ding ³, Pentti Kujala ³, Sen Han ¹, Dinghan Zeng ¹
and Roger Skjetne ⁴

¹ School of Naval Architecture and Ocean Engineering, Jiangsu University of Science and Technology, Zhenjiang, 212100, China

² State Key Laboratory of Ocean Engineering, Shanghai Jiao Tong University, Shanghai, 200030, China

³ School of Ocean and Civil Engineering, Shanghai Jiao Tong University, Shanghai, 200240, China

⁴ Norwegian University of Science and Technology, Trondheim, Norway

* Correspondence: zhouli209@sjtu.edu.cn

Abstract: A large amount of sea ice exists in the polar regions, which is prone to ship-ice collisions leading to ship breakage and sinking events. Complex environmental factors pose challenges to navigation safety, and it is particularly important to perceive the polar environment and make decisions in advance. In this paper, You only look once (YOLOv5) is targeted to be optimized according to the characteristics of ice remote sensing images, which includes adding Squeeze-and-Excitation (SE) attention mechanism, improving spatial pyramid pooling, and replacing Flexible Rectified Linear Unit (FReLU) activation function. In order to verify the effect of the optimized algorithm, ablation experiments are set up, and the accuracy of the optimized algorithm is improved by 3.5% compared with the original YOLOv5. Comparison experiments with other target identification algorithms are conducted, and the results show that the optimized YOLOv5 has higher accuracy and better robustness, and can identify small target sea ice that is difficult to identify in large-size remote sensing images. Polar ship path planning based on the identification results. The identification output of remote sensing ice images is used as the input for the construction of path planning maps, and raster maps corresponding to the actual polar scenes are built. Taking the path length, the number of ship turns, and the sea ice risk value as the objective function, the Any-Angle Path Planning on Grids (Theta*) algorithm is used to simulate the path planning under different sea conditions with different sea ice concentration, and a path with a shorter length, a smoother path, and a higher safety coefficient is found.

Keywords: computer vision; You Only Look Once (YOLOv5); remote sensing image; sea ice; ship path planning

1. Introduction

Arctic sea ice has declined by about 40% since the 1980s. With the gradual melting of sea ice, attempts have been made to navigate the Arctic shipping lanes, which include the Northeast Passage, the Northwest Passage and the Central Passage [1]. Compared with traditional shipping path, Arctic shipping path significantly reduce shipping time and costs, reduce the load on traditional shipping path, reduce the dependence of ships on fixed path, avoid complex shipping areas, improve safety of navigation, and provide more path options for maritime trade [2]. Observation and research on sea ice and polar targets can help to improve the safety of navigation.

Obtaining sea ice information is mainly based on field measurement method, ship's walkover observation method and satellite remote sensing observation method. Lu et al. [3] proposed a two-stream radiative transfer model for ponded sea ice; Weissling et al. [4] used a video camera to photograph sea ice and study the image sequences; Worby et al. [5] analyzed the distribution characteristics of Antarctic sea ice based on more than 20,000 samples observed by Antarctic ship's walking voyage; Cai et al. [6] employed convolutional neural networks to segment sea ice by instances using a simulation ice pool dataset, and estimated ice size and ice concentration; Ressel et al. [7] utilized an artificial neural network to classify ice, and the results demonstrated that the method is resistant to image noise.

After collecting sea ice data, accurately extracting sea ice feature parameters has become a focus of current research, which mainly focuses on recognizing sea ice targets in images, and the methods used mainly include threshold segmentation, target recognition and instance segmentation [8]. In order to improve the accuracy of sea ice recognition and extract sea ice feature parameters, sea ice recognition based on computer vision has become an increasingly popular research direction. Currently, the mainstream target recognition algorithms include YOLO (You Only Look Once) algorithm and SSD (Single Shot MultiBox Detector) algorithm for target classification, Mask R-CNN (Mask Region-based Convolutional Neural Network) [9–11]. Lu et al. [12] proposed a two-stream radiative transfer model for ponded sea ice. The upwelling irradiance from the pond surface is determined and then its spectrum is transformed into RGB color space. Cai et al. [13] utilized Convolutional Neural Networks for instance segmentation of sea ice by using the photos of ice pools as a dataset, and proposed an ice-breaking radius fitting method to fit the radius and angle of the ship when breaking ice. Dong et al. [14] developed a two-staged ice channel identification method based on image segmentation and corner point regression. They employed the image segmentation method to extract channel regions, and proposed an intelligent corner regression network to extract the channel boundary lines from the channel region.

In the process of polar navigation, influenced by high latitude, low temperature, sea ice and other unique environmental factors, the safe navigation of ships is facing a greater threat, and the demand for reliable path planning is becoming higher and higher. Unlike path planning for other conventional sea areas, path planning for polar ships requires comprehensive consideration, which involves the concentration and thickness of sea ice and other key ice parameters.

Choiet al. [15] obtained sea ice concentration information by remote sensing means, used thermodynamic model to invert the ice thickness, through the analysis of sea ice concentration and sea ice thickness information, and based on this data for Arctic ship path planning. Kotovirta et al. [16] studied the Baltic Sea region, introduced ship transportation model, ice model and optimization model, and designed an ice navigation system based on them. Frederking [17] used the Keinonen method to predict the effects of ship characteristics and ice conditions on ship navigational performance, and the sea ice ice conditions were described using the sea ice code published by the World Meteorological Organization. Wu et al. [18] developed an online interactive route planning system for ships sailing in the Arctic based on big Earth data to enhance the safety of shipping navigation and improve information extraction methods for ARs (Arctic Routes).

Current research on polar ship path planning mostly focuses on optimizing traditional algorithms or combining multiple algorithms to achieve improved algorithm performance. In this paper, the combination of target recognition technology and polar ship path planning can extract rich sea ice parameter information from remote sensing images, which provides more reliable inputs for the path planning algorithm and makes the generated path scheme more in line with the actual sea ice conditions.

2. Remote Sensing Sea Ice Image Detection

2.1. Introduction of YOLOv5

YOLO is an end-to-end target detection algorithm based on deep learning, which has a faster detection speed and can meet the real-time requirements when ships are sailing in polar regions [19]. The model structure is simple and efficient, with strong scalability. The detection accuracy of small targets is relatively high, which is suitable for the identification of small sea ice targets in remote sensing images in complex polar environments. The network structure of YOLOv5 mainly contains Input, Backbone, Neck and Head, and the network structure is shown in Figure 1.

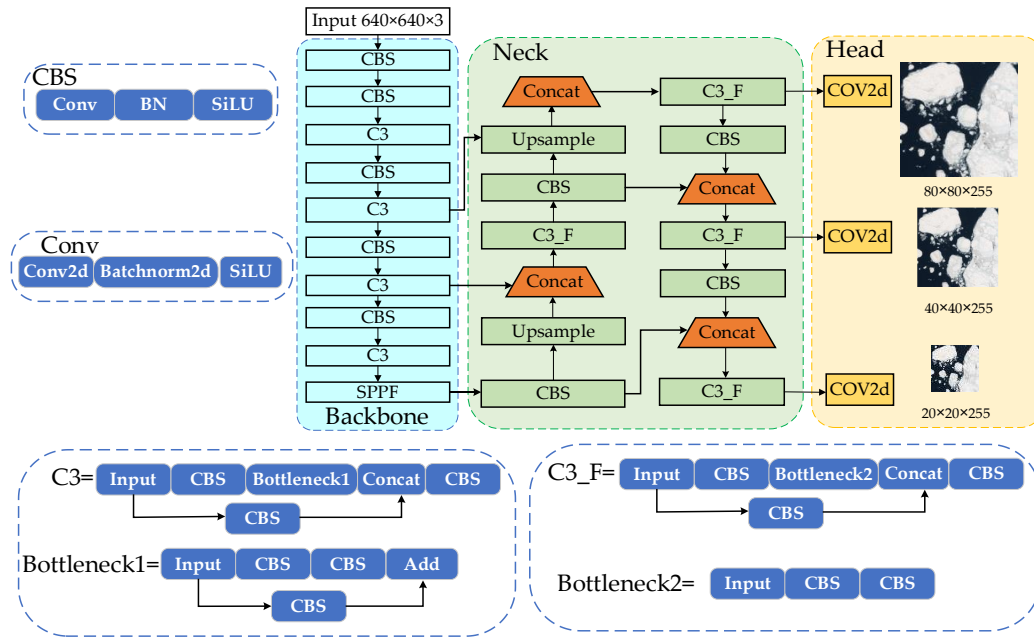


Figure 1. Structure of the YOLOv5 model.

The input side is responsible for pre-processing the input image, including operations such as resizing and normalization, to ensure the consistency of the input data. The input images are scaled uniformly to the standard input size to meet the input requirements of the network. In order to avoid distortion or information loss introduced during the scaling process, this preprocessing step adopts the method of maintaining the aspect ratio of the image by first calculating the aspect ratio of the image, and then scaling it to the standard size uniformly according to this ratio, and the blank area is filled with gray scale bars.

The backbone network contains CBS structure, C3 structure and SPPF (Spatial Pyramid Pooling with Features) structure. These modules together form an efficient and less computationally intensive feature extraction network. The CBS consists of three components, namely Convolution, Batch Normalization, and SiLU (Sigmoid-Weighted Linear Unit) activation function. The C3 structure consists of three standard Convolution and Bottleneck layer modules. Layer and Bottleneck Layer modules, with the C3 structure, features can be extracted and fused efficiently while maintaining a low computational complexity. SPPF uses different sizes of maximal pooling to increase the sensory field. The feature map is manipulated using pooling layers of different scales and used for the construction of the feature pyramid to obtain a multi-scale feature representation.

