基于机器学习的高植被覆盖区降雨滑坡识别

亓莉宁[1,2,3,4]，郭朝旭[2,3]，孟华君[1,5*]，楚克磊[2,3,6]，王园园[1,4]

1. 中国地质科学院地质力学研究所，北京 100081；
2. 自然资源部丘陵山地地质灾害防治重点实验室，福州 350002；
3. 福建省地质灾害重点实验室，福州 350002；
4. 中国地质大学（北京）地球科学与资源学院，北京 100083；
5. 自然资源部陕西宝鸡地质灾害野外科学观测研究站，陕西 宝鸡 721000；
6. 中国地质科学院，北京 100037

第一作者，E-mail：qilining19990228@163.com

* 通信作者，E-mail：mhjun-521@foxmail.com / mhjun521@cags.ac.cn

2024-06-03 收稿；2024-10-18 修回；2024-10-21 接受

基金项目：福建省自然资源科技创新项目(KY-07000-04-2022-023)；国家自然科学重点基金项目(42130720)；中国地质科学院地质力学研究所项目(所科研 87)

中图分类号：TP181；P642.22　　文章编号：2096-8523(2026)01-0185-14　　doi: 10.19509/j.cnki.dzkq.tb20240298

摘　要：近年来，受全球气候变暖影响，我国东南沿海省份频繁遭受超强台风暴雨和梅雨侵袭，诱发了大量群发性浅层滑坡灾害，严重威胁当地人民群众的生命和财产安全。为第一时间获取较准确灾害信息支撑应急救灾科学决策，针对高植被覆盖区强降雨诱发花岗岩强风化层形成的群发性浅层滑坡，采用机器学习方法提出了一种基于决策树轻量级梯度提升机(light gradient boosting machine，简称 LightGBM)算法的浅层滑坡自动识别模型，并利用福建省武平县高精度航拍影像和数字正射影像(DOM)数据选取坡度、高程、平面曲率(*SOA*)、剖面曲率(*SOS*)和 R、G、B 三波段等特征参数搭建识别模型进行训练和测试。结果表明：①多特征组合模型的滑坡识别准确率优于单特征模型；②由坡度、高程、平面曲率、剖面曲率和 R、G、B 三波段组成的快速识别模型能准确识别已发育滑坡体，准确率(*Accuracy*)和召回率(*Recall*)值均大于 0.99；③去除可视化图层后，采用该模型开展高植被覆盖区滑坡早期识别时准确率(*Accuracy*)和召回率(*Recall*)值仍达 0.86；推广至武平县全县范围时，准确率(*Accuracy*)达 0.80，表明识别准确度较高；④识别模型总体得分(*F*1*score*)高达 0.86，推广至武平县全县范围时，*F*1*score* 高达 0.84，表明该模型泛化能力较强，能够较准确对相似地质环境孕育的滑坡灾害实现自动识别。研究成果为高植被区降雨型滑坡灾害早期识别提供了一种新的思路，可为海西台风暴雨型滑坡地质灾害风险防控和救灾决策提供依据和科技支撑。

关键词：降雨诱发型；滑坡；机器学习；高植被覆盖区

近年来，受全球气候变暖影响，我国极端强降雨天气多发、频发，极大增加了降雨滑坡发生的概率[1-2]。如 2021 年河南郑州“7·20”特大暴雨[3]、2021 年山西晋中“21·10”特大暴雨[4]、2022 年湖南麻阳“6·2”特大暴雨、2023 年“23·7”京津冀特大暴雨[5]、2023 年福建“杜苏芮”“海葵”超强台风暴雨等事件都在高植被覆盖区诱发了大量山体滑坡(或滑坡-泥石流)，造成了较大财产损失。这种由超强降雨诱发的滑坡(包括滑坡–泥石流)往往呈“规模小、数量多、密度大、滑面浅、流态化”的特点，滑源区植被覆盖率较高又增加了灾害的隐蔽性，导致灾前难以进行有效预判，且会对灾后救援造成较大困难。因此，如

著录信息：亓莉宁，郭朝旭，孟华君，等. 基于机器学习的高植被覆盖区降雨滑坡识别 [J]. 地质科技通报，2026，45(1)：185-198.
QI Lining，GUO Chaoxu，MENG Huajun，et al. Recognition of rainfall-induced landslide in high vegetation coverage area based on machine learning[J]. Bulletin of Geological Science and Technology，2026，45(1)：185-198.

© Editorial Office of Bulletin of Geological Science and Technology. This is an open access article under the CC BY-NC-ND license.　https://dzkjtb.cug.edu.cn

何实现这类滑坡的灾前早期识别、灾后精准评估是第一时间开展科学救援的关键。

当前，识别滑坡主要是通过合成孔径雷达干涉测量(interferometric synthetic aperture radar，简称 InSAR)技术和光学遥感识别变形区实现的[6]，然后结合地质调查复核、筛查进行综合判识[5]。但 InSAR 技术易受观测角度、植被覆盖、水汽等因素制约，仍难以在高植被覆盖区适用[7-9]；而基于光学遥感的人机视觉解译又耗时费力，且存在输出结果不确定、主观性强、过度依赖专家经验等问题[10]，往往无法第一时间给出准确的灾害体及危险区信息，更难以对强降雨诱发型滑坡灾害进行早期识别。因此，要开展高植被覆盖区的降雨型滑坡准确识别、甚至早期识别，亟需引入新的技术手段和方法。机器学习的快速发展为滑坡灾害自动化、智能化识别提供了新的思路[11]。机器学习具有强大的非线性处理能力和鲁棒性，与地理信息系统融合后，能够基于贝叶斯[12-13]、逻辑回归[14-15]、支持向量机[16-17]、随机森林[18]、人工神经网络[19-21] 等算法开展地质体自动识别，展现出了巨大潜力[22]。如冀全伟等[23] 提出了一种基于梯度提升决策树算法的岩性单元填图方法，识别准确性高达 87%。滑坡作为一种特殊的地质体，在遥感影像上具有典型的纹理特征和形态特征[22]，具备采用机器学习方法识别的可能性。国内外一些学者也在这方面进行了尝试，如 DING 等[24] 采用卷积神经网络(convolutional neural network，简称 CNN)算法开展了深圳地区滑坡的自动识别，WANG 等[25] 采用逻辑回归(logistic regression，简称 LR)、支持向量机(support vector machine，简称 SVM)、随机森林(random forest，简称 RF)、Boosting 方法和卷积神经网络 5 种算法开展了香港大屿山滑坡识别研究，KEYPORT 等[26] 基于影像像素分类和面向对象分类 2 种方法对比了危地马拉 San Juan La Laguna 山体滑坡的识别效率，GHORBANZADEH 等[27] 评价了不同机器学习方法的识别效果等。这些工作都为本研究开展高植被覆盖区强降雨诱发滑坡的自动化识别提供了宝贵的理论、方法和经验参考，但如何提升高植被覆盖区强降雨诱发型滑坡早期识别的准确性仍是业内较大的难题。

鉴于此，本研究拟选择东南台风暴雨区的福建省武平县为研究区，尝试基于机器学习融合梯度提升决策树算法开展该地区降雨滑坡的早期识别研究，为我国汛期地质灾害防御和防灾减灾工作提供科技支撑。

1　研究区概况

研究区选在我国东南台风暴雨区的福建省龙岩市武平县(图 1)，位于闽、粤、赣三省交界处，区域地理坐标为 115°51′−116°23′E，24°47′−25°29′N。全县东西最宽处约 54.5 km，南北最长处约 77 km，国土面积约 2 630 km^2。

