

DOI:10.12403/j.1001-1498.20230195

基于深度学习的内蒙古大兴安岭林区 火灾预测建模研究

张金钰¹, 彭道黎^{1*}, 张超珺², 贺丹妮¹, 杨灿灿^{1,3}

(1. 北京林业大学森林资源和环境管理国家林业和草原局重点实验室, 北京市 100083; 2. 山西农业大学林学院, 山西 晋中 030801;
3. 滁州学院地理信息与旅游学院, 安徽 滁州 239000)

摘要: [目的] 对内蒙古大兴安岭地区的森林火灾进行预测, 为森林防火工作的开展提供重要支持。[方法] 以内蒙古大兴安岭林区为研究对象, 结合 MCD64 A1 月度火点产品、地形、气候等数据, 构建森林火灾潜在影响因子数据集, 分别利用卷积神经网络、随机森林、支持向量机模型对研究区森林火灾的发生概率进行预测与可视化, 在此基础上对模型效果进行评价并分析森林火灾空间分布特征。[结果] 大兴安岭的主要林火驱动因子按重要性值由高到低排序为海拔、平均气温、总降水量、与水域的距离等; CNN、RF、SVM 预测森林火灾发生概率的 AUC 值分别为 0.838、0.794、0.788, CNN 的精度最高; CNN 能够有效划分出森林火灾易感性极高、极低的区域, 有利于划分森林火灾的警示区。[结论] CNN 模型比 RF、SVM 模型更适用于大兴安岭林区火灾发生概率的预测; 大兴安岭林火风险的空间分布有明显的区域性, 主要发生在东南地区。

关键词: 森林; 火灾预测; 卷积神经网络; 森林火灾敏感性

中图分类号: S762.2

文献标识码: A

文章编号: 1001-1498(2024)01-0031-10

深度学习具有空间邻域性, 可以从输入的数据中发现更深层的特征, 能高效、准确地对森林火灾进行预测, 因此运用深度学习进行林火预测逐渐成为国内外研究热点。Dimuccio 等^[1] 使用地理信息系统技术, 利用地形、道路密度、人口密度等与火灾相关的 8 个因素, 使用训练好的反向传播人工神经网络频率-概率程序, 计算每个因素类别的评级并分析, 生成森林火灾敏感性指数图, 利用烧毁面积评价得到其结果一致率为 78%。Radke^[2] 基于美国落基山脉地区, 利用深度学习进行野火蔓延趋势的预测, 形成了 FireCast 模型。Zhang 等^[3] 提出了使用 Convolutional Neural Network (CNN) 的森林火灾敏感性空间预测模型, 采用受试者工作特征曲线下面积 (即 AUC) 指标, 证实了所提出的 CNN 模型 (AUC = 0.86) 比随机森林、支持向量机、多层感知器神经网络和核逻辑回归基准分类器的精度

都更高。骆开菁^[4] 以中国西南地区为研究区域, 通过整合可燃物诱发因子、气象诱发因子、地形诱发因子和森林火灾参考信息因子, 构建了完整的历史森林火灾事件及诱发因子数据库, 通过深度学习模型完成历史森林火灾风险评估, 结果表明, 构建的深度学习模型对森林火灾风险有很好的评估性能。Naderpour^[5] 提出一个空间框架来量化悉尼北部海滩地区的森林火灾风险, 具有 MLP 架构的深度学习模型提高了适应性和决策能力, 可以适应澳大利亚的不同地区, 对加权程序的本地化采用要求也很小。Prapas^[6] 在希腊每日野火风险预报的研究中, 提出了能够捕捉时空数据的深度学习模型。

大兴安岭森林资源丰富, 是我国重要的生态保护地, 也是典型的寒温带森林区域, 曾遭遇严重的森林火灾侵害, 可以作为研究寒温带森林生态系统林火的重要区域, 研究该地区林火事件能够为相似

收稿日期: 2023-05-10 修回日期: 2023-06-29

作者简介: 张金钰, 硕士。研究方向: 林业遥感与信息技术。E-mail: zjy19980406@126.com

* 通讯作者: 彭道黎, 教授, 博士。研究方向: 林业遥感与信息技术、森林资源监测与评价。E-mail: dlpeng@bjfu.edu.cn

的林火研究提供借鉴^[7]。因此开展大兴安岭的林火预测研究是十分必要的。

本研究以内蒙古大兴安岭为研究区,从 MCD64A1 月度火点数据产品获取研究区森林火灾点,构建森林火灾潜在影响因子数据集,通过建立卷积神经网络、随机森林、支持向量机模型对研究区 2018 年森林火灾的发生概率进行预测和森林火灾危险区划,并对模型效果进行评价,以期为大兴安岭森林防火工作开展提供支持。

1 研究地区概况及研究方法

1.1 研究区概况

内蒙古大兴安岭林区,在内蒙古自治区的东北方向(119°36'26"~125°24'10" E, 47°03'26"~53°20'00" N)^[8],海拔 425~1 760 m^[9],是我国 4 大国有林区之一。林区地处高纬地带,南北跨 7°,地域辽阔,是中国最大的集中连片的国有林区^[10]。内蒙古大兴安岭地区位于寒温带大陆性季风气候带,全年平均气温在-4℃至-2℃之间,平均气温小于 10℃的时间长达 9 个月。全年降水量在 300 mm 以上,相对湿度为 70%~75%,林区总面积约 1.07×10^7 hm²,森林覆盖率达 79.82%。2009—2018 年期间,研究区发生过多次森林火灾,火灾分布情况如图 1。

