文章编号:1000 - 4653(2026)04 - 0164 - 11

基于残差-状态空间模型混合架构的北极海冰分割网络

郁纪成, 曲嘉铭*, 郝宇驰
(中交疏浚技术装备国家工程研究中心有限公司, 上海 200082)

摘 要:针对北极海冰多任务分割中局部与全局特征提取难以兼顾的挑战,提出一种融合了残差网络与状态空间模型的混合架构。该架构以残差网络为编码器提取多尺度局部特征,解码器则采用增强型视觉状态空间模块,通过结合二维选择性扫描与双重注意力机制实现全局的上下文建模;在跳跃连接部分,设计了跨尺度上下文融合模块以整合编码器多层级特征,利用注意力机制强化关键区域表达;在训练阶段,引入了解码器分级监督策略优化损失函数计算;最后通过并行分割头同步输出海冰密集度、发展阶段及浮冰尺寸预测。在 AI4Arctic 数据集上的试验表明,所提模型综合评分达 84.014,在所有的 3 种预测任务上的表现均优于 Swin Transformer、SegFormer 等主流方法,实现了海冰分布的精细分割,其成果可为后续识别并划定可通航水域范围提供技术支撑。所获分割结果在北极航道规划与船舶航行安全评估中具有重要的应用价值。

关键词:航行安全;海冰分割网络;状态空间模型;混合架构;注意力机制;损失函数优化
中图分类号:U697 **文献标志码**:A **DOI**:10.3969/j.issn.1000 - 4653.2026.04.018

A residual-state space model hybrid architecture for arctic sea ice segmentation

Yu Jicheng, *Qu Jiaming**, *Hao Yuchi*
(CCCC National Engineering Research Center of Dredging Technology and Equipment Co., Ltd., Shanghai 200082, China)

Abstract: To address the challenge of simultaneously extracting local and global features in multi-task segmentation of Arctic sea ice, a hybrid architecture integrating residual networks and state space models was proposed in this paper. In this architecture, residual networks were employed as the encoder to extract multi-scale local features. Enhanced Visual State Space (EVSS) modules were adopted in the decoder to achieve global context modeling through the combination of 2D Selective Scan and dual attention mechanisms. In the skip connections, Cross-Scale Contextual Fusion (CSCF) modules were designed to integrate multi-level features from the encoder. Attention mechanisms were utilized to enhance the representation of key regions. In the training phase, a Decoder Stage-Wise Supervision (DSWS) was introduced to optimize the calculation of loss functions. Finally, sea ice concentration, development stages, and floe size predictions were generated simultaneously through parallel segmentation heads. Experiments on the AI4Arctic dataset demonstrate that the proposed model achieves a comprehensive score of 84.014. Its performance exceeds that of mainstream methods such as Swin Transformer and SegFormer in all three prediction tasks. Fine-grained segmentation of sea ice distribution is achieved in this paper. The results provide technical support for subsequent identification and delimitation of navigable water areas. The obtained segmentation results hold significant application value in Arctic route planning and ship navigation safety assessment.

收稿日期:2025-11-14
基金项目:国家重点研发计划(2023YFC3208500);广西科技重大专项(AA23023016)资助
第一作者:郁纪成(2000—),男,助理工程师,研究方向为深度学习与海洋遥感,(E-mail)yujicheng@ ccccltd. cn
* **通信作者**:曲嘉铭(1994—),女,工程师,研究方向为深度学习与海洋环境预测,(E-mail)qujiaming@ ccccltd. cn
引用格式:郁纪成,曲嘉铭,郝宇驰. 基于残差-状态空间模型混合架构的北极海冰分割网络[J]. 中国航海,2026,49(4):164 - 174.
Yu J C, Qu J M, Hao Y C. A residual-state space model hybrid architecture for arctic sea ice segmentation[J]. Navigation of China, 2026, 49(4):164 - 174. (in Chinese)

Key words: navigation safety; sea ice segmentation network; state space model; hybrid architecture; attention mechanism; loss function optimization

北极海冰覆盖范围的持续缩减已成为全球气候变化的显著标志。研究数据显示,过去 40 a 间北极夏季海冰面积每 10 a 减少 12.5%[1],直接影响极地反照率与热量交换过程,加剧了北极地区的增温效应[2]。同时由此导致的北极航道通航窗口期的逐年延长,使进入北极的船舶数量较 10 a 前增长近 40%[3],这对海冰冰情的监测提出了更高要求。海冰分割的核心目的之一在于通过实时识别冰区边界确定可通航水域,为极地航运提供准确的导航与避险依据。能准确地识别海冰密集度、发展阶段和浮冰尺寸等多维度参数,不仅帮助航道规划者优化航线设计,也为船舶避险提供重要参考。

早期的遥感图像海冰分割方法主要依赖人工特征提取和传统机器学习分类器的结合。例如基于灰度共生矩阵(Gray-level Co-occurrence Matrix, GLCM)的纹理特征被广泛用于区分海冰和开放水域[4-5],随后利用支持向量机(Support Vector Machines, SVM)[6-7]、随机森林(Random Forest)[8-9]等分类器进行像素级分类。然而这类方法过度依赖人工特征提取与样本选择,处理流程效率较低,难以应对遥感大数据下复杂多变的海冰环境,使其在分类精度与泛化能力方面存在一定局限性[10-12]。

随着深度学习的发展,卷积神经网络(Convolutional Neural Networks, CNN)逐渐成为海冰图像分割的主流技术框架。以 ResNet、U-Net 和 DeepLabv3 + 为代表的网络结构通过局部特征提取,在海冰识别任务中取得了显著进展[13-15]。例如 Feng 等[16]基于 DeepLabv3 + 构建双分支编码结构,分别处理不同来源的遥感数据,并引入自适应多任务损失加权机制,实现北极区域内海冰的高精度识别。尽管 CNN 在局部模式识别方面表现优异,但其感受野受限于网络结构,难以充分捕捉图像中存在的长程依赖与多尺度信息,所以在复杂海冰场景下的表现仍存在不足[17]。

近年来,全局特征建模技术为解决上述问题提供了可能。Transformer 架构通过自注意力机制实现全局的上下文感知,在自然语言处理和计算机视觉任务中展现出强大的性能[18-19]。例如 Ma 等[17]提出 MFGC-Net 模型,引入基于多头自注意力的上下文 Transformer 模块,有效建模合成孔径雷达(Synthetic Aperture Radar, SAR)海冰图像中的远程依赖关系,提升多尺度全局上下文信息的提取能力。然而其计算复杂度随序列呈平方增长,制约高分辨率 SAR 图像处理效率[20]。为解决该问题,Gu 等[21]提出的 Mamba 作为一种新型的状态空间模型(State Space Model, SSM),通过改进的状态更新机制,能在保持线性计算复杂度的同时有效捕捉长距离依赖信息,适用于大规模数据处理。Liu 等[22]则进一步将 Mamba 扩展至视觉任务领域,通过多方向的扫描策略以更好地适配二维图像结构。在遥感场景中,Zhu 等[23]将 SSM 引入语义分割任务,其编码器-解码器架构在多个遥感数据集上刷新性能记录,验证 SSM 在处理遥感图像方面的独特优势。Zhao 等[20]和 Chen 等[24]则优化 Mamba 的扫描方式,在遥感图像分类与分割任务中表现出优于 Transformer 的精度与效率。

