

中图法分类号: TP391.41 文献标识码: A 文章编号: 1006-8961(XXXX)XX-0001-19

论文引用格式: Zhang Benshuo, Ke Feng, An Zhiyong, Yu Xiaoning, Han Zhongwei, Zhao Feng. Himawari-8/9 satellite fire detection integrating multi-scale context and Mamba temporal modeling[J/OL]. Journal of Image and Graphics, XXXX: 1-19. DOI: 10.11834/jig.260158. (张本硕, 柯枫, 安志勇, 于晓宁, 韩忠伟, 赵峰. 融合多尺度上下文与 Mamba 时序建模的 Himawari-8/9 卫星火点检测[J/OL]. 中国图象图形学报, XXXX: 1-19. DOI: 10.11834/jig.260158. ) [DOI: 10.11834/jig.260158]

# 融合多尺度上下文与 Mamba 时序建模的 Himawari-8/9 卫星火点检测

张本硕<sup>1</sup>, 柯枫<sup>2</sup>, 安志勇<sup>1</sup>, 于晓宁<sup>1</sup>, 韩忠伟<sup>3</sup>, 赵峰<sup>1,3\*</sup>

1. 山东工商学院计算机科学与技术学院, 山东省烟台市 264005; 2. 烟台理工学院科研处, 山东省烟台市 264005; 3. 烟台理工学院信息工程学院, 山东省烟台市 264005

**摘要:** 目的 针对 Himawari-8/9 卫星影像中早期微弱火点难以准确识别的问题, 提出一种融合多尺度空间特征与长时序动态建模的火点检测方法, 以提升早期微弱火点识别能力和复杂场景下的检测稳定性。方法 构建 Fire-Mamba 火点检测模型。设计多尺度上下文热异常感知模块 (Multi-Scale Contextual Thermal Anomaly-aware Module, MCTAM), 通过双分支深度可分离卷积联合提取局部热异常信息与大尺度背景上下文信息, 增强微弱热异常的空间表征能力; 在此基础上, 引入 Mamba 时序建模模块, 对逐像素时序特征进行长程依赖建模, 以刻画火点从初始热异常到持续扩展的动态演化过程; 同时结合焦点损失函数缓解火点像素稀疏带来的类别不平衡问题, 并采用滑动时间窗口实现逐时刻火点预测。结果 在 6 起真实森林火灾事件上的实验结果表明, 所提方法在火点识别率、总体准确率和漏检率方面优于各对比方法; 但其误报率高于较为保守的 JAXA (Japan Aerospace Exploration Agency, JAXA) WLF L2 (Himawari Wildfire Product Level 2, WLF L2) 产品, 说明模型在误报抑制方面仍存在进一步优化空间。模型火点准确率 (FA) 达到 90.33%, 总体准确率 (OA) 为 99.60%, 漏检率 (OFR) 为 9.67%, 误报率 (FAR) 为 18.80%, 在复杂背景、弱小火点和火灾早期阶段均表现出较好的检测能力与时序稳定性。可视化结果进一步表明, 该方法能够获得边界更完整、连续性更强的火点响应, 对早期火情具有更高敏感性。结论 所提 Fire-Mamba 方法能够有效融合多尺度空间信息与长时序演化信息, 提升 Himawari-8/9 卫星早期火点检测的准确性, 可为近实时野火监测与早期预警提供技术支持。

**关键词:** Himawari-8/9; 火点识别; Mamba; 多尺度特征提取; 时序建模; 近实时监测; 卫星遥感; 火灾预警

## Himawari-8/9 satellite fire detection integrating multi-scale context and Mamba temporal modeling

Zhang Benshuo<sup>1</sup>, Ke Feng<sup>2</sup>, An Zhiyong<sup>1</sup>, Yu Xiaoning<sup>1</sup>, Han Zhongwei<sup>3</sup>, Zhao Feng<sup>1,3\*</sup>

1. School of Computer Science and Technology, Shandong Technology and Business University, Yantai, 264005, Shandong, China; 2. Office of Scientific Research, Yantai Institute of Technology, Yantai 264005, Shandong, China; 3. School of Information Engineering, Yantai Institute of Technology, Yantai 264005, Shandong, China

收稿日期: 2026-03-27; 修回日期: 2026-05-27

\* 通信作者: 赵峰 zhaofeng1016@126.com

基金项目: 国家自然科学基金项目 (62176140); 烟台市科技创新发展计划项目 (2023XDRH001); 山东省自然科学基金项目 (ZR2022QF151, ZR2023MF088); 烟台智慧城市创新实验室资助项目

Supported by: National Natural Science Foundation of China (62176140); Yantai City Technology Innovation Program (2023XDRH001); Natural Science Foundation of Shandong Province, China (ZR2022QF151, ZR2023MF088); Yantai Smart City Innovation Laboratory

**Abstract:** Wildfires occur frequently worldwide and pose severe threats to ecosystem stability, regional climate systems, and human life and property. In recent years, against the backdrop of global warming, adverse meteorological conditions such as extreme heat, prolonged drought, and anomalous wind fields have become more frequent, further increasing the risk of wildfires. In particular, under the combined influence of complex terrain, hot and dry winds, and abrupt meteorological changes, fires can rapidly evolve from weak initial thermal anomalies into large-scale disasters, causing forest resource loss, increased carbon emissions, ecological degradation, and severe socioeconomic impacts. Therefore, achieving rapid and accurate detection of early wildfire events over large spatial extents has become an important research objective in remote-sensing-based disaster monitoring and a key issue for improving early warning and emergency response capabilities. Compared with ground-based monitoring and airborne observations, satellite remote sensing provides broad spatial coverage, high observation efficiency, and continuous monitoring capability over large regions. Among available platforms, geostationary meteorological satellites are particularly valuable for near-real-time wildfire monitoring because they can repeatedly observe the same region at short temporal intervals. Compared with polar-orbiting satellites, Himawari-8/9 provides continuous observations with a temporal resolution of 10 min and 16 multispectral channels, making it an important data source for operational wildfire monitoring. Benefiting from this high-frequency temporal sampling capability, Himawari-8/9 can capture thermal anomaly responses at the moment of fire ignition. Furthermore, it continuously characterizes the dynamic evolution of fire pixels from emergence to intensification and expansion. This capability is especially important for early wildfire detection. Many early fire pixels are extremely small, exhibit weak thermal anomaly signals, and are not visually salient in a single image. Their identification often depends on variation trends and cumulative characteristics across continuous time series. Existing fire detection methods based on Himawari-8/9 data mainly include traditional threshold-based methods, spatial contextual approaches, and, more recently, deep-learning-based methods. Traditional multi-temporal threshold methods and contextual algorithms usually rely on manually designed decision rules derived from brightness temperature, band differences, and neighborhood background statistics. These methods are physically interpretable, easy to implement, and computationally efficient, and therefore perform well in some typical scenarios. However, they remain prone to false alarms and missed detections. This is especially true under complex surface backgrounds, cloud contamination, smoke interference, and extremely weak thermal anomalies in the early stage of wildfire development. For example, bare land, high-albedo surfaces, urban heat sources, cloud edges, and terrain shadows may produce radiometric responses similar to those of fire pixels, thereby weakening the discriminative capability of rule-based methods. In addition, fixed thresholds or semi-empirical rules often lack sufficient generalization across different regions, seasons, and land-cover types. With the rapid development of deep learning, convolutional neural networks, recurrent neural networks, and Transformer-based models have gradually been introduced into wildfire detection tasks. Compared with traditional methods, these approaches can automatically learn more discriminative spatial and temporal features in a data-driven manner, thereby reducing dependence on handcrafted rules and improving adaptability in complex scenes. Nevertheless, they still exhibit clear limitations for early wildfire detection based on long time-series Himawari-8/9 imagery. Convolutional neural networks mainly emphasize local spatial representations and remain limited in modeling long-term dynamic dependencies during wildfire evolution. Recurrent neural networks and their variants possess sequence modeling capability, but when handling long sequences at high temporal resolution, they are prone to gradient vanishing and long-term memory decay. This issue becomes more pronounced in scenarios involving 24 h of continuous observation and 144 temporal steps. Transformers can model global dependencies and long-range feature interactions. However, their computational complexity generally increases quadratically with sequence length. This imposes considerable efficiency pressure on near-real-time geostationary satellite monitoring. Therefore, early wildfire detection from long time-series Himawari-8/9 imagery still requires an efficient framework that can simultaneously capture weak spatial thermal anomalies and model long-range temporal evolution. To address these challenges, this study proposes Fire-Mamba, an early wildfire detection framework that integrates multi-scale spatial representation with long-sequence dynamic modeling. Fire-Mamba is designed for early wildfire targets in Himawari-8/9 imagery. These targets are typically extremely small, weak in thermal radiation variation, and highly susceptible to complex background interference. The framework aims to improve the accuracy and stability of near-real-time wildfire detection while maintaining manageable computational cost. Its design is motivated by two

observations. First, early wildfire pixels usually manifest as subtle local thermal anomalies and lack salient shape or texture cues. Second, genuine fire pixels often cannot be reliably distinguished from transient background disturbances using a single image. Instead, their identification depends on variation trends across continuous observations. Specifically, the B07, B14, and B03 bands are selected to construct multispectral temporal inputs, and day-night cloud masking is performed using multi-band threshold rules to reduce interference from clouds and complex backgrounds. According to the 10 min observation interval of Himawari-8/9, a continuous sequence of length 144 is constructed to represent a 24 h temporal window, thereby providing sufficient support for modeling wildfire dynamics. Compared with short-window strategies, such a long time-series input can reflect not only the instantaneous thermal anomaly state of candidate fire pixels, but also their continuous intensification, spatial expansion, and temporal persistence. In the spatial feature extraction stage, a Multi-Scale Contextual Thermal Anomaly-aware Module (MCTAM) module is designed to enhance the representation of weak fire-induced thermal anomalies under complex backgrounds. The MCTAM module employs a  $3 \times 3$  depthwise separable convolution to extract local thermal-gradient features induced by weak fire signals. This enables the capture of fine-grained local brightness-temperature variations and spatial discontinuities around candidate fire pixels. Meanwhile, a  $7 \times 7$  convolution branch is introduced to capture broader background contextual information. This branch suppresses pseudo-anomalous responses caused by non-fire heat sources, heterogeneous land-cover patterns, or complex textures. Global average pooling and a multilayer perceptron are further used to generate channel-wise weights for adaptive feature recalibration, enabling the model to enhance fire-related features while suppressing background interference and isolated noise. In the temporal modeling stage, the selective state-space model Mamba is introduced to perform long-range dynamic modeling on pixel-wise temporal features, thereby improving the recognition of weak early thermal anomalies and their evolutionary trends. Compared with conventional recurrent structures, Mamba maintains strong temporal modeling capability while processing long sequences more efficiently. This makes it highly suitable for high-temporal-resolution continuous observations from geostationary satellites. By modeling long-range dependencies in pixel-wise temporal features, Fire-Mamba better captures the persistence, intensification, and diffusion characteristics of fire signals over time. Consequently, it improves the discrimination between genuine wildfire signals and short-term background fluctuations. Finally, to address class imbalance caused by the sparsity of fire pixels, Focal Loss is adopted at the prediction stage to encourage the model to focus more on hard samples and improve fire-class detection performance. Experimental results demonstrate that Fire-Mamba achieves superior performance across multiple quantitative metrics. Its average fire detection rate (FA) reaches 90.33%, significantly outperforming the second-best CNN-LSTM model (82.28%) and the official JAXA WLF L2 product (40.66%). The model also achieves an overall accuracy (OA) of 99.60%, while reducing the average omission rate to 9.67%, indicating high sensitivity to small fire pixels. Compared with most deep-learning baselines, Fire-Mamba reduces isolated false alarms while maintaining higher fire detection sensitivity. However, its FAR remains higher than that of the conservative JAXA WLF L2 product, indicating that false-alarm suppression still requires further improvement. In summary, Fire-Mamba achieves a lightweight integration of multi-scale spatial thermal-gradient perception and long-range temporal evolution modeling, providing strong technical support for large-scale near-real-time wildfire monitoring and early warning.