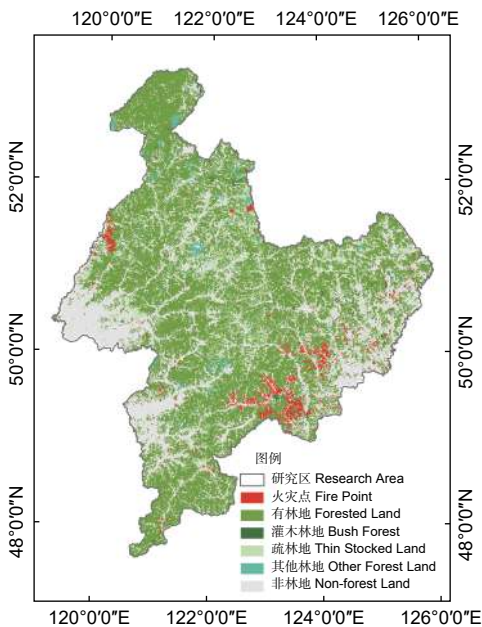


图 1 研究区及森林火灾分布(2009—2017)
Fig. 1 Research area and forest fire distribution(2009—2017)

1.2 数据来源及预处理

1.2.1 森林火灾预设影响因子理论依据 气温直接影响森林可燃物的含水量,高温且降水量少、强风且持续干旱的气候条件下森林火灾高发^[11-12]。地形条件对森林火灾的作用是间接的。地形条件影响局部小气候进而影响到森林可燃物含水率变化,并对森林火灾的发生产生影响^[13-14]。区域火灾历史数据可以反映区域的历史森林火灾状况,发生过森林火灾的地方有再发生森林火灾的潜力,因此需要关注。植被指数能够反映森林覆盖,是对地表植被状况的简单、有效和经验的度量,植被指数越高的地方,发生森林火灾的概率也相应较大。人为因素对森林火灾发生具有“增殖”效应^[15],通常考虑与道路的距离、与河流的距离等。

1.2.2 数据来源 火点数据为 MCD64A1 月火点数据产品,获取自基于云的地理空间处理平台谷歌地球引擎(Google Earth Engine, <https://earthengine.google.com/>),其空间分辨率为 500 m,时间分辨率为 8 d,时间范围为 2009—2018 年,本研究同时选取低、中、高 3 个置信等级等级的火点数据。地形数据为 ASTER GDEM 数据,空间分辨率为 30 m,下载自地理空间数据云(<https://www.gscloud.cn/>),在地理数据处理软件中进一步计算得到坡度、坡向、平面曲率数据。气象数据(平均气温、总降水量、平均风速)源自国家地球系统科学数据中心,空间分辨率为 1 km,时间分辨率为月。NDVI 来源于中国科学院资源环境科学数据中心(<https://www.resdc.cn/DataSearch.aspx>)。人为因素矢量数据(与建筑物的距离、与道路的距离、与水域的距离)获取自开源网站 OpenStreetMap(<https://www.openstreetmap.org/>),时间分辨率为 5 a。

1.2.3 数据预处理 在输入模型前,对每个森林火灾影响因子图层进行最大最小归一化处理,将图层像素点取值控制在 [0,1] 范围。在 ArcGIS 10.8 中,将论文中涉及的森林火点及火灾影响因素图层像素大小统一为 1 km × 1 km,坐标系统一为 WGS_1984_UTM_Zone_51N。最后,使用波段合成工具将 1 a 内所有的影响因子图层合并为 1 个栅格图层,最终建立 11 个多变量栅格数据集。为解决原有火点及非火点数据不平衡的问题,本研究通过创建随机点工具生成非火灾点,与已知火灾点合并生成 5 560 个训练样本(图 2)。设定

2009—2017 年所有火点的森林火灾发生概率为 1, 重复出现的火点只记 1 次, 非火点的森林火灾发生概率为 0, 作为后续模型的标签数据。

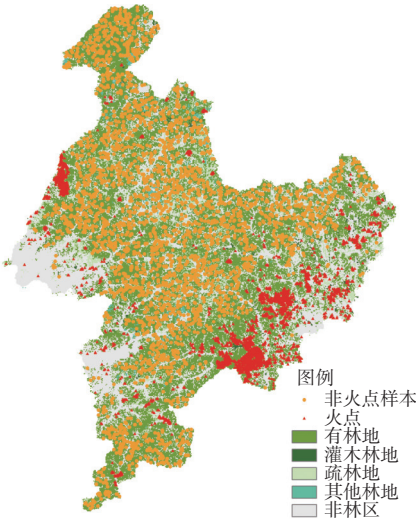


图 2 样本点分布情况

Fig. 2 The distribution of sample points

1.3 研究方法

1.3.1 特征选择方法 本研究通过计算容差值和 VIF 值来进行特征选择, 同时计算 Pearson 相关系数来度量变量之间的线性关系强度, 以及计算 Spearman 秩相关系数来度量变量之间的非线性关系 ($p \leq 0.05$)^[16]。

1.3.2 支持向量机 支持向量机 (Support Vector Machines, SVM) 有很强的数学基础和严谨的理论支持, 是常用的监督学习模型, 可运用于数据分类、回归分析和模式识别模型。支持向量机算法的最优分类面求解问题, 本质是求样本分类间隔最大的二次函数的目标解。使用网格搜索法, 最终确定 γ 为 1, 正则化参数设置为 80。

