ICJK
Chinese Journal of Information Fusion
http://dx.doi.org/10.62762/CJIF.2025.695812

RESEARCH ARTICLE

An Improved YOLOv8-Based Detection Model for Multi-Scale Sea Ice in Satellite Imagery

Yang Liu[1], **Qiang Guo**[1,*], **Chengguo Dong**[2] **and Yiping Luo**[3]

[1] School of Computer and Control Engineering, Yantai University, Yantai 264005, China
[2] School of Architectural Engineering, Weifang University of Science and Technology, Weifang 261000, China
[3] Deep Space Exploration Laboratory, Hefei 230000, China

Abstract

Sea ice detection is of vital importance for maritime navigation. Satellite imagery is a crucial medium for conveying information about sea ice. Currently, most sea ice detection models mainly rely on texture information to identify sea ice in satellite imagery, while ignoring sea ice size information. This research presents an improved YOLOv8-Based detection algorithm for multi-scale sea ice. First, we propose a fusion module based on the attention mechanism and use it to replace the Concat module in the YOLOv8 network structure. Second, we conduct an applicability analysis of the bounding box regression loss function in YOLOv8 and ultimately select Shape-IoU, which is more suitable for sea ice, as the loss function for bounding box regression. Third, we analyze the distribution characteristics of sea ice with different sizes in the NWPU-RESISC45 dataset. Based on these distribution characteristics, the bounding box information predicted by YOLOv8 are converted into evidence vectors for uncertainty quantification. Subsequently, information fusion is achieved by fusing these vectors with the probability of sea ice categories. Compared to YOLOv8 and other detection algorithm, our improved YOLOv8 achieves better detection accuracy on both the NWPU-RESISC45 and the Landsat-8-derived Sea Ice datasets.

Keywords: satellite imagery, YOLO, attention mechanism, loss function, information fusion, evidential reasoning.

1 Introduction

In recent years, due to the continuous global warming [1, 2], the sea ice in high-latitude regions has been persistently melting [3, 4]. The resulting high-latitude waterways can shorten the sailing distances between major trading powers and are urgently in need of development as future maritime routes [5–7]. Specifically, sea ice detection has always been the focus of research in high-latitude seas, which is devoted to accurately locating the positions of sea ice and identifying the scales of sea ice [8, 9].

A multitude of technologies are emerging in the domain of real-time object detection. They are

Academic Editor:
Fengbao Yang

Submitted: 05 March 2025
Accepted: 23 March 2025
Published: 29 March 2025

Vol. 2, **No.** 1, 2025.
10.62762/CJIF.2025.695812

***Corresponding author:**
Qiang Guo
guoqiang@ytu.edu.cn

Citation
Liu, Y., Guo, Q., Dong, C., & Luo, Y. (2025). An Improved YOLOv8-Based Detection Model for Multi-Scale Sea Ice in Satellite Imagery. *Chinese Journal of Information Fusion*, 2(1), 79–99.

© 2025 by the Authors. Published by Institute of Central Computation and Knowledge. This is an open access article under the CC BY license (https://creativecommons.org/licenses/by/4.0/).

extensively adopted in diverse industries, including the identification of suspicious behavior [10], the detection of anomalies in medical images [11], and fish detection [12], and other applications. In recent years, researchers have been concentrating on designing CNN-based object detectors [13–16], with further developments in [17, 18]. Among them, YOLOs achieve accurate classification and positioning of objects with low latency, and they are increasingly gaining popularity through successive versions including [19–23] and later improvements in [24–27].

Furthermore, for an extended period, considerable efforts have been directed towards obtaining high-quality sea ice satellite remote sensing information and detecting sea ice from a diverse range of satellite remote sensing data [28–31]. Hu et al. [32] detected sea ice using GNSS bidirectional radar reflections, where the local linear embedding (LLE) algorithm was employed for sea ice feature extraction. Liu et al. [33] proposed a Bayesian method with consideration of geometric characteristics of China France Oceanography Satellite scatterometer(CSCAT) for sea ice detection. The method operationally produced daily polar sea ice mask throughout its mission duration from 2019 to 2022. Jafari et al. [34] developed an automated method for iceberg detection and classification in complex sea conditions. Using the RADARSAT Constellation Mission (RCM), they collected seasonal sea ice data from the east coast of Canada.

To obtain more abundant spectral information, researchers have explored diverse types of optical remote sensing data [35–38]. Researchers have focused on studying sea ice with visible remote sensing data, as the human eye can intuitively perceive the difference between sea ice and seawater in the visible wavelength band. Advancements in deep convolutional neural networks have achieved automated sea ice detection using visible remote sensing data. Ding et al. [39] proposed a detection model based on YOLOv5. They added Squeeze-and-Excitation Networks (SE) [40] to backbone of YOLOv5. The SE module computes channel-wise attention weights through global average pooling and multilayer perceptron, which are then applied to recalibrate feature map channels by element-wise multiplication. However, over-dependence on channel attention mechanisms (e.g., SE) inevitably discards spatially fine-grained features in imagery, particularly ice-water interface textures and areal extent variations that are critical for sea ice detection.

In this paper, we aim to address these questions precisely and further broaden the application scope of YOLOs. Refining the details of YOLOv8, we aim to enhance its capability in identifying sea ice across a variety of sizes.

Our contributions are as follows:

1. First, we propose a fusion module based on the attention mechanism and use it to replace the Concat module in the YOLOv8 network structure. This module can effectively help YOLOv8 extract the characteristic information of sea ice.
2. Second, we conduct an applicability analysis of the bounding box regression loss function in YOLOv8 and ultimately select Shape-IoU, which is more suitable for sea ice, as the loss function for bounding box regression. YOLOv8 utilizing Shape-IoU [41] not only demonstrates superior detection accuracy across all three categories of sea ice, but it also significantly reduces convergence time.
3. Third, we analyze the distribution characteristics of sea ice with different sizes in the NWPU-RESISC45 dataset. Based on these distribution characteristics, the bounding box information predicted by YOLOv8 are converted into evidence vectors for uncertainty quantification. Subsequently, evidence fusion [42] is achieved by fusing these vectors with the probability of sea ice categories.
4. Based on Landsat-8 satellite data, we have created a sea ice dataset and made it publicly available on this website: https://github.com/LiuYang0911/A-Proprietary-Visible-Light-based-Sea-Ice-Dataset.
5. By comparing with current mainstream object detection algorithms, our improved YOLOv8 achieves better detection accuracy and faster convergence speed.

We are optimistic that the outcomes of our efforts can act as a catalyst for the progress of fellow researchers in this domain.

2 Related Work

2.1 Attention Mechanism

Initially, attention mechanisms were utilized in machine translation tasks. This mechanism enables the model to focus on different parts of the input sentence

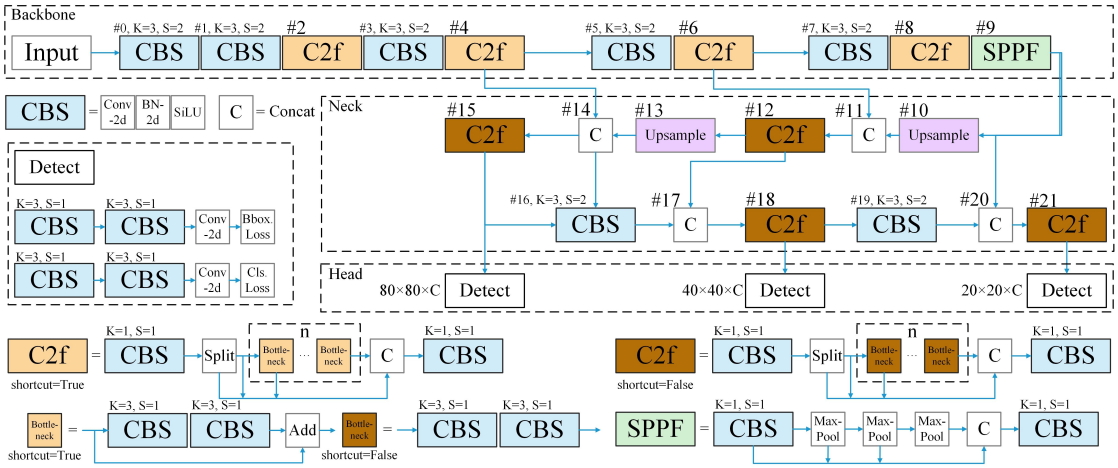


Figure 1. Schematic representation of the network architecture for YOLOv8.

when translating a word, which significantly enhances the translation quality [43]. Over the past years, considerable efforts have been devoted to developing attention mechanism modules that are more applicable to the domain of computer vision [40, 44, 45].

1. Channel attention mechanism: This type of attention mechanism, which concentrates on the channel dimension of the feature map, aims to enhance significant channel information while suppressing less important data. It accomplishes this by learning weights for each channel, akin to the methodology employed in SENet [40]. In SENet [40], the input feature map is first compressed along the spatial dimension before calculating weights for each channel. Finally, these weights are applied to multiply with the input feature map to produce the final output.

2. Spatial attention mechanism: In contrast to the channel attention mechanism, the spatial attention mechanism emphasizes the locations of valid information within the feature map, as exemplified by STN [44]. STN [44] is capable of extracting characteristics from significant regions across various deformation data to produce final prediction results.

3. Hybrid Attention Mechanism: Compared to the aforementioned two attention mechanisms, this particular attention mechanism comprehensively leverages both channel information and spatial information from feature maps, as exemplified by CBAM [45]. It sequentially employs the channel attention module followed by the spatial attention module to generate attention weights, ultimately producing the final feature map.

2.2 YOLOs

2.2.1 Modules and Network Architecture

In 2016, Joseph Redmon introduced YOLOv1 [18], a real-time object detector built upon the deep learning framework Darknet. When compared to other object detectors [13–17], YOLOs [18–20] demonstrate superior detection performance while maintaining high detection speeds.

Over the past few years, significant efforts have been dedicated to exploring more efficient modules and network architectures for the YOLO series. YOLOv4 [21] and YOLOv5 [22] investigated the impact of various activation functions on detection accuracy and speed. It is essential to recognize that YOLOv5 [22] has been widely adopted across numerous sectors as a highly effective object detector. Building upon RepVGG, YOLOv6 [23] introduced RepBlock to replace the CSPDarknet53 [46] architecture used in YOLOv5, which allows the model to better integrate multi-scale features. Furthermore, based on YOLOv5 [22], YOLOv7 [24] proposed E-ELAN [47], which enhances the network's learning capability while preserving the original gradient path.

Based on the C3 module of YOLOv5 [22], YOLOv8 [25] has developed the C2f module, as illustrated in Figure 1. This module dynamically adjusts the number of channels according to the model's size, enabling it to flexibly adapt to various scenarios.

YOLOv9 [26] introduced G-ELEN, a network architecture that integrates the features of CSPNet [46] and ELEN [47], aiming to enhance detection accuracy while preserving detection speed. Building on the foundation established by YOLOv8 [25], YOLOv10 [27] presented several improvements,

including the use of classification heads with reduced parameters and the incorporation of a partial self-attention module, etc., all designed to further transcend the accuracy-speed trade-offs inherent in YOLO models.

2.2.2 Loss Function Utilized in Bounding Box Regression

The object detector based on convolutional neural networks employs a loss function to update the network weights [48]. Historically, iterations of the YOLO object detectors have been engaged in an unrelenting pursuit of optimizing the loss function for bounding box regression, aiming to achieve superior performance [48–50]. Simultaneously, a variety of loss functions for boundary box regression are continuously being developed and refined [41, 51–54], thereby enabling the improved YOLOv8-based object detector to be applied across an increasingly diverse range of scenarios.