**Key words:** Himawari-8/9; fire detection; Mamba; multi-scale feature extraction; time-series modeling; near real-time monitoring; satellite remote sensing; early fire warning

论文引用格式: DOI:10.11834/jig.260158

## 0 引言

由于自然因素或人为因素的影响,森林野火在全球范围内频繁发生,对农业生产、空气污染、气候变化以及人民生命财产安全造成了严重威胁(张国将等,2026;韩琛惠等,2023;Rogers等,2015;Seydi

等,2021;Hering等,2009)。随着时间推移,一旦火势不断增强,受复杂地形和突变气象条件影响,极易演变为难以控制的重特大火灾,导致灭火成本与风险呈几何级数增长,因此在大范围地表区域内快速、准确地探测火灾事件具有重要意义(Soro等,2021;Fusco等,2019)。

常用的野火探测方法包括人工巡逻监测(Kulakowski等,2013)、视频监控(Li等,2022;Guo

等,2022)、无人机监测(Yuan等,2015;Zhang等,2022)。人工巡检监控需要大量人力资源。同时,视频、无人机监控也存在监控范围小、成本高等问题。这些问题使得它们难以应用于大规模野火检测(Liu等,2023)。相比之下,卫星遥感因具备宏观、连续的大范围观测能力,长期以来一直是野火监测中的关键技术(Xiao等,2016;Xiao等,2019;Hong等,2022)。在多种卫星平台中,静止气象卫星因时间分辨率高、覆盖范围广且能够持续稳定监测大尺度区域,在野火探测中具有明显优势(Zhang L等,2025)。Himawari-8/9作为第三代静止气象卫星,主要服务于天气与灾害监测,其观测频率和空间分辨率均较以往有了明显提升(Chen等,2022)。特别是,该卫星可实现最高10分钟一次的高时间分辨率观测,覆盖东亚和东南亚等广大区域。与此同时,卫星搭载先进葵花成像仪(Advanced Himawari Imager, AHI)提供了16个波段的丰富光谱信息(Bessho等,2016;Da等,2015)。凭借高时间分辨率与多光谱成像能力,Himawari-8/9已成为支持野火事件高频率、精细化观测的重要数据来源(Chen等,2023;Xie等,2018;Zheng Z等,2024)。

野火发展过程受温度、湿度、风速风向及可燃物条件等多种物理因素共同影响,具有显著的复杂性。受此影响,野火在发展过程中通常表现出明显的空间异质性、时间动态性和光谱响应差异。围绕这些特征,国内外学者基于遥感数据开展了大量野火检测方法研究。早期研究主要包括两类方法:一类是多时相阈值方法,利用不同观测时刻之间的亮温差异来识别潜在火点,Filizzola等基于多时相卫星红外观测,通过检测当前像元相对于历史正常热辐射背景的异常偏离实现火点识别(Filizzola等,2016),Yan等利用Himawari-8 AHI热红外通道的时相变化特征构建了晴空条件下的火点检测方法,并对晨昏交界、冰雪下垫面和太阳耀斑等复杂场景进行了针对性处理(Yan等,2020)。后者则强调利用目标像元与周围背景像元之间的空间差异进行异常判别,Hally等评估了Himawari-8背景温度的上下文估计方法,表明背景温度估计结果对邻域有效背景像元的完整性较为敏感(Hally等,2018)。另一类是空间上下文信息的阈值方法。通过比较目标像元与邻域背景在中红外和热红外波段上的辐射差异来判定火点,其中Giglio等针对MODIS数据提出了增强型上

下文算法,通过利用目标像元与周围背景像元在热红外响应上的差异,实现了对活动火点的识别(Giglio等,2003)。随后,Maeda等针对Himawari-8 AHI图像在改进MOD14算法基础上引入上下文判别与机器学习策略,以提升早期火点检测能力(Maeda等,2023)。尽管这些方法在实际业务中得到广泛应用,但其性能仍高度依赖人工设定的阈值规则,在复杂地表背景或早期微弱火情条件下仍容易出现漏检或误报问题。

为了提高野火检测的准确性和鲁棒性,深度学习作为一种强大的特征提取工具被广泛引入遥感监测中(Barmoutis等,2020)。对于光学遥感图像中的小目标检测任务,目标尺度小、背景复杂和特征响应弱等因素会显著增加检测难度,深度学习已成为提升遥感小目标检测性能的重要技术路线(高广帅等,2025)。卷积神经网络(convolutional neural networks, CNN)能够通过卷积与池化操作从原始遥感影像中自动学习中高层空间特征,已被应用于火点识别与热异常检测等遥感任务(Krizhevsky等,2012)。Zhu等指出,CNN能够从原始图像中提取中高层抽象特征,并在遥感影像分析中得到广泛应用(Zhu等,2017);Hong等进一步基于Himawari-8卫星影像构建FireCNN,实现了主动火点检测,并利用多尺度卷积提取火点特征(Hong等,2022)。随后,一些研究引入循环神经网络(recurrent neural network, RNN)及其变体长短期记忆网络(long short-term memory, LSTM),通过建模火点数量、火势蔓延速率及环境因子的时间序列变化,刻画火灾发生与发展的动态过程。已有研究分别基于无人机红外影像和MODIS火点数据,验证了LSTM在森林火灾蔓延预测和野火热点预测中的有效性(Elman,1990;Hochreiter和Schmidhuber,1997;Li等,2021;Kadir等,2023)。近年来,Transformer架构由于具备全局依赖建模能力,也逐渐被应用于遥感时序分析任务(Yuan等,2022;Miao等,2023;Zhao和Ban,2025)。然而,针对Himawari-8/9卫星10分钟级高频多光谱观测数据的特性,现有基于该卫星的深度学习火点检测方法仍存在两点明显的局限性(Hong等,2022;Zheng Z等,2024;Zhang L等,2025):首先,现有模型普遍忽视了多尺度时空关系的充分整合,缺乏对局部热异常与宏观上下文背景的联合建模。由于早期野火通常仅表现为极微弱的单像素或亚像素级热异

常,这种空间提取层面的局限导致模型在面对复杂背景(如高反射裸地、云边缘干扰)时,对早期微弱火点的敏感性严重不足。其次,现有方法往往将空间特征提取与时序演化建模相割裂,时空结合不紧密。在处理 Himawari-8/9 长时序连续观测数据(24 小时连续 144 个时间步)时,基于 RNN 的结构面临严重的长时记忆衰减问题,而基于 Transformer 的方法其计算复杂度随序列长度呈二次方增长,难以在满足近实时火险监测计算效率需求的同时,实现长程动态演化的有效建模(Vaswani 等,2017;Gu 和 Dao,2023)。

为解决长序列建模的效率瓶颈,近年来提出的状态空间模型(State Space Model, SSM)及其代表架构 Mamba(Gu 和 Dao,2023)在保持线性计算复杂度的同时,能够有效建模长程时序依赖,在视觉与时序任务中表现出良好潜力(Zhu 等,2024;Liu 等,2024)。对于 Himawari-8/9 卫星 10 分钟级连续观测数据而言,Mamba 能够以较低计算开销捕获长时间序列中的火点演化规律,从而缓解 RNN 长时记忆衰减和 Transformer 计算复杂度较高的问题。目前,Mamba 架构已逐渐被引入遥感图像处理与地球观测任务中。Bao 等人系统梳理了 Vision Mamba 在遥感领域的应用进展,指出 SSM 具有线性计算复杂度和长程上下文建模能力,可缓解 CNN 感受野有限和 Transformer 计算复杂度较高的问题(Bao M 等,2026)。Li 等人进一步从遥感场景出发总结了 Mamba 的架构设计、扫描路径和混合建模范式,认为其更适用于长时间序列、大范围影像块以及高维遥感数据建模任务(Li Z 等,2026)。Zhang 等人将 Mamba 引入遥感影像时空融合任务,验证了 Mamba 在遥感时序信息传播和空间-时间特征融合中的应用潜力(Zhang Q 等,2026)。然而,早期火点通常具有尺度小、响应弱、易受复杂背景干扰等特点,仅依赖长程时序建模难以充分识别局部微弱热异常。因此,有必要将 Mamba 的长序列建模优势与多尺度空间特征提取机制相结合,以同时增强模型对火灾时序演化和早期微弱火点空间特征的感知能力。

针对上述问题,本文提出一种融合多尺度空间特征提取与长程时序建模的野火检测模型 Fire-Mamba。本文的主要贡献包括:(1)提出一种结合多尺度空间特征提取与长程时序建模的野火检测框架 Fire-Mamba;(2)设计 MCTAM 模块以增强微小热

异常的空间表征能力;(3)引入 Mamba 状态空间模型实现高效长时序建模,提高复杂背景下火点检测的稳定性。

## 1 研究区域和数据集

### 1.1 研究区域

为了覆盖更广泛的区域并验证所提方法的泛化能力,分析了中国西南及邻近南方地区的六起代表性森林火灾事件。中国西南地区以其高森林覆盖率和丰富的生态系统多样性著称,涵盖热带雨林到高山森林。季风气候导致旱季降雨稀少,进而提高野火发生风险(Tian 等,2012)。此外,研究区域复杂的地形也为高效的火灾响应和救援作业带来了挑战。关于六场火灾事件的详细信息如图 1 所示。

四川西昌(2020 年 3 月)、四川凉山(2020 年 5 月)、云南玉溪(2023 年 4 月)以及云南大理漾濞(2024 年 3 月)的四起火灾均发生在我国西南横断山区及云贵高原地带。该地形高差大火灾多爆发于干热季风期,导致火势极易形成树冠火并呈立体式蔓延,给早期探测和扑救带来了极大阻碍。重庆火灾(2022 年 8 月)和湖南新田火灾(2022 年 10 月)则属于极端高温干旱背景下的典型森林火灾事件。受罕见的持续性异常高温少雨气候影响,使得火情呈现出多点集中爆发且扑救周期极长的特点。这六起火灾事件涵盖了极端气象、复杂地形以及多变的火灾特征,为验证本模型在早期微弱火情识别与复杂演化场景下的检测性能提供了测试基础。

### 1.2 Himawari-8/9 卫星 AHI 数据

Himawari-8/9 卫星旨在以 10 分钟时间分辨率提供近实时地球观测数据。Himawari-8/9 数据具有高时间分辨率,已被广泛应用于气象监测、环境监测以及自然灾害探测等领域(Bessho 等,2016)。该卫星的观测范围覆盖东亚和澳大利亚地区,空间范围约为南纬 60°至北纬 60°、东经 80°至西经 160°。其搭载的先进 Himawari 成像仪(Advanced Himawari Imager, AHI)可获取 16 个光谱波段的数据,包括 3 个可见光波段、3 个近红外/短波红外波段和 10 个红外波段,为提高表格表达的规范性与一致性,将各波段中心波长及波段范围统一保留至小数点后两位。本文所采用的 Himawari-8/9 AHI 传感器各波段详细信息见表 1。