1.3.3 随机森林 随机森林 (Random Forest,

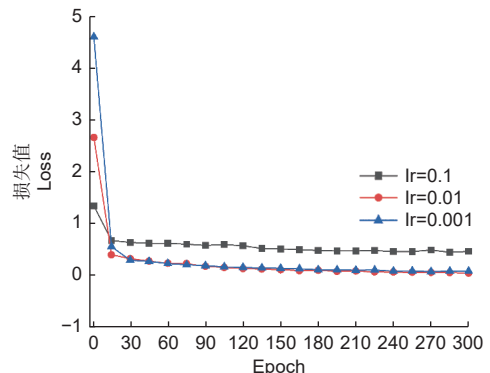
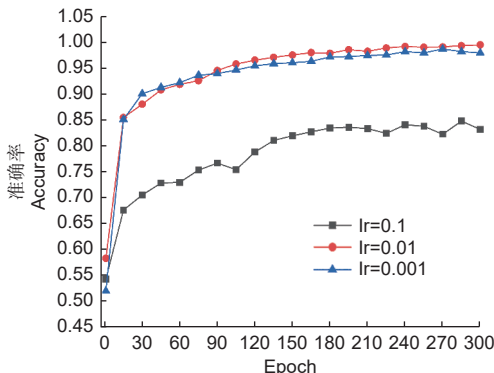
RF) 是一种基于 Bagging 算法和决策树方法的集成学习方法。RF 利用重抽样技术对训练样本进行随机采样, 训练出多个不同的决策树, 将多个决策树的分类结果进行组合, 投票或取平均得出最终的预测结果。同样使用网格搜索法, 最终确定生成树个数为 450, 树的最大深度为 18。

1.3.4 卷积神经网络 卷积神经网络 (Convolutional Neural Networks, CNN) 是一种前馈神经网络, 参数采用经典的随机梯度下降法 (Stochastic Gradient Descent, SGD) 不断调整。本研究所用的卷积神经网络 (CNN) 模型的框架是以 Alexnet 为参考, 使用 TensorFlow 作为后端。为解决特征过多而产生的过拟合问题, 加入 Dropout 算法和增加 Batch Normalization 层。

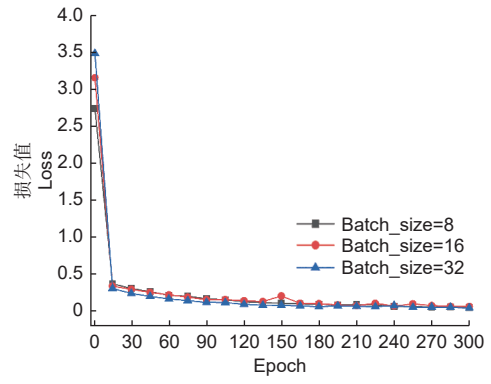
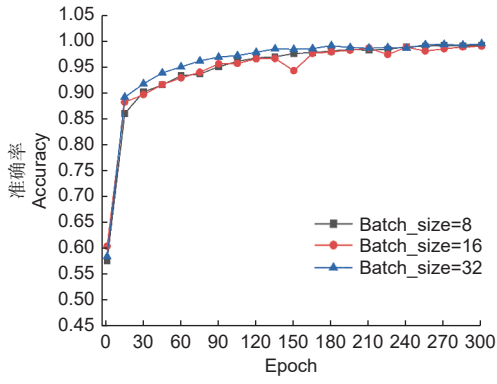
每个输入 patch 都是 1 个大小为 $15 \times 15 \times 11$ 的三维数据 (像元大小 15, 波段数为 11)。本研究的主要架构包括 3 个卷积层、3 个最大池化层和 3 个全连接层。卷积层的核数均为 64, 核大小均为 3×3 , 步长为 1, 填充方式选择“same”。每个卷积层后都有一个 ReLU 激活函数、一个 BN 层和一个最大池化层。最大池化层的核大小为 2×2 , 步长为 2, 3 个全连接层均有 32 个神经元, 在全连接层中调整权重衰减, 并在全连接层后添加 1 个双向分类器的激活函数, 即 Softmax, 可以计算出该点属于火点/非火点的概率值。参数总量为 84 706 个, 参数量设置较为合理。

本研究基于前人经验选择超参数取值并根据训练过程中的准确率和损失值图 (图 3), 确定超参数的最优取值为: 学习率为 0.001, 批处理大小为 32, Dropout 值为 0.5, 优化方法选择 RMSprop, 卷积核数量为 64。

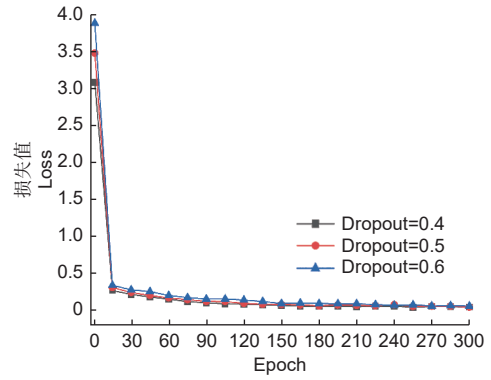
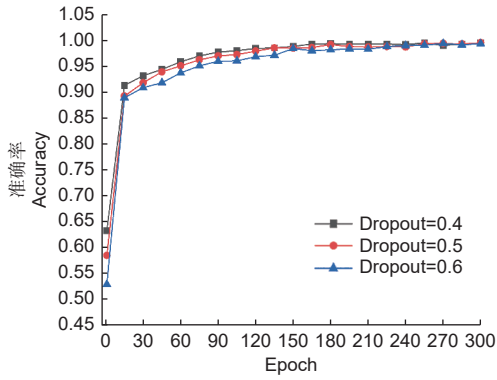
1.3.5 模型评估方法及标准 本研究采用 4 个统



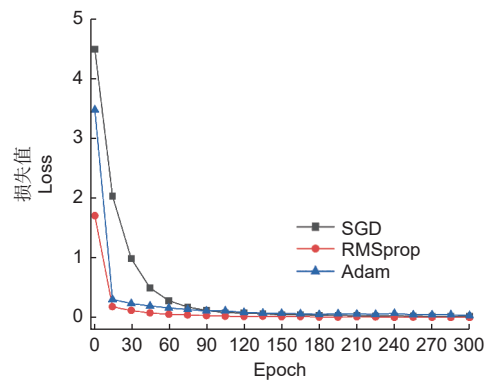
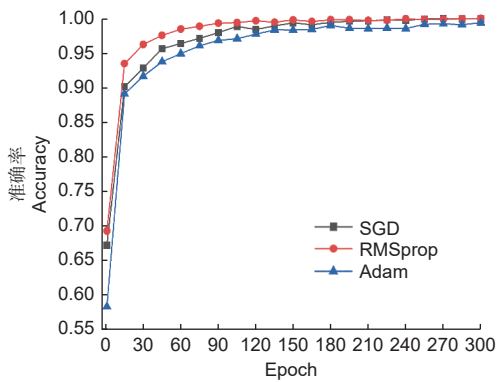
(1)