在此基础上,研究者开始探索融合 CNN 的局部特征提取能力与 Mamba 的全局建模能力的混合架构,并将其成功应用于医学图像处理领域,显著提升图像分割性能。得益于该架构的优势,一系列以 Mamba 为核心的 UNet 混合模型[25-28]在多类器官数据中取得显著的精度提升,验证了其可行性。近期,该架构在遥感领域亦取得显著进展:Wang[29]通过设计密集空间金字塔池化与选择性状态空间解码器,有效提升城市地物分割的精度与效率;Ma 等[30]构建 CNN 主支与 Mamba 辅支的双编码器结构,利用其互补特性增强对复杂城市场景的全局感知和局部细节捕捉能力;Zheng 等[31]则引入边缘引导机制,专注于提升各类农田边界的分割准确性。这些研究充分证明了 CNN-Mamba 混合架构在应对多尺度、复杂场景遥感问题时的优越性。

经过对领域内公开文献的梳理可知,医学图像的成功应用验证了 CNN 与 Mamba 混合架构(如器官边界清晰、纹理相对均匀),还使用在城市或农业遥感(地物结构规则、空间上下文稳定)等场景。相比之下,北极海冰环境具有高度的动态性和复杂性,具体表现为:海冰纹理随温度、风速和洋流作用持续变化,导致冰水边界模糊且形态不规则;冰型与浮冰尺度差异显著,呈现突出的多尺度结构与语义层次;加之极区监测以 SAR 为主,在极地气象干扰与昼夜光照差异影响下噪声增强,且需与被动微波、再分析气象等多源信息在异尺度上对齐,使模型的长程依赖建模与多尺度特征融合等能力面临更为艰巨的考验。而在海冰遥感领域,当前主流方法仍集中于 CNN 或 Transformer 的混合设计,CNN-Mamba 混合模型在北极海冰分割领域的研究尚属空白,其独特

的“局部-全局”协同机制与线性计算优势在极地遥感多任务场景中的潜力亟待探索。

基于此,本文提出一种基于残差-状态空间模型混合架构的北极海冰多任务分割网络。在 AI4Arctic[32] 数据集上的试验结果表明,该方法较现有模型在综合指标及 3 项子任务中均处于领先地位,填补了该架构在极地遥感关键场景中的应用空白。其主要贡献如下:

1)提出增强型视觉状态空间模块,利用二维选择性扫描模拟多方向长程依赖,结合通道与空间双注意力机制,增强关键特征的表达能力,解决了传统卷积网络全局建模能力不足的问题。

2)设计跨尺度上下文融合模块,利用多分支卷积整合编码器多层级特征,通过多分支卷积与注意力加权策略整合编码器不同层级的特征,抑制背景噪声并聚焦关键区域,缓解了跳跃连接中浅层和深层特征的语义差异导致的边界模糊问题。

3)引入解码器分级监督策略,通过在多个解码阶段叠加辅助输出参与损失函数计算,优化梯度传播,提升模型对多尺度海冰特征的敏感度。

1 方 法

1.1 整体架构概述

本文提出的北极海冰多任务分割模型采用编码器-解码器架构,其核心设计融合增强型视觉状态空间模块(Enhanced Visual State Space,EVSS)、跨尺度上下文融合模块(Cross Scale Contextual Fusion,CSCF)和解码器分级监督(Decoder Stage Wise Supervision,DSWS),见图 1。编码器基于 ResNet18 骨干网络,通过 4 个阶段分层提取特征,逐步扩大感受野并压缩空间分辨率,生成包含局部纹理和全局语义的多层级特征。解码器由堆叠的 EVSS 模块构成,结合 CSCF 模块引入的跨尺度跳跃连接,逐步恢复空间分辨率并融合编码器传递的多尺度信息。模型末端通过多任务分割头并行输出海冰密集度(Sea Ice Concentration,SIC)、海冰发展阶段(Stage of Development,SOD)和浮冰(FLOE)尺寸的预测结果,并引入 DSWS 机制的设计来优化损失函数计算。

1.2 EVSS

EVSS 是解码器中的核心单元,模型整体架构见图 2。设输入 EVSS 的特征为 $X \in R^{B\times H\times W\times C}$,其首先通过一个下降维度的 1×1 线性层把通道数映射到中间维度 C_{mid},然后对得到的 X_p 做层归一化得到 X_{norm}。该步骤通过降维不仅减少计算量,还保留关键信息;层归一化则用于稳定训练过程,提升模型的

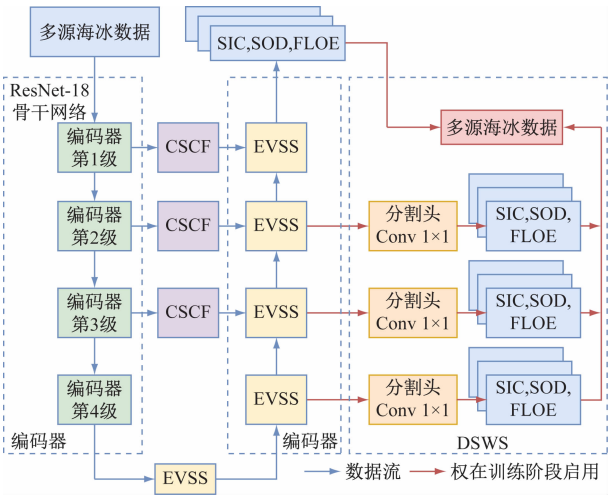


图 1 模型整体架构

Fig. 1 Overall model architecture

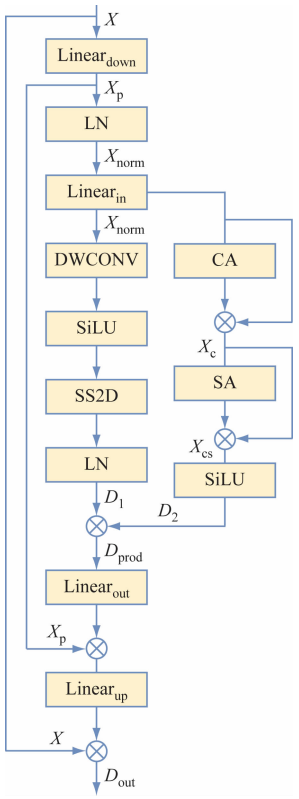


图 2 EVSS 模块结构

Fig. 2 Structure of the EVSS block

收敛性,具体表达式为

$$X_p = \text{Linear}_{\text{down}}(X),\ X_p \in R^{B\times H\times W\times C_{\text{mid}}} \quad (1)$$

$$X_{\text{norm}} = \text{LN}(X_p) \quad (2)$$

X_{norm} 会被 $\text{Linear}_{\text{in}}$ 层拆分,同时进入两条分支。

1)第一条分支先利用深度可分离卷积提取每条通道上的局部空间特征,处理完毕后,所得的局部特征会被送入二维选择性扫描(2D Selective Scan,SS2D)[22] 中。SS2D 在二维平面上模拟多方向的递归扫描,并将来自不同方向的输出合并为一个带有

