

马蕴涵, 陈晓东, 赵观辉, 等. 基于深度学习的船基数字图像处理及海冰密集度识别研究[J]. 海洋学报, 2025, 47(3): 118–128, doi:10.12284/hyxb2025007

Ma Yunhan, Chen Xiaodong, Zhao Guanhui, et al. Research on ship based digital image processing and sea ice concentration recognition based on deep learning[J]. Haiyang Xuebao, 2025, 47(3): 118–128, doi:10.12284/hyxb2025007

基于深度学习的船基数字图像处理 及海冰密集度识别研究

马蕴涵¹, 陈晓东^{1*}, 赵观辉^{2,3}, 季顺迎¹, 杨海天¹

(1. 大连理工大学 工业装备结构分析优化与 CAE 软件全国重点实验室, 辽宁 大连 116023; 2. 浙江大学 计算机科学与技术学院, 浙江 杭州 310027; 3. 中国舰船研究设计中心, 湖北 武汉 430064)

摘要: 海冰是极地海域的典型环境特征, 对船载视频图像进行像素级分类可获取高分辨率的海冰信息。由于极地场景中的光照条件与海冰形态较为复杂, 传统计算机图形学方法的泛化性难以满足海冰要素的智能识别需求。因此, 本文采用基于 DeeplabV3+语义分割网络结构的深度学习对极地场景中海冰要素进行识别。将“雪龙”号科考船在冰区航行中的实测海冰图像制作为数据集, 并对深度学习模型进行训练与验证。根据海冰要素的识别需求与走航观测视频图像特点, 将像素信息划分为海冰、天空、海水与船体 4 种语义类别。根据训练集中的图像信息与语义信息间关联构建深度学习模型, 并通过所训练模型对验证集或其他图像中像素点的语义信息进行预测, 从而实现海冰信息的自动识别。为了研究该方法的鲁棒性进一步分析了海冰密集度、光照条件以及海冰类型对识别结果的影响。此外, 研究了数据集规模与迭代次数对识别精度的影响。图像识别结果显示, 4 类语义信息识别结果的平均交并比高于 95%。这表明深度学习能够在极地复杂的环境中较为准确地获取各类要素分类信息。

关键词: 海冰; DeeplabV3+; 语义分割; 图像识别; 深度学习

中图分类号: P714

文献标志码: A

文章编号: 0253-4193(2025)03-0118-11

1 引言

作为高纬度海洋的典型环境特征, 海冰在气-冰-海热量交互过程中的作用使其成为全球气候模式中的重要组成部分。海冰监测是深入研究海冰生消规律与防治海冰灾害发生的必要手段, 为有效监测海冰以获取南北极海冰分布特性, 通常采用卫星遥感、雷达与走航观测等方式^[1-2]。其中, 卫星遥感图片的覆盖区域较大, 更适合对大面积海域的海冰总体冰情信息进行提取, 但由于两极地区无法采用同步卫星进行观

测, 且所得到的海冰卫星图像采样频率较低。采用船载摄像机对海冰冰情进行图像采集, 所提取的海冰信息空间分辨率与频率均优于卫星遥感图像^[3]。但视频图像中的光线条件与冰缘线更为复杂, 对图像识别的算法的要求也更高。

基于船舶冰区走航视频图像对海冰形态、分布状况等进行实时分析与及时预警可有效保障航行安全^[4]。海冰全球气候变化模型的重要输入参数, 海冰的体量与其密集度 (Sea Ice Concentration, SIC) 密切相关^[5-7]。海冰密集度在船舶路径规划、海上天气预报、海洋灾

收稿日期: 2024-08-20; 修订日期: 2024-12-20。

基金项目: 工信部极地船舶专项 (CBG2N21-2-3) 和国家自然科学基金 (42176241, 52101300, 52101331 U20A20327)。

作者简介: 马蕴涵 (2000—), 男, 辽宁省沈阳市人, 主要从事极地海洋工程研究。E-mail: mayunhan100027@163.com

* 通信作者: 陈晓东, 博士, 高级工程师, 主要从事海冰力学及极地海洋工程研究。E-mail: chenxiaodong@dlut.edu.cn

害预报与海水资源规划等应用方面都非常重要。因此获取海冰密集度能够对海冰的存在状况与运动状态等进行深层次的了解,进而可以对各类海上活动进行指导与帮助^[8-9]。以船基海冰图像视频进行密集度识别需要准确、高效的方法,识别的即时性与准确性对于保证船舶航行安全、人为规划航线和动态修正航线具有重要指导意义。

目前海冰密集度识别的传统方法主要分为人工标注方法与阈值分割方法。基于人工标注的方法可将视频图像二值化并进一步获取海冰密集度信息^[10-12]。根据冰、水介质在图像上显现的不同特征,采用分水岭方法可进一步避免简单二值化引起的阈值不均匀问题^[13-14]。对于阈值分割方法中碎冰边缘粘连的问题,Zhang等基于梯度向量流 (gradient vector flow, GVF) 法提出了一种浮冰轮廓提取算法,规避了平整冰、碎冰边缘的不准确识别^[15]。Kalke与Loewen在河冰数字图片的研究中指出,目前识别冰较为广泛应用的技术是设置常量阈值并应用于多个图像的识别,但阈值方法与人工标注方法均有缺陷^[16]。采用深度学习方法,可进一步降低图像处理中的人工参度较高的问题。

近些年来深度学习方法在计算机视觉领域发展迅猛,卷积神经网络能够自动学习图像中的特征、挖掘深层语义信息,从而对不同图片进行准确的场景理解^[17]。Long等提出了全卷积神经网络 (FCN),利用全卷积层替换以往卷积神经网络 (CNN) 中的全连接层,使得神经网络可以接受任意大小的输入图像并以端对端的训练方式输出像素级别的分割图^[18]。Ronneberger等提出了一种轻量级的语义分割网络 U-Net,其网络延续了自编码的结构并在解码端结合了更多底层特征图的信息,为分割图提供更加精细的特征^[19]。Chen等基于 FCN 网络进行改进,提出了 DeeplabV1 模型,该网络引入的空间金字塔池模块能够充分利用图像中不同尺度的上下文信息,从而有效提高对不同尺寸物体的识别精度^[20]。对于图像这种语义特征不明显,即需要人工输入语义且没有直观物理含义的学习对象,深度学习模型能够在大规模训练数据的基础上取得更好的效果。相比于传统的神经网络,深度神经网络做出了重大的改进,在训练上的难度可以通过逐层预训练有效降低。这些语义分割网络在自动驾驶、遥感以及医疗影像等领域取得了很好的应用效果^[21-24]。而对于海冰图像的识别研究,多数学者基于卷积神经网络对合成孔径雷达 (SAR) 图像中海冰的面积以及密集度进行预测^[25-27]。相比之下,对于现场监测中海冰图像识别方法的研究相对较少,且