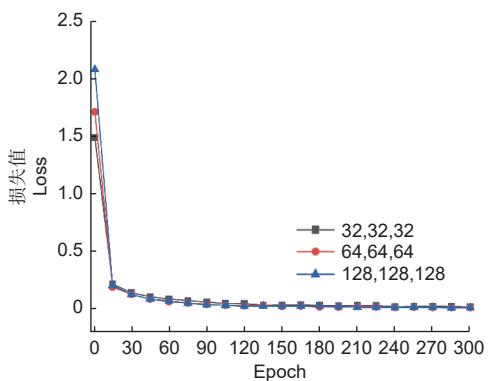
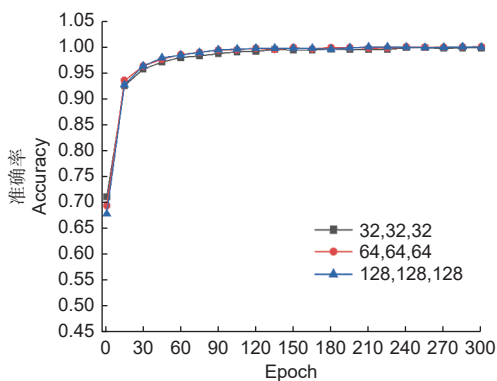
(2)



(3)



(4)



(5)

注：(1) 学习率；(2) 批处理大小；(3) Dropout；(4) 优化方法；(5) 卷积核数量

Notes: (1) Learning Rate; (2) Batch Size; (3) Dropout; (4) Optimization method; (5) Number of convolution kernels

图 3 训练中的准确率和损失值变化曲线

Fig. 3 Accuracy and loss curve during training

计指标, 包括总体准确率、特异性、召回率、精确率^[17]。这 4 个统计指标按以下方式计算:

$$\text{总体准确性} = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$\text{特异性} = \frac{TN}{FN + TN} \quad (2)$$

$$\text{召回率} = \frac{TP}{TP + FN} \quad (3)$$

$$\text{精确率} = \frac{TP}{FP + TP} \quad (4)$$

式中 TP 表示实际为火点, 预测结果也为火点, FP 表示实际为非火点, 预测结果为火点, TN 表示实际为非火点, 预测结果也为非火点, FN 表示实际为火点, 预测结果为非火点。在分类模型中, 预测结果一般以概率值输出, 通常设置阈值来将概

率转化为具体类别。因此, 阈值的设定与模型的准确率十分相关。受试者工作特征 (Receiver Operating Characteristic, ROC) 曲线和曲线下面积 (Area Under ROC Curve, AUC) 用于判断预测结果的准确性, 并表达了模型的泛化能力。

2 结果与分析

2.1 模型的变量选择

基于林火影响因子相关性热力图 (图 4) 结果表明, 坡度与粗糙度呈现明显的正相关性, 而平面曲率和剖面曲率则呈现较明显的负相关性。为了防止模型过拟合, 将影响火灾发生的因子中的粗糙度和剖面曲率剔除, 而保留坡度和平面曲率用于后续研究。

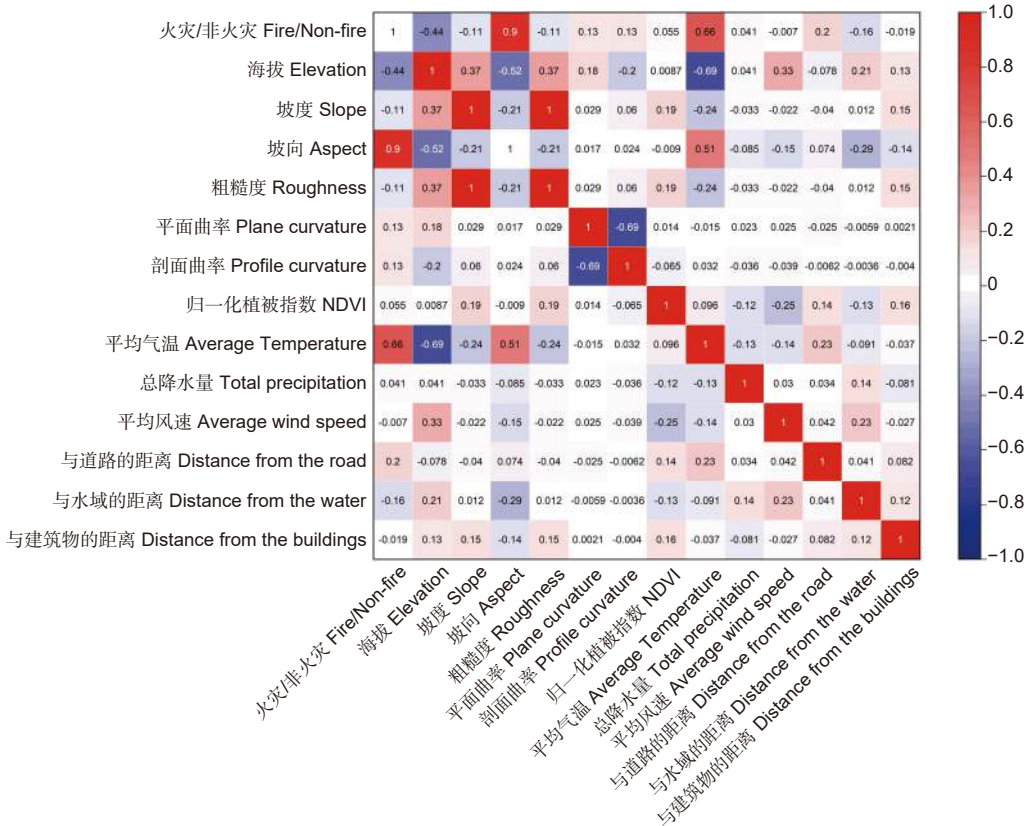


图 4 林火影响因子相关性热力图

Fig. 4 Correlation heat map of forest fire influencing factors

随机森林计算后, 得到 11 个影响林火影响因子的重要性值。气温、海拔、与水域的距离、总降水量是影响林火发生最为重要的 4 个因素。此外, 与建筑物的距离、月均风速和 NDVI 等因素也对森林火灾的发生产生了较为显著的影响, 其重要性值

