0254-6124/2025/45(1)-0082-09 *Chin. J. Space Sci.* 空间科学学报

LI Ming, CUI Yanmei, LUO Bingxian. Machine Learning Solar Full Disk Flare Operational Forecasting (in Chinese). *Chinese Journal of Space Science*, 2025, **45**(1): 82–90. CSTR:32142.14.cjss.2024–0021. DOI:10.11728/cjss2025.01.2024–0021

全日面机器学习耀斑业务预报*

李 铭[1,2,3] 崔延美[1,3] 罗冰显[1,2,3]

1(中国科学院国家空间科学中心 空间天气学国家重点实验室 北京 100190)

2(中国科学院大学 北京 100049)

3(中国科学院空间环境态势感知技术重点实验室 北京 100190)

摘 要 太阳耀斑预报是空间环境预报中的一项重要内容. 目前所构建的深度学习耀斑预报模型大都是基于活动区磁图构建的. 受到投影效应的影响, 这类模型仅能对太阳中心区域的活动区进行预报, 难以满足全日面耀斑业务预报的需求. 基于太阳活动传统参量, 研究统计了活动区磁类型、面积, 耀斑爆发历史以及 $F_{10.7}$ 与耀斑发生的关系, 利用全连接神经网络构建了适用于全日面活动区的太阳耀斑业务预报模型, 该模型可以预报未来 48 h 内活动区≥M 级耀斑的爆发情况. 该模型与已往搭建的深度学习预报模型进行比较, 结果表明本文建立的业务预报模型预报性能略优. 同时, 结果表明投影效应对本研究搭建的耀斑预报模型影响不大. 该模型为耀斑业务预报提供了有效工具.

关键词 耀斑, 业务预报, 机器学习, 全日面活动区

中图分类号 P353

Machine Learning Solar Full Disk Flare Operational Forecasting

LI Ming[1,2,3] CUI Yanmei[1,3] LUO Bingxian[1,2,3]

1(*State Key Laboratory of Space Weather*, *National Space Science Center*, *Chinese Academy of Sciences*, *Beijing* 100190)

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

3(*Key Laboratory of Science and Technology on Environment Space Situation Awareness*, *Chinese Academy of Sciences*, *Beijing* 100190)

Abstract Solar flare forecasting is an essential component in space environment forecasting. Most of the deep learning flare forecasting models constructed are based on the magnetograms of active regions. Affected by the projection effect, these models can only forecast the active region in the center of the Sun. It is difficult to meet the need of operational flare forecasting of the solar full disk. Based on the traditional solar activity parameters, in this study, the relationships between the magnetic type of the

* 中国科学院战略性先导科技专项 (XDB0560000), 国家自然科学基金项目 (42074224) 和中国科学院国家空间科学中心“攀登计划”项目共同资助

2024-02-05 收到原稿, 2024-09-03 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/)

active region, area of the active region, the history of the flare outburst, the 10 cm radio flux and flares from January 1996 to December 2022 were statistically analyzed. By using the fully connected neural network, an operational flare forecasting model for solar full disk active regions was constructed. This model can forecast the eruption of the M-class or above flares of the full solar disk active regions in the next 48 h. The F1 score of the model is 0.4304, the TSS is 0.3689, and the HSS is 0.3906. The model is compared with the deep learning flare forecasting model constructed in the previous work, and the results show that the operational forecasting model constructed in this paper has a better forecasting performance. Meanwhile, in order to explore the influence of the projection effect, the solar full disk active regions flare forecasting model constructed was tested for test data within 30 degrees from the center of the solar disk, within the interval from 30 degrees to 60 degrees, and over 60 degrees, respectively. The results show that the projection effect has little influence on the flare forecast model constructed in this study. The model can be used to forecast flare in the active region of the full solar disk, and provide an effective tool for operational solar flare forecasting.

Key words Flare, Operational forecasting, Machine learning, Solar full disk active regions

0 引言

太阳耀斑是太阳大气局部区域的能量瞬间释放事件, 伴随着强烈的电磁辐射和粒子喷射[1], 可能对通信、导航、卫星运行产生不良影响. 为减小太阳耀斑及伴随事件对地球空间技术系统造成的负面影响, 空间天气业务机构均会提供耀斑预报的相关服务, 例如美国空间天气预报中心 SWPC*、中国空间环境预报中心 SEPC**、韩国空间天气中心 KSWC***等均提供耀斑预报服务.

在过去的 10 多年中, 太阳观测已经积累了大量高时空分辨率的数据, 特别是太阳动力学观测卫星 (SDO) 发射后. SDO 卫星上搭载了三台科学仪器, 分别是大气成像仪 (AIA), 日震与磁成像仪 (HMI) 和极紫外线变化观测仪 (EVE), 可以得到太阳大气成像、磁场、极紫外线辐照度等多种高精度数据. 例如, HMI 每 12 min 采集 1 arcse 空间精度日面磁场, 每天收集的数据量达 1 TByte.

