

中图法分类号: TP18; TP391 文献标识码: A 文章编号: 1006-8961(2026)05-1557-12

论文引用格式: Gong Z, Zhang J L, Gao F, Gan Y H and Dong J Y. 2026. Wavelet-based multiscale residual aggregation network (WRANet) for Arctic sea ice forecasting. Journal of Image and Graphics, 31(5):1557-1568(弓政, 张家亮, 高峰, 甘言海, 董军宇. 2026. 小波多尺度残差聚合北极海冰预测网络 WRANet. 中国图象图形学报, 31(5):1557-1568)[DOI:10.11834/jig.250318]

# 小波多尺度残差聚合北极海冰预测网络 WRANet

弓政, 张家亮, 高峰\*, 甘言海, 董军宇

1. 中国海洋大学海洋动力—物理环境与智能感知全国重点实验室, 青岛 266100;
2. 中国海洋大学计算机科学与技术学院, 青岛 266100

**摘要:** 目的 针对现有北极海冰预测模型在特征提取中容易丢失高频细节信息、导致精细结构预测不准确, 以及对深度网络产生的丰富中间层特征利用不充分从而限制模型表达能力这两大难题, 提出一种新的海冰预测网络以提升预测精度, 即基于小波多尺度残差聚合的海冰预测网络(wavelet-based multiscale residual aggregation network, WRANet)。方法 首先, 设计了小波多尺度特征提取模块, 该模块融合二维离散小波变换与多尺度卷积策略, 在频域上分离并保留图像的高频细节与低频轮廓信息, 以获得更鲁棒的特征表达。在此基础上, 引入一种轻量化的像素级注意力机制, 对特征图进行动态空间校准, 以增强关键区域的特征响应。其次, 引入一种渐进残差聚合结构, 通过聚合网络在不同处理阶段输出的中间层特征, 形成一个更全面、多层次的特征表示, 以捕捉复杂的时空动态。结果 在公开的 OSI-450-a(global sea ice concentration climate data record, release 3)与 AMSR2(ASI-AMSR2 sea ice concentration)海冰数据集上, 本文将所提 WRANet 模型与 ConvLSTM(convolutional LSTM network)、SimVP(simpler yet better video prediction)等6种主流模型进行了综合对比。实验结果表明, 在 OSI-450-a 数据集上, WRANet 在所有关键评估指标上均达到最优, 其均方根误差(root mean square error, RMSE)与平均绝对误差(mean absolute error, MAE)分别低至 6.44% 和 2.02%。与当前性能领先的 SimVP 模型相比, WRANet 在 RMSE 指标上降低了 0.45%, 同时计算复杂度更低。在更高分辨率的 AMSR2 数据集上的实验同样验证了模型的卓越性能与泛化能力。此外, 消融实验定量证明了所提出的模块与结构均是提升模型预测精度的有效设计。结论 本文所提出的 WRANet 网络, 通过有效融合频域分析与多尺度特征提取, 并高效利用网络中间层特征, 成功改善了海冰时空特征的捕捉效果, 提升了模型对复杂时空依赖的建模能力, 为高精度北极海冰预测提供了一种有效的解决方案。

**关键词:** 北极海冰预测; 时空预测; 深度学习; 小波变换; 残差聚合; 海冰密集度(SIC)

## Wavelet-based multiscale residual aggregation network (WRANet) for Arctic sea ice forecasting

Gong Zheng, Zhang Jialiang, Gao Feng\*, Gan Yanhai, Dong Junyu

1. State Key Laboratory of Physical Oceanography, Ocean University of China, Qingdao 266100, China;
2. School of Computer Science and Technology, Ocean University of China, Qingdao 266100, China

**Abstract: Objective** Arctic sea ice forecasting is a critical and challenging task with profound implications for global cli-

收稿日期: 2025-07-14; 修回日期: 2025-09-25; 预印本日期: 2025-10-08

\* 通信作者: 高峰 gaofeng@ouc.edu.cn

基金项目: 科技创新 2030-新一代人工智能重大项目(2022ZD0117202); 山东省自然科学基金项目(ZR2024MF020)

Supported by: National Science and Technology Major Project of China (2022ZD0117202); Natural Science Foundation of Shandong Province, China (ZR2024MF020)

mate research, polar ecosystem protection, resource development, and international strategic interests. As a core indicator of sea ice distribution, sea ice concentration (SIC)—the fraction of a given ocean area covered by ice—provides essential parameters for studying ice-ocean-atmosphere interactions. Improving the accuracy of SIC prediction is paramount for refining global climate models, understanding the mechanisms of sea ice formation and evolution, and enhancing the early warning capabilities for extreme climate events. However, existing spatiotemporal forecasting models, particularly those based on deep learning, face two significant challenges that limit their performance. First, a substantial loss of high-frequency detail information exists. Mainstream models, particularly those reliant on convolutional neural networks, often use convolution and pooling operations that inadvertently discard fine-grained structural and textural details while reducing computational load. This loss is detrimental to the accurate prediction of intricate sea ice features, such as edges and small ice floes. Second, feature information is insufficiently utilized. Deep networks generate a wealth of rich feature representations at various intermediate layers containing crucial information at different semantic levels and scales. Relying solely on the final, most abstract feature map for prediction, many current models fail to leverage these intermediate features, thereby constraining the model's expressive power and its ability to capture complex dynamics. This study aims to develop a novel deep learning architecture for Arctic sea ice forecasting; in this way, the limitations mentioned can be overcome. The primary objective is to design a network that can robustly express spatiotemporal features by preserving multiscale information and can efficiently utilize hierarchical features from all stages of the network to capture complex dependencies and improve overall prediction accuracy. **Method** We propose the wavelet-based multiscale residual aggregation network (WRANet), which enhances a standard encoder-decoder framework with three synergistic innovations, to address the key challenges in sea ice prediction. The cornerstone of our approach is the wavelet multiscale feature extraction module, which is designed to combat the loss of fine-grained information. Instead of using standard convolutions alone, this module first employs a discrete wavelet transform to decompose the input reversibly into distinct low-frequency (coarse structure) and high-frequency (fine details) sub-bands. Then, this frequency-separated representation is processed by a specialized multiscale convolution block, which uses parallel branches with varied kernel sizes to capture local and contextual information. After processing, a lightweight pixel-wise attention mechanism dynamically recalibrates the feature map to amplify important spatial regions while suppressing noise before an inverse wavelet transform reconstructs a feature-rich spatial map. This mechanism complements our novel progressive residual aggregation structure, which tackles the problem of insufficient feature utilization in deep networks. In this structure, the output from each feature extraction module is combined not with the previous layer's output but always with the initial feature map from the encoder. This approach ensures a constant flow of foundational information and prevents its decay. Furthermore, instead of relying on only the final output, the structure aggregates all intermediate feature maps produced throughout the network, thereby creating a comprehensive, multilevel representation that captures a full spectrum of learned dynamics for the final prediction. The entire network is trained end-to-end using the AdamW optimizer. **Result** Our model was evaluated on two public SIC datasets, the OSI-450-a (global sea ice concentration climate data record, release 3) and the AMSR2 (ASI-AMSR2 sea ice concentration). Our training set covered the period from 2000 to 2010, with a validation from 2011 to 2015. We evaluated model performance using four accuracy metrics: root mean square error (RMSE), mean absolute error (MAE), Nash-Sutcliffe efficiency (NSE), and balanced accuracy (BACC), along with giga floating point operations per second (GFLOPs) to measure computational complexity. Our proposed WRANet demonstrated superior performance against six state-of-the-art baseline models, including ConvLSTM (convolutional LSTM network), PredRNNv2 (predictive recurrent neural network), SimVP (simpler yet better video prediction), TAU (temporal attention unit), WaST (wavelet-driven spatiotemporal predictive learning), and PastNet (physics-assisted spatio-temporal network). Quantitative evaluation shows that on the OSI-450-a benchmark, WRANet achieved the best results across all accuracy metrics, with the RMSE, MAE, NSE, and BACC being 6.44%, 2.02%, 97.03%, and 96.96%, respectively. These results represent a significant improvement over all baselines. For instance, on the OSI-450-a benchmark, WRANet reduced the RMSE by 2.33 percentage points compared with ConvLSTM and by approximately 0.5 percentage points compared with the highly competitive SimVP and TAU models. These accuracy gains were achieved with high computational efficiency, as WRANet required only 118.09 GFLOPs; thus, WRANet becomes more efficient and more competitive than other models. Additionally, our model demonstrated consistent