工程应用中的识别精度与智能化程度也有所不足。在应用 DeeplabV3+网络处理冰水分离任务的研究方面,宋巍等为获取准确实时的冰间水道分布情况,提出了一种基于 DeeplabV3+的轻量化合成孔径雷达图像冰间水道分割模型^[28]。孙士昌等选取 DeeplabV3+网络,以格陵兰海为实验区对该海域 2020–2022 年间冬季的 10 景 Sentinel-1A 双极化 SAR 影像进行处理^[29]。Zhang 等基于 Deeplab 网络对海冰图像进行海水、海冰和天空三分类识别^[30]。

为进一步增加视频图像中海冰识别的准确度与适用方面,本文提出了一种以 DeeplabV3+语义分割网络框架为基础的卷积神经网络方法,应用于“雪龙”号北极科考的实测图像数据中。采用该方法在不同海冰场景中分别对海冰、海水、天空与船体对象进行了识别,并通过平均交并比对识别结果的准确性进行评估,进而计算海冰密集度,同时也对于影响识别精度的因素也进行了探究。

2 基于深度学习的海冰图像语义分割方法

海冰图像识别本质上是将图像信息进行像素级分类,从而进一步获取海冰密集度等不同的海冰要素信息。为此本文采用基于 DeeplabV3+的深度学习模型框架对海冰视频图像进行语义分割。采用极地科考采集的真实海冰冰情视频图像作为研究对象,并依据相应评价标准对识别结果的准确性进行评估。

2.1 DeeplabV3+深度学习模型架构

在图像语义分割任务中,通常采用空间金字塔池化 (Spatial Pyramid Pooling, SPP) 或基于编码-解码器的结构。本文 DeeplabV3+基于 Tensorflow 框架来实现功能,主要在模型的架构上做出提升,引入了可控制编码器提取任意特征分辨率的功能的同时通过空洞卷积 (Dilated Convolution) 平衡精度和耗时^[31]。在深度学习模型中,主干网络的作用是提取对象特征,DeeplabV3+的网络结构详见图 1。DeeplabV3+提供了 Xception 和 MobileNetv2 两个主干网络,出于计算效率的考虑本文采用 MobileNetv2 主干网络。由于网络特性,在 DeeplabV3 网络中一般不会进行 5 次下采样,本文采用了 4 次下采样。在完成特征提取后可获得两个有效特征层,一个有效特征层是输入图片高和宽压缩两次的结果,另一个有效特征层是输入图片高和宽压缩 4 次的结果。

针对复杂场景内的海冰图像特征,DeeplabV3+特有的加强特征网络结构能够更好地提取海冰的深度

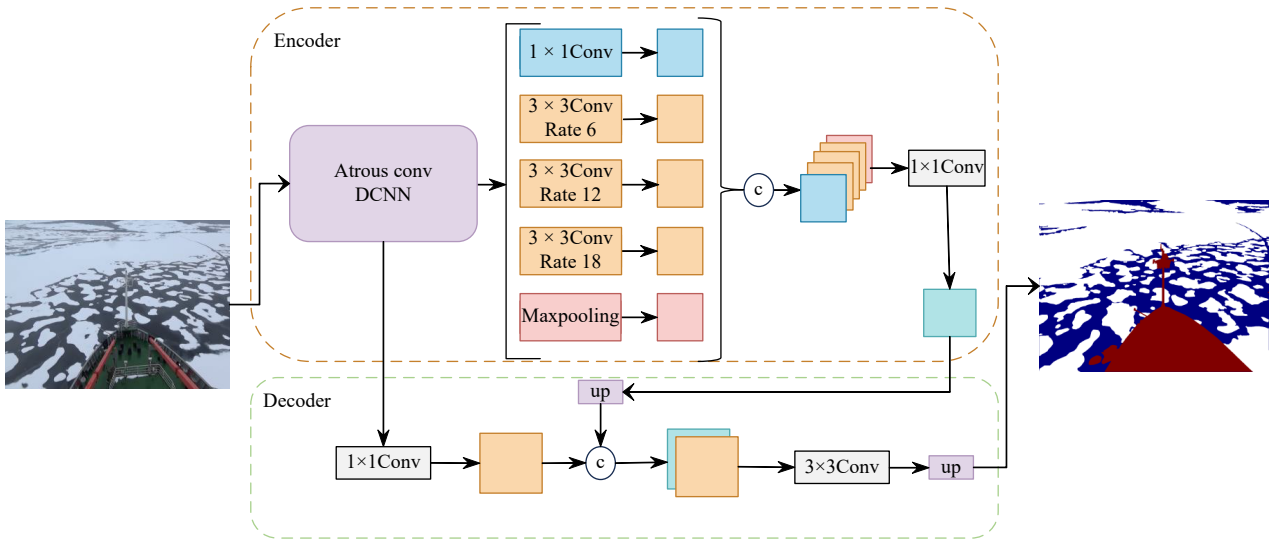


图 1 基于 DeeplabV3+ 的海冰识别模型

Fig. 1 Sea ice recognition model based on DeeplabV3+

语义信息。DeeplabV3+的加强特征提取网络可以分为编码与解码两个部分。通过在编码器中引入了大量空洞卷积,在特征点提取时进行跨像素提取,使每个卷积模型在不损失信息的情况下增加了图像感知区域。同时,编码器会在不同学习率下采用并行空洞卷积核提取压缩4次的初级特征层。而在解码器会利用 1×1 卷积对空洞卷积后的有效特征层进行上采样结果堆叠。同时DeepLabV3+网络新引入的多尺度特征提取模块(ASPP)通过不同的空洞卷积提取不同尺度的特征,从而能够捕捉更广泛尺度的上下文信息。使模型在捕捉大目标特征(海冰)的同时,也能准确识别小目标(船只),从而提升模型整体分割性能。

经过前两步可以获取输入图片的特征,此时需要利用特征获得预测结果,过程分为2步:第1步利用一个 1×1 卷积进行通道调整,调整成类别数(Num_Classes);第2步再利用resize函数进行上采样使得最终输出层,宽高和输入图片一样;最后在进行多次迭代训练后得到图像识别结果。这种新的神经网络在语义图像分割数据集上高度有效,并且在没有任何后处理的情况下能够在测试集上实现89%的性能,特别是针对沿着对象边界方面有着良好的效果。

2.2 海冰视频图像数据集来源

基于深度学习方法对海冰图像进行识别的任务需要人为对图像中的信息进行分类,制作为数据集。一般来说,数据集中的大部分图像用于训练深度学习模型,其余图像用于验证模型的识别准确率。