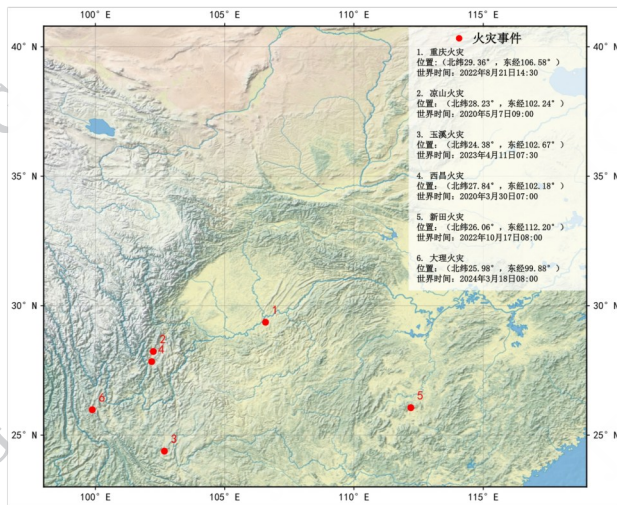


图1 研究区域概况

Fig. 1 Overview of the study area

本文使用 Himawari-8/9 AHI 数据作为主要数据来源。日本气象厅 (Japan Meteorological Agency, JMA) 提供 HSD 和 NetCDF4 两种标准数据格式。考虑到长时间连续观测序列对时序建模的重要性,本

文选用 10 min 分辨率的满盘 NetCDF4 数据作为模型训练与测试的基础数据。

鉴于不同光谱波段对地表热异常与复杂背景敏感度差异,热红外遥感信息提取已成为地表温度反演、材料识别和异常目标探测中的重要研究方向(曹丽琴等,2024)。因此,本文从 AHI 传感器的 16 个波段中,重点选取了对野火检测最具判别力的 3 个核心波段构建多光谱输入张量:第 7 波段(B07)、第 14 波段(B14)以及第 3 波段(B03)。为保证输入时间序列的连续性,针对卫星观测过程中可能出现的数据缺失问题,采用相邻时刻观测值均值进行插补处理。

为降低云层对火点检测结果的干扰本文采用基于阈值判别的云掩膜方法并将识别出的云像素在后续火点检测过程中予以剔除。考虑到白天与夜间云层物理特征存在差异,本文结合可见光、近红外反射率与红外亮温信息,分别构建昼夜分时段云判别规则。在白天观测条件下,云像素满足以下判定条件:

表1 Himawari-8/9 AHI传感器各波段详细信息

Table 1 Detailed information of each band of the Himawari-8/9 AHI sensor

波段编号	波段中心波长 ( $\mu\text{m}$ )	波段范围 ( $\mu\text{m}$ )	类型	空间分辨率 (km)
B01	0.47	0.43-0.48	可见光	1.00
B02	0.51	0.50-0.52	可见光	1.00
B03	0.64	0.63-0.66	可见光	0.50
B04	0.86	0.85-0.87	近红外	1.00
B05	1.61	1.60-1.62	近红外	2.00
B06	2.26	2.25-2.27	短波红外	2.00
B07	3.83	3.74-3.96	中波红外	2.00
B08	6.25	6.06-6.43	水汽通道	2.00
B09	6.96	6.89-7.01	水汽通道	2.00
B10	7.34	7.26-7.43	水汽通道	2.00
B11	8.59	8.44-8.76	红外	2.00
B12	9.63	9.54-9.72	红外	2.00
B13	10.41	10.30-10.60	红外	2.00
B14	11.21	11.10-11.30	红外	2.00
B15	12.36	12.20-12.50	红外	2.00
B16	13.31	13.20-13.40	红外	2.00

注:波段中心波长和空间分辨率依据 JMA 官方公布的 Himawari-8/9 AHI 成像仪规格整理,中心波长采用 Himawari-9 AHI 实测中心波长并统一保留至小数点后两位;波段范围依据 JMA AHI SRF 文件/相关文献整理,所有数值统一保留至小数点后两位。Himawari-8 与 Himawari-9 的实测中心波长存在轻微差异。

$$BT_{15} < 265 \text{ and } R_3 + R_4 > 1.2 \#(1)$$

$$R_3 + R_4 > 0.7 \text{ and } BT_{15} < 285K \#(2)$$

在夜间观测条件下,由于缺乏稳定的可见光与近红外反射率信息,云像素判定切换为基于中波红外与热红外亮温的纯红外阈值规则:

$$BT_{15} < 265K \text{ and } BT_{07} < 285K \#(3)$$

式中,  $R_{03}$  和  $R_{04}$  分别表示 Himawari-8/9 AHI 第 3 波段和第 4 波段的表观反射率,用于描述可见光和近红外波段上的反射特征;  $T_{07}$  和  $T_{15}$  分别表示第 7 波段和第 15 波段的亮温,单位为 K,用于表征中红外和热红外波段上的辐射温度响应。上述阈值规则的物理依据在于,云层相较于地表通常具有更高的短波反射率和更低的云顶亮温,因此可通过反射率和亮温联合约束将云像素从潜在火点候选区域中剔除。

### 1.3 实验数据集

**测试集:**为了真实地评估模型的泛化能力与检测精度,本文针对 1.1 节所述的中国西南地区 6 起典型森林火灾事件,构建了纯手工标注的真实测试集。在标注标准方面,首先以官方发布的火情通报中的起火时间与空间位置为依据确定感兴趣区域,随后结合 Himawari-8/9 AHI 影像的多光谱时序响应进行像元级判定。仅当候选像素在 B07 中波红外通道上表现出随时间连续增强的异常升温趋势,且 B07-B14 差值显著高于其空间邻域背景时,才将其判定为火点像素,否则标记为背景。在标注人员配置方面,本文邀请 3 名具有遥感影像解译经验的研究人员进行背靠背独立标注。完成初次标注后,对三份标注结果进行一致性检查;对于标注一致的区域,直接作为最终像素级标签;对于存在分歧的像元,则结合亮温时序曲线、B07-B14 差值变化及邻域上下文信息进行集中复核,并采用多数投票与人工复核相结合的方式确定最终标签。

**训练集:**由于真实卫星遥感火灾样本的精细标注成本较高,且长时间连续标注序列获取困难,采用生成仿真训练集,仿真训练集的构建主要包括 4 个步骤:(1)从历史无云观测数据中筛选未发生火灾的区域,裁剪得到连续  $T$  个时间步的背景时序张量,记为  $X_{bg} \in \mathbb{R}^{T \times H \times W \times C}$ ,  $T$  为时间窗口长度,  $H$ 、 $W$  分别为空间尺寸,  $C$  为输入波段数。该背景样本保留了真实地表在昼夜周期下的时序变化特征,可为后续仿真火点嵌入提供自然背景基底。(2)在背景张量的空

间范围内,随机选取若干空间位置  $(x_i, y_i)$  作为初始火点种子。(3)以火种像素为中心,在其  $3 \times 3$  的局部空间邻域内随机将其余 4 至 8 个相邻像素同步标记为火点,从而生成形态各异的微小火灾斑块。(4)为增强仿真样本在时间维上的异常表达能力,本文对每个火点区域随机赋予起火时刻  $t_{ignite} \in [1, T)$ 。当  $t < t_{ignite}$  时,像素保持真实背景观测信息;当  $t \geq t_{ignite}$  时,则在背景序列基础上叠加火点异常增量及随机扰动,从而形成由正常背景向异常热响应转变的时序特征。其表达式为:

$$F(t, x_i, y_i) = \begin{cases} F_{bg}(t, x_i, y_i), & t < t_{ignite} \\ F_{bg}(t, x_i, y_i) + \Delta F_{fire} + \varepsilon, & t \geq t_{ignite} \end{cases} \#(4)$$

式中,  $F_{bg}(t, x_i, y_i)$  表示时刻  $t$  下背景像素的多光谱观测值。  $\Delta F_{fire}$  表示异常增量项,其获取方式为:对独立于测试集的历史真实火点样本进行统计分析,提取火点像素与其邻域背景在各波段上的温度/反射率增量分布。在仿真时,从中波红外(B07)、热红外(B14)和可见光(B03)波段的差值经验分布中随机采样获取  $\Delta F_{fire}$  (例如, B07 波段的增量范围设定为 5 K ~ 25 K, B14 波段为 1 K ~ 5 K, B03 波段依据烟雾特性设定反射率增量)。  $\varepsilon \sim \mathcal{N}(0, \sigma^2)$  为随机扰动,用于模拟传感器噪声并提高模型鲁棒性。

## 2 方法

### 2.1 总体架构

为了充分利用高时间分辨率卫星数据中的空间与时间信息,本文构建了一种融合多尺度空间特征提取与长程时序建模的野火检测框架 Fire-Mamba。模型总体架构图如图 2 所示。

在输入阶段,将连续时间观测的多光谱卫星影像构建为时间序列数据。对于任意观测区域,在时间窗口长度为  $T$  的条件下,将连续  $T$  个时刻的多光谱影像按照时间顺序进行堆叠,形成输入张量。每一帧影像包含对火灾检测具有较高敏感性的关键波段信息,包括中波红外波段 B07(3.9  $\mu\text{m}$ )、热红外波段 B14(11.2  $\mu\text{m}$ )以及可见光波段 B03(0.64  $\mu\text{m}$ )。在空间特征提取阶段,输入序列中的每一帧影像首先进入多尺度上下文热异常感知模块(Multi-Scale Contextual Thermal Anomaly-aware Module,

MCTAM)。该模块通过双分支深度可分离卷积结构同时捕获局部热异常响应特征与大尺度背景上下文信息。局部分支主要用于识别微小火点产生的细粒度温度突变,而上下文分支则通过更大的感受野学习火点周围的背景辐射分布,从而为异常热源判定提供参考基准。经过 MCTAM 模块处理后,原始多光谱影像被映射为具有更强判别能力的空间特征表示。模型进一步通过 Mamba 状态空间模型进行长程时序建模。MCTAM 输出的空间特征按照像素位置提取对应的时间序列特征,并输入 Mamba 模块进行状态更新与序列建模,从而捕获火点从初始阴燃到快速扩散过程中的动态演化规律。模型通过线性分类层与 Sigmoid 激活函数将时序特征映射为当前时刻的火点概率图,从而得到研究区域中各像素发生野火的概率分布。在模型训练阶段,为缓解火点像素极度稀疏所带来的类别不平衡问题,本文采用焦点损失函数(Focal Loss)作为优化目标,对预测概率图与真实火点标签之间的误差进行监督学习。

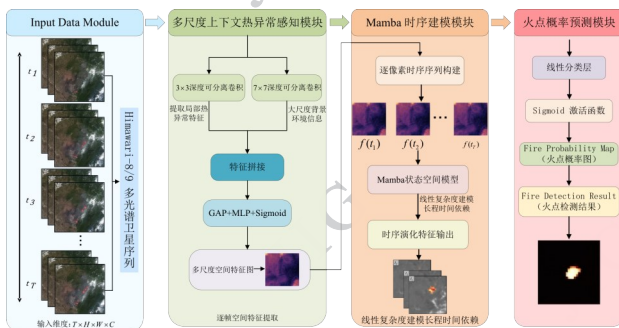


图2 模型总体架构图

Fig. 2 Overall architecture of the proposed Fire-Mamba model

## 2.2 多尺度上下文热异常感知模块

在野火发生初期,火点区域通常极小,甚至仅表现为单像素或亚像素级热异常,其在卫星图像上的主要判别依据并非稳定的形状结构,而是相对于周围背景所形成的局部热辐射突变。复杂地表环境下,裸地、水体、云边缘及高反射地物均可能产生与火点相似的辐射响应,仅依赖单一尺度的局部卷积特征难以准确区分真实火点与背景伪异常。为此,本文设计了多尺度上下文热异常感知模块(Multi-Scale Contextual Thermal Anomaly-aware Module, MCTAM),用于在较低计算开销下联合提取局部热异常信息与大尺度背景上下文信息,为后续时序演化建模提供更加紧凑且判别性更强的空间表征。