state-of-the-art performance on the AMSR2 dataset, thereby confirming its strong generalization capability and robustness across different data sources. **Conclusion** The WRANet network proposed in this paper successfully improves the capture of sea ice spatiotemporal features and enhances the model's ability to model complex spatiotemporal dependencies by effectively integrating frequency domain analysis with multiscale feature extraction and efficiently utilizing intermediate network layer features. The proposed network provides an effective solution for high-precision Arctic sea ice prediction.

**Key words:** Arctic sea ice prediction; spatio-temporal prediction; deep learning; wavelet transform; residual aggregation; sea ice concentration(SIC)

## 0 引言

北极海冰预测对全球气候研究、极地生态保护、资源开发利用以及我国国家利益等具有重要的经济价值与战略意义,已经成为国内外海洋科学领域研究的热点(Parkinson和DiGirolamo,2021;Stroeve等,2012)。受全球变暖和北极放大效应(Tepes等,2021;Pithan和Mauritsen,2014;Serreze和Barry,2011)的双重驱动,本世纪中叶前北极海冰覆盖面积或将持续呈现加速缩减态势。作为表征海冰分布特征的核心指标,海冰密集度(sea ice concentration, SIC)通过量化单位海域的冰盖占比,为研究极区冰—海—气相互作用提供了关键参数支撑。持续监测海冰密集度动态变化不仅能优化全球气候预测模型,更有助于揭示海冰生成机制与演化规律,从而提升极端气候事件的预警效能。

当前北极海冰监测体系主要依托多源遥感数据融合技术,具体包括:美国国防气象卫星计划搭载的扫描多通道微波辐射计、微波成像仪(Cavalieri等,1996)获取的微波遥感数据,欧洲中期天气预报中心(European Centre for Medium-range Weather Forecasts, ECMWF)发布的ERA5(ECMWF reanalysis v5)再分析数据集(Hersbach等,2020),以及新兴的合成孔径雷达和激光雷达观测网络。这种天地一体化的数据采集体系通过时空分辨率互补,为构建高精度海冰预测模型奠定了多维度数据基础。

在北极海冰预测模型研究领域,学界主要有统计模型、数据模型和深度学习模型3类方法。统计模型(Horvath等,2020)依靠历史数据建立回归关系,具有运算效率高的优势,但难以解析海冰热力—动力耦合机制。数值模型基于Navier-Stokes方程构建冰—海—气耦合系统,虽然能精确刻画海冰动力学过程,却受限于计算复杂度与硬件成本。深度学

习模型大多构建基于长短时记忆网络(long short-term memory, LSTM)(Hochreiter和Schmidhuber,1997)和Transformer(Vaswani等,2017)架构预测系统,通过特征自提取机制突破传统方法局限,在短期(7~15天)预测中展现出卓越的非线性建模能力,其时空注意力机制可有效捕捉海冰迁移的混沌特性。实际上,深度学习技术已在更广泛的海冰遥感分析任务中展现出巨大潜力,例如,有研究成功利用卷积神经网络(convolutional neural network, CNN)和深度置信网络(deep belief network, DBN)对合成孔径雷达(synthetic aperture radar, SAR)影像进行高精度的冰—水分类,验证了其自动学习海冰图像特征的有效性(黄冬梅等,2018)。深度学习在分类任务上的成功,为将其应用于更具挑战性的海冰时空预测任务奠定了基础。

现有时空预测模型可分为循环与非循环神经网络架构。基于循环神经网络的模型(如ConvLSTM(convolutional LSTM network)(Shi等,2015)、PredRNN(predictive recurrent neural network)(Wang等,2023))通过引入时空LSTM单元和高阶非平衡性学习机制,实现对时空信息的联合建模;基于非循环神经网络的模型(如SimVP(simpler yet better video prediction)(Gao等,2022)、TAU(temporal attention unit)(Tan等,2023))则通过时空分离架构和专用注意力机制,在计算效率与预测精度间取得平衡。这些创新架构共同推动了计算机视觉与时空分析领域的发展。

尽管当前国内外学者针对时空序列预测已开展多方向研究,同时在国内外公开数据集上已取得较好效果,但仍面临两个关键难题。1)高频细节信息损失。以PredRNNv2为代表的循环网络架构,或以TAU为代表的纯卷积网络,在通过标准的卷积和池化操作提取空间特征时,虽然有效降低了计算成本,但也往往以牺牲图像高频细节为代价。这种信息损

失对于精确预测海冰边缘、裂缝和小型浮冰等精细结构是致命的,往往导致预测结果过于平滑和模糊。这类挑战在其他计算机视觉任务中也普遍存在,在医学图像分割领域,研究人员发现连续的卷积与池化是导致精细结构分割不准确的关键因素(梅华威等,2024)。2)特征信息利用不充分。深度网络能产生层级化特征:浅层保留空间细节,深层捕捉宏观趋势。然而,许多主流编码器—解码器架构存在信息瓶颈,未能有效融合这些分层特征。以先进的非循环模型TAU为例,其解码器仅能基于时序模块输出的最终特征图进行重建,导致所有宝贵的浅层和中间特征未得到有效利用。这种因信息瓶颈造成的不可逆细节损失,使得解码器难以从高度抽象的表征中恢复精细结构,这是限制模型捕捉复杂时空动态、导致预测缺乏真实感的关键原因之一。

为应对上述挑战,本文提出了小波多尺度残差聚合北极海冰预测网络(wavelet-based multiscale residual aggregation network for Arctic sea ice forecasting, WRANet)。该网络旨在通过以下两方面的核心设计来克服现有方法的局限性:1)设计了小波多尺度特征提取模块,通过将二维离散小波变换与多尺度卷积策略进行巧妙结合,该模块能够充分利用小波变换在频域分析上的优势,显式地分离并保留图像中的高频细节与低频轮廓,从而获得更为鲁棒的特征表示;2)引入一种渐进残差聚合结构,该结构不仅通过残差连接保障梯度的高效传播与特征的逐层精炼,更关键的是,它能够有效聚合来自网络不同深度的特征信息,融合成一个更全面、多层次的表示以用于最终的预测,从而更精准地捕捉到隐藏在不同特征层级中的关键动态。