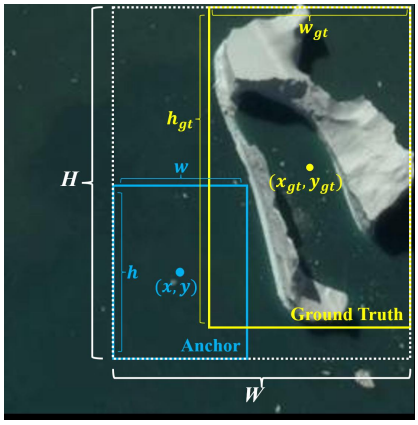


Figure 2. Schematic representation of the loss function for bounding box regression.

The YOLOv8 model employs C-IoU [50] as the loss function for bounding box regression, as shown in Figure 2. The following presents the mathematical expression for C-IoU [50]:

$$CIoU = IoU - \frac{(x - x_{gt})^2 + (y - y_{gt})^2}{W^2 + H^2} - \alpha\upsilon \quad (1)$$

$$IoU = \frac{B_{Anchor} \cap B_{Ground\ Truth}}{B_{Anchor} \cup B_{Ground\ Truth}} \quad (2)$$

$$\alpha = \frac{\upsilon}{1 - IoU + \upsilon} \quad (3)$$

$$\upsilon = \frac{4}{\pi^2} * (\tan^{-1}\frac{w}{h} - \tan^{-1}\frac{w_{gt}}{h_{gt}})^2 \quad (4)$$

From the formulas, it is evident that C-IoU [50] takes into account both the position and shape of the bounding box in a comprehensive manner. This allows the model to learn the characteristics of the ground truth box more thoroughly.

3 Methodology

3.1 An Attention-Based Fusion Module

We categorize these sea ice instances into three distinct groups, as elaborated in section 4.1. Although satellite imagery offers relatively high resolution, actual ice conditions can be highly complex. As shown in Figure 3, several factors make it challenging for object detection models to accurately identify sea ice of varying scales. These include the wide range of ice floe sizes, irregular shapes, and reduced contrast between ice and seawater caused by melting and accumulation of sea ice.

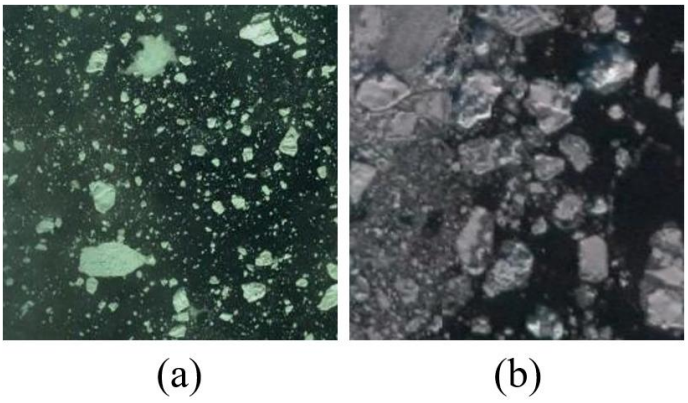


Figure 3. The two key issues: (a) Numerous small-scale sea ice; (b) Ambiguous demarcation between sea ice and seawater.

Besides, as the network deepens, the detection model progressively enhances semantic information in feature maps while inevitably sacrificing spatial details, particularly size characteristics crucial for sea ice analysis. When handling this task, YOLOv8 uses the Concat module to combine deep and shallow feature maps. However, we observe that relying on the feature maps after direct stitching is not sufficient for accurate size classification. To address this limitation, we propose an attention-based fusion module that can effectively enhance the spatial detail information in the feature map, so as to be able to accurately distinguish between sea ice size categories, as shown in Figure 4.

First, we calculate the channel attention weight M_C of the feature map F_1 and multiply it with the feature map F_1 to obtain F_1'. Secondly, we calculate the Spatial attention weights M_S of the feature map F_2 and multiply it with the feature map F_2 to obtain F_2'. Thirdly, we concatenate the feature map F_1' and F_2'

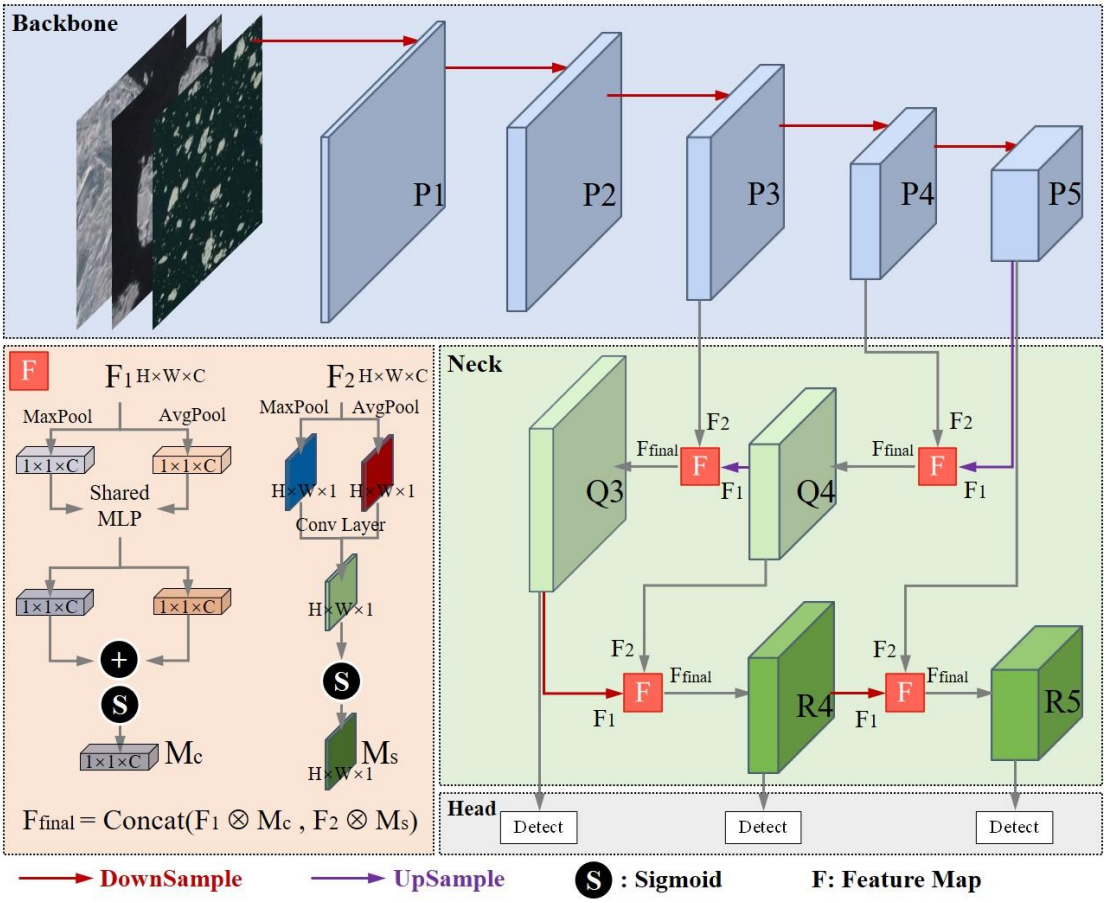


Figure 4. Schematic representation of the attention-based fusion module.

to obtain the final feature map F_{final}. The detailed process is as follows:

$$M_C = \sigma\left(MLP\left(\text{AvgPool}\left(F_1\right)\right) + MLP\left(\text{MaxPool}\left(F_1\right)\right)\right) \tag{5}$$

$$F_1' = M_C \otimes F_1 \tag{6}$$

$$M_S = \sigma(Conv(AvgPool(F_2); MaxPool(F_2))) \tag{7}$$

$$F_2' = M_S \otimes F_2 \tag{8}$$

$$F_{final} = Concat(F_1'; F_2') \tag{9}$$

where the feature map F_1 is derived from deeper layers and contains more semantic information, such as the overall shape and categories of sea ice; the feature map F_2 is derived from shallower layers and includes more detailed information, such as edges, textures, colors, and other low-level features of sea ice. In this way, we optimize the fusion process of different feature maps in YOLOv8, enabling the model to balance attention between the semantic information and detailed information of different feature maps, thereby improving the detection accuracy of the model for sea ice of varying scales.

3.2 Selection of Boundary Regression Loss Function Based on Sea Ice Size Characteristics

The C-IoU loss [50] employed in YOLOv8 considers the geometric relationship between the ground truth box and the predicted box, utilizing both their relative positions and shapes to compute the loss. However, in contrast to general objects, sea ice presents a more diverse aspect ratio and possesses an irregular shape that lacks any fixed pattern. In this study, if C-IoU [50] is utilized as the loss function for bounding box regression, two critical questions arise.

1. As shown in Figure 5, targets ① and ② represent the same sea ice. When two bounding boxes exhibit the same absolute deviation from the ground truth box, the bounding box that regresses from the direction of the shorter side of the rectangle tends to demonstrate a lower Intersection over Union (IoU) value. Our research indicates that this variation in IoU is more pronounced when regression occurs from the direction of the shorter side during bounding box adjustment. Consequently, it is crucial for models to effectively balance the

regression impact of bounding boxes originating from different directions.

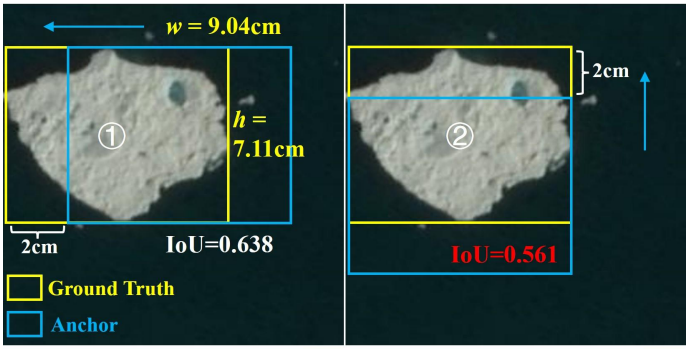


Figure 5. Schematic diagram of the Question 1.

2. As illustrated in Figure 6, the center points of the prediction boxes ③ and ④ have shifted closer to the position of the ground truth box. Furthermore, both prediction boxes maintain an equal distance from the ground truth box along both the long and short sides. However, it is noteworthy that target ④, which regressed from the short side, corresponds to a lower IoU value, indicating reduced overlap with the ground truth.

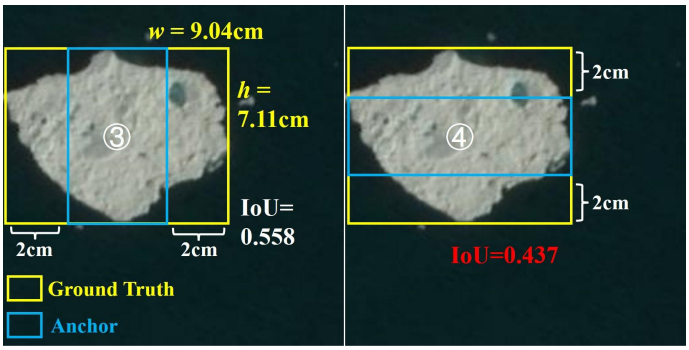


Figure 6. Schematic diagram of the Question 2.

During the bounding box regression process, the variation in IoU is particularly significant when the regression occurs along the short side of the ground truth box. Therefore, it is essential to ensure a balanced regression effect for bounding boxes across various directions throughout this process.