给定输入的多光谱张量  $X \in \mathbb{R}^{C \times H \times W}$  ( $H$  和  $W$  分别表示空间高度与宽度,  $C$  表示通道数), MCTAM 模块首先将其并行输入至两个具有不同感受野的卷积分支中。局部热异常感知分支采用  $3 \times 3$  的深度可分离卷积,凭借较小的感受野专注于提取像素级别的细粒度光谱异常。该分支主要用于捕获处于阴燃或刚起火阶段的微小火点所引发的局部剧烈升温,其特征映射过程可表示为:

$$F_l = \phi(\text{DSConv}_{3 \times 3}(X)) \# (5)$$

式中,  $\phi(\cdot)$  表示 Swish 激活函数,用于增强非线性表达能力并缓解深层网络中的梯度消失问题。由于 Himawari-8/9 卫星影像中的早期火点通常仅表现为单像素或 1~2 个像素范围内的弱热响应,  $3 \times 3$  卷积能够较好地描述中心像素与邻域像素之间的亮温突变和光谱差异,同时避免引入过多无关背景信息。与此同时,全局上下文分支采用  $7 \times 7$  的深度可分离卷积。较大的感受野旨在学习火点周围宏观的背景环境特征。通过获取广泛的上下文信息,该分支为模型判断中心像素的辐射值是否属于真实的异常热事件提供了对比基准,其计算过程表示为:  $F_g = \phi(\text{DSConv}_{7 \times 7}(X)) \# (6)$  公式(6)的物理意义是,通过更大范围的背景辐射信息构建环境参考基准,从而帮助模型从宏观尺度上区分真实火点与云边缘、高反射裸地等伪异常现象。相比  $3 \times 3$  卷积,  $7 \times 7$  卷积具有更大的感受野,能够获取火点周围更充分的背景辐射分布信息,为模型区分真实火点与云边缘、裸地、高反射地表及非火热源等伪异常提供上下文参考。本文选择  $3 \times 3$  与  $7 \times 7$  的双分支卷积结构,是基于早期火点“小尺度局部突变”和复杂背景“较大范围上下文判别”两类特征需求的综合考虑。同时,本文通过不同卷积核组合的对比实验进一步验证了该设置的有效性,结果表明  $3 \times 3 + 7 \times 7$  组合能够在检测精度和计算开销之间取得较好平衡。

在获取多尺度空间特征后,为了突出真实的火灾辐射信号并有效抑制高反射率裸地或水面等背景噪声, MCTAM 引入了一种轻量级通道动态重标定机制。具体而言,我们得到融合特征  $F_c \in \mathbb{R}^{H \times W \times 2C}$ 。并通过全局平均池化(global average pooling, GAP)和多层感知机(multilayer perceptron, MLP)聚合通道级信息,从而计算出动态的上下文权重掩码  $W_{MCTAM}$ :

在获取多尺度空间特征后,为了突出真实的火灾辐射信号并有效抑制高反射率裸地或水面等背景噪声, MCTAM 引入了一种轻量级通道动态重标定机制。具体而言,我们得到融合特征  $F_c \in \mathbb{R}^{H \times W \times 2C}$ 。并通过全局平均池化(global average pooling, GAP)和多层感知机(multilayer perceptron, MLP)聚合通道级信息,从而计算出动态的上下文权重掩码  $W_{MCTAM}$ :

$$W_{MCTAM} = \delta \left( MLP \left( GAP \left( [F_l, F_g] \right) \right) \right) \# (7)$$

式中,  $[\cdot, \cdot]$  表示通道级联 (Concatenation) 操作,  $\delta$  为 Sigmoid 激活函数, 用于将掩码权重归一化至  $(0, 1)$  区间。公式 (7) 的物理意义在于提取通道级重标定权重, 使模型自适应地聚焦于对火点热异常最敏感的光谱通道, 同时主动抑制包含复杂地表干扰的冗余背景通道。具体而言,  $[F_l, F_g]$  表示由局部热异常感知分支和全局上下文分支输出特征级联得到的融合特征, 其通道数为  $2C$ 。GAP 首先将  $F_c$  的空间维度压缩, 得到通道描述向量  $z \in \mathbb{R}^{2C}$ 。随后, MLP 采用两层全连接瓶颈结构: 第一层全连接层将通道维度由  $2C$  压缩至  $2C/r$ , 第二层全连接层再将其恢复至  $2C$ 。本文将降维系数设置为  $r = 16$ , 因此隐藏层神经元数量为  $2C/16$ 。两层全连接层之间采用 Swish 激活函数, 最后通过 Sigmoid 函数生成通道权重  $W_{MCTAM} \in \mathbb{R}^{2C}$ 。该设计能够在较低参数量下建模不同通道之间的依赖关系, 使模型自适应增强与火点热异常相关的通道响应, 并抑制复杂背景中的冗余干扰信息。

最终, MCTAM 模块的输出特征  $F_{MCTAM}$  是通过将融合后的多尺度特征与动态权重掩码进行逐元素相乘 (Element-wise multiplication) 得到的, 并结合残差连接以保留原始多光谱输入中的底层物理信息:

$$F_{MCTAM} = \left( [F_l, F_g] \otimes W_{MCTAM} \right) + X \# (8)$$

式中,  $\otimes$  表示逐通道加权操作。由于级联后的多尺度特征  $[F_l, F_g]$  的通道数为  $2C$ , 而输入特征  $X$  的通道数为  $C$ , 为保证残差连接时通道维度一致, 本文在逐通道加权后引入  $1 \times 1$  卷积  $Conv_{1 \times 1}(\cdot)$  对特征进行融合与降维, 将通道数由  $2C$  映射回  $C$ 。同时,  $W_{MCTAM}$  在空间维度上进行广播, 以实现融合特征的逐通道重标定。

通过上述精简且高效的结构设计, MCTAM 模块不仅有效避免了常规自注意力机制在空间维度上的二次计算灾难, 还将原始多光谱张量完美转化为了富含空间热异常特征和上下文基准信息的特征图。这为后续的 Mamba 长程时序演化建模提供了高质量、去冗余的有效空间表征。

### 2.3 基于 Mamba 的长程时序演化建模

在获取由 MCTAM 模块提取的空间特征后, 模型需要进一步捕获目标区域随时间演化的动态规

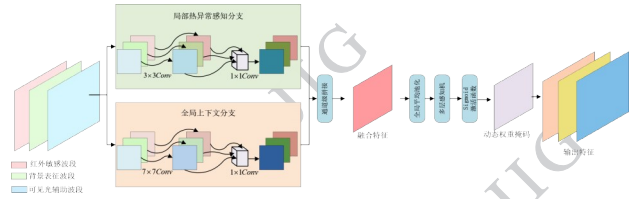


图3 多尺度上下文热异常感知模块

Fig. 3 Multi-Scale Contextual Thermal Anomaly-aware Module

律。由于 Himawari-9 卫星具备 10 min 级时间分辨率, 长时间序列观测不可避免地会引入海量帧数。传统的循环神经网络在处理超长序列时面临严重的梯度消失与遗忘问题, 而 Transformer 架构虽然具备全局感受野, 但其计算复杂度随序列长度  $T$  呈二次方增长, 这在算力受限的近实时野火监测任务中具有较大局限性。

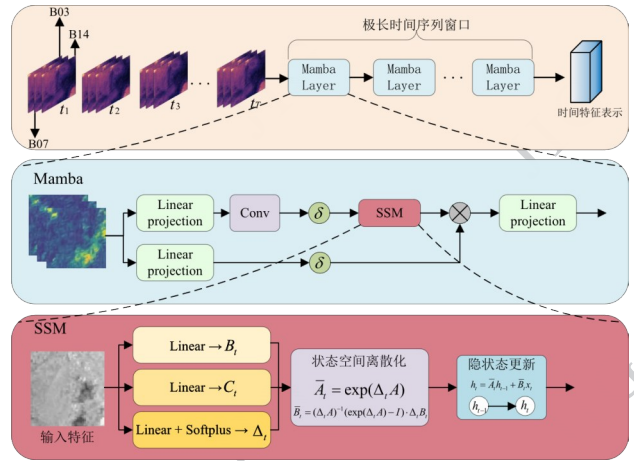


图4 基于 Mamba 的长程时序演化建模模块架构

Fig. 4 Architecture of the Mamba-based long-range temporal evolution modeling module

为了打破长时序建模的计算瓶颈, 本文引入了具有线性计算复杂度  $O(T)$  的 Mamba 架构。Mamba 的核心在于选择性状态空间模型 (Selective State Space Model, S6), 它为区分背景温度变化与异常热事件提供了更强的时序表征能力。

对于连续  $T$  个时间步的多光谱卫星输入序列, MCTAM 模块首先以共享权重的方式逐帧进行空间特征提取, 生成具有时空结构的高维特征张量  $\mathcal{F}_{MCTAM} \in \mathbb{R}^{T \times H \times W \times C}$ 。为了适配 Mamba 模块对一维序列输入的格式要求, 本文采用逐像素时序建模策略。由于空间邻域关系已在 MCTAM 模块中通过多尺度卷积充分编码, Mamba 阶段主要针对单个空间位置的跨时序动态进行建模, 从而避免时空联合建

模带来的额外计算开销。对于任意空间位置 $(i, j)$ , 提取其跨越 $T$ 个时间步的特征序列:

$$X^{(ij)} = F[:, i, j, :] \in \mathbf{R}^{T \times C} \#(9)$$

针对 Himawari-8/9 卫星 10 min 级高时间分辨率观测特性, 本文将序列长度设定为  $T = 144$ , 对应一个完整观测日内的连续时间序列。该设置能够覆盖约 24 h 的昼夜地表辐射变化周期, 使模型学习真实地表背景在太阳辐射驱动下的周期性变化规律, 从而建立更加稳定的背景辐射基准, 并抑制由晨昏交替、高反射率地表日间增温等因素引起的伪异常响应。

随后, 该序列经线性投影映射为维度为  $D$  的输入序列  $\tilde{X}^{(i, j)} \in \mathbf{R}^{T \times D}$ 。考虑到卫星多光谱特征主要表现为连续的光谱响应与辐射变化, 其特征维度相对紧凑, 本文将输入投影维度设定为  $D = 64$ , 以在保留火点多光谱时序变化信息的同时, 降低逐像素长序列建模带来的显存占用和计算开销。为简化符号, 下文省略空间坐标上标, 将状态空间模型在时间步  $t$  的输入记为多维特征向量  $x_t \in \mathbf{R}^D$ 。

经典的状态空间模型通过隐状态  $h(t) \in \mathbf{R}^N$  将输入  $x(t)$  映射为输出  $y(t)$ , 其连续时间动态过程可表示为常微分方程:

$$h'(t) = Ah(t) + Bx(t) \#(10)$$

$$y(t) = Ch(t) \#(11)$$

式中,  $A \in \mathbf{R}^{N \times N}$  为状态转移矩阵,  $B \in \mathbf{R}^{N \times D}$  和  $C \in \mathbf{R}^{D \times N}$  分别为适应多维输入的投影矩阵。本文将状态维度设定为  $N = 16$ 。对于早期火点检测任务, 火点热响应通常表现为沿时间维度的缓慢增强、短时突变或持续扩展, 而非高层语义关系的复杂跳变。因此,  $N = 16$  能够为状态空间模型提供必要的历史信息记忆能力, 同时避免过大的状态维度在 144 个时间步和逐像素密集预测条件下引入额外的计算与显存开销。