从地形地貌看，武平县北部中低山纵横，最高海拔 1 538.4 m，主要由武夷山脉进入武平县境内形成的大致呈 NE-SW 走向的东、西 2 列山系组成，其中西列属武夷山脉南段主体，为闽赣两省水系分水岭，东列为武夷山脉南段东列山地；武平县南部则低山丘岭起伏，多低丘河谷。总体来看，地形复杂，海拔高度相差很大；全县地势由西北向东南倾斜，北高南低，是典型的中低山丘陵地貌。

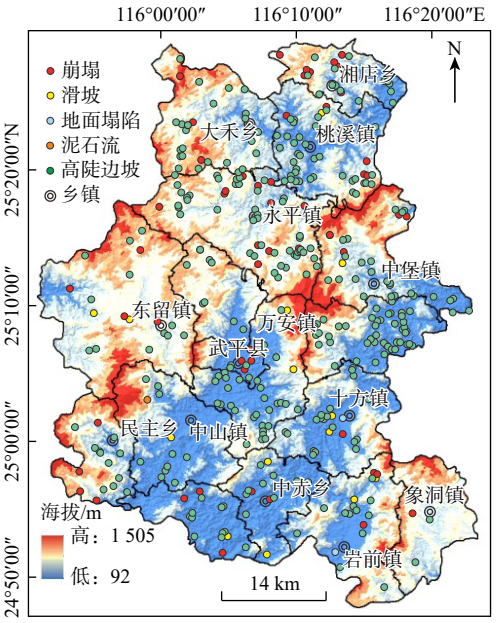


图 1　研究区地理位置

Fig. 1　Geographical location of the study area

从地质条件看，武平县地处于南平−宁化构造−岩浆带南端、政和−大埔断裂西侧的闽西南拗陷带，区内有光泽−武平断裂、云宵−上杭断裂等几条中小型断裂经过，构造环境较为复杂。境内地层发育齐全，从新元古界震旦系到新生界第四系均有出露，主要出露地层为上侏罗统的花岗斑岩、二长花岗岩、黑云母花岗岩，占全县面积的 35%，其次为震旦系楼子坝组，占全县面积的 27.5%，主要分布灰绿变质砂岩、粉砂岩、板岩、千枚岩、石英片岩；白垩系沙县组占 3.5%，为一套紫红色中厚层砾岩、砂砾岩、泥质粉砂岩等；其它地层占比均不超过 3.5%。总体来看，

武平县复杂的地理、地形以及地质构造构成了有利于滑坡孕育的基础环境条件。

从诱发条件看，武平县属中亚热带季风气候区，雨量充沛且相对集中，多发生在 4−8 月。根据武平县 2010−2019 年降雨量资料，全县年降雨量在 1 380.5 ~ 2 841.0 mm 之间，多年平均降雨量 1 768.6 mm。其中 5−6 月是每年的梅雨时节，期间日照时间少、高湿多云且风力较小，常发生连续降雨或暴雨，平均累计降雨量达 570 mm，具有“雨区广、强度大、雨时长”的特点，历史最大日降雨量达 242.00 mm。7−8 月多台风暴雨，多年平均降雨量 393.06 mm，雨量大且集中，降雨强度大。这种连续降雨或暴雨是该地区滑坡灾害的主要诱发因素。

根据 2020 年武平县 1∶5 万地质灾害调查数据[28]，武平县发育滑坡、崩塌、泥石流和地面塌陷 4 种地质灾害共 106 处。其中，滑坡是武平县最主要的地质灾害类型，共 69 处，主要为坡体表层松散的块石土、碎石土、含砾黏性土（强风化层）等物质在连续降雨或短历时强降雨作用下发生滑塌、滑移、滑溜等，滑体厚度多介于 1.5 ~ 4 m 之间，结构较松散、泥化程度高，属于典型的降雨诱发型浅表层碎石土滑坡。2022 年 5 月 26 日−2022 年 5 月 27 日，武平县再次遭遇极端强降雨天气，导致多个乡镇、9 700 名群众不同程度受灾，全县范围出现 272 处崩滑灾害[29]，其中，十方镇 24 h 降雨量达 221.3 mm，过程雨量、小时雨量均超历史极值[30]，出现群发性降雨滑坡灾害，这些滑坡即为笔者的研究对象。

2 研究方法

基于通过已知事件预测未知事件的基本思路，本研究分两步实现降雨诱发型滑坡的早期识别。首先基于已发生滑坡的数据信息训练滑坡快速识别模型，达到模型能够精准识别其他已发生滑坡灾害的目的；在此基础上去掉可视化因子图层，即可获取到除影像纹理、光谱等信息之外能够影响滑坡发育的其他信息，基于这些信息构建滑坡早期识别模型。

2.1 模型选择

鉴于本研究中样本数据有限，选择集成学习方法开展滑坡识别研究。集成学习（ensemble learning）是按照某种集成策略通过组合多个基学习器进行学习任务（图 2），从而得到比单一学习器性能表现更佳的强学习器[31]的机器学习方法，其核心步骤是“采样−学习−组合”形成最优模型。基于学习过程中不同的“采样”理念，逐渐形成了 Bagging[32] 和 Boosting 2 种集成学习算法。其中，Bagging 关注的是多个基学习器的投票组合，由多个基学习器投票产生最终的强学习器[33]，为了降低偏差保证模型的稳定，Bagging 中的每一个基学习器要相对复杂一些，也因此会导致计算量较大。Boosting 算法第一个分类器之后的每一轮学习都是对上一轮中未被正确分类的样本进行学习，因此每一次学习都减少了上一轮的偏差，即可通过增加迭代次数，逐渐得到一个最优的强学习器。随着 2 种学习算法的不断发展，又进一步细分出了更深层次的算法。结合研究区的数据特点，选择 Boosting 算法中的 LightGBM 模型进行滑坡识别。

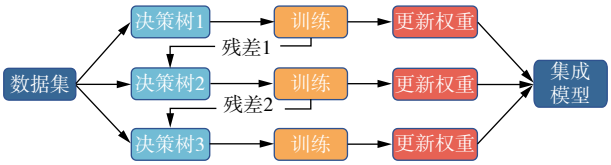


图 2 集成学习模型架构

Fig. 2 Architecture of the ensemble learning model

LightGBM 算法是在梯度提升决策树（gradient boosting decision tree，简称 GBDT）算法的基础上改进而来。GBDT 是 Boosting 算法中的一种[34]，是通过迭代训练基学习器来得到最终的强学习器，即通过损失函数的负梯度去拟合前一轮学习的基学习器残差，这样可以逐渐减小每一轮学习的残差，保证损失函数通过每一轮的训练都在尽可能减小，加速生成最优强学习器，具有训练效果好、不易过拟合的优点。但 GBDT 训练过程每一次迭代都需要多次遍历整个训练数据，对硬件内存要求极高，且反复读写训练数据要耗费很长时间。在不降低结果准确率的情况下，为了减小对内存的消耗，提升训练速度，对传统的 GBDT 算法进行改进，即可得到 LightGBM 算法。与 GBDT 其他的算法不同，LightGBM 采用的是 Leaf-wise 生长策略（图 3），即在每一层中找到分裂效益最大的叶子，进行分裂，并循环重复这个过程，具有生长策略效率更高、误差更小的特点。

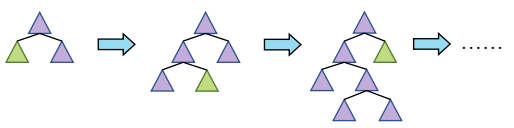


图 3 Leaf-wise 生长策略示意图

Fig. 3 Schematic diagram of the Leaf-wise growth strategy

2.2 方法流程

本研究拟基于 LightGBM 算法的二分类筛选出

针对研究区滑坡识别效果最好的特征因子组合进而构建滑坡快速识别模型。主要包括数据准备、特征选取、特征优化和模型构建几个方面。

（1）数据准备　首先对训练区的所有滑坡进行人机目视解译，并选取训练区的部分滑坡作为数据样本，制作滑坡和非滑坡标签。基于数字高程模型(digital elevation model，简称 DEM)数据，构建研究区的坡度、平面曲率和剖面曲率特征图层。

