

# 基于深度区域提取网络对中尺度涡的检测

陈扬,杨琛,刘博文

中国海洋大学

## 摘要

中尺度涡是海洋科学领域一个重要的研究课题.通常中尺度涡分为两种:气旋涡和反气旋涡.中尺度涡的检测是中尺度涡研究中非常重要的研究方向,有着非常重要的科学意义和渔业、军事价值,目前我国在涡旋检测领域已经做了很多研究工作。

近年以来,人工智能领域中的深度神经网络高速发展,其被广泛应用于模式识别和计算机视觉等许多实际问题当中,并取得了很好的效果。本文将深度学习中的目标检测算法应用于中尺度涡检测,相比较于传统的中尺度涡检测方法只能检测中尺度涡的位置和大小,其能更好地利用多模态信息进行检测、分类和实例分割.本文基于Mask-RCNN 算法,提出一种结合多模态(SSH,SST)卫星遥感图像数据的目标检测算法,对海洋中的中尺度涡进行识别、分类和分割。

关键字:中尺度涡,目标检测,实例分割,多模态,Mask-RCNN

## abstract:

## 正文

在海洋领域,中尺度涡是指在海洋中半径有10-100千米、持续时间为2-10个月的涡旋。相比于海洋中肉眼可见的涡旋,中尺度涡有着更大的半径以及更长的持续时间,但是其规模远小于海洋中的持续存在的大环流.中尺度涡携带极大的动能,涡旋内海水运动速度非常快,相对洋流平均流速而言,要快几倍到一个量级。径向涡旋涉及几十米到上百米,甚至上千米的深海,海洋深层的营养盐和冷水会被径向涡旋带到海洋表面。中尺度涡还能够将海表暖水压到较深的海洋中,在全球海洋物质、能量和热量等的运输和分配中起着不可忽视的作用。因此,海洋涡旋的研究具有非常重要的科学意义和渔业、军事价值。

近年来,深度神经网络在图像分类、人脸识别、文本识别等方面取得了非凡的成就,特别是在图像分类领域,一些新型的技术比如随机丢弃技术、正则化、批归一化等,使得深度卷积神经网络模型在复杂的计算机视觉问题中展现出高效的性能。而在目标检测和分割领域,深度网络模型(R-CNN, Faster R-CNN, Mask R-CNN 等)的很多改进版本,识别准确率达到了很高的水准,应用到了生产生活的很多领域。

