

极光亚暴爆发时点机器识别方法*

蒋家楠^{1,2} 邹自明^{1,3} 陆 阳^{1,3}

1(中国科学院国家空间科学中心 北京 100190)

2(中国科学院大学 北京 100049)

3(国家空间科学数据中心 北京 101407)

摘 要 极光亚暴是地球磁场与太阳风相互作用产生的一种地磁扰动现象, 对于其爆发时点的准确识别有助于深入理解其背后的物理机制. 现有的极光亚暴爆发时点机器识别方法在标准上与人工识别存在差异, 且通常需要经过复杂的图像预处理和人工调参. 为了实现与人工识别标准一致的极光亚暴爆发时点机器识别方法, 设计了两种识别策略, 旨在解决在复现人工识别标准时遇到的图像序列长度不定长问题. 研究采用深度学习方法, 并提出了一种基于 CBAM 注意力的 EfficientNet, 以此作为重要组件来构建模型. 使用 Polar 卫星 1996–1998 年的紫外极光图像对模型进行训练, 并在 1999–2000 年的图像数据上进行测试. 模型识别精确率可达 0.98, 识别效率可达 36.93 frame·s⁻¹. 该模型不仅摆脱了现有模型对于图像预处理的依赖, 还能够适用于真实观测下图像序列不等长以及暴时序列与非暴时序列样本数量极端不均衡的情况, 具有较高的实用性.

关键词 极光亚暴, 爆发时点, 机器识别, 深度学习

中图分类号 P352

Machine Identification Method of Auroral Substorm Onset Time

JIANG Jianan^{1,2} ZOU Ziming^{1,3} LU Yang^{1,3}

1(National Space Science Center, Chinese Academy of Sciences, Beijing 100190)

2(University of Chinese Academy of Sciences, Beijing 100049)

3(National Space Science Data Center, Beijing 101407)

Abstract Auroral substorm is a geomagnetic disturbance resulting from the interaction between Earth's magnetic field and the solar wind. The accurate identification of the onset times is crucial for a deep understanding of the underlying physical mechanisms. The existing machine methods for auroral substorm identification differ from manual identification standards and typically require complex image preprocessing and parameter tuning by manual. To achieve a machine model consistent with manual

* 国家重点研发计划专项项目 (2022YFF0711400), 中国科学院“十四五”网络安全和信息化专项项目 (CAS-WX2022SDC-XK15) 和中国科学院网信专项项目 (CAS-WX2022SF-0103) 共同资助

2024-03-12 收到原稿, 2024-04-26 收到修定稿

©The Author(s) 2025. This is an open access article under the CC-BY 4.0 License
(<https://creativecommons.org/licenses/by/4.0/>)

identification standards, this paper designs two identification strategies aimed at addressing the issue of variable image sequence lengths encountered in replicating manual standards. Based on deep learning methods, this paper proposes an EfficientNet model featuring CBAM attention as a key component for model construction. The model is trained using ultraviolet auroral images from the Polar satellite between 1996 and 1998 and tested on image data from 1999 to 2000. The model achieves an identification accuracy of up to 0.98 and an efficiency of 36.93 frames per second. This model not only eliminates the reliance on image preprocessing present in existing models but also adapts to real-world observations with unequal image sequence lengths and extreme imbalances in the number of samples between sub-storm and non-substorm sequences, demonstrating its high practicality.

Key words Auroral substorm, Onset time, Machine identification, Deep learning

0 引言

亚暴是一种由磁层能量释放引发的地磁扰动现象,其间会伴随着极光的剧烈活动^[1],Akasofu^[2]将极光从平静经过各种活跃阶段最后重新恢复平静的过程称作极光亚暴。作为空间天气变化的一个重要现象,极光亚暴的相关数据能够为空间天气模型提供关键信息,进而帮助空间物理学家更好地理解太阳活动对地球磁层和大气层的影响并对其他空间天气事件进行研究。因此,实现对极光亚暴的准确识别成为空间物理学领域内的一个重要课题。

对于极光亚暴的识别,目前包括两种方法。一种方法是利用空间环境参数对亚暴的指示作用来对其进行识别。Juusola等^[3]提出了一种基于极光电集流指数(Auroral Lower Electrojet Index, AL)指数的亚暴起始时刻确定方法。Partamies等^[4]结合行星际磁场改进了基于AL指数的亚暴识别方法。Ohtani等^[5]使用SuperMAG极光电集流指数SML指数开发了一种更严谨的亚暴识别算法,不仅能够实现亚暴爆发时刻的检测,还可以对亚暴不同阶段进行一定程度的估计。然而,鉴于极光亚暴的发生与空间环境参数之间存在的时间延迟,这些识别方法不可避免地会引入一定的误差。另一种方法则是使用卫星观测的紫外极光图像。在早期,学者主要依靠肉眼观察。极光亚暴爆发时,极光卵中的某个区域会突然变亮。在此之后的数十分钟内,极光强度持续增强,极光卵显著扩张,这些特征成为人工识别的主要依据。Frey等^[6]通过对磁层-极光全球探测卫星(Imager for Magnetopause-to-Aurora Global Exploration, IMAGE)远紫外宽波带成像相机(Far Ultraviolet Wideband

Imaging Camera, WIC)图像的观察和统计,发布了2000年5月至2002年12月的极光亚暴列表。Liou^[7]通过观察极地卫星Polar紫外成像仪(Ultraviolet Imager, UVI)的图像数据,给出了1996—2008年的极光亚暴爆发时点及位点列表。然而,随着卫星传回数据的日益累积,人工识别方法已经无法满足对海量数据进行高效、精确分析的需求。为了应对这一问题,研究开始转向自动化技术。Yang等^[8]通过模糊C均值聚类法提取极光图像中的亮斑,再借助极光强度等特征来确定极光亚暴的起始时刻,但识别的精确率不高,并且需要人工设置多个阈值。Yang等^[9]提出了一种基于形状约束的稀疏低秩矩阵分解算法,初步实现了大规模图像数据的极光亚暴自动识别,但该方法需人工参与且运行效率低。Han等^[10]构建了基于EfficientNet的亚暴事件识别模型,大大提高了极光亚暴事件识别的准确性。然而该方法涉及的眼动信息收集和序列指纹构建过程对硬件要求过高,实现起来较为复杂。Yang等^[11]建立了一种基于双流网络的亚暴自动识别方法,此方法是一种多阶段的识别方法,但其涉及的光流图计算对硬件资源同样有较高的要求,这些因素可能会影响该方法在实际应用中的实用性。由此可见,目前已有的基于图像的极光亚暴爆发时点机器识别方法在准确性和实用性上都有待提高。

根据Liou^[7]的识别标准,确认极光亚暴的发生至少需要观察10 min的图像序列。在Polar卫星的实际观测中,10 min的图像序列包含2~17 frame的图像。这对极光亚暴的机器识别提出了新的挑战,即能够应对变长的图像序列。然而,现有的识别方法大多仅依赖于亚暴爆发时刻前后的图像,这种做法与人工识别的标准存在差异。为了克服这一局限性,研究提

出了两种针对变长图像序列的极光亚暴爆发时点识别策略, 包括通道填充策略和短时记忆策略, 分别通过对图像通道维度进行填充和对前序图像的时空特征进行记忆的方式来实现对变长图像序列的处理. 基于以上两种策略, 设计了一种基于 CBAM 注意力的 EfficientNet, 用于构建识别模型. 提出的极光亚暴爆发时点识别模型不仅实现了极光亚暴的高精度识别, 还克服了现有识别技术在处理变长图像序列方面的局限性, 为极光亚暴的实时监测和分析提供了一种更为可靠和高效的解决方案.

1 数据

1.1 极光图像数据

研究使用的数据集来源于美国宇航局 (National Aeronautics and Space Administration, NASA) 公开的 Polar/UVI 数据, 卫星在轨期间拍摄了 1996 年 3 月—2000 年 2 月的北半球紫外极光图像. 紫外成像仪^[12]UVI 是一个 2 D 快拍类型的照相机, 能够拍摄远紫外极光图像. 在紫外波段范围内, 由于 LBHL (Lyman-Birge-Hopfield-Long) 波段的强度与沉降粒子的能通量成正比^[13], 因此研究使用了 LBHL 波段的极光图像. Polar 卫星官网^{*}公布了.cdf 格式的文件, 其中内容包含了 LBHL 波段的图像. 首先使用官方的处理

软件^{**}从.cdf 格式的文件中读取 LBHL 波段的图像, 同时对其进行初步的预处理, 包括视野校正和坐标变换. 图 1(a) 是一张由官方处理软件处理过后的图像, 其中极光图像已经被映射到了 APEX^[14] 坐标系 (Apex Coordinate System) 下, 并且归一化到对数尺度下^[15]. 之所以使用对数尺度下的图像, 是因为相比于线性尺度下的图像, 对数尺度能够对中低强度的极光进行一定程度的放大, 从而更清晰地展示出极光强度的分布.