（2）LightGBM 模型构建　利用无人机影像的 R、G、B 三波段作为特征，输入标签和训练区影像进行 LightGBM 模型的构建和测试。

（3）特征优选与模型构建　通过将不同的特征组合代入已构建好的 LightGBM 模型，得出不同的滑坡识别结果；通过迭代优化不同特征的权重，选取该地区表现最好的特征组合完成滑坡识别模型的构建。

（4）结果评价　通过准确率(*Accuracy*)、精确率(*Precision*)、召回率(*Recall*)以及 *F*1*score* 等经典指标对识别结果及模型进行精度评价。各指标的计算方法及代表意义如下列公式所示。

$$P = TP + FN, \tag{1}$$

$$N = TN + FP, \tag{2}$$

$$Accuracy = \frac{TP + TN}{P + N}, \tag{3}$$

$$Precision = \frac{TP}{TP + FP}, \tag{4}$$

$$Recall = \frac{TP}{TP + FN}, \tag{5}$$

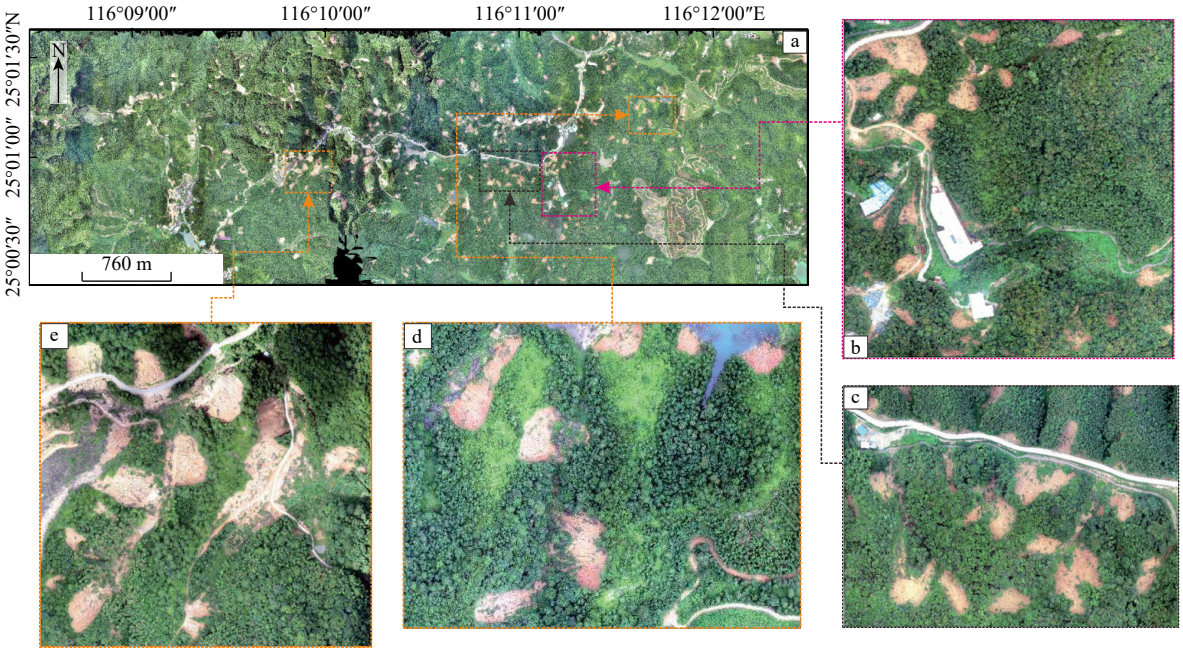
$$F1score = \frac{2 \times Rrecision \times Recall}{Precision + Recall}, \tag{6}$$

式中：*P* 为正样本数量；*N* 为负样本数量；*TP* 为实际为正被预测为正的样本数量；*FP* 为实际为负但被预测为正的样本数量；*FN* 为实际为正但被预测为负的样本数量；*TN* 为实际为负被预测为负的样本数量；*Accuracy* 为准确率，分类正确的样本占总样本的比例；*Precision* 为精确率，模型预测为正的样本中实际为正的样本占被预测为正的样本的比例；*Recall* 为召回率，实际为正的样本中被预测为正的样本占实际为正的样本的比例；*F*1*score* 为总体得分，精确率和召回率的调和平均值，表示模型的稳健。

2.3　实验数据及特征选取

（1）实验数据

2022 年，福建省武平县十方镇发生强降雨诱发滑坡群发事件。2023 年 7 月，采用大疆精灵 4 无人机在十方镇滑坡群发区域开展了低空摄影航拍，获取了十方镇主要滑坡分布区的航拍影像和数字正射影像图(digital orthophoto map，简称 DOM)数据，水平和垂直分辨率均为 96 dpi，构成本研究机器学习的基础数据(图 4a)，影像精度高于 0.1 m，包含 R、G、



a.十方镇滑坡群发区航拍影像；b. 快速识别模型训练区；c. 快速识别模型测试区；d, e. 早期识别模型测试区

图 4　滑坡航拍影像、训练区与测试区

Fig. 4　Aerial image, training area and test area for landslide

B 三波段属性。各种地物在影像上具有典型的色彩特征，容易分辨。其中，植被茂盛呈浅绿色或深绿色，建筑物为亮白色或浅蓝色，滑坡灾害则表现为裸露的土黄色，且大多滑坡在影像上形态完整。辅以 DOM 数据，可以获取滑坡滑源区的坡度、高程等地质环境信息。图 4a 为航拍范围整体影像，图 4b 为滑坡识别模型训练区，即提取图 4b 区域中滑坡和非滑坡地物的各种信息供模型进行学习；图 4c 为滑坡识别模型测试区，即对训练出来的识别模型进行识别效果验证，进而评价模型的性能和准确性；图 4d, e 则是对新构建的早期识别模型进行识别效果测试。

（2）特征选取

根据地质灾害调查报告和相关文献 [28]，结合现有数据可挖掘程度及特点，选取高程、坡度、剖面曲率（*SOS*）、平面曲率（*SOA*）以及影像自带的 3 个波段（R、G、B）作为识别特征进行学习。影像自带的 3 个波段（R、G、B）可直接从光学影像中提取。高程、坡度、平面曲率（*SOA*）、剖面曲率（*SOS*）4 个特征图层可通过对研究区 DOM 数据进行处理和相应的空间分析获取。通过对这些特征因子进行选择、组合，选出最适合研究区的特征组合，即可完成滑坡识别模型的构建。

其中，高程指标可反映滑源区的绝对高度和滑坡体蕴含的势能；高程相差越大，即表示孕灾斜坡具有更大的临空面，越容易形成滑坡灾害（图 5a）。坡度是对地面高程的变化率的求解（图 5b），该指标可反映斜坡倾斜程度对滑坡发育的贡献。剖面曲率是指微分空间的地面坡度的变化率（图 5c），即坡度之梯度（slope of slope，简称 SOS），该指标可反映坡形对孕育滑坡的贡献。平面曲率又叫地面坡向变率（图 5d），它是坡向之坡度（slope of aspect，简称 SOA），该指标可反映的坡向变化对滑坡发育的贡献。