全局上下文的特征图。再次经过层归一化后,此分支的最终输出定义为 D_1,有

$$D_1 = \mathrm{LN}(\mathrm{SS2D}(\mathrm{SiLU}(\mathrm{DWConv}(X_{\mathrm{norm}})))) \quad (3)$$

式(1)~式(3)中:$\mathrm{Linear}_{\mathrm{down}}$ 为下降维度的线性;B 为批量大小,即每批次输入的样本数量;H 和 W 分别为特征图的高度和宽度;C_{mid} 为经 $\mathrm{Linear}_{\mathrm{down}}$ 降维后的中间通道;LN 为层归一化;DWConv 为深度可分离卷积;SiLU 为激活函数;SS2D 为二维选择性扫描模块。该分支的核心设计在于通过深度可分离卷积高效提取局部特征,再借助 SS2D 机制实现全局上下文建模,以线性计算量捕捉长距离依赖,从而在保持计算效率的同时增强模型对复杂海冰场景的全局信息感知能力。

2) X_{norm} 还会并行地进入一个包含双注意力的分支。

①该分支先通过一个通道注意力模块 CA,用于学习不同特征通道的重要性,并对通道进行加权。

②输出 X_{c} 再经过一个空间注意力模块 SA,识别并增强空间上的关键区域信息。

③输出 X_{cs} 通过 SiLU 激活得到这个双注意力分支的输出 D_2,也就是对通道和空间位置重要性的加权特征,有

$$X_{\mathrm{c}} = \mathrm{CA}(X_{\mathrm{norm}}) \cdot X_{\mathrm{norm}} \quad (4)$$

$$X_{\mathrm{cs}} = \mathrm{SA}(X_{\mathrm{c}}) \cdot X_{\mathrm{c}} \quad (5)$$

式(4)和式(5)中:·为元素级乘法。该分支通过通道注意力突出重要特征维度,空间注意力则聚焦于图像中的关键区域,二者的结合有效增强了模型对多尺度、多位置特征的表达能力,尤其适用于多任务海冰分割任务中冰水边界的精细化识别。

以上两条分支的输出会进行逐元素相乘。相乘后的融合特征 D_{prod} 再通过一个线性层 $\mathrm{Linear}_{\mathrm{up}}$ 调整通道数,得到最终的特征表示。最后通过一个残差连接,将 EVSS 模块最初的输入与经过上述一系列处理后的输出相加,形成该模块的最终输出 D_{out},有

$$D_{\mathrm{prod}} = D_1 \cdot D_2 \quad (6)$$

$$D_{\mathrm{out}} = \mathrm{Linear}_{\mathrm{up}}(X_p + \mathrm{Linear}_{\mathrm{out}}(D_{\mathrm{prod}})) + X \quad (7)$$

式(6)和式(7)中:$\mathrm{Linear}_{\mathrm{out}}$ 为线性投影层,用于恢复被拆分为两路的输入特征的维度;逐元素相乘旨在使来自两条分支的信息深度融合,从而使局部细节与全局上下文得以互补;残差连接则促进了梯度传播,缓解深层网络的训练困难,进一步提升模型的表达能力和稳定性。

1.3 CSCF

为在网络的编码器和解码器之间有效地传递和整合不同尺度的特征信息,引入 CSCF,其模块结构见图 3。设编码器第 i 级经过通道映射后的特征为 $E_i \in R^{B\times H_i\times W_i\times C_i}$。CSCF 模块接收来自编码器多个不同分辨率的调整后特征 T_j 作为输入。这些特征图在空间维度上可能存在差异,因此,在融合前需通过插值操作将其调整至统一的空间分辨率。对齐后的多尺度特征 $\tilde{T}_{j\to i}$ 随后在通道维度上进行拼接,形成一个包含了多个尺度信息的聚合特征 U_i。这个聚合特征随后会通过一个 1×1 卷积进行初步融合与降维。具体表达式为

$$T_i = \mathrm{Conv}_{1\times1}(E_i) \quad (8)$$

$$\tilde{T}_{j\to i} = \mathrm{Interp}(T_j, H_i, W_i), \quad j > i \quad (9)$$

$$U_i = \mathrm{Conv}_{1\times1}(\mathrm{Concat}(T_i, \tilde{T}_{i+1\to i}, \tilde{T}_{i+2\to i})) \quad (10)$$

式(8)~式(10)中:$\mathrm{Conv}_{1\times1}$ 为 1×1 卷积;Interp 为插值操作;Concat 为拼接操作。

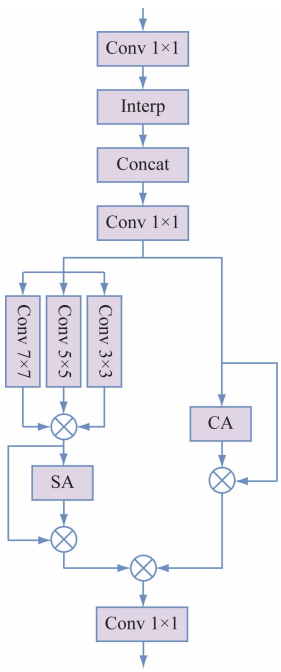


图 3 CSCF 模块结构

Fig. 3 Structure of the CSCF block

为进一步增强特征的表达能力,CSCF 也集成注意力机制。通道注意力和空间注意力机制对经过多分支卷积处理的特征进行动态加权,使模块能自动强调那些真正与海冰分割相关的通道和区域并抑制背景噪声。上一步被整合后的特征 U_i 会并行地进入通道注意力分支和多尺度融合-空间注意力分支。

经过通道注意力分支的输出 U_{c} 为

$$\mathrm{CA}(Y) = \sigma(\mathrm{FC}(\mathrm{GAP}(Y)) + \mathrm{FC}(\mathrm{GMP}(Y))) \quad (11)$$

$$U_{\mathrm{c}} = U_i \cdot \mathrm{CA}(U_i) \quad (12)$$

式(11)和式(12)中:GAP 和 GMP 分别为全局平均池化和全局最大池化;FC 为全连接;σ 为 Sigmoid 激

活函；·为元素级乘法。

进入多尺度融合-空间注意力分支的特征是首先进入一个包含多分支卷积路径的并行单元。每个分支采用不同大小的卷积核 $\mathrm{Conv}_{k\times k}$，从不同感受野尺度捕捉特征信息。随后这些多分支卷积的输出将被结合得到 S_{sum} 并输入下一层。

$$S_k = \mathrm{Conv}_{k\times k}(U_i),\ k \in 3,5,7 \tag{13}$$

$$S_{\mathrm{sum}} = S_3 + S_5 + S_7 \tag{14}$$

S_{sum} 经过空间注意力机制增强得到 U_{s} 为

$$\mathrm{SA}(Y) = \sigma(\mathrm{Conv}_{7\times 7}(\mathrm{Concat}(\mathrm{Mean}_{\mathrm{ch}}(Y), \mathrm{Max}_{\mathrm{ch}}(Y)))) \tag{15}$$

$$U_{\mathrm{s}} = S_{\mathrm{sum}} \cdot \mathrm{SA}(S_{\mathrm{sum}}) \tag{16}$$