深度学习方法由于其出色的处理大数据的能力被广泛应用于耀斑预报中[2–10]. 例如, 2018 年 Huang 等[3] 使用 SDO/ HMI 活动区 (SHARP) 磁图搭建了卷积神经网络耀斑预报模型; 2019 年 Chen 等[8] 提取了 SDO SHARP 矢量磁图数据中的时序信息搭建了长短时记忆神经网络 (LSTM) 耀斑预报模型; 2021 年 Guastavino 等[9] 使用 SDO SHARP 视频磁图搭建了结合卷积神经网络和长短时记忆神经网络的混合深度学习耀斑预报模型; 2023 年 Li 等[10] 使用 SDO SHARP 磁图数据建立融合了耀斑爆发先验知识的深度学习预报模型. 然而, 考虑投影效应的影响, 这些基于活动区磁图的深度学习模型只能对位于日面中心区域附近的活动区进行预报, 不能完全满足全日面耀斑业务预报的需求. 因此, 面向业务预报的全日面耀斑预报模型还有待继续研究.

业务上, 耀斑短期预报仍以人工经验预报为主, 模型参考主要基于太阳活动传统参量 (例如黑子分类、射电流量、磁特征参数等) 为预报因子的模式. 这类模型可以对全日面活动区是否发生耀斑预报, 但通常建立在 10 年前. 例如, 1989 年 Miller 等[11] 开发了经验模型 WOLF, 实现了基于活动区的 McIntosh 分型进行耀斑预报; 2009 年 Li 等[12] 使用活动区磁分类、黑子面积以及太阳 10 cm 射电流量数据, 建立了基于支持向量机与近邻法相结合的耀斑预报模型; 2012 年 Shaun 等[13] 使用地球静止轨道环境业务卫星 (GOES) 的 X 射线观测数据及 McIntosh 分类计算耀斑发生的泊松概率, 进而实现耀斑预测. 在近 10 年的时间里已经积累了更多的太阳活动传统参量数据, 重新统计这些参量与耀斑的爆发关系, 统计关系将具有更高的可靠性, 以此搭建的耀斑预报模型的

* https://www.swpc.noaa.gov/products/forecast-discussion

** http://www.sepc.ac.cn/dailyForecast_chn.php

*** http://www.spaceweather.go.kr

预报结果将更准确, 性能将更好.

在本研究中, 考虑到太阳活动水平, 活动区爆发历史, 活动区磁复杂性等因素, 重新统计 1996—2022 年太阳活动区磁类型、面积, 耀斑爆发历史和 $F_{10.7}$ 指数与耀斑的爆发关系, 基于全连接神经网络方法构建了适用于全日面活动区的太阳耀斑业务预报模型, 模型可以预报活动区未来 48 h 内≥M 级耀斑的爆发情况.

1 数据选择与处理

活动区信息和耀斑爆发信息由 NOAA/SWPC* 提供的相应列表给出. 活动区列表在每日 00:30 UTC 发布, 提供了时间 (年月日)、活动区编号、磁类型、面积等信息. 耀斑列表提供了耀斑起始时间、等级及对应的活动区编号等信息. $F_{10.7}$ 指数信息来自 OMNI 数据集, 由 NASA 提供**. 数据选取的时间范围为 1996 年 1 月至 2022 年 12 月.

1.1 耀斑样本与非耀斑样本定义

当一个活动区在未来 48 h 内产生 M 级或以上级别耀斑时, 该样本被定义为耀斑样本, 标记为 1. 反之, 如果一个活动区未来 48 h 内没有产生 M 级或以上级别的耀斑, 相应的样本则被定义为非耀斑样本, 标注为 0. 最后, 共获得 2415 个耀斑样本和 33976 个非耀斑样本.

1.2 太阳活动参量与耀斑发生的统计关系

统计耀斑爆发历史、活动区磁类型、活动区面积以及 $F_{10.7}$ 指数与活动区未来 48 h 内爆发 M 级或以上等级的耀斑的发生关系. 耀斑爆发历史是指被预测活动区在前 24 h 内是否有 M 级或以上耀斑爆发. 若一个活动区在预测前 24 h 内有 M 级或以上的耀斑爆发, 则记为过去 24 h 内有耀斑爆发的样本. 反之则被记为过去 24 h 内未爆发耀斑的样本. 活动区的威尔逊山磁类型在这里被分为 4 类, 分别是 α 类、β 类、γ 类 (γ, β-γ)、δ 类 (δ, δ-γ, β-γ-δ, β-δ). 在统计耀斑爆发历史、活动区磁类型与耀斑的发生关系时, 将这两个参数同时考虑, 即统计考虑耀斑爆发历史的活动区磁类型与耀斑的发生关系. 首先, 按照爆发历史对样本进行分类, 将样本分为过去 24 h 内未爆发≥M 级耀斑和过去 24 h 内爆发≥M 级耀斑两类. 之后, 在每类样本中按照活动区磁类型分别统计 α, β, γ, δ 这 4 类样本的总样本数和耀斑样本数, 根据耀斑样本数及总样本数计算耀斑发生率.

表 1 和图 1 给出了考虑耀斑爆发历史后的活动区磁类型与耀斑的产生关系, 包含总样本数、耀斑样本数、耀斑发生率信息.

从统计结果上看, 过去 24 h 内未爆发≥M 级耀斑的活动区中, α 类、β 类、γ 类、δ 类的活动区对应的耀斑发生率分别为 0.0109, 0.0391, 0.1844, 0.3995. 过去 24 h 内爆发≥M 级耀斑的活动区中, α 类、β 类、γ 类、δ 类的活动区对应的耀斑发生率分别为 0.2105, 0.2633, 0.5042, 0.7307. 随着活动区的磁结构复杂性提高, 耀斑发生率上升. 同时, 如果 α 类、β 类、γ 类、δ 类的活动区在过去 24 h 内爆发≥M 级耀斑, 在未来 48 h 内爆发≥M 级耀斑的可能性更高.