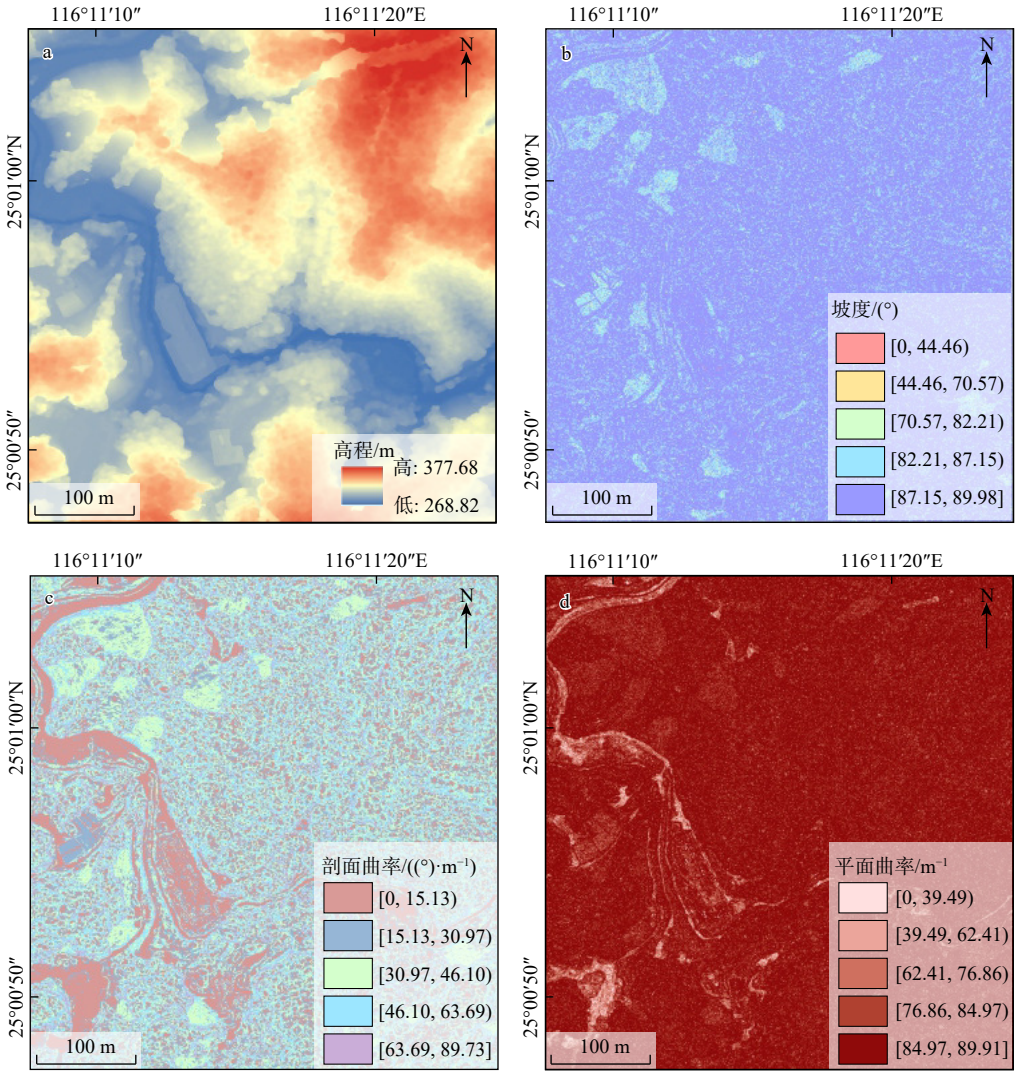


图 5　各特征因子示意图

Fig. 5　Schematic diagram of each feature factor

3　模型训练及验证

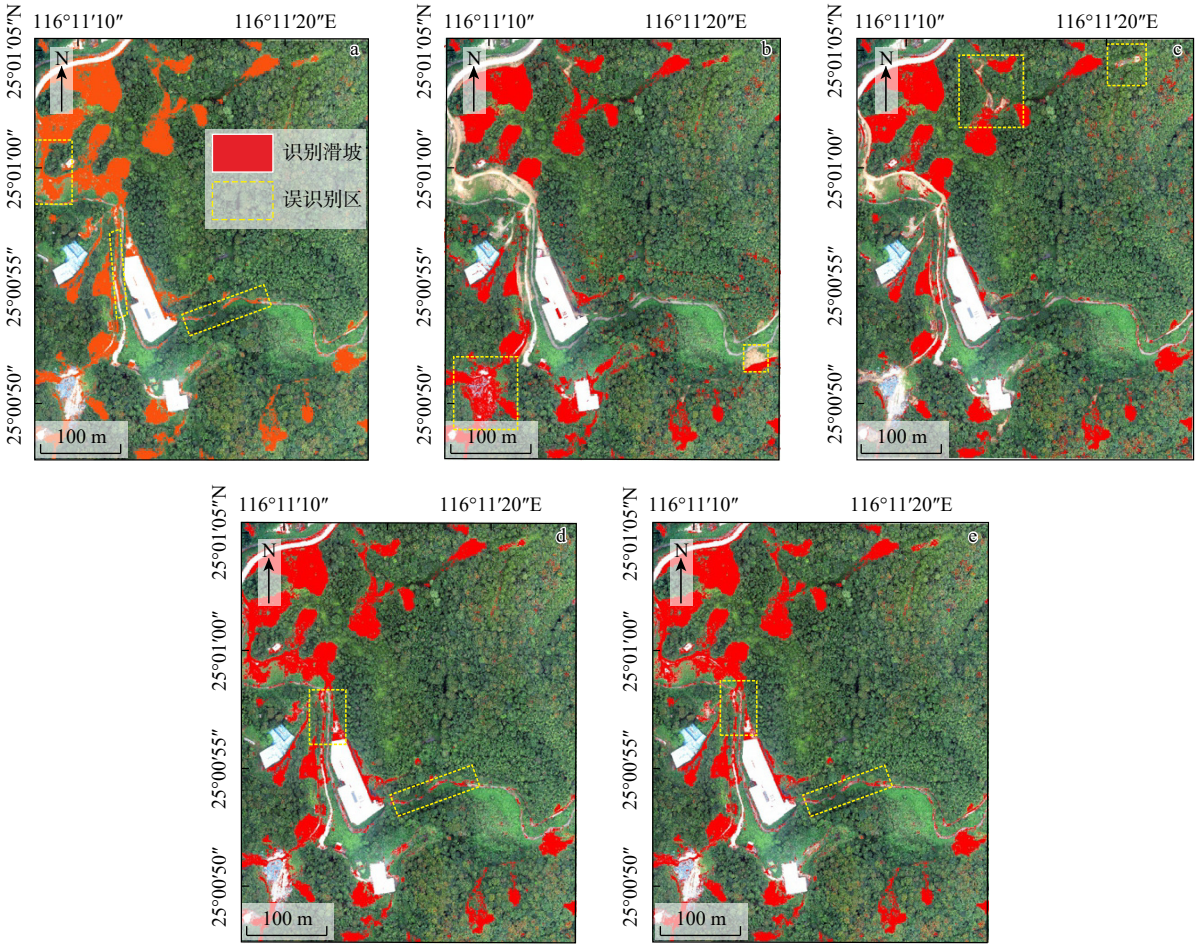
3.1　单特征叠加模型训练结果分析

为分析单一特征在滑坡识别中的贡献，优化该特征因子在识别模型中的参数选取，分别构建“影像+单一地质环境特征”组合进行训练，并对训练区滑坡进行识别，结果见图 6。图 6a ~ e 分别为基于影像自带 R、G、B 三波段识别效果、基于影像叠加高程等各单一特征组合的识别效果。通过比较各图的差异性，分析单一特征的加成对识别效果的影响。