为了处理离散的卫星观测序列, Mamba 引入了步长参数  $\Delta$ , 并采用零阶保持对系统进行离散化处理。传统状态空间模型中的参数通常是固定不变的, 难以灵活应对复杂的地表动态。Fire-Mamba 通过引入选择性扫描机制, 使得离散化过程具备了数据依赖性。具体而言, 状态转移矩阵  $A$  保持共享参数形式, 以维持序列演化的稳定动力学结构; 而输入相关的投影矩阵  $B_i, C_i$  和时间步长参数  $\Delta_i$  则根据当前

观测  $x_t$  进行自适应动态调整:

$$B_i = \text{Linear}_{B_i}(x_t) \#(12)$$

$$C_i = \text{Linear}_{C_i}(x_t) \#(13)$$

$$\Delta_i = \text{Softplus}(\text{Linear}_{\Delta_i}(x_t) + \Delta_{base}) \#(14)$$

式(12)~(14)中,  $\text{Linear}(\cdot)$  表示线性映射层。在式(14)中,  $\Delta_{base}$  为模型可学习的基础时间步长参数, 而  $\text{Linear}_{\Delta_i}(x_t)$  则是基于当前瞬时卫星观测输入生成的动态调整量。Softplus 激活函数用于确保生成的时间步长始终为正值。从火点演化的物理角度来看, 这种设计的意义在于: 当卫星观测到地表温度发生剧烈突变时, 模型可根据输入动态调整时间步长参数  $\Delta_i$ , 从而迅速将异常热量输入到系统的隐状态中进行长程记忆。

基于动态生成的参数, 离散化后的状态转移矩阵  $\bar{A}_i$  和输入矩阵  $\bar{B}_i$  定义为:

$$\bar{A}_i = \exp(\Delta_i A) \#(15)$$

$$\bar{B}_i = (\Delta_i A)^{-1} (\exp(\Delta_i A) - I) \cdot \Delta_i B_i \#(16)$$

进而, 离散时间步  $t$  下的隐状态更新与输出计算转化为:

$$h_t = \bar{A}_i h_{t-1} + \bar{B}_i x_t \#(17)$$

$$y_t = C_i h_t \#(18)$$

上述隐状态更新过程(公式 17 和 18)结合野火热力学演变规律, 在物理层面上刻画了火点像素从初始热异常突变、阴燃状态逐渐向持续扩展蔓延的长程时序动态特征。通过公式(12)至(16)生成的数据依赖性动态参数, 赋予了模型依据瞬时观测状态自适应调整时序特征传递与遗忘的能力, 从而在复杂地表的高频连续观测中, 牢固捕捉真实野火信号的时序连贯性并抑制短暂的背景波动扰动。这种数据依赖的参数生成机制使模型能够根据当前输入自适应调整状态更新强度, 从而提高对不同时间演化模式的表征灵活性, 并增强对异常热事件的建模能力。最终, 对于每个空间位置  $(ij)$ , 其特征序列被输入至堆叠层数为  $L = 2$ 、内部扩展因子为  $E = 2$  的 Mamba 模块中。采用两层堆叠结构能够在增强长程时序建模能力的同时, 降低过深网络在火点正样本稀疏条件下产生过拟合的风险, 并减少前向推理开销, 从而更好地满足近实时火点监测与灾害预警的效率需求。经过 Mamba 模块的多层处理后, 取其在当前观测时刻  $T$  的输出向量  $y_T^{(ij)} \in \mathbf{R}^D$ 。该向量经线性分类层与 Sigmoid 激活函数映射为该位置发生野

火的概率值  $p^{(i,j)} \in (0, 1)$ 。遍历所有空间位置后, 模型重组并输出当前时刻  $T$  的最终火点概率图  $P \in \mathbf{R}^{H \times W}$ 。

#### 2.4 焦点损失函数

由于火点像素所占的比例通常极低, 模型训练面临极端的类别不平衡的挑战。若直接采用标准的二元交叉熵损失函数, 海量且易于分类的负样本所产生的累积梯度将主导网络的权重更新, 导致模型倾向于将所有像素预测为背景, 从而引发严重的漏报现象。为了解决这一问题, 本文引入了焦点损失函数 (Focal Loss) 作为模型优化的核心目标。Focal Loss 通过在标准交叉熵的基础上增加调制系数, 动态降低易分样本的权重, 迫使模型将优化重心转移至难以判别的样本上。

由 2.3 节可知, 模型在当前观测时刻  $T$  输出空间分辨率为  $H \times W$  的火点概率图

$$P \in \mathbf{R}^{H \times W}$$

式中, 空间位置  $(ij)$  处的预测概率记为  $p^{(ij)} \in (0, 1)$ 。设该时刻对应的真实火点二值标签图为

$$Y \in \{0, 1\}^{H \times W}$$

式中,  $y^{(ij)} = 1$  表示该像素为真实火点,  $y^{(ij)} = 0$  表示背景。由于本文采用 many-to-one 的时序预测策略, 损失函数仅在末端时刻  $T$  的输出概率图上逐像素计算。

对于空间位置  $(ij)$  处的像素, 其焦点损失定义为

$$L_{\text{Focal}}^{(ij)} = -\alpha y^{(ij)} (1 - p^{(ij)})^\gamma \log(p^{(ij)}) - (1 - \alpha) (1 - y^{(ij)}) (p^{(ij)})^\gamma \log(1 - p^{(ij)}) \quad (19)$$

$0 < \alpha < 1$  为类别平衡因子, 用于提高稀缺正样本真实火点的权重;  $\gamma \geq 0$  为聚焦参数, 用于降低易分类样本对总损失的贡献并增强模型对难分类样本的关注能力。对正类火点像素而言, 当  $p^{(ij)}$  较小时, 调制项  $(1 - p^{(ij)})^\gamma$  较大, 其损失不会被显著削弱; 对负类背景像素而言, 当  $p^{(ij)}$  接近 0 时, 调制项  $(p^{(ij)})^\gamma$  接近 0, 从而显著抑制大量易分类背景像素对全局梯度的贡献。

对于单个输入时序样本, 其末端时刻概率图上的总监督损失定义为

$$\mathcal{L}_{\text{Total}} = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W L_{\text{Focal}}^{(ij)} \quad (20)$$

## 3 实验与结果

### 3.1 实施细节和评估

本节介绍 Fire-Mamba 模型的实验环境配置、关键参数设定以及训练实施细节。主要实验环境与核心超参数设置见表 2。

表 2 Fire-Mamba 模型主要实验参数设置

Table 2 Main experimental parameter settings of the Fire-Mamba model

类别	参数名称	取值
硬件环境	CPU	Intel Core i7-12700F
硬件环境	GPU	NVIDIA RTX 4090 GPU (24GB)
软件环境	深度学习框架	PyTorch
输入设置	时间窗口长度	$T = 144$
MCTAM 模块	局部分支/上下文分支	$3 \times 3/7 \times 7$
Mamba 模块	状态维度 $N$ 扩展因子 $E$	16/2
训练设置	Batch Size/Epochs	4/50
优化器	AdamW 初始/最小学习率	$1 \times 10^{-4} \sim 1 \times 10^{-6}$
优化器	AdamW 动量参数 $(\beta_1, \beta_2)$	(0.9, 0.999)
损失函数	损失函数类型	Focal Loss ( $\alpha = 0.75, \gamma = 2.0$ )

在模型实现方面, Fire-Mamba 采用 PyTorch 框架实现,并在配备 Intel Core i7-12700F 处理器和 NVIDIA RTX 4090 GPU(24 GB 显存)的工作站上完成训练与推理。模型训练过程中, MCTAM 模块与 Mamba 模块的关键超参数均通过参数敏感性分析与消融实验进行确定。对于 MCTAM 模块,局部热异常感知分支和宏观背景建模分支的卷积核尺寸分别设置为  $3 \times 3$  和  $7 \times 7$ ,以分别提取细粒度局部热异常响应特征与大尺度背景上下文信息。对于 Mamba 模块,考虑到 Himawari-8/9 卫星 10 min 级时间分辨率下的长时间序列特性,本文采用滑动时间窗口方式构造输入样本,并将时间窗口长度设为  $T = 144$ ,对应连续 24h 的观测序列; Mamba 的状态维度  $N$  设为 16,扩展因子设为 2。模型训练采用 AdamW 优化器,初始学习率设置为  $1 \times 10^{-4}$ ,权重衰减系数设置为  $1 \times 10^{-2}$ ,并结合余弦退火策略动态调整学习率。批大小设为 4,最大训练轮数设为 50。针对火点像素极度稀疏所带来的类别不平衡问题,本文采用 Focal Loss 作为监督目标,其中类别平衡系数  $\alpha$  设为 0.75,聚焦参数  $\gamma$  设为 2.0。

为了对模型的野火检测结果进行全面的定量评估,本文选用了遥感火点检测领域广泛使用的四项像元级评价指标:火点准确率(Fire Accuracy, FA)、总体准确率(Overall Accuracy, OA)、误报率(False Alarm Rate, FAR)以及漏检率(Omission Fire Rate, OFR)。上述指标基于预测结果与真实标签对比产生的真阳性(true positive, TP)、假阴性(false negative, FN)、假阳性(false positive, FP)和真阴性(true negative, TN)构建(Zhang 等, 2025)。

为进一步验证 Fire-Mamba 的性能优势,本文选取多种代表性方法作为对比模型。考虑到 Himawari-8/9 AHI 数据具有高时间分辨率和多光谱特征,传统仅依赖空间信息的卷积神经网络难以充分利用火点的长时序演化规律,因此对比方法既包括基于空间特征提取的经典卷积网络,也包括面向时序建模的循环神经网络和 Transformer 类方法。为了保证对比实验的公平性与可重复性,本文进一步明确各基线模型的具体结构与引用依据。CNN (Krizhevsky 等, 2012; Hong 等, 2022)采用 3 层二维卷积提取空间特征,并通过 Sigmoid 函数输出火点概率图; RNN (Elman, 1990)采用 2 层循环结构对逐像素时间序列进行建模; LSTM (Hochreiter 和

Schmidhuber, 1997; Kadir 等, 2023)采用 2 层 LSTM 结构,隐藏维度设置为 64,用于捕获火点演化过程中的时序依赖关系; Transformer (Vaswani 等, 2017; Zhao 和 Ban, 2025)采用 2 层 Encoder、4 个注意力头和 128 维前馈网络,用于评估自注意力机制对长时序火点变化的建模能力; CNN-LSTM (Li 等, 2021)先利用 CNN 提取逐帧空间特征,再通过 2 层 LSTM 进行时序融合。所有基线模型均采用相同的输入波段、时间窗口、数据划分和训练策略,并在统一实验条件下与 Fire-Mamba 进行对比分析,以验证 MCTAM 模块在空间热异常表征方面的有效性以及 Mamba 模块在长程时序建模方面的优势。