活动区面积以太阳半球的百万分之一 (MH) 为单位, 在统计活动区面积与耀斑的产生关系时, 按照面积大小分为 5 个区间: 0～100, 100～300, 300～600, 600～1300, ≥1300. 分别统计这 5 个面积区间下

表 1 考虑耀斑爆发历史后的活动区磁类型与耀斑爆发的统计关系

Table 1 Statistical relationship between the magnetic types of active regions and flares considering the history of flare eruptions

磁类型	过去24 h内未爆发≥M级耀斑			过去24 h内爆发≥M级耀斑		
	总样本数	耀斑样本数	耀斑发生率	总样本数	耀斑样本数	耀斑发生率
α	10275	112	0.0109	38	8	0.2105
β	19008	743	0.0391	281	74	0.2633
γ (γ, β-γ)	2559	472	0.1844	357	180	0.5042
δ (δ, δ-γ, β-γ-δ, β-δ)	791	316	0.3995	375	274	0.7307

* https://www.swpc.noaa.gov/

** https://omniweb.gsfc.nasa.gov/form/dx1.html

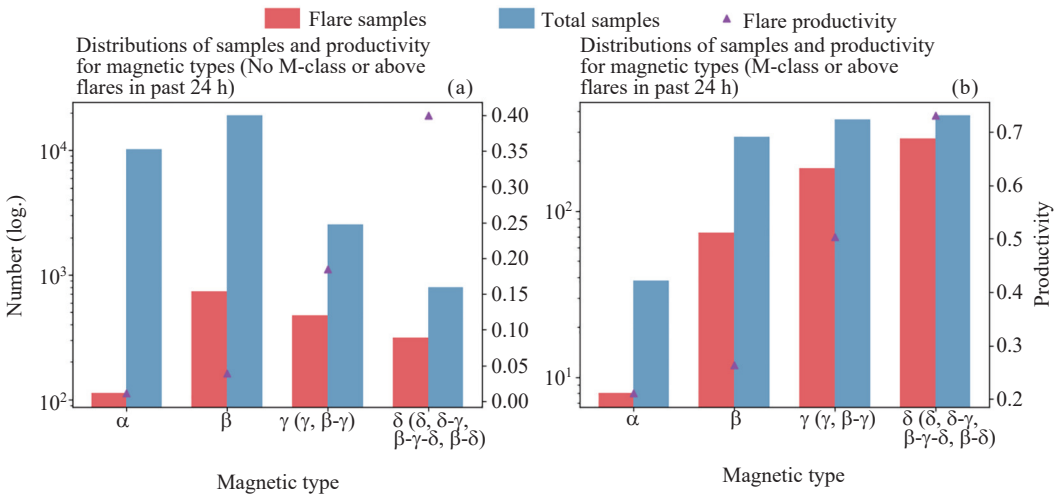


图 1　考虑耀斑爆发历史的活动区磁类型与耀斑爆发的统计关系

Fig. 1　Statistical relationship considering the history of flare eruptions between the magnetic types of active regions and flares

的总样本数、耀斑样本数和耀斑发生率. 表 2 和图 2 给出了活动区面积与耀斑的产生关系, 包含总样本数、耀斑样本数、耀斑发生率信息. 对于不同面积区间的活动区, 0～100, 100～300, 300～600, 600～1300, ≥1300 区间对应的耀斑发生率分别为 0.0167, 0.0938, 0.2349, 0.4473, 0.8440. 即耀斑发生率随着活动区面积的增大而增加.

在统计 $F_{10.7}$ 指数与耀斑的产生关系时, 按照数值大小分为 4 个区间, 分别为 0～91 sfu, 91～169 sfu, 169～221 sfu, ≥221 sfu. 分别统计这 4 个区间下的总样本数、耀斑样本数和耀斑发生率. 表 3 和图 3 给出了 $F_{10.7}$ 指数与耀斑的产生关系, 包含总样本数、耀斑样本数、耀斑发生率信息. 区间为 0～91 sfu, 91～169 sfu, 169～221 sfu, ≥221 sfu 的 $F_{10.7}$ 指数对应的耀斑发生率分别为 0.0146, 0.0588, 0.0934, 0.1450, 即 $F_{10.7}$ 指数越大, 太阳活动水平越高, 耀斑发生率也越高.

将考虑耀斑爆发历史的活动区磁类型、活动区面积以及 $F_{10.7}$ 指数对应的耀斑发生率作为输入搭建耀斑业务预报模型.

2　全日面活动区的机器学习耀斑业务预报模型

2.1　数据集划分

数据集划分主要考虑了两个方面. 一方面是机器学习中积累的历史经验[14], 通常将 3/5 的数据划分为训练集, 1/5 的数据为验证集, 1/5 的数据为测试集; 另一方面是为了确保训练集、验证集和测试集中有完全不同的活动区样本, 数据集划分考虑了时间顺序的原则. 在这两个原则下, 对数据集进行了整年划分. 训练集的时间范围为 1996 年 1 月至 2010 年 12 月, 验证集的时间范围为 2011 年 1 月至 2013 年 12 月, 测试集的时间范围为 2014 年 1 月至 2022 年 12 月. 各数据集中的正负样本数目在表 4 中给出.