本文的主要贡献如下:1)提出小波多尺度特征提取模块:在图像特征鲁棒表达方面,创新性地融合频域分析(小波变换)与多尺度卷积,实现了对不同粒度级别特征(高频细节与低频结构)的高效提取与保留;2)设计了一种新颖的像素级空间注意力模块:该模块通过一种轻量级的相似性度量方式生成空间注意力图,能够动态地增强信息丰富区域的特征响应,同时抑制无关背景与噪声;3)设计了渐进残差聚合结构:在模型架构层面,通过聚合模型在特征提取过程中不同阶段的输出,有效利用了中间特征信息,提升了模型捕捉复杂时空依赖的能力;4)模型有效性验证:通过在公开的OSI-450-a(global sea ice con-

centration climate data record, release 3)上进行充分的实验评估,验证了本文提出的WRANet相较于现有方法的优越性和有效性。

## 1 本文方法

本文提出的WRANet架构如图1所示,包含标准的SimVP编码器、解码器,以及小波多尺度特征提取模块和渐进残差聚合结构。接下来逐步介绍网络的整体架构、编码器与解码器、特征提取模块与渐进残差聚合结构。

### 1.1 网络整体架构

如图1所示,WRANet采用编码器—解码器架构处理海冰时空序列数据。输入序列 $I \in \mathbf{R}^{B \times T \times C \times H \times W}$ 。首先通过编码器转换为特征表示,其中, $B$ 为批次大小, $T$ 为时间序列长度, $C$ 为通道数, $H$ 和 $W$ 分别为特征图高度和宽度。网络由多个交替连接的小波多尺度特征提取模块和渐进残差聚合结构组成。二维卷积层将输入特征映射至隐藏维度 $C_{hid}$ ,经过特征拼接后维度扩展为 $(C_{hid} \times 4)$ ,以增强表示能力。

特征提取模块的结构如图1(b)所示,首先对输入特征进行小波变换,然后输出到多尺度卷积模块进行特征加工,最后使用小波逆变换恢复特征。其中,多尺度卷积能够捕获不同尺度的海冰特征,增强模型对多尺度特征的建模能力。最后,使用像素级注意力进一步处理,增强模型对关键特征的处理能力。

图2展示了多尺度卷积模块,采用并行与串行结合的设计增强模型对多尺度变化的感知能力。其中,每一个Conv2D都是Conv2D+instancenorm2D+GeLU(Gaussian error linear unit)的组合。图3展示了像素级注意力机制,通过残差连接强化重要区域的特征表示。最终,处理后的特征通过解码器重建为与输入相同维度的预测结果。

### 1.2 编码器与解码器

本文选择了与SimVP网络结构相同的编码器与解码器模块。编码器用于对图像进行下采样,而解码器则用于对图像进行上采样,从而恢复图像。编码器由4个Conv2D,GroupNorm2D(Wu和He,2018)与SiLU(sigmoid linear unit)激活函数(Elfwing等,2018)堆叠而成,通过步长卷积逐步对空间维度进行

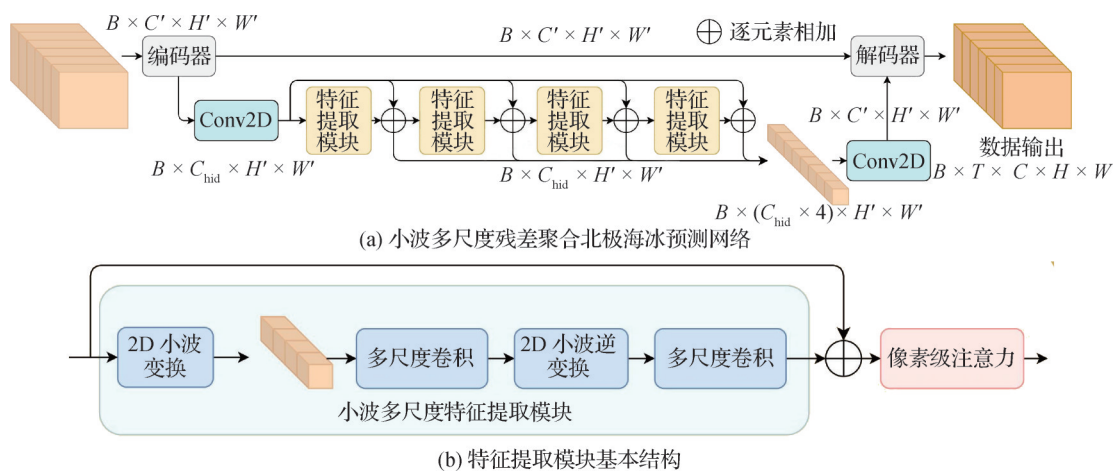


图1 小波多尺度残差聚合北极海冰预测网络 (WRANet) 整体架构

Fig. 1 The overall architecture of WRANet ((a) WRANet; (b) basic structure of the feature extraction module)

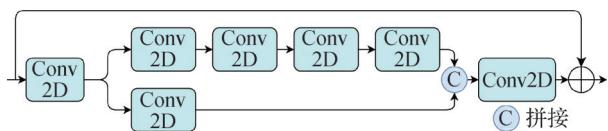


图2 多尺度卷积模块结构图

Fig. 2 Multi-scale convolution module structure diagram

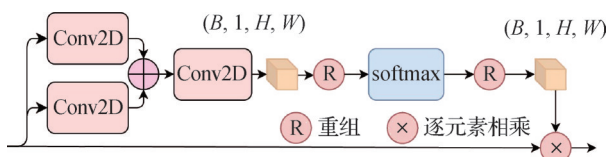


图3 像素注意力模块结构图

Fig. 3 Pixel-wise attention module structure diagram

下采样,提取层次化特征。相对地,解码器也包含4个堆叠的 Conv2D, GroupNorm2D 与 SiLU 激活函数,利用像素重组进行上采样,以恢复特征图的空间分辨率。

输入视频张量  $(B, T, C, H, W)$  首先融合批次与时间维度,转换为四维结构  $(B \times T, C, H, W)$  进行空间特征编码。随后再将编码特征的时间维度与通道维度融合,形成  $(B, T \times C', H', W')$  表示,这可以使得二维卷积在空间计算中隐式地捕获时序动态模式。最终通过逆向维度转换重建标准五维输出,实现时空特征的有效解耦与重建。该策略在保留二维卷积计算优势的同时,建立了通道维度与时间信息的参数化映射机制。具体为

$$X_{hid}, skip = encoder(X) \quad (1)$$

$$X_{hid'} = extraction(X_{hid}) \quad (2)$$

$$Y = decoder(X_{hid'}, skip) \quad (3)$$

式中,  $X, Y \in \mathbf{R}^{B \times T \times C \times H \times W}$  表示模型的输入与输出特征,  $skip$  变量是编码器中间层特征,  $X_{hid} \in \mathbf{R}^{B \times C_{hid} \times H' \times W'}$  表示模型的隐藏特征,  $encoder(\cdot)$  表示编码器,  $decoder(\cdot)$  表示解码器,  $extraction(\cdot)$  表示中间一系列特征提取模块。