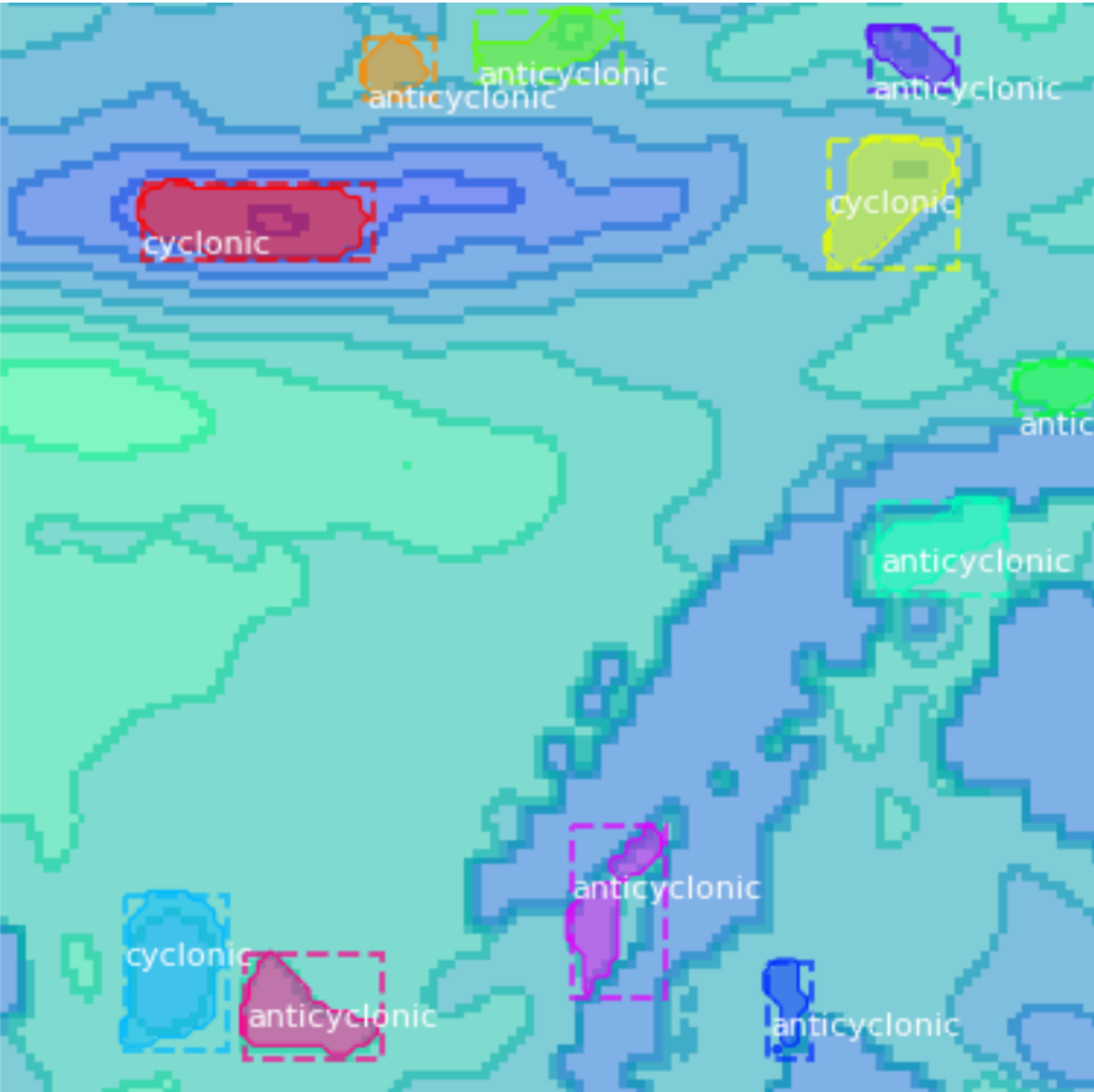
在卫星资料、浮标资料和高分辨模式的辅助下,中尺度涡检测已成为物理海洋领域的研究热点。但从卫星遥感数据中识别中尺度涡的主要方法还是依赖专家分析,这种方法劳动强度大,很难满足快速、准确检测的需求,对于现在的数据量而言更难以做到全面的检测和分析。另外,很多算法利用海表高度、温度等数据,基于流场几何特征、边缘检测以及拉格朗日随机模型进行中尺度涡检测,这些算法误检率较高。还有一些基于图像和深度神经网络的方法,将海洋卫星遥感数据转换成图像作为网络输入,这一类方法往往会损失卫星遥感数据的细节信息,且已有的深度学习模型相对简单,因此检测结果也不够精确。

为解决上述问题，我们将探索使用人工智能领域中的深度学习方法，对涡旋进行检测和分析，也就是利用深度卷积网络和多模态海洋卫星遥感数据，实现对中尺度涡的准确检测。我们将使用多模态数据融合，对卫星遥感数据，如海洋表面高度、温度及流速数据进行融合学习，摒弃以往将单一类型数据转换成图像数据再进行检测分析的策略。同时，借鉴经典的目标检测、语义分割网络的优点，设计区域提取网络的用于中尺度涡检测。其中，残差网络部分学习中尺度涡特征表示，区域生成网络生成含有中尺度涡的区域并提取特征，头网络部分负责中尺度涡类别和范围预测，网络能够通过端到端的方式进行学习。

本文主要从设计应用于中尺度涡检测的深度神经网络入手，探索神经网络在提取卫星遥感数据特征中的有效性。本文通过建立特定中尺度涡检测数据集，训练深度神经网络模型，并对比本文提出的基于深度神经网络的方法和其同类方法之间的有效性以及优缺点。