The Neck network of YOLOv5 is located between the backbone network and the output, which is responsible for fusing feature maps of different scales to enhance the feature expression capability of the network. The Neck module is composed of a Feature Pyramid Network (FPN) as well as a Pyramid Attention Network (PAN). The FPN structure transmits high-level semantic features from top to bottom, and the PAN structure realizes comprehensive information coverage by transmitting low-level spatial features downward, so that the feature maps of each size contain both semantic and spatial information of the target.

The output side is the last part of YOLOv5, which is responsible for outputting the target category and correcting the position of the candidate box according to the position offset to get more accurate detection results. The output side employs a convolutional layer to convert the feature map into a prediction result, and also introduces anchor frames as a priori information for predicting the target, the size of which is adaptively calculated based on the statistical information of the training set. The output covers the loss function and NMS, which work together to improve the stability and reliability of the network. The NMS algorithm is utilized to filter a best result from multiple prediction frames and eliminate redundant detection [20].

2.2. YOLOv5 Optimization

The main difference between remote sensing satellite image recognition and general image recognition is that its image size is huge, while the target size in it is small and usually clustered together, resulting in recognition difficulties, as shown in Figure 2 [21]. The sea ice targets to be recognized are very small and difficult to recognize relative to the large size of the remote sensing image, and in areas of dense sea ice, multiple pieces of sea ice may also be close to each other or partially obscured.

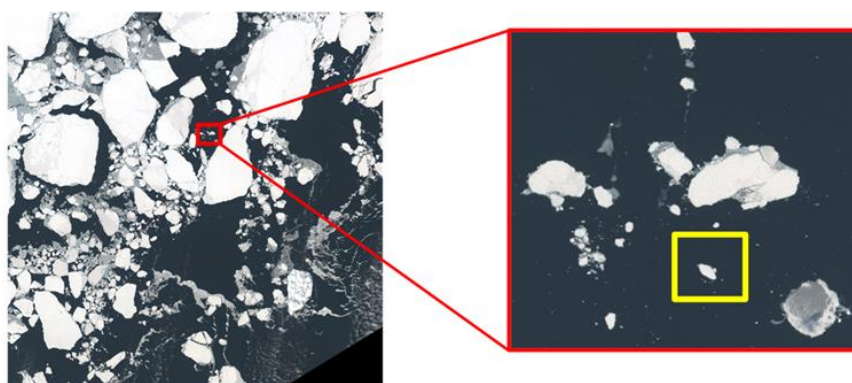


Figure 2. Recognition task for Remote sensing scenarios.

In this paper, the YOLOv5 algorithm is improved in three aspects. First, the SE (Squeeze-and-Excitation Network) attention mechanism is added to the Backbone of the original model, which enables the neural network to better capture the features of the sea-ice images, and improves the performance and generalization ability of the network. Secondly, the structure of SPPF spatial pyramid pooling is improved, and the SPPCSPC-F (Fast Spatial Pyramid Pooling and Cross Stage Partial Network) module is used to enhance the feature representation of sea ice. Finally, the FReLU (Flexible Rectified Linear Unit) activation function, which is more suitable for the target recognition task, is used to replace the SiLU (Sigmoid Gated Linear Unit) activation function in the original model to improve the accuracy of sea ice recognition.

2.2.1. Squeeze-and-Excitation Networks (SE) attention mechanism

Attention mechanisms mainly contain three types: spatial, channel and hybrid domains. The SE model is a typical representative of the channel domain attention mechanism, focusing on the adaptive weight assignment of feature channels. In the working mechanism of this model, the feature map first undergoes compression in the spatial dimension, and then different weight values are applied to each feature channel, which characterize the relative importance of the information carried by each channel [22]. The initial feature map is recalibrated according to the obtained weights to realize the enhancement of critical feature channels and the suppression of non-critical channels.

Due to the large size of the remote sensing image and the small size of the target sea ice, it is easy to lose some key information when performing the identification, and there is a leakage of small target sea ice. For this reason, the SE attention mechanism is added in layer 9 of the YOLOv5 backbone network Backbone. The core part of this module is the two steps of Squeeze and Excitation, which adaptively adjusts the importance of each channel by learning the weights, so that the neural network can better capture the features of the sea ice image, and improve the performance of the network without adding too much computational burden. The structure of SE Attention Mechanism is shown in Figure 3.

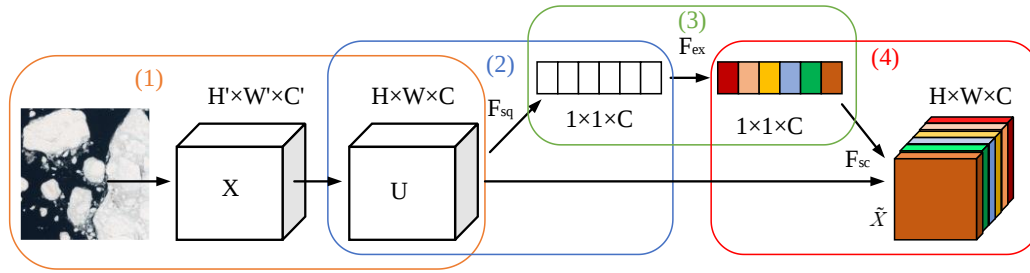


Figure 3. Structure of the SE attention mechanism.

2.2.2. SPPCSPC-F Spatial Pyramid Pooling

SPPF (Fast Spatial Pyramid Pooling) spatial pyramid pooling is used in the YOLOv5 model. SPPF is an improvement on the structure of SPP. The convolutional kernel CBL in the internal structure of SPP consists of Convolution, Standard Normalization, and Leaky ReLU, respectively. The structure of SPPF is superior to that of SPP, in that the maximal pooling layers will be in chunks by connecting them in series, thus speeding up the computation. The structure of SPPF is shown in Figure 4. shown, the convolution kernel CBS consists of convolution, standard normalization, and SiLU activation function, and its structure contains three consecutive maximal pooling layers of size 5×5 , which can acquire three different scales of sensory fields.

Although the maximum pooling operation can expand the sensory field to obtain rich contextual information, this nonlinear downsampling reduces the spatial resolution of the feature maps, and may lose some of the discriminative information in the original feature maps, making the detection of small targets ineffective. And this pooling structure easily leads to overfitting, which needs to be avoided by using more training data or regularization.

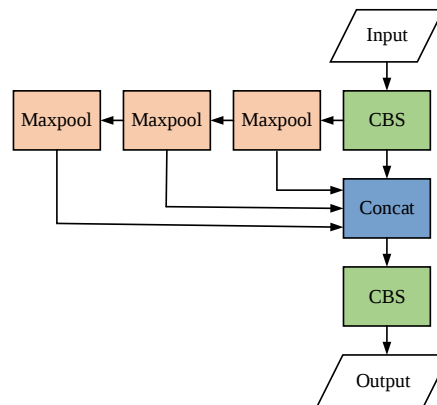


Figure 4. Structure of the SPPF module.

In order to solve the above problems of SPPF, the advantages of the SPPCSPC structure in the YOLOv7 network model are integrated to improve the structure of SPPF, and the SPPCSPC-F module is obtained, and the SPP structure in the SPPCSPC is changed into the SPPF structure, which is due to the fact that the SPPF structure has better accuracy and speed, and it is placed in layer 10 of the YOLOv5 backbone network. layer 10 of the YOLOv5 backbone network, and its structure is shown in Figure 5.

The SPPCSPC-F module first splits the input feature map into two branches by channel. One part of the branch undergoes maximum pooling at three different scales to obtain multi-scale sea ice information. The other part of the branch goes through 1×1 convolution directly to maintain the original resolution. Then the two parts of the features are connected by channel, so that the multi-scale features are obtained and the original detail information is retained, and finally the two convolutions are further feature fused. The SPPCSPC-F module modifies the order of the maximum pooling, which retains the details while keeping the sensory field unchanged, and enhances the feature expression capability with stronger feature fusion ability.

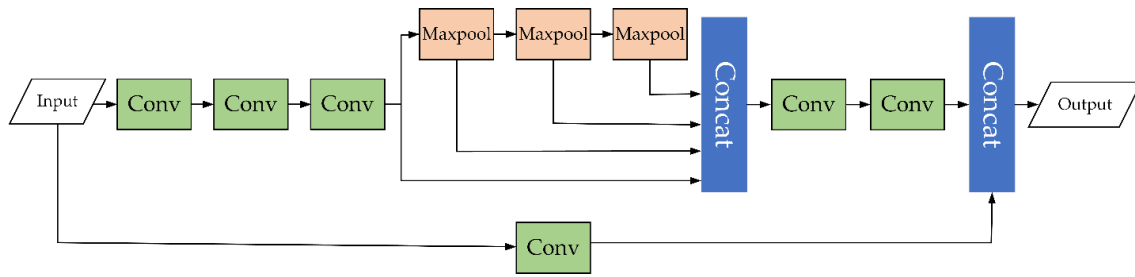


Figure 5. Structure of the SPPCSPC-F module.

2.2.3. FReLU Activation Function

The original YOLOv5 algorithm uses the SiLU activation function. This traditional activation function has some limitations, such as when the input value is far away from 0, the derivative of the SiLU activation function will tend to 0, which will lead to the problem of vanishing gradient. The vanishing gradient will make it difficult for the model to perform effective backpropagation, which will make it difficult for the network to converge or cause instability in training, reducing the efficiency of training and even leading to loss of information.

In this paper, we use the FReLU activation function, which is more suitable for the target recognition task, to replace the SiLU activation function. FReLU is a kind of funnel function, which is obtained by the improvement of the ReLU activation function, and extends the ReLU function by adding a spatial condition to expand the space to two dimensions, which is a relatively simple process to realize, and only adds a small computational overhead [23]. The structure of the two activation functions is shown in Figure 6.