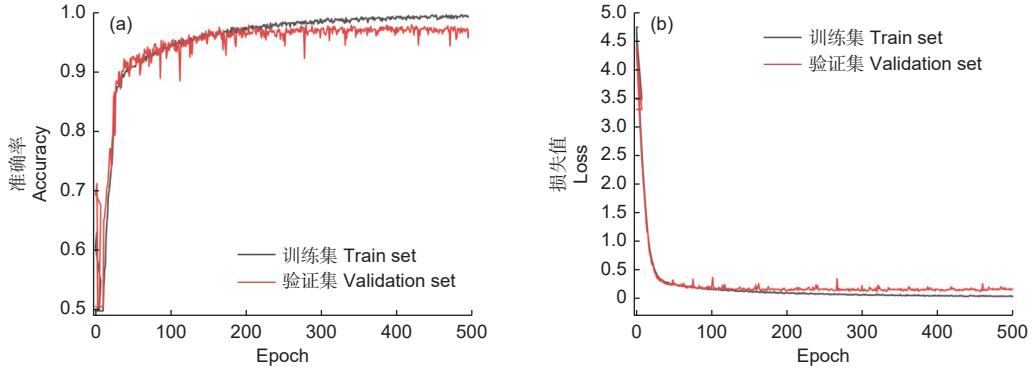
均大于 0.05。这些变量均与森林火灾有着明显的关系, 因此可以用于后续的建模研究。

2.2 卷积神经网络训练结果

训练过程包括训练和验证 2 个阶段。训练样本和验证样本分别占训练集的 80%、20%。在每个

epoch 结束后, 使用验证样本数据集对超参数进行微调。结果表明 (图 5), 每次训练后, 训练样本数据集和验证样本数据集的准确性均达到 95% 以

上, 验证集的损失值低于 0.05, 而验证集的损失值在 0.15 左右。最终, 选择验证集的损失值最小时 (epoch=300) 对应的模型作为最终的分类模型。



注: (a) 准确率; (b) 损失值

Notes: (a) accuracy; (b) loss

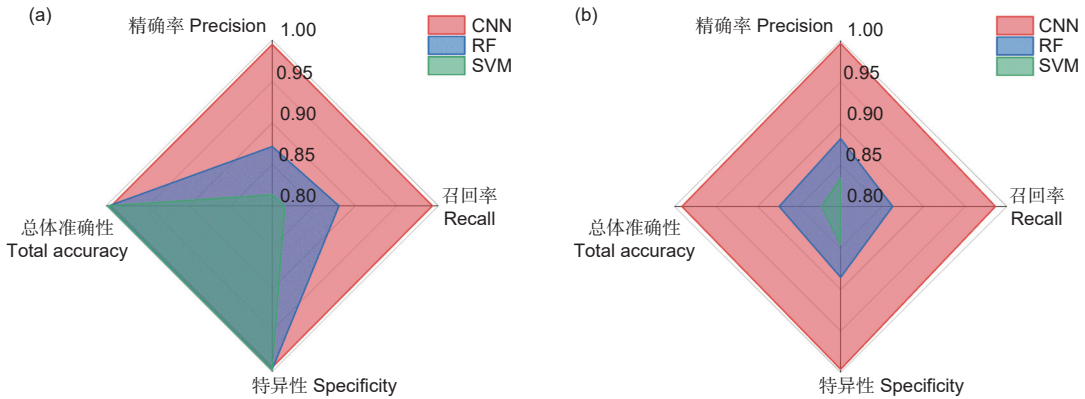
图 5 训练过程中准确率和损失值变化

Fig. 5 The change of accuracy and loss value during graph training

2.3 模型评价指标及模型比较

本研究将 CNN 与 RF 及 SVM 这两种常见的机器学习方法进行比较, 并通过绘制评价指标雷达图 (图 6) 对这 3 种方法进行比较。在预测样本

数据集中, CNN 精确率和召回率远高于 RF 及 SVM, 达到 90% 以上, CNN 特异性和总体准确率略低于 RF 及 SVM, 但可明显观察到 CNN 总体性能优于 RF 及 SVM, 故 CNN 预测结果最为可靠。



(a) 训练集; (b) 验证集

(a) Train Set; (b) Validation set

图 6 3 种模型评价指标雷达

Fig. 6 Three kinds of model evaluation index radar map

AUC 值是评价模型分类性能的一个重要指标, 用来评估模型的平均性能。其中, CNN 的 AUC 值为 0.838, 表示 CNN 对 2018 年大兴安岭林区森林火灾预测的整体拟合度达到了 83.8%, 高于 RF (0.788) 和 SVM (0.794) (图 7)。

2.4 模型预测及生成森林火灾敏感性图

经过训练, 最终得到分类模型, 使用测试数据

集 (2018 年的多变量栅格数据集) 来评估最终模型性能。因为滑动窗口在预处理阶段密集重叠并覆盖了整个栅格数据集, 故模型可以对所需预测图中的每个像素得出火灾和非火灾类别的概率, 选择火灾类别的概率作为最终的预测概率值, 可视化生成森林火灾敏感性图 (图 8)。3 种模型均识别出在大兴安岭东南方向森林火灾易感性的值较大, 中部