表 2　活动区面积与耀斑爆发的统计关系

Table 2　Statistical relationship between the area of active regions and flares

面积/MH	总样本数	耀斑样本数	耀斑发生率
0～100	21979	368	0.0167
100～300	8397	788	0.0938
300～600	2354	553	0.2349
600～1300	845	378	0.4473
≥1300	109	92	0.8440

2.2　耀斑预报模型搭建

在神经网络领域, 全连接神经网络作为多层感知机的基础结构, 展现了其灵活性和强大的表示学习能力[15]. 这一模型由输入层、隐藏层和输出层组成, 其独特之处在于每个神经元都与相邻层中的所有神经元相连接, 从而形成全局性的权重关系. 这种全连接的架构使得网络能够适应各种复杂的数据关系, 通过

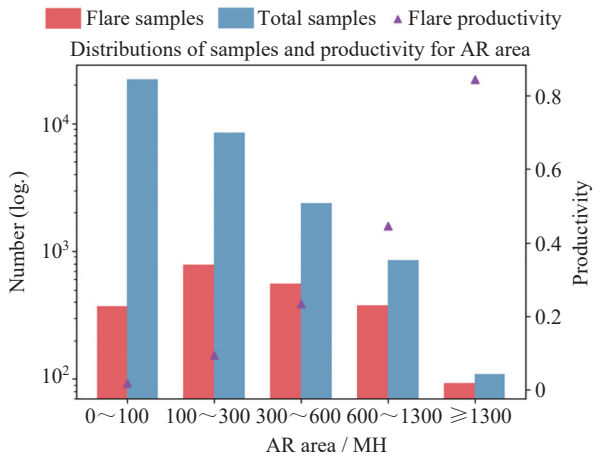


图 2 活动区面积与耀斑爆发的统计关系

Fig. 2 Statistical relationship between the area of Active Regions (ARs) and flares

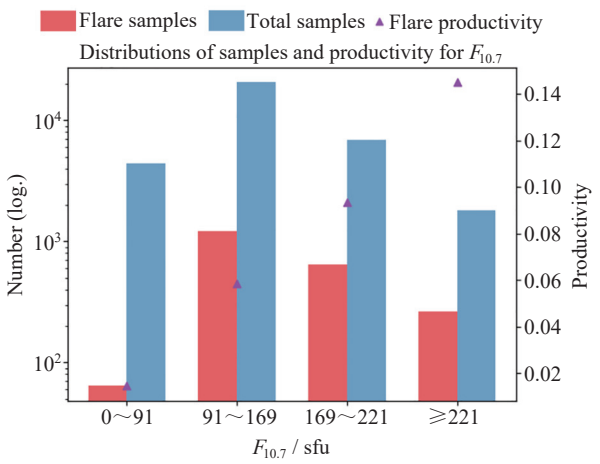


图 3 $F_{10.7}$ 指数与耀斑爆发的统计关系

Fig. 3 Statistical relationship between $F_{10.7}$ index and flares

表 3 $F_{10.7}$ 指数与耀斑爆发的统计关系

Table 3 Statistical relationship between $F_{10.7}$ index and flares

$F_{10.7}$/sfu	总样本数	耀斑样本数	耀斑发生率
0～91	4386	64	0.0146
91～169	20602	1211	0.0588
169～221	6849	640	0.0934
≥221	1821	264	0.1450

多层次的非线性变换, 可以有效捕捉输入数据的抽象特征.

使用包含一个隐藏层的全连接神经网络搭建耀斑业务预报模型, 其结构如图 4 所示. 其中输入层由三个神经元组成, 分别为考虑耀斑爆发历史的活动区磁类型、活动区面积以及 $F_{10.7}$ 指数的耀斑发生率. 输出层由两个神经元组成, 分别对应未来 48 h 内有无≥M 级耀斑爆发.

在搭建网络时, 需要对超参数[16] 进行设置. 由于耀斑样本的数量远少于非耀斑样本的数量, 这种不平衡可能导致模型在预测时偏向于预测数量较多的类别, 而忽略数量较少的类别. 解决样本不平衡问题的常见方法包括过采样、欠采样、改变样本权重和修改决策阈值等方法. 这里选取改变样本权重的方法来减少样本不平衡问题带来的影响. 为了寻找最佳的权重, 分别使用不同的样本权重进行实验并统计了 F_1 分数[17] 的变化, 其中横坐标为模型正例样本权重,

表 4 数据集分布情况

Table 4 Distribution of data sets

	训练集	验证集	测试集
耀斑样本数	1484	350	581
非耀斑样本数	20836	5388	7752
总样本数	22320	5738	8333

纵坐标为在此权重下模型的 F_1 分数. 实验结果如图 5 所示. 从图 5 可以看出, 实验结果在 0.8 左右时 F_1 分数最高, 即模型的效果最好. 因此, 在设置网络参数时, 将正例样本的权重设置为 0.8, 负例样本的权重设置为 0.2. 同时, 在本文建模过程中所使用的其他超参数设置如表 5 所示.

在对模型的超参数设置完成后, 将训练集数据输入至搭建好的网络结构中进行训练. 验证集在训练模型的过程中用于调整模型参数如学习率等, 进而提高模型的性能并避免过拟合或欠拟合的问题. 训练过程中模型的损失函数变化曲线如图 6 所示, 模型在 Epoch 接近 20 左右时收敛速度下降, 且接近收敛.