在训练模型之前还需要对图像进行进一步的预处理. 首先需要从图像上去除坐标系、标题和色条, 这样可以得到一张只有极光的图像, 以上操作均在官方处理软件中完成. 接着对图像进行剪裁, 仅保留大小为 240×240 像素的极光部分. 由于图像中的地磁纬度范围为 $60^\circ \sim 90^\circ$, 因此预处理后图像的纬度分辨率是 0.25° . 为了确保神经网络能够快速收敛, 还需要将图像中的像素强度归一化到 $[0, 1]$ 范围内. 最终预处理后的图像如图 1(b) 所示.

1.2 极光亚暴数据

极光亚暴爆发时点数据来自 Liou 提供的极光亚暴时点列表, 在 SuperMAG 网站^{***}下载. 当极光在某一时刻突然变亮, 并且之后出现长达 10 min 的强度增强和区域扩张时, Liou 的识别标准认为该时刻是极光亚暴的爆发时点^[7]. 由于 Liou 使用了多个波段的图

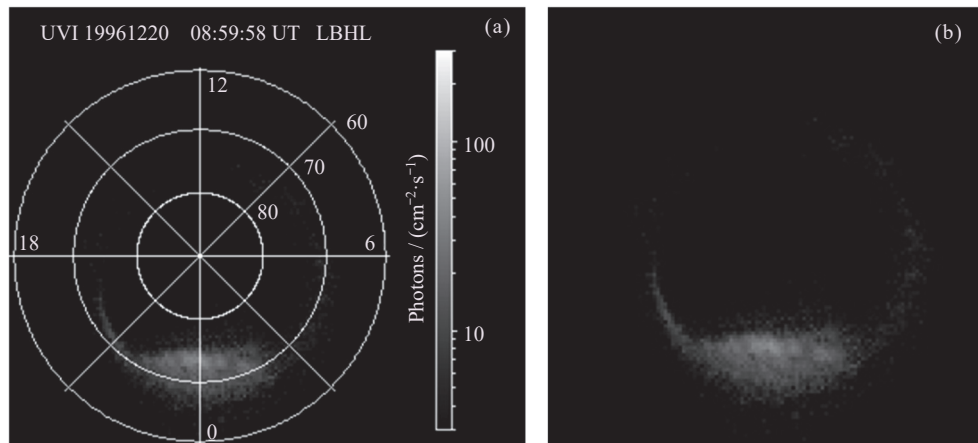


图 1 Polar/UVI 样图. (a) 由官方处理软件预处理后的图像, (b) 去除坐标系、标题和色条的图像

Fig. 1 An example of Polar/UVI image. (a) Image preprocessed by official studio software.

(b) Image with the coordinate system, title, and color bar removed

* <https://spdf.gsfc.nasa.gov/pub/data/polar/uvi/>

** https://figshare.com/articles/dataset/POLAR_UVI_calibration_toolkit/5197084

*** <https://supermag.jhuapl.edu/substorms>

像进行识别, 不是仅使用了 LBHL 波段图像, 因此
在一些爆发时点之后的 10 min 内, 可能不存在
LBHL 波段的图像. 由于需要使模型学习处于不同时期 (暴时或者非暴时) 的极光图像序列的特征, 因此
应保证暴时的极光图像序列包含超过 1 帧图像, 根据
该标准剔除了 21 个爆发时点, 使用剩余的 1982 个爆
发时点数据来进一步构建用于极光亚暴爆发时点识
别的图像序列数据集.

1.3 样本构造与数据集构建

由于 Liou 在识别极光亚暴爆发时点时要求至少
存在 10 min 的极光强度增强及区域扩张, 因此以
10 min 的图像序列为单位来对极光亚暴爆发时点进
行识别, 最大程度上还原了人工识别的标准. 通过图
像序列来识别极光亚暴爆发时点的问题等价于判断
图像序列是否是亚暴爆发后的序列, 本质上是一个二
分类问题. 因此, 极光亚暴列表中爆发时点之后
10 min 内的图像就可以构成一个极光亚暴爆发后的
图像序列样本, 也可称为暴时图像序列样本, 即二
分类问题中的正样本, 被标注为 1. 与正样本相反, 负样
本应为不处在爆发时点之后的图像序列, 也可称为非
暴时图像序列样本. 由此设计两种负样本构建方法
来探究样本构造对于模型识别性能的影响, 对应两
组训练/测试数据集.

第一组数据集中, 负样本由极光亚暴爆发时点
之前 30 min 内的图像组成. 在爆发时点前的 30 min
内, 每 10 min 的图像构成一个非暴时序列样本, 即
负样本, 被标注为 0. 这种负样本的构成能够使模
型具有区分爆发时点前后图像序列的能力, 但由于
模型只学习到了爆发时点附近图像序列的特征, 因
此对于不在爆发时点附近的图像序列, 模型可能无
法做出准确的判断. 由此, 设计第二组数据集, 其
中的负样本由所有不处在爆发时点之后的非暴时图
像序列组成. 具体而言, 对于所有不在爆发时点后
2 h 内的图像, 每

10 min 的图像将被组合成一个非暴时图像序列, 作
为负样本, 被标注为 0. 这一负样本构造过程排除
了爆发时点后 2 h 内的图像, 是因为极光亚暴从开
始到结束通常需要大约 2 h. 这样的筛选标准确保
了所选取的负样本不会包含在极光亚暴活动的任
何阶段内, 保证了模型所学习到的样本特征的准确
性和可靠性.

两组数据集中 1996—1998 年的正负样本组成
训练集, 随机选取其中的 80% 用作训练, 剩余的
20% 用作验证, 1999—2000 年的图像序列样本
则构成测试集. 根据以上样本构造和数据集划分方
法, 得到的训练集和测试集中的序列数量和图像数
量的情况列于表 1.

在两组数据集中, 非暴时序列数量都多于暴时
序列数量, 二者比例分别约为 3:1 和 55:1. 尤其
在第二组数据集中, 正负样本数量处于极度不均
衡的状态, 最大程度地还原了真实观测中极光亚
暴发生的低频性和非亚暴状态的普遍性. 使用这
组数据集进行实验, 有助于评估模型在真实极光
亚暴发生情况下的识别能力.

图 2 给出了两组数据集上序列长度的分布情
况. 在两组数据集中, 图像序列的长度为 2~17
frame, 其中以 4 frame 和 17 frame 的序列为主.
尽管在训练集中短序列的数量较多, 但长序列也
有足够的样本量; 同样, 在测试集中虽然长序列
占多数, 但短序列的数量也是足够的, 能够确保
模型在不同长度的图像序列上进行学习和评估.

2 极光亚暴爆发时点识别模型

2.1 针对变长图像序列的识别策略

在构建神经网络模型时, 通常需要统一输入
张量的形状, 但是在 1.3 节构建的数据集中, 图
像序列在时间这一维度上的大小是不一致的, 这
就导致无法直

表 1 两组数据集的序列与图像数量
Table 1 Number of sequences and images in two datasets

	训练集1	测试集1	训练集2	测试集2
暴时序列数量	1367	615	1367	615
非暴时序列数量	3740	1772	72028	37666
暴时图像数量	7382	8545	7382	8545
非暴时图像数量	17926	23712	337442	447926

接将图像序列输入到模型当中. 针对这一问题, 设计了两种策略, 分别是通道填充策略和短时记忆策略.

2.1.1 通道填充策略

通道填充策略 (Channel Filling Strategy, CFS) 指的是将时间维度调换到通道维度上, 并将所有的样本填充至相同的通道数, 该策略的流程如图 3 所示. 对于一个图像序列样本, 其形状为 (Frame, Height, Weight, Channel), 其中 Frame 为帧数, Height 为图像的高, Weight 为图像的宽, Channel 为图像的通道数, 对于 1.3 节所建立的数据集, 高和宽的值固定为 240, 通道数固定为 1, 而帧数是不固定的. 由于图像通道数为 1, 因此可以将帧数维度调整到通道维度上, 如此每个样本被重塑为 (Height, Weight, Frame) 的形状. 接着用 0 将新的通道维度填充到 20, 这样得到的样本的形状就都变成 (240, 240, 20). 以一个帧数为 10 的图像序列为例, 其初始的形状为 (10, 240, 240, 1), 首先将其重塑为 (240, 240, 10), 再用 0 将最后一个维度填充到 20, 所得到的图像序列的尺寸就是 (240, 240, 20). 使用的图像序列的帧数最多为 17 frame, 这里预留了 3 frame, 以防出现更多帧数的

情况. 最终填充后的图像序列将被输入到识别模型 (Identification Model) 中, 识别模型中的主干网络 (Backbone Network) 可由任意结构的神经网络组成, 用于提取特征. 之后再通过一个多层感知输出器 (MLP Head) 输出其所属类别 (Class), 即是否为暴时序列.