### 3.2 可视化结果分析

为了直观地评估模型性能,图 5 展示了各对比模型在火灾发生后 24 小时内,对研究区域检测到的火灾像素频率分布。颜色越深表示该像素在连续观测序列中被判定为火点的次数越多,可用于反映模型对火点区域的持续响应能力和空间稳定性。从图 5 可以看出, Fire-Mamba 的检测结果在形状和空间分布上与真实标签最为接近,能够准确覆盖火灾蔓延的主要区域。相比之下, JAXA WLF L2 产品在多数场景中仅能检测到少量高强度火点,对早期弱小火点响应不足,在大理、凉山等事件中存在明显漏检现象。LSTM 模型的检测频率图产生大量孤立的误报像素点,表明其在复杂背景下容易产生伪火点响应; Transformer 模型虽然在核心区域表现较好,但在边缘地带仍存在识别缺失,说明仅依赖全局时序关系难以稳定刻画弱小火点的空间结构。CNN-LSTM 模型相比单一时序模型具有更好的空间聚集性,但在火点边界区域仍存在一定程度的扩散和不完整问题。Fire-Mamba 通过 MCTAM 模块提取的空间热异常特征,有效地抑制了地表高反射噪声,使得频率图的背景区域保持极高的纯净度,显著优于 JAXA WLF L2 产品及其他深度学习基线模型。同时, Mamba 模块对长时序演化信息的建模进一步增强了模型对持续性热异常的判别能力,使得检测结果在空间上更加连续、边界更加清晰。

为了进一步验证模型对极早期阴燃及微小火点的敏感度,本文选取了六个火灾事件的关键时间节点进行逐帧分析。图 6 中红色表示正确检测火点,蓝色表示非火区域,浅绿色表示误报,浅黄色表示漏检,从而更清晰地地区分不同模型的检测行为。在重

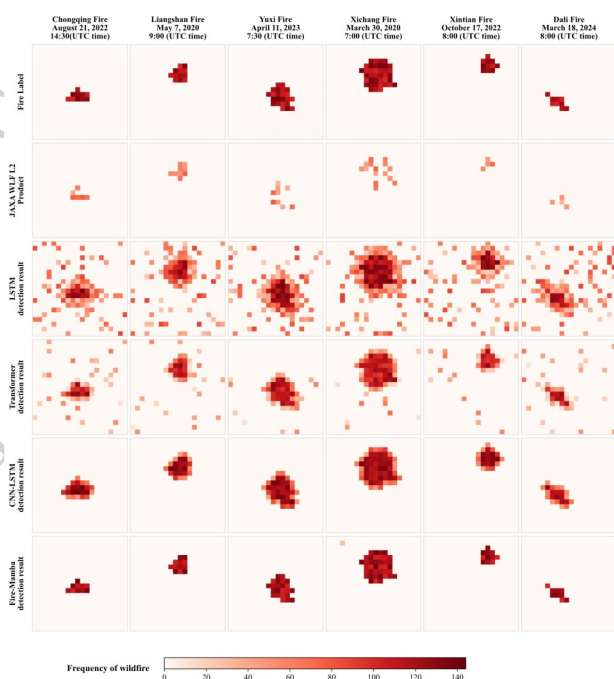


图5 各对比模型检测到的火灾像素频率分布图

Fig. 5 Frequency distribution of detected fire pixels by different comparison models

庆火灾(14:30)和大理火灾(08:00)等样本中,火点在  $BT_{07}-BT_{14}$  差值图中仅表现为极微弱的亮度突变。在该阶段,JAXA 产品及部分深度学习模型存在明显漏检,而 Fire-Mamba 能够准确捕捉到这些亚像素级或少像素级的热异常,表现出更高的早期火点敏感性。在凉山火灾与西昌火灾的快速蔓延期,火灾斑块形态复杂。LSTM 和 Transformer 方法在多个时刻出现明显的误报点或漏检区域,尤其在云干扰或背景亮温波动较强时,易产生离散噪声响应;CNN-LSTM 虽能保持主体区域,但在边界细节上仍存在缺失。得益于 Mamba 模块的长程时序建模能力,Fire-Mamba 能够利用前序帧的背景基准自适应调整当前的检测阈值,从而生成边界清晰的火点预测掩膜。在多个连续时间步中,Fire-Mamba 的检测结果与真实标签在空间位置与形态结构上保持高度一致,误报显著减少,漏检情况也得到有效缓解。

### 3.3 数据结果分析

Fire-Mamba 的 FA 达到 0.9033,为所有对比方法中的最高值,较次优的 CNN-LSTM 提高了 0.0805,较 JAXA WLF L2 Products 提高了 0.4967。这一结果表明,Fire-Mamba 在真实火点像素识别方面具有显著优势,能够更有效地检测出火灾区域。同时 OFR 结果为 0.0967,为全部方法中的最低值,表明

模型在降低漏检方面表现最优,进一步说明 Fire-Mamba 对微弱火点和早期火情具有更高的敏感性。Fire-Mamba 的 OA 为 0.9960,略高于 CNN-LSTM (0.9946)和 JAXA WLF L2 Products (0.9924)。这说明本文方法不仅能够较准确地区分火点与背景,而且在全图尺度上保持了稳定的分类能力。在误报率 FAR 指标上,Fire-Mamba 的平均值为 0.1880,并非所有方法中的最低值,但明显优于 RNN (0.6901)、LSTM (0.7752)和 Transformer (0.4526),也优于 CNN (0.2202)和 CNN-LSTM (0.1967)。

从不同火灾事件的检测结果来看,Fire-Mamba 在六个案例中均保持了较高的 FA 水平,分别达到 0.8376、0.9368、0.9127、0.9274、0.9189 和 0.8863。Fire-Mamba 在六起火灾事件中均取得最高 FA,表明模型在不同区域、不同背景和不同火势发展阶段下均具有较强的泛化能力。

### 3.4 消融实验分析

为验证本文所提模型中各模块(多尺度上下文热异常感知模块 MCTAM、状态空间建模模块 Mamba 以及 Focal Loss)的独立有效性,以及空间提取卷积核参数选取的合理性,本文在统一的实验环境下设计了详细的消融实验。评价指标采用 6 起火灾事件的平均表现。

#### (1) 核心模块有效性验证

为了明确各模块的具体贡献,实验以基础 Mamba 网络(不含多尺度空间特征提取与焦点损失)作为基线模型(Baseline),并分别设计了引入 MCTAM 模块、引入 Focal Loss 以及使用传统 CNN-LSTM 替换 Mamba 模块的对比实验。实验结果如表 4 所示。

由表 4 可见,与基线模型相比,仅引入 MCTAM 模块即可将平均火点准确率(FA)由 0.8536 提升至 0.8854,表明双分支空间特征提取机制对早期微小火点的辐射异常具有更高的敏感性。当仅引入 Focal Loss 时,漏检率(OFR)显著下降,进一步表明该损失函数在缓解火点像素极度稀疏问题方面的有效性。此外,在保留 MCTAM 与 Focal Loss 用于空间特征提取与损失约束的条件下,将长程时序建模模块由 Mamba 替换为传统 LSTM 后,模型 FA 下降至 0.8615,漏检率(OFR)上升至 0.1385。这表明,在处理 Himawari-8/9 极长时序连续观测数据(144 个时间步)时,Mamba 较 LSTM 更能缓解长时记忆衰减,从

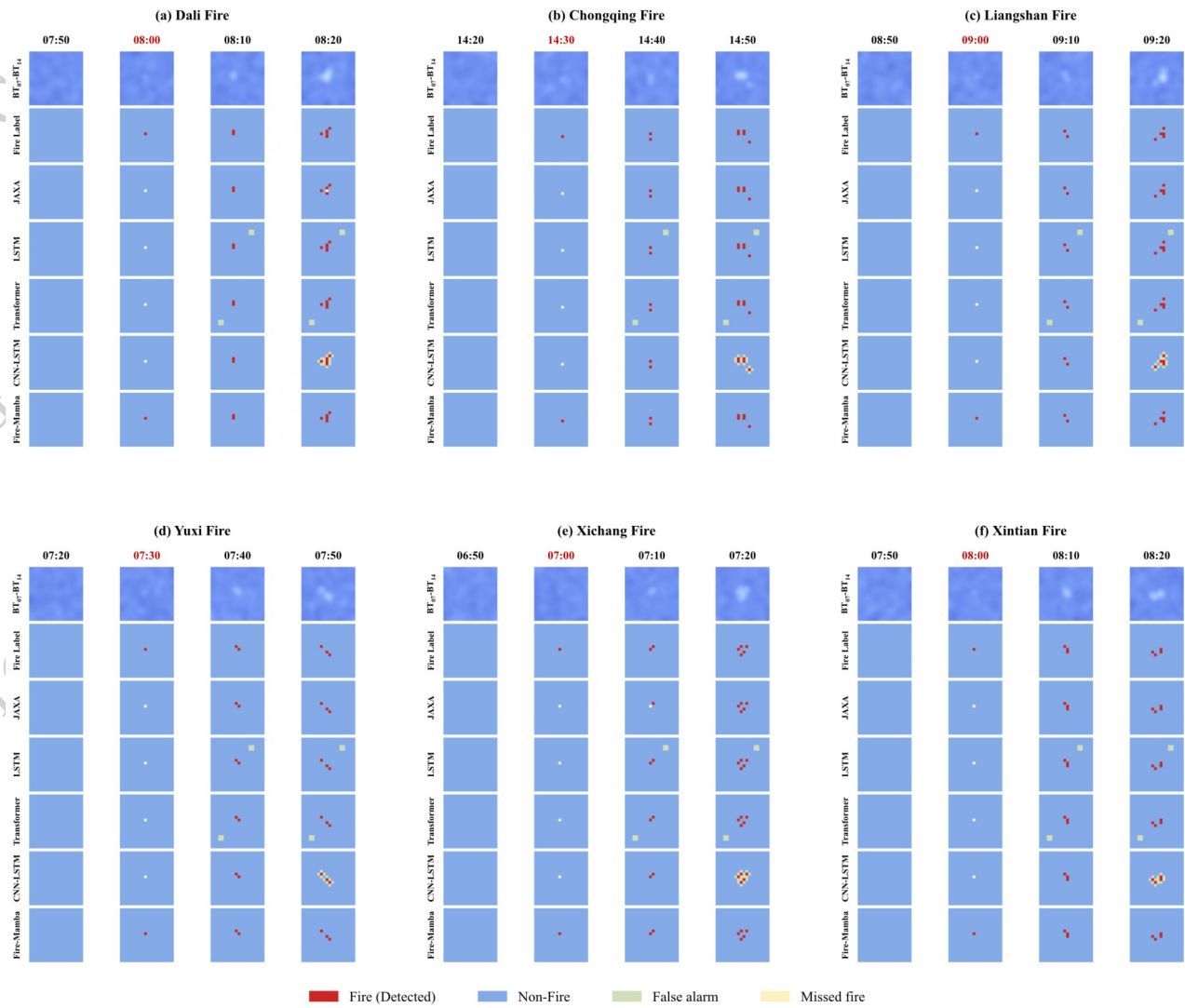


图6 典型火灾事件关键时间节点检测结果对比图

Fig. 6 Comparison of frame-by-frame detection results at key time nodes of typical fire events

而有效捕获火点演化规律并降低漏检。最终,融合上述三种关键设计的 Fire-Mamba 模型取得了最佳的综合检测性能。

### (2) MCTAM 模块卷积核尺寸消融分析

为进一步验证 MCTAM 模块中双分支卷积核尺寸设置的合理性,在保持模型整体架构(包括 Mamba 与 Focal Loss)不变的条件下,对局部感知分支与宏观上下文分支的不同卷积核尺寸组合进行了交叉对比实验。实验结果见表 5。

如表 5 所示,不同卷积核尺寸组合对检测性能有影响。当移除  $3 \times 3$  局部感知分支时(采用  $5 \times 5$  与  $7 \times 7$  组合),由于对单像素及亚像素级热梯度表征能力不足,火点准确率(FA)下降。在保留  $3 \times 3$  分支的情况下,较小的上下文感受野( $3 \times 3$  与