本文采用“雪龙”号在2017年第8次北冰洋科学考察与2018年第9次北冰洋科学考察任务中的走航视频图像作为数据集。这两个航次中进行海冰图像

信息采集时,视频摄像机均安装于罗经甲板正前方。其中,摄像机水平视场角为 90° ,竖直视场角为 45° ,焦距为4 mm。根据摄像机的成像范围,所拍摄的图像中主要包含天空、海水、海冰与船身4类信息。因此在制作数据集的过程中采用Labelme图像分类工具将每张图像以4类要素进行标注。本文采用的“雪龙”冰区走航视频总时长为1440 h,采样频率为30帧/s,确保能够捕捉到足够细节以支撑分析结果;根据视频中参照物“雪龙”号船宽与采样像素估测,图像分辨率从约为 $5.76 \sim 28.8$ 像素/m;在筛选数据时由于能见度、天气状况、云雾影响等因素会严重影响标签的制作工作,因此为确保数据质量对可视情况较差的样本进行了剔除。总数据集共包括在视频中截取的670张图片,其中603张(90%)图片用作训练集,67张图片(10%)用作验证集。图2中列出了训练集中的海冰图像以及对应的标注结果示例,共有海冰、海水、天空、船身4个语义类别。

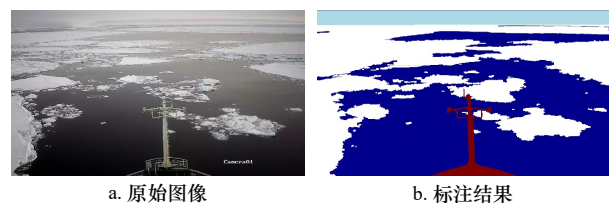


图 2 数据集标签图像

Fig. 2 Original image and annotated image of the dataset

2.3 海冰图像识别的训练流程

本文以DeeplabV3+网络对海冰图像进行识别。首先进行数据集的建立,通过搭载于船体的摄像机对冰区航行过程进行视频信息采集,从航行视频中截取

不同条件下的海冰图像后对其进行人工标注,制作数据集。之后进行训练,将数据集输入 DeeplabV3+网络中迭代训练,迭代次数取决于损失函数是否收敛,训练结束保存相应的模型权值参数文件。最后进入验证阶段:将验证数据集中的测试图像输入已训练好的海冰识别模型中,调用训练得出的权值文件,根据输出预测图与真实结果进行比对来评估模型的预测精度。

2.3.1 深度学习模型的参数设置及运行环境

海冰语义分割试验基于 Tensorflow-GPU 2.6 环境搭建,由于软硬件配置对模型的训练与识别效率具有一定影响,因此将采用的软硬件配置参数列于表 1。

表 1 海冰图像识别的软硬件配置

Table 1 Configurations of hardwares and softwares for identification of sea ice images

硬件参数	CPU	Intel i7
	RAM	16 GB
	Graphics card	Nvidia GeForce RTX 3 060 Laptop
	GPU Memory	6 GB
软件参数	Operating system	Windows 11
	Python	Python 3.8
	Framework	Tensorflow-gpu 2.6
	CUDA	11.2

除硬件配置外,网络中超参数的设置也影响整个训练过程的速度及最终模型的精度,重要的超参数有迭代次数(epoch)、学习率(Lr)与动量参数(Momentum)。迭代次数即进行训练的次数,表征所有图片被输入神经网络中进行学习的次数。学习率决定经过一次训练后模型更新权重的步长,迭代次数与学习率会影响模型的训练速度和性能。动量参数被用以优化模型训练的效果,在深度学习中进行梯度下降算法(Momentum SGD)时可以去噪声,并使权值更新的过程更加平滑,引入动量参数能够让模型在学习过程中更多保留一些先前训练的信息,进而提升权值模型的准确度并加快收敛速度。

对于本文数据集,海冰复杂场景分类任务涉及高层次特征学习的任务在使用 adam 优化器训练时震荡较大、收敛趋势不明显,因此采用了动量因子为 0.9 的动量梯度下降法作为优化器来更新神经元的权重。总迭代次数(I_{max}) 10 000 次,初始学习率(Lr_{init}) 设置为 $7e^{-3}$,遵从 poly 学习策略进行多项式误差衰

减,其计算式可写为

$$r = Lr_{init} \left(1 - \frac{N_i}{I_{max}}\right)^m, \quad (1)$$

式中: i 为当前迭代的次数; N_i 为动量参数,一般设为 0.9; m 为误差衰减的指数,随着迭代次数的增加,学习率会不断降低以寻找模型的最优解,保证在训练后期网络不会有过大波动。由于本文数据集是从零开始训练,训练输入数据受到计算机硬件内存限制,因此训练时的输入大小(input_shape)遵从该数据集并相应设置为 512 像素 \times 512 像素。同样由于显卡性能的限制,每一个训练批次能够处理 4 张图片(batchsize = 4),权重衰减模式(lr_decay_type)为 cos 型,下采样倍数(downsample_factor)为 16,该参数取值越小会使特征图保留更多的空间信息,在一些情况下会使得训练效果越好,但同时要求更大显存。

2.3.2 深度学习模型的权值输出

权值衡量模型中神经元间输入输出的影响程度,作用是连接各个神经元并使得模型网络能够在获得输入图像后给出预测结果,也是模型对图片进行语义分割的最终输出结果。本文中权值文件选择以 HDF5 的格式进行输出。基于以上参数对数据集进行训练,会每 5 次训练迭代出一次权值并进行评估,该评估过程以训练损失率(epoch_loss)作为判别依据,选择当前最佳训练周期中得到的权值(best_epoch)。损失率由深度学习模型的真实值和预测值确定,表征二者间差异程度,在本文中以验证集输入分割图与输出分割图得出。同时每 5 次迭代后还会输出验证集的平均交并比(mIoU),以作为判断权值优良的第二判据。在训练结束后,将最佳权值代入预测程序,查看识别结果。

2.4 深度学习方法准确性的评价方法

评价指标能够对预测结果优良进行评估,是衡量模型准确性的主要依据。模型权重训练完成后,需要将验证集的图像作为输入,根据预测图的结果来验证模型的预测精度。在图像语义分割任务中,常用精确率(P_r)、召回率(recall)、平均交并比(IoU)、加权交并比(mIoU)等指标作为评估标准。其中 P_r 表示在所有被预测为正的样本中实际为正的样本的概率,表达式为

$$P_r = \frac{T_p}{T_p + F_p}, \quad (2)$$

式中: T_p 表示真实标签为正类且预测也为正类的像素数目, F_p 表示真实标签为负类但被预测为正类的像素数目。以 recall 代表真实标签为正样本中被正确分割

的比例,其可写为

$$\text{recall} = \frac{T_p}{T_p + F_N}, \quad (3)$$

式中: T_N 表示真实标签为负类且预测也为负类的像素数目。

交并比 (IoU), 为语义分割任务中最常用的评价标准, 它表示每一类别预测的结果和真实值的交集与并集的比值, 其表达式为

$$\text{IoU} = \frac{n_{ii}}{t_i + \sum_j n_{ji} - n_{ii}}. \quad (4)$$

平均交并比 (mIoU) 是在 IoU 的基础上并对所有类别的交并比求平均值的结果, 表达式为

$$\text{mIoU} = \frac{1}{n} \sum_i \frac{n_{ii}}{t_i + \sum_j n_{ji} - n_{ii}}, \quad (5)$$

