

中图法分类号: TP301.6 文献标识码: A 文章编号: 1006-8961(2023)11-3509-11

论文引用格式: Du Y L, Wu T Y, Chen K, Chen G and Song W. 2023. Small object detection for ocean eddies using contextual information and attention mechanism. Journal of Image and Graphics, 28(11):3509-3519(杜艳玲, 吴天宇, 陈括, 陈刚, 宋巍. 2023. 融合上下文和注意力的海洋涡旋小目标检测. 中国图象图形学报, 28(11):3509-3519)[DOI:10.11834/jig.220944]

## 融合上下文和注意力的海洋涡旋小目标检测

杜艳玲<sup>1</sup>, 吴天宇<sup>1</sup>, 陈括<sup>2</sup>, 陈刚<sup>3</sup>, 宋巍<sup>1\*</sup>

1. 上海海洋大学信息学院, 上海 200136; 2. 自然资源部东海信息中心, 上海 200136; 3. 国家海洋信息中心, 天津 300171

**摘要:** 目的 海洋涡旋精准检测是揭示海洋涡旋演变规律及其与其他海洋现象相互作用的基础。然而, 海洋涡旋在其活跃海域呈现小尺度目标、密集分布的特点, 导致显著的检测精度低问题。传统方法受限于人工设计参数缺乏泛化能力, 而深度学习模型的高采样率在检测小目标过程中底层细节和轮廓等信息损失严重, 使得目标检测轮廓与目标真实轮廓相差甚远。针对海洋涡旋小目标特点导致检测精度低, 高采样率深度学习模型检测轮廓不精确的问题, 提出一种改进的U-Net网络。**方法** 该模型基于渐进式采样结构, 为获取上下文信息提升不同极性海洋涡旋目标的检测精度, 增加上下文特征融合模块; 为增加该模块对海洋涡旋小目标的关注, 在特征融合前对最底层特征嵌入残差注意力模块, 使模型可以更多关注海洋涡旋的轮廓信息。最后引入数据扩充方法缓解模型存在的过拟合问题。**结果** 本文以南大西洋的卫星海表面高度数据集开展实验, 结果表明, 本文模型检测准确率达到了93.24%, 同时在海洋涡旋的检测数量上与真实结果更加接近, 验证了模型在小目标检测方面的性能更加优秀。**结论** 本文提出的海洋涡旋小目标检测模型, 在检测海洋涡旋的性能与海洋涡旋目标轮廓精准度方面均显著优于全卷积神经网络(fully convolutional network, FCN)等深度学习模型。

**关键词:** 海洋涡旋; 小目标检测; 语义分割; 注意力机制; 特征融合

## Small object detection for ocean eddies using contextual information and attention mechanism

Du Yanling<sup>1</sup>, Wu Tianyu<sup>1</sup>, Chen Kuo<sup>2</sup>, Chen Gang<sup>3</sup>, Song Wei<sup>1\*</sup>

1. College of Information Technology, Shanghai Ocean University, Shanghai 200136, China;

2. East China Sea Information Center, Ministry of Natural Resources, Shanghai 200136, China;

3. National Marine Data and Information Service, Tianjin 300171, China

**Abstract: Objective** Ocean eddies are responsible for most of the material transportation and energy transfers in the ocean. The accurate detection of these eddies serves as the basis for revealing the evolution of ocean eddies and their interactions with other marine phenomena. However, small-scale objects and dense distribution are often observed in the active area of ocean eddies, which leads to problem of low detection accuracy. Traditional detection methods are limited by the poor generalizability of the artificial parameter design. These methods also have poor ocean eddy detection accuracy compared with deep learning methods. However, a deep learning model with high sampling rate loses the underlying details and contour information in the process of small target detection. The target detection contour is located far from the real con-

收稿日期: 2022-10-10; 修回日期: 2023-01-14; 预印本日期: 2023-01-21

\* 通信作者: 宋巍 wsong@shou.edu.cn

基金项目: 国家自然科学基金项目(41906179, 61972240)

Supported by: National Natural Science Foundation of China (41906179, 61972240)

tour of the target. To address the low detection accuracy caused by the loss of low-level detail information and contour information of small-scale ocean eddy targets, this paper proposes an improved U-Net network. **Method** Based on the U-shaped progressive sampling network, a context feature fusion module is added to fuse the features of each coding layer, and a residual attention mechanism is added to the target features before the feature fusion in order for the model to pay attention to the contour information of the ocean eddies. A data augmentation method is then introduced to reduce the overfitting problem of the model. Feature fusion is carried out through the context feature fusion module, which takes the three-layer feature map of the U-shaped structure coding layer of the U-Net network as input, the lowest-level feature map as the target feature, and the last two-layer feature map as the context and target features. The context feature map is initially upsampled to the same size as the lowest-level feature through the deconvolution structure, and the number of channels is reduced to 1/2 of the lowest-level feature in order to prevent the amount of information of the context feature from exceeding that of the target feature. L2 norm and ReLU are then used to achieve the fusion of context and target features. The proposed model uses two contextual feature fusion modules, which take the first to third layer feature maps of the encoding layer as input and the second to fourth layer feature maps as input, respectively. The residual attention mechanism consists of two processing channels. The first channel has a residual structure (batch norm, conv of  $1 \times 1$  kernel and multiple concatenation of ReLU) that prevents gradient disappearance and extracts certain contour information, while the second channel comprises a down-up sampling layer and a sigmoid layer to extract high-level semantic information. To effectively reduce the over-fitting phenomenon, random region sampling and random mask processing are used for data augmentation. In the experiment, the model is trained in the NVIDIA GTX 1080Ti GPU environment, where its initial learning rate is set to  $1 \times 10^{-3}$ , the loss function is optimized by the Adam optimizer, the batch size of the model training is set to 16, and the number of iterations is set to 200. **Result** The satellite sea surface height dataset of the South Atlantic is used for the experiments. Ablation experiments are carried out to test the influence of each module on the performance of the ocean eddies detection model. The effects of adding the context feature fusion module, adding the attention mechanism module, and adding both modules at the same time are compared, and the detection effect after adding the data augmentation method is analyzed. In the ablation experiment, due to the introduction of the contextual feature fusion module and the residual attention mechanism, the model can fuse the contextual features of the ocean eddies in different feature layers, and the network can extract additional low-level spatial details of the ocean eddies. Each module improves the detection performance of the model, and the optimal detection accuracy of the model after using the data augmentation method reaches 93.27%. Compared with other deep learning models, the proposed model has a detection accuracy of up to 93.24%, and its detected number of ocean eddies is closer to the truth, thereby verifying its excellent performance in small target detection. Meanwhile, compared with the fully convolutional network (FCN) model, the proposed model can detect more small-scale ocean eddy targets, and the detected ocean eddy target contour is closer to the truth, thereby verifying the positive effect of progressive sampling on small target detection. **Conclusion** The proposed model significantly outperforms the other deep learning models in detecting ocean eddies. Compared with the state-of-the-art, the proposed model achieves a higher small target detection accuracy, and the detected contour of ocean eddies is closer to the truth.

