

基于 Multiloss-SAM-ConvLSTM 的北极航道独立海冰运动预测

郑付强^{1,2}, 匡定波², 胡勇^{2*}, 巩彩兰², 黄硕²

(1. 中国科学院大学, 北京 100049;

2. 中国科学院上海技术物理研究所 中国科学院红外探测与成像技术重点实验室, 上海 200083)

摘要: 北极航道海冰运动的准确预测对于保证航行安全、评估航道可通行性和动态修正航线具有重要的指导意义。传统的光流法无法满足海冰运动预测任务中“时空预测+语义分割”的要求。为此, 基于 MERSI-II 影像制作了海冰运动数据集 SeaiCeMoving, 提出了一种基于 Multiloss-SAM-ConvLSTM 的海冰运动预测算法, 该算法在 SAM-ConvLSTM 的基础上引入加权的 FDWloss, 强化了各节点空间语义的获取。针对样本分布不平衡, 讨论了后端分割阈值的偏移效应, 通过网格搜索确定最佳分割阈值, 提高了海冰整体预测结果。实验结果表明, 该方法的 Kappa 系数为 0.75, IOU 为 0.61, Dice 系数为 0.76, 相较于 SAM-ConvLSTM, 分别提高了 0.1、0.12 和 0.1, 对运动后海冰的位置预测和形状提取能力均有提升, 减少了海冰“粘连”的情况。此外, 该算法对薄云干扰下的海冰运动依然具备良好的预测能力, 可以为北极航线的动态规划和航线修正提供较为准确的技术支撑。

关键词: 北极航道; 风云三号卫星; 独立海冰; Multiloss-SAM-ConvLSTM; 运动预测
中图分类号: TP7 文献标识码: A

Prediction of independent sea ice motion in Arctic channel based on Multiloss-SAM-ConvLSTM

ZHENG Fu-Qiang^{1,2}, KUANG Ding-Bo², HU Yong^{2*}, GONG Cai-Lan², HUANG Shuo²

(1. University of Chinese Academy of Sciences, Beijing 100049, China;

2. Key Laboratory of Infrared System Detection and Imaging Technology, Shanghai Institute of Technical Physics Chinese Academic of Sciences, Shanghai 200083, China)

Abstract: Accurate prediction of independent sea ice motion in arctic shipping lanes is of great guiding significance for ensuring navigation safety, assessing navigation navigability and dynamically correcting shipping lanes. However, the traditional optical flow method cannot meet the requirement of "spatio-temporal prediction + semantic segmentation". In this paper, the sea ice motion data set SeaiCeMoving was made based on MERSI-II image and a sea ice motion prediction algorithm based on Multiloss-SAM-ConvLSTM was proposed, introducing weighted FDWloss based on SAM-ConvLSTM to enhance the acquisition of spatial semantics of each node. Aiming at the imbalanced sample distribution, we discussed the offset effect of back-end segmentation threshold. The optimal segmentation threshold is determined by grid search method, which improves the overall prediction result of sea ice motion. The experimental results indicate that the Kappa coefficient, IOU coefficient and Dice coefficient of the proposed method are 0.75, 0.61 and 0.76 respectively. Compared with SAM-ConvLSTM, they are improved by 0.1, 0.12 and 0.1 respectively. Furthermore, the proposed method can improve the position prediction and shape recovery ability of sea ice after motion and reduce the "adhesion" of sea ice. In addition, the algorithm can still effectively predict the sea ice motion under the interference of thin clouds, which can provide more accurate technical support for the dynamic planning and route correction of the Arctic route.

收稿日期: 2022-01-28, 修回日期: 2022-07-02

Received date: 2022-01-28, revised date: 2022-07-02

基金项目: 国家自然科学基金(31970378); 江苏省水利科技项目(2020068)

Foundation items: Supported by the National Natural Science Foundation of China (31970378) and the Water Conservancy Science and Technology Project of Jiangsu Province (2020068)

作者简介 (Biography): 郑付强 (1993—), 男, 浙江三门县人, 博士研究生, 主要研究方向为极地海冰分割与预测、基于深度学习的遥感图像应用. E-mail: zhengfq@whu.edu.cn

* 通讯作者 (Corresponding author): E-mail: huyong@mail.sitp.ac.cn

Key words: Arctic route, FY-3 satellite, independent sea ice, Multiloss-SAM-ConvLSTM, motion prediction

引言

海上航运是全球贸易不可或缺的“血管”,近年来传统海运线路由于航运压力的增大而逐渐拥挤甚至发生大规模的堵塞事件。因此,作为联系亚、欧、美三大洲的最短航线,北极航线的战略价值与航运价值显得尤为重要^[1]。与此同时,随着全球变暖愈演愈烈,整个北极的海冰覆盖面积逐年减少,使得开辟北极航线成为可能。但是如何及时掌握航线冰情,并有效规避航道内形状与位置不断变化的海冰仍然是全线开辟北极航线亟待解决的问题。由于一年中北极航线海冰消融的时间段具有周期性,航行时一味地绕过海冰区域,会增加航线的总长度以及缩短航道通航的窗口期,最终导致航线的基础航行成本呈跳跃式增长。为了科学地动态规划航线,有效地降低航行成本,除了要精准区分海冰区域的大型独立海冰、碎屑冰以及冰水混合物等,更重要的是结合独立海冰的运动轨迹,预测独立海冰在未来短时间内的走势,以便船员重新评估前方海域的可通行性,从而对航线进行动态修正。

目前针对海冰运动的预测研究较少,其中 Petrou Z I 等人通过被动微波数据计算包含海冰运动信息的光流,再利用 ConvLSTM 预测光流走势,通过光流走势来表征海冰的运动,并将结果与 IABP 的浮标信息进行比对以评估准确性^[2]。此方法需要大量的预处理操作来提取光流图,同时使用的被动微波数据 AMSR 的空间分辨率为 10 km,不能表达北极航道上独立海冰等地物的细节特征。除了被动微波数据之外,SAR 影像也是海冰相关研究中常用的数据,但是 SAR 影像重访周期过长,无法满足北极航道中海冰运动短时预测的时间分辨率要求。

独立海冰运动的短时预测本质上是时空序列预测问题。计算机视觉领域的光流法是这类问题的传统解决方法,即利用光流向量中像素点所携带的运动信息与物体结构信息对图像序列外推得到下个时刻的预测图像^[3]。根据光流约束条件的不同,光流法还分为全局光流算法与局部光流算法,两者均已被应用到不同的时空预报场景之中^[4-6]。然而,由于光流估计步骤与预测图像生成步骤分离导致模型参数难以确定,无法预测目标的消亡趋势,受噪声影响大等原因,光流法在实际的时空序

