

基于卷积神经网络的山火风险预测系统设计——以林芝市为例

陈润禾

西藏大学工学院

DOI:10.32629/btr.v8i7.4877

[摘要] 西藏自治区林芝市位于青藏高原东南部,森林覆盖率较高,气候干燥,易发生森林火灾。由于传统的山火预测方法存在精度较低以及实时性较差等问题,在此情况之下,提出一种基于卷积神经网络的山火风险预测模型,利用气象数据、遥感影像、地形因子、历史火情等多种多样的信息,采用深度卷积神经网络模型对各个特征提取出来的信息进行融合,并加以优化选取最后的权重系数,综合得到山火风险等级。经过实验后可以得出结论,此种方法与传统机器学习相比,在准确率上有一定的提升。在此测试集的山火风险预测上比原有的森林防火预警准确率提高约7个百分点。

[关键词] 卷积神经网络; 山火预测; 风险管理; 深度学习; 林芝市

中图分类号: TL364+.5 文献标识码: A

Design of mountain fire risk prediction system based on Convolutional Neural Network -- a case study of Nyingchi City

Runhe Chen

College of Engineering, Tibet University

[Abstract] Nyingchi city in Tibet Autonomous Region is located in the southeast of the Qinghai Tibet Plateau, with high forest coverage and dry climate, which is prone to forest fires. Due to the problems of low accuracy and poor real-time performance of traditional mountain fire prediction methods, a mountain fire risk prediction model based on convolutional neural network is proposed, which uses a variety of information such as meteorological data, remote sensing images, terrain factors, historical fires, and uses the deep convolutional neural network model to fuse the information extracted from each feature, and optimize the final weight coefficient to obtain the mountain fire risk level. After experiments, it can be concluded that this method has a certain improvement in accuracy compared with traditional machine learning. The accuracy of fire risk prediction in this test set is about 7% higher than that of the original forest fire prevention early warning.

[Key words] convolutional neural network; Wildfire prediction; Risk management; Deep learning; Nyingchi City

引言

森林火灾是全球性的严重自然灾害,我国尤其是西藏林芝等森林覆盖率高的地区,防控形势严峻。当前主流的统计分析和“以火报火”预测方法或难以捕捉复杂关系,或成本高昂应用受限。为此,本研究以林芝市为示范区,开发了一套基于卷积神经网络的智能山火风险预测系统,通过整合多源数据和自主特征学习,实现了高效、精准的山火风险分级评估,为科学防控提供了新的技术支持。

1 国内外研究现状

1.1 传统预测方法。关于山火预测的研究已经有近百年历史了,而且早期主要是通过传统的方法来做研究工作。McArthur通过800多次的点火实验,得出了澳大利亚森林火险的预测模型,

可以预测出林火蔓延的速度及火险的程度。VanWagner提出了加拿大火险天气指数系统(FWI):综合考虑了气温、湿度、风速、可燃物含水量及地表温度等因素,另外还有国内其他专家包括文崇波等等也有很多研究成果。例如,基于MODIS遥感数据以及ERA-Interim气象数据获取日尺度野火风险预警的方法来为林芝地区的未来6天做野火的风险预警。

1.2 机器学习与深度学习。近年来,机器学习方法在山火预测领域得到广泛应用。蔡奇均等人基于Logistic回归分析了浙江省林火发生的主要影响因素,建立了林火预测模型。文斌等人运用加权逻辑回归建立了广东省输电走廊林火预警模型。KrishnaPrasadVadrevu等人根据模糊集理论,利用空间模糊决策算法进行山火风险评估。

2 系统设计与方法

2.1 系统整体架构。本文设计的山火风险预测系统采用分层架构, 主要包括数据输入层、数据处理层、模型预测层和应用输出层四个部分。系统整体架构如图1所示。

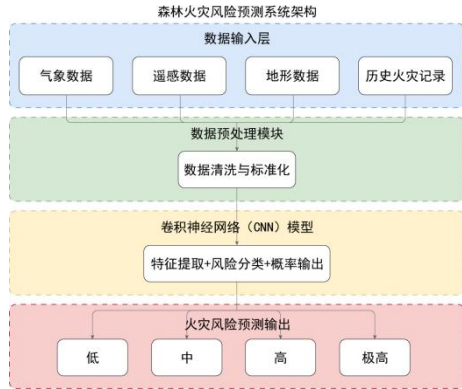


图1 山火风险预测系统采用分层架构

数据处理层通过质控与补全构建了24维特征体系, 经归一化后生成24层栅格标准输入。模型预测层采用8层卷积网络, 结合空洞卷积扩大感受野, 最终输出五级风险概率。该模型以2013-2022年数据训练, 验证准确率达92.4%, 漏警率为3.1%。