式中: n 表示类别数, n_{ii} 表示被正确预测为 i 类的像素数目, t_i 表示类别 i 在标签中的像素总量, n_{ji} 表示被预测为 i 类的总像素数目。

本文不仅关注全局指标 (如 mIoU), 还单独关注小类别的指标表现 (如船的 IoU)。主要采用 IoU 以及 mIoU 对海冰识别模型的精度进行评估, IoU 值越高说明模型预测图中的某一个特定类别越接近于标签中的真实结果, 而 mIoU 能够反映模型的整体预测能力。如果小类别指标过低, 会进一步优化加权函数。

3 海冰视频图像的分类结果与分析

3.1 海冰图像的分类结果及评估

采用前文所述的深度学习模型对网络训练后, 将验证集中的海冰图像作为输入以检验海冰识别模型的分类效果。部分预测结果如图 3 所示, 图中第 1 列表示输入进网络中的原始图像, 第 2 列表示对应的标注图像, 第 3 列表示原始图像在本文网络上的预测结果。通过将结果与标注结果相对比, 可以看出训练的海冰语义分割模型能够有效识别输入的 RGB 图像, 对图像中的海冰、海水、天空以及船体对象有较好的识别效果。

本文模型准确分割出了不同形状浮冰, 特别是对碎冰的预测结果与真实值较为接近。该模型在碎冰尺寸较小且冰块边缘杂乱时也能够较为准确地对碎冰要素进行预测, 如图 3 1a-c。在图像中海水要素的占比较大时, 远端海冰的总体分布识别结果与标签图较为接近。而在冰-水交界的轮廓处存在一定的识别误差, 特别是大面积海冰内出现区域较小的水道与融池时, 有时无法准确识别, 如图 3 g-i 中的场景 3 所

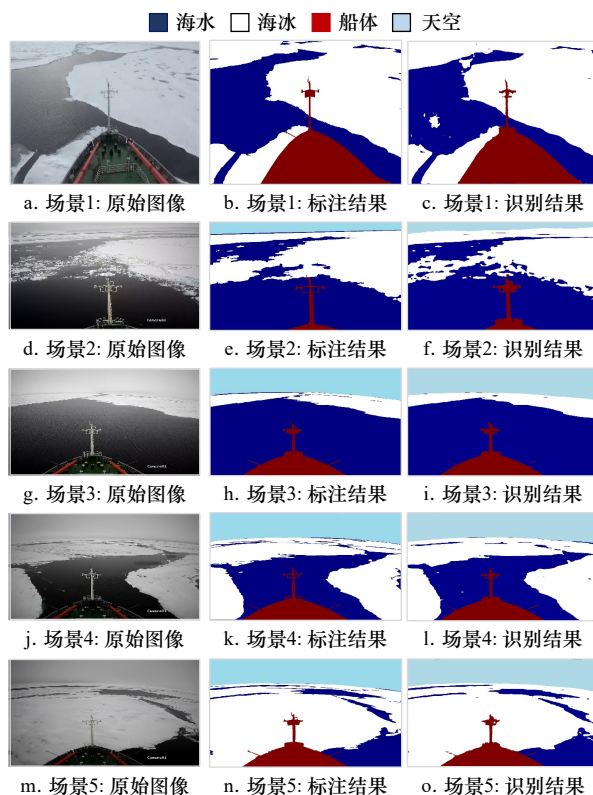


图 3 海冰图像分类识别结果

Fig. 3 Segmentation results of sea ice images

示。当视野中尤其近处出现海冰占比较高时, 模型对浮冰和海水的分割效果较为稳定, 如图 3 5m-o 所示。根据 2.1 节中所采用的图像精度评估指标, 将不同场景中 4 类要素的 IoU 及 mIoU 计算结果列于表 2。在评估标准中, 各对象单项交并比可以反映各识别对象的识别准确度; 平均交并比能够对各单项交并比进行加权, 进而客观反映出模型对图像总体的识别准确度。

表 2 针对不同场景识别结果的单项交并比与平均交并比

Table 2 IoU and mIoU of different samples

	IoU-Sea/%	IoU-Ice/%	IoU-Sky/%	IoU-Ship/%	mIoU/%
场景1	99.9	97.6	-	97.4	98.3
场景2	99.8	93.5	87.9	87.7	90.8
场景3	99.9	92.1	97.1	95.7	96.2
场景4	96.9	95.3	98.8	93.3	96.9
场景5	95.9	96.2	98.8	93.9	97.2
验证集(67张)	96.9	97.4	96.5	95.8	96.0

验证集中选取的 67 张图片考虑到了海冰不同的分布情况, 可以更好地作为评估模型分割性能的依据。从验证集精度数据可以看出, 本文中所训练的海

冰识别模型有较高的精度, 验证集整体的 mIoU 值达到了 96.9%, 从各类标签的 IoU 值可以看出, 模型对于海水和天空目标识别结果较为稳定, 其 IoU 值分别达到了 90.7% 以及 98.4%。从变化趋势上看, 海冰的交并比预测值与图像中冰的分布情况有关, 冰在占比下的单项交并比 (IoU-Ice) 与平均交并比 (mIoU) 的数值规律如下: 在进行图像矫正前, 冰所占比例较低时, 如场景 3 (占比 15.2%)、标签中海冰类别的像素数目较少, 分割图中所预测海冰存在的误差会在很大程度上影响冰交并比 (92.1%), 进而影响平均交并比 (96.2%)。反之, 当视野中冰较多、比例较高时, 如场景 4、5 (占比 55.0%、80.5%), 少部分海冰的识别错误并不会对冰的 IoU 值有过大影响, 冰单项交并比能达到 95.3% 与 96.3%, 平均交并比分别达到 96.9% 与 97.2%。

3.2 海冰密集度对识别精度的影响

由 3.1 节输出结果可知, 图像内各类要素的占比对识别结果存在一定影响, 因此需要针对不同类别要素占比对识别结果的影响做进一步分析, 对模型输出的分割图像进行矫正可进一步对计算得到海冰密集度进行讨论。具体矫正过程如图 4, 对原始图像和分类图像分别进行畸变矫正和倾斜矫正。其中畸变矫正可以去除相机镜头造成的广角畸变, 倾斜将倾斜拍摄的图像转换为俯视图, 以确保后续计算时, 图像的像素面积与实际海面上各要素面积比例一致。

分类图像经矫正后能够准确地得到真实的冰水比例, 计算海冰的面积占比, 从而获取海冰密集度,

$$\text{Ice_concentration} = \frac{\text{Pixels_ice}}{\text{Pixels_ice} + \text{Pixels_sea}}, \quad (6)$$

式中: Pixels_ice 为矫正后的海冰像素数, Pixels_sea 为矫正后的海水像素数, Ice_concentration 为海冰像素占海水海冰像素总和的比例, 即海冰密集度。