列预测任务中预测性能并不尽如人意^[7-8]。此外,对运动后海冰的精确语义分割是航道可通行性评估的基础,而光流法并不能兼顾“时空预测”与“语义分割”。

近年来,深度学习不断地被应用在时空序列预测领域。Marc' Aurelio Ranzato 提出了第一个基于递归神经网络(Recurrent Neural Networks, RNN)的时空序列预测模型^[9]。Srivastava 提出了基于全连接长短期记忆(Fully Connected-Long Short Term Memory, FC-LSTM)结合编、解码的时空序列预测模型,利用两个 LSTM 结构对输入帧与输出帧分别进行编码与解码^[10]。不论是基于 RNN 还是 LSTM 的方法都无法满足时空序列预测中对时空相关性的要求。因此在 FC-LSTM 的基础上,施健行引入了对空间信息敏感的卷积从而构建了卷积长短期记忆(Convolutional LSTM, ConvLSTM)^[8],弥补了 FC-LSTM 空间特征捕捉能力差的缺点,并将其应用于降水预报中。此后还出现了各种 ConvLSTM 的变体^[11-13],但是单纯通过堆叠卷积不仅不能提高模型对旋转运动特征的获取能力,过深的模型还会导致梯度弥散现象的产生^[14]。为了提高对全局特征的获取能力以及对不规则运动的预测能力,施健行提出了轨迹门循环单元(Trajectory Gate Recurrent Unit, TrajGRU)^[15],在继承 ConvLSTM 对时空特征敏感的基础上,通过学习 GRU 模块隐藏状态的偏移特征,并加入生成光流分量的过程,提高了模型对不规则运动的预测能力。Lin Z 在 ConvLSTM 的基础上嵌入了自注意力记忆模块,构建了自注意力卷积长短期记忆(Self-attention memory Convolutional LSTM, SAM-ConvLSTM)^[16],不仅提高了模型对全局时空特征的提取能力,还减少了预测过程的计算参量,在 MovingMNIST 和 KTH 等标准数据集中都取得了最好的结果,是目前时空预测能力最好的网络之一。与 KTH 和 MovingMNIST 数据集中前后景差别大的样本分布不同,北极场景中还有雪地、浮冰、冰盖等与独立海冰相似的地物干扰,独立海冰运动的短时预测不仅要求对海冰运动轨迹的精准建模,还需要对运动后海冰进行精确提取,对时空预测模型提出了更高的要求。

为研究北极航道中海冰运动的预测问题,本研究基于风云三号卫星 MERSI-II 影像,经过一系列预处理操作,制作了独立海冰运动数据集(Seaice-

Moving)。在 SAM-ConvLSTM 的基础上,提出了一种 Multiloss-SAM-ConvLSTM 模型,实现了独立海冰运动的短时预测。以西北航道中靠近格陵兰岛的区域为实验区,将实验结果与 SAM-ConvLSTM 进行对比。结果表明,本算法不仅提高了对海冰位置与旋转特性的预测能力,对运动后海冰形状的恢复能力也更加出色。针对数据集样本不平衡引起的分割阈值偏移效应,本研究通过网格搜索确定最佳分割阈值,提高了海冰整体预测结果。此外,本研究在一定程度上可以抑制薄云对海冰运动预测时的干扰,对薄云下的海冰依然具有良好的预测与提取能力,可以为北极航线的动态规划和航线纠正提供技术支持。

1 相关工作

1.1 研究区概况与数据集制作

相比微波数据,光学遥感影像具有尺度大、低成本的优势,能够提供大面积的冰情信息。其中,NASA 的 MODIS 数据根据海冰在可见光与红外波段的辐射特性,已经生成了成熟的海冰区域轮廓产品^[17]。我国风云三号 D 星上搭载的中分辨率光谱成像仪(Medium Resolution Spectral Imager-II, MERSI-II)在同为 250 m 分辨率的尺度下,仪器性能相比 MODIS 更加优越,可以提供更精细的海冰分布与运动信息,对于同一地区的相邻两次成像间隔为 100 分钟,可以为海冰运动的预测研究提供更加精准的数据支撑^[18]。

真实极区航道的可通行性会随着船舶的航行及区域内独立海冰自身形态变化和洋流与海风的影响下发生的位移变化而动态变化。因此,只有对极区航道内海冰的运动进行监测并准确预测一定时间内海冰的运动趋势,再与船舶的既定航线相结合才可以分析接下来船舶驶入的海域的可通行性,根据分析结果指导船舶针对海冰的运动趋势提前

做出反应,实现航线的动态规划。周颖^[19]在研究中得出弗雷姆海峡的海冰流速在 1~2 km/h 左右,而远洋船只的一般航速为 12 节(约 22 km/h)。因此,只要掌握船舶既定航线上未来 1~2 小时的海冰运动状态与范围,船只就可以在保持航速的前提下提前修正航线来动态规避有相撞风险的海冰,从而提高船舶在流冰区的机动通行能力、降低成本。

结合海冰流速、船舶航行速度以及预测时间间隔三者对研究数据的硬性要求,本文选取了我国风云三号 D 星上搭载的 MERSI-II L1 产品作为海冰运动的预测研究的数据源。MERSI-II 是国际上最先进的宽幅成像遥感仪器之一,可以每日无缝隙获取全球 250 米分辨率真彩色影像,实现云、气溶胶、水汽、陆地表面特性、海洋水色等大气、陆地、海洋参量的高精度定量反演^[20],同时,每天可以对同一地区进行多次重访。MERSI-II 总共有 25 个探测通道,其中 band1-band19 是 0.4~2.1 μm 的 19 个太阳反射通道(RSB),band20~band25 是 3.8~12.5 μm 的 6 个热发射通道(TEB)。MERSI-II 的性能指标如表 1 所示。

研究区位于加拿大北部北极群岛到阿拉斯加北岸的西北航道中靠近格陵兰岛的区域。航道融冰期正常存在的独立海冰、薄云覆盖下的或者与薄冰区域共存的独立海冰等西北航道中可能会出现各类海冰场景在此研究区均有分布,保证了实验结果的泛用性。研究区位置以及数据采集时间分布如图 1。

MERSI-II L1 数据采用无符号整型记录观测值,本身没有实际的物理意义,经过辐射校正、地理定位、几何校正、坐标系变换和配准等一系列预处理后,可以从几何与辐射两个维度准确表征图像所蕴含的地物信息。在此基础上,再经过通道选择、数据标注与划分以及数据裁剪得到最终的海冰运动预测数据集 SeaiCeMoving。

表 1 MERSI-II 性能指标

Table 1 Performance indicators of MERSI-II

性能参数	指标
量化等级	12 bit
扫描范围	$\pm 55.1^\circ \pm 0.1^\circ$
每条扫描线采样点数	2 048(1 000 m), 8 192(250 m)
波段范围	0.4~12.5 μm
通道间像元配准	<0.3 个像元
定标精度	可见光和近红外通道:5%(反射率),星上定标器实现可见光星上定标(相对和绝对辐射)。 红外通道(星上黑体):0.5 K(270 K),分裂窗两个通道定标误差一致性<小于 0.5 K