2.1.2 短时记忆策略

在短时记忆策略 (Short Memory Strategy, SMS) 中, 一个序列中的图像将被逐帧输入到识别模型中的主干网络中, 并将主干网络提取到的当前时刻图像的空间特征图与融合层在上一时刻得到的时空特征图进行融合, 实现对前序图像特征的记忆. 图 4 给出了短时记忆策略的流程图, 短时记忆策略对应的识别模型包括三个部分: 主干网络、融合层 (Fuse Layers) 和多层感知输出器. 某一时刻的图像输入到主干网络中得到该时刻图像的空间特征图 f_t , 同时将上一时刻通过融合层获得的时空特征图 h_{t-1} 与 f_t 一起输入到融合层中进行融合, 获得当前时刻的时空特征图 h_t , 并将在下一时刻送入融合层与下一时刻图像的空间特征图进行融合. 融合层在最后一个时刻上的输

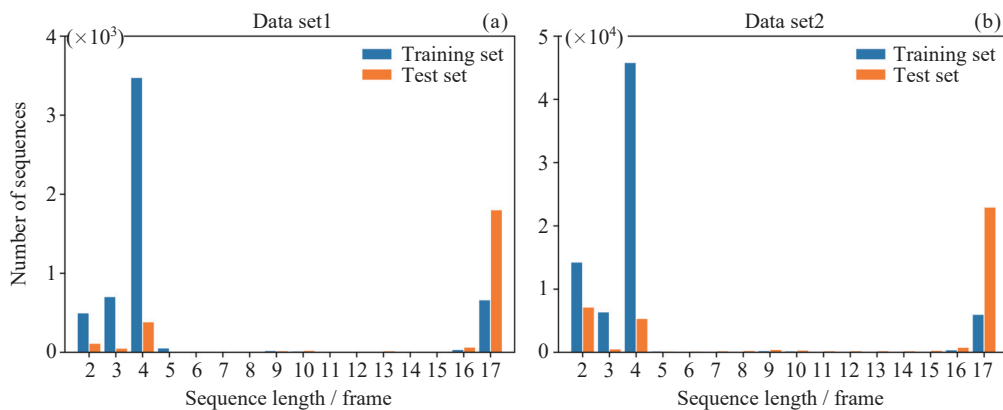


图 2 两组数据集上序列长度分布情况

Fig. 2 Distribution of sequence frames in two datasets

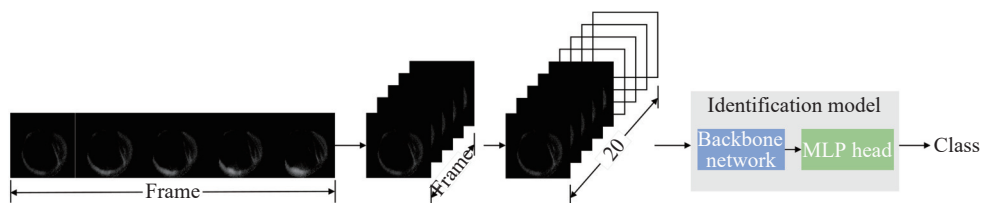


图 3 通道填充策略流程

Fig. 3 Flowchart of channel filling strategy

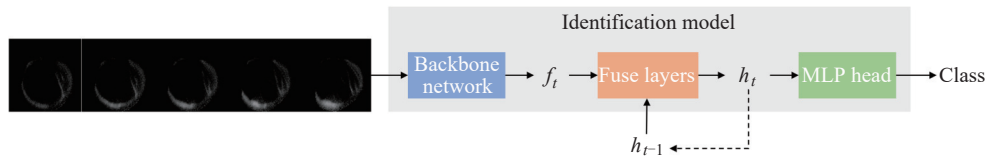


图 4 短时记忆策略流程

Fig. 4 Flowchart of short memory strategy

出将会被输入到一个多层感知输出器中获得当前图像序列所对应的类别. 在该策略中, 主干网络和融合层的权重在所有时刻都是共享的. 与通道填充策略类似, 短时记忆策略中的主干网络也可以由任意结构的神经网络组成.

2.2 基于 CBAM 注意力的 EfficientNet

综上设计了两种针对变长图像序列的识别策略, 两种策略中涉及的识别模型在结构上虽然有所不同, 但都包含一个用于提取空间特征的主干网络. 以 EfficientNet^[16] 系列网络和 CBAM 注意力机制^[17,18] 为参考, 提出一种基于 CBAM 注意力的 EfficientNet (CBAM-EfficientNet) 主干网络结构, 用于构建两种策略下的识别模型.

CBAM-EfficientNet 中的主要模块 CBAM-MB 的结构如图 5 所示. 输入的特征首先经过一个由卷积层 (Conv2d)、批归一化层 (BN) 和 Swish 激活函数^[19] 组成的模块, 再经过一个由深度可分离卷积^[20] (DConv2d)、批归一化层和 Swish 激活函数组成的模块, 这两个模块会对输入进行初步的特征提取. 接着将经过 CBAM 中的通道注意力模块 (Channel Attention Module, CAM). 该模块将首先对输入的特征进行全局平均池化 (Global Average Pooling) 和全局最大池化 (Global Max Pooling), 二者的结果将分别输入到两个共享的卷积层中, 再将卷积后的两个结果相加, 最后通过 Sigmoid 激活函数获得输入特征每一个通道的权重, 即为 CAM 的输出. 该输出将会与深度可分离卷积的结果相乘, 之后输入到空间注意力模块 (Spatial Attention Module, SAM) 中. SAM 会对该输入的特征在每个通道上取最大值和平均值, 再将结果进行拼接 (Concat), 然后经过一个卷积层调整通道数, 最后再通过 Sigmoid 激活函数获得每个输入特征在每个空间特征点上的权重, 即为 SAM 的输出. 将该输出与输入特征相乘后就能得到 CBAM 重标定后的特征. 最终经由一个卷积层处理, 再与最初的输

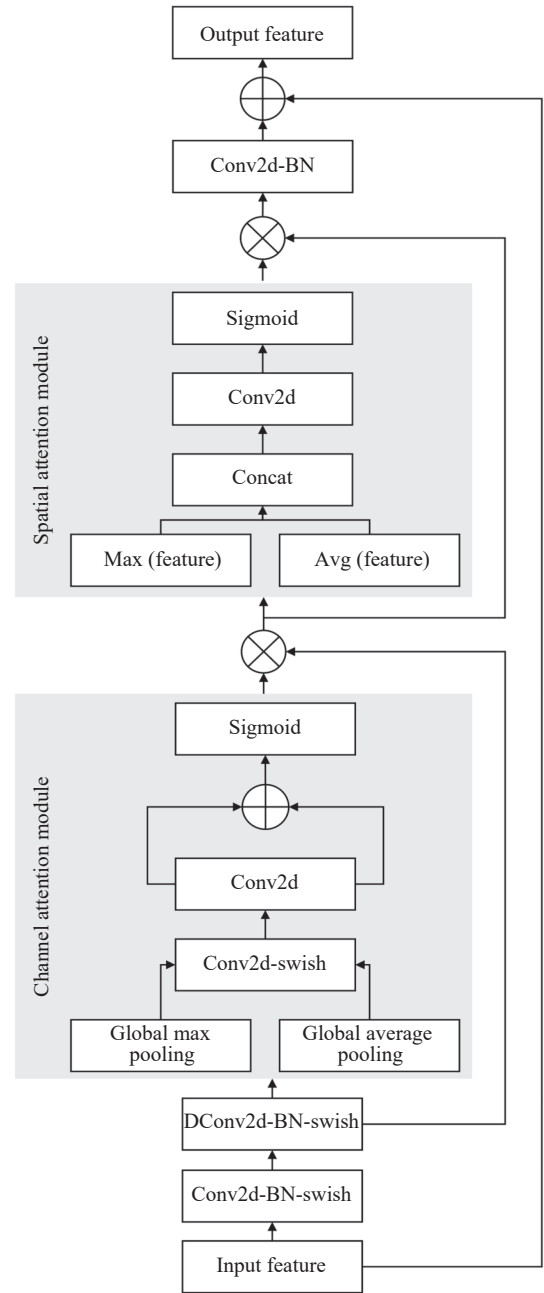


图 5 CBAM-MB 模块结构

Fig. 5 Structure of CBAM-MB module

入特征相加就可以得到 CBAM-MB 模块的输出特征.