在测试集中, 模型的预报结果由二分类混淆矩阵给出, 如表 6 所示. 二分类混淆矩阵包含 4 个参数, 分别为 N_{TP}, N_{TN}, N_{FP}, N_{FN}. 这里 N_{TP} 和 N_{TN} 表示被正确预报的耀斑样本量和非耀斑样本量, N_{FP} 和 N_{FN} 表示未被正确预报的耀斑样本量和非耀斑样本量. 搭建的全日面活动区耀斑预报模型 N_{TP}, N_{TN}, N_{FP}, N_{FN} 分别为 235, 7476, 346, 276.

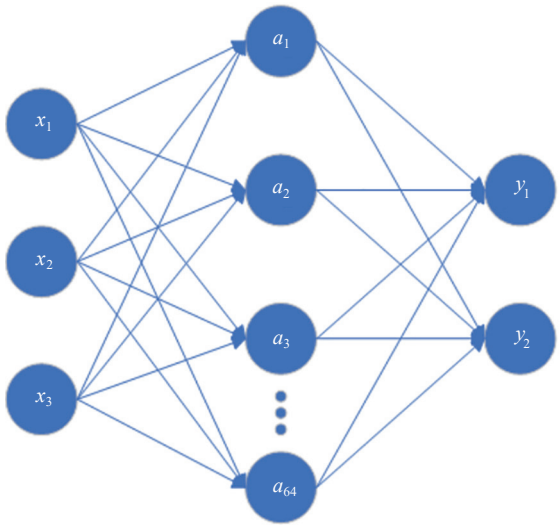


图 4　全连接神经网络结构. 该网络由输入层、隐藏层和输出层组成. x_1～x_3表示网络的三个输入参量, a_1～a_{64}表示神经网络隐藏层中的 64 个神经元, y_1和y_2表示网络的两个输出

Fig. 4　Structure of fully connected neural network. The network consists of input, hidden and output layers. x_1～x_3 denote the three input parameters of the network, a_1～a_{64} denote the 64 neurons in the hidden layer of the neural network, and y_1 and y_2 denote the two outputs of the network

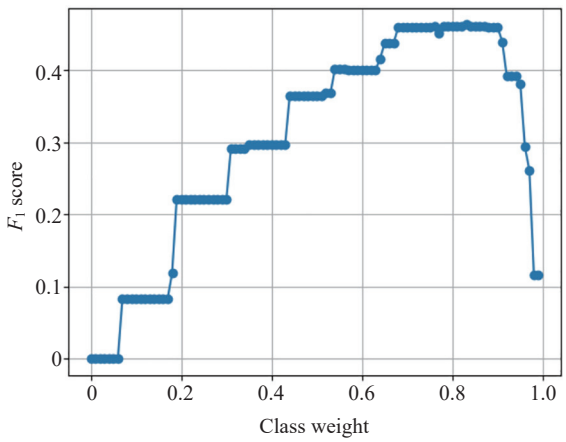


图 5　不同权重下模型的 F_1 分数

Fig. 5　F_1 score of the model under different class weights

3　模型评估

为了评估模型的预报性能, 计算了几个常用的评估指标: 精确率 (P)、召回率 (R)、F_1 分数[17]、准确率 (ACC)、关键成功指数 (CSI)[18]、虚报率 (FAR)、海德克技能评分 (HSS)[19] 和真实技能统计 (TSS)[20], 计算公式及模型在测试集对应的相关评分如表 7 所示. P 表示在模型预测为耀斑的样本中, 有多少样本真正属于耀斑样本. R 表示在实际耀斑样本中, 模型成功预测为耀斑样本的比例. ACC 表示模型正确预测的样本比例, 是模型在整个数据集上的整体表现. CSI 将精确率和召回率结合, 旨在提供模型对耀斑预测性

表 5　全连接神经网络超参数设置

Table 5　Hyperparameter settings for fully connected neural networks

超参数名称	设置值
Epoch	100
隐藏层数量	1
隐藏层中神经元数量	64
优化算法	Adam
隐藏层激活函数	Relu
输出层激活函数	Sigmoid
损失函数	二元交叉熵

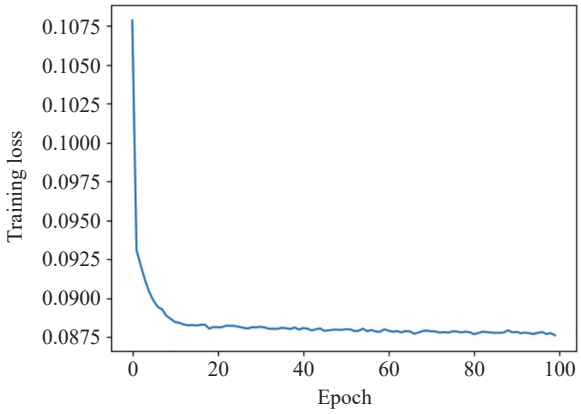


图 6　训练过程中模型的损失函数变化情况

Fig. 6　Loss function variation of the model during training

表 6　全日面活动区预报模型的二分类混淆矩阵

Table 6　Binary confusion matrices for active region forecasting models of the full solar disk

预测	观测	
	真	假
真	235 (N_{TP})	346 (N_{FP})
假	276 (N_{FN})	7476 (N_{TN})