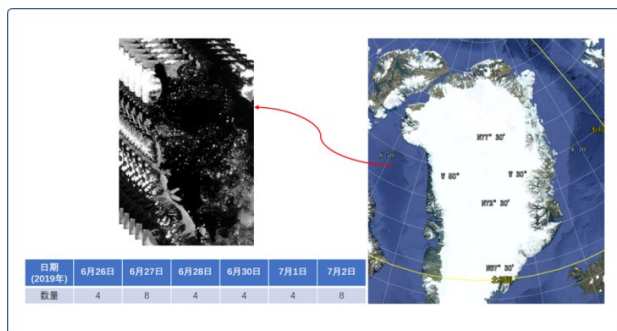


图1 研究区位置以及数据采集时间分布:Seaicemoving 数据集制作

Fig. 1 Location of research distribution and data acquisition time distribution

图像配准

为了更好地解决独立海冰运动的短时预测问题,我们在原始数据的预处理以及后续数据集的构造过程中必须进行通道配准,保证每个时刻影像之间同一位置的像元所对应的实际地理位置是高度一致的。图2给出了数据集中同一研究区在2019年6月27日11时35分、13时15分和14时55分时刻图像第一通道配准前后的RGB合成图,可以看到配准后边缘不再有明显的色彩差异,最终本底数据集图像对的配准误差都小于1个像元,能够满足海冰运动短时预测的需要。

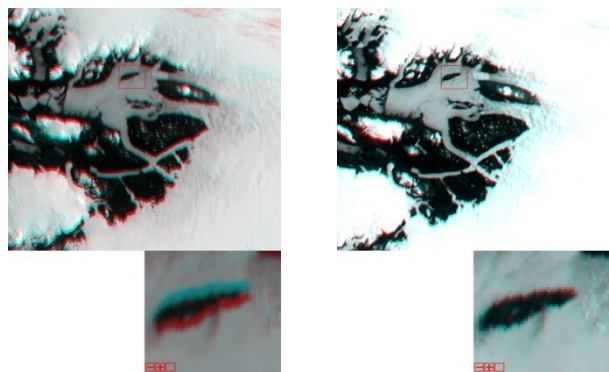


图2 不同时刻海冰影像配准前后对比图:通道选择

Fig. 2 Comparison before and after sea ice image registration at different times

为解决依靠空间结构特征样本区分难度大的问题,本文考虑到北极航道的地物类型以及MERSI-II 对应的光谱响应特性,本文对预处理之后的MERSI-II L1 数据通道进行重组,保留了对海冰、海水、云层等极地地物敏感的可见近红外波段与短波红外波段(1.640 μm)通道,剔除了其余的冗余通

道。其中,分辨率为1 000 m、中心波长为1.640 μm 的短波红外通道数据需要重采样到250 m之后再与其他通道数据进行融合。本文最终的海冰运动本底数据通道特性如表2所示。

表2 海冰运动本底数据集通道特性

Table 2 Channel characteristics of Seaicemoving datasets

通道号	对应传感器通道号	中心波长 (μm)	光谱带宽 (μm)	空间分辨率(m)
1	1	0.470	0.05	250
2	2	0.550	0.05	250
3	3	0.650	0.05	250
4	4	0.865	0.05	250
5	6	1.640	0.02	250(重采样)

数据划分与标注

本文把对同一研究区连续成像的4幅MERSI-II L1 影像(同一研究区两次成像之间时间间隔一致,为100分钟)定义为最小实验影像对(数据空间维度分布为[成像时刻间隔数 T ,通道分布 C ,长 L ,宽 W])。设每个最小实验影像对的第一幅MERSI-II L1 影像为 T_0 ,因此,北极航道独立海冰运动的短时预测就是在 T_0 、 T_{100} 、 T_{200} (下标数值单位为分钟)三幅影像的基础上生成 T_{300} 的海冰分布。最终,本文挑选了不同基准成像时刻、不同研究区的18个最小实验影像对(单个最小实验影像对的覆盖面积约16万平方公里),经过上文提到的相关预处理后形成最终的本底数据集,并将其中的3个最小实验影像对挑选为测试集,剩下的15个最小实验影像对为训练集和验证集。为了验证算法的有效性和准确性,邀请了海洋与极地研究领域的专业解译人员对经过预处理之后的最小实验影像对的每一幅影像进行人工解译,并将解译的结果图作为独立海冰运动短时预测可参考的独立海冰分布真值图。本文定义的最小实验影像对及其对应的真值图影像对如图3所示。

为丰富样本量,扩充实验数据规模,本文采用随机重叠度与长宽双向随机滑动裁剪方式对最小实验影像对进行裁剪,并对裁剪结果在图像层面随机进行数据增强(随机翻转、旋转、平移、缩放、添加噪声等),最终得到独立海冰运动预测数据集 Seaicemoving, Seaicemoving 数据集总体分布如表3所示。

1.2 基于 Multiloss-SAM-ConvLSTM 的独立海冰运动预测方法

针对船舶航线动态规划的实际需求以及独立

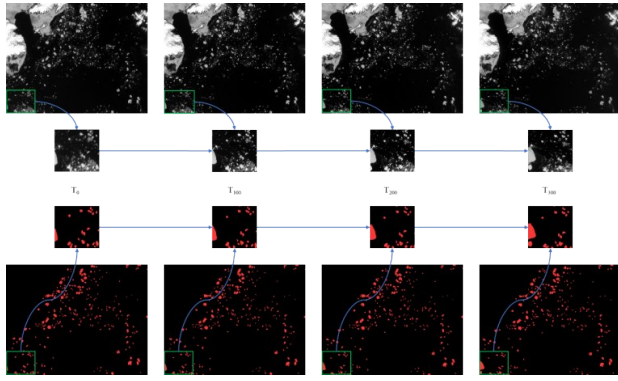


图3 实验数据示意图,从左到右,从上到下分别对应 $T_0 - T_{300}$ 时刻实验区海冰运动时序数据可视化图与人工解译标签图:数据裁剪

Fig. 3 The schematic diagram of experimental data, from left to right, from top to bottom, they correspond to the schematic of data and label at T_0 to T_{300}

表3 海冰运动预测数据集分布

Table 3 Distribution of SeaiiceMoving datasets

样本空间维度分布	[4, 5, 64, 64]
训练样本数量	16 560
验证样本数量	3 310
独立海冰占比	0. 03

海冰的实际运动特点,本文在 SAM-ConvLSTM 的基础上,提出了一种 Multiloss-SAM-ConvLSTM 模型,实现了独立海冰运动的短时预测。此外,针对 SeaiiceMoving 数据集的样本不平衡问题,本文还从损失函数构建和模型后端分割阈值偏移效应两个方面入手,提高了海冰运动整体预测结果。

1. 2. 1 Multiloss-SAM-ConvLSTM 网络结构

SAM-ConvLSTM 是 Lin 在 ConvLSTM 的基础上提出的全新时空预测网络,通过引入自我注意力机制^[21]构建了自注意力记忆模块(Self-Attention Memory, SAM)。SAM 通过特征聚合、记忆更新和整合输出三个步骤增强了对时空预测过程中长期的空间依赖特征的提取能力,从而可以让模型对复杂的动态时空关系进行更好地表征。引入 SAM 固然可以提高时空预测过程中时间维度上空间特征的传递与记忆能力,但是同一个时刻的空间特征提取过程仍然是卷积层的简单堆叠,这种特征提取方式限制了 SAM-ConvLSTM 对单一时刻图像内各类物体本身特征的提取能力。综合分析此类时空预测模型的应用场景可以看出, KTH 和 MovingMNIST 数据集中前景和背景本身就有比较大的区分度;而基于雷达影像的降雨预测场景对预测结果的细节以及边