CBAM-MB 模块是 CBAM-EfficientNet 的核心组成部分, 能够自适应地调整通道和空间信息的权重, 提高网络对重要特征的敏感度. 实现的 CBAM-EfficientNet 共包含 11 个阶段, 如表 2 所示.

以 CBAM-EfficientNet 为主干网络的识别模型结构如图 6 所示. 图 6(a) 为通道填充策略下以 CBAM-EfficientNet 为主干网络的模型 CBAM-EfficientNet-CFS, 图 6(b) 为短时记忆策略下以 CBAM-EfficientNet 为主干网络的模型 CBAM-EfficientNet-SMS. 在两种策略的模型中, 输入都将首先经过 CBAM-EfficientNet 来提取时空特征. 在通道填充策略下, 该特征将会输入到一个由全连接层和 Soft-

max 激活函数组成的多层感知输出器中, 从而输出序列所属类别. 在短时记忆策略下, 主干网络提取到的当前时刻的空间特征图将在融合层中与上一时刻的时空特征图在通道维度上拼接, 再经过一个由卷积层、Swish 激活函数和批归一化层构成的模块进行融合, 得到当前时刻的时空特征图, 并将参与下一时刻的空间特征图融合. 当处理到最后一帧图像时, 融合层的输出将被输入到多层感知输出器中, 从而获得当前序列所属的类别.

在完成模型构建后, 将使用两个训练集对模型分别进行训练. 训练过程采用交叉熵函数作为损失函数, 并选用 Adam^[21] 优化器进行参数优化, 初始学习

表 2 CBAM-EfficientNet 组成架构
Table 2 Architecture of CBAM-EfficientNet

阶段	模块	卷积核大小	通道数	步长	模块数量
1	Conv2d-BN-Swish	3×3	32	2	1
2	CBAM	7×7	32	1	1
3	CBAM-MB	3×3	16	1	1
4	CBAM-MB	3×3	24	2	2
5	CBAM-MB	5×5	40	2	2
6	CBAM-MB	3×3	80	2	3
7	CBAM-MB	5×5	112	1	4
8	CBAM-MB	5×5	192	2	5
9	CBAM-MB	3×3	320	1	1
10	Conv2d-BN-Swish	1×1	1280	1	1
11	Global Average Pooling	—	—	—	1

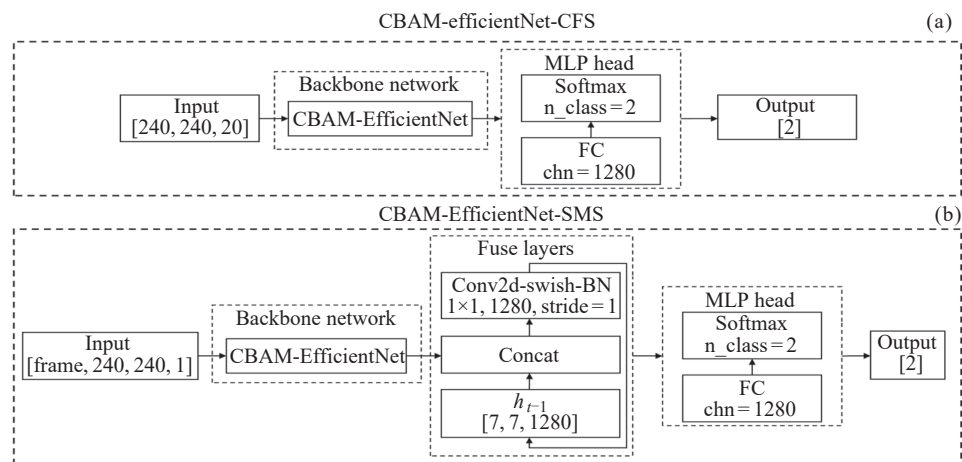


图 6 CBAM-EfficientNet 为主干网络的识别模型结构

Fig. 6 Structure of the models with CBAM-EfficientNet as the backbone network

率设定为 0.0001. 最大训练轮次设置为 200, 同时监控模型每一轮在验证集上的表现, 并采取早停法根据验证集上的损失函数值提前中止训练. 最优模型参数将取模型在训练损失收敛后在验证集上损失最小的轮次所对应的参数, 并用于最终的模型测试过程. 模型训练与测试过程均使用了国家空间科学中心公共技术服务中心空间科学数据融合计算平台提供的计算服务.

3 实验结果与分析

3.1 评估指标

由于极光亚暴爆发时点识别属于二分类问题, 因此本章将使用分类问题常用的几个指标来评估模型识别的准确性, 包括精确率 (Precision, P)、召回率 (Recall, R) 和 F_1 . 根据真实情况和预测结果的组合, 可以将测试样本划分为真正例 (True Positive, TP)、假正例 (False Positive, FP)、假反例 (False Negative, FN)、真反例 (True Negative, TN) 四种情形, 令 N_{TP} , N_{FP} , N_{FN} , N_{TN} 分别表示其对应的样本数. 可以得到三种指标的计算公式如下:

$$P = \frac{N_{TP}}{N_{TP} + N_{FP}}, \quad (1)$$

$$R = \frac{N_{TP}}{N_{TP} + N_{FN}}, \quad (2)$$

$$F_1 = \frac{2PR}{P + R}. \quad (3)$$

其中, P 为指被识别为暴时图像序列中正确的比例, 反映了模型不误判的准确性. R 是指所有真实为暴时图像序列中被正确识别的比例, 反映了模型不漏判的敏感性. F_1 为基于 P 和 R 的调和平均值, 是模型综

合性能反映.

使用每秒处理帧数 (Frames Per Second, FPS) 对比评估提出的模型和其他识别方法的识别效率.

3.2 模型识别准确性评估

为了对比分析提出的网络结构的优劣势, 还使用了 InceptionV3^[22] 和 Vision Transformer^[23,24] (ViT) 作为主干网络构建了模型. 对于每种策略, 都使用三种网络构建了对应的模型, 并使用两组数据集进行了训练和测试. 测试结果如表 3 所示. 表中 InceptionV3-CFS 指以 InceptionV3 为主干网络的通道填充策略模型, InceptionV3-SMS 指以 InceptionV3 为主干网络的短时记忆策略模型, 其他模型缩写以此类推. 测试集 1 的结果是使用训练集 1 训练得到的模型在测试集 1 上测试得到, 测试集 2 的结果是使用训练集 2 训练得到的模型在测试集 2 上测试得到. 每种评估指标最高的结果均以粗体显示.

由表 3 可见, 所有六种模型在各项评估指标上都有优异的表现, 指标值皆高于 0.82, 这表明根据提出的两种策略构建的模型对极光亚暴的爆发时点具有足够的识别能力. 通道填充策略下的模型在 Precision 指标上普遍高于短时记忆策略下的模型, 而在 Recall 指标上二者的表现差别甚微. 在六种模型中, CBAM-EfficientNet-CFS 在 Precision 指标上达到了最高, 而 CBAM-EfficientNet-SMS 则是在 Recall 指标上表现最佳. 在 F_1 指标上, ViT-CFS 与 CBAM-EfficientNet-CFS 两种模型表现较优, 且二者差异甚微. 在同样的模型构建策略下, 以 CBAM-EfficientNet 为主干网络的模型在多个指标上达到了最优. 以上结果说明, 提出的 CBAM-EfficientNet 网络结构能够有效提高模型的识别性能.

表 3 以不同网络结构为主干网络的模型在测试集上的识别准确性评估结果

Table 3 Identification accuracy evaluation results of the models with different networks as backbone

模型	测试集1			测试集2		
	P	R	F_1	P	R	F_1
InceptionV3-CFS	0.98	0.90	0.94	0.96	0.87	0.91
ViT-CFS	0.98	0.93	0.96	0.93	0.88	0.90
CBAM-EfficientNet-CFS	0.99	0.91	0.95	0.98	0.87	0.92
InceptionV3-SMS	0.91	0.87	0.89	0.82	0.88	0.85
ViT-SMS	0.93	0.90	0.91	0.86	0.88	0.87
CBAM-EfficientNet-SMS	0.94	0.94	0.94	0.84	0.89	0.86

对比测试集 1 与测试集 2 的评估结果, 可以观察到测试集 2 上的各项指标普遍低于测试集 1. 这一现象的主要原因是测试集 2 中负样本的数量显著高于测试集 1. 尽管第二组数据的负样本数量是第一组的 20 倍, 但 CBAM-EfficientNet-CFS 模型的精确率 P 仅下降了 0.01, 召回率 R 下降了 0.04, F_1 下降了 0.03. 在所有 6 种模型中, CBAM-EfficientNet-CFS 的指标下降幅度相对较小, 同时维持了较高的指标水平. 这一结果充分证明了该模型在极光亚暴爆发时点识别任务上的稳健性, 同时凸显了其在真实观测情况中应对样本极端不均衡问题的适应性.