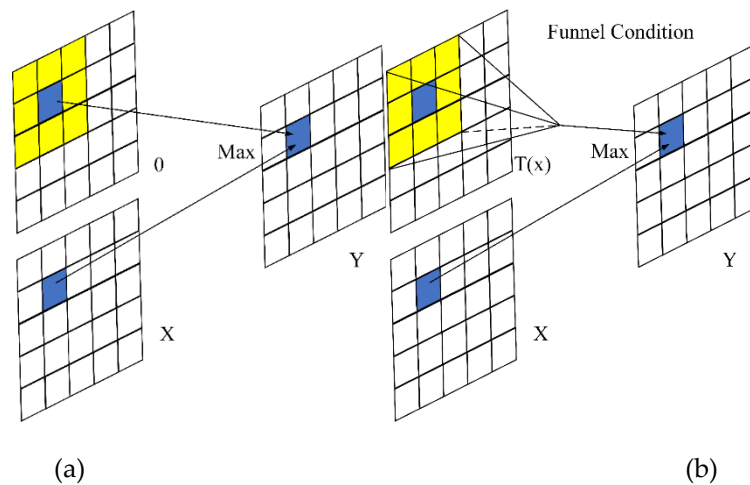


Figure 6. Activation function of ReLU and FReLU. (a) ReLU activation function conditioned on 0; (b) FReLU activation functions with visualization conditions.

The FReLU activation function incorporates learnable parameters that allow the network to adaptively adjust the shape of the function through learning. This flexibility enhances the learning ability of the model and better adapts to the characteristics of sea ice images, and the advantages of the FReLU activation function in nonlinear transformation and feature enhancement can improve the performance of the target recognition model. Combining the SE attention mechanism and the FReLU activation function enables the model to learn more important feature representations, focus on key object regions, mitigate the overfitting problem and improve the generalization ability of the model when recognizing sea ice with smaller sizes in remote sensing images. The synergistic effect of the two enables the model to obtain better recognition performance under limited data conditions.

2.2.4. Optimizing the Overall YOLOv5 Framework

The SE attention mechanism is added to the 9th layer of the backbone network, and the spatial pyramid pooling structure is improved in the 10th layer, while the SiLU activation function is replaced with the FReLU activation function, and correspondingly, the CBS layer of the network is changed to the CBF layer to optimize the structure. The YOLOv5 model is shown in Figure 7.

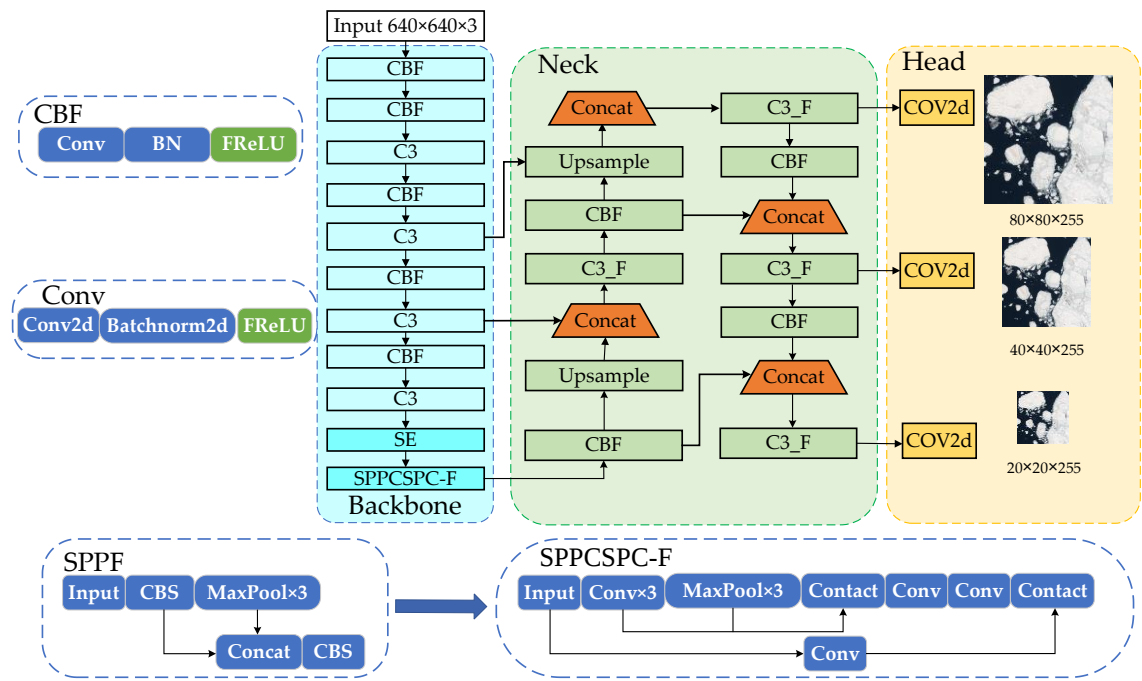


Figure 7. Structure of the improved YOLOv5 model.

The optimization of YOLOv5 increases the depth and size of the network, which adds some computational overhead, but as the depth of the network increases, the model is able to learn more complex and abstract representations of sea ice features, improving the accuracy of the algorithm. Table 1. and Table 2. show the parameters of the network before and after YOLOv5 optimization.

Table 1. Original YOLOv5 network parameters.

Layer	Module	f	n	Params
0	Conv	-1	1	3520
1	Conv	-1	1	18560
2	C3	-1	1	18816
3	Conv	-1	1	73984
4	C3	-1	2	115712
5	Conv	-1	1	295424
6	C3	-1	3	625152
7	Conv	-1	1	1180672
8	C3	-1	1	1182720
9	SPPF	-1	1	656896
10	Conv	-1	1	131584
11	Upsample	-1	1	0
12	Concat	[-1,6]	1	0
13	C3_F	-1	1	361984
14	Conv	-1	1	33024
15	Upsample	-1	1	0

16	Concat	[-1,4]	1	0
17	C3_F	-1	1	90880
18	Conv	-1	1	147712
19	Concat	[-1,14]	1	0
20	C3_F	-1	1	296448
21	Conv	-1	1	590336
22	Concat	[-1,10]	1	0
23	C3_F	-1	1	1182720

Table 2. Optimized YOLOv5 network parameters.

Layer	Module	f	n	Params
0	Conv	-1	1	3872
1	Conv	-1	1	19264
2	C3	-1	1	20928
3	Conv	-1	1	75392
4	C3	-1	2	121344
5	Conv	-1	1	298240
6	C3	-1	3	639232
7	Conv	-1	1	1186304
8	C3	-1	1	1199616
9	SE	-1	1	32768
10	SPPCSPC-F	-1	1	7124480
11	Conv	-1	1	134400
12	Upsample	-1	1	0
13	Concat	[-1,6]	1	0
14	C3_F	-1	1	370432
15	Conv	-1	1	34432
16	Upsample	-1	1	0
17	Concat	[-1,4]	1	0
18	C3_F	-1	1	95104
19	Conv	-1	1	149120
20	Concat	[-1,15]	1	0
21	C3_F	-1	1	304896
22	Conv	-1	1	593152
23	Concat	[-1,11]	1	0
24	C3_F	-1	1	1199616

where: f denotes the input layer, -1 represents the input from the previous layer, n denotes the connection of the layer to the input layer, and Params denotes the number of parameters of the network.

2.3. Experimental Results

2.3.1. Construction and Parameterization of the Data Set

The remote sensing sea ice dataset used in this paper comes from three main sources. The first is the sea ice images in the NWPU Dataset released by the Northwestern Polytechnical University [24], the second is the Arctic sea ice images downloaded from Google Earth (<http://earthengine.google.com/>), and the third is the remote sensing images provided by the

Norwegian University of Science and Technology (NTNU). The three different sources of remote sensing images are in different environments and weather conditions, which can improve the model's ability to learn sea ice features under different sea conditions, reduce the risk of overfitting, and obtain a more three-dimensional and comprehensive representation of data features.

The remote sensing sea ice dataset is constructed after the steps of de-weighting, manual labeling, and auditing. The label name is "ice", the number of images is 600, and the number of labels is 15948. The number of labels is large relative to the number of images, which is due to the large number of sea ice targets in a large-size remote sensing image and is very dense. The dataset is randomly divided into training set and test set according to the ratio of 8:2. Some of the data in the sea ice remote sensing dataset are shown in Figure 8.

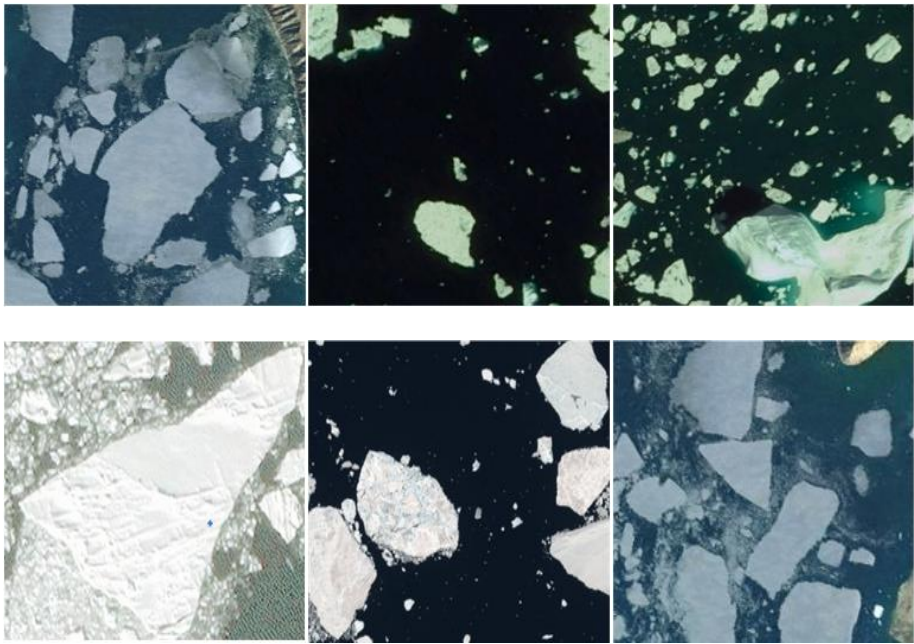


Figure 8. Remote sensing sea ice dataset.