界要求相对较低,且数据集本身并没有明显的前后景之分^[15],这些场景均偏向解决“时空预测+图像恢复”类问题。不同于上述应用场景,北冰洋的独立海冰分布不仅形态差异巨大,还会受到薄云以及碎屑冰、小冰和冰水混合的薄冰区的干扰,独立海冰运动的短时预测场景属于“时空预测+语义分割”问题,不仅要求对海冰运动轨迹的精准建模,还需要对运动后海冰进行精确提取。因此,预测过程中各个时刻语义特征的提取程度对最终生成的预测结果影响更大。针对这一问题,在 SAM-ConvLSTM 的基础上,本文提出了 Multiloss-SAM-ConvLSTM,即在网络后端各个“时间节点”处构建多个损失函数,在保留了 SAM-ConvLSTM 模块对时间维度上空间依赖特征有较好提取能力的同时更加注重独立海冰语义特征的提取。其整体结构如图 4 所示:

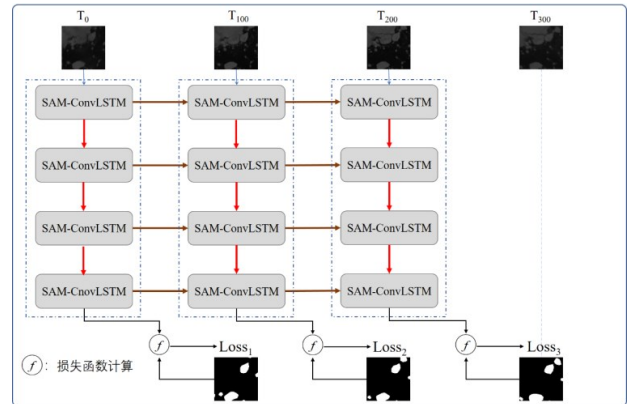


图4 Multiloss-SAM-ConvLSTM 整体架构: SAM-ConvLSTM 模块

Fig. 4 Architecture of Multiloss-SAM-ConvLSTM

标准的 SAM-ConvLSTM 模块是将基于自注意力记忆模块 SAM 嵌入到 ConvLSTM 模块形成的,具体结构图如图 5 所示^[15],其基本模型如式(1)所示。

$$\begin{aligned} \hat{\chi}_t &= SAM(\hat{\chi}_t), \hat{\mathcal{H}}_{t-1} = SAM(\mathcal{H}_{t-1})i_t = \sigma(W_{xi} * \hat{\chi}_t + \\ & W_{hi} * \hat{\mathcal{H}}_{t-1} + b_i) \\ f_t &= \sigma(W_{xf} * \hat{\chi}_t + W_{hf} * \hat{\mathcal{H}}_{t-1} + b_f) \\ g_t &= \tanh(W_{xc} * \hat{\chi}_t + W_{hc} * \hat{\mathcal{H}}_{t-1} + b_c) \quad , \quad (1) \\ \mathcal{L}_t &= f_t \circ \mathcal{L}_{t-1} + i_t \circ g_t \\ o_t &= \sigma(W_{xo} * \hat{\chi}_t + W_{ho} * \hat{\mathcal{H}}_{t-1} + b_o) \\ \mathcal{H}_t &= o_t \circ \tanh(\mathcal{L}_t) \end{aligned}$$

其中 SAM 表示自注意力记忆模块的运算过程。

自注意力机制通过计算二元关系函数中不同位置的特征图之间的成对关系来捕捉图像的全局特征。此外,过去时刻的相关特征对当前时间步的

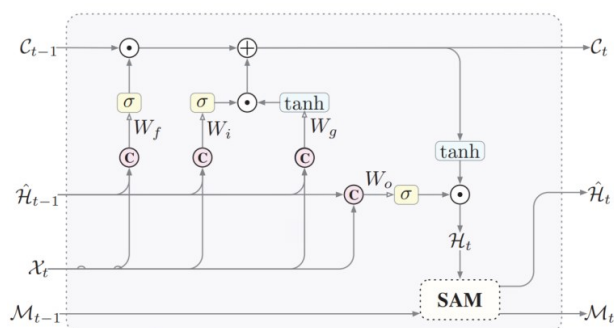


图5 SAM-ConvLSTM模块

Fig. 5 SAM-ConvLSTM Module

预测有促进作用。因此,在自注意力机制的基础上,SAM引入了用来表征全局上下文信息的记忆单元,通过结合当前时间步特征 \mathcal{H}_t 与上一步记忆特征 \mathcal{M}_{t-1} 得到聚合特征的方式,提高了模块对当前时刻的预测能力。具体操作为通过定义多个卷积核方式将 \mathcal{H}_t 与 \mathcal{M}_{t-1} 映射到不同特征空间,分别进行加权与相似度计算,把结果与 \mathcal{H}_t 进行通道合并得到最终的聚合特征。采用门控机制,结合聚合特征将上一步记忆特征 \mathcal{M}_{t-1} 更新为当前时间步特征 \mathcal{M}_t ,最后把 \mathcal{M}_t 与聚合特征整合成 $\hat{\mathcal{H}}_t$ 并输出。SAM的整体结构如图6所示。

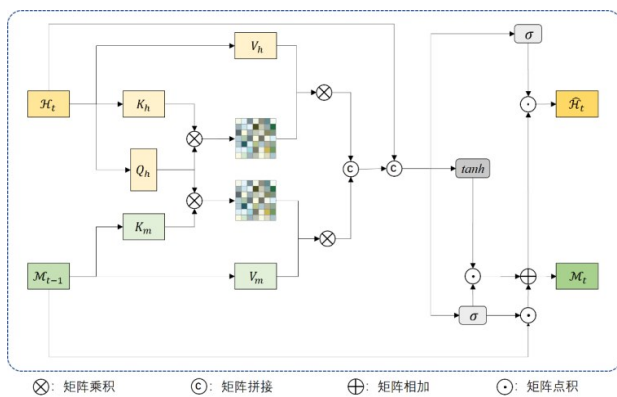


图6 Self-Attention Memory结构

Fig. 6 Architecture of Self-Attention Memory

1.2.2 样本不平衡问题

不同于常规区域的地物分布,北极航道地物种类相对简单,主要包括海水、裸陆地、冰盖陆地、云、独立海冰与浮冰等地物类型。其中,海水分布相对集中,分布特征较为简单,其它地物种类地物形态各异且分布比例悬殊。整体而言,独立海冰在其中占比少,且空间纹理与冰盖陆地、浮冰等地物纹理非常相似,依靠空间结构进行区分难度较大。因此,北极航道独立海冰运动的短时预测在数据分布

上存在样本区分难易和样本数量不平衡两大困难。本研究将从损失函数选择和分割阈值偏移效应两个方面进行改进。

损失函数选择

样本区分难度大与数量分布的不平衡导致传统的标准交叉熵损失函数(CrossEntropyLoss)在模型训练时会使得易分的负样本在总体loss中占的比重过大,进而对占比较少但是较难区分的样本产生淹没,导致模型的性能与优化方向不稳定。因此,本文选择了对样本不平衡效果较好的 $FDWloss^{[18]}$,来优化训练模型的参数,该函数结合了 focalloss^[22] 与 dice 系数特点,函数形式如式(2)所示。

$$FDWloss(pre, tar) = W_1 * Focalloss(pre, tar) + DiceWloss(pre, tar) \quad (2)$$