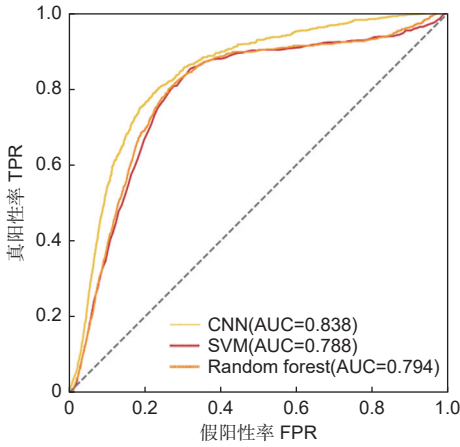


图 7 3 种模型的 ROC 曲线

Fig. 7 ROC curves of the three models

和西部位置火灾发生概率较小, 这表明大兴安岭东南地区更易发生森林火灾, 中部和西部地区较不易发生森林火灾。

通过分析图 9 可以发现, CNN 模型中极高和极低的森林火灾类别占总研究区森林面积的 90% 以上, 在 3 种模型中所占比例最高, 高、中、低 3 类分别占总面积的 2.12%、1.89% 和 2.74%。因此, CNN 模型能够有效划分出森林火灾易感性极高及极低的区域, 有利于在实践中加强对森林火灾易感区的火灾预防。相比之下, RF 和 SVM 的预测图中存在较大范围的中高易感区, 对林火高易感性区域没有明确的判定。因此, CNN 模型的结果更适合该地区森林火灾的预测。

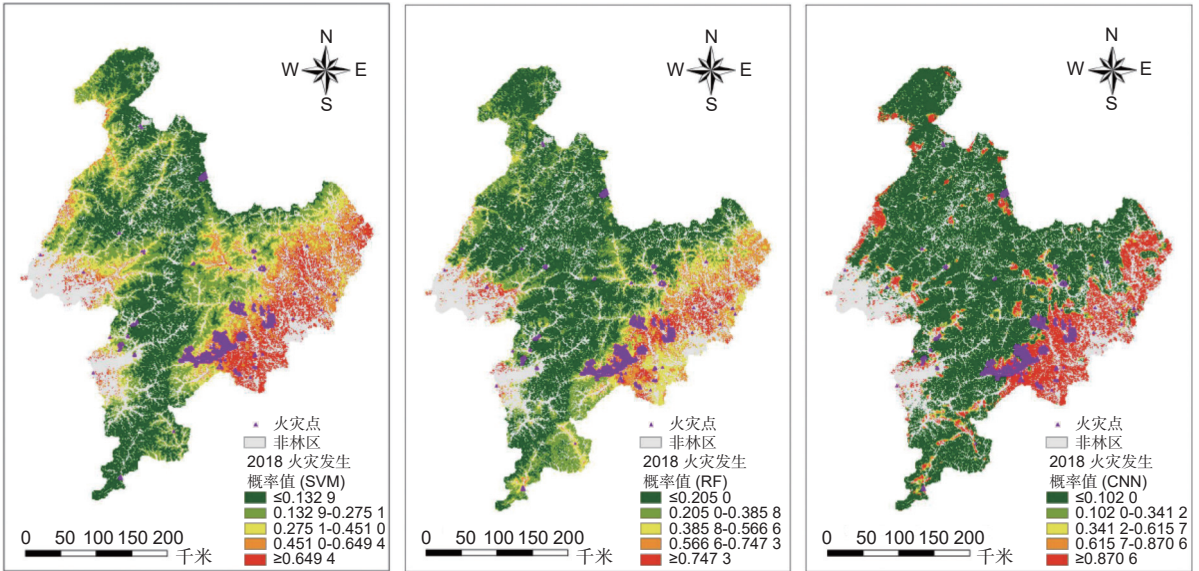


图 8 RF、SVM、CNN 模型预测 2018 年森林火灾敏感性

Fig. 8 RF, SVM and CNN model prediction of forest fire sensitivity in 2018

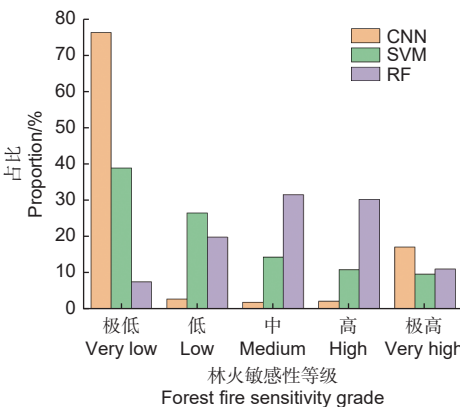


图 9 3 种模型下林火敏感性等级占比情况

Fig. 9 The proportion of forest fire sensitivity grades under the three models

3 讨论

3.1 森林火灾主要潜在影响因素

针对变量选择, 本研究将粗糙度和剖面曲率这两个因子剔除, 保留剩余的 11 个特征用于后续建模。气温、海拔、与水域的距离、总降水量是影响林火发生较为重要的 4 个因素, 这与 Najafabadi、Pourtaghi、马文苑、Nami、Abdi 等的研究结果基本一致^[18-22]。随着海拔的增加, 林火发生概率明显下降, 这是因为海拔的增加使得植被和土壤的湿度增加^[22], 且受人为干扰较少^[23], 不利于森林火灾的发生。气象条件大多是通过影响可燃物的含水量和温度对林火的发生时间和区域产生影响^[24], 气温升