式(15)和式(16)中：$\mathrm{Mean}_{\mathrm{ch}}$ 和 $\mathrm{Max}_{\mathrm{ch}}$ 分别为沿通道维度的平均和最大池化。

通道增强特征 U_{c}、多尺度和空间增强特征 U_{s} 相加后，通过一个 1×1 卷积调整通道数，得到 CSCF 的输出 C_{out} 为

$$C_{\mathrm{out}} = \mathrm{Conv}_{1\times 1}(U_{\mathrm{c}} + U_{\mathrm{s}}) \tag{17}$$

通过在编码器的不同阶段部署 CSCF 作为跳跃连接，提供给解码器的对应层级，可确保解码器在进行上采样和特征恢复时，能融合来自编码器不同深度的特征图，形成上下文感知能力更强的特征表示，以缓解编码器不同阶段特征语义层次差异造成的上下文信息割裂。如此，即使面对多变的海冰纹理和背景干扰，也能让分割结果与边界更加准确。

1.4 多任务输出与 DSWS

为实现对北极海冰多种参数的同步分割，本模型采用多任务输出的策略，如图 1 右侧部分所示。解码器最终会产出一个高分辨率特征图，该特征图随后被送入多个并行的分割头。每个分割头由一个 1×1 卷积层构成，负责将共享的特征表示映射到 SIC、SOD 和 FLOE 的预测概率图上。每个分割头独立地为相应的任务生成像素级分类结果。

为减轻梯度消失问题以进一步提升模型性能，本模型在损失函数计算过程中实现 DSWS 的机制。在解码器的第二级～第四级 EVSS 的输出特征经过通道维度的重排后，分别通过专用的辅助卷积层进行处理，生成与主输出特征相同任务维度的辅助预测图。这些辅助输出分支仅在训练阶段启用，对训练过程提供额外的监督，与最终输出共同参与损失函数的计算。这一设计使网络不必等到反向传播到最后一层才获得学习信号，而是在多个解码阶段的同时接收不同尺度的特征信息，提升模型对多尺度海冰特征的捕捉能力。

损失函数的计算定义如下，记第 i 级解码器输出特征为

$$F_i \in R^{B\times H_i\times W_i\times C_i},\quad i = 0,2,3,4 \tag{18}$$

式(18)中：$i=0$ 对应最终输出，$i=2,3,4$ 分别对应解码器的 3 层中间输出。

每个分割头都是一个 1×1 卷积，将通道数映射到对应任务类别数 C_t，再经 SoftMax 归一化得到像素级概率分布 $P_{t,i}$ 为

$$P_{t,i} \in R^{B\times C_t\times H_i\times W_i},\quad t \in \mathrm{SIC},\mathrm{SOD},\mathrm{FLOE} \tag{19}$$

主输出采用交叉熵损失，具体定义 L_0 为

$$L_0 = \sum_t \omega_t \frac{1}{H_0W_0}\sum_{n=1}^{H_0W_0} -\ln(P_{t,0}[n, y_{t,n}]) \tag{20}$$

式(20)中：ω_t 为任务权重；$y_{t,n}$ 为像素 n 的真实类别索引。

相似地，对于第 i 级辅助输出，其交叉熵损失 L_i 定义为

$$L_i = \sum_t \omega_t \frac{1}{H_iW_i}\sum_{n=1}^{H_iW_i} -\ln(P_{t,i}[n, y_{t,n}]),\quad i = 2,3,4 \tag{21}$$

将主输出和辅助输出的交叉熵损失加权合并得到最终的解码器分级监督损失 L_{DSWS} 为

$$L_{\mathrm{DSWS}} = L_0 + \sum_{i=2}^{4} \lambda_i L_i \tag{22}$$

式(22)中：λ_i 为在第 i 级施加的辅助监督权重；λ_2、λ_3、λ_4 按照试验经验分别调整为 0.20、0.15、0.10。

2 试验与结果

2.1 数据集

本文采用 AI4Arctic[32] 海冰遥感数据集作为试验基准，该数据集覆盖格陵兰岛与加拿大北极区域 2018—2021 年的观测数据。数据集大小约 50 GB，包含 513 个训练样本和 20 个测试样本，每个样本以 NetCDF 格式存储多维遥感参数与冰图标签。

输入数据由多源的遥感观测与气象参数构成：

1）Sentinel-1 双极化（HH/HV）SAR 图像，经 NERSC 算法噪声校正后分辨率为 80 m。

2）AMSR2 被动微波辐射计数据，包含 18.7/36.5 GHz 双极化亮温观测，分辨率为 2 km。

3）ERA5 再分析数据包含 10 m 风速、2 m 气温、总水汽量和液态云水含量这些参数，经高斯插值统一至 2 km 空间分辨率。

4）辅助数据包含海岸距离图、经纬度网格和观测时间编码等。所有输入参数均归一化至[-1,1]区间，缺失值以 -1 填充，经纬度信息经归一化后融

入模型以增强空间感知能力。为提升试验效率,本文选用 RTT 预处理版本,该版本已对原始数据进行噪声校正、空间对齐和格式统一处理。

标签数据基于丹麦气象局和加拿大冰勤署的冰图产品,包含 3 类分割任务:

1) SIC,按 10% 间隔划分为 0 ~ 100% 共 11 类。

2) SOD,包含开放水域、新冰、年轻冰、薄一年冰、厚一年冰和多年冰 6 类。

3) FLOE,涵盖开放水域、碎冰、小/中/大/巨型浮冰和冰山等 7 类。训练阶段将原始训练集随机划分为 492 个训练样本和 20 个验证样本,测试集采用官方提供的 20 个无标签样本,确保模型评估的客观性。

本文将数据集中基于官方冰图所产生的标签数据作为评测所用的真实值。该标签数据来源于丹麦气象局与加拿大冰勤署的分析专家手工绘制的业务化冰图产品。制图过程基于多源卫星观测数据,其中 SAR 影像因具有高分辨率及全天时全天候成像能力成为主要数据源,并辅以光学影像、被动微波辐射计等数据支持。冰情分析专家通过专业地理信息系统(Geographic Information System,GIS)平台,依据季节、SAR 影像中的海冰纹理和形态等特征进行人工判断,标注出海冰状况相对均一的多边形区域,并为每个多边形分配遵循世界气象组织海冰代码的标准化海冰参数信息。最终,这些包含每个多边形 ID 及其对应海冰参数(SIC、SOD、FLOE)的矢量冰图被转换为与 SAR 影像尺寸对齐的 NetCDF 格式网格化标签数据。这些冰图标签产品的本质是基于专家经验的解释性成果,在内部冰区细节及中间浓度区间可能存在一定的主观性和不确定性,并非绝对意义上的"真值",但仍被国际认可为当前业务化海冰监测中最具权威性的参考数据,多年来广泛应用于极地遥感与航海保障领域。本研究在此前提下将其作为训练与评估的基准参考。