其中 W_1 为数量级校正因子,作用是保证前者与后者的值处在相同数量级上, pre 是经过 sigmoid 激活函数作用之后的预测概率, tar 是真值图对应的类别概率, $DiceWloss$ 为加权的 dice 系数损失函数。

$Focalloss(pre, tar)$ 的定义如式(3)所示。

$$Focalloss(pre, tar) = \begin{cases} -\alpha(1-pre)^\gamma \log(pre) & tar = 1 \\ -(1-\alpha)pre^\gamma \log(1-pre) & tar = 0 \end{cases} \quad (3)$$

其中 α 为类别权重参数,其值与正负样本比例相关, $(1-pre)^\gamma$ 为样本可分性调节因子。

$DiceWloss$ 是在 Dice 系数的基础上通过分离目标类别与背景类别的 Dice 系数计算方式,引入类别权重参数 W 以消除类别不平衡带来的影响,定义如式(4)所示。

$$DiceWloss(pre, tar) = 1 - W_p \frac{Comm(pre, tar)}{Sum(pre) + Sum(tar)} - W_{1-p} \frac{Comm((1-pre), (1-tar))}{Sum(1-pre) + Sum(1-tar)} \quad (4)$$

其中 $1-pre$ 与 $1-tar$ 分别表示背景预测结果和背景真值, $Comm(pre, tar)$ 表示的是预测结果与真值之间的交集, W_p 是目标类别权重, W_{1-p} 是背景类别权重, $Sum(x)$ 表示对 x 逐点求和的结果。

因此,最终在 Multiloss-SAM-ConvLSTM 中 Multiloss 的定义如式(5)所示。

$$Multiloss = \sum_{k=1}^{Frames} W_{T_k} * FDWloss(pre_{T_k}, tar_{T_k}) \quad (5)$$

其中 pre_{T_k} , tar_{T_k} 分别为 T_k 时刻对应的 pre , tar , W_{T_k} 为时刻 T_k 对应的权重系数, $Frames$ 为数据集所对应的

时序长度。

分割阈值偏移效应

样本区分难度大与数量分布的不均衡除了影响模型的训练过程外,还会导致网络后端的结果输出与正常情况不同。对于二分类问题,通常会把 Sigmoid 函数作为激活层添加到网络的最后一层,将最后的结果归一化到 $[0, 1]$ 的概率分布空间中,再设置一个分割阈值对结果二值化得到最终分类结果。显然,分割阈值的取值会影响模型的最终分类表现,对于绝大多数数据集而言,由于正负样本比例差距不大,最佳分割阈值在 0.5 左右浮动,分割阈值一般设为 0.5。针对此类数据集,可以将训练集和测试集的激活函数输出值整理成新的数据集,再利用传统机器学习方法(如支持向量机等)来探究最佳的分割阈值取值。但是如果数据集中的正负样本比例悬殊,最佳分割阈值就会发生严重偏移。由于比例悬殊的现象在激活函数输出结果所在的数据空间中依然存在,无法用一般的机器学习方法来寻找最佳分割阈值。

针对实验区样本分布不均衡的情况,本文通过网格法得到网络后端分割阈值的最佳取值。分割阈值 α 的取值范围为 $[0, 1]$,在这个范围中以 p 为间隔等间隔地对 α 采样得到数组 $\alpha = [\alpha_1, \alpha_2, \dots, \alpha_n]$,对于一个模型而言,每个 α 都有对应的评价指标输出,通过比较评价指标的优劣,就可以得到数组 α 在以 p 为采样频率下最佳的 α ,即最佳的分割阈值取值。

2 实验与讨论

2.1 精度评价指标

上文提到基于 MERSI-II 的北极航道独立海冰运动的短时预测其实是“时空预测+语义分割”的混合问题。不论是独立海冰位置预测的准确性还是预测时刻海冰形状恢复的精确性,最终预测效果还是体现在预测结果与标签图对应像元的分类准确度上。因此本文引入了全局准确率(Overall Accuracy, OA)、Kappa 系数和均值交并比(Intersection Over Union, IOU)三种语义分割领域常用的评价指标^[18]。此外,由于 SeaiiceMoving 数据集存在样本分布不平衡的问题,为了更全面客观地评价预测效果,本文同时引入了与 F1score 等价的 Dice 系数,一方面可以更加全面地论证实验结果的准确性和可行性,另一方面还可以根据各个评价指标变化趋势的相关性判断它们对样本分布的敏感程度。

OA 表示所有判断正确的结果占总体的比重,其计算式定义如式(6),其中 TP 为预测为正类的正样本, TN 为预测为负类的负样本, FP 为预测为正类的负样本, FN 为预测为负类的正样本,下同。

$$OA = \frac{TP + TN}{TP + FP + FN + TN} \quad (6)$$

Kappa 是用来评价图像分割算法精确度的参数,其计算式定义如式(7)。式中, p_0 为整体分类精度, pe 为期望分类精度。Kappa 的值越接近 1,表示分割效果越理想。

$$Kappa = \frac{p_0 - pe}{1 - pe} \quad (7)$$

IOU 代表图像分割算法某一类别的分割结果和真实值的交集与并集的比值,IOU 的值越接近 1,表示海冰预测恢复的效果越理想,具体计算方法如式(8)。

$$IOU = \frac{TP}{TP + FP + FN} \quad (8)$$

Dice 系数是医学图像分割领域常用的评价指标,主要是用来计算分割结果与真实值的相似性,其计算式定义如式(9)。Dice 系数的值越接近 1,表示海冰预测恢复的效果越理想。

$$Dice = \frac{2*TP}{2*TP + FP + FN} \quad (9)$$

2.2 实验环境

本文实验采用的深度学习框架为 pytorch,版本号为 1.3.0,模型训练所使用的 GPU 为单卡 NVIDIA 2080 TI, GPU 显存为 12 G。整个训练过程的各项超参数设置如下,批处理尺寸(Batch Size)为 24,总 Epoch 为 1 000,采用 Early Stopping 机制避免模型出现过拟合,初始学习率设置为 10^{-3} ,最小学习率设置为 10^{-4} ,采用余弦函数进行学习率衰减,并通过 Adam 优化器对模型参数进行更新。模型的整体训练流程可以从图 2 中体现,即将 T_0 、 T_{100} 、 T_{200} 三个时刻的数据作为输入,经过网络运算后分别计算 T_{100} 、 T_{200} 、 T_{300} 对应输出与标签真值图的 Multiloss 后,更新网络权重。

2.3 实验结果讨论

如前文所述,由于北极航道独特的地物分布情况,SeaiiceMoving 数据集中存在明显的独立海冰与背景分布不均衡现象。样本分布不均衡会导致网络后端最佳分割阈值的取值出现偏移效应进而影响模型最终的预测准确度。为了得到当前模型下最优的独立海冰运动短时预测结果,本文探究了最佳分割阈值偏移效应对最终预测结果的影响,以

Multiloss-SAM-ConvLSTM 的最优模型输出结果作为待验证数据,利用重叠度为 45% 的重叠消边法^[18]分别计算全局 OA、Kappa 系数、IOU 和 Dice 系数,优化得到了 SeaiiceMoving 数据集空间下的最佳分割阈值取值。实验结果如图 7 所示。