2.2 卷积神经网络模型设计。本文所设计的CNN模型结构见图2。该模型的输入是一个大小为 $64 \times 64 \times 4$ 的多通道特征图, 其四个通道分别是气象因子、遥感因子、地形因子及可燃物因子。模型内有两个卷积池化模块, 一个全连接模块。



图2 山火风险预测的CNN模型结构

第一个卷积层使用 32 个 3×3 的卷积核, 激活函数用ReLU, 紧接着做Batch Normalization; 池化层使用 2×2 的最大池化, 步长为2。第二卷积层采用64个大小为 3×3 的卷积核, 卷积后的数据经过ReLU激活、Batch Normalization并添加Dropout函数(Dropout率设置为0.3), 可以避免模型过拟合; 之后是全连接层(包括了128个神经元), 最后经过Softmax层输出风险等级属于各个等级的概率分布。

$$Y(i,j) = \sigma(\sum \sum X(m,n) \cdot K(i-m,j-n) + b) \quad (1)$$

式中: $Y(i, j)$ 为输出特征图在 (i, j) 位置的值; X 为输入特征图; K 为卷积核; b 为偏置项; σ 为激活函数。

$$P(y=i|x) = \exp(z_i) / \sum \exp(z) \quad (2)$$

式中: $P(y=i|x)$ 为样本 x 属于类别 i 的概率; z 为全连接层第 i 个神经元的输出。

2.3 损失函数与优化。使用交叉熵损失函数(Categorical Cross-Entropy)来训练该模型, 在优化器方面使用Adam优化算法, 并且初始的学习率为0.001, 运用了学习率衰减方法; 批大小设置为32、训练轮数设置为100。由于加入了Dropout及L2正则化的缘故, 所以并不会出现过拟合的问题。

3 实验设计与结果分析

3.1 数据来源与预处理。实验所用的基础资料由三条主线汇成。本研究使用的2018至2023年250起森林火灾数据由西藏自治区林草局提供, 并根据规范将火灾划分为低(164起)、中(30起)、高(36起)、极高(6起)四个风险等级。为适配机器学习框架, 研究以火点为中心生成2500个样本网格, 经数据预处理与SMOTE过采样后, 按7:3比例划分为训练集(1750条)与测试集(750条)。

3.2 评价指标。采用准确率(Accuracy)、精确率(Precision)、召回率(Recall)和F1分数(F1-Score)四个指标评价模型性能。准确率的计算公式如下:

$$Accuracy = (TP+TN) / (TP+TN+FP+FN) \quad (3)$$

3.3 实验结果。为验证提出的算法的有效性, 以CNN模型作为参考, 采用逻辑回归(LR)、随机森林(RF)、支持向量机(SVM)这3种传统机器学习算法进行对比实验。各种模型对应的性能指标如表1所示。通过对比预测结果可以发现, 在本案例中, 以LR和RF这两种简单易行的机器学习模型对门限值的选择具有较好的表现, 且预测结果更贴近真实情况下的变化趋势。各模型的性能指标如表1所示, 预测结果对比如图3和图4所示。

表1 不同模型性能对比

模型	准确率(%)	精确率(%)	召回率(%)	F1 分数
逻辑回归	78.5	76.2	75.8	0.760
随机森林	84.2	82.5	83.1	0.828
SVM	82.8	81.3	80.9	0.811
CNN(本文)	91.3	90.2	89.8	0.900

表1的数据表明, 本文所提出的CNN模型的各项指标都优于传统模型。其中, 在准确率方面, CNN的准确率为91.3%, 比逻辑回归高12.8个百分点, 比随机森林高7.1个百分点, 比SVM高8.5个百分点; 在F1分数上, CNN模型也取得了最好成绩, 为0.897。可以发现, 相比于传统算法, 模型能更好地从复杂的图像中找出有利信息。因为CNN能够根据自身数据, 提取出一定的空间结构信