高会使林区内可燃物的含水量减少,也会使可燃物本身的温度升高,减少了外部因素达到其燃点所需的热量,可燃物容易被点燃^[25-27];降水可以反映出可燃物湿度的变化情况,降水量较多时林区内可燃物含水量处于饱和状态,不易被点燃,发生森林火灾的可能性和严重程度都会较低^[26-27]。人为因素也会对森林火灾的发生有一定影响,距离人为活动的距离越远,森林火灾发生的概率也就越小^[28-29]。而坡向对森林火灾发生的解释度较小,这与 Li 等^[30]的研究结果一致。

3.2 深度学习在森林火灾预测方面的优势

本研究证实了 CNN 在森林火灾预测方面具有比 RF 和 SVM 更为突出的优势,能够有效地划分出森林火灾易感性极高和极低的区域,因此 CNN 模型的结果更加适合森林火灾预测。CNN 模型 AUC 值达到 0.8,其精度介于 Dimuccio 等^[1]、Zhang 等^[3]研究的 AUC 值之间,基本满足预测精度。在 Bisquert^[31]关于西班牙西北部加利西亚森林火灾概率的预测研究中也证实,人工神经网络的准确性和精度都高于逻辑回归。Bergado^[32]运用 29 个量化特征,采用深度学习生成 2006—2017 年澳大利亚维多利亚州未来 7 天内野火燃烧概率的每日地图,该地图更平滑、更正则化,在各项定量和定性指标中都显示出很大的优越性。Muhammad^[33]探索了 CNN 并设计了一种微调架构,用于在有效灾害管理系统的监视期间进行早期火灾检测,研究证实该框架有高精度性和低火灾误报率。总的来看,与机器学习 SVM、RF 相比,CNN 具有以下优势:首先,CNN 可以考虑相邻空间区域的相关性,因此适用于空间和地理相关的研究。其次,深度学习可以揭示更深层的特征。通过多次卷积和池化操作,CNN 可以提取更高级和抽象的特征,这些特征对于森林火灾的发生起着决定性的作用。最后,CNN 通过权值共享减少了训练过程中的超参数数量,从而降低了 CNN 的结构复杂度。

4 结论

本研究选取地形、气候、植被和人为方面共 11 个因素形成多变量栅格数据集,构建出 3 个林火敏感性模型,并将大兴安岭森林火灾的风险进行可视化展示。对大兴安岭森林火灾进行预测,有助于相关部门重点盯防,做到早预防、早发现,对大

兴安岭森林火灾预防工作有重要意义。

(1) 通过特征选择方法,最终确定 11 个森林火灾潜在影响因子,分别是地形、平面曲率、坡度、坡向、平均气温、总降水量、平均风速、与道路的距离、与水域的距离、与建筑的距离、NDVI 等。其中,气温、海拔、与水域的距离、总降水量是影响林火发生最为重要的 4 个因素。

(2) CNN 模型的 AUC 值高于 RF、SVM,说明 CNN 对林火的预测更准确,模型可靠性更高。

(3) 空间上,大兴安岭森林地区东南地区森林火险等级较高,较易发生火灾。

参考文献:

- [1] Dimuccio L A, Ferreira R, Cunha L, et al. Regional forest-fire susceptibility analysis in central Portugal using a probabilistic ratings procedure and artificial neural network weights assignment[J]. *International Journal of Wildland Fire*, 2011, 20(6): 776-791.
- [2] Radke D, Hessler A, Ellsworth D. Firecast: Leveraging deep learning to predict wildfire spread[C], *International Joint Conferences on Artificial Intelligence*, Macao, China, 2019.
- [3] Zhang G L, Wang M, Liu K. Forest fire susceptibility modeling using a convolutional neural network for Yunnan Province of China[J]. *International Journal of Disaster Risk Science*, 2019, 10(3): 386-403.
- [4] 骆开苇. 全球森林火灾风险时空挖掘及预警预测方法研究[D]. 成都: 电子科技大学, 2020.
- [5] Naderpour M, Rizeei H M, Ramezani F. Forest fire risk prediction: A spatial deep neural network-based framework[J]. *Remote Sensing*, 2021, 13(13): 2513.
- [6] Prapas I, Kondylatos S, Papoutsis I, et al. Deep Learning Methods for Daily Wildfire Danger Forecasting[J]. *arXiv e-prints*, 2021. DOI: 10.48550/arXiv.2111.02736.
- [7] 苏漳文. 基于地理信息系统的大兴安岭林火发生驱动因子及预测模型的研究[D]. 黑龙江: 东北林业大学, 2020.
- [8] 毕杰和. 内蒙古大兴安岭林区森林生态状况及保护建议[J]. *内蒙古林业调查设计*, 2018, 41(6): 6-8.
- [9] 岳永杰, 韩军, 李玉柱, 等. 内蒙古大兴安岭森林涵养水源和保育土壤功能评估[J]. *中南林业科技大学学报*, 2013, 33(12): 91-95.
- [10] 张秋良, 边玉明, 代海燕, 等. 内蒙古大兴安岭林区极端气温事件变化特征[J]. *地理科学*, 2017, 37(12): 1909-1916.
- [11] Long S, Wang Q, Wei S, et al. Response characteristics and prospect of forest fire disasters in the context of climate change in china[J]. *Journal of Catastrophology*, 2014, 29(1): 12-17.
- [12] Wang S, Li H, Niu S. Empirical research on climate warming risks for forest fires: a case study of grade I forest fire danger