2.2 训练设置

模型基于 PyTorch 2.0.1 框架搭建,试验环境配置为 Ubuntu 20.04 操作系统、CUDA 11.8 库和单张 RTX 3090 显卡。模型训练采用 AdamW 优化器,初始学习率设为0.000 1,并配合余弦退火学习率调度器动态地调整参数更新步长,以平衡训练稳定性和收敛效率。输入数据每批次包含 16 张图像,训练总周期设置为 100 轮,每轮包含 500 代迭代。

2.3 评估指标

模型性能采用3 类指标综合评估。针对 SIC 的回归预测特性,使用决定系数 R^2 衡量预测值和真实值的拟合精度,计算公式为

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_{i,\text{true}} - y_{i,\text{pred}})^2}{\sum_{i=1}^{n}(y_{i,\text{true}} - \hat{y}_{i,\text{true}})^2} \tag{23}$$

式(23)中:$y_{i,\text{true}}$为第 i 个像素的真实标签值;$\hat{y}_{i,\text{true}}$为所有像素真实值的均值;$y_{i,\text{pred}}$为第 i 个像素的模型预测值。

对于 SOD 和 FLOE 的离散分类任务,采用 F_1 分数评估,该指标通过精确率 p(Precision)与召回率 r(Recall)的调和平均数综合反映分类性能,特别适用于类别数量不均衡的遥感数据,有

$$F_1 = 2 \cdot \frac{p \cdot r}{p + r} \tag{24}$$

式(24)中:

$$p = \frac{T_\text{P}}{T_\text{P} + F_\text{P}} \tag{25}$$

$$r = \frac{T_\text{P}}{T_\text{P} + F_\text{N}} \tag{26}$$

式(25)和式(26)中:T_P 为真正例数量(True Positives);F_P 为假正例数量(False Positives);F_N 为假负例数量(False Negatives)。

最终的综合评分通过加权量化模型在多任务的整体表现,SIC 与 SOD 因在业务场景和制图过程中的重要性更高,权重均为 2/5,而 FLOE 权重为 1/5。计算时,陆地等非海冰区域的像素均被排除,确保评估聚焦于有效的目标区域。

2.4 对比试验

将所提基于残差-状态空间模型混合架构的海冰多任务分割模型与 U-Net[33]、DeepLabV3 +[34]、Swin Transformer[35]和 SegFormer[36]等 4 种主流方法在测试集的 SIC、SOD 和 LOE 等 3 个子任务上的性能进行全面对比,以验证所提方法在多任务分割中的综合表现。

不同模型在测试集上的对比结果见表 1。

1) 在 SIC 任务中,所提模型 R^2 得分为 84.197%,略高于 U-Net 的 84.176%,同时高出 SegFormer(82.449%)、Swin Transformer(80.030%)和 DeepLabV3 +(73.413%)。

2) 在 SOD 任务中,所提模型 F_1 分数达到 88.682%,对比 SegFormer 的 87.467%、DeepLabV3 +的 87.234%、Swin Transformer 的 86.535% 和 U-Net 的 83.998%,分别提升 1.215%、1.448%、2.147% 和 4.684%。

表 1 不同模型在测试集上的对比结果

Tab. 1 Comparison results of different models on the test set 单位:%

任务	U-Net	DeepLabV3 +	Swin Transformer	SegFormer	所提模型
SIC R^2	84.176	73.413	80.030	82.449	84.197
SOD F_1	83.998	87.234	86.535	87.467	88.682
FLOE F_1	71.501	70.062	71.521	73.870	74.310
综合评分	81.570	78.271	80.930	82.740	84.014

3) 在 FLOE 任务中,所提模型 F_1 分数为 74.310%,高于 SegFormer 的 73.870%、Swin Transformer 的 71.521%、U-Net 的 71.501% 和 DeepLabV3 + 的 70.062%,最大提升 4.248%。

4) 综合评分方面,所提模型以 84.014 分排名第一,领先 SegFormer (82.740%)、U-Net (81.570%)、Swin Transformer (80.930%) 和 DeepLabV3 + (78.271%)。

5 种模型在选取场景的 SIC 任务上的预测结果对比见图 4。图 4 中:DeepLabV3 + 未能有效识别场景中部的 80%、50% 和 20% 密集度海冰区域,其预测结果大面积误判为开放水域;Swin Transformer 在陆地边缘区域的预测精度略有提升,但在场景中部的密集海冰区域仍存在密集的误判,将大部分中高密集度海冰错误地归类为开放水域;SegFormer 与 U-Net 在中部海冰区域的识别表现较好,均能成功检测出主要冰区,U-Net 对各类冰与水边界的形状判断更为准确,SegFormer 在 20% 低密集度区域的识别存在缺失;所提模型在整体冰水边界划分方面与 U-Net 表现相当,但在 50%、70% 和 80% 中高密集度区域的预测准确性更优,其预测结果与参考标签的空间分布更加吻合。

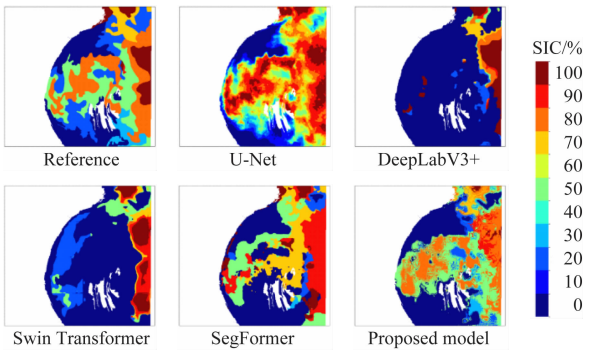


图 4 不同模型的 SIC 任务对比结果

Fig. 4 Comparison results of different models on the SIC task

5 种模型在选取场景的 SOD 任务上的分类效果差异见图 5。图 5 中:U-Net 与 DeepLabV3 + 在场景底部区域的多年冰分类中出现局部的误判,将其部分区域错误地识别为厚一年冰,但在裂缝等狭长地段的分类表现较好;Swin Transformer 的分类性能显著下降,不仅存在与 U-Net 类似的多年冰误判问题,还将裂缝区域的开放水域错误归类为厚一年冰;SegFormer 在场景底部多年冰区域的分类精度亦有所波动,部分区域被错误识别为差异更大的新冰;所提模型在该场景中展现出最佳的分类性能,其预测结果与参考标签的空间分布高度一致,成功识别出主要多年冰区域,且在裂缝等复杂地段的开放水域分类中未出现频繁误判。