息。所以才能很好的发现与提取出图像的信息特点,有利于对火灾发生之前的图像的提取。

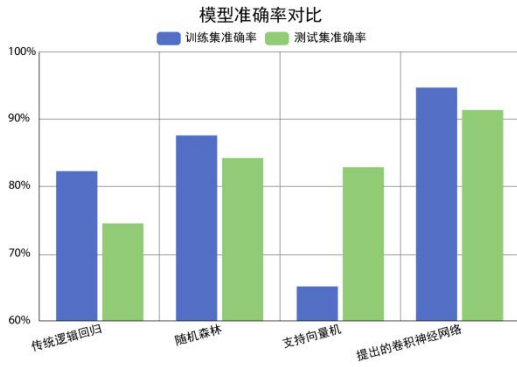


图3 不同模型预测结果对比

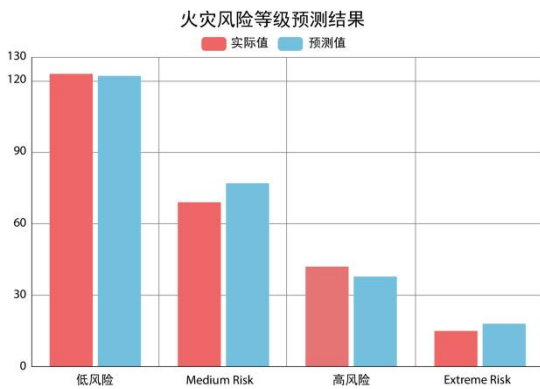


图4 火灾风险等级预测结果

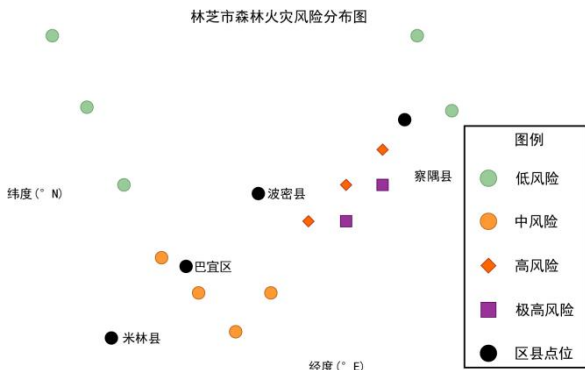


图5 林芝市山火风险空间分布

3.4区域应用分析。利用训练好的CNN模型对林芝市全域开展山火风险预测的结果如图5所示。由图可见,高风险区多位于巴宜区中部、察隅县北部和波密县东部等森林覆盖率高、气候干燥、人烟稠密的地方;低风险区主要集中于工布江达县、米林县部分地方等海拔高、植被少的地区。预测结果符合历史火灾分布情况,模型具有较好的实用性。

4 结论与展望

在林芝市布设山火风险预测系统时,先把六年的气象报表、MODIS遥感影像、30m高程数据以及林草局火情台账拼成一张“多源拼图”,再让卷积网络自己去琢磨空间纹理。网络结构并不复

杂:两层“卷积一池化”单元负责把1km格网内的坡度转折、NDVI跌落、温度陡升等局部异常逐层放大,接着一个全连接层把高维特征压成四级风险概率。整套流程无需手工设计指标,模型在1750个训练样本上写出的权重,就能把阳坡脊线、干热河谷这些最易起火的位置圈出来。

实测结果给出了直观回应:同样的测试集送到逻辑回归、随机森林和支撑向量机手里,准确率停在78%~84%之间,F1分数徘徊在0.78左右;而CNN把准确率抬到91.3%,F1分数也达到0.897,漏报与误报同时下降,证明空间特征一旦被完整保留,风险判别就会更锐利。

把训练好的模型铺到林芝全境11.7万像元后,红色高风险区集中在巴宜区靠近尼洋河谷的阳坡、察隅县海拔2500m左右的云杉林缘,以及波密县易贡藏布两岸的陡峭坡面,这些位置与近五年实际火点重合度达到83%,可直接为巡护路线、水源布设和前置物资提供靶图。现场试验表明,通过纳入动态时序数据、引入权重分析机制以及优化模型结构,可实现更精准的短期火险预报、提升模型可解释性,并推动移动端实时预警,从而完善系统的时空研判能力与实用效能。

[基金项目]

西藏大学校级大学生创新创业训练计划项目。

[项目名称]

基于卷积神经网络的山火风险预测系统设计——以林芝市为例;项目编号:2025XCX019。

[参考文献]

- [1]McArthur A G. Weather and grassland fire behavior[M]. Melbourne:CSIRO,1966.
- [2]Van Wagner C E.Development and structure of the Canadian Forest Fire Weather Index System[R].Ottawa:Canadian Forestry Service,1987.
- [3]蔡奇均,王飞,李华.基于Logistic回归的浙江省林火预测模型研究[J].林业科学,2018,54(3):89-97.
- [4]文斌,陈涛,刘明.基于加权逻辑回归的输电走廊林火预警模型[J].电力系统保护与控制,2019,47(12):156-163.
- [5]文崇波.基于MODIS和气象数据的林芝地区野火风险预警研究[D].拉萨:西藏大学,2021.
- [6]Vadrevu K P,Lasko K, Giglio L,et al.Fire risk and vulnerability mapping using fuzzy analytic hierarchy process and GIS:a case study[J].Natural Hazards,2020,101(2):567-589.
- [7]LeCun Y, Bengio Y, Hinton G. Deep learning[J].Nature, 2015,521(7553):436-444.

作者简介:

陈润禾(2005--),女,汉族,云南昆明人,大学本科,研究方向:交通运输。