In the end, we select Shape-IoU [41] as our loss function for bounding box regression, as it effectively addresses the two types of issues mentioned above. The formula for Shape-IoU [41] is presented below:

$$\begin{cases} ww = \frac{2*(w_{gt})^{scale}}{(w_{gt})^{scale}+(h_{gt})^{scale}} \\ hh = \frac{2*(h_{gt})^{scale}}{(w_{gt})^{scale}+(h_{gt})^{scale}} \end{cases} \tag{10}$$

$$distance_{shape} = hh*\frac{(x-x_{gt})^2}{W^2+H^2}+ww*\frac{(y-y_{gt})^2}{W^2+H^2} \tag{11}$$

$$\begin{cases} \omega_w = hh * \frac{|w-w_{gt}|}{max(w,w_{gt})} \\ \omega_h = ww * \frac{|h-h_{gt}|}{max(h,h_{gt})} \end{cases} \tag{12}$$

$$\Omega_{shape} = \sum_{t=w,h} (1-e^{-\omega_t})^{\theta}, \theta = 4 \tag{13}$$

where $scale$ represents the scale factor, which can be adjusted based on the dimensions of the target. Taking the ground truth box in Figure 5 as an example(where $w_{gt} > h_{gt}$), when $scale = 0$ and $ww = hh = 1$, the bounding box regression lacks directional prioritization. by increasing the value of scale, the regression effectiveness can be enhanced. In this study, we set $scale = 1$, resulting in $ww > hh$, which indicates that higher regression weight is assigned to the vertical dimension (height adjustment).

Equation 14 calculates the loss value for bounding box regression.

$$L_{Shape-IoU} = 1 - IoU + distance_{shape} + 0.5 * \Omega_{shape} \tag{14}$$

It is noteworthy that Shape-IoU dynamically emphasizes the gradient update path of bounding box parameters (e.g., center offsets and aspect ratios) during model convergence. Unlike C-IoU, which indirectly guides optimization through geometric penalties (center distance and aspect ratio matching), Shape-IoU explicitly introduces a directional weighting coefficient, thereby clarifying the prioritization of regression targets with higher shape discrepancies. This property of Shape-IoU shortens the convergence time of the model.

3.3 An Evidence Fusion Module for the Correction of Sea Ice Categories

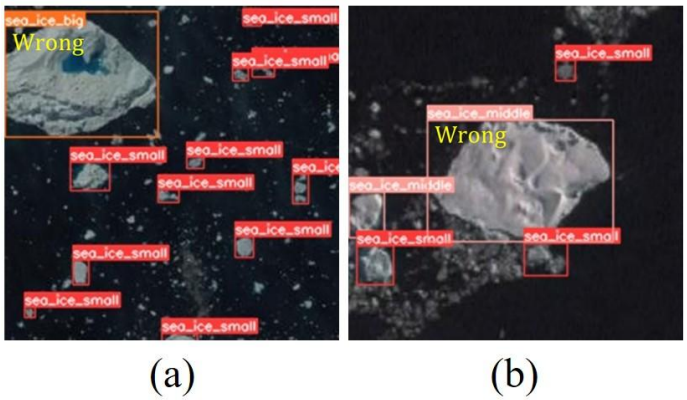


Figure 7. Typical cases of misclassification.

After extensive experimentation, we discovered that YOLOv8 is capable of accurately predicting the

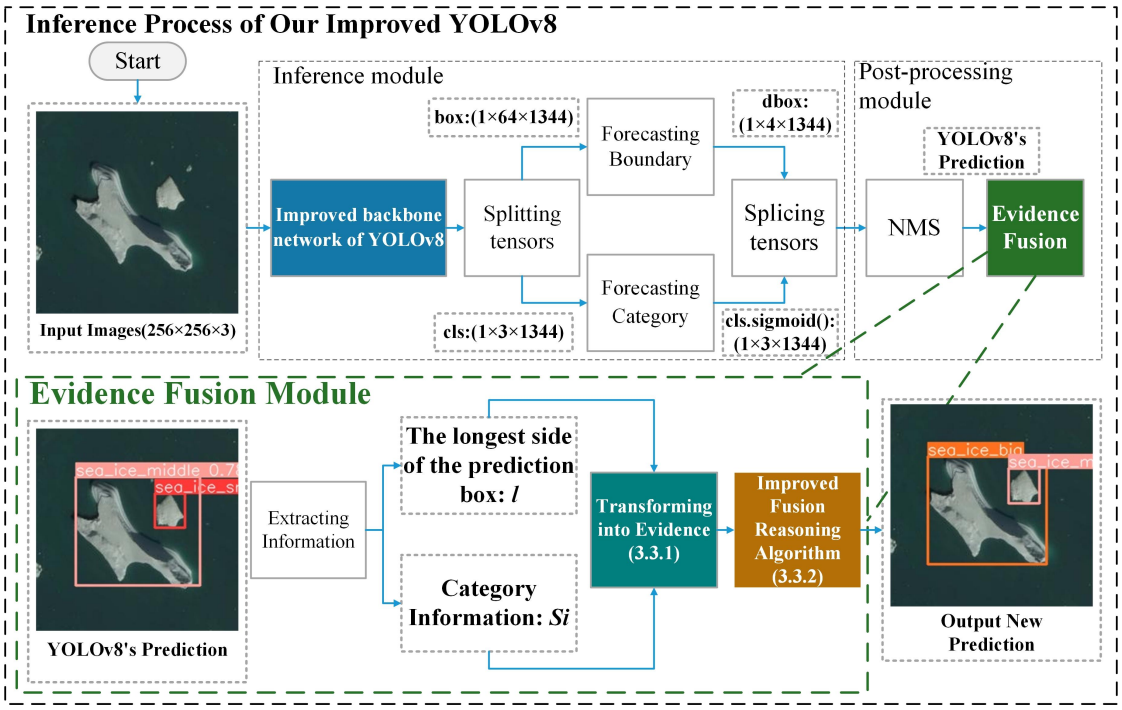


Figure 8. Schematic diagram of evidence fusion module.

bounding boxes of sea ice; however, it occasionally misclassifies the categories of sea ice. More specifically, YOLOv8 occasionally misclassifies medium-scale sea ice as large-scale sea ice and conversely misclassifies large-scale sea ice as medium-scale sea ice, As illustrated in Figure 7.

In Figure 7 (a), the sea ice located in the upper left corner is classified as medium-scale, as the longest side of its circumscribed rectangle measures less than 128 pixels. However, it was incorrectly identified by YOLOv8 as large sea ice. Meanwhile, in Figure 7 (b), the sea ice situated at the center is categorized as large-scale, given that the longest side of its circumscribed rectangle exceeds 128 pixels. However, it was inaccurately classified by YOLOv8 as medium-scale sea ice.

With these issues in consideration, we conducted a more thorough examination of YOLO. The YOLO algorithm is designed to extract features from images and classify targets based on these extracted characteristics. The features encompass various types of information, including texture, color, shape, and more. More specifically, YOLO relies more heavily on the aforementioned features for target classification than on the scale information of the targets.

However, in our task of classifying sea ice, the scale information of the target cannot be overlooked. We aim to improve YOLOv8 so that the scale information of the target can serve as a more significant feature for predicting categories of sea ice.

In the inference process of YOLOv8, the role of non-maximum suppression (NMS) is to eliminate redundant prediction boxes and produce the final output. Based on this, we propose the Evidence Fusion module to address the aforementioned issues, the details of the Evidence Fusion module are illustrated in Figure 8.

First, we convert the prediction box information and prediction category information provided by YOLOv8 into evidences. Utilizing an enhanced DSmT fusion inference algorithm [42], we subsequently integrate these two types of evidence to establish a new prediction category. The algorithmic model is primarily composed of the following two components.

3.3.1 *Convert the Information Predicted by YOLOv8 into Evidence Characterizing Uncertainty*

1. The bounding box information predicted by YOLOv8: We begin by counting the instances of sea ice larger than 8 pixels in the satellite image dataset NWPU-RESISC45 [29]. Subsequently, we categorize these sea ice instances into three distinct groups based on their scale, as elaborated in section 4.1. Finally, we generate histograms to illustrate the frequency distribution of the longest side of the circumscribed rectangles for each type of sea ice and fit distribution curves to these

histograms, as depicted in Figure 9.

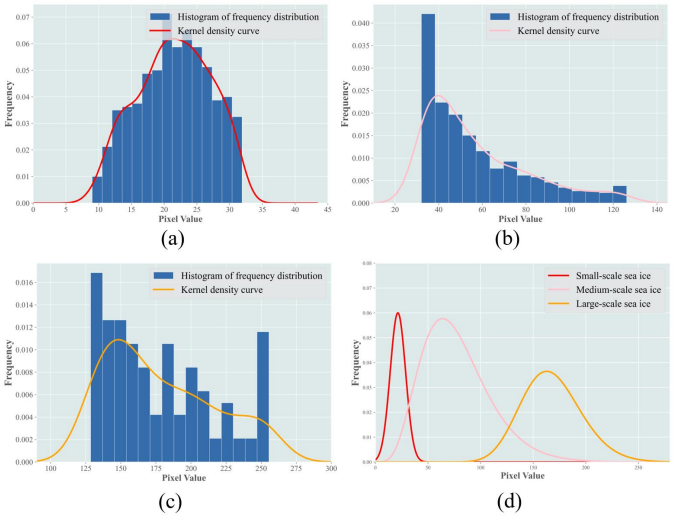


Figure 9. Distribution of the size of three types of sea ice in the NWPU-RESISC45 dataset: (a), (b), and (c) Frequency histograms and probability density curves showing the distribution of *Small-scale sea ice*, *Medium-scale sea ice* and *Large-scale sea ice*, respectively; (d) A schematic diagram is presented, showing the three curves.

Based on the distribution shown in Figure 9, we fit the curve, as shown in equation 15, 16, and 17.

$$f_{small}(l) = \frac{1}{7.32 * \sqrt{2\pi}} * e^{-\frac{(l-21.42)^2}{2*7.32^2}}, l > 0 \quad (15)$$

where 21.42 represents the mean μ of the normal distribution, 7.32 represents the standard deviation σ of the normal distribution. Here, l denotes the pixel value corresponding to the longest side of the circumscribed rectangle for sea ice, and $f_{small}(l)$ is the probability of occurrence of sea ice.

$$f_{medium}(l) = 0.64 * \frac{1}{2^{\frac{11.43}{2}} * \Gamma\left(\frac{11.43}{2}\right)} * \left(\frac{l}{6.78}\right)^{\left(\frac{11.43}{2}-1\right)} * e^{-\frac{l}{2*6.78}}, l > 0 \quad (16)$$

where 0.64 serves as the scaling parameter for the function, 6.78 is the scale parameter, 11.43 denotes the degrees of freedom for the chi-square distribution, and $\Gamma(\cdot)$ is the symbol for the gamma function. Here, l denotes the pixel value corresponding to the longest side of the circumscribed rectangle for sea ice, and $f_{medium}(l)$ is the probability of occurrence of sea ice.

$$f_{large}(l) = 1.06 * \frac{1}{2^{\frac{69.06}{2}} * \Gamma(\frac{69.06}{2})} * (\frac{l}{2.43})^{(\frac{69.06}{2}-1)} * e^{-\frac{l}{2*2.43}}, l > 0 \quad (17)$$

where 1.06 serves as the scaling parameter for the function, 2.43 is the scale parameter, 69.06 denotes the degrees of freedom for the chi-square distribution, and $\Gamma(\cdot)$ is the symbol for the gamma function. Here, l denotes the pixel value corresponding to the longest side of the circumscribed rectangle for sea ice, and $f_{large}(l)$ is the probability of occurrence of sea ice.

We normalize the distribution rules mentioned above, as shown in equations 18, and finally transform the bounding box information predicted by YOLOv8 into evidence that describes the uncertainty.

$$\begin{cases} a_1 = \frac{f_{small}(l)}{f_{small}(l)+f_{middle}(l)+f_{large}(l)} \\ a_2 = \frac{f_{medium}(l)}{f_{small}(l)+f_{medium}(l)+f_{large}(l)} \\ a_3 = \frac{f_{large}(l)}{f_{small}(l)+f_{medium}(l)+f_{large}(l)} \end{cases} \quad (18)$$

where a_1, a_2, a_3 represent the scale reliability for the three types of sea ice, respectively.