## 相关工作

### 中尺度涡数据集来源



目前具有全球覆盖率可用于海洋中尺度涡研究的观测资料有:1. 卫星高度计(altimetry)的海面高度异常场(sea surface height anomaly, SSHA)网格化资料, 从 1992 年至今, 已经有 23 年全球覆盖的资料;2. 海表温度资料(sea surface temperature, SST)从 20 世纪 80 年代初至今, 已有接近 35 年的全球覆盖资料;3. 水色卫星遥感资料;4.星载合成孔径雷达数据(synthetic aperture radar, SAR);5.全球海洋浮标资料(global drifter program, GDP)。

## eddynscan

EddyScan 解决了上述中尺度涡检测过程中存在的挑战。首先, 为了解决中尺度涡识别过程中的叠加挑战, 作者在-100cm 到+100cm 的区间进行 1cm 间隔的重复计算。在每个阈值  $i$ , 作者识别出数量具有至少为  $i$  的海平面高度异常的所有连通组件。然后, 算法去除属于所识别的连通分量的所有像素, 并且以  $i$  递增。为了识别反气旋涡, 作者在-100cm 初始化并以 1cm 步长增加到+100cm。相反, 通过将  $i$  从+100cm 减小到-100cm 来实现对气旋涡的检测。通过这种方式, 作者确定了涡流最大可能的闭合轮廓。CH11[33]提出了逐步阈值法, CH11 虽然也测量了亚厘米阈值增量, 但没有显著提升模型准确度。其次, 为了解决纬度原因造成的涡旋大小不同, 作者使用基于理论[34-36]的二次函数和经验研究[33, 37]来限制基于纬度的合理涡旋半径。最后, 如果 涡旋在纬度上大于预期, 则有可能错误地将两个或更多涡旋合并在一起。

对于大于正常规模的涡旋, 用凸包函数<sup>38</sup>来确定包含涡旋的所有网格数据的最小凸集的大小。如果凸包的面积远大于连接区域的面积, 则可能是多个涡旋合并的情况, 并且连接的部分未标记为涡旋, 丢弃连接的区域, 它将保留在稍后要检查的网格数据中, 增加了再次检测的几率, 这样能够保证不漏检, 运行流程见算法 1。

## 构建深度学习数据集

众所周知, 训练神经网络需要一个可靠的数据集, 而人工标注如此大量的数据也不太可能实现。本文采取的方法是使用EddyScan 对卫星雷达高度计数据进行中尺度涡检测, 并将此结果作为数据标注, 然后对本文提出的算法进行有效性检测。目前用于涡旋检测的多数研究工作中, 普遍采用高度参数化的连通分量算法, 使用专家过滤数据, 在重现性和扩展性方面有一定的不足。在 EddyScan 中, 改进了最优秀的连接组件涡旋监控算法, 能够实现全局涡流跟踪。这项工作有三个主要贡献:第一, 没有进行数据预处理, 最大限度地减小数据中重要信号的损失风险;第二, 基于理论和实际研究对涡旋采用物理一致的凸性要求, 提高方法的准确度和计算复杂度。最后, 作者准确地分离了空间上邻近的涡旋, 这也是同期方法无法实现的。

因此运行 EddyScan 算法, 能够以相对较高的准确率检测到涡旋的位置, 这样就能够顺利的进行涡旋的标注。另外, 本文使用的卫星遥感数据来自于哥白尼海洋环境监测服务中心。哥白尼海洋环境监测服务中心的全球海洋物理再分析数据包(GLOBAL OCEAN PHYSICS REANALYSIS GLORYS2V4)包含海表温度, 盐度, 洋流, 高度以及海冰参数的日常平均值, 分辨率为 0.25 度, 覆盖时间为 1993 年~2015 年。由于其精度高, 数据类型全面, 所以本数据集的建立使用了全球海洋物理再分析数据包中能够辅助中尺度涡检测的温度、高度以及流速数据。

1)我们下载了GLORYS2V4中2000年01月16日到2009年12月16日共计十年的温度、高度以及流速数据, 其中这三种数据的维度分别为 $681 \times 1440 \times 120$ , 681 为纬度的维度, 1440 为经度的维度, 120 表示数据来自连续的 120 个月。需要注意的是, 海水流速数据包含 2 个方向, 分别是东向海水的流速以及北向海水的流速, 这可以理解为海洋中某一个点的速度矢量分解到东西和南北这两个方向。