表 7 评价指标及其对应得分

Table 7 Evaluation metrics and corresponding scores

评价指标	计算公式	预报模型
F_1 分数	$\frac{2P\times R}{P+R}$	0.4304
P	$\frac{N_{\rm TP}}{N_{\rm TP}+N_{\rm FP}}$	0.4599
R	$\frac{N_{\rm TP}}{N_{\rm TP}+N_{\rm FN}}$	0.4045
ACC	$\frac{N_{\rm TP}+N_{\rm TN}}{N_{\rm TP}+N_{\rm FP}+N_{\rm FN}+N_{\rm TN}}$	0.9245
CSI	$\frac{N_{\rm TP}}{N_{\rm TP}+N_{\rm FP}+N_{\rm FN}}$	0.2472
FAR	$\frac{N_{\rm FP}}{N_{\rm TP}+N_{\rm FP}}$	0.5401
TSS	$\frac{N_{\rm TP}}{N_{\rm TP}+N_{\rm FN}}-\frac{N_{\rm FP}}{N_{\rm FP}+N_{\rm TN}}$	0.3689
HSS	$\frac{2[(N_{\rm TP}\times N_{\rm TN})-(N_{\rm FN}\times N_{\rm FP})]}{(N_{\rm TP}+N_{\rm FN})(N_{\rm FN}+N_{\rm TN})+(N_{\rm TP}+N_{\rm FP})(N_{\rm FP}+N_{\rm TN})}$	0.3906

能的全面评估. FAR 衡量了模型在非耀斑样本上的误报程度, 即将实际为非耀斑类别的样本错误地预测为耀斑样本的比例. TSS 综合考虑了 $N_{\rm TP}$, $N_{\rm TN}$, $N_{\rm FP}$, $N_{\rm FN}$, 旨在提供对模型在正负类别上的全面评估. HSS 旨在衡量模型相对于随机猜测的性能表现. 该指标考虑了模型的正确分类、误分类以及随机猜测的因素, 提供了对模型能力的全面评估. 精确率、召回率、准确率、CSI 和 F_1 分数的值范围在 0～1 之间, 越接近 1 表示模型的性能越好. FAR 的值范围在 0～1 之间, 越接近 0 表示模型在非耀斑样本上的误报率越低, 性能越好. TSS 和 HSS 的值范围为–1～1, 值越高, 表示模型性能越好. 由于耀斑是小概率事件且会对日地空间环境造成危害, 因此在耀斑业务预报中预报员更关注耀斑事件报对 ($N_{\rm TP}$)、虚报 ($N_{\rm FN}$) 和漏报 ($N_{\rm FP}$) 情况, 在评估模型时需要同时考虑这三个属性. 由于 F_1 分数综合考虑了 $N_{\rm TP}$, $N_{\rm TN}$, $N_{\rm FN}$ 等方面的影响, 因此选择 F_1 分数作为主要评价指标, 其他评分如 ACC, CSI, HSS 和 TSS 等作为参考指标.

从表 7 可以看出, 全日面活动区耀斑预报模型 F_1 分数为 0.4304, ACC 为 0.9245, CSI 为 0.2472, FAR 为 0.5401, TSS 为 0.3689, HSS 为 0.3906. 从 ACC 上分析, 在测试集中有 90% 的样本都被正确预报. 但其他分数例如 F_1 分数、TSS、HSS 等都不高, 均未超过 0.5. 这主要是因为耀斑事件为小概率事件, 其发生存在很大的偶然因素. 另外, 耀斑样本数量较少, 而非耀斑的样本数量较多, 这会使模型可能更倾向于预报为非耀斑事件, 导致 ACC 提高, 但其他指标不高. 如果只关注数据集中的耀斑样本被正确预报, 模型可能会倾向于预测更多样本为耀斑样本, 以确保不错过任何真正的耀斑样本. 这可能导致模型的精度降低, 因为其更倾向于将非耀斑样本错误分类为耀斑样本; 如果只关注预测为耀斑的样本中有多少是真正的耀斑样本, 模型可能会更加谨慎地进行耀斑样本的预测, 但这可能导致漏掉一些真正的耀斑样本. 这本来就是互相矛盾的, 所以在建模时需要综合考虑这两方面的影响. 因此, 在调整样本权重时选用了考虑 $N_{\rm TP}$, $N_{\rm FP}$ 和 $N_{\rm FN}$ 的 F_1 分数作为标准.

为了评估投影效应对全日面活动区耀斑业务预报模型预报效果的影响, 采取了分区间测试的方法. 将测试集中的活动区样本按照日面位置 ($L_{\rm O}$) 分为三个区间: 距离日面中心 30°、30°～60°和超过 60°, 对应的样本数量分别为 3467, 2990 和 1876. 全日面活动区耀斑预报模型在这三个区间测试数据上的预报结果和评估指标 F_1 分数在表 8 给出. 在这三个区间测试数据中, 模型的 F_1 分数分别为 0.4247, 0.4422, 0.4219, 即数值相差不大; 模型在三个区间测试数据的 F_1 分数与总测试样本中的 F_1 分数相近. 由此, 认为投影效应对本研究中搭建的耀斑预报模型影响不大, 该模型可以预报全日面活动区的耀斑爆发情况.

4 模型比较

为了进一步比较该模型的预报性能, 将该模型与之前工作中[9] 利用 CNN 方法搭建的 Fusion model 2 进行对比. Fusion model 2 是以 SDO SHARP 视向磁图与活动区面积作为输入, 基于卷积神经网络搭建的深度学习耀斑预报模型, 该模型可以预测日面中心 30°的活动区未来 48 h 内是否爆发≥M 级耀斑.