**Key words:** ocean eddy; small object detection; semantic segmentation; attention mechanisms; feature fusion

## 0 引言

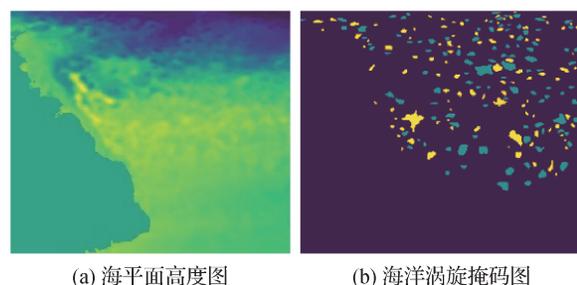
海洋涡旋普遍存在于世界大洋和边缘海中,空间尺度通常为几公里到几百公里不等。根据海洋涡旋旋转方向可分为气旋涡和反气旋涡,将逆时针旋转(在北半球)归为气旋涡,反之为反气旋涡(Faghmous等,2013)。海洋涡旋在海洋中动能、质量、热量和营养物质等的传输中起着重要作用,对海洋的

循环和大规模水分布(Chen等,2011;Faghmous等,2012)影响显著,因此海洋涡旋的检测具有重要意义。

国内外学者已开展了丰富的海洋涡旋检测研究。传统的检测方法主要分为3大类:1)基于物理参数的方法,包括Okubo-Keiss参数法(Chelton等,2007;Henson和Thomas,2008)、卷绕角度法(Chaigneau等,2008;Ali Sadarjoen和Post,2000)和二维小波法(Doglioli等,2007);2)基于流场几何特

征的方法,其中有基于速度矢量几何的涡流检测算法(Nencioli等,2010)和水平速度分量检测方法(Williams等,2011);3)基于混合方法,其中有通过结合速度矢量几何以及物理参数法的非线性涡旋检测方法(Chelton等,2011)以及McWilliam(1984)最早提出的一种基于相对涡旋同时考虑流场几何特征的涡旋检测方法。但是现有的海洋涡旋检测方法存在对参数设置敏感、涡核的初始值设置以及检测方法的泛化性差等问题。随着深度学习的不断发展,衍生出许多将深度学习应用到海洋涡旋检测的方法。Lguensat等人(2018)在U-Net网络的基础上降低了原有模型结构,同时根据先验知识改进了多分类的交叉熵损失函数来提高对涡旋的检测。Fan等人(2020)认为中尺度涡旋检测与其他模态数据密切相关,因此提出一种使用多模态数据的涡旋检测网络(eddy detection network, EDNet),通过对涡旋的多模态数据进行融合来提高涡旋检测的精度。Xu等人(2019)为了充分利用全局上下文信息来捕获更多的细节,使用金字塔场景解析网络(pyramid scene parsing network, PSPNet)(Zhao等,2017)对涡旋进行检测,PSPNet的金字塔模块通过不同尺度的池化模块来获取不同比例的子区域特征,进而提高对不同尺度涡旋的检测精度,同时对检测到的涡旋数量以及大小等进行统计与分析。沈飙等人(2020)使用Mask-RCNN(mask region-based convolutional neural network)(He等,2017)网络来检测进行多模态数据融合的海洋涡旋。杜艳玲等人(2022)在RetinaNet(Lin等,2017)基础上做出改进,提高多尺度海洋涡旋的检测精度。

如图1所示,对大尺度空间区域进行涡旋检测时,由于遥感成像空间分辨率的原因,导致原本空间尺度较大的海洋涡旋在数据表示中呈现为较小的目标,尤其在海洋涡旋活跃海域小尺度目标、密集分布普遍存在,导致其检测精度低问题显著。基于U-Net网络(Ronneberger等,2015)在小目标检测方面具有优势,本文采用其渐进式下采样和上采样的网络结构,同时融合上下文信息提取模块,通过获取更加丰富的目标轮廓和尺度特征信息提升不同极性海洋涡旋目标检测精度。此外,为了增强小目标关注度,加入注意力机制,在获取丰富的细粒度底层空间信息的同时减少采样过程中轮廓信息丢失,提升小目标检测精度。



(a) 海平面高度图 (b) 海洋涡旋掩码图

图1 海洋涡旋数据集样例

Fig. 1 Sample of ocean eddy dataset

((a) sea surface height map; (b) ocean eddy mask map)

## 1 相关工作

### 1.1 基于深度学习的语义分割

采用语义分割技术实现目标检测是计算机视觉领域的研究热点,深度学习极大地提高了语义分割的准确性。全卷积神经网络(fully convolutional network, FCN)(Long等,2015)的提出,掀起了深度学习在语义分割领域应用的热潮。FCN结构将原有的全连接层进行替换,将输入图像尺寸的解码特征作为输出,减少了位置信息的丢失。但由于最后的全卷积层上采样率过高,会引起底层细节和轮廓信息丢失,尤其对小目标的检测效果不理想。Ronneberger等人(2015)提出U-Net模型,以FCN为基础,通过渐进式下采样和上采样减少特征的丢失进而增加小目标的检测性能,最早应用于医学图像分割。Chen等人(2017)提出的用于图像分割的深度卷积网络DeepLabV3以及Zhao等人(2017)提出的PSPNet(pyramid scene parsing network)均通过上下文特征的提取提高模型检测的性能,前者使用不同空洞率的空洞卷积对同一特征图进行操作,一方面通过空洞卷积来增加感受野,提取不同尺度的信息,另一方面学习上下文信息来提高模型的性能。后者使用的金字塔池化模块通过不同采样率的池化模块来获取上下文信息进而提升模型对不同尺度目标的检测性能。DeepLabV3+网络(Chen等,2018)在针对上下文特征提取改进的模型中最具代表性。虽然现有方法针对小目标检测提出了改进,但现阶段小目标精准检测仍是计算机视觉中面临的挑战。

### 1.2 小目标检测

由于遥感影像成像特点导致其小目标问题显

著,现阶段已有不少学者提出小目标检测的一系列方法(Fu等,2017;Jeong等,2017;Lim等,2021;Liu等,2021)。反卷积单步检测器(deconvolutional single shot detector, DSSD)(Fu等,2017)对单步检测器(single shot multibox detector, SSD)中所有特征图使用反卷积进行上采样来获得放大的特征图,但是使用反卷积会增加模型的计算量,降低检测的速度。R-SSD(rainbow single shot multibox detector)(Jeong等,2017)融合通过池化和反卷积获得的不同尺度特征,与DSSD相比获得了更高的精度和速度。Lim等人(2021)通过在SSD的基础上增加上下文与注意力机制的融合结构来提升对小目标的检测精度。贾可心等人(2022)在SSD的基础上增加注意力机制,来提高对海面小目标的检测精度。用于小目标检测的高分辨率网络(high-resolution detection network, HRDNet)(Liu等,2021)使用浅网络来处理高分辨率图像,深网络处理低分辨率图像进而提升高分辨率图像的小目标检测精度。