值得注意的是, 表 3 中的 R 普遍低于 P , 尤其在测试集 1 中. 构建的两个数据集均存在严重的正负样本不均衡的现象. 在这种数据集中, 由于少数类的数量远少于多数类, 因此模型可能会倾向于将更多的样本判别为多数类, 并将更少的样本判别为少数类. 在此情况下, 假反例 (FN) 的数量会增多, 真正例 (TP) 与假正例 (FP) 的总量会减少, 根据式 (1) 与 (2) 可知, R 会因此而降低, P 会因此而增大. 而测试集 1 的样本构造使得正负样本间的特征存在明显差异, 导致模型更容易准确判别正样本, 从而增加了真正例 (TP) 的数量, 导致了 P 的进一步增大.

图 7 给出 CBAM-EfficientNet-CFS 对不同长度序列识别的 F_1 指标结果. 图 7 中的序列长度与图 2 并非完全一致, 原因在于有些长度的序列均为负样本, 因此无法计算 F_1 结果. 在测试集 1 上, 所有长度上的 F_1 值全部高于 0.8. 当长度为 2 frame 和

3 frame 时, F_1 值相对较低. 这可能是由于在长度较短时, 模型能够获取的时序信息过少, 无法做出准确的判断. 又或因填充的通道数量多于原本的帧数, 导致序列真实的时空特征被填充的特征所覆盖, 模型无法准确提取此类序列的特征. 在测试集 2 上, 当长度为 2, 3, 9, 13, 16 frame 时, F_1 值均未达到 0.8. 这是因为测试集 2 内包含了较多的负样本, 模型将一些负样本误判成了正样本, 导致 Precision 偏低, 进而导致 F_1 值偏低. 但整体来看, 模型在训练集中样本数量偏多的 4 帧和 17 frame 上, F_1 值都达到了 0.93 以上, 证明模型通过训练仍然学习到了暴时图像序列的特征, 能够做出较为准确的判断. 在其他样本数量偏少的长度上, 虽然效果差些, 但仍能对多数样本做出准确的识别. 因此, 提出的 CBAM-EfficientNet-CFS 模型在使用变长图像序列对极光亚暴爆发时点进行识别的问题上有较好的适应性.

图 8 给出了 CBAM-EfficientNet-SMS 对不同长度序列识别的 F_1 指标结果. 在一些长度上, 该模型的 F_1 值为 0, 尤其在测试集 2 上. 在训练集中, 这些长度的样本数量普遍偏少, 导致模型无法对这些长度的序列时空特征进行充分学习. 类似于 CBAM-EfficientNet-CFS, CBAM-EfficientNet-SMS 对于数据集中样本数量较多的 4 frame 和 17 frame 能够做出相对准确的判断, 但在其他长度上的效果都明显低于 CBAM-EfficientNet-CFS. 以上结果说明, 短时记忆策略下的模型会受到图像序列长度的显著影响, 在识别鲁棒性上差于通道填充策略下的模型.

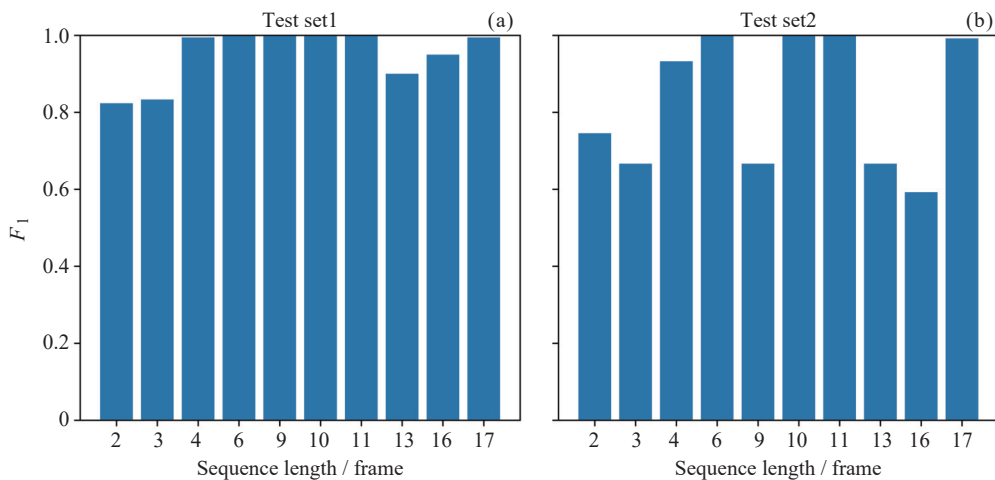


图 7 CBAM-EfficientNet-CFS 在不同长度序列上的 F_1 指标结果

Fig. 7 F_1 scores of CBAM-EfficientNet-CFS on the sequences with different length

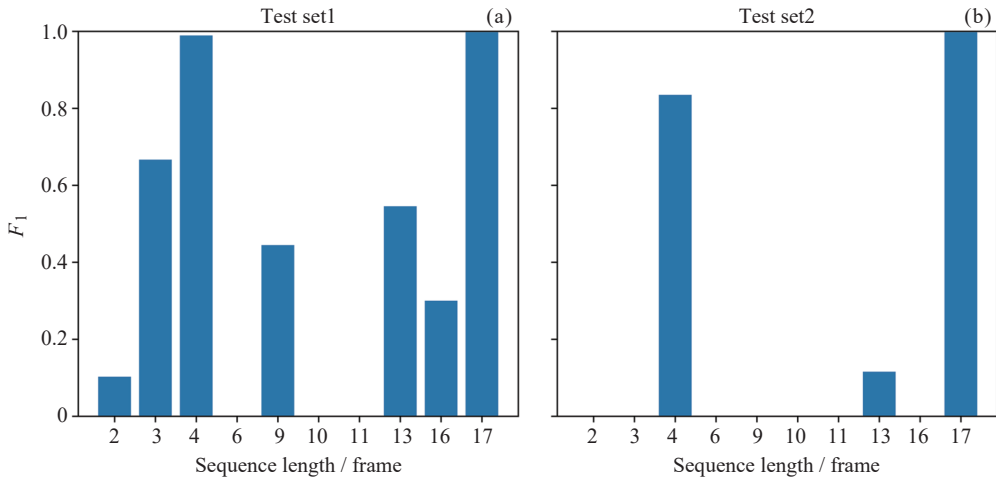


图 8 CBAM-EfficientNet-SMS 在不同长度序列上的 F_1 指标结果

Fig. 8 F_1 scores of CBAM-EfficientNet-SMS on the sequences with different length

3.3 与其他识别方法的对比

表 4 列出了提出的 CBAM-EfficientNet-CFS 模型与已有方法在极光亚暴识别问题上的性能对比结果. 除了对比 P 指标外, 这里还对使用每秒处理帧数来评估识别效率.

相比现有方法, CBAM-EfficientNet-CFS 在识别精确率方面表现最优, 这一结果凸显了提出的方法在极光亚暴爆发时点识别准确性方面的显著优势. 在识别效率方面, 尽管 CBAM-EfficientNet-CFS 的 FPS 远低于 SODN, 但值得注意的是, SODN 的精确率比 CBAM-EfficientNet-CFS 低了 0.1. 与 Han 等^[10]的方法相比, CBAM-EfficientNet-CFS 的 FPS 仅降低了 $1.2 \text{ frame}\cdot\text{s}^{-1}$, 这是因为 CBAM-EfficientNet-CFS 在 EfficientNet 的基础上增加了 CBAM 注意力, 因此运算速度会稍微变慢. 与识别效率最低的 SC-SLD 方法相比, CBAM-EfficientNet-CFS 在 FPS 上提升了近 5 倍. 以上结果表明, 尽管提出的方法在识别效率上的表现适中, 但是在识别精确率上达到了最佳水平.