根据以上公式所计算的海冰密集度, 将识别评价指标与海冰密集度的关系在图 5 中给出。从图 5 中能够看出海冰密集度的变化对识别精度的影响较为显著, mIoU 与 IoU-Ice 均随海冰密集度增加而上升。由于平均交并比包含海冰在内所有识别类别交并比的加权平均值, 因此平均交并比的上升趋势较弱。海冰单项交并比受海冰密集度的影响较大, 上升趋势较为明显。这是由于海冰作为识别对象在图中被标注后输入到模型网络中, 对在网络中海冰像素的增加使得海冰像素预测值与真实值的交集 (n_{ii}) 与并集 ($t_i + \sum_j n_{ji} - n_{ii}$) 含有的像素值等大小增加, 因此计算交并比 ($n_{ii}/(t_i + \sum_j n_{ji} - n_{ii})$) 时分子、分母项会随着密集度的升高同时等值增加, 从而提升交并比的数值。同时由于海冰对象在图像中占据更大的比例, 使得更多的海冰信息需要被输入到深度学习模型中, 也需要更大的感受野来捕获更广泛的上下文信息。本文模型基于 DeepLabV3+ 架构, 通过空洞卷积有效地扩展感受野, 能够在不增加计算负担的情况下提取多尺度的特征。这种多层次特征提取能力, 使得模型能够从训练集中的像素信息中提取更深层的海冰语义特征。此外由于采用编码器-解码器结构, 其中编码器使用了空洞卷积来捕捉多尺度的上下文信息, 解码器则通过对低层次特征的逐步融合, 恢复精细的分割边界, 帮助模型在分割边缘和细节上取得了显著的改进。对于海冰密集度较高的场景可以更好地理解海冰对象的复杂特征, 从而提高识别的准确性。

3.3 观测距离对识别精度的影响

考虑到倾斜摄影测量的特点, 分析观测距离对冰水分类精度的影响。由于倾斜摄影测量的成像角度与视距呈以下关系: 非垂直角度下的斜拍图像通常会使得图像中的物体可能呈现不同的形态变形。因此

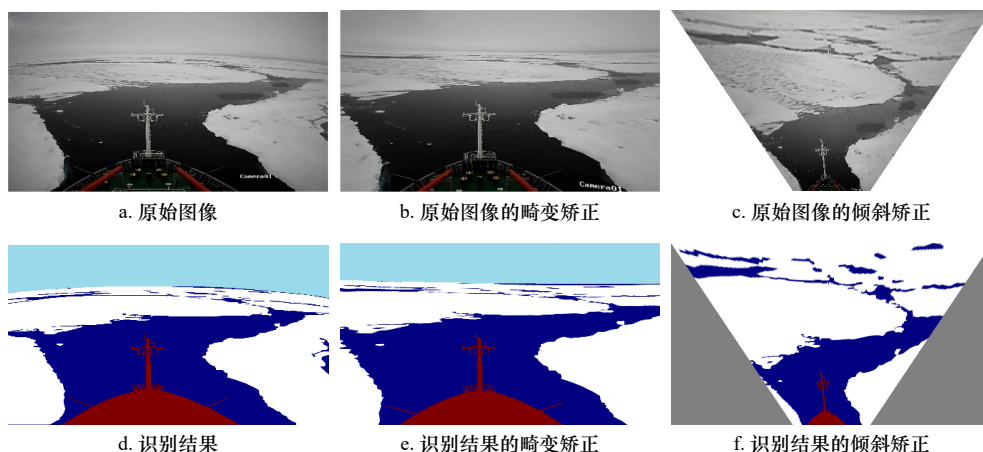


图 4 海冰图像的畸变矫正与倾斜矫正

Fig. 4 Distortion and perspective correction of sea ice images

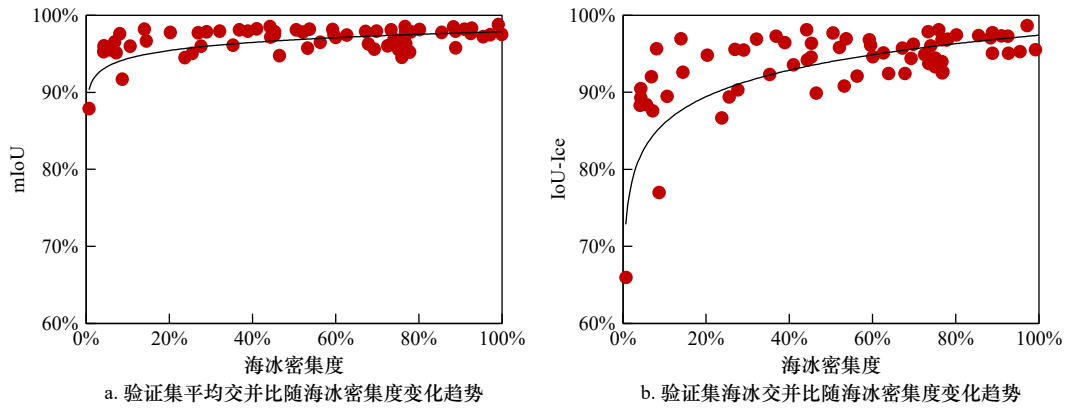


图5 识别精度随海冰密集度的变化趋势

Fig. 5 The trend of recognition accuracy with ice concentration

与直拍图像相比,斜拍图像的观测角度使得物体的尺寸、形状、和相对位置发生变化,在分割任务中直接影响到分割精度。近距离观测时,图像的分辨率较高,细节更加清晰,分类精度较高。由于较小的视角差异,冰水等区域的划分较为准确。而随着观测距离的增加,图像分辨率下降,对象细节丢失,分类精度可能降低。尤其在冰水的分割上,细微的特征变化(如冰与水的边界)变得不再明显,导致误分类。因此随着观测距离的增大,分类精度会逐渐下降。可以通过分析不同观测距离下模型的误差定量评估这一影响。利用固定数量的标注数据集,比较模型在不同距离下的分类性能。通过计算每个像素距离下的 mIoU 指标,得到图像分类精度的量化结果,如图 6 所示。

图中横坐标是观测距离,在图中随着倾斜角度变化由近到远,纵坐标是在该高度下各个识别对象的平均交并比。能够看出分类精度与观测距离存在显著的相关性。在观测距离约为 0~230 m 的范围内由于像素精度高,在最近处可达到约 28.8 像素/m,分类效果也较好。在 230~250 m 范围内,此时平均交并比最低值降至 0.62。这是由于处于较远位置(天际线附

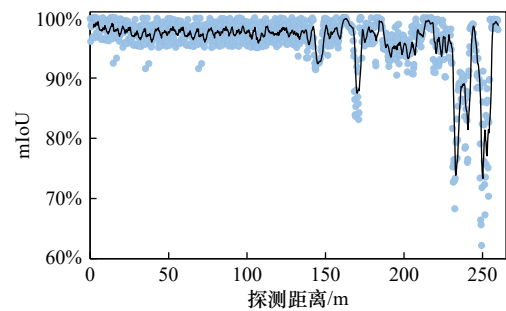


图6 验证集平均交并比随观测距离的变化趋势

Fig. 6 The correlation between image mIoU and observation distance

近)的像素精度较低,约 5.76~6.26 像素/m,且标签制作误差较大,因此此处分类效果较差。通过对比不同观测距离下的分类精度,可以认为在不大于 200 m 的尺度下本文模型可以保证较高的分类精度。