2)由于中尺度涡的生命周期短则几天长可达一年左右,为了让取到数据差异更大,多样性更强,本文选择了120个月中间隔为3个月的40个月的数据。

3)为了保证神经网络的训练速度,本文使用了区域大小为 $128 \times 128$ 的遥感数据。数据采集的方式为随机不重复的在每个月的各类数据中取13次,并保证温度、高度以及流速的对应位置相同。此外为了保证训练效果,本文进行了多次随机取值,最终使用的数据中有超过90%的数据含有中尺度涡(陆地以及某些海域没有中尺度涡)。之所以选择没有中尺度涡的区域,是为了提升算法的泛化能力。

4)在标签(label)的设计中,气旋涡标注为1.0,反气旋涡标注为-1.0,没有涡的区域标注为0.0。这样能将三种情况的差异尽可能地变大,在识别的时候尽可能降低学习的难度。

5)神经网络的训练过程中,为了保证训练效果,也就是高准确率,同时提升泛化能力,本文选择按照3:7的比例将所有数据分为测试数据和训练数据。

## 多模态数据融合

卫星遥感数据中的海洋表面温度、高度以及流速都会从不同的角度描述中尺度涡的特征,所以理想的情况下,使用尽可能多的数据类型进行中尺度涡检测是最有效的。多模态数据的融合不仅有利于卷积网络进行中尺度涡特征的提取,更有利于进行中尺度涡的预测。

数据融合也存在一些难点,由于不同类型数据表示的海洋特性的不同,所以不同数据的单位也各不相同,例如,海洋表面高度的单位是米(meters),海洋表面温度的单位是卡式温标(Kelvin),流速的单位是米每秒(meters per second),这样就导致神经网络计算过程中存在很多问题。首先,不同数据的数值差异很大,在神经网络的前向传播中,会造成前向传播过程中的数据计算异常。其次,近似为0的数值频繁在很多数据中出现,会造成预测结果难以学习。最后,不同类型的数据对中尺度涡检测起到的作用是不同的,换句话说,数据需要按比重进行缩放。

为了解决上述问题,并提升深度模型的学习效率,我们采用数据预处理进行数据整合。在数据预处理过程中,最重要的是解决数据分布不均匀的问题,本研究尝试使用不同的归一化方式进行处理。一种归一化方式是min-max标准化[60](Min-max normalization),这种方式是通过线性计算,把原始的遥感数据进行线性缩放,最终落到[a, b]之间。首先需要找到原始数据分布中的最大值(max)和最小值(min),然后计算缩放系数,

$$k = \frac{b - a}{\max - \min} \circ \quad (1)$$

这里我们一般将区间设置为[0,1],因此计算方式也很简单:

$$x^* = \frac{x - \min}{\max - \min} \quad (2)$$

为了找到最好的中尺度涡表示数据,便于神经网络提取特征,我们使用多层卷积网络,并利用通过端到端的训练方式进行权重的学习。整体的计算方式可以简单表示为:

$$x^* = ([H, T, V] \times x_1 + b_1) \times x_2 + b_2 \circ \quad (3)$$

在图3-3中,H表示海洋表面高度数据,T表示海洋表面温度数据,V表示

海洋表面流速数据(双通道二维数据)。

# MASK-RCNN

Mask-RCNN 的输入是经过数据融合网络得到的多模态融合数据

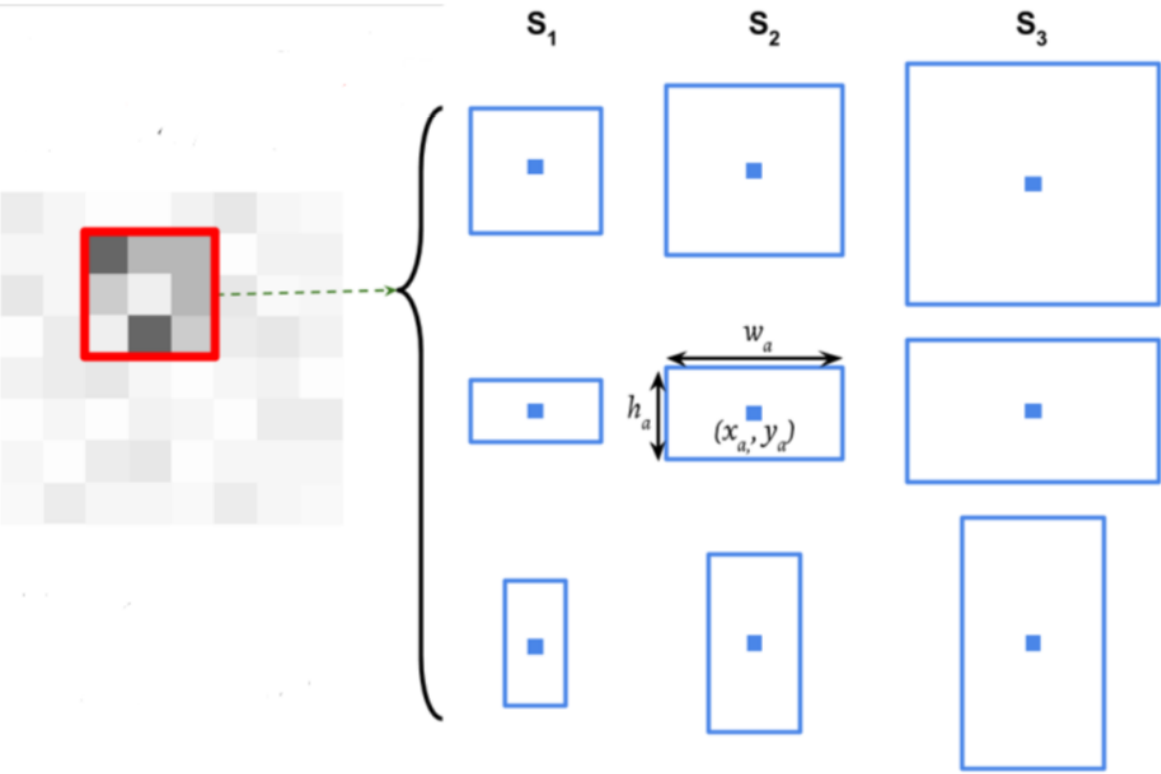
经过深度残差网络得到高维特征图

通过 RPN 网络计算出ROI 区域

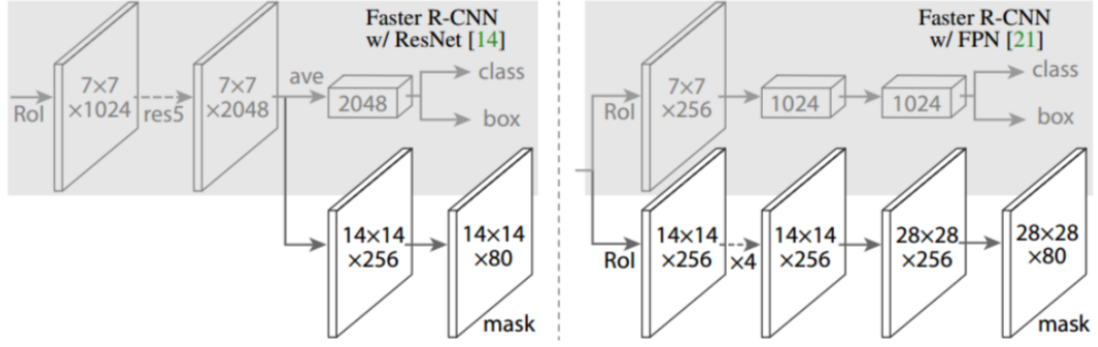
ROI经过区域生成网络得到特征向量

特征向量经过一个头网络得到中尺度涡的定位,类别,

在对其中另一个头网络进行反卷积,得到中尺度涡的实例分割图



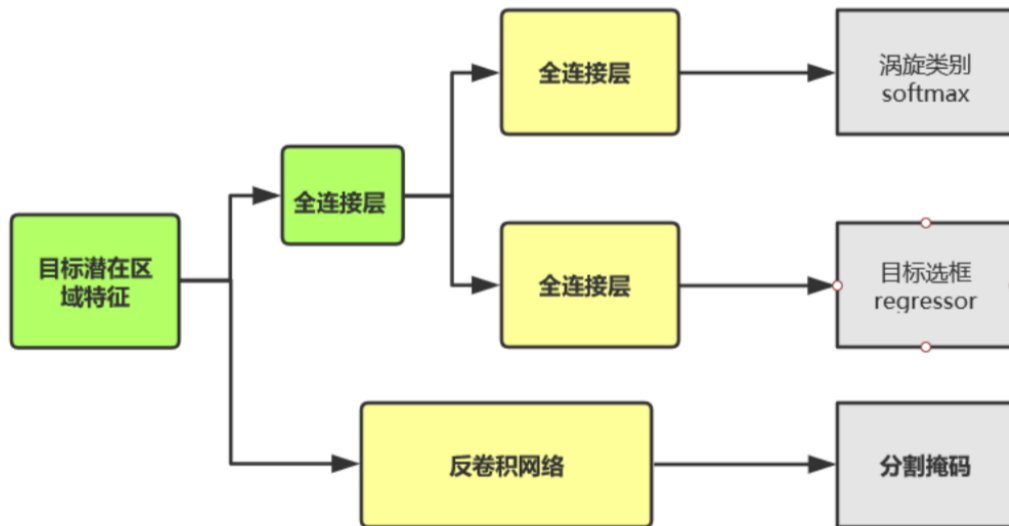
每个滑动窗口产生 9 个锚点区域，然后进行分类和回归计算。每个锚点区域 需要回归的值有四个:可能含有中尺度涡区域的中心坐标，以及区域长和宽。分 类任务主要是对区域内是否含有中尺度涡进行判 别，也就是一个二分类任务，含 有中尺度涡为 1，不含中尺度涡只是背景为 0。含有中尺度涡的区域将会在下 一节中进行第二次回归，对位置进行校正，并判定最终的中尺度涡类别。



掩码分支是一个卷积网络，取 ROI 分类器选择的正区域为输入，并生成它们的掩码。其生成的掩码是低分辨率的：28x28 像素。但它们是由浮点数表示的软掩码，相对于二进制掩码有更多的细节。掩码的小尺寸属性有助于保持掩码分支网络的轻量性。在训练过程中，我们将真实的掩码缩小为 28x28 来计算损失函数，在推断过程中，我们将预测的掩码放大为 ROI 边框的尺寸以给出最终的掩码结果，每个目标有一个掩码。