2. The category information predicted by YOLOv8: The prediction values for the three types of sea ice—cls_{small}, cls_{medium}, cls_{large}—are included in the prediction information provided by YOLOv8. The category with the highest prediction value indicates the model's predicted target. We convert this set of data into category evidence, as shown in equations 19.

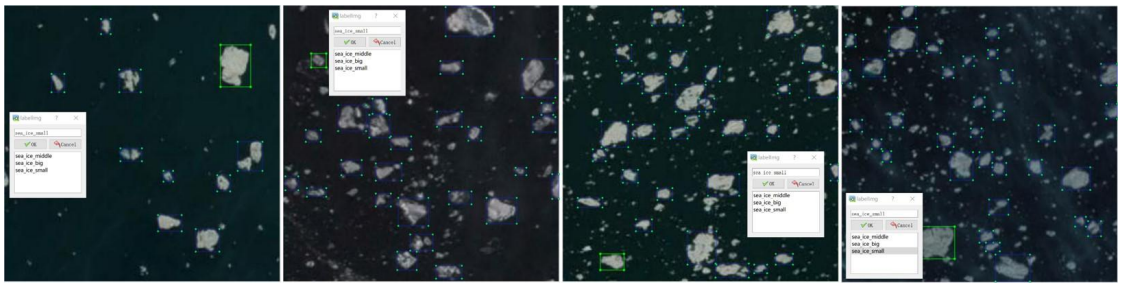
$$\begin{cases} b_1 = cls_{small} \\ b_2 = cls_{medium} \\ b_3 = cls_{large} \end{cases} \quad (19)$$

where b_1, b_2, b_3 represent the category reliability for the three types of sea ice, respectively.

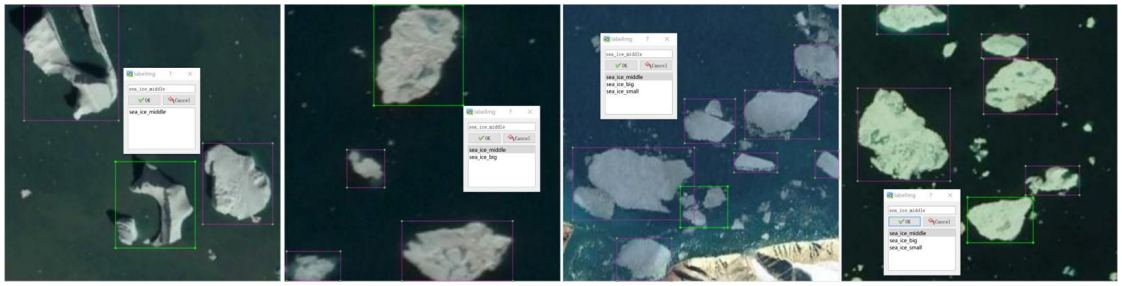
4 Experiments

4.1 Data Collection

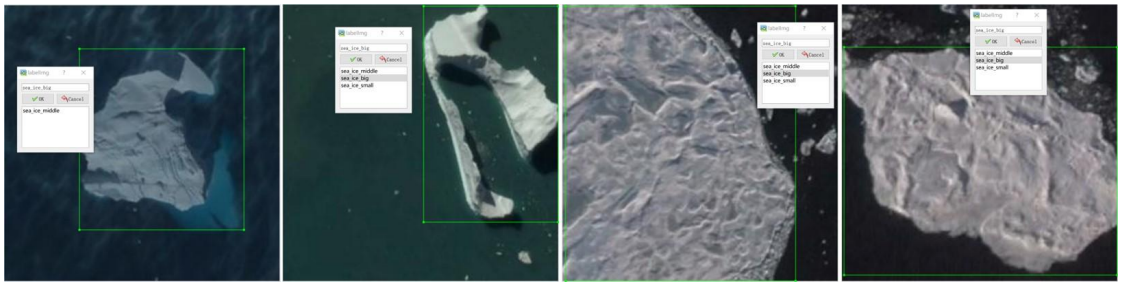
We currently employ two distinct sea ice datasets to assess the detection accuracy of our improved YOLOv8 model. The first dataset is the widely recognized



(a) Small-scale sea ice: the longest side of the circumscribed rectangle for this type of sea ice ranges between 8 pixels and 32 pixels.



(b) Medium-scale sea ice: the longest side of the circumscribed rectangle for this type of sea ice ranges between 32 pixels and 128 pixels.



(c) Large-scale sea ice: the longest side of the circumscribed rectangle for this type of sea ice exceeds 128 pixels.

Figure 10. Three distinct categories of sea ice.

Algorithm 1: Optimized DSmT fusion inference algorithm

Input: The bounding box evidence a_i, The category evidence b_i;
Output: New category result New_cate_i;
for $i = 1, 2, 3$ **do**

$$ma_i = 1 - b_i;$$
$$mb_i = 1 - a_i;$$
$$B_i = a_i^2 \cdot b_i + \frac{a_i^2 \cdot ma_i}{a_i + ma_i} + \frac{a_i \cdot b_i^2 \cdot mb_i}{b_i + mb_i};$$

end

$sumB = \sum_{i=1}^{3} B_i$;
$New_cate_i = 0$;
for $i = 1, 2, 3$ **do**

$$New_cate_i = \frac{B_i}{sumB};$$

end

NWPU-RESISC45 [29], while the second consists of an exclusive Landsat8-based sea ice dataset that we have developed.

4.1.1 *A Sea Ice Dataset Derived from NWPU-RESISC45*

When traversing sea-ice laden waters, it is imperative for the crew to swiftly discern and precisely locate sea ice of diverse dimensions to enable the vessel to bypass the perilous sea ice. Consequently, we categorize the sea ice in these two datasets into three distinct categories. We employ the software labelimg [55] to annotate the images of three distinct types of sea ice, designated as *Small-scale sea ice*, *Medium-scale sea ice*, and *Large-scale sea ice*, as shown in Figure 10.

4.1.2 *An Landsat8-based Sea Ice Dataset*

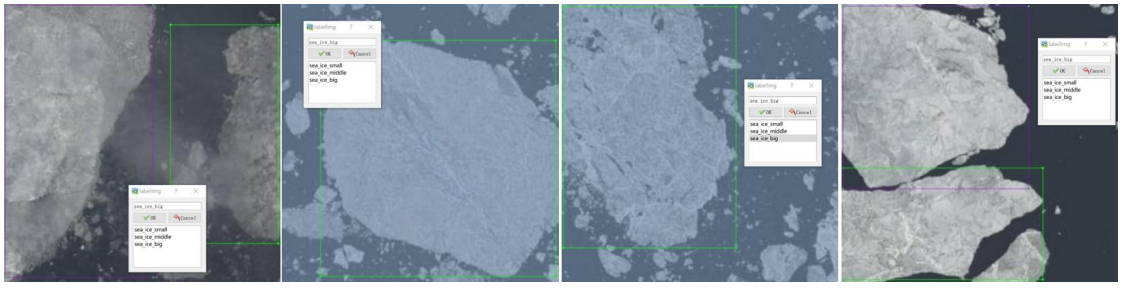
We cropped the satellite data into uniformly sized images and utilized the software labelimg [55] to annotate three distinct types of sea ice present in these images, designated as *Small-scale sea ice*, *Medium-scale*



(a) Small-scale sea ice: the longest side of the circumscribed rectangle for this type of sea ice ranges between 8 pixels and 32 pixels.



(b) Medium-scale sea ice: the longest side of the circumscribed rectangle for this type of sea ice ranges between 32 pixels and 128 pixels.



(c) Large-scale sea ice: the longest side of the circumscribed rectangle for this type of sea ice exceeds 128 pixels.

Figure 11. Three distinct categories of sea ice.

sea ice, and *Large-scale sea ice*, as shown in Figure 11. The specifics of this dataset are provided in the appendix, as shown in Tables 1 and 2.

In addition to the NWPU-RESISC45 [29] dataset, we also explored other datasets to continually validate the performance of our enhanced YOLOv8-based sea ice detector. We identified areas where sea ice occurs at high latitudes and acquired satellite data for these regions, as depicted in Figure 12.

4.2 Implementation Details

We use YOLOv8 as a baseline model. Since the release of YOLOv8 in 2023, it has been deployed on various types of hardware due to its low resource requirements. The following are the experimental configuration, hyper-parameters of improved YOLOv8 object detector and summary of the labels for the three types of sea ice, as shown in Tables 3 and 4.

Table 1. Summary of the labels for the three types of sea ice.

		Small-Scale Sea Ice	Medium-Scale Sea Ice	Big-Scale Sea Ice
NWPU- RESISC45 [29]	Quantity	6820	3710	382
	Percentage	62.5%	34.0%	3.5%
Our Sea Ice Dataset	Quantity	3162	1170	256
	Percentage	69.0%	25.5%	5.5%

Note: We randomly select 70% of the images to train the model, and the remaining 30% of the images were used to verify the training effect.

Table 2. Important information about the exclusive Landsat8-based sea ice dataset.

Attribute	Attribute Value	
	1	2
SPACECRAFT_ID	LANDSAT8	LANDSAT8
ORIGIN	Image courtesy of the U.S. Geological Survey	Image courtesy of the U.S. Geological Survey
LANDSAT_SCENE_ID	LC80482392019215LGN00	LC81300082018179LGN00
LANDSAT_PRODUCT_ID	LC08_L1GT_048239_20190803_20190819_01_T2	LC08_L1TP_130008_20180628_20180704_01_T1
FILE_DATE	2019-08-19T23:27:47Z	2018-07-04T09:14:55Z
OUTPUT_FORMAT	GEOTIFF	GEOTIFF
SENSOR_ID	OLI_TIRS	OLI_TIRS
TARGET_WRS_PATH	48	130
TARGET_WRS_ROW	239	8
DATE_ACQUIRED	2019-08-03	2018-06-28
SCENE_CENTER_TIME	20:32:39.0212890Z	03:26:15.2127540Z
CLOUD_COVER	1.55	2.20
CLOUD_COVER_LAND	0.02	0.12
IMAGE_QUALITY_OLI	9	9
IMAGE_QUALITY_TIRS	9	9

Note: This dataset pertains to the sea ice data corresponding to the aforementioned two scenes. The dataset comprises a total of 430 images, each with dimensions of 256 * 256 pixels.

Table 3. Experimental configuration.

Attribute	Attribute Value
CPU	Core i5 12450H
GPU	NVIDIA GeForce RTX 3050
Running memory	16GB
Storage memory	256GB
Operating system	Win 10
Interpreter	Python 3.9
Deep Learning Frameworks	PyTorch 1.9
IDEA	PyCharm

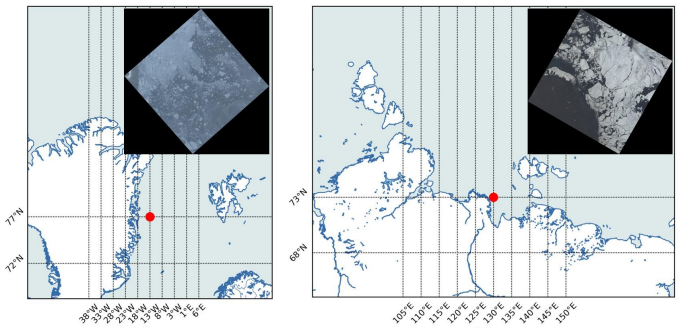


Figure 12. Schematic diagram of the sea ice satellite imagery.