### 1.3 小波多尺度特征提取模块

小波变换是一种强大的图像分析工具,通过多尺度分析弥补了传统信号处理方法在空间局部性分析上的不足。它能够将图像信号分解为不同频率和方向的子带,从而在不同尺度上进行局部化分析。具体而言,二维离散小波变换(2D-DWT)通过对图像的行和列分别应用低通和高通滤波器,可逆地将原始特征图分解为4个具有明确物理意义的子带:1个低频近似子带 LL(low-low)以及3个高频细节子带 LH(low-high)、HL(high-low)和 HH(high-high)。其中,LL子带保留了图像的整体轮廓和概貌信息(即粗粒度特征),而LH、HL和HH子带则分别捕捉了水平、垂直和对角线方向的边缘与纹理细节(即细粒度特征)。小波变换能够显式分离并保留图像的粗、细粒度信息。鉴于此独特优势,本文利用小波变换设计多尺度特征提取模块。这一思想也与近期在其他领域的研究相呼应,例如,LCTNet(lightweight CNN-Transformer combined network)(侯志强等,2025)将小波变换与卷积网络结合,设计了一种轻量级的实时语义分割网络,并验证了其在平衡局部细节和全局上下文信息方面的有效性。

如图1(b)所示,给定一个特征图  $I \in \mathbf{R}^{B \times C \times H \times W}$ , 会先使用2D离散小波变换将原始的特征图分割成

4份可逆的特征图  $\tilde{I} \in \mathbf{R}^{B \times C \times \frac{H}{2} \times \frac{W}{2}}$ 。此时,通道数不变,而特征图的尺寸变成原来的1/4。

本文选择了Haar小波函数,它是一种正交小波,具有紧支撑性。对于2D Haar小波,它通过先对图像的行向进行一维变换,再对结果的列向进行一维变换来实现。

对于输入的特征图  $A \in \mathbf{R}^{B \times C \times H \times W}$ ,会先使用2D Haar小波变换将特征图分解成4个小波子带:  $X_{LL}$ 、 $X_{LH}$ 、 $X_{HL}$ 和 $X_{HH}$ 。具体为

$$\{X_{LL}, X_{LH}, X_{HL}, X_{HH}\} = DWT(A) \quad (4)$$

式中,  $X_i \in \mathbf{R}^{B \times C \times \frac{H}{2} \times \frac{W}{2}}$ 。之后再在通道维度上做拼接,得到了拼接后的特征图  $Y$ 。具体为

$$Y = Concat([X_{LL}, X_{LH}, X_{HL}, X_{HH}], dim = 1) \quad (5)$$

式中,  $Y \in \mathbf{R}^{B \times 4C \times \frac{H}{2} \times \frac{W}{2}}$ 。  $Concat([\cdot, \cdot], dim = 1)$ 表示沿通道维度进行连接。再将拼接后的结果送入多尺度卷积模块进行更深层次的特征提取,得到  $Y'$ 。这里卷积的组数为4,从而保证不同子带的特征在各自的通道组内进行处理,不会相互影响。  $Multi-Scale-Conv()$ 为多尺度卷积模块,具体为

$$Y' = Multi-Scale-Conv(Y) \quad (6)$$

处理以后,再将  $Y'$ 恢复成空间域的  $X'$ 特征图,具体为

$$\{X'_{LL}, X'_{LH}, X'_{HL}, X'_{HH}\} = Split(Y') \quad (7)$$

$$X' = IDWT(X'_{LL}, X'_{LH}, X'_{HL}, X'_{HH}) \quad (8)$$

式中,  $X'$ 与  $X$ 的形状一样,都是  $X' \in \mathbf{R}^{B \times C \times H \times W}$ 。为了进一步提取深层特征,模块将IDWT的输出特征  $X'$ 送入一个新的多尺度卷积层处理,然后再与原始输入的特征相加。最后,应用一个像素级注意力机制自适应地调整特征响应,生成最终的输出特征图。

#### 1.4 多尺度卷积

多尺度卷积模块通过并行处理输入特征提取多尺度信息。输入首先经过一个  $1 \times 1$ 卷积进行通道变换和初步特征提取,具体为

$$X_{init} = f_{1 \times 1}(X) \quad (9)$$

式中,  $X_{init} \in \mathbf{R}^{B \times C_{int} \times H \times W}$ ,  $f_{1 \times 1}$ 表示一个卷积核大小为  $1 \times 1$ 的二维卷积。随后,将变换后的特征送入两个并行分支:一个分支采用大核接小核再接小核接大核(L→S→S→L)的卷积序列,这种设计旨在先捕捉较粗糙的特征(大核),然后细化(小核),再进行细化(小核),最后再次整合较大范围的上下文信息(大

核),从而探索不同尺度特征之间的复杂交互;另一个分支则采用连续的小核卷积(S→S),专注于提取局部精细特征,具体为

$$X_{b1} = f_L(f_S(f_S(f_L(X_{init})))) \quad (10)$$

$$X_{b2} = f_S(f_S(X_{init})) \quad (11)$$

式中,  $f_L$ 表示大卷积核的二维卷积,  $f_S$ 表示小卷积核的二维卷积。将两个分支的输出在通道维度上进行连接,并通过最终的  $1 \times 1$ 卷积进行特征融合,将通道数恢复。具体为

$$X_{concat} = Concat([X_{b1}, X_{b2}], dim = 1) \quad (12)$$

$$Y_{fused} = f_{1 \times 1}(X_{concat}) \quad (13)$$

$$Y = Y_{fused} + X_{init} \quad (14)$$

#### 1.5 像素级注意力

像素级注意力模块实现了一种纯粹的像素级注意力机制,旨在动态地重新校准输入特征图在空间维度(高度  $H$ 和宽度  $W$ )上的重要性。其核心目标是使模型能够聚焦于信息最丰富的空间区域,同时抑制不相关的区域,从而提升特征表示的质量。

该模块通过两个二维卷积层从输入的特征图中生成查询( $Q$ )与键( $K$ )。同时,为了简化计算,这两个卷积操作会先将通道数从  $C$ 压缩到一个中间维度  $C'$ 。具体为

$$Q = f_{3 \times 3}(X) \quad (15)$$

$$K = f_{3 \times 3}(X) \quad (16)$$

式中,  $Q, K \in \mathbf{R}^{B \times C' \times H \times W}$ 。随后,模块采用一种简化的相似性度量方式,将  $Q$ 和  $K$ 进行逐元素相加。这个相加结果接着通过一个  $1 \times 1$ 卷积层将通道维度压缩至1,生成空间注意力对数,记为  $L \in \mathbf{R}^{B \times 1 \times H \times W}$ ,具体为

$$L = f_{1 \times 1}(Q + K) \quad (17)$$

为了得到归一化的空间注意力图,模块先将  $L$ 在空间维度上展平( $B, 1, (H \times W)$ ),然后应用softmax函数在最后一个维度(即展平后的空间维度)上进行归一化,确保所有空间位置的注意力权重之和为1。最后,将归一化后的注意力权重重新塑形回原始形状,其结果为空间注意力图  $A$ ,具体为

$$A = \sigma_{\text{spatial}}(L) \in \mathbf{R}^{B \times 1 \times H \times W} \quad (18)$$