## 深度区域生成网络

顾名思义就是找到多个可能含有中尺度涡的区域，并输出含有中尺度涡概率比较高的区域，然后再针对每个区域进行特征提取、计算，最后进行中尺度涡类别和范围预测。深度区域生成网络也可以被看成一个初始目标检测网络，它的优势是利用更少的计算量对卫星遥感数据的所有位置进行扫描，再加上人工设计的不同形状、大小的锚点，更能准确定位中尺度涡。下面我们将介绍区域生成网络的具体工作流程。



$$L = L_{cls} + L_{box} + L_{mask} \quad (4)$$

上式中， $L_{cls}$  表示bounding box的分类损失值， $L_{box}$  表示bounding box的回归损失值， $L_{mask}$  表示mask部分的损失值

$$L_{cls} = -\log p_u \quad (5)$$

$$L_{\text{mask}}(\text{Cls\_k}) = \text{Sigmoid}(\text{Cls\_k}) \quad (6)$$

$L_{\text{mask}}$ 只定义所预测的类别的掩码损失，不计算其他类别的掩码损失。定义  $L_{\text{mask}}$  的方式使得我们的网络在生成单个类别的掩码时不会受类别竞争影响，解耦了掩码和类别预测。

$L_{\text{mask}}(\text{Cls\_k}) = \text{Sigmoid}(\text{Cls\_k})$ ，也就是平均二值交叉熵(average binary cross-entropy)损失，通过逐像素计算得到。

$$L_{\text{box}} = \sum_i^n g(t_i^u - v_i) \quad (7)$$

$$g(x) = \begin{cases} 0.5x^2 & |x| < 1 \\ |x| - 0.5 & \text{其他情况} \end{cases} \quad (8)$$