由图 6a 可知，采用无人机影像 R、G、B 三波段特征识别效果良好，基本可以识别出所有生物视觉可见的滑坡体。但由于修路、建房等工程活动产生的裸露地表的影像特征跟滑坡体十分相像，也被误识别了出来，如图 6a 下方长条形房屋两侧的道路、以及左上方道路两侧连片识别等，这无疑影响了滑坡识别的准确性。从图 6b ~ e 4 幅识别结果图来看，已发生的滑坡体基本全部被识别出，但叠加每一种特征因子对灾害体的识别程度不一样，同样也存在误识别情况。如图 6b 是叠加高程因子的识别结果，能够完全抵消图 6a 中对道路的错误识别；但却增加了左下角房屋误识别。图 6c 是影像叠加坡度因子的识别结果，其结果较图 6a, b 更好，不仅没有误识别房屋，而且有效降低了对道路的误识别，但在图上方的两处较小灾害被遗漏了。与图 6a ~ c 相比，图 6d, e 没能抵消掉对道路误识别，但能纠正图 6b 对房屋的误识别，也能纠正图 6c 上方的识别遗漏区。这几种识别结果充分说明，所选的特征图层在滑坡体识别中均发挥了相应的作用，且各种贡献各不相同，这也表明在影像自带特征基础上叠加其他地质环境特征对滑坡体进行识别是十分必要的。

3.2　多特征组合模型训练结果分析

为了抵消单特征因子叠加存在的负面影响，同时充分发挥不同特征因子在滑坡体识别过程中的作用，尝试开展多特征因子组合模型训练，并通过不断迭代、优化，最终选出最适合研究区的特征组合构建



a. R、G、B 三波段识别结果图；b. 影像与高程识别结果图；c. 影像与坡度识别结果图；d. 影像与平面曲率识别结果图；e. 影像与剖面曲率识别结果图

图 6　基于无人机影像的 R、G、B 三波段与高程、坡度、平面曲率、剖面曲率的两两组合识别结果及相应评价示意图

Fig. 6　Landslide recognition results and evaluation metrics based on pairwise combinations of R, G and B bands with elevation, slope, plane curvature, profile curvature from UAV imagery

识别模型。为此，共设计 3 种特征组合开展模型训练，分别为：组合 1：RGB 波段+高程(*DEM*)+平面曲率(*SOA*)+剖面曲率(*SOS*)；组合 2：RGB 波段+坡度(*Slope*)+平面曲率(*SOA*)+剖面曲率(*SOS*)；组合 3：RGB 波段+高程(*DEM*)+坡度(*Slope*)+平面曲率(*SOA*)+剖面曲率(*SOS*)。

不同特征组合模型在经滑坡训练区训练后得到的识别结果见图 7，模型识别性能评价参数见表 1。

表 1　不同特征组合模型的识别结果参数

Table 1　Parameters of landslide recognition results for different feature combination models

参数	*Accuracy*	*Precision*	*Recall*	*F1score*
组合1	0.972	0.964	0.971	0.967
组合2	0.993	0.979	0.991	0.985
组合3	0.994	0.982	0.992	0.987

注：组合1. RGB波段+高程(*DEM*)+平面曲率(*SOA*)+剖面曲率(*SOS*)；组合2. RGB波段+坡度(*Slope*)+平面曲率(*SOA*)+剖面曲率(*SOS*)；组合 3. RGB波段+高程(*DEM*)+坡度(*Slope*)+平面曲率(*SOA*)+剖面曲率(*SOS*)；*Accuracy*. 准确率；*Precision*. 精确率；*Recall*. 召回率；*F1score*. 总体得分；下同

从图 7 的识别效果看：3 个特征组合对滑坡体的识别效果都很好，能够识别几乎全部滑坡。其中，图 7a 为组合 1 的识别结果图，可以很好去掉公路沿线的误识别现象，但是没有能够很好地消除部分左下方建筑物的误识别，识别视觉效果较单一特征识别效果更好。图 7b 为组合 2 的识别结果，能够很好地克服组合 1 无法消除建筑物误识别的问题，过滤掉了部分由于人类活动(修路、盖房)产生的裸地以及部分雨水冲积到路面的残土，但部分滑坡未能完整识别，结果大大减小了误识别区域，识别出来的滑坡相对完整；从视觉效果看，该模型识别结果较好。组合 2 各项模型评价指标较组合 1 有所提升，但是个别小型、非典型滑坡没有识别出来。图 7c 结果与图 7a 类似，但降低了图 7a 左下房屋的误识别。

3 个组合识别精度指标见表 1，3 种组合的识别结果参数表现都很好，结果的准确率都高于 0.97，识别精确率都高于 0.96，其中组合 3 准确率和召回率均大于 0.99，且分数最高，组合 1 的参数略低于组合 2 和组合 3，表明识别效果都很优秀，且特征越多，打分效果越全面。加入坡度、平面曲率、剖面曲率的识别模型与仅以 R、G、B 三波段作为特征的识别模型相比，各项模型评价指标表现良好，均有所提升。综合从模型评价参数表和识别效果图来看，组合 3 是地区滑坡识别模型的最优模型。

3.3　模型构建及验证

（1）特征组合选取

相较于组合 1，多特征组合 3 构建的模型不仅减小了误识别区域的面积，还明显看到各项模型评价指标都有了显著的提升；相较于组合 2，组合 3 构建的模型能识别出更多小型、非典型的滑坡，为本研究中表现最好的模型，即组合 3 为该研究区搭建滑坡识别模型的最佳特征组合。

（2）超参数优选

为评估模型泛化能力及优选超参数，采用五折交叉验证方法。将样本数据按类别比例随机划分为 5 个互不重叠的子集，依次以各子集作为验证集，其余作为训练集进行模型训练与验证。在每次训练中，将选定特征输入 LightGBM 模型，该模型基于前向分布算法迭代构建多棵决策树，并通过监控对数损

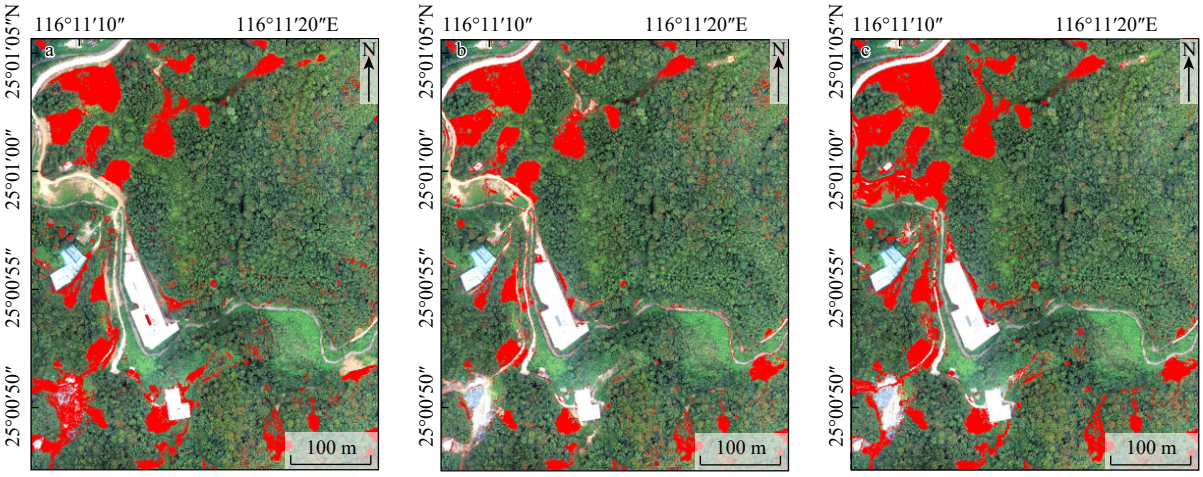


图 7　3 种特征组合模型的滑坡识别结果对比图（a. 组合 1；b. 组合 2；c. 组合 3；红色区域为识别滑坡；下同）

Fig. 7　Comparison of landslide recognition results using three feature combination models