- zone, Sichuan Province, China[J]. *Sustainability*, 2021, 13(14): 7773-7792.
- [13] 覃先林, 张子辉, 易浩若, 等. 一种预测森林可燃物含水率的方法[J]. *火灾科学*, 2001, 10 (3): 159-162.
- [14] Holden Z A, Jolly W M. Modeling topographic influences on fuel moisture and fire danger in complex terrain to improve wild land fire management decision support[J]. *Forest Ecology and Management*, 2011, 262(12): 2133-2141.
- [15] Mollicone D. , Eva H., Achard F. Human role in Russian wild fires[J]. *Nature*, 2006, 440: 436-437.
- [16] 方匡南, 吴见彬, 朱建平, 等. 随机森林方法研究综述[J]. *统计与信息论坛*, 2011, 26 (3): 32-38.
- [17] Bui D T, Bui Q T, Nguyen Q P, *et al.* A hybrid artificial intelligence approach using GIS-based neural-fuzzy inference system and particle swarm optimization for forest fire susceptibility modeling at a tropical area[J]. *Agricultural and Forest Meteorology*, 2017, 233: 32-44.
- [18] Najafabadi A, Gorgani F, Najafabadi M O. Modeling forest fires in Mazandaran Province, Iran[J]. *Journal of Forestry Research*, 2015, 26(4): 851-858.
- [19] Pourtaghi Z S, Pourghasemi H R, Aretano R, *et al.* Investigation of general indicators influencing on forest fire and its susceptibility modeling using different data mining techniques[J]. *Ecological Indicators*, 2016, 64: 72-84.
- [20] 马文苑, 冯仲科, 成竺欣, 等. 山西省林火驱动因子及分布格局研究[J]. *中南林业科技大学学报*, 2020, 40 (9): 57-69.
- [21] Nami M H, Jaafari A, Fallah M, *et al.* Spatial prediction of wildfire probability in the Hyrcanian ecoregion using evidential belief function model and GIS[J]. *International Journal of Environmental Science and Technology*, 2018, 15(2): 373-384.
- [22] Abdi O, Kamkar B, Shirvani Z, *et al.* Spatial-statistical analysis of factors determining forest fires: a case study from Golestan, Northeast Iran[J]. *Geomatics Natural Hazards & Risk*, 2018, 9(1): 267-280.
- [23] Vilar L, Woolford D G, Martell D L, *et al.* A model for predicting human-caused wildfire occurrence in the region of Madrid, Spain[J]. *International Journal of Wildland Fire*, 2010, 19(3): 325-337.
- [24] Guo F T, Su Z W, Wang G Y, *et al.* Understanding fire drivers and relative impacts in different Chinese forest ecosystems[J]. *Science of the Total Environment*, 2017, 605: 411-425.
- [25] Minnichl R A, Bahrez C J. Wildland fire and chaparral succession along the California Baja-California boundary[J]. *International Journal of Wildland Fire*, 1995, 5(1): 13-24.
- [26] 何 诚, 舒立福, 刘柯珍. 大兴安岭地区夏季森林火灾环境因子特征分析[J]. *西南林业大学学报(自然科学)*, 2021, 41 (3): 87-93.
- [27] 黄宝华. 山东森林火灾时空热点识别与气候因子表征关系研究[C]. *中国环境科学学会2022年科学技术年会*, 南昌, 2022.
- [28] 王一番. 长白山区森林火发生风险的多尺度气候驱动机制[D]. 长春: 东北师范大学, 2022.
- [29] Cardile J A, Ventura S J, Turner M G. Environmental and social factors influencing wildfires in the Upper Midwest, United States[J]. *Ecological Applications*, 2001, 11(1): 111-127.
- [30] Li W, Xu Q, Yi J, *et al.* Predictive Model of Spatial Scale of Forest Fire Driving Factors: A Case Study of Yunnan Province, China[J]. *Scientific Reports*, 2022. 12 (1).
- [31] Bisquert M, Caselles E, Sanchez J M, *et al.* Application of artificial neural networks and logistic regression to the prediction of forest fire danger in Galicia using MODIS data[J]. *International Journal of Wildland Fire*, 2012, 21(8): 1025-1029.
- [32] Bergado J R, Persello C, Reinke K, *et al.* Predicting wildfire burns from big geodata using deep learning[J]. *Safety Science*, 2021, 140(105): 276-287.
- [33] Muhammad K, Ahmad J, Baik S W. Early fire detection using convolutional neural networks during surveillance for effective disaster management[J]. *Neurocomputing*, 2018, 288: 30-42.

Deep Learning-based Forest Fire Prediction Model Research in the Daxing'anling Mountains, Inner Mongolia

ZHANG Jin-yu¹, PENG Dao-li¹, ZHANG Chao-jun², HE Dan-ni¹, YANG Can-can^{1,3}

(1. Key Laboratory of Forest Resources & Environmental Management, National Forestry and Grassland Administration, Beijing Forestry University, Beijing 100083, China; 2. College of Forestry, Shanxi Agricultural University, Jinzhong 030801, Shanxi, China; 3. School of Geographic Information and Tourism, Chuzhou University, Chuzhou 239000, Anhui, China)

Abstract: [Objective] To predict forest fires in the Daxing'anling Mountains of Inner Mongolia and provide important support for forest fire prevention. [Method] Based on the Daxing'anling Mountains of Inner Mongolia as the research object, combined with MCD64 A1 monthly fire point products, terrain, climate and other data, the forest fire potential impact factor data set was constructed, and the convolutional neural network, random forest and support vector machine models were used respectively to predict and visualize the probability of forest fires in the study area. The developed models were evaluated and the spatial distribution characteristics of forest fires were analyzed. [Results] The main driving factors of forest fire in the Daxing'anling Mountains were altitude, average temperature, total precipitation and the distance from water area in order of importance. The AUC value of CNN, RF and SVM was 0.838, 0.794 and 0.788, respectively, and the accuracy of CNN was the highest. CNN can effectively divide areas with high and low forest fire susceptibility, which is conducive to dividing forest fire warning areas. [Conclusion] The CNN model is more suitable for predicting the probability of forest fires in the Daxing'anling Mountains than RF and SVM models. The spatial distribution of forest fire risk in the Greater Khingan Mountains is obviously regional, mainly occurring in the southeast region.

Keywords: forest; fire warning; deep learning; sensitivity modeling

(责任编辑: 彭南轩)