为了保持两组模型中样本数据的一致性, 选择每日零点且位于日面中心±30°的活动区进行比较. 所选取数据的时间范围为 2020 年 1 月至 2022 年 12 月. 共得到 905 个样本, 其中耀斑样本数为 73, 非耀斑样本为 832. 两模型在所选取数据中的评估结果在表 9 中给出.

从表 9 中可以看出, Fusion model 2 在 ACC (0.9017), P(0.3333) 和 FAR(0.6667) 的评分优于全日面活动区模型. 全日面活动区模型在 F_1 分数 (0.2879), R(0.2603), CSI(0.1681), TSS(0.2122) 和

表 8　模型预报性能比较

Table 8　Comparison of model forecast performance

指标	$L_O \leqslant 30°$	$30° < L_O \leqslant 60°$	$L_O > 60°$
N_{TP}	93	88	54
N_{FP}	114	126	106
N_{FN}	138	96	42
N_{TN}	3122	2680	1674
F_1分数	0.4247	0.4422	0.4219

表 9　相同数据源模型预报性能比较

Table 9　Comparison of model forecast performance in the same data source

指标	Fusion model 2	全日面活动区模型
N_{TP}	16	19
N_{FP}	32	40
N_{FN}	57	54
N_{TN}	800	792
F_1分数	0.2645	0.2879
P	0.3333	0.3220
R	0.2192	0.2603
ACC	0.9017	0.8961
CSI	0.1524	0.1681
FAR	0.6667	0.6780
TSS	0.1807	0.2122
HSS	0.2142	0.2325

HSS(0.2325) 优于 Fusion model 2. 考虑到主要评估指标 F_1 分数的比较结果, 认为全日面活动区模型的预报性能要略优于 Fusion model 2. 同时, 也发现全日面活动区耀斑预报模型在该数据样本中, F_1 分数要低于第 3 节模型评估中的 F_1 分数 (0.4304). 这可能是由于该数据源的测试数据样本是有限的. 在该时间段内, 总样本数为 905 个, 其中耀斑样本数只有 73 个. 有限的样本数量可能导致模型的性能没有得到充分体现, 最终使得模型在该部分得到的 F_1 分数与第 3 节中结果存在较大偏差. 因此, 随着未来数据的积累, 本文模型与之前基于 CNN 方法搭建的 Fusion model 2 可以在更大的数据集上进行性能对比.

5　讨论与结论

太阳耀斑预报是空间天气预报中的一项重要内容. 目前所构建的深度学习耀斑预报模型大都是基于活动区磁图构建的. 这类模型通常考虑投影效应的影响, 仅对太阳中心区域的活动区进行建模预报, 不能满足全日面耀斑业务预报的需求. 基于太阳活动传统参量为预报因子的模型可以实现对全日面活动区是否发生耀斑进行预报, 但距离这类模型的建立已经过去了 10 多年的时间. 因此重新统计太阳活动传统参量与耀斑的关系发展业务预报模型是非常必要的.

使用 1996 年 1 月至 2022 年 12 月的太阳活动数据, 统计太阳活动区磁类型、面积、耀斑爆发历史以及 $F_{10.7}$ 指数与耀斑的关系, 基于全连接神经网络方法构建了适用于全日面活动区的太阳耀斑业务预报模型, 该模型可以预报未来 48 h 内活动区≥M 级耀斑的爆发情况. 模型的 ACC 为 0.9245, CSI 为 0.2472, TSS 为 0.3689, HSS 为 0.3906, F_1 分数为 0.4304. 同时, 为探究投影效应的影响, 搭建的全日面活动区耀斑预报模型对距离日面中心 30°、30°～60°和超过 60°的测试数据分别测试. 结果表明, 模型在这三个区域的 F_1 分数相差不大并且与使用全日面数据搭建的耀斑预报模型的 F_1 分数相近. 因此, 认为投影效应对本研究中搭建的耀斑预报模型影响不大.

为进一步比较该全日面活动区耀斑预报模型的预报性能, 使用 2020 年 1 月至 2022 年 12 月的数据, 将搭建的全日面活动区模型与之前工作中搭建的深度学习耀斑预报模型 Fusion model 2 在相同的数据源中对比. 结果表明, Fusion model 2 的 F_1 分数为 0.2645, 全日面活动区模型的 F_1 分数为 0.2879. 全日面活动区模型的预报性能要略优于 Fusion model 2.

研究表明, 日面边缘数据的处理还存在着不足. 对于日面东边缘的活动区, 由于在过去 24 h 中未出现在日面可观测范围内, 在标注其耀斑爆发历史时可能存在漏标注的情况. 对于日面西边缘的活动区, 未来 48 h 内可能会转出日面观测范围, 会有部分耀斑样本活动区在存在漏标注的情况. 限于观测条件, 当前还难以对以上日面边缘出现的情况进行解决, 这也是影响模型预报性能提高的重要因素, 期待未来能够有相应的观测进行补充. 基于深度学习方法的耀斑预报建模研究是空间天气预报研究的一个主要方向, 但目前深度学习耀斑预报的优势还没充分发挥出来. 开展深度学习耀斑预报模型与传统业务预报模型的集合预报研究或许是当前满足业务预报需求的一个研究方向.