### 1.3 视觉注意网络

注意力机制往往关注的是面向特定任务的局部特征而非全局特征,因此在提取显著特征方面具有优秀的表现。Wang等人(2017)通过叠加残差注意力模块,提高了ImageNet数据集的分类性能。Woo等人(2018)提出了卷积注意模块(convolutional block attention module, CBAM)。该模块由通道和空间注意力模块组成,轻量通用化使得该模块可以简单有效地集成到现有的模型中,在增加少量参数量以及运算成本的基础上提升模型的特征提取能力。为了降低注意力机制的计算成本,用于语义分割的期望最大化注意力网络(expectation-maximization attention network, EMANet)(Li等,2019)被提出,同时EMANet非局部网络的设计将自注意力机制应用到计算机视觉当中。全局上下文网络(global context network, GCNet)(Cao等,2019)同时在注意力机制与全局上下文的使用中做出改进,使得模型可以更有效地对全局上下文建模。

## 2 方法

### 2.1 融合上下文特征和注意力机制的U-Net

在检测偏差容忍度较小的小尺度目标时,需要检测模型提取更加细粒度的空间信息实现小目标的

精确定位。而现有深度学习方法由于高采样率导致目标底层空间信息与轮廓特征的丢失,引起小目标检测精度的降低。

基于遥感影像的海洋涡旋检测小目标问题是限制当前海洋涡旋检测精度的关键,获取丰富的细粒度底层空间信息、减少采样过程轮廓信息丢失是提升小目标检测精度的重要途经。U-Net网络采用渐进式下采样和上采样方式进行特征提取,同时利用了多尺度特征,一方面可以减少高层语义特征的损失,另一方面可以充分提取和利用底层空间特征,在小目标检测方面具有一定的优势。因此本文在U-Net的基础上改进以提升小目标的检测性能,首先增加上下文特征融合模块来增加对底层空间信息的利用。此外,在网络模型中增加残差注意力模块来提高对特征的提取能力,同时残差注意力模块中的残差结构可以防止网络产生梯度消失。

U-Net网络对上下文信息的利用较少,模型使用卷积神经网络的串行连接,其U型结构主要提取多尺度信息。通常,获取上下文信息的模块会对同一特征进行不同尺度的池化处理(Zhao等,2017;Chen等,2017),通过特征融合来获取上下文信息(Lim等,2021)。上下文信息有利于将不同目标更细致地区分,获得细化轮廓,同时使不同目标被区分得更加明确,降低不同目标重叠或相隔较近而导致检测效果不理想的问题。由于数据集中的海洋涡旋普遍以小目标的形式呈现,小目标检测问题一直都是深度学习任务中有待解决的问题,所以本文为了解决小目标海洋涡旋的检测问题,使用注意力机制与上下文特征融合模块,来达到提高小目标检测效果的目的。针对U-Net网络对小目标检测精度不足,融合注意力机制,提升模型的特征获取和表达能力,进一步提高其对小目标的检测精度。Lguensat等人(2018)表示,对于涡旋检测任务,现有的模型复杂而训练数据样本数量少导致模型存在过拟合,本文方法也存在同样的问题。为了处理过拟合现象,采用随机区域采样以及随机掩码处理的数据扩充方法,增加训练数据样本量,同时增加训练数据的多样性以降低过拟合。

本文基于U型的渐进式采样网络,改进的U-Net结构FA-U-Net(feature-attention-U-Net)如图2所示。其中,左半部分为下采样的编码器结构,右半部分为

上采样的解码器结构。编码器和解码器都由4个相同的编解码块构成,编码器中的每个编解码块由两组相同的卷积核为 $3 \times 3$ 的卷积层、ReLU激活函数构成,最后通过一个 $2 \times 2$ 步长的最大池化模块对得到的特征图进行下采样,下采样后获得一个大小为原特征图 $1/2$ 的特征图,同时增大模型的感受野。编码器用于提取输入图像的特征信息,最终4个下采样层得到4个不同尺度的特征图。解码器中的每个解码块首先对前一层输出的特征图使用步长为 $2 \times 2$ 的转置卷积进行上采样。与编码块不同的是,解码块的第1组 $3 \times 3$ 卷积核的卷积层和ReLU激活函数

是用来对上采样后的特征图与对称的编码结构输出的特征图进行特征融合,最后经过4次上采样后获得与原图像尺寸相同的特征图(董子意等,2022)。同时编码器所得到的前3层特征图,以及第1—4层特征图分别进行特征融合,以获取上下文特征。对输入到上下文特征融合模块的3层特征中最底层的特征图(空间尺度最大的特征图)使用残差注意力模块,以增强对海洋涡旋轮廓信息以及高级语义特征的提取。最后使用SSD的分层预测结构,分别对两个上下文特征融合模块以及解码器最后一层的特征进行预测。