The training parameters are set as follows, the number of iterations is 300, the initial learning rate is 0.001, the momentum parameter is 0.9, the weight decay parameter is 0.0005, and the threshold of the non-great suppression ratio is 0.5. Evaluation is carried out every 30 rounds of training. The detailed hardware and software parameters are shown in Table 3. Considering that the model cannot be adequately evaluated using only the accuracy P or only the recall R, the F1 value is used as a comprehensive index to reconcile the accuracy P and recall R values, which can comprehensively evaluate the quality of the optimization model, and the larger the F1 value indicates that the quality of the model is higher.

Table 3. Hardware and software configurations and versions.

Hardware and software configuration	Models and Versions
operating system	Window10
CPU, Central Processing Unit	Intel Xeon W-2255
Graphics Card GPU	NVIDIA Quadro P620
Deep Learning Platform	Pytorch
Pytorch version	1.10.2
CUDA version	11.3
CUDNN version	8.2.1
Python version	3.9

The value AP (Average Precision) and the mean average precision mAP (mean Average Precision) are generally used in the field of target recognition to evaluate the quality of algorithms. Two metrics, P (Precision) and R (Recall), are used to plot PR curves (Precision Recall Curve) and calculate mAP by integrating them; P quantifies the effectiveness of sample classification and R quantifies the ability to detect positive samples [25].

The precision rate P is the probability of identifying correctly in all positive samples, also known as the check rate in the definition of model prediction. Recall R is the probability of identifying correctly in all positive samples. The average precision AP can be a more comprehensive measure of the model, the recall rate R indicator as the horizontal coordinate, the accuracy rate P indicator as the vertical coordinate, one time to plot the PR curve, the PR curve and the area enclosed by the horizontal coordinate is the average precision AP. The calculation of the mAP is generally divided into two steps: the first step is to calculate the average of each category in the AP of each category in the dataset, and the second step is to take the average value after summing the average accuracies of each category. Overall, a good target recognition model should have both high accuracy P and recall R, and further, a high mAP value. The relevant equations are shown in Equations (1)-(5).

$$\text{Precision} = \frac{TP}{TP + FP} \quad (1)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (2)$$

$$F1 = 2 \frac{P \cdot R}{P + R} \quad (3)$$

$$AP = \int_0^1 P \cdot RdR \quad (4)$$

$$mAP = \frac{\sum_{i=1}^k AP_i}{k} \quad (5)$$

where: the number of samples correctly categorized as positive samples is known as TP (True positives), the number of samples incorrectly categorized as positive samples is known as FP (False Positives), the number of samples The number of correctly categorized negative samples is called TN (True Negatives) and the number of incorrectly categorized negative samples is called FN (False Negatives).

2.3.2. Ablation Experiments

To validate the model, ablation experiments were set up. The results of the ablation experiments are shown in Table 4 and Figure 9.

Table 4. Ablation study results.

Name	P	R	F1	mAP
YOLOv5	0.719	0.684	0.701	0.719
YOLOv5+SE	0.731	0.701	0.716	0.738
YOLOv5+ SPPCSPC-F	0.737	0.688	0.712	0.743

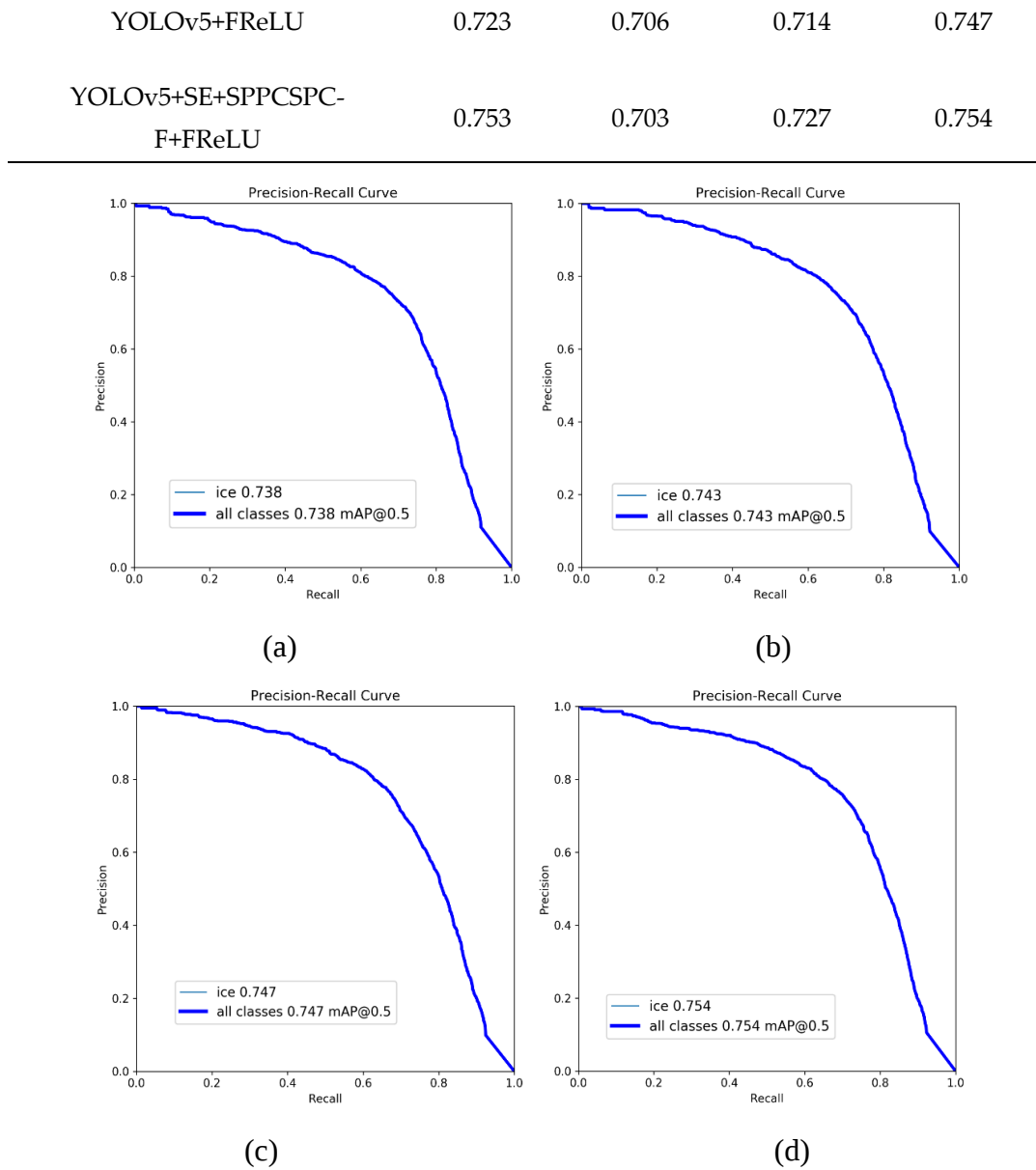


Figure 9. AP values for each model in the ablation experiment. (a) YOLOv5+SE; (b) YOLOv5+SPPCSPC-F; (c) YOLOv5+FReLU; (d) YOLOv5+SE+SPPCSPC-F+FReLU.

Based on the results of the ablation experiments, it can be seen that the original YOLOv5 has a mAP of 0.719, which is the lowest of all the evaluated models. Adding the SE attention mechanism improves the mAP by 1.9% to 0.738. adding SPPCSPC-F spatial pyramid pooling improves the mAP by 2.4% to 0.743, however the R-value is relatively low at 0.688. replacing the FReLU activation function improves the mAP by 2.8% to 0.747. when the three optimizations were used simultaneously, the mAP improved by 3.5% to 0.754, the best performing set of models in the ablation experiments.

Similarly, the P, R, and F1 values of the original YOLOv5 were 0.719, 0.684, and 0.701, respectively. After optimization, the P, R, and F1 values were 0.753, 0.703, and 0.727, which were improved by 3.4%, 1.9%, and 1.8%, respectively. The results show that the improved YOLOv5 has higher accuracy in recognizing remote sensing sea ice images.

2.3.3. Comparison Experiments

To further validate the effectiveness of optimizing YOLOv5, this paper sets up a comparison experiment. The optimized algorithm is compared with the original YOLOv5 and other target

detection models such as Faster-RCNN, YOLOv3, YOLOv4 and the current newer YOLOv8 model, and the Loss value and mAP value of each model are calculated, and the comparison results are shown in Figure 10.

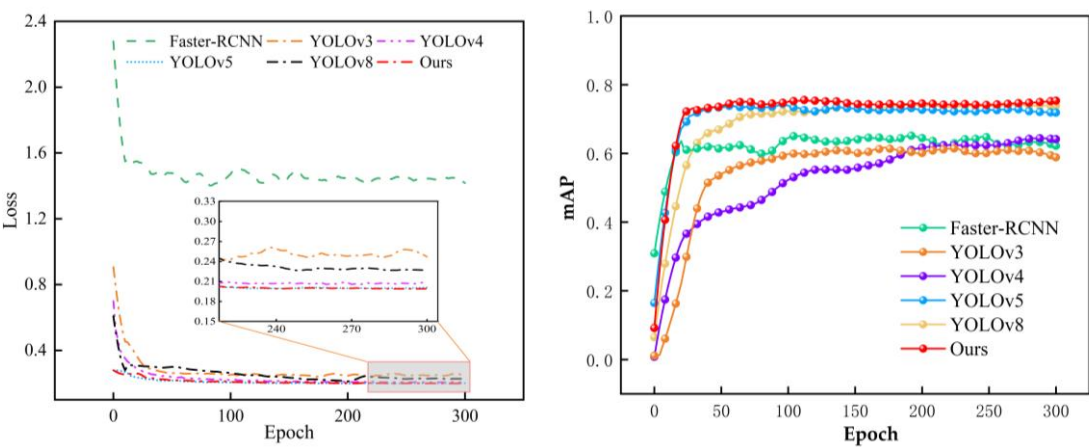


Figure 10. The loss and mAP value of each model in comparison experiment.