式中,  $\sigma_{\text{spatial}}$ 代表在空间维度上进行展平,softmax归一化再重新塑形的操作。最后一步是应用空间注意力。模块将计算得到的空间注意力图  $A$ 与原始输入特征图  $X$ 进行逐元素相乘。得益于广播机制,形状

为  $(B, 1, H, W)$  的注意力图  $A$  会自动扩展至  $(B, C, H, W)$ , 从而能够对输入特征图  $X$  的每个通道应用相同的空间权重。最终输出  $Y$  计算为

$$Y = X \odot A \quad (19)$$

式中,  $\odot$  表示逐元素相乘。输出  $Y$  保留了与输入  $X$  相同的维度, 但其特征值已根据空间注意力图进行了重新加权, 突出了重要空间位置的特征响应。

### 1.6 渐进残差聚合结构

为有效融合网络不同深度的特征并缓解信息衰减, 本文设计了一种渐进残差聚合结构。如图 1(a) 所示, 该结构首先通过一个初始卷积层将编码器输出的特征  $z$  映射为基础特征图  $z_1$ 。随后, 一个隐藏状态  $h$  被初始化为  $z_1$ , 并被依次送入  $N$  个串联的特征提取模块  $F_i$ 。其核心的渐进式更新流程可表示为

$$h_0 = z_1 \quad (20)$$

$$h_i = F_i(h_{i-1}) + z_1, \quad i = 1, \dots, N \quad (21)$$

在每个步骤中, 当前模块的输出并非与上一模块的输入相加, 而是始终与初始的基础特征  $z_1$  进行逐元素相加。这个过程确保了无论网络层多深, 原始的基础信息都能作为一个恒定的参照系, 直接参与到每一阶段的特征学习中, 从而为模型提供了稳定的信息补充和梯度流。

与仅利用最后一层输出的传统残差结构不同, 本结构对学习过程中的所有中间状态进行聚合。在完成  $N$  次迭代后, 所有存储的隐藏状态  $\{h_1, h_2, \dots, h_N\}$  将在通道维度上进行拼接, 形成一个包含多层次信息的聚合特征图  $h_{agg}$ , 最后通过一个融合层进行处理。

$$h_{agg} = \text{Concat}(h_1, h_2, \dots, h_N) \quad (22)$$

这种“恒定残差源”与“全阶段特征聚合”的结合, 是本结构与传统残差网络的核心区别。它不仅能有效防止深层网络中的信息丢失, 还能让模型在解码时充分利用从浅到深的所有特征, 从而提升最终的预测精度。

## 2 实验结果与分析

### 2.1 数据集

本研究采用 OSI-450-a 数据集 (Lavergne 等, 2019) 作为主要数据源。该数据集是由欧洲气象卫星开发组织 (European Organisation for the Exploitation of Meteorological satellites, EUMETSAT) 下属的

海冰卫星应用设施 (ocean and sea ice satellite application facility, OSI SAF) 发布的全球海冰浓度气候数据记录 (climate data record, CDR)。OSI-450-a 数据集时间跨度达 42 年, 涵盖 1978 年—2020 年的连续观测数据, 分辨率为 25 km。凭借其高质量和全面性, 该数据集已成为全球气候研究领域广泛使用的海冰浓度数据产品之一, 为研究人员提供了宝贵的长期海冰变化信息。

本文验证了 WRANet 模型的有效性和泛化能力, 使用了德国不来梅大学发布的 AMSR2 (ASI-AMSR2 sea ice concentration) 高分辨率北极海冰密集度数据集。该数据集基于 GCOM-W1 (global change observation mission—water 1) 卫星的 AMSR2 传感器数据, 通过 ASI 算法处理生成, 分辨率为 3.125 km (n3125 产品), 格式为 NetCDF (network common data form)。相比 25 km 的 OSI-450-a 数据集, AMSR2 能更精细地刻画海冰边缘和冰间水道等结构, 对模型提出了更高要求。实验采用与 OSI-450-a 相同的划分方式, 确保对比公平性。

### 2.2 实验参数

实验中, 本文采用 AdamW 优化器 (Loshchilov 和 Hutter, 2017), 学习率设置为 0.0001, 权重衰减参数为 0.0001。考虑到不同模型的特性, 保留了各自原始的损失函数, 而非使用统一的损失函数。数据集的划分方式为: 训练集包含 2000 年 1 月 1 日至 2010 年 12 月 31 日的数据, 验证集则包含 2011 年 1 月 1 日至 2015 年 12 月 31 日的数据。在训练过程中, 最大训练轮数设为 100, 同时采用了基于验证集上 MAE 指标的早期停止策略。采用有效的训练策略与数据增强技术是提升模型泛化能力的关键, 这在计算机视觉领域已有广泛共识和深入研究 (马愈卓等, 2025)。

### 2.3 对比实验分析

为验证本文提出的 WRANet 模型在北极海冰预测任务中的性能表现, 选取了 6 类具有代表性的基准模型进行对比实验, 包括两类基于循环结构的模型 (ConvLSTM (Shi 等, 2015)、PredRNNv2 (Wang 等, 2023)) 和 4 类基于非循环结构的模型 (SimVP (Gao 等, 2022)、TAU (Tan 等, 2023)、WaST (wavelet-driven spatiotemporal predictive learning) (Nie 等, 2024) 和 PastNet (Wu 等, 2024))。

ConvLSTM 通过将卷积运算引入 LSTM 单元内

部,开创性地实现了对时空序列数据(如视频帧)中时空相关性的同时建模。PredRNNv2通过引入梯度高速架构显著缓解了深度时空循环网络中的梯度消失问题,开创性地提升了模型学习视频长程动态依赖的能力。SimVP证明了在视频预测任务中,简洁的、以卷积为主(通常无显式循环结构)的非循环模型架构也能达到甚至超越复杂循环网络的性能。TAU开创性地将时间注意力机制应用于时空预测模型,使模型能够动态地关注和加权过去不同时刻的信息,以更有效地捕捉长程或关键的时间依赖关系。WaST设计了一个时间-频率感知转换器,通过分别建模空间频率和时间变化来有效地学习短程和长程时空信息。PastNet在傅里叶域中引入谱卷积算子,有效地引入了来自底层物理定律的归纳偏置。

实验从多个维度评估了模型的性能。在预测精度方面,选取平均绝对误差(mean absolute error, MAE)、均方根误差(root mean squared error, RMSE)、纳什效率系数(Nash-sutcliffe efficiency, NSE)以及平衡准确率(balanced accuracy, BACC)作为评价指标。为了方便比较,指标都以百分比呈现。MAE直观反映了预测值与实际观测值之间的平均绝对偏差;RMSE通过度量预测误差的平方均值评估模型,其特性是对较大误差更为敏感。MAE与RMSE这两项指标,其值越小,代表模型的预测精度越高。NSE通过综合考量预测值与观测值之间的偏差及方差评估模型的匹配程度,其值越接近1表明预测性能越优。BACC是一个评估模型空间场分类能力的指标。该指标首先通过一个预设阈值(本研究中为0.15)将连续的预测场和观测场进行二值化处理,即将其转化为“有冰”(值为1)与“无冰”(值为0)的二元空间分布图。随后,通过计算预测图与真实图之间不匹配的像元总数来量化模型的空间预测误差。最终,BACC通过将此误差进行归一化处理并用1减去该值而得到。因此,BACC取值为(0,1),越接近1,表明模型对海冰覆盖范围的二元分类预测越准确、性能越优。在计算效率方面,采用十亿次浮点运算数(giga floating point operations, GFLOPs)衡量模型的计算复杂度。这一综合性评估框架不仅可以精确衡量模型的预测准确性,也旨在考量其在实际应用场景中的计算资源需求。