3.4 训练迭代次数对识别精度的影响

海冰图片识别的精度也与训练迭代次数相关。本文的 10 000 次训练分为两次完成,每次 5 000 次迭代,如图 7 所示。在 0~5 000 次迭代的训练过程中 loss 值逐渐减小,并趋近于 0.45 左右,平均交并比逐

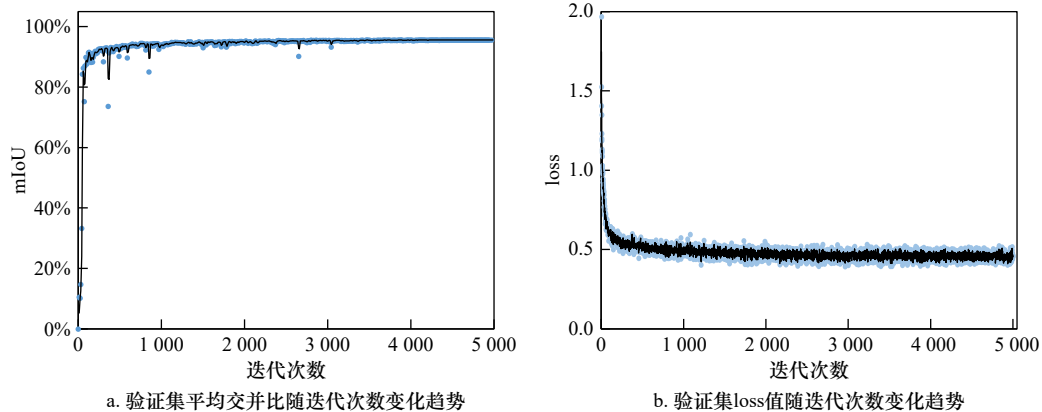


图7 识别精度随迭代次数变化趋势

Fig. 7 The trend of recognition accuracy with the epochs

渐增加,并趋近于95%。随着迭代次数增多识别精度也在逐渐上升,从平均交并比与损失值的收敛状况来看,训练精度在2000次迭代趋于稳定(mIoU达到91%以上),并且趋势逐渐减缓。前期平均交并比随着迭代次数快速增加是由于各个识别对象的预测值与真实值交集集合(n_{ii})逐渐增加的同时并集合($t_i + \sum_j n_{ji} - n_{ii}$)减少,导致各对象交并比 $[n_{ii}/(t_i + \sum_j n_{ji} - n_{ii})]$ 的分子项增大、分母项减小,进而使得平均交并比快速增加。而2000次迭代后上升趋势的减缓主要是由于训练模型精度已经足够高,触发了网络的防过拟合机制,迭代次数的继续增加会使得识别对象边缘继续细化、接近真实值。由于边缘部分的像素相较识别对象本身占比较少,因此继续训练对平均交并比数值上的提升并不明显,但对于对象边缘细节的分割会更加精确。

3.5 数据集规模对识别精度的影响

除上文中讨论的海冰密集度与训练迭代次数,数据集的大小同样会影响图像识别的精度。使用对照组和不同大小的数据集进行对比,可以清晰地评估训练效果。本文使用了670张图片进行训练,在对照组中我们从上述数据集中随机选取134、268、402、536张图片,即原数据集的20%、40%、60%和80%作为训练集,每组训练5000代与本文中前5000代的训练结果进行对比。如图8所示,可以看出随着数据集大小的增加,识别精度也在持续增加。这是因为更大的数据集提供了更多的样本,使模型能够学习更广泛、更具代表性的特征和模式,从而提高了模型的泛化能力。数据集大小与识别精度之间的关系通常呈现出一种学习曲线,随着数据量的增加,模型性能会逐渐提升。然而一旦达到某个阈值,额外增加数据对性能的提升效果会逐渐减弱,在数据集图片数达到268张后平均交并比达到90%以上之后上升趋势变

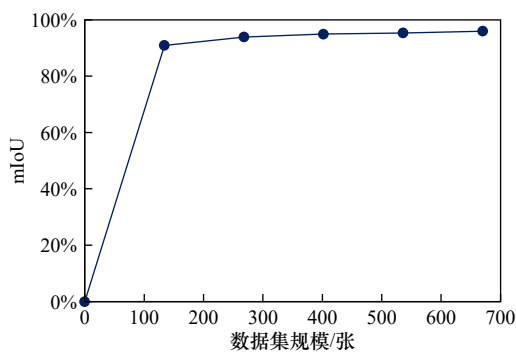


图8 验证集平均交并比随训练集大小的变化趋势
Fig. 8 The trend of mIoU of the validation set with the image amount

缓。原因是随着数据集大小的增加,验证集识别精度呈现出递增的趋势在数据集大小达到一定程度后,模型已经学习到了大部分海冰图像数据中的分布与形态模式,模型的性能提升效果将逐渐减弱,因此平均交并比会在一定范围内趋于稳定。

3.6 海冰语义分割模型的泛化性分析

为进一步分析语义分割模型的泛化性,本文选取了制作数据集以外的海冰场景作为分析验证对象,如图9所示。其中,泛化性分析所采用的图像为第10次北极科考与第39次南极科考的走航观测实测数据。与北极科考不同,陆地的存在使得南极科考过程中会采集大量沿岸固定冰的图像。固定海冰不产生漂移运动,冰区内部一般不存在冰间水道或开阔水域,所采集图像中可能仅有海冰与天空而没有海水,如图9c所示,在这种情况下模型可以准确地识别出海天交界线和准确的冰轮廓,如图9d所示,mIoU达到98.0%。由于每个航次均对摄像机进行重新安装,图像采集设备的角度与型号参数均存在一定差异,从而导致图像信息也产生一定变化。摄像俯仰角的变化导致了海天交界线位置与天空占比明显变化,如图9a与b所示,mIoU达到99.4%。因此尽管天空要素的占比发生大幅度变化,但从分类结果可以看出受到的干扰并不明显。根据数据集外不同场景的预测结果能够看出,深度学习模型在不同相机类型、拍摄角度等因素下仍具备较好的泛化能力与鲁棒性。