From the trend of Loss value, it can be seen that in the first 40 epochs, the Loss value of each model decreases rapidly, which indicates that the network is learning the features of the sea ice rapidly and the training has not yet reached the stable stage. After 200 epochs, the training is gradually stabilized, in which the optimization model has a lower Loss value than the other algorithms, which indicates that the optimization of YOLOv5 has a fast convergence speed. All algorithms converge to stability at 250 epochs. In the stabilization phase, the optimized model has lower Loss value and higher mAP value, which indicates better generalization ability and detection performance of optimized YOLOv5 compared to other models.

The individual trained models are evaluated and the results are compared as follows Table 5. shows. Compared to Faster-RCNN, YOLOv3, YOLOv4, YOLOv5, and YOLOv8, the mAP values of the optimized YOLOv5 model are improved by 15%, 10.6%, 9.9%, 3.5%, and 1.3%, respectively. Among them, YOLOv3 has the lowest mAP of 0.604, and the optimized algorithm has the highest mAP of 0.754. YOLOv3, YOLOv4, and YOLOv8 have higher accuracy value P, but lower recall R. The F1 values of 0.552, 0.636, and 0.617 indicate that there is ice under detection for the identification of targets by the three models, and the optimized YOLOv5 has the highest F1 value of 0.727. From the results of the comparison experiments, it can be seen that the optimized YOLOv5 can better identify sea ice targets in remote sensing images.

Table 5. Contrast study results.

mould	P	R	F1	mAP
Faster-RCNN	0.641	0.632	0.636	0.655
YOLOv3	0.858	0.407	0.552	0.604
YOLOv4	0.757	0.548	0.636	0.648
YOLOv5	0.719	0.684	0.701	0.719
YOLOv8	0.839	0.488	0.617	0.741
Ours	0.753	0.703	0.727	0.754

The large-size sea ice remote sensing images are recognized using the optimized YOLOv5. Since the sea ice targets are too dense, the confidence level is hidden in the result graph in order to better show the recognition effect. The recognition results of the algorithm before and after the optimization and the local zoomed-in area are shown in Figure 11. In the same area, the number of sea ice recognized by the original YOLOv5 is 14 and 55, and the number of sea ice recognized by the

optimized YOLOv5 is 53 and 88. In comparison, the number of recognized sea ice has increased by 39 and 33 respectively, and most of them are small targets that are difficult to detect.

The results of remote sensing sea ice image recognition can provide environmental information for ship path planning, and the detailed sea ice information provides reliable input data for the ship path planning algorithm, which helps to generate the optimal path plan that is more in line with the actual ice conditions.

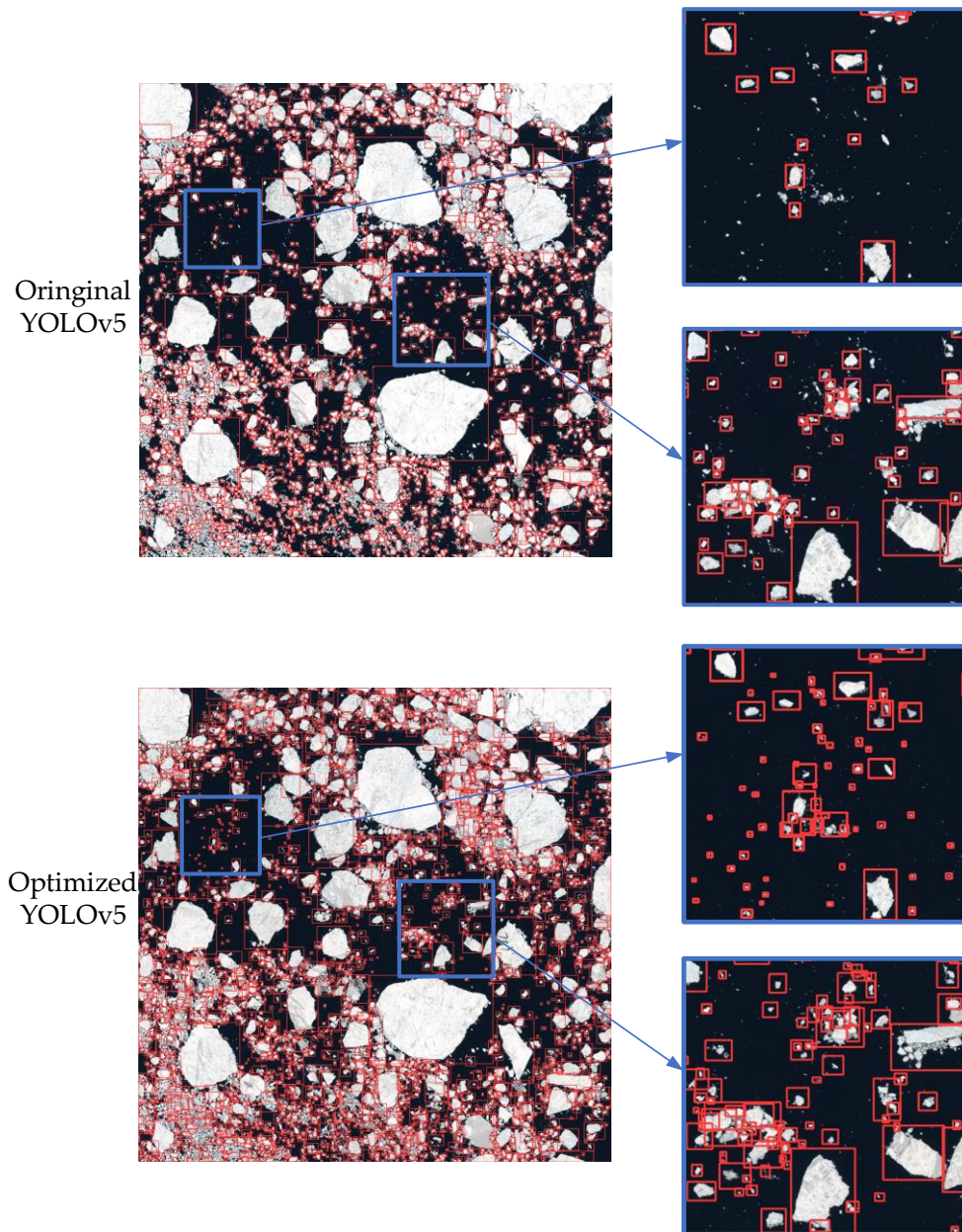


Figure 11. Comparison of the detection results.

3. Polar Ship Path Planning

3.1. Introduction of YOLOv5 Model

The output of target recognition is used as the input for map construction, and the sea ice distribution results are extracted and raster maps are constructed. After optimizing the YOLOv5 algorithm, the sea ice target in the remote sensing image is recognized, and based on the recognition result, the prediction frame of the sea ice is extracted, which contains the distribution information of the sea ice, including the positional coordinate information of the sea ice, and at the same time, the prediction frame also characterizes the size of the sea ice. After obtaining the sea ice distribution

results, the sea ice information is used as an input for the construction of the path planning map, and the process is shown in Figure 12.

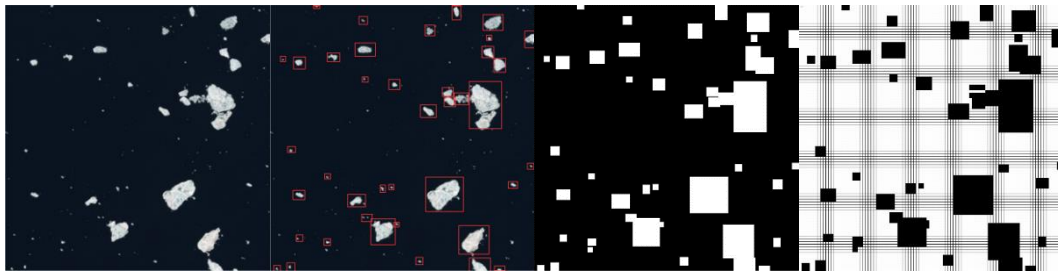


Figure 12. Raster map construction based on recognition results.

The output of optimizing YOLOv5 recognition of remote sensing sea ice is used as the input for the construction of path planning maps, and raster maps corresponding to the actual environment are built within the ice area in which it navigates. The prediction box for target identification is the outer rectangle of the sea ice, which is equivalent to puffing up the sea ice obstacles and completing the unfilled grids to ensure that sufficient safety distance is left between the ship and the sea ice to ensure that the ship carries out collision-free movement and reduce the potential collision risk caused by the edge of the obstacles [26]. The black grid in the figure indicates the location and size of the sea ice, and the white grid indicates the open water area through which the ship can pass. The sea ice obstacle puffing treatment is shown in Figure 13.

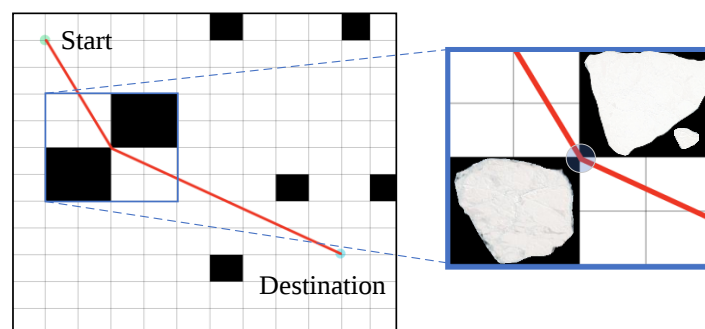


Figure 13. Puffing of ice obstacles.