为了全面评估模型的泛化性能,在OSI-450-a和AMSR2数据集进行了相同的对比实验,结果如表1

和表2所示。可以看出,本文提出的WRANet在RMSE、MAE、NSE和BACC等所有关键指标上均显著优于其他所有基准模型。

表1 OSI-450-a数据集上不同方法的实验结果对比

Table 1 Comparison of experimental results of different methods on the OSI-450-a dataset

方法	RMSE/% ↓	MAE/% ↓	NSE/% ↑	BACC/% ↑	GFLOPs ↓
ConvLSTM	8.77	2.81	94.29	95.49	1 526.26
PredRNNv2	8.51	2.70	94.66	95.78	3 200.50
SimVP	6.89	2.30	96.57	96.46	124.05
TAU	6.93	2.34	96.47	96.47	148.68
WaST	7.21	2.63	96.10	96.71	120.34
PastNet	7.29	2.56	96.07	96.46	119.52
WRANet	<b>6.44</b>	<b>2.02</b>	<b>97.03</b>	<b>96.96</b>	<b>118.09</b>

注:加粗字体表示各列最优结果,↓表示值越小越好,↑表示值越大越好。

表2 AMSR2数据集上不同方法的实验结果对比

Table 2 Comparison of experimental results of different methods on the AMSR2 dataset

方法	RMSE ↓	MAE ↓	NSE ↑	BACC ↑	/%
ConvLSTM	14.12	3.68	88.95	90.21	
PredRNNv2	13.65	3.51	89.41	90.53	
SimVP	10.85	2.89	91.58	91.42	
TAU	10.98	3.01	91.45	91.31	
WaST	11.45	3.38	91.10	91.55	
PastNet	11.70	3.35	90.88	91.25	
WRANet	<b>10.19</b>	<b>2.58</b>	<b>91.94</b>	<b>91.87</b>	

注:加粗字体表示各列最优结果,↓表示值越小越好,↑表示值越大越好。

从实验结果可以看出,WRANet在绝大多数指标上均展现出显著优势。在预测误差方面,RMSE达到6.44%,MAE达到2.02%,显著优于对比方法。在预测精度方面,NSE和BACC分别达到97.03%和96.96%,均显著优于对比方法。特别值得注意的是,得益于卷积网络的高并行性架构设计,本文方法WRANet展现出优异的计算效率,浮点运算次数仅为118.09 GFLOPs。相较于基于循环结构的ConvLSTM和PredRNNv2, RMSE分别降低2.33%和

2.07%。与当前最先进的非循环结构模型 SimVP 和 TAU 相比,WRANet 在保持了计算效率优势的同时,进一步将预测精度提升了约 0.5%。这些实验结果充分验证了 WRANet 在北极海冰预测任务中的卓越性能。

### 2.4 可视化分析

为更直观对比本文模型与现有最优模型的性能差异,选取 2020 年 9 月 1 日—14 日作为输入数据,并将 9 月 15 日—28 日设定为预测目标。该时段选取依据为:2020 年 9 月 15 日是当年北极海冰范围年度最小值出现日期。通过色标地图可视化模型预测与实际值的差异:红色表示高估区域,蓝色代表低估区域。为便于比较分析,本文重点考察 9 月 15 日、21 日

和 28 日 3 个关键时间节点误差分布特征。

为进行系统性的性能对比,选取两种具有代表性的预测模型:基于循环神经网络架构的 PredRNNv2 模型和基于非循环结构的 TAU 模型作为基准。这两种模型分别代表了当前海冰预测领域的主流技术路线。将它们的预测结果与 WRANet 进行可视化对比,如图 4 所示,其中,红色区域表示预测值高于实际值(即高估),蓝色区域表示预测值低于实际值(即低估)。从图 4 可以直观地观察到,WRANet 的预测结果在海冰形态和分布上与真实观测(ground truth)更为接近。特别是在预测值与真实值的差异分布图(右侧)中,WRANet 对应的误差区域(红色和蓝色)范围更小、颜色更浅,表明预测偏差显