4.3 Comparison with State-of-the-Arts

4.3.1 *Experimental Results Utilizing the NWPU-RESISC45 Dataset*

As shown in Table 5, we conduct experiments on the NWPU-RESISC45 Dataset with mainstream object detection algorithms and perform a comparative analysis of our improved YOLOv8. The selected object detection algorithms include: the two-stage object detection algorithm Faster R-CNN, the Transformer-based object detection algorithm RT-DETR, the one-stage object detection algorithms

Table 4. Hyper-parameters of improved YOLOv8 Algorithm.

Attribute	Attribute Value
epochs	500
batch size	16
imgsz	256
workers	8
close mosaic	Last 10 epochs
optimizer	AdamW
initial learning rate	0.01
final learning rate	0.0001
momentum	0.937
weight decay	0.0005
warm-up epochs	3.0
warm-up momentum	0.8
warm-up bias learning rate	0.1
box loss gain	7.5
class loss gain	0.5
DFL loss gain	1.5
hsv hue augmentation	0.015
hsv saturation augmentation	0.7
hsv value augmentation	0.4
translation augmentation	0.1
scale augmentation	0.9
mosaic augmentation	1.0
mixup augmentation	0.1
copy-paste augmentation	0.1

from the YOLO series, and the improved YOLO series algorithms. Compared to YOLOv8, our improved YOLOv8 achieves a 15.6% increase in Recall, a 5.6% improvement in mAP50, a 3.1% enhancement in mAP50-95, and a 6.8% boost in F1 score, while simultaneously reducing the training time by 13.8%.

In addition, compared to other improved YOLO series algorithms, our enhanced YOLOv8 also demonstrates outstanding detection accuracy and relatively faster convergence speed. As Figure 13 illustrates, we present the detection effects of our improved YOLOv8 alongside the baseline model YOLOv8.

4.3.2 *Experimental Results Utilizing the Landsat 8-Based Sea Ice Dataset*

As shown in Table 6, we conduct experiments on the landsat 8-based sea ice dataset with mainstream object detection algorithms and perform a comparative analysis of our improved YOLOv8. The selected object detection algorithms include: the two-stage object detection algorithm Faster R-CNN, the Transformer-based object detection algorithm

Table 5. Comparisons with the baseline model and state-of-the-arts.

Model	Precision (%)	Recall (%)	mAP50 (%)	mAP50-95 (%)	F1 (%)	Training time (h)	FPS
Faster R-CNN	62.7	80.3	79.2	47.1	70.4	6.09	4.4
SSD	62.3	76.0	78.0	47.6	68.5	0.98	6.9
RT-DETR	74.9	66.1	73	54.2	70.2	3.213	68.0
YOLOv3	66.8	86	74.7	50.9	75.2	5.719	59.2
YOLOv5	73.3	73.3	82.2	52.5	73.3	0.728	**208.3**
YOLOv6	72.4	69.3	78.4	47.0	70.8	3.262	69.4
YOLOv7	76.7	64.6	81.5	49.1	70.1	4.529	87.7
YOLOv9	65.6	**82.3**	78.0	44.2	73.0	5.46	108.7
YOLOv10	84.2	68.1	81.9	49.0	75.3	**0.483**	128.2
ASF-YOLO	64.6	77.9	80.3	45.1	70.6	3.983	53.5
GOLD-YOLO	73.4	74.7	81.1	47.8	74.0	6.118	58.5
Hyper-YOLO	69.2	76.9	83.1	50.5	72.8	6.307	44.4
Improved YOLOv5[39]	71.2	75.1	82.6	51.5	73.1	3.52	126.4
YOLOv8	**84.7**	62.4	81.6	56.2	71.9	1.113	82.6
Our Improved YOLOv8	79.4	78.0	**87.2**	**59.3**	**78.7**	0.959	48.3

In the same group of experiments, the best-performing data is highlighted in bold.

Table 6. Comparisons with the baseline model and state-of-the-arts.

Model	Precision (%)	Recall (%)	mAP50 (%)	mAP50-95 (%)	F1 (%)	Training time (h)	FPS
Faster-RCNN	70.7	84.1	83.2	66.2	76.8	14.79	5.8
SSD	72.1	71.4	72.0	53.5	71.7	2.38	6.9
RT-DETR	66.7	67.2	73.6	45.4	66.9	7.871	40.5
YOLOv3	73.1	71.3	81.4	50.2	72.2	6.29	53.8
YOLOv5	72.5	88.9	90.2	69.5	79.9	2.499	156.4
YOLOv6	74.0	77.0	80.4	46.1	75.5	13.668	52.8
YOLOv7	74.9	74.6	82.8	53.0	74.7	2.142	**192.3**
YOLOv9	86.8	70.3	84.7	52.8	77.7	1.921	181.8
YOLOv10	77.5	**91.6**	91.3	62.2	84.0	**1.581**	112.8
ASF-YOLO	77.3	90.0	91.6	68.5	83.2	5.865	61.7
GOLD-YOLO	81.7	65.9	82.1	51.6	73.0	6.919	55.6
Hyper-YOLO	86.2	66.4	86.3	55.0	75.0	8.5	70.9
Improved YOLOv5[39]	75.1	90.3	90.1	65.9	82.0	2.042	134.7
YOLOv8	86.1	85.6	92.7	67.6	85.8	9.435	50.8
Our Improved YOLOv8	**90.6**	78.0	**93.8**	**71.8**	**87.4**	4.624	53.2

In the same group of experiments, the best-performing data is highlighted in bold.

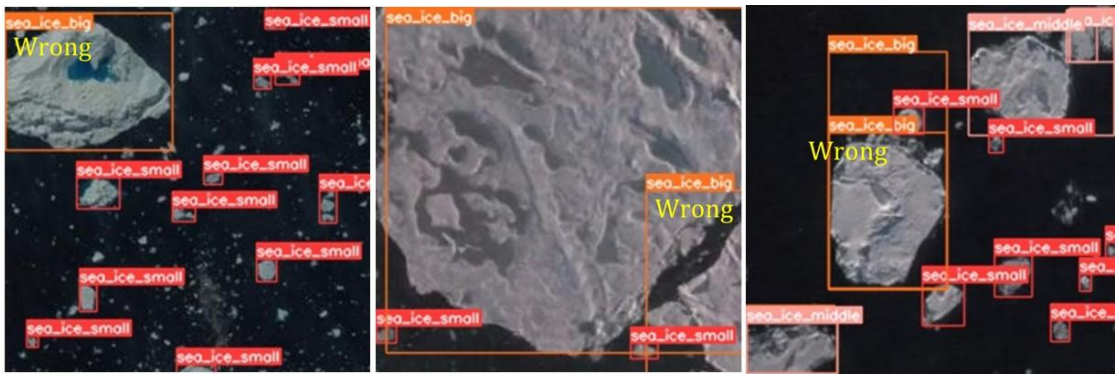
Table 7. Ablation study with improved YOLOv8.

Dataset	**Method**	**Fusion Module**	**Shape-IoU**	**Evidence Fusion**	**mAP50 (%)**	**mAP50-95 (%)**
NWPU-RESISC45 [29]	YOLOv8				81.6	56.2
	Algorithm 1	✓			83.3	57.0
	Algorithm 2		✓		82.8	58.3
	Algorithm 3			✓	85.3	56.8
	Our Improved YOLOv8	✓	✓	✓	**87.2**	**59.3**
Our Sea Ice Dataset	YOLOv8				92.7	67.6
	Algorithm 1	✓			92.9	68.2
	Algorithm 2		✓		93.2	70.1
	Algorithm 3			✓	93.3	68.0
	Our Improved YOLOv8	✓	✓	✓	**93.8**	**71.8**

Note: ✓ denotes an added module based on YOLOv8. In the same group of experiments, the best-performing data is highlighted in bold.

RT-DETR, the one-stage object detection algorithms from the YOLO series, and the improved YOLO series algorithms. Compared to YOLOv8, our improved YOLOv8 achieves a 4.5% increase in Precision, a 1.1% improvement in mAP50, a 4.2% enhancement in mAP50-95, and a 1.5% boost in F1 score, while simultaneously reducing the training time by 51.0%.

In addition, compared to other improved YOLO series algorithms, our enhanced YOLOv8 also demonstrates outstanding detection accuracy and relatively faster convergence speed. As Figure 14 illustrates, we



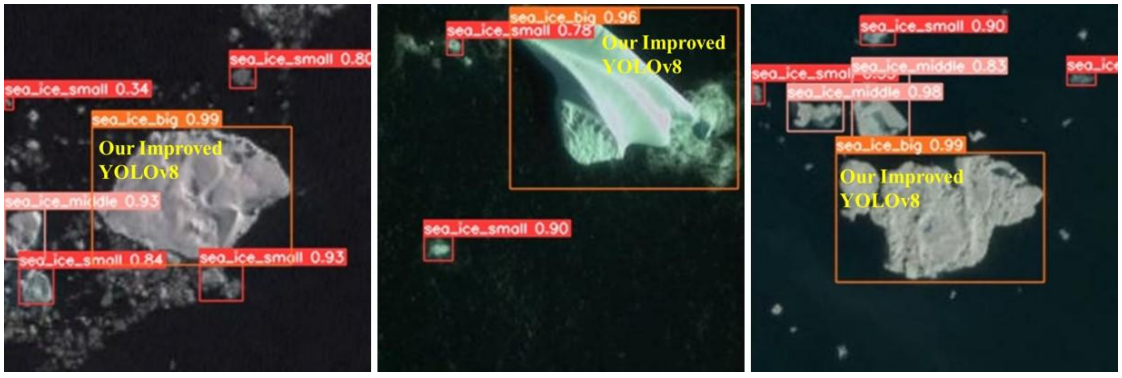
(a) YOLOv8 incorrectly predicts medium-scale sea ice as large-scale sea ice.



(b) Our improved YOLOv8 correctly predicts the results.

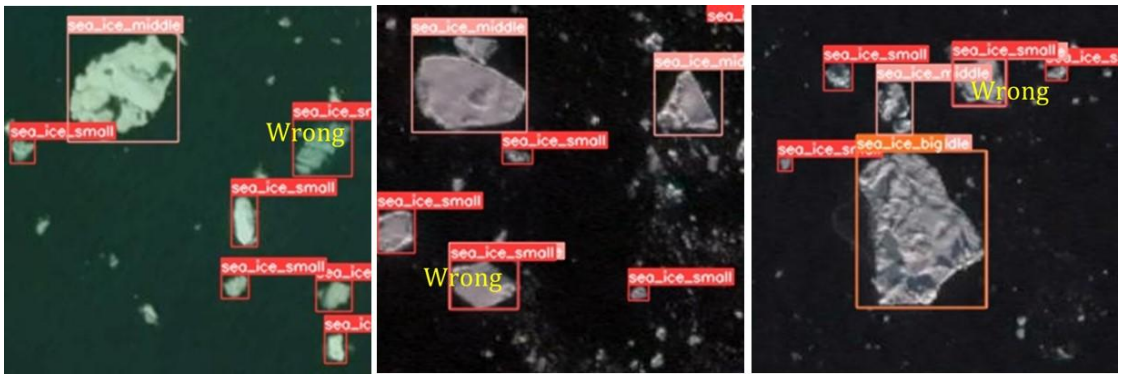


(c) YOLOv8 incorrectly predicts large-scale sea ice as medium-scale sea ice.