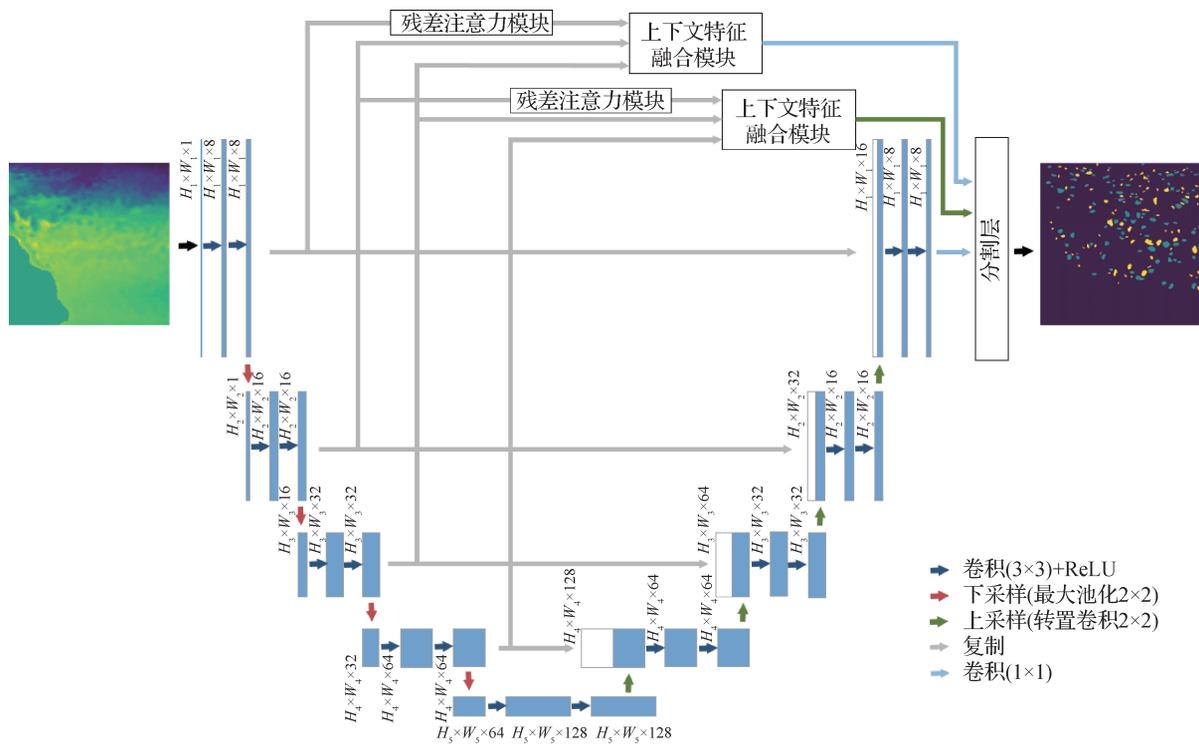


图2 FA-U-Net结构图

Fig. 2 Structure diagram of FA-U-Net

## 2.2 上下文特征融合模块

上下文特征融合模块结构如图3所示,该模块对编码器输出的不同层特征图进行上下文信息提取。加入特征融合模块进行分支预测一方面增加对底层特征的利用,有利于小目标检测。另一方面通过两个不同的特征融合模块可以提高模型对多尺度目标的检测能力。第1个上下文特征融合模块将编码器前3层的特征图作为输入,最底层的特征图作为目标特征,后两层特征图作为上下文特征,通过转置卷积使其具有与目标特征相同的空间尺度、通道数为目标特征一半的上下文特征。通过转置卷积获

得的上下文特征通道数仅有目标特征一半,目的是使上下文信息量不会超过目标特征本身。最后将得到的上下文特征与目标特征堆叠进行单独预测。第2个上下文特征融合模块将编码器的第1—3层特征图作为输入,第1层特征图作为目标特征。在预测时由于第1层特征为原图像尺寸的 $1/2$ ,所以需要堆叠后的特征使用转置卷积进行上采样后再进行预测,由于上下文特征融合模块的目标特征都为底层特征,所以提取了海洋涡旋更丰富的轮廓信息。

## 2.3 残差注意力模块

注意力机制可以帮助网络提取更多有用特征,本

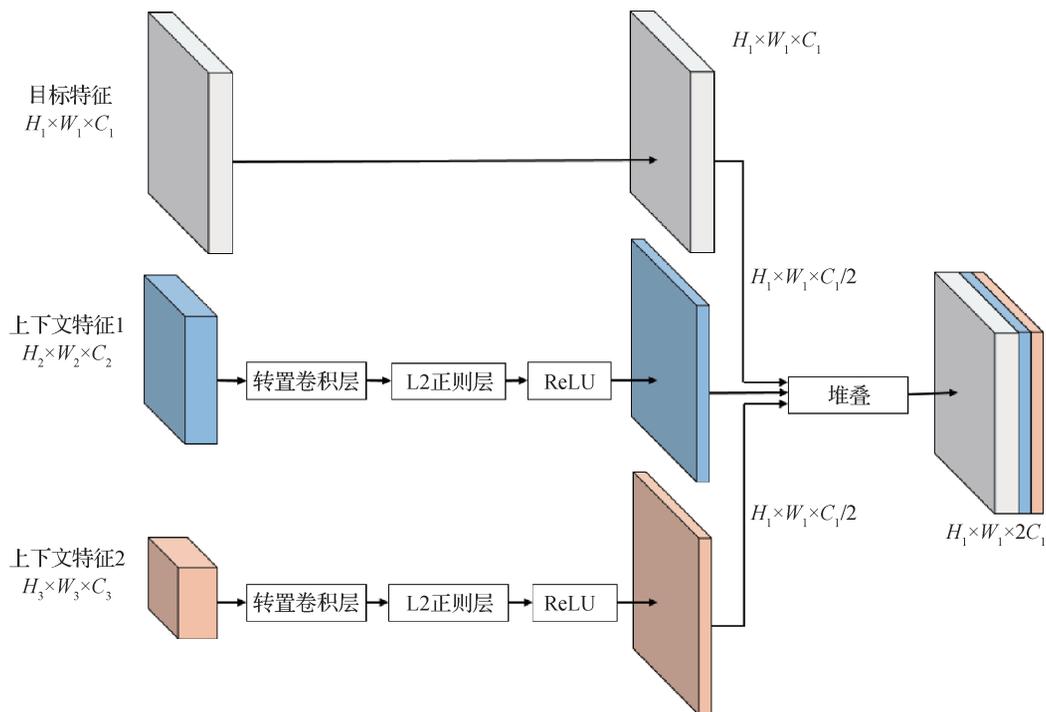


图3 上下文特征融合模块操作过程

Fig. 3 Contextual feature fusion module operation procedure

文在改进后的U-Net网络的上下文特征融合模块前加入残差注意力模块(Lim等,2021),该模块对输入到上下文特征融合模块的最底层特征进行操作,其结构如图4所示。其中,残差注意力模块结构如图4(a)所示,该模块由两个路径组成,其中一条路径由两个级联的残差模块组成(如图4(c)所示),残差模块由3组相同的批量正则层、ReLU激活函数以及卷积核为 $1 \times 1$ 的卷积层进行串联,用以保留底层海洋涡旋的轮廓特征。另一条路径通过下一上采样模块提取更多的语义信息,其后连接的sigmoid函数生成注意力图。下一上采样模块如图4(b)所示,通过级联残差模块以及两个 $2 \times 2$ 步长的最大池化和双线性插值的上采样模块组成。残差结构一定程度上保留了底层特征中海洋涡旋的轮廓信息,而下一上采样结构可以提取更多的高层语义信息,弥补底层特征缺乏高层语义信息的问题。为了降低模型的复杂度,防止训练数据较少带来的过拟合问题,使用更简单的残差注意力模块,将下一上采样模块统一为两次下采样和两次上采样。

### 3 实验

#### 3.1 实验数据

本文采用公开的海洋涡旋数据集 EddyNet-Data

(<https://github.com/redouanelg/EddyNet/tree/master/Data>)(Lguensat等,2018)。该数据集为南大西洋范围 $10^\circ \sim 58^\circ \text{N}$ , $0^\circ \sim 64^\circ \text{W}$ 的子区域。