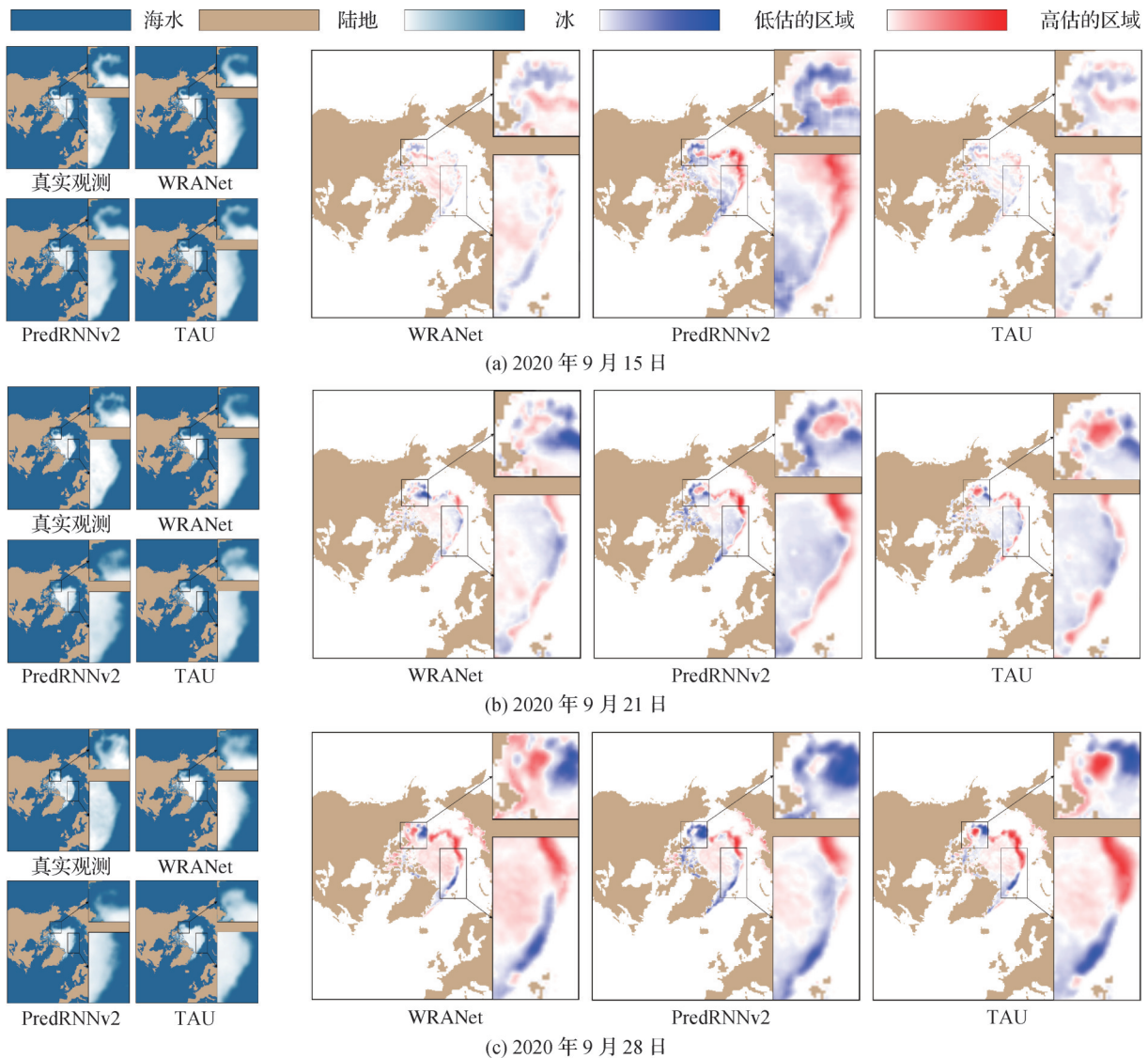


图 4 预测值与真实值差异的分布情况

Fig. 4 The distribution of differences between predicted and true values  
(a) September 15, 2020; (b) September 21, 2020; (c) September 28, 2020

著低于PredRNNv2和TAU模型,展现出更优越的预测性能。

## 2.5 消融实验

为了验证所设计的小波多尺度特征提取模块与渐进残差聚合结构的有效性,在OSI-450-a数据集上进行消融实验。实验包括4种配置:1)主干网络仅包含基础编码器—解码器结构,编码器输出经过简单的二维残差块堆叠4次后直接连接到解码器;2)将基准模型中的二维残差卷积块替换为本文设计的小波多尺度特征提取模块,但保持其直接堆叠的聚合方式不变;3)将直接堆叠的聚合方式替换为渐进残差聚合结构,而内部的特征提取单元使用与基准模型相同的二维残差卷积块;4)本文完整模型(同时集成小波多尺度特征提取模块与渐进残差聚合结构)。同样地,消融实验测量了RMSE、MAE、NSE和BACC等多项关键性能指标,结果如表3所示。可以看出,单独使用小波多尺度特征提取模块时,MAE从2.57%降至2.28%,表明其对高频细节的捕捉有效;而渐进残差聚合结构使BACC提升0.41%(从96.11%升至96.52%),印证了其优化特征一致性的能力。两者的结合使NSE指标达到97.03%,相较于主干网络提升了0.97%。数值上验证了多尺度特征融合与渐进优化策略的协同优势。

表3 小波多尺度特征提取模块和渐进残差聚合结构消融实验结果

Table 3 Ablation experiment results of wavelet multi-scale feature extraction module and progressive residual aggregation structure

网络结构	RMSE ↓	MAE ↓	NSE ↑	BACC ↑
主干网络	7.45	2.57	96.06	96.11
+小波多尺度特征提取模块	6.58	2.28	96.93	96.64
+渐进残差聚合结构	6.68	2.28	96.79	96.52
+小波多尺度特征提取模块+渐进残差聚合结构	<b>6.44</b>	<b>2.02</b>	<b>97.03</b>	<b>96.96</b>

注:加粗字体表示各列最优结果,↓表示值越小越好,↑表示值越大越好。

为验证小波变换在模型中的有效性,进行了一项消融研究。将完整模型与一个移除了小波变换模块的版本进行比较。移除该模块后,模型将仅能从

空间域提取特征,而无法利用频率域中的信息。实验结果如表4所示,移除小波变换以后, RMSE上升0.13%, MAE上升0.2%, NSE下降0.24%, BACC下降0.22%。因此,小波变换能够有效提取关键的频率域信息,对于提升模型的整体预测精度和稳健性至关重要。

表4 小波变换消融实验结果

Table 4 Ablation experiment results of wavelet transform

网络结构	RMSE ↓	MAE ↓	NSE ↑	BACC ↑
小波变换(w/o)	6.57	2.22	96.79	96.74
完整模型	<b>6.44</b>	<b>2.02</b>	<b>97.03</b>	<b>96.96</b>

注:加粗字体表示各列最优结果,↓表示值越小越好,↑表示值越大越好。

为验证像素级注意力模块的有效性,将其从完整模型中移除进行消融实验。如表5所示,移除该模块后,模型性能在所有评估指标上均出现下降: RMSE增加0.16%, MAE增加0.15%, NSE下降0.13%, BACC下降0.22%。这些一致性的性能下降清晰地证明了像素级注意力模块对于提升模型预测准确性的有效性。

表5 像素级注意力模块消融实验结果

Table 5 Ablation experiment results of pixel-wise attention module

网络结构	RMSE ↓	MAE ↓	NSE ↑	BACC ↑
像素级注意力模块(w/o)	6.60	2.17	96.90	96.74
完整模型	<b>6.44</b>	<b>2.02</b>	<b>97.03</b>	<b>96.96</b>

注:加粗字体表示各列最优结果,↓表示值越小越好,↑表示值越大越好。

## 3 结论

针对北极海冰密集度预测中的挑战问题,本文设计并实现了小波多尺度残差聚合北极海冰预测网络WRANet。首先,提出了小波多尺度特征提取模块,基于其频域分析能力(二维离散小波变换)并结合多尺度卷积策略,通过在特征提取过程中区分处理,以完成对高频细节和低频轮廓特征的充分提取

与保留,改善了海冰时空特征的捕捉效果,对预测精度有所提升。然后,设计的渐进残差聚合结构通过有效聚合网络深层处理中不同阶段的中间特征输出,充分挖掘了隐藏在不同特征层级中的关键动态信息,改善了模型对复杂时空依赖关系的建模能力和最终的预测效果。

本文在公开的 OSI-450-a 和 AMSR2 海冰数据集上,通过 RMSE, MAE, NSE, BACC 等多个评价指标对实验结果完成评价和分析,结合预测误差可视化结果,表明了本文所提出的 WRANet 的有效性和相对于基准模型的优越性。在两个来源和分辨率不同的数据集上均取得了一致的领先结果,充分证明了所提模型的鲁棒性和良好的泛化能力,为高精度北极海冰预测提供了可靠的解决方案。

本文主要聚焦于利用现有遥感和再分析数据进行短期的海冰密集度预测。未来可继续探索将 WRANet 应用于更长期的预测任务,或融合更多来源(如更高分辨率的 SAR 数据、海洋动力学数据、大气强迫数据等)的信息,进一步研究不同数据源之间的协同作用与信息融合方法,以扩展 WRANet 的应用场景,在气候变化监测与预警领域为社会带来更多的实际应用价值。