3.4 事件分析

图 9 给出了两个被 CBAM-EfficientNetB0-CFS 模型正确识别爆发时点的极光亚暴示例, 其中使用红色框框出来的图像为真实爆发时点之后的第一张 LBHL 波段图像, 在其之后的图像构成一个暴时序列, 在其之前的图像组成了一个非暴时序列. 在图 9(a) 所给出的示例中, 爆发时点前几乎没有极光出现, 而在爆发时点后极光强度突然增强, 并且极光发生了明

表 4 与其他识别方法的对比结果

Table 4 Comparative results with other identification methods

方法	P	FPS/(frame·s ⁻¹)
SCSLD ^[9]	0.89	6.93
SODN ^[11]	0.88	393
Han's ^[10]	0.93	38.13
CBAM-EfficientNet-CFS	0.98	36.93

显的膨胀. 对于该示例, 模型能够将爆发时点前的序列判别为非暴时序列, 而将爆发时点后的序列判别为暴时序列, 证明了模型对于爆发时点前后时空特征差别较大的序列能够做出准确的区分. 图 9(b) 中的爆发时点与前一个极光亚暴事件的爆发时点相差约 1 h, 爆发时点前的几张图像中极光强度也有所增强, 极光也有所膨胀, 说明前一个极光亚暴事件可能仍处于膨胀期, 但模型依旧能够区分这两种时空特征略有相似的序列, 证明模型对极光亚暴叠加发生情况下的爆发时点同样具有一定的识别能力.

图 10 给出了两个漏检的示例, 其中使用红色框框出来的图像为真实的爆发时点之后的第一张 LBHL 波段图像, 在其之后的图像构成一个暴时序列, 在其之前的图像组成了一个非暴时序列. 对于这两个示例中的暴时图像序列, 模型将其识别为非暴时图像序列. 在这两个暴时序列中, 虽然每个时刻的图像相比前一时刻都有强度增强, 但极光总强度极低, 并且

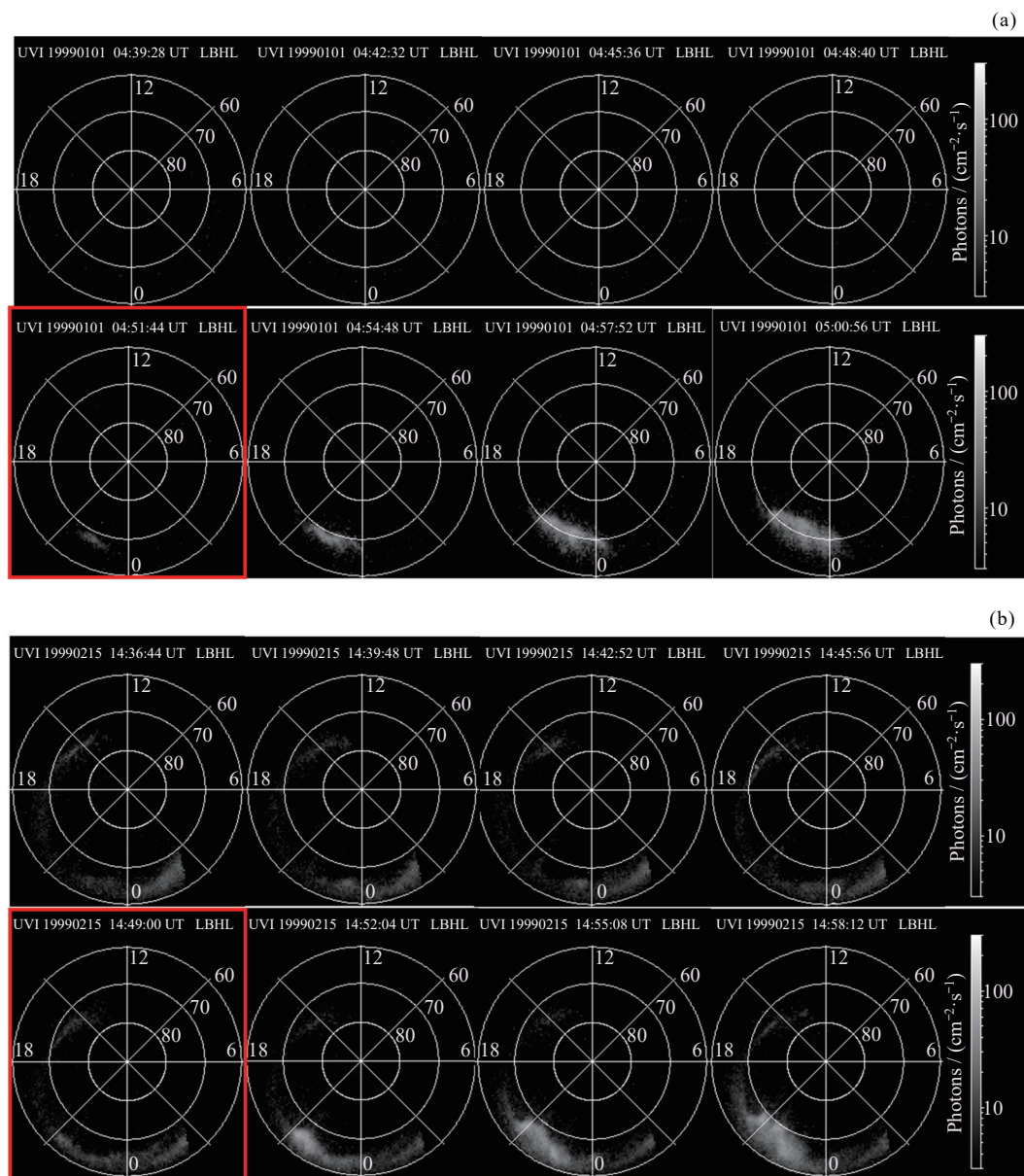


图9 极光亚暴爆发时点正确识别示例

Fig. 9 Examples of correctly identified auroral substorm onsets

极光整体的强度和形态变化不大,甚至在爆发时点之后的5~10 min强度有所下降,且极光有所收缩.由于其极光强度增强的时间不足10 min,不符合人工识别的标准,因此这样的极光亚暴事件有可能是通过其他波段的极光图像观察获得的,也有可能是人工标注过程中产生的误判事件.由于多数暴时序列都满足人工识别的标准,因此对于这类不满足人工识别标准的暴时序列,模型无法做出正确的判断.

图11给出了两个误检的示例,两个示例中所有

的图像均为测试集2中不在极光亚暴爆发时点附近的非暴时序列中的图像,其中使用红色框框出来的图片为模型多检的爆发时点对应的图像,其后的图像构成一个真实非暴时序列,但被模型错误地识别成了暴时序列.在图11(a)的示例中,图像中的极光从无到有,并且在模型识别出来的爆发时点之后的10 min内,极光强度逐渐增强,符合人工识别的标准,但该爆发时点不包含在人工识别的极光亚暴列表中,有可能是人工识别时遗漏的事件.在图11(b)示例中,