EddyNet-Data数据集采用经典的PET(py-eddy-tracker)方法(Mason等,2014)生成了训练所需要的海洋涡旋标签数据。该数据集由2000年1月1日至2011年12月31日共12年,以天为单位的4383幅海表面高度图像组成,空间分辨率为 $0.25^\circ$ 。该数据集将前11年的4018幅海表面高度图像作为训练集,将2011年的365幅海表面高度图像作为测试集(董子意等,2022)。

#### 3.2 实验设置

本文的海洋涡旋图像检测本质上是对图像像素点进行多分类任务,即判断图中像素点是气旋涡、反气旋涡或非涡旋(董子意等,2022)。实验使用的ground-truth数据为PET方法生成的掩码标签,通过与掩码标签进行逐像素对比来计算模型性能指标。评价指标采用通用指标F1-Score、气旋涡与反气旋涡的平均精确度(average precision, AP)以及模型的受试者工作特征曲线(receiver operating characteristic, ROC)来对模型的涡旋检测性能进行评估。

在考虑到模型计算效率、设备硬件条件以及结果精度,采用交叉熵函数作为损失函数,Adam优化

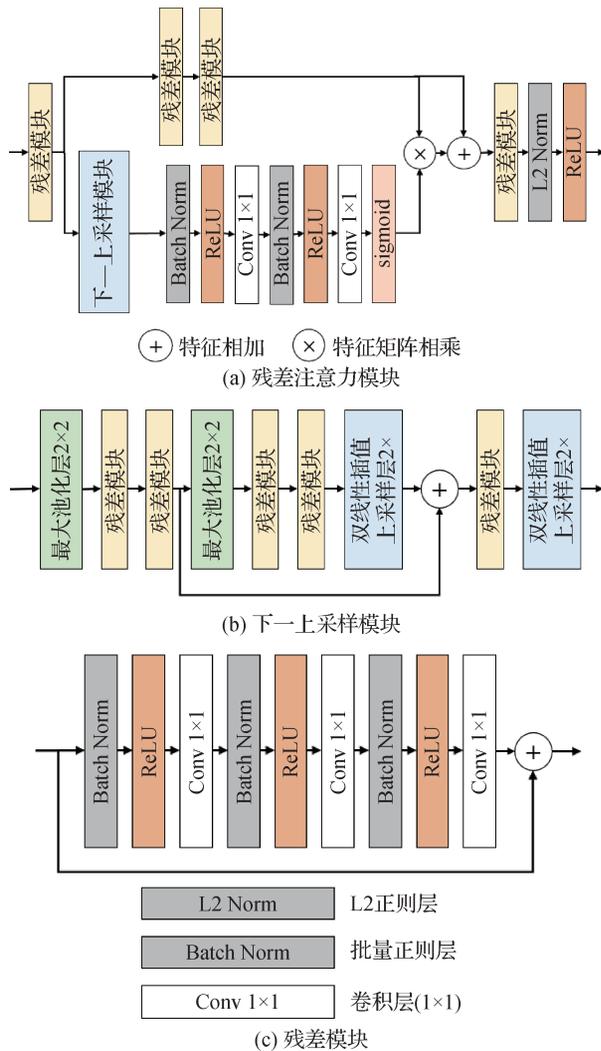


图4 残差注意力模块各部分结构图

Fig. 4 Structural diagram of each part of the residual attention module ((a) residual attention module; (b) down-up sample module; (c) residual module)

器对损失函数进行优化,本文模型采用交叉熵函数作为损失函数,Adam优化器对损失函数进行优化,初始学习率设置为0.001。模型训练时的批量大小设置为16,迭代次数为200。经过200次迭代训练,网络最终收敛。

### 3.3 消融实验

为了测试各模块对海洋涡旋检测性能的影响,比较了单独加入上下文特征融合模块、残差注意力模块以及同时加入二者后的模型性能。最后加入数据扩充方法降低过拟合现象。由表1可知,上下文特征融合模块和残差注意力模块的引入使得模型可以融合不同特征层海洋涡旋的上下文特征,同时使网络可以提取到更丰富的海洋涡旋的底层空间细节

信息,各模块均提升了模型对海洋涡旋的检测性能,其中最优检测精度可达93.27%。

表1 模型消融实验结果

Table 1 Ablation experiments verify the results

模型	AP		F1-Score
	气旋涡	反气旋涡	
U-Net	0.326 8	0.456 4	0.889 8
F-U-Net	0.440 9	0.513 5	0.928 3
A-U-Net	0.462 0	<b>0.657 5</b>	0.923 6
FA-U-Net	0.481 0	0.520 4	0.932 4
FA-U-Net(数据扩充)	<b>0.532 0</b>	0.524 6	<b>0.932 7</b>

注:加粗字体表示各列最优结果。

### 3.4 结果分析

采用 EddyNet-Data 数据集,与 FCN 方法、EddyNet 方法、DeepLabV3+ 方法以及 DANet (dual attention network) 方法 (Fu 等, 2019) 进行了实验分析,并通过 F1-Score 评价指标以及气旋涡与反气旋涡的 AP 来对海洋涡旋检测的性能进行对比。从表2中可以看出,本文方法在 F1-Score 上均比 U-Net、EddyNet、FCN、DeepLabV3+ 方法取得了更好结果,在涡旋的检测精度上比 DANet 有着更好的性能。

表2 不同方法在 EddyNet-Data 数据集上的分割性能对比

Table 2 Comparison of segmentation performance of different methods on EddyNet-Data datasets

模型	AP		F1-Score
	气旋涡	反气旋涡	
FCN	0.229 4	0.216 0	0.870 8
EddyNet	<b>0.627 8</b>	<b>0.653 5</b>	0.892 3
DeepLabV3+	0.447 0	0.453 7	0.924 3
DANet	0.444 6	0.415 5	<b>0.936 5</b>
FA-U-Net(本文)	0.532 0	0.524 6	0.932 7

注:加粗字体为各列最优结果。

本文方法的 ROC 曲线如图5所示,其中类别0为非涡旋的类别,类别1为反气旋涡的类别,类别2为气旋涡的类别,通过 ROC 曲线图可以看出,本文方法在检测反气旋涡的性能更好。