3.2. Path Planning Algorithms and Objective Functions

Theta* is an improved algorithm based on the A* algorithm, both algorithms are path planning algorithms under the raster map. Although the A* algorithm can plan the shortest path, this path is only the shortest relative to the raster map, and does not fully match the shortest path in the actual physical environment [27]. The A* algorithm is constrained by the raster map, the resulting path has only 4 angles of movement, and will follow the edge of the grid when planning the path, so the deviation of the path from the actual shortest path is larger. A* algorithm is constrained by the raster map and produces path with only 4 angular movement directions, which are planned along the edges of the grid, and the deviation of such path from the actual shortest path is larger. The deviation between such a path and the actual shortest path is larger. In ship path planning, to meet the requirements of high efficiency and maneuverability, it is necessary to ensure that the ship makes smooth turns as much as possible, and the Theta* algorithm is not constrained by grid boundaries, which can better solve this problem [28].

The main difference between the Theta* algorithm and the A* algorithm is that the Line Of Sight (LOS) algorithm detection is performed before determining the next waypoint when expanding the node so that the direction of the path is not constrained by the raster when generating waypoints, and the LOS detection algorithm is shown in Figure 14.

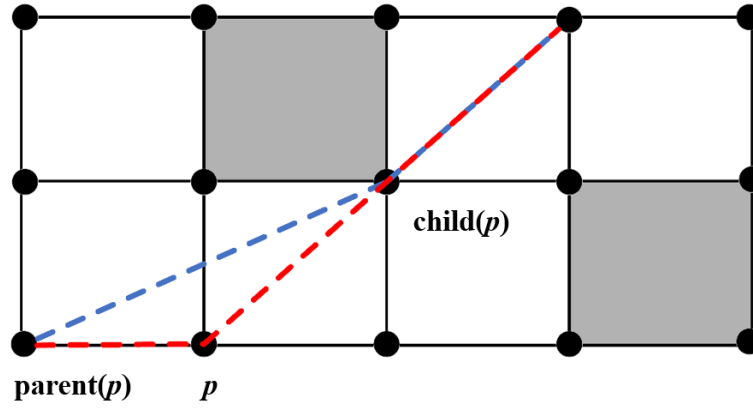


Figure 14. Schematic diagram of LOS algorithm.

The red path in the figure is the path planned by the A* algorithm, and the blue path is the path planned by the Theta* algorithm. It can be seen that the path generated by the Theta* algorithm is smoother, with fewer turns, which is more suitable for the navigation scenarios of the ship in the polar region. The Theta* algorithm ensures the optimality, while the speed of the path searching and the environment adaptability are both improved on the basis of the A* algorithm.

Polar ship path planning should take into account the path distance, operation complexity and sea ice avoidance while ensuring navigation safety. Therefore, considering the ship navigation conditions and ship performance, the objective function is established from three aspects: path distance, the number of ship turns and sea ice risk index. The path length is defined as shown in equation (6) and the sea ice risk value is defined as shown in equations (7)-(8).

$$D(L) = \sum_{n=1}^{N-1} \text{dist}(f_n, f_{n+1}) \quad (6)$$

where: $D(L)$ is the path length, and f_n is the sea ice avoidance waypoint at which the n navigation path is located, $\text{dist}(f_n, f_{n+1})$ is the distance from point f_n to point f_{n+1} .

$$k(i) = \sum_{i=1}^n f_n k \quad (7)$$

$$\sum_{j=1}^n \text{risk}(f_n, f_{n+1}) = \frac{k(i)}{D(L)} \times 100\% \quad (8)$$

where: $k(i)$ is the number of ice floes avoided, k is a variable from 0 to 1, when f_n belongs to the sea ice avoidance waypoint, $k=1$, otherwise $k=0$; $\sum_{j=1}^n \text{risk}(f_n, f_{n+1})$ is the sea ice risk value, which represents the percentage of the sea ice avoidance path length to the total path length.

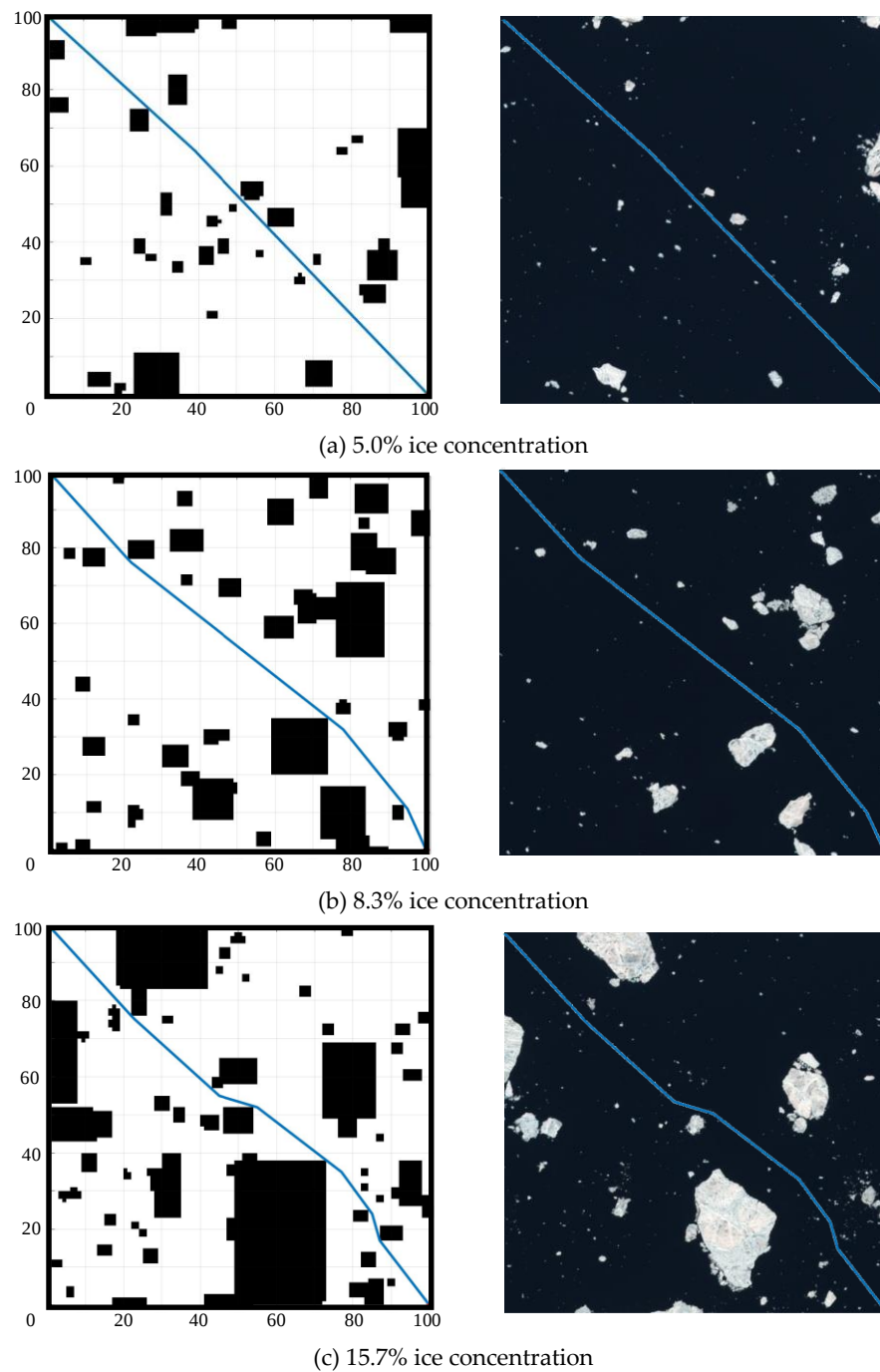
3.3. Path Planning under Different Sea Ice Densities

Remote sensing sea ice images with a fixed spatial resolution of 10km×10km are selected and input into the optimized YOLOv5 algorithm for recognition, and the recognition results are extracted. The corresponding navigation scenarios are divided into 100×100 environmental grids, with a total of 10,000 grids, and the spatial resolution corresponding to one grid is 0.1km. In order to more conveniently display the planned path, 1 large grid in the figure contains 10 small grids.

K-means clustering is used to obtain sea conditions with different sea ice densities, and different navigation scenarios of ships in the polar regions are simulated. The process of K-means begins by inputting a remote sensing image. Firstly, the number of clusters, denoted as k , is determined. Then, k cluster centers are initialized. The next step involves calculating the difference between the RGB values of each image pixel and the cluster centers. This difference is compared against a predefined threshold. If the difference exceeds the threshold, the process returns to the previous step to

recalculate. If the difference is within the threshold, the process continues by outputting the proportion of each color in the k clusters. Subsequently, the sea ice concentration is calculated. Finally, a new remote sensing image can be inputted to repeat the entire process.

After calculating the k value through sea ice dataset, six different navigation scenarios with varying sea ice concentrations were obtained. The sea ice concentrations of the scenarios are calculated in 5.0%, 8.3%, 15.7%, 18.4%, 25.1%, and 40.9%, respectively [29]. Based on the recognition results of remote sensing sea ice images, the corresponding raster maps are constructed, and the Theta* algorithm is used for path planning. Setting the first grid of the map as the starting point and the last grid as the end point, the path planning results are shown in Figure 15.