失函数判断收敛。完成五折交叉验证后，依据各超参数组合在验证集上的平均性能选取最优配置，并使用全部训练数据重新训练最终模型。交叉验证结果显示，模型在约 100 次迭代后损失值稳定于 0.1 以下(图 8a)，表明其具备良好的有效性与稳定性。基于测试集得到的特征重要性如图 8b 所示，该结果对应识别模型中的最佳特征组合。

（3）模型验证

由最佳特征组合及最优权重代入 LightGBM 算法，即得到本研究区滑坡的最佳识别模型。采用这一模型对预先设置的测试区(图 4c)进行滑坡识别，效果见图 9 和表 2。从图 9 可以看出，测试区的滑坡全部被识别出，且识别面积接近 100%。根据评价参数结果，按栅格单元计算，模型识别滑坡的准确率(*Precision*)为 0.97，识别滑坡的召回率(*Recall*)为 0.99，识别滑坡时模型总体得分(*F*1*score*)为 0.98；在识别非滑坡单元时，精度为 1.00，召回率达到 0.99，模型得分为 1.00。这表明，该模型对训练区临近区域内的滑坡有较好的适应性，可移植性较强。

这也表明，基于 LightGBM 算法采用 RGB 波段、高程(*DEM*)、坡度(*Slope*)、平面曲率(*SOA*)以及剖面曲率(*SOS*)构建的综合特征集识别模型，可以实现对研究区已发生降雨型浅表层滑坡灾害的有效识别。

表 2　最佳特征组合模型的测试评价参数

Table 2　Test evaluation parameters for the best feature combination model

参数	*Precision*	*Recall*	*F*1*score*
滑坡	0.97	0.99	0.98
非滑坡	1.00	0.99	1.00

4　滑坡灾害早期识别

现实防灾减灾工作中，除了对滑坡灾害刚发生阶段的快速识别需求外，往往更希望对灾害可能发生在哪里进行提前预判，也即对灾害进行早期识别。但潜在灾害隐患点因为尚未完全发育，往往不具备典型滑坡灾害体的轮廓雏形和形态特征，也就无法使用光学遥感影像进行识别。现阶段行业内通常采用的办法是利用 InSAR 技术对变形体进行变形量监测，圈定正在发生形变的斜坡范围[35-38]。但在东南诸省这种高植被覆盖区，InSAR 技术受限于植被覆盖密度大的影响，难以穿透植被枝叶抵达地表。因此，作为高植被覆盖区，本研究还需解决一个关键问题：如何在滑坡发生前，利用模型有效识别潜在滑坡灾害隐患点(或潜在滑坡发育区)。

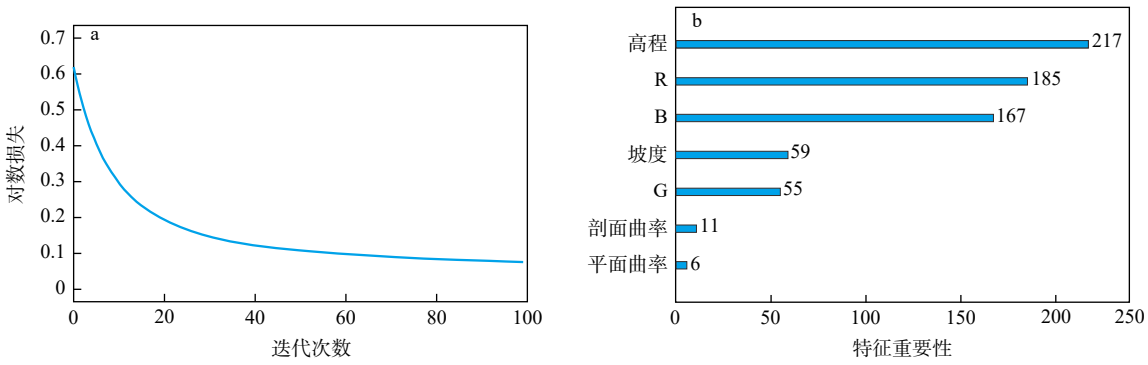


图 8　最佳特征组合模型的损失函数统计图 (a) 和各特征重要性输出图 (b)

Fig. 8　Loss function statistics (a) and each feature importance (b) for the best feature combination model

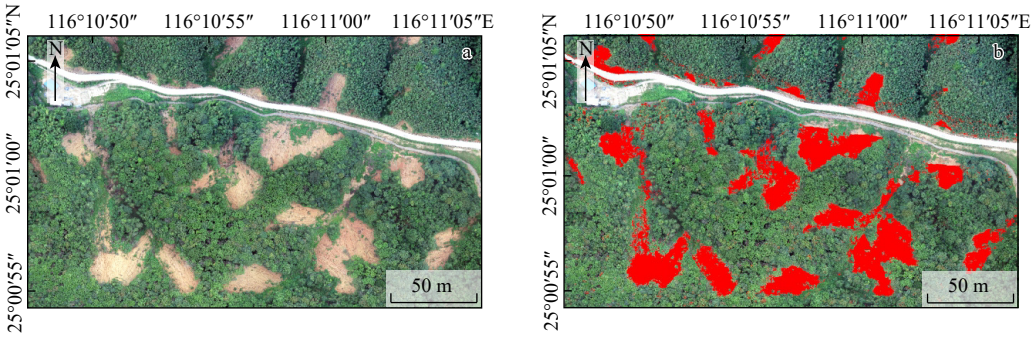


图 9　测试区原始影像 (a) 和最佳识别模型的滑坡自动识别结果 (b)（测试区见图 4c）

Fig. 9　Original image (a) and landslide automatic recognition results of the best recognition model (b) in the test area

4.1 早期识别模型构建思路

基于图 8b 中所展示的各特征及相应权重在已发生滑坡灾害的快速识别中表现优异，且各地质环境特征因子在滑坡灾害孕灾过程中均发挥了重要作用[39-40]，若去除影像（R、G、B 波段）信息后，仅采用剩余特征因子组合也可以对这些滑坡灾害进行识别[41]，说明仅使用滑坡孕灾环境特征因子组合也可实现对滑坡灾害的早期识别。下文即对这一早期识别的效果进行测试和验证。

4.2 识别效果验证

（1）十方镇降雨滑坡识别效果验证

首先选择十方镇的训练区（图 4b）开展识别，观察识别效果，目的是检验该新模型能否很好识别输入图层的信息，也即反推回去看看能否实现 100% 识别，识别效果见图 10。从图 10 中可以看出，新构建的早期识别模型能够识别出潜在的滑坡风险隐患区。对比图 10 和图 7c 识别结果，发现新模型仍能对已发生滑坡区域进行 100% 识别，且灾害体本身的识别精度不低于图 7c，效果良好。其次，选择模型分别对图 4c ~ e 3 个测试区进行测试，验证去除 RGB 特征的模型与原模型对同一地区的识别效果。

图 11 是 2 种模型对图 4c 测试区的识别结果图。结果发现，早期识别模型总体识别效果与训练区快速识别效果相似，总体良好，能够识别出大部分的灾害体，同时存在误报现象。不同的是，新模型在图 4c 测试区未能对全部灾害体实现 100% 识别，如图 11a 左下角的 3 处灾害体未能全部识别。