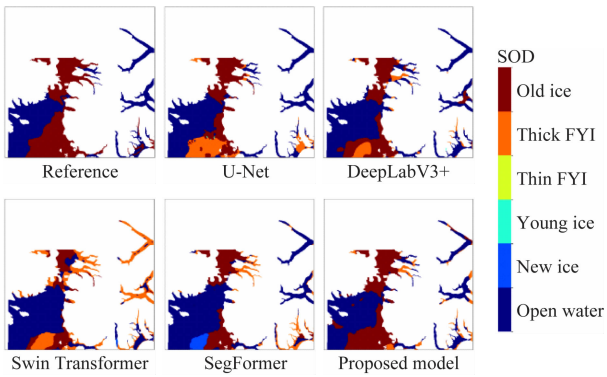


图 5 不同模型的 SOD 任务对比结果

Fig. 5 Comparison results of different models on the SOD task

5 种模型在选取场景的 FLOE 任务中的浮冰尺寸识别能力见图 6。图 6 中:U-Net 对巨型浮冰的识别存在系统性低估,将大部分该类别区域错误地归类为大浮冰;DeepLabV3 + 在浮冰边界划分方面略有改善,但仍存在对巨型浮冰的低估问题;Swin Transformer 与 SegFormer 在此任务中表现较好,都能捕捉浮冰边界的走向与空间分布规律,Swin Transformer 对巨型浮冰的识别范围更接近参考标签,SegFormer 在中型浮冰边界的划分上存在局部偏差;所提模型在此场景中再次展现出最优性能,不仅准确地识别出主导的巨型浮冰区域,还在开放水域边缘的大浮冰区域保持了较高的正确分类率,且没有出现明显的类别低估现象。

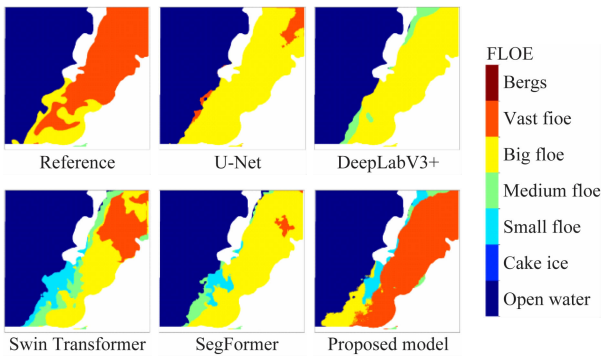


图 6 不同模型的 FLOE 任务对比结果

Fig. 6 Comparison results of different models on the FLOE task

尽管可视化结果显示各模型在关键区域(如复杂冰水边界、细碎浮冰和不同冰型过渡区)的分割效果存在显著差异,整体量化指标差异相对较小如表 1 所示。这一现象源于北极海冰遥感图像中,开阔水域或大面积连续冰区等"背景"类像素通常在数量上占据主导,而决定冰情分析精度与航行安全的关键区域虽在视觉上和实际应用中至关重要,其像素总数却占比有限。因此,基于整体像素统计的评估指标(如 R^2、F_1)易被大面积易分区域的高正确率所主导,难以充分凸显各模型在关键目标上的性能差异。因此,尽管表 1 中数值差异不大,但可视化结果清晰表明,所提模型在细节保持与边界准确性等方面具有明显优势,这对实际业务应用中的航道安全评估至关重要。

总体而言,本文模型在多任务分割中性能的全面提升,归因于其架构设计相较于对比方法的固有优势。所提模型通过 EVSS 将状态空间模型引入解码器以实现对海冰遥感场景中长程依赖的高效建模,并结合通道-空间注意力机制动态增强关键特征,从而提升了 SIC 连续值回归的精度和 SOD 分类的空间一致性;CSCF 通过多分支卷积与跨尺度注意力融合机制,有效缓解了编码器多层级特征间的语义差距大与尺度不匹配问题,增强模型对复杂冰水边界与多类型浮冰的判别能力;DSWS 机制通过多级监督优化了梯度传播路径,缓解了深层网络的训练困难。

相比之下,U-Net 虽凭借编码器-解码器对称结构与跳跃连接在局部特征保持上具有优势,但其核心卷积操作的局部性限制了模型对海冰广域连续分布特性的感知能力,难以捕捉浮冰分布中的远距离空间关联,因此在 FLOE 任务中出现对巨型浮冰的持续性低估。DeepLabV3 + 依赖的空洞空间金字塔池化模块在处理海冰这类纹理多变、形态不规则的目标时,预设的膨胀率难以自适应地捕捉极端多尺度特征,易导致区域误判或细节丢失。Swin Transformer 虽通过移动窗口机制引入一定全局性,但其注意力计算仍局限于窗口内部,在跨窗口连续性建模上存在局限。SegFormer 采用层次化 Transformer 编码器与轻量级(Multi-Layer Perceptron,MLP)解码器,在效率与性能间取得了较好平衡,但其将图像分割至小块再输入分层编码器的处理方式会破坏空间结构的连续性,且解码端完全采用轻量的线性 MLP 层,在特征融合和细粒度复原时或存在不足,在低密集度海冰、破碎冰缘区等细节区域容易产生漏判。

2.5 消融试验

为验证模型中关键组件的有效性,通过逐步引入 CSCF、EVSS 和 DSWS 等 3 个核心设计组件,构建 4 个对比模型进行消融试验。各模型在测试集上的试验性能对比见表 2。表 2 中:试验 1 为未包含本文所提出组件的基准模型,即编码器亦为残差网络构成的 Mamba-UNet[25],后续模型依次叠加新模块直至构成完整架构。

表 2 各模块消融试验对比

Tab. 2 Ablation study comparison of each module

试验	试验 1	试验 2	试验 3	所提模型
CSCF		✓	✓	✓
EVSS			✓	✓
DSWS				✓
SIC R^2/%	82.321	82.720	82.948	84.197
SOD F_1/%	85.711	87.759	88.242	88.682
FLOE F_1/%	72.642	73.536	74.340	74.310
综合评分/%	81.741	82.899	83.344	84.014

由表 2 可知:当仅添加 CSCF 时(试验 2),综合评分从 81.741% 提升至 82.899%,其中 SOD 任务的 F_1 分数提升最为显著(+2.048 个百分点),这表明跨尺度特征融合能有效增强模型对海冰发展阶段分类的判别能力。在试验中发现,在处理多年冰与厚一年冰的过渡区域时,多尺度的上下文信息能帮助网络更好地区分类别间的差异。此外 CSCF 对 FLOE 任务的提升达 0.894%,说明整合不同层级的纹理特征利于浮冰尺寸识别。

在试验 2 基础上引入 EVSS 构成试验 3 后,SIC 任务的 R^2 值提升 0.228 个百分点,SOD 任务 F_1 分数增加 0.483 个百分点,FLOE 任务 F_1 分数提升 0.804%。EVSS 通过状态空间模型捕捉长距离依

赖,配合双重注意力机制强化关键区域特征,使模型在 SIC 估算中能更精确地预测连续值分布,同时改善 FLOE 任务中对浮冰边界的判断准确度。EVSS 对 FLOE 的提升幅度超过其他模块,说明浮冰尺寸的识别还高度依赖全局空间关系的建模。