$5 \times 5$ )难以提供充分的背景参照以抑制地表噪声;而过大的感受野( $3 \times 3$  与  $9 \times 9$ )则引入冗余及干扰信息,导致误报率(FAR)升高(0.2026)。综合来看, $3 \times 3 + 7 \times 7$  的组合在微弱热源表征与背景噪声抑制之间实现了最佳平衡。

### 3.5 模型计算效率与近实时应用分析

静止气象卫星火点监测的核心应用价值在于早期“近实时”预警。Himawari-8/9 卫星以 10 分钟的高频率不断生成观测数据,这对深度学习模型的推理时效性和显存占用提出了严格的工程要求。为了验证 Fire-Mamba 架构在极长序列时序建模任务中的计算效率优势,本文在相同的硬件平台(RTX4090)下,对各对比模型的计算复杂度进行了量化评测,具体包括模型参数量、处理单个 144 帧时间窗口的推

表3 各火点检测模型在6起真实火灾事件上的定量评价结果对比

Table 3 Quantitative evaluation results comparison of different fire detection models on six real fire events

火点事件	评价指标	JAXA WLF L2 Products	CNN	RNN	LSTM	Transformer	CNN- LSTM	Fire- Mamba
重庆火灾	FA	0.7364	0.7812	0.2936	0.4187	0.2014	0.8089	<b>0.8376</b>
	OA	<b>0.9974</b>	0.9968	0.9016	0.8469	0.9598	0.9962	0.9972
	FAR	<b>0.0281</b>	0.1624	0.9682	0.9746	0.9628	0.2146	0.1983
	OFR	0.2636	0.2188	0.7064	0.5813	0.7986	0.1911	<b>0.1624</b>
凉山火灾	FA	0.1248	0.8654	0.8013	0.8429	0.7548	0.9147	<b>0.9368</b>
	OA	0.9917	0.9859	0.9274	0.9088	0.9908	0.9947	<b>0.9971</b>
	FAR	<b>0.0369</b>	0.2463	0.8345	0.9114	0.4471	0.1769	0.1692
	OFR	0.8752	0.1346	0.1987	0.1571	0.2452	0.0853	<b>0.0632</b>
西昌火灾	FA	0.1689	0.8341	0.7562	0.8715	0.8019	0.8864	<b>0.9127</b>
	OA	0.9836	0.9894	0.9886	0.9869	<b>0.9954</b>	0.9938	0.9945
	FAR	0.0887	0.1856	0.2017	0.3842	<b>0.0736</b>	0.1658	0.1327
	OFR	0.8311	0.1659	0.2438	0.1285	0.1981	0.1136	<b>0.0873</b>
玉溪火灾	FA	0.6157	0.8579	0.8265	0.9184	0.8841	0.9016	<b>0.9274</b>
	OA	0.9929	0.9728	0.9863	0.9835	<b>0.9968</b>	0.9942	0.9938
	FAR	<b>0.0418</b>	0.2769	0.3184	0.5017	0.0581	0.2085	0.2214
	OFR	0.3843	0.1421	0.1735	0.0816	0.1159	0.0984	<b>0.0726</b>
新田火灾	FA	0.4521	0.7965	0.7438	0.9056	0.8732	0.8925	<b>0.9189</b>
	OA	0.9952	0.9921	0.9507	0.9194	0.9949	0.9963	<b>0.9974</b>
	FAR	<b>0.0174</b>	0.2317	0.8461	0.9023	0.2865	0.1714	0.1548
	OFR	0.5479	0.2035	0.2562	0.0944	0.1268	0.1075	<b>0.0811</b>
大理火灾	FA	0.3415	0.5018	0.3189	0.4627	0.4368	0.5324	<b>0.8863</b>
	OA	0.9934	0.9911	0.8642	0.8126	0.9582	0.9924	<b>0.9959</b>
	FAR	<b>0.0395</b>	0.2184	0.9715	0.9768	0.8876	0.2431	0.2517
	OFR	0.6585	0.4982	0.6811	0.5373	0.5632	0.4676	<b>0.1137</b>
平均值	FA	0.4066	0.7728	0.6234	0.7366	0.6587	0.8228	<b>0.9033</b>
	OA	0.9924	0.9880	0.9365	0.9097	0.9826	0.9946	<b>0.9960</b>
	FAR	<b>0.0421</b>	0.2202	0.6901	0.7752	0.4526	0.1967	0.1880
	OFR	0.5934	0.2272	0.3766	0.2634	0.3413	0.1772	<b>0.0967</b>

理时间、平均单帧推理时间以及运行时的 GPU 显存占用。测试结果如表6所示。

由表6的对比结果可知,面对包含144个时间步的高维极长特征序列,Transformer架构由于自注意力机制在时间维度上的二次方计算复杂度( $O(T^2)$ ),导致其推理时间最长(达到3.92s)且显存占用最高(6.7GB),难以满足大规模区域的密集监测部署需

求。主流的CNN-LSTM模型虽然在精度上表现较好,但受限于LSTM循环计算的串行瓶颈,其单时间窗口的推理时间也达到了2.35s。单纯的空间CNN虽然速度最快(0.41s),但完全舍弃了火点演化的时序动态信息,导致其在先前的实验中检测精度不佳。

相比之下,本文提出的Fire-Mamba模型展现出了较好的计算效率。得益于MCTAM模块精简的多

表 4 核心模块消融实验结果

Table 4 Ablation study results of core modules

实验类别	模型设置	MCTAM	Mamba	Focal Loss	FA	OA	FAR	OFR
基线模型	Baseline:基础 Mamba		√		0.8536	0.9949	0.2038	0.1464
模块消融	Baseline+MCTAM	√	√		0.8854	0.9957	<b>0.1764</b>	0.1146
模块消融	Baseline+Focal Loss		√	√	0.8769	0.9952	0.2135	0.1231
结构替换	MCTAM+LSTM+Focal Loss	√		√	0.8615	0.9951	0.1912	0.1385
本文方法	Fire-Mamba	√	√	√	<b>0.9033</b>	<b>0.9960</b>	0.1880	<b>0.0967</b>

表 5 MCTAM 模块卷积核尺寸消融实验结果

Table 5 Ablation study results of convolutional kernel sizes in the MCTAM module

实验类别	卷积核组合	FA	OA	FAR	OFR
尺寸消融	3 × 3 + 5 × 5	0.8887	0.9956	0.1849	0.1113
尺寸消融	3 × 3 + 7 × 7	<b>0.9033</b>	<b>0.9960</b>	<b>0.1880</b>	<b>0.0967</b>
尺寸消融	5 × 5 + 7 × 7	0.8938	0.9958	0.1902	0.1062
尺寸消融	3 × 3 + 9 × 9	0.8991	0.9957	0.2026	0.1009

尺度特征提取设计以及 Mamba 模块在长序列上的线性计算复杂度( $O(T)$ ), Fire-Mamba 在保持 3.21M 适中参数量的前提下,将 GPU 显存占用严格控制在 3.6GB 左右。其处理完整的 144 帧时间窗口仅需 1.28s,平均单帧推理时间低至 8.89ms。该推理速度不仅优于 Transformer,且相较于 CNN-LSTM 模型提速近 1 倍。综上所述,Fire-Mamba 不仅在微弱火点检测精度上实现了提升,其高效的推理速度与低显存开销也较好满足了 Himawari-8/9 卫星 10 分钟观测频率的业务化“近实时”动态监测需求。

## 4 结论

本文针对 Himawari-8/9 卫星影像中早期野火火点弱小、识别困难的问题,提出了 Fire-Mamba 火点检测模型。该模型通过 MCTAM 模块提取局部热梯度和背景上下文特征,并结合 Mamba 模块建模长时序演化信息,从而提高了复杂背景下微弱火点的识别能力。在 6 起真实森林火灾事件中展现出了优异的检测性能与计算效率。与现有基于 Himawari-8/9 卫星的深度学习火点检测方法相比,本文方法的核心优势与差异在于:有效打破了传统方法中空间特征提取与时序动态建模相割裂的局限。一方面, MCTAM 模块避免了单一感受野的盲区,极大地提升了对亚像素级微弱热异常的空间敏感度;另一方面, Mamba 架构有效缓解了 RNN 类模型在处理 10 分钟级高频连续观测数据(如 144 帧长序列)时的长时记忆衰减问题,且规避了 Transformer 的二次方计算开销。两者的高效融合,使得 Fire-Mamba 能够在较低的延迟与显存占用下,实现大尺度复杂背景中微小弱目标的近实时、高灵敏度检测。尽管本文方

表 6 各模型计算效率与推理时间对比

Table 6 Comparison of computational efficiency and inference time among models

模型	时序模块	参数量/M	单个 144 帧时间窗口推理时间/s	平均单帧推理时间/ms	GPU 显存占用/GB
CNN	None	1.82	0.41	2.85	2.1
RNN	RNN	2.34	1.76	12.22	3.4
LSTM	LSTM	2.91	2.08	14.44	3.8
CNN-LSTM	LSTM	3.76	2.35	16.32	4.2
Transformer	Self-Attention	4.58	3.92	27.22	6.7
Fire-Mamba	Mamba	3.21	1.28	8.89	3.6

法在多数场景下表现出较强的鲁棒性,但客观上仍存在一定的局限性。首先,受限于真实高精度逐帧标注样本的规模,模型在跨大洲气候区与跨季节条件下的泛化能力仍需进一步验证。其次,作为基于光学与红外波段的检测模型,在面对极端厚云覆盖或强浓烟遮挡导致地表热辐射严重衰减的极端场景时,模型仍不可避免地存在性能短板与漏检风险。此外,Himawari-8/9 卫星红外波段 2 km 的物理空间分辨率上限,也在一定程度上限制了火灾斑块绝对边界的精细化刻画。在后续的研究中,我们将进一步探索无监督或半监督学习范式,以降低模型对大规模人工精细标注数据的依赖。同时,考虑引入多源卫星数据融合机制(如结合极轨卫星的高空间分辨率或微波雷达的穿透能力),以期在极端气象覆盖干扰下进一步提升野火监测系统的全天候业务化应用水平。

**致谢:**本文所使用的 Himawari-8/9 卫星 AHI 观测数据由日本气象厅(Japan Meteorological Agency, JMA)提供,在此对其数据共享服务表示衷心感谢。同时,感谢编辑部和审稿专家对本文提出的宝贵意见和建议。