经过计算，早期识别模型在识别滑坡时的总体得分达到了 0.86，证明了模型在滑坡早期识别领域的优秀表现（表 3）。为了全面检验该模型的普适性和泛化能力，特别选取了与训练区域相隔较远的 2 处灾害较集中的区域作为新的测试区（图 4d, e）。与

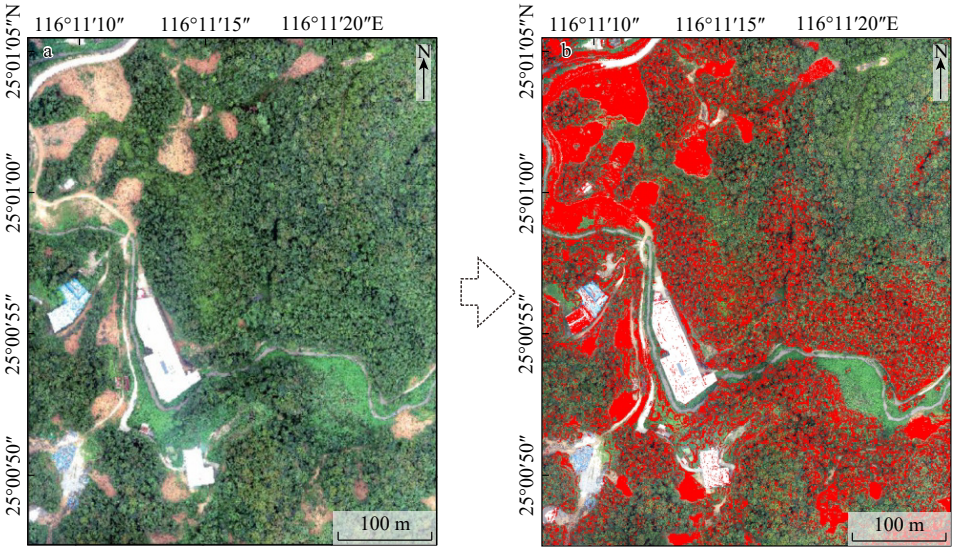
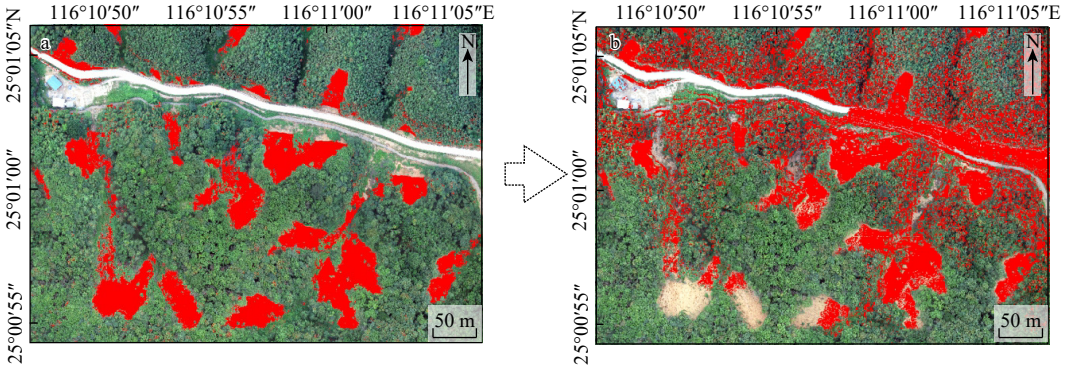


图 10 训练区原始影像 (a) 和早期识别模型的滑坡识别结果 (b)（训练区见图 4b）

Fig. 10 Original image (a) and landslide recognition results of the early recognition model (b) in the training area



a. 未去除影像（R、G、B 波段）信息；b. 去除影像（R、G、B 波段）信息；测试区见图 4c

图 11 测试区早期识别模型的滑坡识别结果对比图

Fig. 11 Comparison of landslide recognition results of the early recognition model in the test area

图 4c 测试区相比，这 2 个区域的地质条件、植被发育程度等可能与训练区域存在的差异更大，但正是这些差异能够更有效地检验模型的泛化能力。

测试效果的损失函数见图 12，经迭代 160 次后，损失函数降到可接受范围，运算结束。

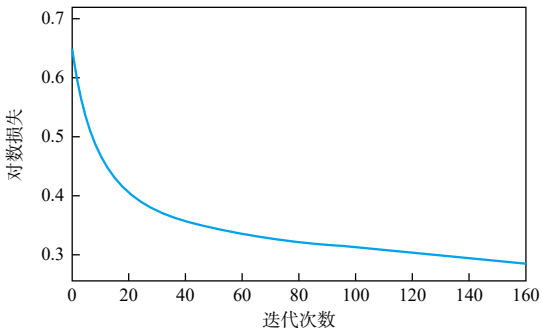


图 12　早期识别模型损失函数统计图

Fig. 12　Loss function statistics of the early identification model

识别效果见图 13。从结果图可看出，与快速识别模型在图 4b 训练区和图 4c 测试区的表现相比，早期识别模型在图 4d 测试区误识别范围较多，而图 4e 测试区则存在单灾害体部分面积识别不全的问题。但从总体识别效果来看，新模型在图 4d 测试区和图 4e 测试区的测试效果均较好，能明显分辨出滑坡灾害的位置和形态，能够识别出全部灾害体。这充分说明在没有遥感影像加持的条件下，该模型是可以仅依靠其他特征参数就能对滑坡体进行提前识别的，也就是达到早期识别的目的。绘制了受试者工作特征曲线(receiver operating characteristic curve，简称 ROC)（图 14)通过 ROC 曲线下的面积(area under the roc curve，简称 AUC)来评价该早期识别模型的识别效果，结果显示 *AUC* 值达到 0.93，说明识别正确率较高、效果良好。

早期识别模型的评价参数见表 3。结果显示，运用本研究构建的滑坡早期识别模型在十方镇滑坡灾害识别中效果显著，准确率(*Accuracy*)达到了 0.86，精确率(*Precision*)达到了 0.93，这一结果显著优于许多同类方法。输出结果显示召回率(*Recall*)值达 0.86，表明有 86% 的滑坡像元能被正确分类，模型总体得分达到了 0.86，说明该模型在研究区滑坡灾害早期识别中表现优秀，具有较好的泛化能力。模型最终的特征重要性输出结果见图 15。

表 3　早期识别模型的评价参数

Table 3　Evaluation parameters of the early recognition model

参数	*Accuracy*	*Precision*	*F1score*	*Recall*
灾害体	0.86	0.93	0.86	0.86

（2）武平县滑坡隐患点识别效果验证

尽管该模型在十方镇降雨型浅层滑坡识别实验中效果较好，但仍需对模型能否在更大范围内是否适用进行检验。为此，进一步将测试区域扩大至武平县全域，对县域范围内潜在滑坡隐患区进行全面系统的识别，并采用 2020 年开展的武平县 1∶5 万地质灾害调查报告[28] 中统计的滑坡灾害隐患点对识别效果进行验证。图 16a 为识别模型识别出的滑坡隐患点分布图，图 16b 为隐患点机器学习早期识别结果数据与 1∶5 万灾害隐患点调查报告数据的叠加图。经统计分析发现，经人工调查识别的 604 处潜在风险边坡隐患点中，有 482 处所处位置被模型判定为滑坡隐患点，比例达 79.8%。这一高比例充分展现了模型在识别潜在风险斜坡隐患点方面的准确性。

武平县全域滑坡隐患点识别效果评价指标如表 4 所示。经栅格单元分析，隐患点栅格单元识别精确率达 0.89，斜坡隐患点真实正确率(召回率)为 0.80。综合两指标，模型滑坡灾害识别总体得分为