当进一步添加 DSWS 机制形成完整模型时,综合评分达到 84.014%。DSWS 对 SIC 任务的优化效果最为突出,使 R^2 值提升 1.249 个百分点,这得益于分级监督机制加强了对多尺度信息的利用,缓解深层网络的梯度消失问题,增强了对 SIC 连续值回归的表现。同时,DSWS 使 SOD 任务 F_1 分数再提升 0.440 个百分点,表明多级监督能有效保持海冰发展阶段分类的空间一致性。尽管 FLOE 任务在完整模型中略有波动(下降 0.030 个百分点),但其综合评分仍保持最高,说明模块间的协同作用能平衡多任务性能。

各模块的贡献度比较如下:CSCF 对综合评分的提升贡献最大(+1.158%),其次为 DSWS(+0.670%)和 EVSS(+0.445%);在任务维度上,CSCF 对 SOD 的提升贡献率达 68.9%,EVSS 对 FLOE 的提升贡献超过了 1/2,而 DSWS 对 SIC 的提升贡献率达 66.6%。这种差异表明:

1)跨尺度特征融合能显著改善需要细粒度分类的 SOD 任务。

2)状态空间建模更适合处理浮冰尺寸识别中的复杂空间模式。

3)分级监督机制通过优化梯度传播路径,有效提升了密集度估算的回归精度。

3 结 语

本文针对北极海冰多任务分割的需求,提出一种基于残差-状态空间模型的混合架构。通过编码器分层提取局部特征,解码器融合 EVSS 的全局状态扫描与双注意力机制,模型有效平衡了局部细节保持与长程依赖建模。CSCF 在跳跃连接中整合多层级特征信息,DSWS 机制的设计则优化了损失函数的计算。

在 AI4Arctic 数据集上的试验表明:模型在 SIC 任务中 R^2 值达 84.197%,较 U-Net 提升 0.021%;SOD 任务 F_1 分数 88.682%,领先 SegFormer 1.215 个百分点;FLOE 任务 F_1 分数 74.310%,较次优模型提升 0.440%。消融试验验证了各模块的有效性:CSCF 对 SOD 任务提升贡献率最高,EVSS 改善了浮冰边界识别,DSWS 机制则使 SIC 回归精度提升 1.249%。可视化结果证实,模型在冰水边界划分、多年冰分类和巨型浮冰识别中均展现出更优的空间一致性。

未来工作的潜在方向如下:

1)拓展模型的对比验证范围,将本文所提架构与近期发布的先进模型在极地场景下进行系统性的公平对比,以更全面地评估其性能的边界与优势。

2)更进一步探索多源遥感数据的动态融合机制,如结合被动微波与红外影像的互补特性解决 SAR 数据云雾遮挡问题。

3)探索模型对南极海冰的迁移性验证,研发适应海冰季节性变化的自训练策略。

4)开发面向混合冰型标注的小样本数据增强策略以解决现有海冰数据集的划分精度不足问题,提升复杂场景下的分割效果。

本文将基于残差-状态空间模型的混合架构应用于北极海冰多任务分割,提出的 EVSS、CSCF 和 DSWS 机制,解决了局部特征提取与全局依赖建模难以兼顾、多尺度信息利用不足和深层网络梯度消失等关键难题,为实时监测北极海冰分布并划定可通航水域提供了技术支撑。未来可将该方法集成至航运安全系统中,以提高航道利用效率和船舶通行安全性。

参 考 文 献

[1] Sticker A, Massonnet F, Fichefet T, et al. Seasonality and scenario dependence of rapid Arctic sea ice loss events in CMIP6 simulations[J]. Cryosphere, 2025, 19(8): 1-21.

[2] Parmentier F J, Christensen T R, Sørensen L L, et al. The impact of lower sea-ice extent on Arctic greenhouse-gas exchange[J]. Nature climate change, 2013, 3(3): 195-202.

[3] Chen Y, Cheng K. Decarbonizing Arctic shipping: governance pathways and future directions[J]. Frontiers in Marine Science, 2025, 12: 1-13.

[4] 朱立先,惠凤鸣,张智伦,等. 基于 Sentinel-1A/B SAR 数据的西北航道海冰分类研究[J]. 北京师范大学学报(自然科学版),2019,55(1):66-76.
Zhu L X, Hui F M, Zhang Z L, et al. Sea ice classification in Northwest Passage based on Sentinel-1A/B SAR data [J]. Journal of Beijing Normal University(Natural Science),2019,55(1):66-76. (in Chinese)

[5] Zhang L, Liu H, Gu X, et al. Sea ice classification using TerraSAR-X ScanSAR data with removal of scalloping and interscan banding [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote

Sensing,2019,12(2):589-598.

[6] Zakhvatkina N, Korosov A, Muckenhuber S, et al. Operational algorithm for ice-water classification on dual-polarized RADARSAT-2 images[J]. The Cryosphere, 2017,11(1):33-46.

[7] Liu H,Guo H,Zhang L. SVM-based sea ice classification using textural features and concentration from RADARSAT-2 dual-pol ScanSAR data[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,2015,8(4):1601-1613.

[8] 王志勇,张梦悦,于亚冉,等. 一种融合纹理特征与NDVI的随机森林海冰精细分类方法[J]. 海洋学报,2021,43(10):149-156.
Wang Z Y, Zhang M Y, Yu Y R, et al. A fine classification method for sea ice based on random forest combining texture feature and NDVI [J]. Haiyang Xuebao, 2021,43(10):149-156. (in Chinese)

[9] Park J,Korosov A A,Babiker M, et al. Classification of sea ice types in Sentinel-1 synthetic aperture radar images[J]. The Cryosphere,2020,14(8):2629-2645.

[10] Li M X,Sun Y,Zhang Q. Extraction of sea ice cover by sentinel-1 SAR based on support vector machine with unsupervised generation of training data [J]. IEEE Transactions on Geoscience and Remote Sensing,2020, 59(4):3040-3053.

[11] Markus R,Gustau C,Bjorn S, et al. Deep learning and process understanding for data-driven earth system science[J]. Nature,2019,566(7743):195-204.

[12] Chen Y, Yan Q, Bashmachnikov I, et al. MFDA: unified multi-task architecture for cross-scene sea ice classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2024, 62: 1-21.

[13] Tianyu Z, Ying Y, Mohammed S, et al. Deep learning based sea ice classification with gaofen-3 fully polarimetric SAR data [J]. Remote Sensing, 2021, 13 (8):1452.

[14] Karvonen J. Baltic sea ice concentration estimation from c-band dual-polarized sar imagery by image segmentation and convolutional neural networks [J]. IEEE Transactions on Geoscience and Remote Sensing, 2021, 60: 1-11.

[15] 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 S C, Wang Z Y, Li Z J, 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. (in Chinese)

[16] Feng T, Yang Y, Jiang P, et al. A dual-branch architecture for adaptive loss multitask mapping based on AI4Arctic sea ice challenge dataset [J]. IEEE Geoscience and Remote Sensing Letters, 2025, 22: 1-5.