参考文献

[1] LI M, CUI Y M, LUO B X, *et al.* Knowledge-informed deep neural networks for solar flare forecasting[J]. *Space Weather*, 2022, **20**(8): e2021SW002985

[2] WOODS T N, KOPP G, CHAMBERLIN P C. Contributions of the solar ultraviolet irradiance to the total solar irradiance during large flares[J]. *Journal of Geophysical Research: Space Physics*, 2006, **111**(A10): A10S14

[3] HUANG X, WANG H N, XU L, *et al.* Deep learning based solar flare forecasting model. I. results for line-of-sight magnetograms[J]. *The Astrophysical Journal*, 2018, **856**(1): 7

[4] (LI Rong, HUANG Xin. Solar flare forecasting model based on automatic feature extraction of sunspots[J]. *Scientia Sinica Physica, Mechanica & Astronomica*, 2018, **48**(11): 119601 (李蓉, 黄鑫. 一种黑子特征自动提取的太阳耀斑模型 [J]. 中国科学: 物理学 力学 天文学, 2018, **48**(11): 119601)

[5] NISHIZUKA N, SUGIURA K, KUBO Y, *et al.* Deep Flare Net (DeFN) model for solar flare prediction[J]. *The Astrophysical Journal*, 2018, **858**(2): 113

[6] LIU H, LIU C, WANG J T L, *et al.* Predicting solar flares using a long short-term memory network[J]. *The Astrophysical Journal*, 2019, **877**(2): 121

[7] SUN H, MANCHESTER W, JIAO Z B, *et al.* Interpreting LSTM prediction on solar flare eruption with time-series clustering[OL]. arXiv preprint arXiv: 1912.12360, 2019

[8] CHEN Y, MANCHESTER W B, HERO A O, *et al.* Identifying solar flare precursors using time series of SDO/HMI images and SHARP parameters[J]. *Space Weather*, 2019, **17**(10): 1404-1426

[9] GUASTAVINO S, MARCHETTI F, BENVENUTO F, *et al.* Implementation paradigm for supervised flare forecasting studies: a deep learning application with video data[J]. *Astronomy & Astrophysics*, 2022, **662**: A105

[10] LI M, CUI Y M, LUO B X, *et al.* Deep neural networks of solar flare forecasting for complex active regions[J]. *Frontiers in Astronomy and Space Sciences*, 2023, **10**: 1177550

[11] MILLER R W. WOLF—A computer expert system for sunspot classification and solar flare prediction[M]//HECK A, MURTAGH F. Knowledge-Based Systems in Astronomy. Berlin, Heidelberg: Springer, 1989

[12] LI Rong, CUI Yanmei. Solar flare forecasting method of combining support vector machine with nearest neighbors[J]. *Computer Engineering and Design*, 2009, **30**(15): 3605-3607,3611 (李蓉, 崔延美. 结合支持向量机和近邻法的太阳耀斑预报方法 [J]. 计算机工程与设计, 2009, **30**(15): 3605-3607,3611)

[13] SHAUN BLOOMFIELD D, HIGGINS P A, JAMES MCATEER R T, *et al.* Toward reliable benchmarking of solar flare forecasting methods[J]. *The Astrophysical Journal Letters*, 2012, **747**(2): L41

[14] (ZHOU Zhihua. Machine Learning[M]. Beijing: Tsinghua University Press, 2016 (周志华. 机器学习 [M]. 北京: 清华大学出版社, 2016)

[15] ROSENBLATT F. The perceptron: a probabilistic model for information storage and organization in the brain[J]. *Psychological Review*, 1958, **65**(6): 386-408

[16] BENGIO Y. Practical recommendations for gradient-based training of deep architectures[M]//MONTAVON G, ORR G B, MÜLLER K R. Neural Networks: Tricks of the Trade: 2nd ed. Berlin, Heidelberg: Springer, 2012: 437-478

[17] GOUTTE C, GAUSSIER E. A probabilistic interpretation of precision, recall and *F*-Score, with implication for evaluation[C]//Proceedings of the 27th European Conference on IR Research on Advances in Information Retrieval. Santiago de Compostela, Spain: Springer, 2005

[18] DONALDSON R, DYER R M, KRAUS M J. An objective evaluator of techniques for predicting severe weather events [C]// 9th Conference on Severe Local Storms, Texas: American Meteorological Society, 1975: 321-326

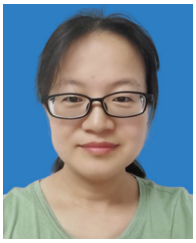
[19] HEIDKE P. Berechnung des erfolges und der güte der windstärkevorhersagen im sturmwarnungsdienst[J]. *Geografiska Annaler*, 1926, **8**(4): 301-349

[20] HANSSEN A W, KUIPERS W J A. On the relationship between the frequency of rain and various meteorological parameters[J]. *Koninklijk Nederlands Meteorologisch Instituut.*, 1965, **81**: 2-15

作者简介

李 铭 男, 1997 年 5 月出生于河北省张家口市, 现为中国科学院国家空间科学中心特别研究助理, 主要研究方向为深度学习在太阳爆发预报建模中的应用.

E-mail: liming01@nssc.ac.cn



崔延美（通信作者） 女, 1980 年 1 月出生于山东省德州市, 现为中国科学院国家空间科学中心副研究员, 硕士生导师, 主要研究方向为空间环境预报、空间灾害性天气事件建模与预报等.

E-mail: ymcui@nssc.ac.cn