不同方法在 EddyNet-Data 数据集上的检测结果如图6所示,由于海洋涡旋的小目标数量较多,FCN 方法最后一层使用上采样层的上采样率较大,导致

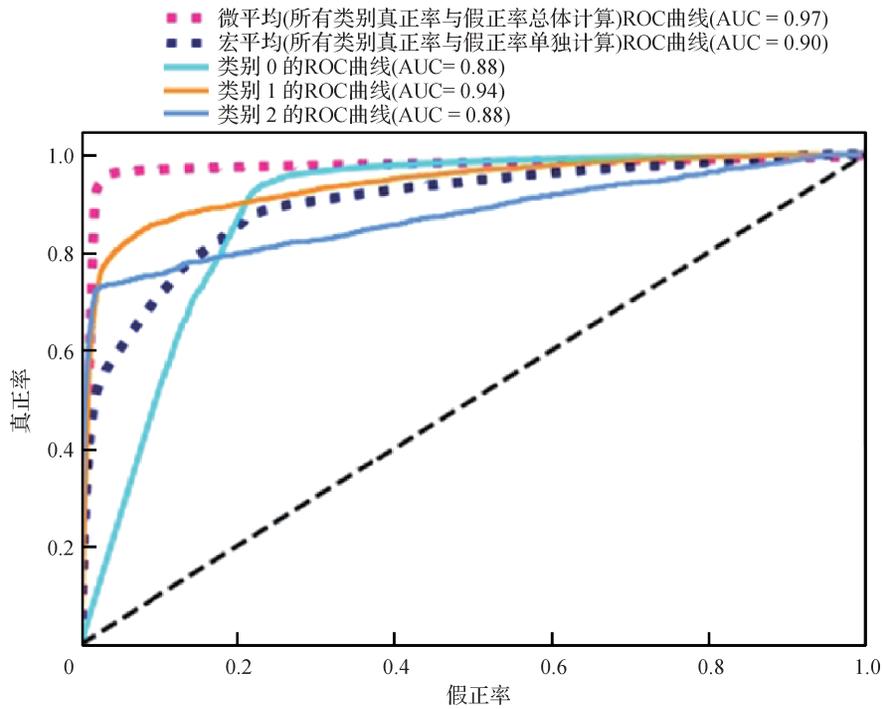


图5 本文方法的ROC曲线图  
Fig. 5 ROC plot of the method in this document

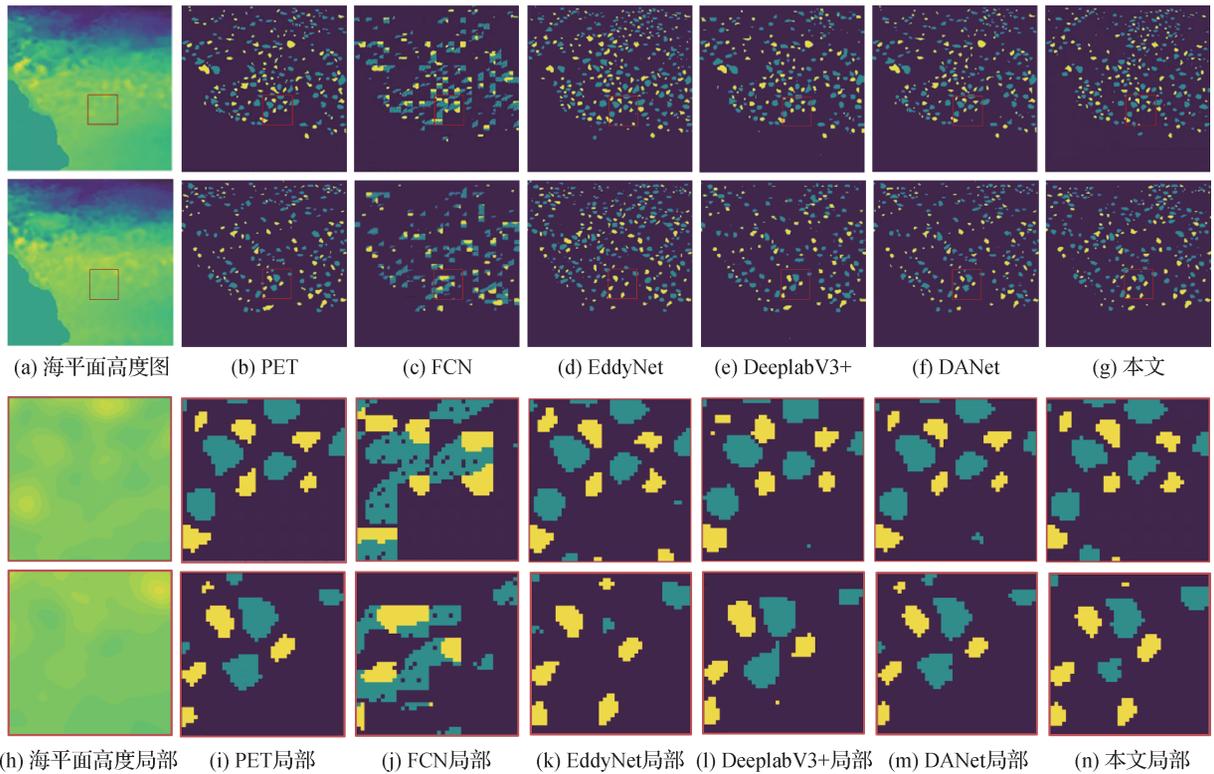


图6 不同方法在EddyNet-Data数据集上的检测结果

Fig. 6 The results of different methods in the EddyNet-data dataset ((a) sea surface height map; (b) PET; (c) FCN; (d) EddyNet; (e) DeeplabV3+; (f) DANet; (g) ours; (h) local map of sea level height; (i) local map of PET; (j) local map of FCN; (k) local map of EddyNet; (l) local map of DeeplabV3+; (m) local map of DANet; (n) local map of ours)

对小目标的预测结果较差。获取底层空间信息和轮廓信息是提升小目标检测精度的关键,但FCN模型在频繁采样过程中会导致这些底层特征的丢失,所以对较小的海洋涡旋的检测效果不理想,并且检测到的海洋涡旋轮廓与PET方法获取的真实数据差别较大。本文方法在利用多尺度特征的同时进一步融合上下文特征,预测较小尺度海洋涡旋的效果更好。同时上下文特征融合模块也增加了对底层特征的利用,即提取了更多的轮廓信息,如图6所示,本文方法预测的目标轮廓与真实目标轮廓更加接近。

不同方法检测的海洋涡旋数量如图7所示,对如图6所示的第1个区域海洋涡旋检测结果进行统计发现,本文方法与PET方法检测到的数量接近,由于EddyNet的误检率高,所检测到的海洋涡旋数量远超过PET方法。由于FCN方法检测小目标的效果不理想,所以检测到的数量要比真实数量少很多,同时较小尺度的海洋涡旋检测到的数量更少,较大尺度海洋涡旋的轮廓也与PET方法相差很大。由此可见,本文方法对于小目标的海洋涡旋检测具有较好的检测效果,同时海洋涡旋检测数量更接近PET方法,误检率更低。