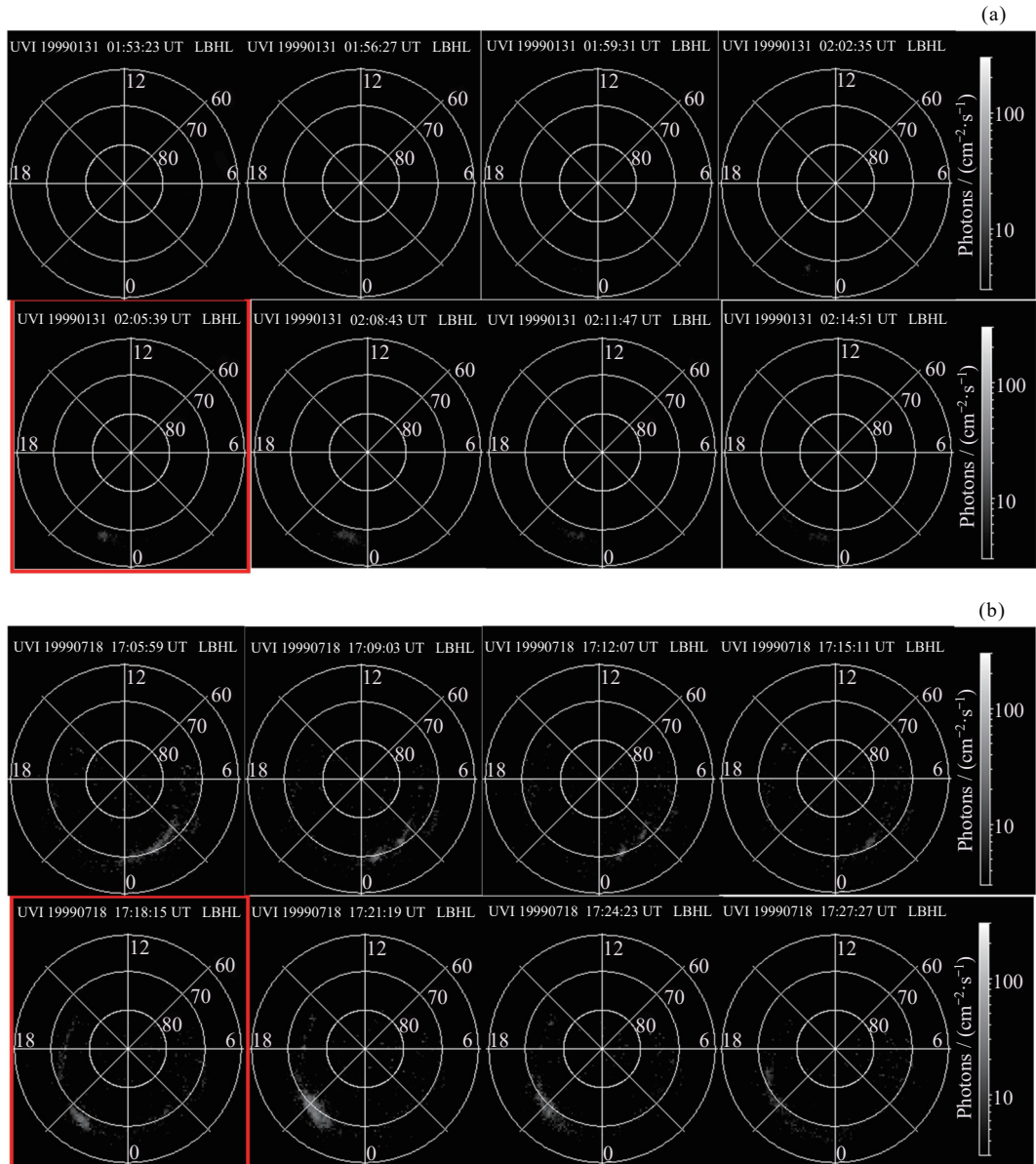


图 10 极光亚暴爆发时点漏检示例

Fig. 10 Examples of missing auroral substorm onsets

极光同样从无到有, 在模型识别出来的爆发时点之后的 6 min 内, 极光强度有明显的增强, 但在之后的 3 min 内出现了减弱. 该非暴时序列与真实的暴时图像序列非常相似, 因此模型出现了误检.

4 结论

为了实现与人工识别标准完全一致的极光亚暴爆发时点的机器识别模型, 设计了两种策略, 旨在解决在复现人工识别标准时遇到的图像序列长度不定

长问题. 两种策略分别通过填充图像通道维度和记忆前序图像特征的方式来实现对变长图像序列的处理. 基于这两种策略, 进一步开发了一种基于 CBAM 注意力的 EfficientNet, 作为识别模型中的主干网络. 通过与以其他网络作为主干网络的模型对比, 提出的 CBAM-EfficientNet 能够有效提升两种策略下模型对于极光亚暴爆发时点的识别准确性. 两种模型构建策略均能够实现对变长图像序列的准确识别, 尤其在样本数量较多的序列长度上 F_1 指标可达 0.99. 相较于其他的极光亚暴识别方法, 研究提出的方法在识别

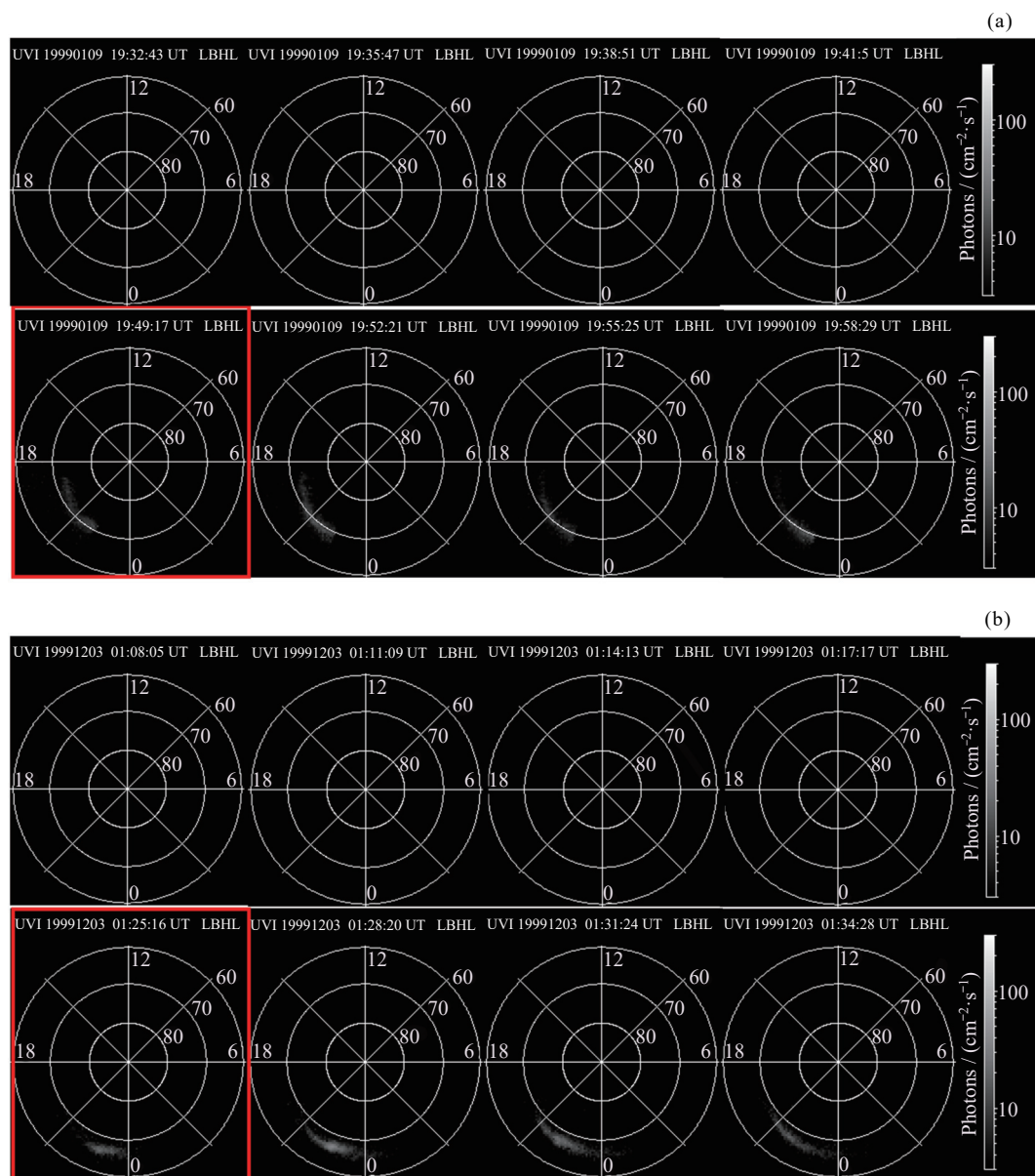


图 11 极光亚暴爆发时点误检示例

Fig. 11 Examples of false identified auroral substorm onsets

精确率上达到了最优,而在识别效率方面处于适中的水平.

研究提出的极光亚暴爆发时点机器识别模型是一种端到端的单阶段模型,无需复杂的数据预处理步骤,仅需要输入一段原始的极光图像序列,即可判断其是否是极光亚暴爆发时点之后的图像序列.此方法不仅成功打破了现有方法对繁琐的图像预处理步骤的依赖,还解决了多数方法在实际操作中需要人工调参和判别的问题.此外,本文的模型还能够很好地应对真实观测中暴时序列样本与非暴时序列样本数量

极端不均衡的情况,在实际应用场景下具有较高的实用性.