## 参考文献(References)

- Bao M, Lyu S, Xu Z, Zhou H, Ren J, Xiang S, et al. 2026. Vision Mamba in remote sensing: a comprehensive survey of techniques, applications and outlook. *Remote Sensing*, 18(4): 594 [DOI: 10.3390/rs18040594]
- Barmpoutis P, Papaioannou P, Dimitropoulos K and Grammalidis N. 2020. A review on early forest fire detection systems using optical remote sensing. *Sensors*, 20(22): 6442 [DOI: 10.3390/s20226442]
- Bessho K, Date K, Hayashi M, Ikeda A, Imai T, Inoue H, et al. 2016. An introduction to Himawari-8/9—Japan's new-generation geostationary meteorological satellites. *Journal of the Meteorological Society of Japan*, 94(2): 151-183 [DOI:10.2151/jmsj.2016-009]
- Chen J, Zheng W, Wu S, Liu C and Yan H. 2022. Fire Monitoring Algorithm and Its Application on the Geo-Kompsat-2A Geostationary Meteorological Satellite. *Remote Sensing*, 14(11): 2655 [DOI: 10.3390/rs14112655]
- Chen J, Lv Q, Wu S, Zeng Y, Li M, Chen Z, et al. 2023. An adapted hourly Himawari-8 fire product for China: principle, methodology and verification. *Earth System Science Data*, 15: 1911-1931 [DOI:10.5194/essd-15-1911-2023]
- 曹丽琴,汪都,熊海洋,钟燕飞. 2024. 热红外高光谱遥感影像信息提取方法综述. *中国图象图形学报*, 29(08): 2089-2112 [DOI: 10.11834/jig.230738]
- Da C. 2015. Preliminary assessment of the Advanced Himawari Imager (AHI) measurement onboard Himawari-8 geostationary satellite. *Remote Sensing Letters*, 6(8): 637-646 [DOI: 10.1080/2150704X.2015.1066522]
- Elman J L. 1990. Finding structure in time. *Cognitive Science*, 14(2): 179-211 [DOI:10.1016/0364-0213(90)90002-E]
- Filizzola C, Corrado R, Marchese F, Mazzeo G, Paciello R, Pergola N, et al. 2016. RST-FIRES, an exportable algorithm for early-fire detection and monitoring: description, implementation, and field validation in the case of the MSG-SEVIRI sensor. *Remote Sensing of Environment*, 186: 196-216 [DOI:10.1016/j.rse.2016.08.008]
- Fusco E J, Finn J T, Abatzoglou J T, Balch J K, Dadashi S and Bradley B A. 2019. Detection rates and biases of fire observations from MODIS and agency reports in the conterminous United States. *Remote Sensing of Environment*, 220: 30-40 [DOI: 10.1016/j.rse.2018.10.028]
- Giglio L, Descloitres J, Justice C O and Kaufman Y J. 2003. An enhanced contextual fire detection algorithm for MODIS. *Remote Sensing of Environment*, 87(2-3): 273-282 [DOI:10.1016/S0034-4257(03)00184-6]
- Gu A and Dao T. 2023. Mamba: Linear-Time Sequence Modeling with Selective State Spaces[EB/OL].[2026-03-17]. <https://arxiv.org/abs/2312.00752>
- Guo Y Q, Chen G, Wang Y N, Zha X M and Xu Z D. 2022. Wildfire identification based on an improved two-channel convolutional neural network. *Forests*, 13(8): 1302 [DOI:10.3390/f13081302]
- 高广帅,尚云琦,董燕. 2025. 光学遥感图像小目标检测深度学习算法综述. *中国图象图形学报*, 30(11): 3479-3505 [DOI: 10.11834/jig.240740]
- Hally B, Wallace L, Reinke K, Jones S, Engel C and Skidmore A. 2018. Estimating fire background temperature at a geostationary scale—an evaluation of contextual methods for AHI-8. *Remote Sensing*, 10(9): 1368 [DOI:10.3390/rs10091368]
- Han C H, Yang Q, He X H, Mei Y, Ding M, Li Y, et al. 2023. Research on fire spot detection algorithm based on the new generation of geostationary meteorological satellite. *Remote Sensing Technology and Application*, 38(1): 173-181 (韩琛惠, 杨倩, 贺晓慧, 梅焱, 丁敏, 李妍, 等. 2023. 基于 Himawari-8 卫星的火点判别算法研究. *遥感技术与应用*, 38(1): 173-181) [DOI: 10.11873/j.issn.1004-0323.2023.1.0173]
- Hering A S, Bell C L and Genton M G. 2009. Modeling spatio-temporal wildfire ignition point patterns. *Environmental and Ecological Statistics*, 16(2): 225-250 [DOI:10.1007/s10651-007-0080-6]
- Hochreiter S and Schmidhuber J. 1997. Long short-term memory. *Neural Computation*, 9(8): 1735-1780 [DOI: 10.1162/neco.1997.9.8.1735]
- Hong Z, Tang Z, Pan H, Zhang Y, Zheng Z, Zhou R, et al. 2022.

- Active fire detection using a novel convolutional neural network based on Himawari-8 satellite images. *Frontiers in Environmental Science*, 10: 794028 [DOI:10.3389/fenvs.2022.794028]
- Kadir E A, Kung H T, AlMansour A A, Irie H, Rosa S L and Fauzi S S. 2023. Wildfire hotspots forecasting and mapping for environmental monitoring based on the long short-term memory networks deep learning algorithm. *Environments*, 10 (7) : 124 [DOI: 10.3390/environments10070124]
- Krizhevsky A, Sutskever I and Hinton G E. 2012. ImageNet classification with deep convolutional neural networks//Proceedings of the 25th International Conference on Neural Information Processing Systems. Lake Tahoe: Curran Associates, Inc.: 1097-1105
- Kulakowski P, Calle E and Marzo J L. 2013. Performance study of wireless sensor and actuator networks in forest fire scenarios. *International Journal of Communication Systems*, 26(4) : 515-529 [DOI: 10.1002/dac.2311]
- Li A, Zhao Y and Zheng Z. 2022. Novel recursive BiFPN combining with Swin Transformer for wildland fire smoke detection. *Forests*, 13(12) : 2032 [DOI:10.3390/f13122032]
- Li X, Gao H, Zhang M, Zhang S, Gao Z, Liu J, et al. 2021. Prediction of forest fire spread rate using UAV images and an LSTM model considering the interaction between fire and wind. *Remote Sensing*, 13(21) : 4325 [DOI:10.3390/rs13214325]
- Li Z, Zhao L, Lu Y, Ma Y and Li G. 2026. Mamba for remote sensing: architectures, hybrid paradigms, and future directions. *Remote Sensing*, 18(2) : 243 [DOI:10.3390/rs18020243]
- Liu C, Chen R and He B. 2023. Integrating machine learning and a spatial contextual algorithm to detect wildfire from Himawari-8 data in Southwest China. *Forests*, 14(5) : 919 [DOI:10.3390/f14050919]
- Liu Y, Tian Y, Zhao Y, Yu H, Xie L, Wang Y, et al. 2024. VMamba: visual state space model. *Advances in Neural Information Processing Systems*, 37: 103031-103063
- Maeda N and Tonooka H. 2023. Early stage forest fire detection from Himawari-8 AHI images using a modified MOD14 algorithm combined with machine learning. *Sensors*, 23(1) : 210 [DOI: 10.3390/s23010210]
- Rogers B M, Soja A J, Goulden M L and Randerson J T. 2015. Influence of tree species on continental differences in boreal fires and climate feedbacks. *Nature Geoscience*, 8: 228-234 [DOI:10.1038/ngeo2352]
- Seydi S T, Akhoondzadeh M, Amani M and Mahdavi S. 2021. Wildfire damage assessment over Australia using Sentinel-2 imagery and MODIS land cover product within the Google Earth Engine cloud platform. *Remote Sensing*, 13 (2) : 220 [DOI: 10.3390/rs13020220]
- Soro TD, Koné M, N'Dri AB, N'Datchoh ET. 2021. Identified main fire hotspots and seasons in Côte d'Ivoire (West Africa) using MODIS fire data. *South African Journal of Science*, 117(1/2) [DOI: 10.17159/sajs.2021/7659]
- Tian X R, Shu L F, Zhao F J and Wang M Y. 2012. Forest fire danger changes for Southwest China under future scenarios. *Scientia Silvae Sinicae*, 48(1) : 121-125 (田晓瑞, 舒立福, 赵凤君, 王明玉. 2012. 未来情景下西南地区森林火险变化. *林业科学*, 48(1) : 121-125) [DOI:10.11707/j.1001-7488.20120120]
- 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; Curran Associates, Inc.: 5998-6008
- Xiao J, Chevallier F, Gomez C, Guanter L, Hicke J A, Huete A R, et al. 2019. Remote sensing of the terrestrial carbon cycle: A review of advances over 50 years. *Remote Sensing of Environment*, 233: 111383 [DOI:10.1016/j.rse.2019.111383]
- Xiao Z, Liang S, Wang J, Xiang Y, Zhao X and Song J. 2016. Long-Time-Series Global Land Surface Satellite Leaf Area Index Product Derived From MODIS and AVHRR Surface Reflectance. *IEEE Transactions on Geoscience and Remote Sensing*, 54 (9) : 5301-5318 [DOI:10.1109/TGRS.2016.2560522]
- Xie Z, Song W, Ba R, Li X and Xia L. 2018. A spatiotemporal contextual model for forest fire detection using Himawari-8 satellite data. *Remote Sensing*, 10(12) : 1992 [DOI:10.3390/rs10121992]
- Yan J J, Qu J H, Ran M N and Zhang F F. 2020. Himawari-8 AHI fire detection in clear sky based on time-phase change. *Journal of Remote Sensing*, 24(5) : 571-577 (鄢俊洁, 瞿建华, 冉茂农, 张芳芳. 2020. 基于时相变化的晴空条件下 Himawari-8 AHI 火点检测. *遥感学报*, 24 (5) : 571-577) [DOI: 10.11834/jrs.20209122]
- Yuan C, Liu Z and Zhang Y. 2015. UAV-based forest fire detection and tracking using image processing techniques//Proceedings of the 2015 International Conference on Unmanned Aircraft Systems. Denver: IEEE: 639-643 [DOI:10.1109/ICUAS.2015.7152345]
- Zhang G J, Bo Y C, He S, Ye J X, Yang S, Kou W L, et al. 2026. Himawari-9 and FY-3D satellite wildfire monitoring performance analysis. *Journal of Southwest Forestry University*, 46(1) : 182-188 (张国将, 柏盈川, 贺帅, 叶江霞, 杨松, 寇卫利, 等. 2026. Himawari-9 及 FY-3D 卫星野火监测的性能分析. *西南林业大学学报(自然科学)*, 46(1) : 182-188) [DOI: 10.11929/j.swfu.202408010]
- Zhang L, Wang M, Fu Y and Ding Y. 2022. A forest fire recognition method using UAV images based on transfer learning. *Forests*, 13 (7) : 975 [DOI:10.3390/f13070975]
- Zhang Q, Zhang X, Quan C, Zhao T, Huo W and Huang Y. 2025. Mamba-STFM: a Mamba-based spatiotemporal fusion method for remote sensing images. *Remote Sensing*, 17(13) : 2135 [DOI: 10.3390/rs17132135]
- Zheng Z, Hu H, Huang W, Zhou F, Ma Y, Liu Q, et al. 2024. Wildfire detection based on the spatiotemporal and spectral features of Himawari-8 data. *IEEE Transactions on Geoscience and Remote Sensing*, 62: 5408213 [DOI:10.1109/TGRS.2024.3434434]

Zhu L, Liao B, Zhang Q, Wang X, Liu W and Wang X. 2024. Vision Mamba: efficient visual representation learning with bidirectional state space model[EB/OL].[2026-05-03].

<https://arxiv.org/abs/2401.09417> [DOI: 10.48550/arXiv. 2401.09417]

Zhu X X, Tuia D, Mou L, Xia G S, Zhang L, Xu F, et al. 2017. Deep learning in remote sensing: a comprehensive review and list of resources. IEEE Geoscience and Remote Sensing Magazine, 5(4): 8-36 [DOI:10.1109/MGRS.2017.2762307]

### 作者简介

张本硕,男,硕士研究生,主要研究方向为计算机视觉与遥感

图像处理。E-mail:zbenhuo@163.com

赵峰,通信作者,男,教授,博士,主要研究方向为机器学习与遥感图像处理。E-mail:zhaofeng1016@126.com

柯枫,女,讲师,主要研究方向为遥感图像处理。E-mail:kefeng@yitsd.edu.cn

安志勇,男,副教授,博士,主要研究方向为遥感图像处理。E-mail:201513351@sdtbu.edu.cn

于晓宁,女,讲师,博士,主要研究方向为计算机视觉与模式识别。E-mail:yuxiaoning@sdtbu.edu.cn

韩忠伟,男,助教,主要研究方向为计算机视觉与遥感图像处理。E-mail:hzw\_han@163.com