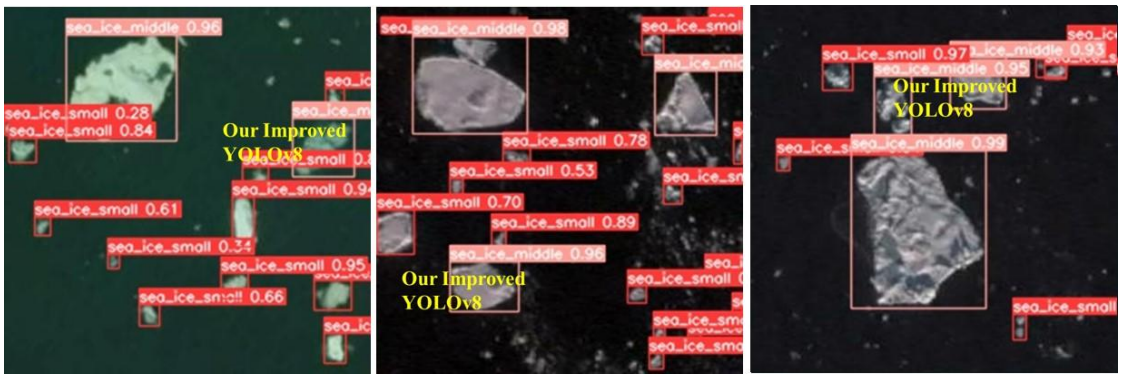


(d) Our improved YOLOv8 correctly predicts the results.

Figure 13. Detection results comparison between YOLOv8 and our improved model.



(e) YOLOv8 incorrectly predicts medium-scale sea ice as small-scale sea ice.



(f) Our improved YOLOv8 correctly predicts the results.



(g) YOLOv8 outputs redundant prediction boxes.



(h) Our improved YOLOv8 correctly predicts the results.

Figure 13. Continued detection results comparison between YOLOv8 and our improved model.

Table 8. Ablation study with fusion module.

Dataset	Method	AP50(%)			mAP50 (%)	mAP50-95 (%)
		small	medium	big		
NWPU-RESISC45 [29]	YOLOv8	80.3	80.8	83.7	81.6	56.2
	+ SE	52.3	**87.2**	**88.3**	75.9	51.7
	+ EMA	69.0	78.9	81.3	76.4	48.0
	+ CA	62.0	74.7	76.2	71.0	52.2
	+ CBAM	79.5	76.1	83.9	79.8	53.5
	+ Fusion Module (Algorithm 1)	**84.5**	81.0	84.4	**83.3**	**57.0**
Our Sea Ice Dataset	YOLOv8	89.1	94.0	95.0	92.7	67.6
	+ SE	79.7	91.8	93.2	88.2	65.6
	+ EMA	84.4	91.1	92.2	89.3	60.3
	+ CA	85.9	91.8	94.3	90.7	66.7
	+ CBAM	80.9	89.3	91.3	87.2	64.4
	+ Fusion Module (Algorithm 1)	**89.3**	**94.1**	**95.3**	**92.9**	**68.2**

In the same group of experiments, the best-performing data is highlighted in bold.

present the detection effects of our improved YOLOv8 alongside the baseline model YOLOv8.

4.4 Model Analyses

4.4.1 Ablation Study

As shown in Table 7, we exhibit the results of ablation experiments based on our improved YOLOv8. On the basis of YOLOv8, we replace the Concat module with a fusion module to obtain Algorithm 1, we substitute the C-IoU with Shape-IoU to obtain Algorithm 2, and we add an evidence fusion module to obtain Algorithm 3.

The experimental data from the NWPU-RESISC45 dataset reveal that our improved network architecture, incorporating a fusion module, improves the mAP50 of YOLOv8 by 1.7%. In addition, we propose an evidence fusion module that improves the mAP50 of YOLOv8 by 3.7%. The experimental data from Our Sea Ice Dataset reveal that our improved network architecture, incorporating a fusion module, improves the mAP50 of YOLOv8 by 0.2%. In addition, we propose an evidence fusion module that improves the mAP50 of YOLOv8 by 0.6%.

4.4.2 Analyses for An Attention-Based Fusion Module

As depicted in Table 8, we present the outcomes of the ablation experiments conducted on YOLOv8, utilizing various mainstream attention mechanisms.

As shown in Table 8, the attention module affects the detection accuracy improvement of YOLOv8, while the fusion module compensates for the negative impact of using only the attention mechanism and enhances the overall detection performance.

To intuitively perceive the positive impact of the fusion module, we use heatmaps to visualize the feature extraction effects of YOLOv8 after introducing our fusion module, as shown in Figure 15. Among them, Figure 15 (a) shows the experimental results of our improved YOLOv8 on the NWPU-RESISC45 dataset, while Figure 15 (b) presents the experimental results of our improved YOLOv8 on our sea ice dataset.

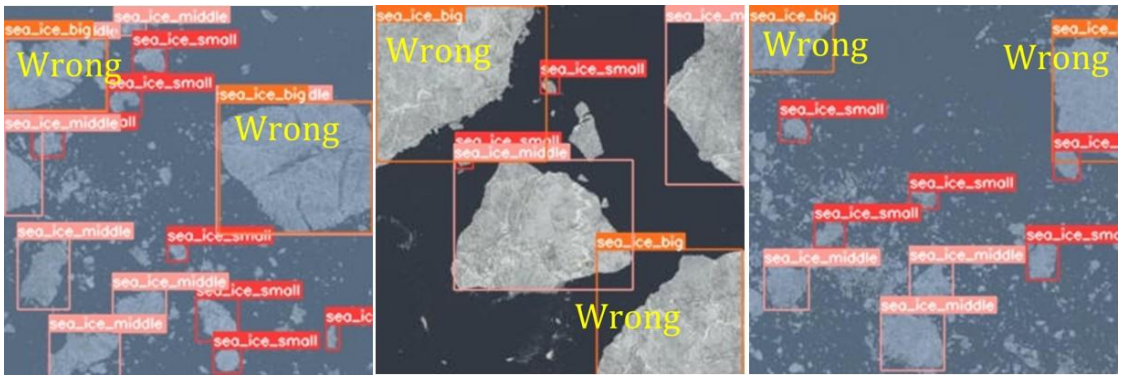
4.4.3 Analyses for Loss Function

We substitute the C-IoU [50] loss with a contemporary mainstream loss function for bounding box regression. As demonstrated in Table 9, we present the results of ablation experiments conducted using YOLOv8.

As illustrated in Table 9, when compared to other enhanced methods, YOLOv8 utilizing Shape-IoU [41] not only demonstrates superior detection accuracy across all three categories of sea ice, but it also significantly reduces convergence time.

Compared to YOLOv8, our Algorithm 2, equipped with Shape-IoU, achieves 1.2% mAP50 improvement, 2.1% mAP50-95 improvement along with a reduction in training time of 23.8%, as validated by the NWPU-RESISC45 dataset. Furthermore, this YOLOv8 variant equipped with Shape-IoU realizes a 0.5% increase in mAP50 and a 2.5% rise in mAP50-95 while concurrently reducing training time by 11.8%, as confirmed by our sea ice dataset.

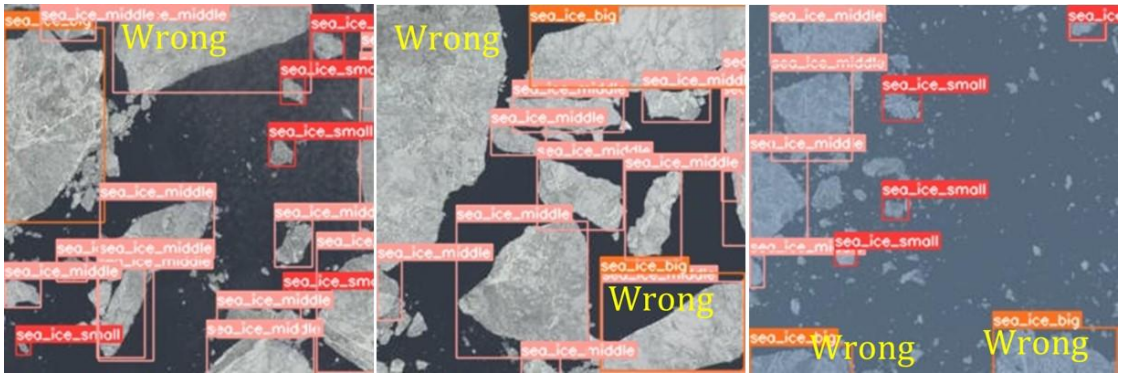
To intuitively perceive the positive impact of the



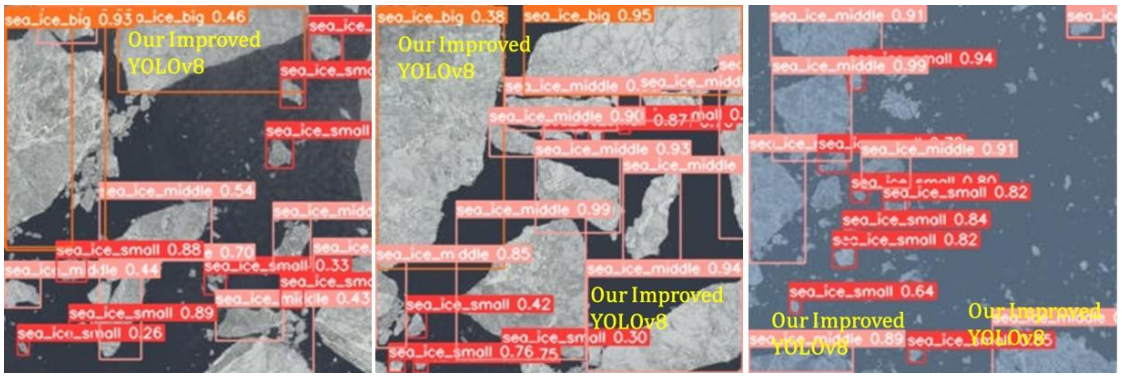
(a) YOLOv8 incorrectly predicts categories of sea ice.



(b) Our improved YOLOv8 correctly predicts the results.



(c) YOLOv8 incorrectly predicts categories of sea ice.



(d) Our improved YOLOv8 correctly predicts the results.

Figure 14. Schematic diagram of experiment results.