## 实验

---

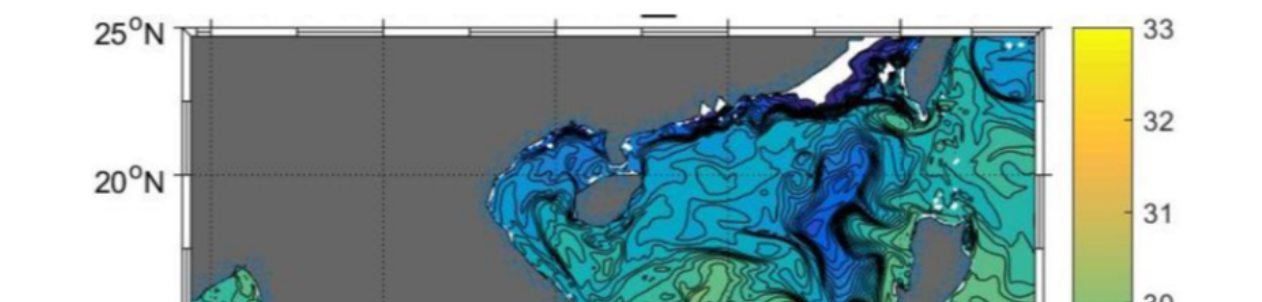


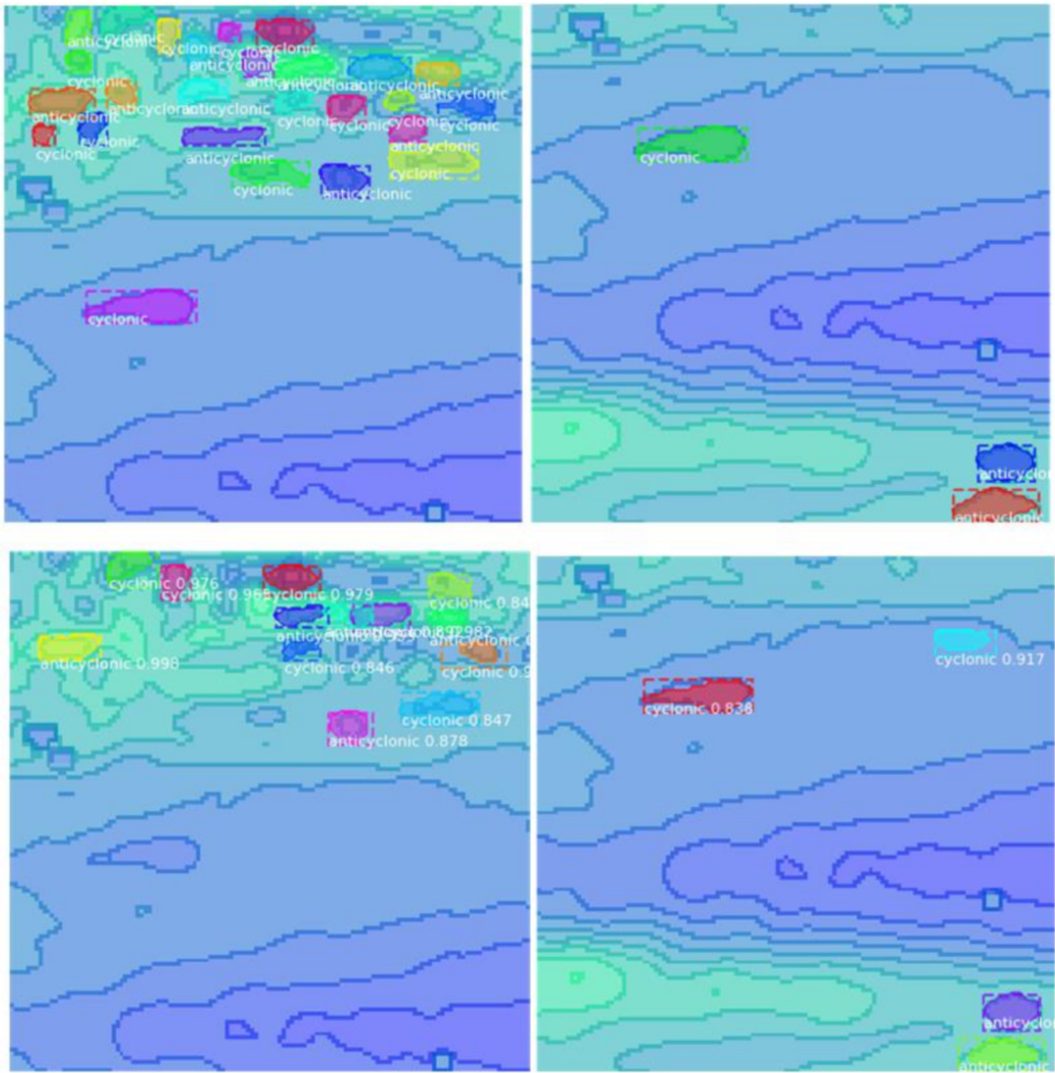


表 4-1 实验结果

锚点规模	学习率 (头网络)	学习率 (整体训练)	mAP (%)
(8,32,128)	0.001	0.001	45.50
(4,16,64)	0.001	0.001	53.89
(8,32,128)	0.001	0.0001	52.62
<b>(4,16,64)</b>	<b>0.001</b>	<b>0.0001</b>	<b>56.35</b>

SSH

由上述检测结果可得，在锚点规模为 (4,16,64)，训练头网络的初始学习率为 0.001，整体训练的初始学习率为 0.0001 时，mAP 最高为 56.35。



## 多模态数据融合实验

通过上述实验结果可知，当 $\alpha$ 、 $\beta$ 、 $\gamma$ 的比例为 5:3:2 时，实验结果最优。由于 实验硬件资源有限，没有 尝试更多的不同比例组合，但从现有的实验设置可以看 到，多模态数据融合的效果普遍好于仅使用单一 数据类型进行实验。通过实验结 果分析，多种类型的数据融合涵盖的中尺度涡特征信息要多于单一类型 数据。

功能实现方面，DeepEddy 仅能实现涡旋图片分类，而 Eddynet 在图片分类 的基础上实现了涡旋的分割，本方法能基于图片和遥感数据实现涡旋分类、分割 以及位置预测。