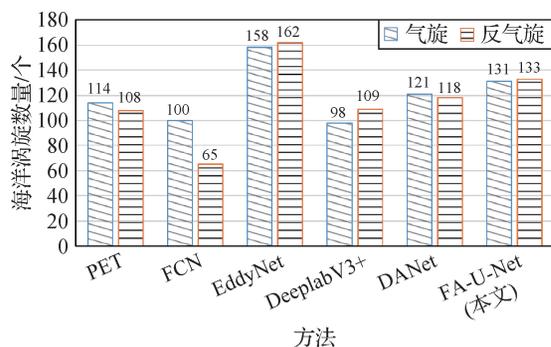


图7 不同方法检测的海洋涡旋数量

Fig. 7 The number of ocean vortices detected by different methods

## 4 结论

本文提出了一种基于U-Net网络改进的海洋涡旋小目标检测模型,该模型引入上下文融合模块对编码器输出的特征进行融合提取更多海洋涡旋的上下文信息,在利用底层特征的同时融入上下文信息进一步提高海洋涡旋的检测精度。同时,在上下文

特征融合模块前加入残差注意力模块,增强底层特征中海洋涡旋的轮廓以及高层语义特征。此外,在模型训练时对训练样本集采用数据扩充,一定程度上缓解了模型过拟合的问题。实验结果表明,本文模型与对比方法相比,对小尺度海洋涡旋目标检测效果更优,同时在轮廓以及数量上都与PET方法更加接近,因此在海洋涡旋的检测中具有更高的精确性以及合理性。与FCN方法相比,本文方法对海洋涡旋检测的轮廓和数量更接近PET方法,比EddyNet方法和DeeplabV3+方法的F1-Score更高,本文方法在海洋涡旋检测的AP比DANet方法更高。与EddyNet方法相比,在检测海洋涡旋中存在不足,产生这种问题的主要原因是由于比EddyNet进行更多的下采样,导致小目标的特征丢失,进而导致对海洋涡旋检测效果不理想。在未来工作中,为了更加明确不同海洋涡旋类别和目标,将对海洋涡旋同时进行目标检测与实例分割任务,使检测出的海洋涡旋不同的类别和目标更加明确。

## 参考文献 (References)

- Ali Sadarjoen I and Post F H. 2000. Detection, quantification, and tracking of vortices using streamline geometry. *Computers and Graphics*, 24 (3) : 333-341 [DOI: 10.1016/S0097-8493 (00) 00029-7]
- Cao Y, Xu J R, Lin S, Wei F Y and Hu H. 2019. Gcnnet: non-local networks meet squeeze-excitation networks and beyond//Proceedings of 2019 IEEE/CVF International Conference on Computer Vision Workshops. Seoul, Korea (South) : IEEE: 1971-1980 [DOI: 10.1109/ICCVW.2019.00246]
- Chaigneau A, Gizolme A and Grados C. 2008. Mesoscale eddies off Peru in altimeter records: identification algorithms and eddy spatio-temporal patterns. *Progress in Oceanography*, 79 (2-4) : 106-119 [DOI: 10.1016/j.pocean.2008.10.013]
- Chelton D B, Schlax M G, Samelson R M and De Szoeke R A. 2007. Global observations of large oceanic eddies. *Geophysical Research Letters*, 34(15) : #L15606 [DOI: 10.1029/2007GL030812]
- Chelton D B, Schlax M G and Samelson R M. 2011. Global observations of nonlinear mesoscale eddies. *Progress in Oceanography*, 91 (2) : 167-216 [DOI: 10.1016/j.pocean.2011.01.002]
- Chen G X, Hou Y J and Chu X Q. 2011. Mesoscale eddies in the South China Sea: mean properties, spatiotemporal variability, and impact on thermohaline structure. *Journal of Geophysical Research: Oceans*, 116 (C6) : #C06018 [DOI: 10.1029/2010jc006716]