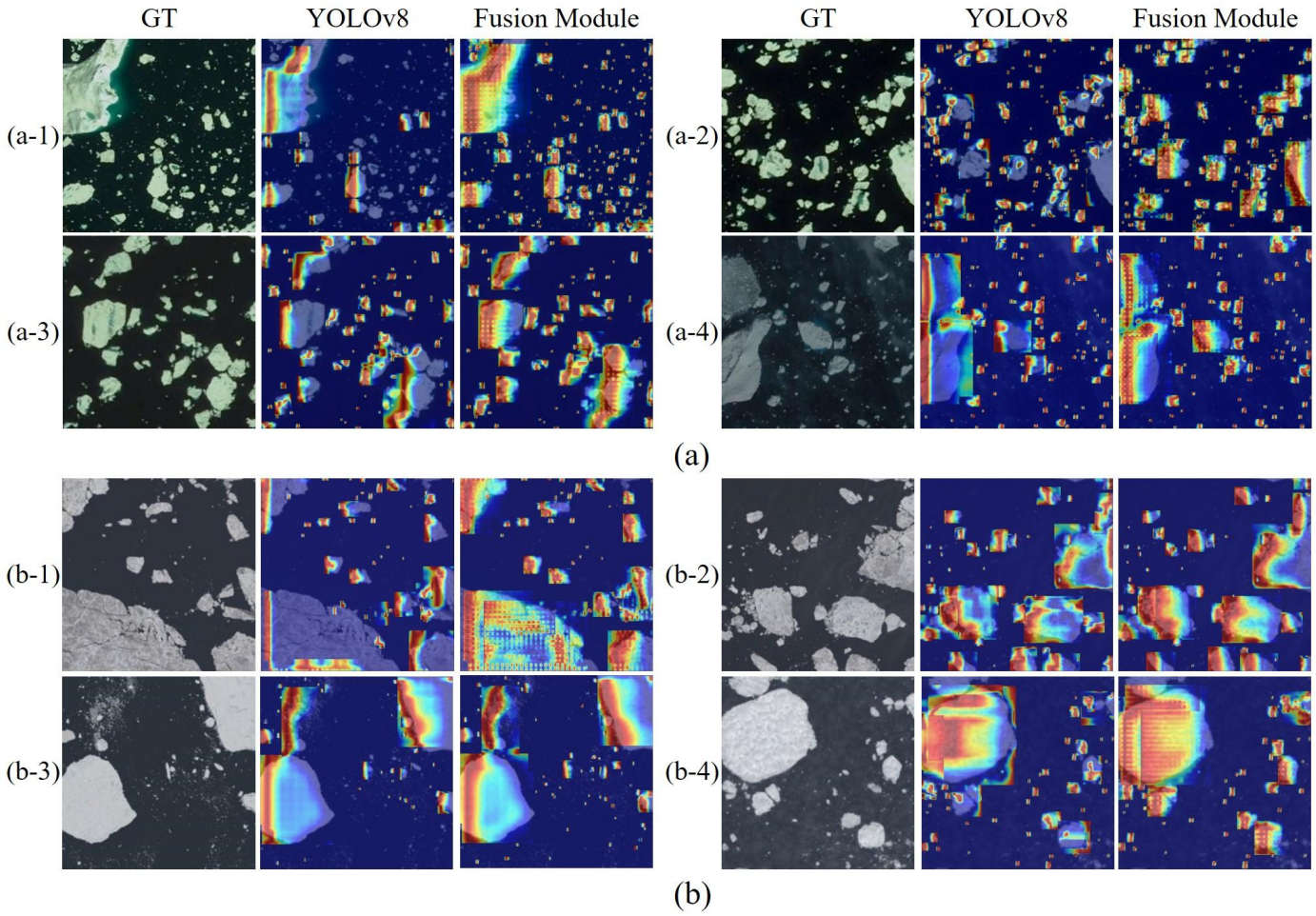


Figure 15. Schematic diagram of the heatmap.

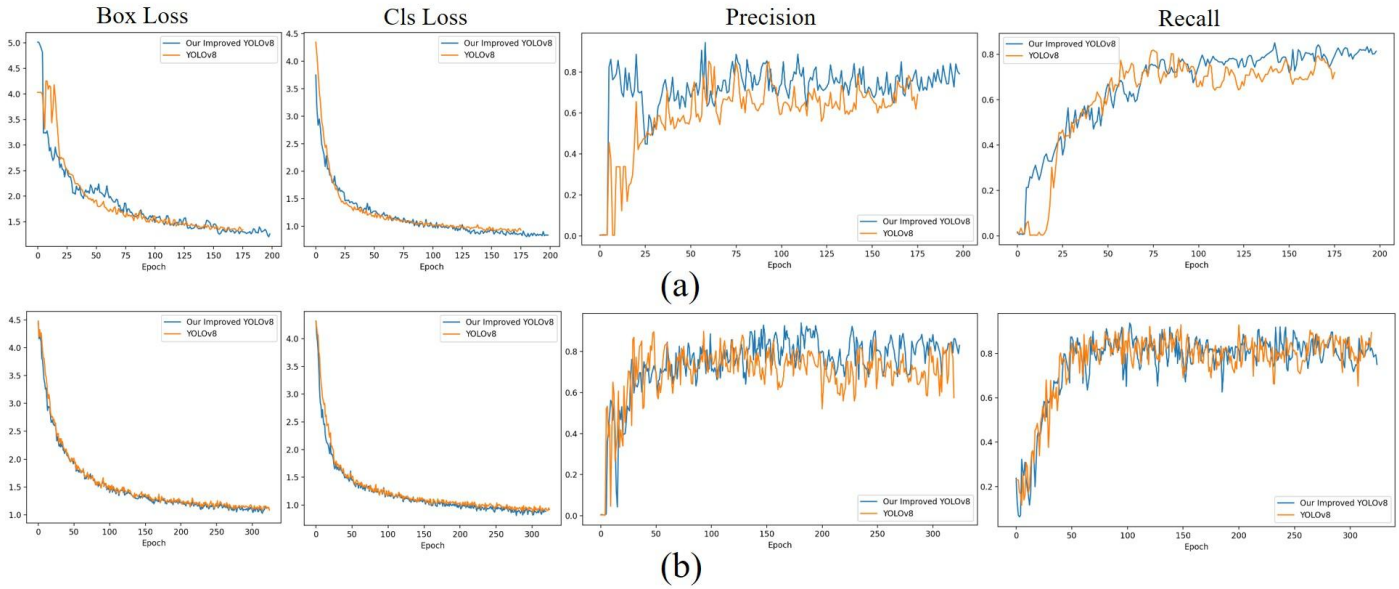


Figure 16. Schematic diagram of the loss function curve.

Shape-IoU, we use Loss function curve to visualize the convergence process of our improved YOLOv8 and YOLOv8, as shown in Figure 16. Among them, Figure 16 (a) shows the convergence process of our improved YOLOv8 and YOLOv8 on the NWPU-RESISC45 dataset, while Figure 16 (b) presents the convergence

Table 9. Ablation study with loss function for bounding box regression.

Dataset	Method	AP50(%)			mAP50 (%)	mAP 50-95 (%)	Training time (h)
		small	medium	big			
NWPU-RESISC45 [29]	YOLOv8 (C-IoU)	80.3	80.8	83.7	81.6	56.2	1.113
	+ G-IoU	72.5	75.9	77.8	75.4	53.1	1.058
	+ D-IoU	72	72.9	75.3	73.4	56.2	1.710
	+ F-IoU	74	75.1	81.1	76.7	56	1.071
	+ S-IoU	74.8	76	82.9	77.9	55.6	0.958
	+ W-IoU	78.5	79.8	82	80.1	56.4	1.000
	+ Inner-IoU	75.3	78.4	78.8	77.5	53.9	1.839
	+ Shape-IoU (Algorithm 2)	**82.3**	**81.4**	**84.7**	**82.8**	**58.3**	**0.848**
Our Sea Ice Dataset	YOLOv8 (C-IoU)	**89.1**	94.0	95.0	92.7	67.6	9.435
	+ G-IoU	80.9	84.7	83.7	83.1	66.4	10.810
	+ D-IoU	84.3	85.5	86	85.3	67.4	10.138
	+ F-IoU	82.9	84.4	83	83.4	68	9.234
	+ S-IoU	82	87	86.6	85.2	68.9	10.465
	+ W-IoU	83.7	93.5	93	90.1	68.7	9.911
	+ Inner-IoU	81.8	86.5	84.8	84.4	69	11.868
	+ Shape-IoU (Algorithm 2)	84	**96.2**	**99.5**	**93.2**	**70.1**	**8.325**

Table 10. Ablation study with evidence fusion.

Dataset	Method	Small-Scale Sea Ice	Medium-Scale Sea Ice	Big-Scale Sea Ice	mAP50 (%)	mAP50-95 (%)	Training Time (h)
NWPU- RESISC45 [29]	YOLOv8 (C-IoU)	80.3	80.8	83.7	81.6	56.2	1.113
	+ Evidence Fusion (Algorithm 2)	**83.9**	**84.5**	**87.5**	**85.3**	**56.8**	-
Our Sea Ice Dataset	YOLOv8 (C-IoU)	89.1	94.0	95.0	92.7	67.6	9.435
	+ Evidence Fusion (Algorithm 2)	**89.7**	**94.5**	**95.7**	**93.3**	**68.0**	-

process of our improved YOLOv8 and YOLOv8 on our sea ice dataset.

4.4.4 Analyses for Evidence Fusion

YOLOv8 utilizes a detection architecture which separates the tasks of classification and localization. This architecture disassembles the tensors, enabling independent predictions for both the bounding box and the category of each target.

The bounding boxes and categories of the targets in sea ice dataset are closely related. Consequently, the design of the detection architecture may result in inconsistencies between the predicted bounding boxes and their corresponding sea ice categories. For instance, a bounding box that represents large-scale sea ice might be inaccurately associated with a predicted category of medium-scale sea ice.

In summary, we simultaneously transform the bounding box and category information predicted by YOLOv8 into multiple pieces of evidence that characterize uncertainty. Subsequently, we utilize an enhanced DSmT fusion inference algorithm to predict the new category. As shown in Table 10, we exhibit the results of ablation experiments based on YOLOv8.

From Table 10, it is intuitively clear that YOLOv8, when using evidence fusion, achieves better detection accuracy on all three types of sea ice.

5 Conclusion

In this paper, we propose a YOLOv8-based sea ice detection algorithm designed to identify sea ice of various sizes in satellite imagery. Firstly, we incorporate an attention-based fusion module into the concatenation component of the YOLOv8 neck network. Secondly, we substitute the C-IoU loss function in YOLOv8 with the more recent Shape-IoU as the boundary regression loss for the detection head. Thridly, we convert the inference results of the YOLOv8's output into uncertain multiple evidences according to the size distribution of sea ice in the

dataset. Subsequently, we fuse multiple pieces of evidences and infer new results based on the improved DSmT fusion inference algorithm. These bring our improved YOLOv8, an sea ice detection algorithm for detecting sea ice of multiple sizes in satellite imagery. The results show that our improved YOLOv8 achieves the state-of-the-art performance in two aspects: identifying sea ice and dividing sea ice size compared with the baseline model and other advanced detection algorithms.

Data Availability Statement

The datasets used in this work include: (1) the publicly available NWPU-RESISC45 remote sensing benchmark dataset, accessible via Baidu Wangpan at http://pan.baidu.com/s/1mifR6tU; and (2) our Landsat-8-derived Sea Ice detection dataset processed from Landsat 8 OLI/TIRS imagery, which has been made publicly available on GitHub at https://github.com/LiuYang0911/A-Proprietary-Visible-Light-based-Sea-Ice-Dataset.

Funding

This work was supported in part by the National Natural Science Foundation of China under Grant 62072392 and Grant 62272405; in part by the Shandong Natural Science Foundation of China under Grant ZR2020QF010; in part by the Yantai City Science and Technology Innovation Development Program - Basic Research Category Projects under Grant 2024JCYJ038.

Conflicts of Interest

Yiping Luo is an employee of Deep Space Exploration Laboratory, Hefei 230000, China.

Ethical Approval and Consent to Participate

Not applicable.