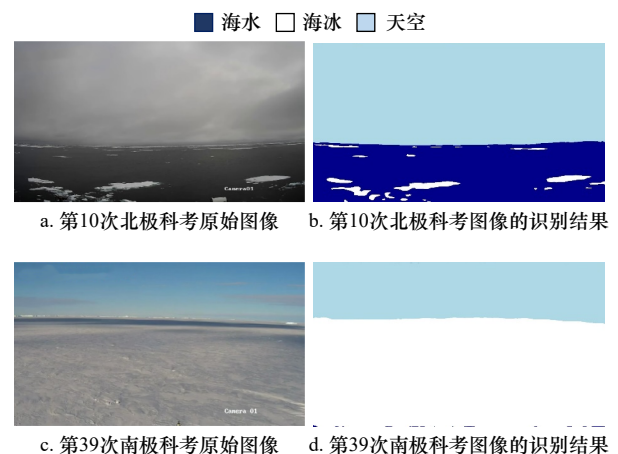


图9 不同场景海冰图像识别结果

Fig. 9 Identification results of sea ice images in different scenarios

不同模型对于船基海冰图像所包含的高、低频特征的学习能力是有差异的, DeeplabV3+模型较强的泛化能力主要与其网络结构特点相关。图像中的高频特征是图像灰度、颜色与形状信息随像素移动发生高频变化的部分,如对象边界、纹理及细小特征等,

在船基海冰图像中一般为海冰的边缘、裂缝、船体轮廓或天际线等语义特征。而低频特征是图像中信息变化较缓、灰度梯度较小或者整体性较强的部分,比如图像背景、全局特征、平滑着色区域等,在海冰图像中一般代表海冰的形状、分布和船体、天空、海水、平整冰等大型对象的整体特征。这些信息可以帮助模型理解海冰的大体结构和空间分布,从而对海冰图像进行整体的识别和分析。尽管 DeeplabV3+采用了空间金字塔池化模块来提取不同大小目标的特征,但该模块的输入较多由提取到的高频特征图构成。因此,在随后的特征融合阶段网络本身并没有加入更多浅层特征图的权重,是直接上采样到原图尺寸输出预测结果。而低频特征,如原图 1/2、1/4 大小的特征图中含有更为丰富的轮廓内信息。这也是 DeeplabV3+网络对于识别边界较为圆润,导致复杂形状碎冰与裂缝端点的预测效果仍不够精细的一个原因。

4 结论

采用基于深度学习的方法,以 DeeplabV3+语义分

割框架作为基础,本文研究了卷积神经网络应用于船载摄像机拍摄海冰图像的识别能力。通过对海冰、海水、以及天空目标的预测结果,证明了语义分割模型在海冰图像识别中的可行性。本文主要研究结果主要体现为:所训练的海冰语义分割模型能够有效识别视频图像中的海冰信息,预测结果识别精度较高,其中验证集的 mIoU 值超过了 95%;本文采用的迁移学习方法能够使网络以较小规模的训练集本达到较为理想的学习效果,当数据集规模超过 150 张时,增加数据集规模对提升模型识别精度的效果不再显著;基于 DeeplabV3+语义分割框架的深度学习模型在识别海冰图片时具有较好的泛化性,对不同场景下与不同航次的海冰图像均得到了较好识别结果。

DeeplabV3+网络结构中没有考虑引入浅层特征,使得预测图中部分碎冰的预测结果仍然不够精细,因此在后续的工作中应评估其他主流的语义分割网络应用于海冰图像的预测能力,同时应加入更多类别的海冰图像,以增量学习的方式进一步提高模型可靠性以及准确性。

参考文献:

- [1] 苏洁,郝光华,叶鑫欣,等.极区海冰密集度 AMSR-E 数据反演算法的试验与验证[J]. *遥感学报*, 2013, 17(3): 495-513.
Su Jie, Hao Guanghua, Ye Xinxin, et al. The experiment and validation of sea ice concentration AMSR-E retrieval algorithm in polar region[J]. *Journal of Remote Sensing*, 2013, 17(3): 495-513.
- [2] 魏彬航,李宝辉,刘煜,等.辽东湾海冰分布面积历史数据重构及其影响因素分析[J]. *海洋学报*, 2023, 45(11): 20-33.
Wei Binhang, Li Baohui, Liu Yu, et al. Reconstruction of sea ice extent in the Liaodong Bay and analysis of its impact factors[J]. *Haiyang Xuebao*, 2023, 45(11): 20-33.
- [3] 崔洪宇,胡大士,孔帅,等.基于正则化方法的雪龙号破冰船冰载荷反演的研究[J]. *中国造船*, 2020, 61(1): 109-119.
Cui Hongyu, Hu Dashi, Kong Shuai, et al. Study on inversion of ice load for Xue Long icebreaker based on regularization method[J]. *Shipbuilding of China*, 2020, 61(1): 109-119.
- [4] 邓娟,柯长青,雷瑞波,等.2009 年春夏季北极海冰运动及其变化监测[J]. *极地研究*, 2013, 25(1): 96-104.
Deng Juan, Ke Changqing, Lei Ruibo, et al. Monitoring the motion of arctic sea-ice and its changes in summer and winter 2009[J]. *Chinese Journal of Polar Research*, 2013, 25(1): 96-104.
- [5] 张培宣,陈晓东,孔帅,等.基于 Hough 变换原理的海冰厚度识别方法[J]. *海洋学报*, 2022, 44(7): 161-169.
Zhang Peixuan, Chen Xiaodong, Kong Shuai, et al. Research on sea ice thickness identification method based on Hough transform principle[J]. *Haiyang Xuebao*, 2022, 44(7): 161-169.
- [6] Laxon S W, Giles K A, Ridout A L, et al. CryoSat-2 estimates of Arctic sea ice thickness and volume[J]. *Geophysical Research Letters*, 2013, 40(4): 732-737.
- [7] Aldenhoff W, Berg A, Eriksson L E B. Sea ice concentration estimation from Sentinel-1 Synthetic Aperture Radar images over the Fram Strait[C]//Proceedings of 2016 IEEE International Geoscience and Remote Sensing Symposium. Beijing, China: IEEE, 2016: 7675-7677.
- [8] Liu Yinghui, Key J, Mahoney R. Sea and freshwater ice concentration from VIIRS on suomi NPP and the future JPSS satellites[J]. *Remote Sensing*, 2016, 8(6): 523.
- [9] 郑付强,匡定波,胡勇,等.基于 U-ASPP-Net 的北极独立海冰精细识别方法[J]. *红外与毫米波学报*, 2021, 40(6): 798-808.
Zheng Fuqiang, Kuang Dingbo, Hu Yong, et al. Refined segmentation method based on U-ASPP-Net for Arctic independent sea ice[J]. *Journal of Infrared and Millimeter Waves*, 2021, 40(6): 798-808.
- [10] 周嘉儒,卢鹏,王庆凯,等.基于视频图像获取冰面特征的自动检测算法研究[J]. *水利科学与寒区工程*, 2021, 4(5): 60-65.
Zhou Jiaru, Lu Peng, Wang Qingkai, et al. Research on automatic detection algorithm based on video image acquisition for ice surface feature[J]. *Hydro Science and Cold Zone Engineering*, 2021, 4(5): 60-65.
- [11] Toyota T, Haas C, Tamura T. Size distribution and shape properties of relatively small sea-ice floes in the Antarctic marginal ice zone in late winter[J]. *Deep Sea Research Part II: Topical Studies in Oceanography*, 2011, 58(9/10): 1182-1193.