研究提供了新的极光亚暴爆发时点的机器识别方法,未来的研究工作将从以下几个方面进行优化和拓展.首先,使用的数据集是根据人工识别的极光亚暴爆发时点建立的.但是,人工识别的结果存在一定的主观偏差,这可能使得据此构建的数据集不够准确.为了提高数据的可靠性,后续研究将结合由 *AL* 指数和 *SML* 指数检测得到的亚暴事件数据,构建更全面且客观的极光亚暴爆发时点数据集.其次,提

出的方法在极短的序列和训练样本数量较少的长度上表现稍差. 后续工作将尝试借助一些处理长尾分布数据的方法^[25,26]来解决模型在不同序列长度上识别准确性差异较大的问题. 另外, 研究提出的方法仍会出现漏检和误检的问题, 其原因一方面是只使用了一个波段的极光图像导致先验信息缺失, 另一方面可能是模型本身对于暴时图像序列的时空特征的提取还不充分. 未来的研究将致力于整合多波段图像数据, 以构建更为全面的数据集, 并采用更先进的神经网络来优化模型. 此外, 还计划将模型应用于其他卫星 (例如 SMILE 卫星^[27]) 的紫外观测数据, 以评估模型的泛化能力.

致谢 数据资源和计算服务由国家科技资源共享服务平台-国家空间科学数据中心 (<https://www.nssdc.ac.cn>) 提供. 美国宇航局空间物理数据设施 (NASA's Space Physics Data Facility, SPDF) 和 SuperMAG 提供了数据支持. Polar/UVI 团队 (PI: G. Parks) 提供了相关数据处理工具.

数据可用性声明 本篇论文的关联数据 (DOI:10.57760/sciencedb.space.01829) 可在空间科学领域社区 (<https://doi.org/10.57760/sciencedb.space.01829>) 中访问获取.

参考文献

- [1] TU Chuangzhi, ZONG Qingang, ZHOU Xuzhi. Solar-Terrestrial Physics: Volume 2 - Magnetospheric Physics[M]. 2nd edit. Beijing: Science Press, 2020 (涂传诒, 宗秋刚, 周煦之. 日地空间物理学: 下册·磁层物理 [M]. 2 版. 北京: 科学出版社, 2020)
- [2] AKASOFU S I. Dynamic morphology of auroras[J]. *Space Science Reviews*, 1965, **4**(4): 498-540
- [3] JUUSOLA L, ØSTGAARD N, TANSKANEN E, *et al.* Earthward plasma sheet flows during substorm phases[J]. *Journal of Geophysical Research: Space Physics*, 2011, **116**(A10): A10228
- [4] PARTAMIES N, JUUSOLA L, TANSKANEN E, *et al.* Statistical properties of substorms during different storm and solar cycle phases[J]. *Annales Geophysicae*, 2013, **31**(2): 349-358
- [5] OHTANI S, GJERLOEV J W. Is the substorm current wedge an ensemble of wedgelets: revisit to midlatitude positive bays[J]. *Journal of Geophysical Research: Space Physics*, 2020, **125**(9): e2020JA027902
- [6] FREY H U, MENDE S B, ANGELOPOULOS V, *et al.* Substorm onset observations by IMAGE-FUV[J]. *Journal of Geophysical Research: Space Physics*, 2004, **109**(A10): A10304
- [7] LIOU K. Polar ultraviolet imager observation of auroral breakup[J]. *Journal of Geophysical Research: Space Physics*, 2010, **115**(A12): A12219
- [8] YANG Qiuju, LIANG Jimin, LIU Junming, *et al.* A method for automatic identification of substorm expansion phase onset from UVI images[J]. *Chinese Journal of Geophysics*, 2013, **56**(5): 1435-1447 (杨秋菊, 梁继民, 刘俊明, 等. 一种基于紫外极光图像的亚暴膨胀期起始时刻的自动检测方法 [J]. 地球物理学报, 2013, **56**(5): 1435-1447)
- [9] YANG X, GAO X B, TAO D C, *et al.* Shape-constrained sparse and low-rank decomposition for auroral substorm detection[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2016, **27**(1): 32-46
- [10] HAN Yiyuan, HAN Bing, GAO Xinbo. Auroral Substorm event recognition method combining eye movement Information and sequence fingerprint[J]. *Journal of Frontiers of Computer Science and Technology*, 2023, **17**(1): 187-197 (韩怡园, 韩冰, 高新波. 结合眼动信息和序列指纹的亚暴事件识别方法 [J]. 计算机科学与探索, 2023, **17**(1): 187-197)
- [11] YANG Qiuju, REN Jie, XIANG Han. Spatio-temporal detection of auroral substorm based on deep learning[J]. *Chinese Journal of Geophysics*, 2022, **65**(3): 898-907 (杨秋菊, 任杰, 向晗. 基于深度学习的极光亚暴时-空自动检测 [J]. 地球物理学报, 2022, **65**(3): 898-907)
- [12] TORR M R, TORR D G, ZUKIC M, *et al.* A far ultraviolet imager for the international solar-terrestrial physics mission[J]. *Space Science Reviews*, 1995, **71**(1): 329-383
- [13] LUMMERZHEIM D, LILENSTEN J. Electron transport and energy degradation in the ionosphere: evaluation of the numerical solution, comparison with laboratory experiments and auroral observations[J]. *Annales Geophysicae*, 1994, **12**(10): 1039-1051
- [14] RICHMOND A D. Ionospheric electrodynamics using magnetic apex coordinates[J]. *Journal of Geomagnetism and Geoelectricity*, 1995, **47**(2): 191-212
- [15] JIANG J N, ZOU Z M, LU Y. A ConvLSTM-based prediction model of aurora evolution during the substorm expansion phase[J]. *Earth and Space Science*, 2023, **10**(4): e2022EA002721
- [16] TAN M X, LE Q V. EfficientNetV2: Smaller models and faster training[OL]. arXiv preprint arXiv: 2104.00298, 2021
- [17] FU Guodong, HUANG Jin, YANG Tao, *et al.* Improved lightweight attention model based on CBAM[J]. *Computer Engineering and Applications*, 2021, **57**(20): 150-156 (付国栋, 黄进, 杨涛, 等. 改进 CBAM 的轻量级注意力模型 [J]. 计算机工程与应用, 2021, **57**(20): 150-156)
- [18] WEI Ya'na, WANG Zhibin, QIAO Xiaojun, *et al.* Lightweight rice disease identification method based on attention mechanism and EfficientNet[J]. *Journal of Chinese Agricultural Mechanization*, 2022, **43**(11): 172-181 (卫雅娜, 王

- 志彬, 乔晓军, 等. 基于注意力机制与 EfficientNet 的轻量化水稻病害识别方法 [J]. 中国农机化学报, 2022, **43**(11): 172-181)
- [19] RAMACHANDRAN P, ZOPH B, LE Q V. Searching for activation functions[OL]. arXiv preprint arXiv: 1710.05941, 2017
- [20] CHOLLET F. Xception: Deep learning with depthwise separable convolutions[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, HI, USA: IEEE, 2017: 1800-1807
- [21] KINGMA D P, BA J. Adam: A method for stochastic optimization[OL]. arXiv preprint arXiv: 1412.6980, 2017
- [22] SZEGEDY C, VANHOUCKE V, IOFFE S, *et al.* Rethinking the inception architecture for computer vision[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA: IEEE, 2016: 2818-2826
- [23] ARNAB A, DEHGHANI M, HEIGOLD G, *et al.* ViViT: A video vision transformer[C]//2021 IEEE/CVF International Conference on Computer Vision (ICCV). Montreal, QC, Canada: IEEE, 2021
- [24] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, *et al.* An image is worth 16x16 words: Transformers for image recognition at scale[OL]. arXiv: 2010.11929, 2021
- [25] CHEN X H, ZHOU Y C, WU D Y, *et al.* Imagine by reasoning: A reasoning-based implicit semantic data augmentation for long-tailed classification[C]//Thirty-Sixth AAAI Conference on Artificial Intelligence. Palo Alto, CA, USA: AAAI, 2022
- [26] ZHANG S Y, LI Z M, YAN S P, *et al.* Distribution alignment: A unified framework for long-tail visual recognition [C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville, TN, USA: IEEE, 2021
- [27] LIU Xiaohong, WANG Yongmei, WANG Tianfang. Design and realization of the ultra-violet imager communication and control system based on packet utilization standard[J]. *Chinese Journal of Space Science*, 2022, **42**(6): 1204-1209 (刘晓红, 王咏梅, 王天放. 基于包应用标准的紫外极光成像仪通信控制系统的设计与实现 [J]. *空间科学学报*, 2022, **42**(6): 1204-1209)

作者简介



蒋家楠 女, 1996年5月出生, 现为中国科学院国家空间科学中心在读博士生, 主要研究方向为机器学习在极光科学领域的应用。

E-mail: jiangjianan18@mailsucas.edu.cn



邹自明 男, 1971年5月出生, 现为中国科学院国家空间科学中心研究员, 博士生导师, 主要研究方向为日地空间信息系统技术、机器学习在空间物理的应用等。

E-mail: mzou@nssc.ac.cn