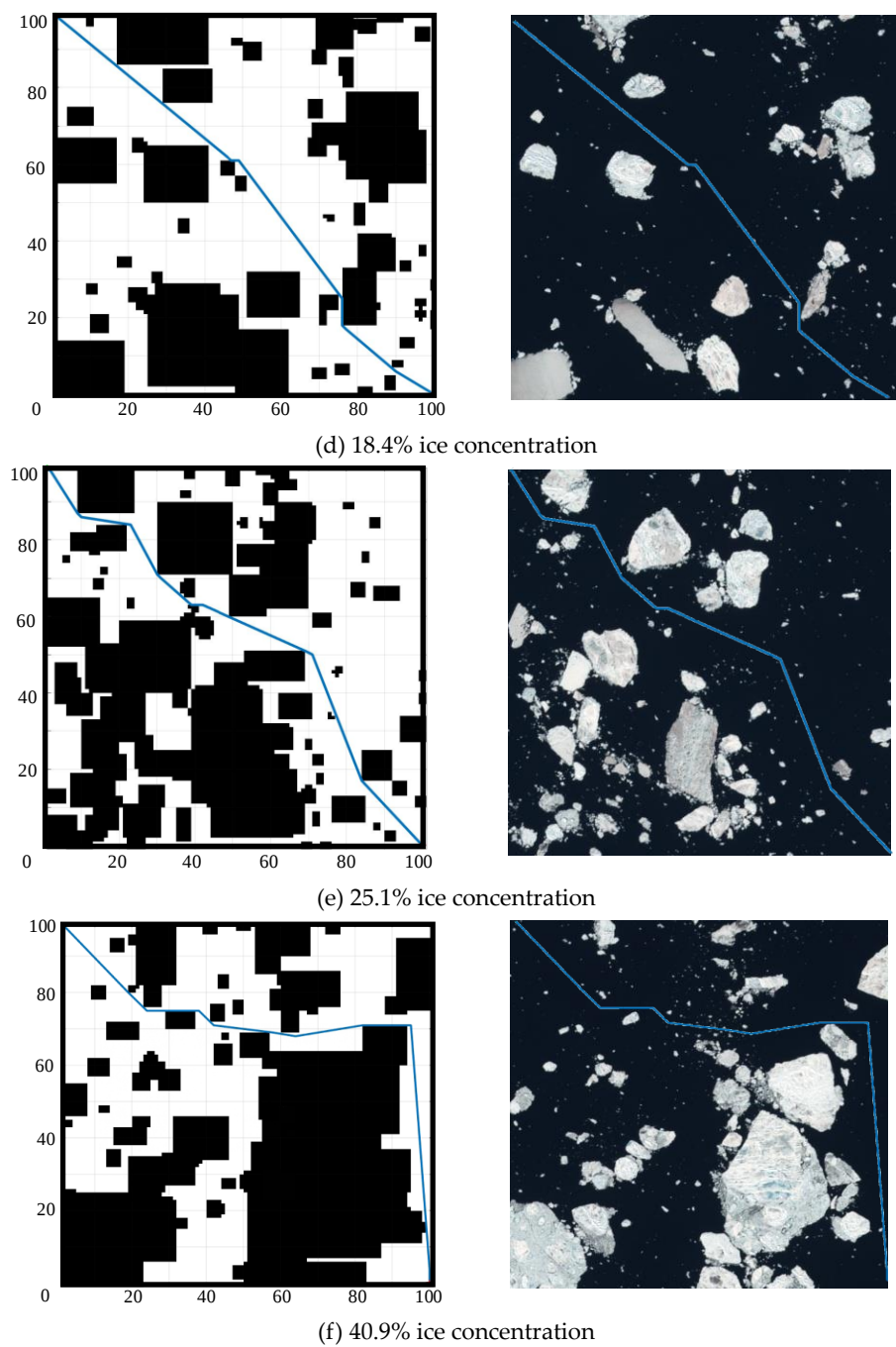


Figure 15. Path planning results for different sea ice concentration scenarios.

The results of the planned path for different sea ice densities sailing scenarios are shown in Table 6, which contains the path length, the number of ship turns, the number of sea ice avoidance, and the sea ice risk value. The parameter comparison is visualized in Figure 16.

Table 6. Path planning results for different sea ice concentrations.

Ice concentration	Path length	Number of ship turn	Number of ice avoidance	Sea ice risk value
5.0%	14.23 km	1	8	5.6%
8.3%	14.44 km	3	6	4.2%
15.7%	14.59 km	5	12	8.3%
18.4%	14.67 km	5	16	10.9%

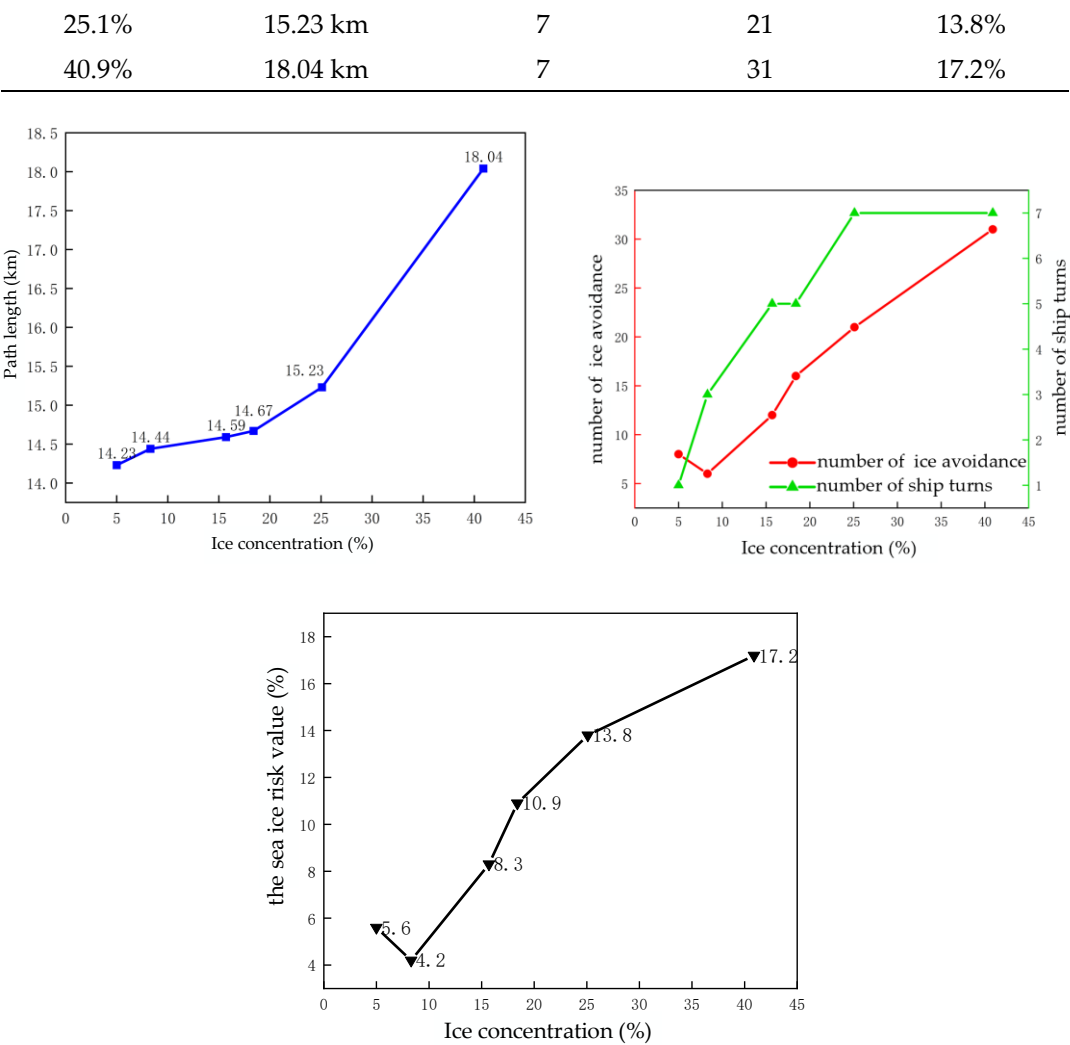


Figure 16. Results of each parameter of path planning.

From the results, it can be seen that the sea ice concentration parameter significantly affects the ship's path selection and the safety of navigation. Specifically, when the ship navigates under different sea conditions, with the increase of sea ice concentration, its path shows a more complex direction, the number of turns and the number of sea ice obstacle avoidance show an increasing trend, and the value of the sea ice risk increases. The numerical analysis results show that when the sea ice concentration increases from 5.0% to 40.9%, the length of the ship's path increases by 3.81km, the number of turning adjustments to avoid sea ice obstacles increases by 23 times, and the overall sea ice risk value increases by 11.6%. Changes in sea ice concentration can have a significant impact on ship navigation safety and path optimization.

Target recognition provides information about the distribution and morphology of sea ice, and path planning avoids dense sea ice areas according to this information. This paper utilizes the algorithms of target recognition and path planning, combines perception and decision-making, and finds a shorter, smoother and safer path in the complex sea conditions in the polar region, which provides technical support for the navigation of polar ships [30].

4. Conclusions

In this paper, remote sensing sea ice identification is performed using computer vision methods. Based on the identification results, polar ship route planning is conducted to assist in decision-making for navigation. The following conclusions can be drawn.

- (1) Targeted optimization of YOLOv5 is carried out according to the characteristics of remote sensing images. This optimization includes three improvements: adding the SE attention

mechanism, improving the spatial pyramid pooling structure, and replacing the activation function with FReLU, which is more suitable for the target identification task. The model structure and network parameters before and after the optimization are then compared.

- (2) Experimental results and analysis of the optimization algorithm are presented. Ablation experiments are conducted to compare the effects of different improvement methods. When the optimization methods are applied simultaneously, the mAP improves by 3.5%. Comparison experiments are conducted to verify the effectiveness of the optimization algorithm by comparing Faster-RCNN, YOLOv3, YOLOv4, YOLOv5, and YOLOv8. The optimized YOLOv5 achieves a mAP of 75.4%, making it the best-performing model in the comparison experiments. In the same region of the remote sensing image, the number of detected sea ice instances increased by 39 and 33, most of which are small targets that are difficult to detect.
- (3) Path planning for polar ships is based on the identification results. The sea ice identification output of the optimized YOLOv5 was used as input for constructing a path planning map, and a raster map corresponding to the actual polar scene was created. Using path length, the number of ship turns, and the sea ice risk value as the objective functions, simulations were carried out under different sea conditions, with ice concentrations ranging from 5.0% to 40.9%. Finally, a shorter, smoother, and safer path was found.