实验结果表明,当分割阈值的取值范围为 [0.1, 0.9] 之间时, Kappa 系数、IOU 和 Dice 系数的变化趋势基本一致。同时, Kappa 系数、IOU 和 Dice 系数均在分割阈值取值为 0.3 时达到最优。与上述三种指标不同的是,随着分割阈值取值的增大, OA 变化趋势为先加速上升再缓慢下降,整体变化幅度很小约 0.3% 左右,且值域分布在 98~98.5% 之间。独立海冰在 SeaiiceMoving 数据集中占比只有 0.03 左右。根据 OA 的定义,测试集中预测正确的背景像元数 TN 在分子与分母中占比都非常大,导致 OA 的值非常接近 1。同时,独立海冰运动的短时预测任务对独立海冰的位置定位与形状恢复要求高,在样本属性中属于难分正样本;而样本中绝大部分的背景地物是较易区分的海水,属于易分负样本。样本难易分布不均衡对 Multiloss-SAM-ConvLSTM 预测能力的影响可以通过 *FDWloss* 削弱,但是不能完全消除。当分割阈值在 [0.1, 0.9] 范围取值时, TP+TN 的像元绝对变化量比较小,进而导致 OA 的浮动范围只有 0.3% 左右。因此在样本不平衡时, OA 无法客观地对 Multiloss-SAM-ConvLSTM 的预测能力进

行评价,不适合作为评价指标。综上所述,本实验论证了样本不平衡时网络后端最佳分割阈值存在偏移效应,针对本文的 SeaiiceMoving 数据集,最佳的分割阈值取值为 0.3。

为了验证 Multiloss-SAM-ConvLSTM 的有效性,本文以 SAM-ConvLSTM 为对比方法进行实验,两者的各项实验条件一致。训练时均使用 *FDWLoss* 作为损失函数,对预测的结果使用 45% 重叠度进行有消边地拼接,实验结果如表 4 所示。从表 4 可以看到,本文提出的 Multiloss-SAM-ConvLSTM 在各项指标上均优于 SAM-ConvLSTM,其中 OA 达到了 98.46%, Kappa 系数提高了 0.1, IOU 提高了 0.12, Dice 系数提高了 0.1。以上实验结果充分证明了本文提出的通过增加中间层的损失函数加强 SAM-ConvLSTM 在时空预测过程中对空间维的约束,来强化网络深层对独立海冰空间特征的获取,从而减少了独立海冰空间特征信息在时间维传递过程中的损失是有效的。本文方法在保持了网络对独立海冰短时运动预测时位置信息预测能力的前提下,进一步提高了预测结果中独立海冰形状的恢复水平。

海冰运动预测实验的部分结果如图 8 所示。其中 (a-d) 为 T_0 、 T_{100} 、 T_{200} 和 T_{300} 时刻海冰分布的真值图, (e) 为使用 Multiloss-SAM-ConvLSTM 训练的模型对应结果图, (f) 为使用 SAM-ConvLSTM 训练的模

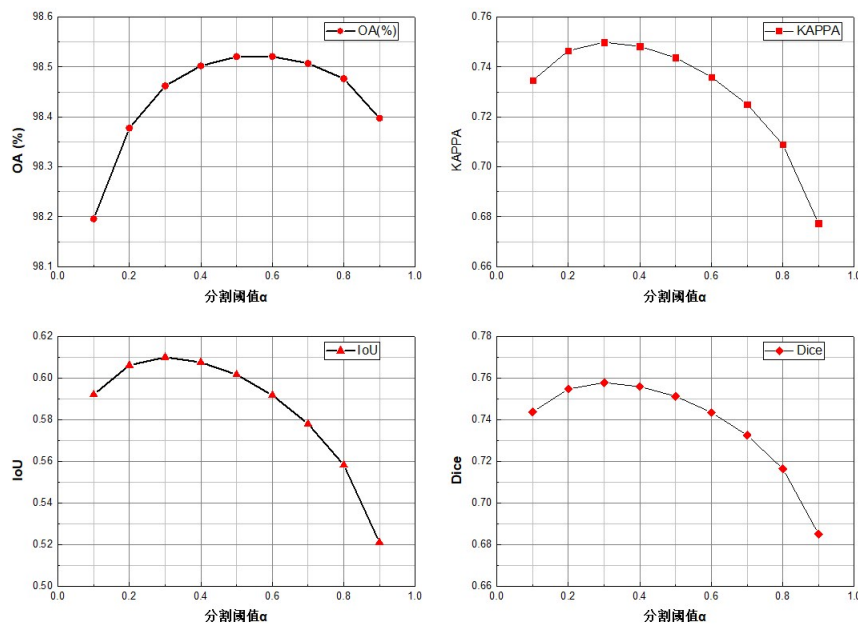


图 7 不同分割阈值下各个指标的变化趋势图

Fig. 7 The changing trend of each indicator under different segmentation threshold

表4 不同海冰运动预测算法的精度对比

Table 4 Comparison of accuracies for motion prediction of sea ice with different methods

	准确率	Kappa	IOU	Dice系数
Multiloss-SAM-ConvLSTM	98.46%	0.75	0.61	0.76
SAM-ConvLSTM	97.88%	0.65	0.49	0.66

型对应结果图,(g)为对应区域的可视化图。从实验结果图中可得,深度学习方法对北极航道中独立海冰运动的短时预测并提取是行之有效的,在具备对独立海冰运动短时预测能力的同时,还能有效抵抗北极航道中与独立海冰相似的地物背景(雪地,带有冰盖的海岸线附近等)的干扰(图8中第一行)。相较于SAM-ConvLSTM,本文提出的Multiloss-SAM-ConvLSTM在大、小型独立海冰的位置预测和形状恢复能力上均有提升。从图8第二行的结果中可以看到Multiloss-SAM-ConvLSTM对下个时刻独立海冰的整体位置以及运动过程中由于自身旋转产生角度偏移的预测能力更强。从图8第三至五行的结果图中可以看到Multiloss-SAM-ConvLSTM对独立海冰运动后的整体形状以及边缘细节恢复效

果更优。此外,还可以减少SAM-ConvLSTM结果中小型海冰发生“黏连”的情况(图8中第五行的(e-f))。

北极多云雾天气,抛开厚云遮挡导致完全无法看到地物的情况,大部分薄云场景下的地物虽然受到云干扰,但是仍然具备可识别性。若能克服薄云干扰,准确预测薄云下独立海冰的短时运动轨迹,就可以大大拓宽北极航线动态规划的时间窗口。此外,本底数据集范围比较大,SeaiceMoving数据集采用随机裁剪得到小图像块的方式制作,这种制作方式会导致大量独立海冰被分割,随着独立海冰不断运动, T_0 、 T_{100} 、 T_{200} 和 T_{300} 4个时刻影像的边缘部分会出现独立海冰逐渐消失或者出现的“位形联动”现象(图9)。针对这两个问题,本文探讨了Multi-