References

[1] Samset, B. H., Zhou, C., Fuglestvedt, J. S., Lund, M. T., Marotzke, J., & Zelinka, M. D. (2023). Steady global surface warming from 1973 to 2022 but increased warming rate after 1990. *Communications Earth & Environment*, 4(1), 400. [CrossRef]

[2] McKay, D. I. A., Staal, A., Abrams, J. F., Winkelmann, R., Sakschewski, B., Loriani, S., ... & Lenton, T. M. (2022). Exceeding 1.5°C global warming could trigger multiple climate tipping points. *Science*, 377(6611), eabn7950. [CrossRef]

[3] Screen, J. A., Deser, C., Smith, D. M., Zhang, X., Blackport, R., Kushner, P. J., ... & Sun, L. (2018). Consistency and discrepancy in the atmospheric response to Arctic sea-ice loss across climate models. *Nature Geoscience*, 11(3), 155-163. [CrossRef]

[4] Li, H., & Fedorov, A. (2021). Persistent freshening of the Arctic Ocean and changes in the North Atlantic salinity caused by Arctic sea ice decline. *Climate Dynamics*, 57(11), 2995-3013. [CrossRef]

[5] Cao, Y., Liang, S., Sun, L., Liu, J., Cheng, X., Wang, D., ... & Feng, K. (2022). Trans-Arctic shipping routes expanding faster than the model projections. *Global Environmental Change*, 73, 102488. [CrossRef]

[6] Min, C., Zhou, X., Luo, H., Yang, Y., Wang, Y., Zhang, J., & Yang, Q. (2023). Toward quantifying the increasing accessibility of the Arctic Northeast Passage in the past four decades. *Advances in Atmospheric Sciences*, 40(12), 2378-2390. [CrossRef]

[7] Kapsar, K., Gunn, G., Brigham, L., & Liu, J. (2023). Mapping vessel traffic patterns in the ice-covered waters of the Pacific Arctic. *Climatic Change*, 176(7), 94. [CrossRef]]

[8] Rodriguez Alvarez, N., Holt, B., Jaruwatanadilok, S., Podest, E., & Cavanaugh, K. (2019). An Arctic sea ice multi-step classification based on GNSS-R data from the TDS-1 mission. *Remote Sensing of Environment*, 230, 111201. [CrossRef]]

[9] Cai, Y., Wan, F., Hu, S., & Lang, S. (2022). Accurate prediction of ice surface and bottom boundary based on multi-scale feature fusion network. *Applied Intelligence*, 52(14), 16370-16381. [CrossRef]

[10] Qaraqe, M., Yang, Y. D., Varghese, E. B., Elzein, A., & Basaran, E. (2024). Crowd behavior detection: Leveraging video swin transformer for crowd size and violence level analysis. *Applied Intelligence*, 54(21), 10709-10730. [CrossRef]

[11] Li, X., Zhou, Y., Du, P., Lang, G., Xu, M., & Wu, W. (2021). A deep learning system that generates quantitative CT reports for diagnosing pulmonary Tuberculosis. *Applied Intelligence*, 51(6), 4082-4093. [CrossRef]

[12] Knausgård, K., Wiklund, A., Sørdalen, T., Halvorsen, K., Kleiven, A., Jiao, L., & Goodwin, M. (2022). Temperate fish detection and classification: A deep learning based approach. *Applied Intelligence*, 52(6), 6988-7001. [CrossRef]

[13] Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. In *2014 IEEE Conference on Computer Vision and Pattern Recognition* (pp. 580-587). [CrossRef]

[14] Girshick, R. (2015). Fast R-CNN. In *2015 IEEE International Conference on Computer Vision* (pp. 1440-1448). [CrossRef]

[15] Ren, S., He, K., Girshick, R., & Sun, J. (2017). Faster R-CNN: Towards real-time object detection with

region proposal networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(6), 1137-1149. [CrossRef]

[16] He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask R-CNN. In *2017 IEEE International Conference on Computer Vision* (pp. 2961-2969). [CrossRef]

[17] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., & Berg, A. C. (2016). Ssd: Single shot multibox detector. In *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14* (pp. 21-37). Springer International Publishing. [CrossRef]

[18] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *2016 IEEE Conference on Computer Vision and Pattern Recognition* (pp. 779-788). [CrossRef]

[19] Redmon, J., & Farhadi, A. (2017). YOLO9000: Better, faster, stronger. In *2017 IEEE Conference on Computer Vision and Pattern Recognition* (pp. 7263-7271). [CrossRef]

[20] Redmon, J., & Farhadi, A. (2018). YOLOv3: An incremental improvement. *arXiv preprint arXiv:1804.02767*.

[21] Bochkovskiy, A., Wang, C. Y., & Liao, H. Y. M. (2020). YOLOv4: Optimal speed and accuracy of object detection. *arXiv preprint arXiv:2004.10934*.

[22] Jocher, G., Stoken, A., Borovec, J., Changyu, L., Hogan, A., Diaconu, L., ... & Dave, P. (2020). ultralytics/yolov5: v3. 0. Zenodo. Retrieved from https://ui.adsabs.harvard.edu/link_gateway/2020zndo...3983579J/doi:10.5281/zenodo.3983579

[23] Li, C., Li, L., Geng, Y., Jiang, H., Cheng, M., Zhang, B., ... & Chu, X. (2023). YOLOv6 v3.0: A full-scale reloading. *arXiv preprint arXiv:2301.05586*.

[24] Wang, C. Y., Bochkovskiy, A., & Liao, H. Y. M. (2023). YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. In *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 7464-7475). [CrossRef]

[25] Jocher, G., Chaurasia, A., & Qiu, J. (2023). Ultralytics YOLOv8. GitHub repository. Retrieved from https://github.com/ultralytics/ultralytics

[26] Wang, C. Y., Yeh, I. H., & Liao, H. (2024). YOLOv9: Learning what you want to learn using programmable gradient information. *arXiv preprint arXiv:2402.13616*.

[27] Wang, A., Chen, H., Liu, L., Chen, K., Lin, Z., Han, J., & Ding, G. (2024). YOLOv10: Real-time end-to-end object detection. *arXiv preprint arXiv:2405.14458*.

[28] Li, W., Hsu, C. Y., & Tedesco, M. (2024). Advancing Arctic sea ice remote sensing with AI and deep learning: Opportunities and challenges. *Remote Sensing*, 16(20), 3764. [CrossRef]

[29] Cheng, G., Han, J., & Lu, X. (2017). Remote sensing image scene classification: Benchmark and state of the art. *Proceedings of the IEEE*, 105(10), 1865-1883. [CrossRef]

[30] Rogers, M., Fox, M., Fleming, A., Zeeland, L., Wilkinson, J., & Hosking, S. (2024). Sea ice detection using concurrent multispectral and synthetic aperture radar imagery. *Remote Sensing of Environment*, 305, 114073. [CrossRef]

[31] Sandven, S., Spreen, G., Heygster, G., Girard-Ardhuin, F., Farrell, S., Dierking, W., & Allard, R. (2023). Sea ice remote sensing—Recent developments in methods and climate data sets. *Surveys in Geophysics*, 44(5), 1653-1689. [CrossRef]

[32] Hu, Y., Hua, X., Yan, Q., Liu, W., Jiang, Z., & Wickert, J. (2024). Sea ice detection from GNSS-R data based on local linear embedding. *Remote Sensing*, 16(14), 2621. [CrossRef]

[33] Liu, L., Dong, X., Lin, W., & Lang, S. (2023). Polar sea ice detection using a rotating fan beam scatterometer. *Remote Sensing*, 15(20), 5063. [CrossRef]

[34] Jafari, Z., Bobby, P., Karami, E., & Taylor, R. (2025). Machine learning-based detection of icebergs in sea ice and open water using SAR imagery. *Remote Sensing*, 17(4), 702. [CrossRef]

[35] Xiong, Y., Wang, D., Fu, D., & Huang, H. (2023). Ice identification with error-accumulation enhanced neural dynamics in optical remote sensing images. *Remote Sensing*, 15(23), 5555. [CrossRef]

[36] Chai, Y., Ren, J., Hwang, B., Wang, J., Fan, D., Yan, Y., & Zhu, S. (2021). Texture-sensitive superpixeling and adaptive thresholding for effective segmentation of sea ice floes in high-resolution optical images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14, 577-586. [CrossRef]

[37] Qiu, Y., Li, X. M., & Guo, H. (2023). Spaceborne thermal infrared observations of Arctic sea ice leads at 30 m resolution. *The Cryosphere*, 17(7), 2829-2849. [CrossRef]

[38] Liang, S., Zeng, J. Y., Li, Z., Chen, K. S., & Zhang, P. (2020). Assessment of four passive microwave sea ice concentrations by using automatic MODIS sea ice classification. In *IGARSS 2020-2020 IEEE International Geoscience and Remote Sensing Symposium* (pp. 3039-3042). [CrossRef]

[39] Ding, S., Zeng, D., Zhou, L., Han, S., Li, F., & Wang, Q. (2023). Multi-scale polar object detection based on computer vision. *Water*, 15(19), 3431. [CrossRef]

[40] Hu, J., Shen, L., & Sun, G. (2018). Squeeze-and-excitation networks. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 7132-7141). [CrossRef]

[41] Zhang, H., & Zhang, S. (2023). Shape-IoU: More accurate metric considering bounding box shape and scale. *arXiv preprint arXiv:2312.17663*.

[42] Guo, Q., Pan, X. & Tang, T. (2023). DSmT-DS Multi-Source Uncertainty Reasoning Methodology. *Multi-source Uncertain Information Reasoning Technology*,

(pp. 59-60).

[43] Bahdanau, D., Cho, K., & Bengio, Y. (2014). Neural machine translation by jointly learning to align and translate. *arXiv preprint arXiv:1409.0473*.

[44] Jaderberg, M., Simonyan, K., Zisserman, A., & Kavukcuoglu, K. (2015). Spatial transformer networks. *arXiv preprint arXiv:1506.02025*.

[45] Woo, S., Park, J., Lee, J. Y., & Kweon, I. S. (2018). CBAM: Convolutional block attention module. In *Proceedings of the European Conference on Computer Vision* (pp. 3-19). [CrossRef]

[46] Wang, C. Y., Liao, H. Y. M., Wu, Y. H., Chen, P. Y., Hsieh, J. W., & Yeh, I. H. (2020). CSPNet: A new backbone that can enhance learning capability of CNN. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops* (pp. 1571-1580). [CrossRef]

[47] Zhang, X., Zeng, H., Guo, S., & Zhang, L. (2022). Efficient long-range attention network for image super-resolution. In *European Conference on Computer Vision* (pp. 649-667). [CrossRef]

[48] Yu, J., Jiang, Y., Wang, Z., Cao, Z., & Huang, T. S. (2016). UnitBox: An advanced object detection network. In *Proceedings of the 24th ACM International Conference on Multimedia* (pp. 516-520). [CrossRef]

[49] Rezatofighi, H., Tsoi, N., Gwak, J. Y., Sadeghian, A., Reid, I., & Savarese, S. (2019). Generalized intersection over union: A metric and a loss for bounding box regression. In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 658-666). [CrossRef]

[50] Zheng, Z., Wang, P., Liu, W., Li, J., Ye, R., & Ren, D. (2020). Distance-IoU loss: Faster and better learning for bounding box regression. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(07), 12993-13000. [CrossRef]

[51] Zhang, Y. F., Ren, W., Zhang, Z., Jia, Z., Wang, L., & Tan, T. (2022). Focal and efficient IOU loss for accurate bounding box regression. *Neurocomputing*, 506, 146-157. [CrossRef]

[52] Gevorgyan, Z. (2022). SIoU loss: More powerful learning for bounding box regression. *arXiv preprint arXiv:2205.12740*.

[53] Tong, Z., Chen, Y., Xu, Z., & Yu, R. (2023). Wise-IoU: Bounding box regression loss with dynamic focusing mechanism. *arXiv preprint arXiv:2301.10051*.

[54] Zhang, H., Xu, C., & Zhang, S. (2023). Inner-IoU: More effective intersection over union loss with auxiliary bounding box. *arXiv preprint arXiv:2311.02877*.

[55] Tzutalin. (2021). LabelImg. *PyPI*. Retrieved from https://pypi.org/project/labelImg/



Yang Liu received the B.S. degree in computer science and technology from Suzhou City University, 264005, China, in 2020. (yangliu0911@s.ytu.edu.cn)



Qiang Guo received the PhD. degree in information and communication engineering from Naval Aeronautical University, and is currently an associate professor at the School of Computer and Control Engineering at Yantai University. His research interests include satellite image processing, radar signal processing, and multimodal data fusion, 264005, China, in 2015. (guoqiang@ytu.edu.cn)



Chengguo Dong male, lecturer at the School of Architecture and Engineering of Weifang University of Science and Technology, obtained a Master's degree in Transportation Engineering from Shandong University of Technology in 2012. His research interests include transportation planning and management, as well as engineering surveying. (dongchengguo@wfust.edu.cn)



Yiping Luo female, senior engineer at Deep Space Exploration Laboratory, obtained a doctor degree in Aerospace Remote Sensing Engineering from the Information and Engineering University in 2010. Research direction: deep space exploration and remote sensing. (luoyiping_nj@126.com)