## 参考文献 (References)

- Cavaleri D J, Parkinson C L, Gloersen P and Zwally H J. 1996. Sea Ice Concentrations from Nimbus-7 SMMR and DMSP SSM/I-SSMIS Passive Microwave Data. NSIDC-0051. NASA National Snow and Ice Data Center Distributed Active Archive Center [DOI: 10.5067/8GQ8LZQVLOVL]
- Elfving S, Uchibe E and Doya K. 2018. Sigmoid-weighted linear units for neural network function approximation in reinforcement learning. *Neural Networks*, 107: 3-11 [DOI: 10.1016/j.neunet.2017.12.012]
- Gao Z Y, Tan C, Wu L R and Li S Z. 2022. SimVP: simpler yet better video prediction//Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans, USA: IEEE: 3160-3170 [DOI: 10.1109/CVPR52688.2022.00317]
- Hersbach H, Bell B, Berrisford P, Hirahara S, Horányi A, Muñoz-Sabater J, et al. 2020. The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146 (730): 1999-2049 [DOI: 10.1002/qj.3803]
- Hochreiter S and Schmidhuber J. 1997. Long short-term memory. *Neural Computation*, 9 (8): 1735-1780 [DOI: 10.1162/neco.1997.9.8.1735]
- Horvath S, Stroeve J, Rajagopalan B and Kleiber W. 2020. A Bayesian logistic regression for probabilistic forecasts of the minimum September Arctic sea ice cover. *Earth and Space Science*, 7 (10): #e2020EA001176 [DOI: 10.1029/2020EA001176]
- Hou Z Q, Qu M J, Li J G, Ma S G, Wang Y C and Yang X B. 2025. Lightweight CNN-Transformer combined network for real-time semantic segmentation. *Journal of Image and Graphics*, 30 (7): 2437-2450 (侯志强, 屈敏杰, 李俊歌, 马素刚, 王昀琛, 杨小宝. 2025. 轻量级 CNN-Transformer 相结合的实时语义分割网络. *中国图象图形学报*, 30 (7): 2437-2450) [DOI: 10.11834/jig.240527]
- Huang D M, Li M H, Song W and Wang J. 2018. Performance of convolutional neural network and deep belief network in sea ice-water classification using SAR imagery. *Journal of Image and Graphics*, 23(11): 1720-1732 (黄冬梅, 李明慧, 宋巍, 王建. 2018. 卷积神经网络和深度置信网络在 SAR 影像冰水分类的性能评估. *中国图象图形学报*, 23 (11): 1720-1732) [DOI: 10.11834/jig.180226]
- Lavergne T, Sørensen A M, Kern S, Tonboe R, Notz D, Aaboe S, et al. 2019. Version 2 of the EUMETSAT OSI SAF and ESA CCI sea-ice concentration climate data records. *The Cryosphere*, 13 (1): 49-78 [DOI: 10.5194/tc-13-49-2019]
- Loshchilov I and Hutter F. 2017. Decoupled weight decay regularization//Proceedings of the 7th International Conference on Learning Representations. New Orleans, USA: ICLR: 1-10
- Ma Y Z, Zhang Y F, Jia W, Liu J Y, Gan T, Yang W H, et al. 2025. Recent advances in data generation and its applications in computer vision. *Journal of Image and Graphics*, 30 (6): 1872-1952 (马愈卓, 张永飞, 贾伟, 刘家瑛, 甘甜, 杨文瀚, 等. 2025. 面向计算机视觉的数据生成与应用研究进展. *中国图象图形学报*, 30(6): 1872-1952) [DOI: 10.11834/jig.250085]
- Mei H W, Shang H L, Su P and Liu Y P. 2024. Optic disc and cup segmentation with combined residual context encoding and path augmentation. *Journal of Image and Graphics*, 29(3): 637-654 (梅华威, 尚虹霖, 苏攀, 刘艳平. 2024. 融合残差上下文编码和路径增强的视杯视盘分割. *中国图象图形学报*, 29(3): 637-654) [DOI: 10.11834/jig.230140]
- Nie X S, Yan Y F, Li S Y, Tan C, Chen X, Jin H Y, et al. 2024. Wavelet-driven spatiotemporal predictive learning: bridging frequency and time variations//Proceedings of the 38th AAAI Conference on Artificial Intelligence. Vancouver, Canada: AAAI Press: 4334-4342 [DOI: 10.1609/aaai.v38i5.28230]
- Parkinson C L and DiGirolamo N E. 2021. Sea ice extents continue to set new records: Arctic, Antarctic, and global results. *Remote Sensing of Environment*, 267: #112753 [DOI: 10.1016/j.rse.2021.112753]
- Pithan F and Mauritsen T. 2014. Arctic amplification dominated by temperature feedbacks in contemporary climate models. *Nature Geoscience*, 7(3): 181-184 [DOI: 10.1038/ngeo2071]

- Serreze M C and Barry R G. 2011. Processes and impacts of Arctic amplification: a research synthesis. *Global and Planetary Change*, 77(1/2): 85-96 [DOI: 10.1016/j.gloplacha.2011.03.004]
- Shi X J, Chen Z R, Wang H, Yeung D Y, Wong W K and Woo W C. 2015. Convolutional LSTM network: a machine learning approach for precipitation nowcasting//Proceedings of the 29th International Conference on Neural Information Processing Systems. Montreal, Canada: MIT Press: 802-810
- Stroeve J C, Serreze M C, Holland M M, Kay J E, Malanik J and Barrett A P. 2012. The Arctic's rapidly shrinking sea ice cover: a research synthesis. *Climatic Change*, 110(3/4): 1005-1027 [DOI: 10.1007/s10584-011-0101-1]
- Tan C, Gao Z Y, Wu L R, Xu Y J, Xia J, Li S Y, et al. 2023. Temporal attention unit: towards efficient spatiotemporal predictive learning//Proceedings of 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver, Canada: IEEE: 18770-18782 [DOI: 10.1109/CVPR52729.2023.01800]
- Tepes P, Gourmelen N, Nienow P, Tsamados M, Shepherd A and Weissgerber F. 2021. Changes in elevation and mass of Arctic glaciers and ice caps, 2010-2017. *Remote Sensing of Environment*, 261: #112481 [DOI: 10.1016/j.rse.2021.112481]
- Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez A N, et al. 2017. Attention is all you need//Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach, USA: Curran Associates Inc.: 6000-6010
- Wang Y B, Wu H X, Zhang J J, Gao Z F, Wang J M, Yu P S, et al. 2023. PredRNN: a recurrent neural network for spatiotemporal predictive learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(2): 2208-2225 [DOI: 10.1109/TPAMI.2022.3165153]
- Wu H, Xu F, Chen C, Hua X S, Luo X and Wang H X. 2024. PastNet: introducing physical inductive biases for spatio-temporal video prediction//Proceedings of the 32nd ACM International Conference on Multimedia. Melbourne, Australia: ACM: 2917-2926 [DOI: 10.1145/3664647.3681489]
- Wu Y X and He K M. 2018. Group normalization//Proceedings of the 15th European Conference on Computer Vision. Munich, Germany: Springer: 3-19 [DOI: 10.1007/978-3-030-01261-8\_1]

### 作者简介

- 弓政,男,硕士研究生,主要研究方向为人工智能海洋学、遥感图像处理与分析。E-mail:ghost\_him@126.com
- 高峰,通信作者,男,副教授,主要研究方向为人工智能海洋学和高光谱图像处理。E-mail:gaofeng@ouc.edu.cn
- 张家亮,男,硕士研究生,主要研究方向为人工智能海洋学、遥感图像处理与分析。E-mail:zhangjia\_liang@foxmail.com
- 甘言海,男,讲师,主要研究方向为人工智能海洋学、遥感图像处理与分析。E-mail:ganyanhai@ouc.edu.cn
- 董军宇,男,教授,主要研究方向为海洋大数据、水下图像处理与分析。E-mail:dongjunyu@ouc.edu.cn