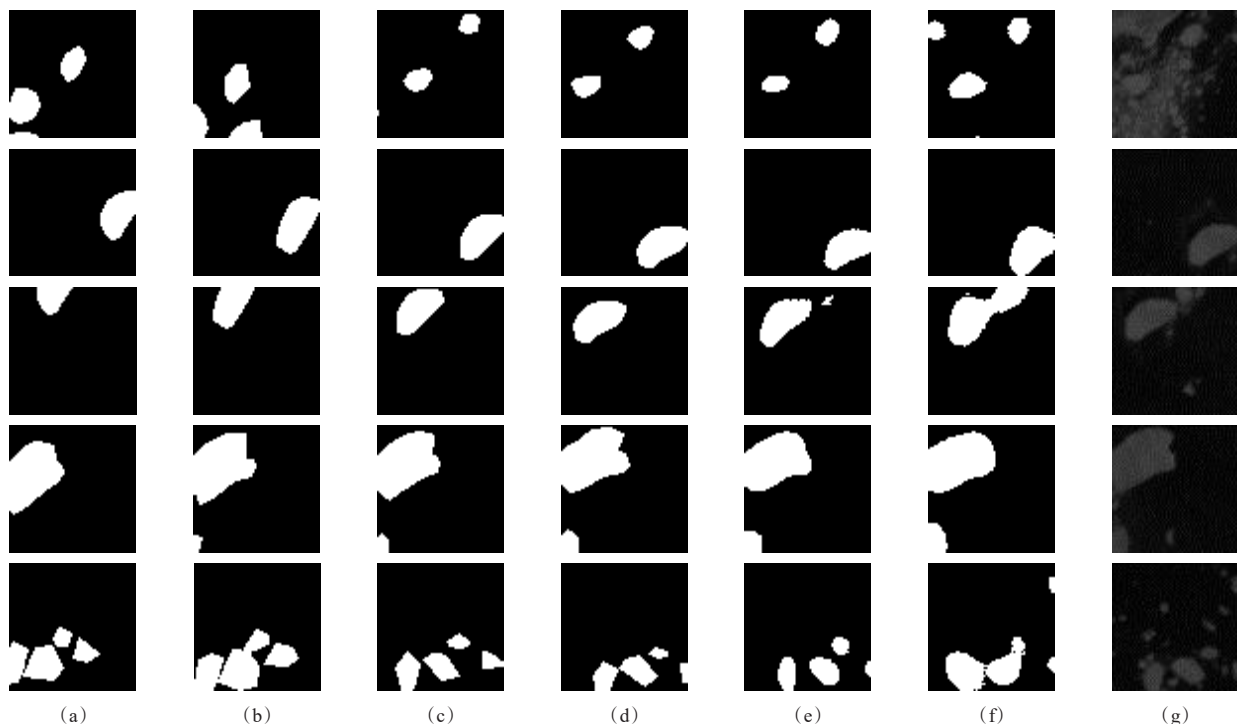


图8 不同方法的海冰运动预测结果(a)-(d) $T_0 - T_{300}$ 人工解译真值图(e) Multiloss-SAM-ConvLSTM 预测的 T_{300} 时刻海冰分布结果(f)SAM-ConvLSTM 预测的 T_{300} 时刻海冰分布结果(g) T_{300} 时刻的真实海冰分布图

Fig. 8 Independent sea ice motion prediction results by different methods (a-d) manually interpreted images at T_0 to T_{300} , (e) results of sea ice distribution at T_{300} predicted by Multiloss-SAM-ConvLSTM, (f) results of sea ice distribution at T_{300} predicted by SAM-ConvLSTM, (g) real distribution of sea ice at T_{300}

loss-SAM-ConvLSTM 对薄云下独立海冰运动的预测能力以及当部分数据集边缘处特有的“位形联动”现象发生时,对独立海冰的定位和形状恢复能力。具体结果如图 9 所示。从前三行的(f-g)可以看到,当 T_{300} 时刻受到云干扰时,独立海冰的可分性大大下降,但是从对应的(e)中可见, Multiloss-SAM-ConvLSTM 对运动后受到云干扰的独立海冰仍然可以准确定位并恢复形状。此外,第四、五行的结果表明本文算法的预测能力不会随数据集边缘处特有的“位形联动”现象而下降,可以精确捕捉此类海冰运动后的位置与旋转变化并恢复形状。

3 结论

基于风云三号 MERSI-II L1 时序影像,在 SAM-ConvLSTM 的基础上提出了 Multiloss-SAM-ConvLSTM 网络对北极航道独立海冰运动进行短时预测。考虑到 SeaiceMoving 数据集中独立海冰占比少,采用加权 FDWloss 来减弱样本不平衡对网络预测能力的影响。针对网络后端最佳分割阈值的偏移效应,为了得到最优的海冰运动预测结果,探究了分割阈值变化对各项指标的影响,并在实验中论证了 Sea-

iceMoving 数据集的最佳分割阈值为 0.3。以格陵兰岛经加拿大北部北极群岛到阿拉斯加北岸的西北航道为研究区,将 Multiloss-SAM-ConvLSTM 与 SAM-ConvLSTM 的预测结果进行对比,实验结果表明 Multiloss-SAM-ConvLSTM 各项评价指标均更加优异,不仅提高了对独立海冰位置与旋转特性的预测,对独立海冰形状的恢复能力也更加出色,同时,对薄云干扰下的独立海冰运动仍然可以有效预测。

综上,本研究可以为北极航线规划和动态修正提供较为全面且精准的技术支撑。但是,基础网络 SAM-ConvLSTM 在特征提取的过程中缺少多尺度的采样,导致网络缺少对大块海冰的边缘与极小块海冰本身敏感的高阶语义特征的获取,对小型海冰的预测效果一般,对其他独立海冰预测过程中边缘等细节部分的恢复能力有待提高。后续将通过基础网络 SAM-ConvLSTM 本身的结构进行改进,进一步提高北极航道独立海冰运动预测能力。

References

- [1] WU Zhan-Kai, WANG Xing-Dong, WANG Cheng. Sea Ice Distribution Detection Based on FY-3 Satellite MWRI