- Chen L C, Papandreou G, Schroff F and Adam H. 2017. Rethinking atrous convolution for semantic image segmentation [EB/OL]. [2022-10-10]. <https://arxiv.org/pdf/1706.05587.pdf>
- Chen L C, Zhu Y K, Papandreou G, Schroff F and Adam H. 2018. Encoder-decoder with atrous separable convolution for semantic image segmentation//Proceedings of the 15th European Conference on Computer Vision. Berlin, Germany: Springer: 833-851 [DOI: 10.1007/978-3-030-01234-2\_49]
- Doglioli A M, Blanke B, Speich S and Lapeyre G. 2007. Tracking coherent structures in a regional ocean model with wavelet analysis: application to Cape Basin eddies. *Journal of Geophysical Research: Oceans*, 112(C5): #C05043 [DOI: 10.1029/2006JC003952]
- Dong Z Y, Du Z H, Wu S S, Li Y D, Zhang F, Liu R Y. 2022. An automatic marine mesoscale eddy detection model based on improved U-Net network. *Acta Oceanologica Sinica*, 44(2): 123-131 (董子意, 杜震洪, 吴森森, 李亚东, 张丰, 刘仁义. 2022. 基于改进 U-Net 网络的海洋中尺度涡自动检测模型. *海洋学报*, 44(2): 123-131) [DOI: 10.12284/hyxb2022038]
- Du Y L, Liu Q Q, Wang L L, Xu X, Wei Q M and Song W. 2022. Multi-scale rotating anchor mechanism based automatic detection of ocean mesoscale eddy. *Journal of Image and Graphics*, 27(10): 3092-3101 (杜艳玲, 刘倩倩, 王丽丽, 徐鑫, 魏泉苗, 宋巍. 2022. 融合多尺度旋转锚机制的海洋中尺度涡自动检测. *中国图象图形学报*, 27(10): 3092-3101) [DOI: 10.11834/jig.210286]
- Faghmous J H, Le M, Uluoyol M, Kumar V and Chatterjee S. 2013. A parameter-free spatio-temporal pattern mining model to catalog global ocean dynamics//Proceedings of the 13th IEEE International Conference on Data Mining. Dallas, USA: IEEE: 151-160 [DOI: 10.1109/ICDM.2013.162]
- Faghmous J H, Styles L, Mithal V, Boriah S, Liess S, Kumar V, Vikebø F and Dos Santos Mesquita M. 2012. EddyScan: a physically consistent ocean eddy monitoring application//Proceedings of 2012 Conference on Intelligent Data Understanding. Boulder, USA: IEEE: 96-103 [DOI: 10.1109/CIDU.2012.6382189]
- Fan Z L, Zhong G Q, Wei H X and Li H T. 2020. EDNet: a mesoscale eddy detection network with multi-modal data//Proceedings of 2020 International Joint Conference on Neural Networks. Glasgow, UK: IEEE: 1-7 [DOI: 10.1109/IJCNN48605.2020.9206613]
- Fu C Y, Liu W, Ranga A, Tyagi A and Berg A C. 2017. DSSD: deconvolutional single shot detector [EB/OL]. [2022-10-02]. <https://arxiv.org/pdf/1701.06659.pdf>
- Fu J, Liu J, Tian H J, Li Y, Bao Y J, Fang Z W and Lu H Q. 2019. Dual attention network for scene segmentation//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, USA: IEEE: 3141-3149 [DOI: 10.1109/CVPR.2019.00326]
- He K M, Gkioxari G, Dollar P and Girshick R. 2017. Mask R-CNN//Proceedings of 2017 IEEE International Conference on Computer Vision. Venice, Italy: IEEE: 2980-2988 [DOI: 10.1109/ICCV.2017.322]
- Henson S A and Thomas A C. 2008. A census of oceanic anticyclonic eddies in the Gulf of Alaska. *Deep Sea Research Part I: Oceanographic Research Papers*, 55(2): 163-176 [DOI: 10.1016/j.dsr.2007.11.005]
- Jeong J, Park H and Kwak N. 2017. Enhancement of SSD by concatenating feature maps for object detection [EB/OL]. [2022-10-02]. <https://arxiv.org/pdf/1705.09587.pdf>
- Jia K X, Ma Z H, Zhu R and Li Y G. 2022. Attention-mechanism-based light single shot multiBox detector modelling improvement for small object detection on the sea surface. *Journal of Image and Graphics*, 27(4): 1161-1175 (贾可心, 马正华, 朱蓉, 李永刚. 2022. 注意力机制改进轻量 SSD 模型的海面小目标检测. *中国图象图形学报*, 27(4): 1161-1175) [DOI: 10.11834/jig.200517]
- Iguensat R, Sun M, Fablet R, Tandeo P, Mason E and Chen G. 2018. EddyNet: a deep neural network for pixel-wise classification of oceanic eddies//Proceedings of IGARSS 2018 IEEE International Geoscience and Remote Sensing Symposium. Valencia, Spain: IEEE: 1764-1767 [DOI: 10.1109/IGARSS.2018.8518411]
- Li X, Zhong Z S, Wu J L, Yang Y B, Lin Z C and Liu H. 2019. Expectation-maximization attention networks for semantic segmentation//Proceedings of 2019 IEEE/CVF International Conference on Computer Vision. Seoul, Korea (South): IEEE: 9166-9175 [DOI: 10.1109/ICCV.2019.00926]
- Lim J S, Astrid M, Yoon H J and Lee S I. 2021. Small object detection using context and attention//Proceedings of 2021 International Conference on Artificial Intelligence in Information and Communication. Jeju Island, Korea (South): IEEE: 181-186 [DOI: 10.1109/ICAIIIC51459.2021.9415217]
- Lin T Y, Dollar P, Girshick R, He K M, Hariharan B and Belongie S. 2017. Feature pyramid networks for object detection//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, USA: IEEE: 936-944 [DOI: 10.1109/CVPR.2017.106]
- Liu Z M, Gao G Y, Sun L and Fang Z Y. 2021. HRDNet: high-resolution detection network for small objects//Proceedings of 2021 IEEE International Conference on Multimedia and Expo. Shenzhen, China: IEEE: 1-6 [DOI: 10.1109/ICME51207.2021.9428241]
- Long J, Shelhamer E and Darrell T. 2015. Fully convolutional networks for semantic segmentation//Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, USA: IEEE: 3431-3440 [DOI: 10.1109/CVPR.2015.7298965]
- Mason E, Pascual A and McWilliams J C. 2014. A new sea surface height-based code for oceanic mesoscale eddy tracking. *Journal of Atmospheric and Oceanic Technology*, 31(5): 1181-1188 [DOI: 10.1175/JTECH-D-14-00019.1]
- McWilliams J C. 1984. The emergence of isolated, coherent vortices in

- turbulent flow. *AIP Conference Proceedings*, 106 (1) : 205-221 [DOI: 10.1063/1.34273]
- Nencioli F, Dong C M, Dickey T, Washburn L and McWilliams J C. 2010. A vector geometry - based eddy detection algorithm and its application to a high-resolution numerical model product and high-frequency radar surface velocities in the Southern California Bight. *Journal of Atmospheric and Oceanic Technology*, 27(3) : 564-579 [DOI: 10.1175/2009JTECH0725.1]
- Ronneberger O, Fischer P and Brox T. 2015. U-Net: convolutional networks for biomedical image segmentation//*Proceedings of the 18th International Conference on Medical Image Computing and Computer-Assisted Intervention*. Munich, Germany: Springer: 234-241 [DOI: 10.1007/978-3-319-24574-4\_28]
- Shen B, Chen Y, Yang C and Liu B W. 2020. Computer vision detection and analysis of mesoscale eddies in marine science. *Frontiers of Data and Computing*, 2(6) : 30-41 (沈飙, 陈扬, 杨琛, 刘博文. 2020. 海洋科学中尺度涡的计算机视觉检测和分析方法. *数据与计算发展前沿*, 2(6) : 30-41) [DOI: 10.11871/jfdic.issn.2096-742X.2020.06.004]
- Wang F, Jiang M Q, Qian C, Yang S, Li C, Zhang H G, Wang X G and Tang X O. 2017. Residual attention network for image classification//*Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu, USA: IEEE: 6450-6458 [DOI: 10.1109/CVPR.2017.683]
- Williams S, Hecht M, Petersen M, Strelitz R, Maltrud M, Ahrens J, Hlawitschka M and Hamann B. 2011. Visualization and analysis of eddies in a global ocean simulation. *Computer Graphics Forum*, 30(3) : 991-1000 [DOI: 10.1111/j.1467-8659.2011.01948.x]
- Woo S, Park J, Lee J Y and Kweon I S. 2018. CBAM: convolutional block attention module//*Proceedings of the 15th European Conference on Computer Vision*. Munich, Germany: Springer: 3-19 [DOI: 10.1007/978-3-030-01234-2\_1]
- Xu G J, Cheng C, Yang W X, Xie W H, Kong L M, Hang R L, Ma F R, Dong C M and Yang J S. 2019. Oceanic eddy identification using an AI scheme. *Remote Sensing*, 11(11) : #1349 [DOI: 10.3390/rs11111349]
- Zhao H S, Shi J P, Qi X J, Wang X G and Jia J Y. 2017. Pyramid scene parsing network//*Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu, USA: IEEE: 6230-6239 [DOI: 10.1109/CVPR.2017.660]

### 作者简介

- 杜艳玲,女,副教授,硕士生导师,主要研究方向为海洋中尺度涡旋自动检测与轨迹追踪。E-mail: yldu@shou.edu.cn
- 宋巍,通信作者,女,教授,主要研究方向为机器视觉、图像/视频处理和海洋大数据分析。E-mail: wsong@shou.edu.cn
- 陈刚,男,高级工程师,主要研究方向为海洋信息化和深度学习。E-mail: cg232@126.com
- 吴天宇,男,硕士研究生,主要研究方向为语义分割和深度学习。E-mail: 1392508773@qq.com
- 陈括,男,工程师,主要研究方向为海洋遥感处理与分析 and 海洋信息化。E-mail: chenkuo@ecs.mnr.gov.cn