- [12] 季顺迎, 王安良, 王宇新, 等. 渤海海冰现场监测的数字图像技术及其应用[J]. 海洋学报, 2011, 33(4): 79–87.
Ji Shunying, Wang Anliang, Wang Yuxin, et al. A digital image technology and its application for the sea ice field observation in the Bohai Sea[J]. Haiyang Xuebao, 2011, 33(4): 79–87.
- [13] Blunt J D, Garas V Y, Matskevitch D G, et al. Image analysis techniques for high arctic, deepwater operation support[C]//Proceedings of the OTC Arctic Technology Conference. Houston: OTC, 2012.
- [14] Ijtona T B, Ren Jinchang, Hwang P B. SAR sea ice image segmentation using watershed with intensity-based region merging[C]//Proceedings of 2014 IEEE International Conference on Computer and Information Technology. Xi'an, China: IEEE, 2014: 168–172.
- [15] Zhang Qin, Skjetne R, Metrikin I, et al. Image processing for ice floe analyses in broken-ice model testing[J]. *Cold Regions Science and Technology*, 2015, 111: 27–38.
- [16] Kalke H, Loewen M. Support vector machine learning applied to digital images of river ice conditions[J]. *Cold Regions Science and Technology*, 2018, 155: 225–236.
- [17] 田萱, 王亮, 丁琪. 基于深度学习的图像语义分割方法综述[J]. 软件学报, 2019, 30(2): 440–468.
Tian Xuan, Wang Liang, Ding Qi. Review of image semantic segmentation based on deep learning[J]. Journal of Software, 2019, 30(2): 440–468.
- [18] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation[C]//Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, USA: IEEE, 2015: 3431–3440.
- [19] Badrinarayanan V, Kendall A, Cipolla R. SegNet: a deep convolutional encoder-decoder architecture for image segmentation[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(12): 2481–2495.
- [20] Chen L C, Papandreou G, Kokkinos I, et al. DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, 40(4): 834–848.
- [21] Everingham M, Van Gool L, Williams C K I, et al. The Pascal visual object classes (VOC) challenge[J]. *International Journal of Computer Vision*, 2010, 88(2): 303–338.
- [22] Garcia-Garcia A, Orts-Escobedo S, Oprea S, et al. A survey on deep learning techniques for image and video semantic segmentation[J]. *Applied Soft Computing*, 2018, 70: 41–65.
- [23] Zhu Xiaoxiang, Tuia D, Mou Lichao, et al. Deep learning in remote sensing: a comprehensive review and list of resources[J]. *IEEE Geoscience and Remote Sensing Magazine*, 2017, 5(4): 8–36.
- [24] Hesamian M H, Jia Wenjing, He Xiangjian, et al. Deep learning techniques for medical image segmentation: achievements and challenges[J]. *Journal of Digital Imaging*, 2019, 32(4): 582–596.
- [25] Cooke C L V, Scott K A. Estimating sea ice concentration from SAR: training convolutional neural networks with passive microwave data[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2019, 57(7): 4735–4747.
- [26] Wang Lei, Scott K A, Xu Linlin, et al. Sea ice concentration estimation during melt from dual-pol SAR scenes using deep convolutional neural networks: a case study[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2016, 54(8): 4524–4533.
- [27] Wang Yiran, Li Xiaoming. Arctic sea ice cover data from spaceborne synthetic aperture radar by deep learning[J]. *Earth System Science Data*, 2021, 13(6): 2723–2742.
- [28] 宋巍, 祝敏, 石少华, 等. 基于改进 DeepLabV3+ 的轻量化 SAR 图像冰间水道分割[J/OL]. 计算机工程与应用, 1–13[2024-08-20]. <http://kns.cnki.net/kcms/detail/11.2127.TP.20240430.1016.002.html>.
Song Wei, Zhu Min, Shi Shaohua, et al. Lightweight SAR image lead segmentation based on improved DeepLabV3+[J/OL]. Computer Engineering and Applications, 1–13[2024-08-20]. <http://kns.cnki.net/kcms/detail/11.2127.TP.20240430.1016.002.html>.
- [29] 孙士昌, 王志勇, 李振今, 等. 基于改进 DeepLabV3+ 模型的海冰提取方法——以北极格陵兰海为例[J]. 海洋学报, 2024, 46(8): 131–142.
Sun Shichang, Wang Zhiyong, Li Zhenjin, et al. An extraction method for sea ice based on improved DeepLabV3+ model: taking the arctic Greenland sea as an example[J]. Haiyang Xuebao, 2024, 46(8): 131–142.
- [30] Zhang Chengqian, Chen Xiaodong, Ji Shunying. Semantic image segmentation for sea ice parameters recognition using deep convolutional neural networks[J]. *International Journal of Applied Earth Observation and Geoinformation*, 2022, 112: 102885.
- [31] Chen L C, Zhu Yukun, Papandreou G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[C]//Proceedings of the 15th European Conference on Computer Vision. Munich: Springer, 2018. DOI: 10.1007/978-3-030-01234-2_49.

Research on ship based digital image processing and sea ice concentration recognition based on deep learning

Ma Yunhan¹, Chen Xiaodong¹, Zhao Guanhui^{2,3}, Ji Shunying¹, Yang Haitian¹

(1. State Key Laboratory of Structural Analysis, Optimization and CAE Software for Industrial Equipment, Dalian University of Technology, Dalian 116023, China; 2. College of Computer Science and Technology, Zhejiang University, Hangzhou 310027, China; 3. China Ship Development and Design Center, Wuhan 430064, China)

Abstract: Sea ice is a typical environmental feature of polar sea areas, and pixel-level classification of ship-borne video images can provide high-resolution sea ice information. Due to the complex light conditions and sea ice morphology in polar scenes, traditional computer graphics methods lack the generalization needed for intelligent identification of sea ice elements. Therefore, this paper deploys a deep learning approach based on the DeeplabV3+ semantic segmentation network structure to identify sea ice elements in polar scenes. The dataset consists of sea ice images captured by the icebreaker *Xuelong* during its navigation in ice-covered regions, and also is used to train and validate the deep learning model. To meet the requirements of sea ice element identification and the characteristics of the underway observation video images, the pixel information is divided into four semantic categories: sea ice, sky, seawater, and ship. The deep learning model is built based on the correlation between image information and semantic information in the training set. The model trained is used to predict the semantic information of pixels in the validation set or other images, thereby achieving automatic identification of sea ice information. To study the robustness of this method, the influences of sea ice concentration, lighting conditions, and sea ice types on the identification results was further analyzed. Additionally, the effects of dataset size and the number of iterations on identification accuracy were examined. The recognition results for images show that the mean Intersection over Union (*mIoU*) for the four types of semantic information exceeds 95%, indicating that the deep learning method can accurately classify various elements in the polar environment.

Key words: sea ice; DeeplabV3+; semantic segmentation; image recognition; deep learning