具体实现方面，DeepEddy 先将 SAR 图像切分成固定大小的小图像，然后人 工标注小图像是否为涡旋 图像，最后进行算法训练，分类准确率在 97%左右，但 是 DeepEddy 仅实现了 SAR 切分图像的分类， 判断图像内是否含有涡旋，因此 DeepEddy 不如本文提出的算法创新性高。另外，本文提出的基于深度 区域 提取网络的中尺度涡检测算法能够在全 球范围内的遥感数据上对中尺度涡进行 查找定位，并进行 分类。因此，本文提出的算法相对于 DeepEddy 而言，实现 的功能多，易用性强，创新性也更高。

EddyNet 首先将海洋表面高度数据转化成为图像，然后在图像的基础上进行 像素级别的标注，最后使 用类 U-Net 网络进行目标分割。最终实验结果对比如表

使用多模态融合数据进行训练时检测结 果最优，不但明显好于 EddyNet，也好于基于图像数据集训练 的结果。因此可知， 使用图像进行模型训练时，本方法要好于 EddyNet;使用本文提出的模型，用 多模 态融合数据进行训练的检测结果好于使用图像数据集进行训练的结果。因此， 本节实验不但证明本模型 有效性优于 EddyNet，也证明使用多模态融合数据训练 模型能够显著提升模型准确度。

## 总结

1)提出了基于深度区域提取网络的涡旋检测模型。在功能方面，实现了卫星 遥感数据上的涡旋分割，包 含了分类，精细定位。另外，模型本身也具有 一定的创新性，针对涡旋检测的复杂程度对模型的深度进 行设计。

2)为了验证涡旋检测模型的有效性，本文提出了涡旋检测数据集。数据集 制作采用了传统非深度学习方 法，对遥感数据中每个数据点的类别进行 了精准标注;然后在大型数据集 ImageNet 上进行预训练，并 在涡旋检测 数据集上进行训练，得到了优秀的检测精度。

3)本文使用了多模态数据融合。在涡旋检测的过程中使用到了海洋表面 高度、海洋表面温度，以及海水 流速数据。经过对比实验表明，融合数据 在涡旋检测问题上要优于单一类型的数据，并且通过实验找到 了数据融 合的最佳策略。