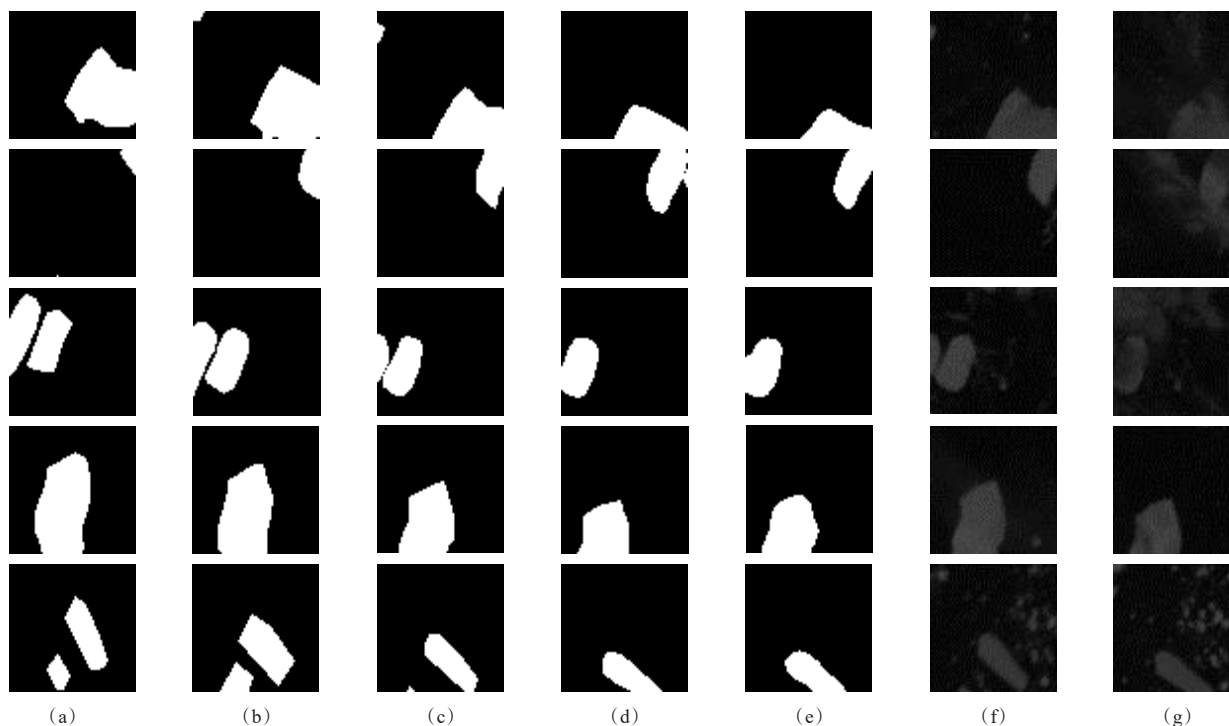


图 9 薄云与“位形联动”现象干扰下的预测结果 (a)-(d) $T_0 - T_{300}$ 人工解译真值图 (e) Multiloss-SAM-ConvLSTM 预测的 T_{300} 时刻海冰分布结果 (f) T_{200} 时刻的真实海冰分布图 (g) T_{300} 时刻的真实海冰分布图

Fig. 9 (a-d) Manually interpreted images at T_0 to T_{300} , (e) results of sea ice distribution at T_{300} predicted by Multiloss-SAM-ConvLSTM, (f) real distribution of sea ice at T_{200} , (g) real distribution of sea ice at T_{300}

- Data [J]. *Bulletin of Surveying and Mapping*, 2018, **0** (10): 56–60, 65. (吴展开, 王星东, 王成. 利用 FY-3 卫星 MWRI 数据探测海冰分布[J]. *测绘通报*), 2018, **0** (10): 56–60, 65.
- [2] Petrou Z I, Tian Y. Prediction of sea ice motion with convolutional long short-term memory networks[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2019, **57**(9): 6865–6876.
- [3] Georgy Ayzel, Maik Heistermann, Tanja Winterrath. Optical flow models as an open benchmark for radar-based precipitation nowcasting[J]. *Geoscientific Model Development*, 2019, **12**(4).
- [4] Cheung P, Yeung H Y. Application of optical-flow technique to significant convection nowcast for terminal areas in Hong Kong[C]. In the 3rd WMO International Symposium on Nowcasting and Very Short-Range Forecasting (WSN12), 2012:6–10.
- [5] Hidetomo Sakaino. Spatio-Temporal Image Pattern Prediction Method Based on a Physical Model With Time-Varying Optical Flow[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2013, **51**(5–2).
- [6] Liu Y, Xi D G, Li Z L, *et al.* A new methodology for pixel-quantitative precipitation nowcasting using a pyramid Lucas Kanade optical flow approach [J]. *Journal of Hydrology*, 2015, **529**(529):354–364.
- [7] SHI Yu-li, SHI Sheng-wei. Research on Accuracy Evaluation of Optical Flow Algorithm in FY4A Infrared Image Extrapolation [J]. *Journal of Ordnance Equipment Engineering*, 2021, **42**(09): 150–158, 224. (石玉立, 施声伟. 光流算法在 FY4A 红外图像外推中的精度评价研究[J]. *兵器装备工程学报*), 2021, **42**(09): 150–158, 224.
- [8] Xingjian SHI, Chen Z, Wang H, *et al.* Convolutional LSTM network: A machine learning approach for precipitation nowcasting [C]. *Advances in neural information processing systems*. 2015: 802–810.
- [9] Marc'Aurelio Ranzato, Arthur Szlam, Joan Bruna, *et al.* Video (language) modeling: a baseline for generative models of natural videos [J]. *arXiv preprint arXiv:1412.6604*, 2014.
- [10] Srivastava N, Mansimov E, Salakhudinov R. Unsupervised Learning of Video Representations using LSTMs [C]. *International conference on machine learning*. PMLR, 2015: 843–852.
- [11] Wang Y, Long M, Wang J, *et al.* Predrnn: Recurrent neural networks for predictive learning using spatiotemporal lstms[C]. *Proceedings of the 31st International Conference on Neural Information Processing Systems*. 2017: 879–888.
- [12] Wang Y, Gao Z, Long M, *et al.* Predrnn++: Towards a resolution of the deep-in-time dilemma in spatiotemporal predictive learning [C]. *International Conference on Machine Learning*. PMLR, 2018: 5123–5132.
- [13] Wang Y, Zhang J, Zhu H, *et al.* Memory in memory: A predictive neural network for learning higher-order non-stationarity from spatiotemporal dynamics [C]. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2019: 9154–9162.
- [14] He K, Zhang X, Ren S, *et al.* Deep residual learning for image recognition [C]. *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016: 770–778.
- [15] Shi X, Gao Z, Lausen L, *et al.* Deep learning for precipitation nowcasting: A benchmark and a new model[J]. *arXiv preprint arXiv:1706.03458*, 2017.
- [16] Lin Z, Li M, Zheng Z, *et al.* Self-attention convlstm for spatiotemporal prediction [C]. *Proceedings of the AAAI Conference on Artificial Intelligence*. 2020, **34** (07) : 11531–11538.
- [17] Kang D, Im J, Lee M I, *et al.* The MODIS ice surface temperature product as an indicator of sea ice minimum over the Arctic Ocean[J]. *Remote sensing of environment*, 2014, **152**: 99–108.
- [18] ZHENG Fu-Qiang, KUANG Ding-Bo, HU Yong, *et al.* Refined segmentation method based on U-ASPP-Net for Arctic independent sea ice[J]. *J. Infrared Millim. Waves*, 2021, **40**(6): 798–808. (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.
- [19] ZHOU Ying. The Research on Monitoring the Motion and Change of the Arctic Sea Ice Based on the FY-3 Meteorological Satellite Data [D]. *University of Chinese Academy of Sciences (中国科学院大学)*, 2013.
- [20] ZHOU Ying, GONG Cai-Lan, HU Yong, *et al.* Extraction of Sea Ice Information in the Arctic Ocean by Using FY-3/MERSI Data [J]. *Journal of Atmospheric and Environmental Optics*, 2013, **8**(1): 53–59. (周颖, 巩彩兰, 胡勇, 等. 风云三号 MERSI 数据提取北冰洋海冰信息方法研究[J]. *大气与环境光学学报*), 2013, **8**(1): 53–59.
- [21] Vaswani A, Shazeer N, Parmar N, *et al.* Attention is all you need [C]. *Advances in neural information processing systems*. 2017: 5998–6008.
- [22] Lin T Y, Goyal P, Girshick R, *et al.* Focal loss for dense object detection [C]. *Proceedings of the IEEE international conference on computer vision*, 2017: 2980–2988.