[17] Ma T, Chen X, Xu L, et al. MFGC-net: bridging and fusing multiscale features and global contexts for multitask sea ice fine segmentation[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2025, 18: 14015-14035.

[18] Wolf T, Debut L, Sanh V, et al. Transformers: state-of-the-art natural language processing[C]//Proceedings of the 2020 conference on empirical methods in natural language processing: system demonstrations. [S. l.]: Association for Computational Linguistics, 2020: 38-45.

[19] Xiao H, Li L, Liu Q, et al. Transformers in medical image segmentation: a review [J]. Biomedical Signal Processing and Control, 2023, 84: 1-17.

[20] Zhao S, Chen H, Zhang X, et al. Rs-mamba for large remote sensing image dense prediction [J]. IEEE Transactions on Geoscience and Remote Sensing, 2024, 62: 1-14.

[21] Gu A, Dao T. Mamba: linear-time sequence modeling with selective state spaces [J/OL]. arXiv preprint, 2023-12-01. https://doi.org/10.48550/arXiv.2312.00752.

[22] Liu Y, Tian Y, Zhao Y, et al. Vmamba: visual state space model [J]. Advances in neural information processing systems, 2024, 37: 103031-103063.

[23] Zhu Q, Cai Y, Fang Y, et al. Samba: semantic segmentation of remotely sensed images with state space model[J]. Heliyon, 2024, 10(19):1-10.

[24] Chen K, Chen B, Liu C, et al. Rsmamba: remote sensing image classification with state space model[J]. IEEE Geoscience and Remote Sensing Letters, 2024, 21: 1-5.

[25] Wang Z, Zheng J Q, Zhang Y, et al. Mamba-unet: unet-like pure visual mamba for medical image segmentation [J/OL]. arXiv preprint, 2024-02-07. https://doi.org/10.48550/arXiv.2402.05079.

[26] Ma J, Li F, Wang B. U-mamba: enhancing long-range dependency for biomedical image segmentation[J/OL]. arXiv preprint, 2024-01-19. https://doi.org/10.48550/arXiv.2401.04722.

[27] Xing Z, Ye T, Yang Y, et al. SegMamba: long-range sequential modeling mamba for 3D medical image segmentation[C]//International Conference on Medical Image Computing and Computer-Assisted Intervention. Marrakech:MICCAI,2024: 578-588.

[28] Wang J, Chen J, Chen D, et al. LKM-UNet: large kernel vision mamba unet for medical image segmentation[C]//International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham: Springer Nature Switzerland, 2024: 360-370.

[29] Wang L, Li D, Dong S, et al. PyramidMamba: rethinking pyramid feature fusion with selective space state model for semantic segmentation of remote sensing imagery[J]. International Journal of Applied Earth Observation and Geoinformation, 2025, 144: 1-11.

[30] Ma X, Zhang X, Pun M O. Rs 3 mamba: visual state space model for remote sensing image semantic segmentation[J]. IEEE Geoscience and Remote Sensing Letters, 2024, 21: 1-5.

[31] Zheng J, Fu Y, Chen X, et al. EGCM-UNet: edge guided hybrid CNN-Mamba UNet for farmland remote sensing image semantic segmentation[J]. Geocarto International, 2025, 40(1): 1-20.

[32] Stokholm A, Buus-Hinkler J, Wulf T, et al. The AutoICE challenge[J]. The Cryosphere, 2024, 18(8): 3471-3494.

[33] Ren Y, Xu H, Liu B, et al. Sea ice and open water classification of SAR images using a deep learning model[C]//IGARSS 2020-2020 IEEE International Geoscience and Remote Sensing Symposium. New York: IEEE, 2020: 3051-3054.

[34] Chen L C, Zhu Y, Papandreou G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[C]//Proceedings of the European Conference on Computer Vision (ECCV). Munich: ECCV, 2018: 801-818.

[35] Liu Z, Lin Y, Cao Y, et al. Swin transformer: hierarchical vision transformer using shifted windows[C]//Proceedings of the IEEE/CVF international conference on computer vision. Seattle: IEEE/CVF, 2021: 10012-10022.

[36] Xie E, Wang W, Yu Z, et al. SegFormer: simple and efficient design for semantic segmentation with transformers[J]. Advances in Neural Information Processing Systems, 2021, 34: 12077-12090.

(上接第163页)

[9] 李召辉, 王博岩, 狄仲捷, 等. 基于智能调度系统的琼州海峡客滚运输班期优化[J]. 航海技术, 2025(3): 38-42.
Li Z H, Wang B Y, Di Z J, et al. Optimization of schedules for Ro-Ro passenger transportation in the Qiongzhou strait based on an intelligent scheduling system[J]. Marine Technology, 2025(3): 38-42. (in Chinese)

[10] 张新宇, 李瑞杰, 林俊, 等. 双向通航港口船舶调度优化模型与算法[J]. 中国航海, 2018, 41(2): 36-40.
Zhang X Y, Li R J, Lin J, et al. Vessel scheduling optimization in two-way traffic ports[J]. Navigation of China, 2018, 41(2): 36-40. (in Chinese)

[11] Liu B L, Li Z C, Sheng D, et al. Integrated planning of berth allocation and vessel sequencing in a seaport with one-way navigation channel[J]. Transportation Research Part B: Methodological, 2021, 143: 23-47.

[12] 李润佛, 张新宇, 李俊杰, 等. 基于强化学习的自学习遗传算法在船舶调度中的应用[J]. 大连海事大学学报, 2022, 48(3): 20-30.
Li R F, Zhang X Y, Li J J, et al. Application of self-learning Genetic Algorithm based on reinforcement learning in ship scheduling[J]. Journal of Dalian Maritime University, 2022, 48(3): 20-30. (in Chinese)

[13] Lalla-Ruiz E, Shi X N, Voss S. The waterway ship scheduling problem[J]. Transportation Research Part D: Transport and Environment, 2018, 60: 191-209.

[14] Xu F Q, Ren H X, Duan Y T, et al. Joint optimization of ship traffic organization and berth allocation in a seaport with one-way channel[J]. Ocean Engineering, 2025, 326: 120843.

[15] Jia Q, Li R H, Li J W. Departure vessel scheduling optimization considering traffic restrictions in turning basin: a case study for Xuwen terminal[J]. Journal of Marine Science and Engineering, 2023, 11(7): 1311.

[16] Jia Q, Li R H, Li J W, et al. Vessel traffic scheduling optimization for passenger RoRo terminals with restricted harbor basin[J]. Ocean and Coastal Management, 2023, 246: 106904.

[17] 贾琼. 操纵水域受限的客滚运输港口船舶调度优化研究[D]. 湛江: 广东海洋大学, 2024.
Jia Q. Vessel scheduling optimization for semi-enclosed passenger RoRo terminals with restricted harbor basin[D]. Zhanjiang: Guangdong Ocean University, 2024. (in Chinese)

[18] Zhao R Q, Shu S N, Wu L X, et al. Robust vessel traffic scheduling with uncertain berth service times in a seaport[J]. Transportation Research Part B: Methodological, 2025, 200: 103294.