Although the study has successfully detected sea ice in remote sensing images and performed path planning based on it, there are still some limitations. This study focused on rectangular detection boxes. If more detailed sea ice information is needed, instance segmentation could be applied to process the ice images in future work. Additionally, this paper focuses on the combination of target recognition and path planning, with less consideration given to environmental factors in the path planning process.

For future research, a refined classification of sea ice can be established, such as distinguishing between first-year ice and multi-year ice, to provide more detailed information about the polar environment. Additionally, more complex environmental factors can be considered in path planning, such as the impact of wind, waves, and currents on the ship's course, as well as accounting for the ship's own maneuverability, so that the planned path is more aligned with the actual conditions [30,31].

Acknowledgments: This study is funded by the National Key Research and Development Program (2022YFE0107000), General Projects of National Natural Science Foundation of China (52171259), High-tech ship research project of Ministry of Industry and Information Technology ([2021]342), CSSC-SJTU joint prospect funding (ZCJDQZ202307A01), Science and Technology Commission of Shanghai Municipality Project (22DZ1204400), Foundation of State Key Laboratory of Ocean Engineering in Shanghai Jiao Tong University (GKZD010086-2), Science and Technology Commission of Shanghai Municipality Project (23YF1419900), Young Scientists Fund of National Natural Science Foundation of China (52301331).

References

1. Chuah, L.F.; Mokhtar, K.; Ruslan, S.M.M.; Abu Bakar, A.; Abdullah, M.A.; Osman, N.H.; Bokhari, A.; Mubashir, M.; Show, P.L. Implementation of the energy efficiency existing ship index and carbon intensity indicator on domestic ship for marine environmental protection. *Environ. Res.* 2023, 222, 115348. <https://doi.org/10.1016/j.envres.2023.115348>
2. Zuo, Q.; Qian, L.; Xu, X.; Yan, J.; Cheng, L.; Zhang, Z. Navigation strategy and economic research of the northeast passage in the Arctic. *Chin. J. Polar Res.* 2015, 27, 203.
3. Lu, P.; Leppäranta, M.; Cheng, B.; Li, Z.; Istomina, L.; Heygster, G. The color of melt ponds on Arctic sea ice. *Cryosphere* 2018, 12, 1331–1345. <https://doi.org/10.5194/tc-12-1331-2018>
4. Weissling, B.; Ackley, S.; Wagner, P.; Xie, H. EISCAM—Digital image acquisition and processing for sea ice parameters from ships. *Cold Reg. Sci. Technol.* 2009, 57, 49–60. <https://doi.org/10.1016/j.coldregions.2009.01.001>
5. Worby, A.; Comiso, J. Studies of the Antarctic sea ice edge and ice extent from satellite and ship observations. *Remote. Sens. Environ.* 2004, 92, 98–111. <https://doi.org/10.1016/j.rse.2004.05.007>
6. Zhou, L.; Cai, J.; Ding, S. The Identification of Ice Floes and Calculation of Sea Ice Concentration Based on a Deep Learning Method. *Remote. Sens.* 2023, 15, 2663. <https://doi.org/10.3390/rs15102663>

7. Ressel, R.; Frost, A.; Lehner, S. A Neural Network-Based Classification for Sea Ice Types on X-Band SAR Images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 2015, 8, 3672–3680. <https://doi.org/10.1109/JSTARS.2015.2436993>
8. Anderson, S. Remote Sensing of the Polar Ice Zones with HF Radar. *Remote Sens.* 2021, 13, 4398. <https://doi.org/10.3390/rs13214398>
9. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE Trans. Pattern Anal. Mach. Intell.* 2017, 39, 1137–1149. <https://doi.org/10.1109/TPAMI.2016.2577031>
10. Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S.; Fu, C.Y.; Berg, A.C. SSD: Single shot multibox detector. In *Proceedings of the European Conference on Computer Vision, Amsterdam, The Netherlands, 11–14 October 2016*. https://doi.org/10.1007/978-3-319-46448-0_2
11. Dong, W.; Zhou, L.; Ding, S.; Ma, Q.; Li, F. Fast and Intelligent Ice Channel Recognition Based on Row Selection. *J. Mar. Sci. Eng.* 2023, 11, 1652. <https://doi.org/10.3390/jmse11091652>
12. Lu, P.; Leppranta, M.; Cheng, B.; Li, Z.; Heygster, Georg. The color of melt ponds on Arctic sea ice. *The Cryosphere*, 2018, 12(4): 1331–1345. <https://doi.org/10.5194/tc-12-1331-2018>
13. Cai, J.; Ding, S.; Zhang, Q.; Liu, R.; Zeng, D.; Zhou, L. Broken ice circumferential crack estimation via image techniques. *Ocean Engineering*, 2022, 259. <https://doi.org/10.1016/j.oceaneng.2022.111735>
14. Li, F.; Chen, J.; Zhou, L.; Kujala, P. Investigation of ice wedge bearing capacity based on an anisotropic beam analogy. *Ocean Engineering*, Volume 302, 2024, 117611.
15. M. Choi, H. Chung, H. Yamaguchi, et al. Arctic Sea route path planning based on an uncertain ice prediction model[J]. *Cold Regions Science and Technology*, 2015:10961–69. <https://doi.org/10.1016/j.coldregions.2014.10.001>
16. V. Kotovirta, R. Jalonen, L. Axell, et al. A system for route optimization in ice-covered waters[J]. *Cold Regions Science and Technology*, 2009, 55(1):52–62. <https://doi.org/10.1016/j.coldregions.2008.07.003>
17. Hass, F.S.; Arsanjani, J.J. Deep Learning for Detecting and Classifying Ocean Objects: Application of YoloV3 for Iceberg–Ship Discrimination. *ISPRS Int. J. Geo-Information* 2020, 9, 758. <https://doi.org/10.3390/ijgi9120758>
18. Wu, A., Che, T., Li, X., & Zhu, X. (2022). Routeview: an intelligent route planning system for ships sailing through Arctic ice zones based on big Earth data. *International Journal of Digital Earth*, 15(1), 1588–1613. <https://doi.org/10.1080/17538947.2022.2126016>
19. Yang, Z.; Yin, Y.; Jing, Q.; Shao, Z. A High-Precision Detection Model of Small Objects in Maritime UAV Perspective Based on Improved YOLOv5. *J. Mar. Sci. Eng.* 2023, 11, 1680. <https://doi.org/10.3390/jmse11091680>
20. Qiao, W.; Guo, H.; Huang, E.; Su, X.; Li, W.; Chen, H. Real-Time Detection of Slug Flow in Subsea Pipelines by Embedding a Yolo Object Detection Algorithm into Jetson Nano. *J. Mar. Sci. Eng.* 2023, 11, 1658. <https://doi.org/10.3390/jmse11091658>
21. Ophoff, T.; Puttemans, S.; Kalogirou, V.; Robin, J.-P.; Goedemé, T. Vehicle and Vessel Detection on Satellite Imagery: A Comparative Study on Single-Shot Detectors. *Remote Sens.* 2020, 12, 1217. <https://doi.org/10.3390/rs12071217>
22. Zhao, M.; Zhou, H.; Li, X. YOLOv7-SN: Underwater Target Detection Algorithm Based on Improved YOLOv7. *Symmetry* 2024, 16, 514. <https://doi.org/10.3390/sym16050514>
23. Bao, Z.; Guo, Y.; Wang, J.; Zhu, L.; Huang, J.; Yan, S. Underwater Target Detection Based on Parallel High-Resolution Networks. *Sensors* 2023, 23, 7337. <https://doi.org/10.3390/s23177337>
24. Cheng G, Han J, Lu X. Remote sensing image scene classification: benchmark and state of the art[J]. *Proceedings of the IEEE*, 2017, 105(10):1865–1883. <https://doi.org/10.48550/arXiv.1703.00121>
25. Redmon, J.; Farhadi, A. YOLO9000: Better, Faster, Stronger. In *Proceedings of the IEEE Conference on Computer Vision & Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017*; pp. 6517–6525.
26. Cao, S.; Fan, P.; Yan, T.; Xie, C.; Deng, J.; Xu, F.; Shu, Y. Inland Waterway Ship Path Planning Based on Improved RRT Algorithm. *J. Mar. Sci. Eng.* 2022, 10, 1460. <https://doi.org/10.3390/jmse10101460>
27. Hu, S.; Tian, S.; Zhao, J.; Shen, R. Path Planning of an Unmanned Surface Vessel Based on the Improved A-Star and Dynamic Window Method. *J. Mar. Sci. Eng.* 2023, 11, 1060. <https://doi.org/10.3390/jmse11051060>
28. S. Han, J. Sun, S. Ding and L. Zhou, "A Potential Field-Based Model Predictive Target Following Controller for Underactuated Unmanned Surface Vehicles," in *IEEE Transactions on Vehicular Technology*, doi: 10.1109/TVT.2024.3410183.
29. Frackiewicz, M.; Mandrella, A.; Palus, H. Fast Color Quantization by K-Means Clustering Combined with Image Sampling. *Symmetry* 2019, 11, 963. <https://doi.org/10.3390/sym11080963>
30. Sun, Q.; Chen, J.; Zhou, L.; Ding, S.; Han, S. A study on ice resistance prediction based on deep learning data generation method. *Ocean Engineering*, Volume 301, 2024, 117467.
31. Xie, C.; Zhou, L.; Ding S.; Liu, R.; Zheng, S. Experimental and numerical investigation on self-propulsion performance of polar merchant ship in brash ice channel. *Ocean Engineering*, 2023, 269. <https://doi.org/10.1016/j.oceaneng.2022.113424>

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.