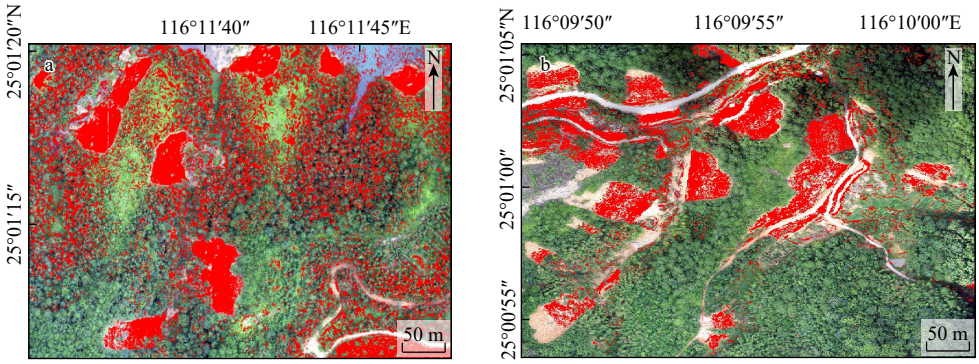


图 13　测试区早期识别模型的滑坡识别结果（a. 图 4d 测试区；b. 图 4e 测试区）

Fig. 13　Landslide recognition results of the early recognition model in the test area

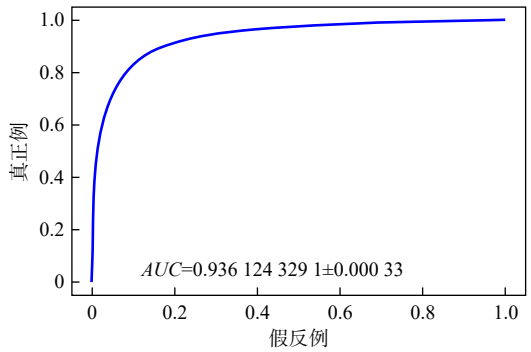


图 14　早期识别模型的受试者工作特征曲线（ROC）

Fig. 14　Receiver operating characteristic curve of the early recognition model

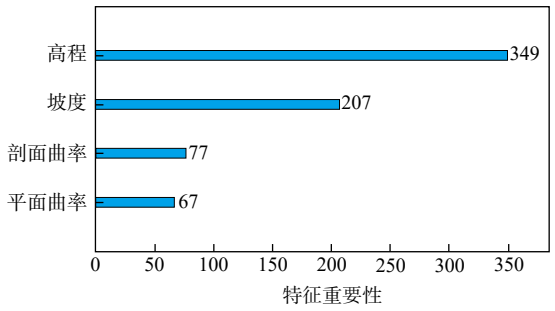


图 15　早期识别模型的特征重要性输出图

Fig. 15　Feature importance of the early recognition model

0.84, 表明模型整体性能优异。另一方面, 非隐患点栅格单元识别精确率为 0.83, 非隐患点召回率为 0.86, 显示模型避免误判能力强, 能够正确区分灾害区域和非灾害区域。

综上所述, 基于 LightGBM 算法构建滑坡灾害识别模型不仅在识别隐患点时表现出色, 而且在识别非隐患点区域时也具有良好的性能。特别是在处理临近训练区的其他区域内时, 模型能够较好地适应并准确识别滑坡灾害发育区, 显示出较强的可移植性和泛化能力。这一研究成果对于提升东南台风暴雨区滑坡–崩塌等斜坡灾害的早期识别和防灾减灾提供了新的技术支持, 并有望在实际应用中发挥重要作用。

表 4　武平县全域模型测试评价参数

Table 4　Model test evaluation parameters in Wuping County

参数	*Accuracy*	*Precision*	*F1score*	*Recall*
隐患点	0.80	0.89	0.84	0.80
非隐患点	0.85	0.83	0.84	0.86

5　结　论

(1)以坡度、高程、平面曲率、剖面曲率和 R、G、B 三波段特征组合构建的多特征组合识别模型在开展降雨型浅层碎石土滑坡滑后识别时, 识别准确性优于单特征识别效果。

(2)由坡度、高程、平面曲率、剖面曲率和 R、G、B 三波段组成的快速识别模型能准确识别已发育滑坡体, 准确率(*Accuracy*)和召回率(*Recall*)值均大于 0.99。

(3)去除可视化图层后, 采用该模型开展高植被覆盖区滑坡早期识别时, 准确率(*Accuracy*)和召回率(*Recall*)仍达 0.86; 推广至武平县全县范围时, 准确率(*Accuracy*)达 0.80, 表明识别准确度较高。

(4)在对十方镇降雨滑坡进行识别时, 基于 LightGBM 算法构建的滑坡灾害早期识别模型

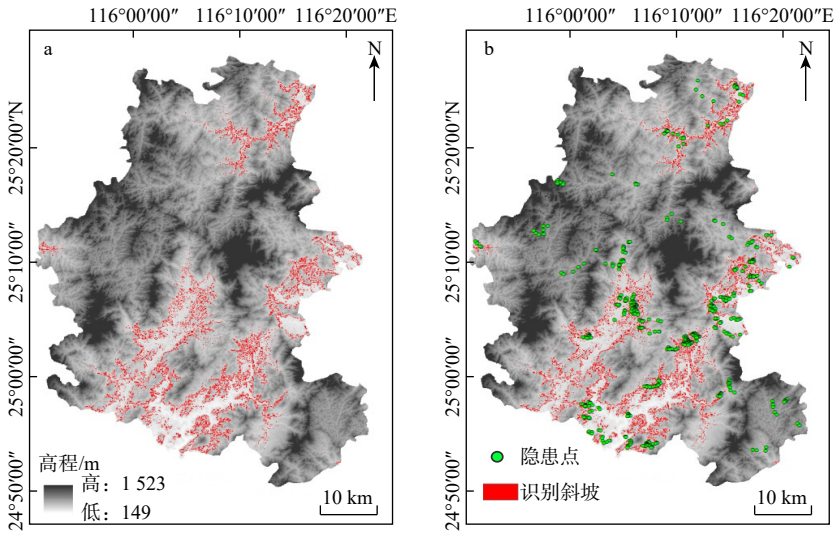


图 16　武平县斜坡灾害隐患点早期识别结果 (a) 与调查隐患点分布图 (b)

Fig. 16　Early recognition results of slope hazard points (a) and distribution of surveyed hazard points (b) in Wuping County

*F*1*score* 高达 0.86, 推广至武平县全县范围时, 模型 *F*1*score* 高达 0.84, 表明该模型泛化能力较强, 能够较准确对相似地质环境孕育的滑坡灾害实现自动识别。

所有作者声明不存在利益冲突。

The authors declare that no competing interests exist.

参考文献:

[1] ZHOU D Y, ZHANG Z, LI J C, et al. Seepage-stress coupled modeling for rainfall induced loess landslide[J]. Theoretical and Applied Mechanics Letters, 2019, 9(1): 7-13.

[2] LI Z, HE Y J, LI H E, et al. Antecedent rainfall induced shallow landslide: A case study of Yunnan landslide, China[J]. Scientia Iranica, Transactions A: Civil Engineering, 2019, 26(1): 202-212.

[3] 佚名. 河南郑州“7·20”特大暴雨灾害调查报告公布 [J]. 中国应急管理, 2022(2): 4.
Anon. The release of the investigation report on the severe rainstorm disaster in Zhengzhou, Henan on July 20[J]. China Emergency Management, 2022(2): 4. (in Chinese with English abstract)

[4] 吴伟伟, 胥涛, 高志强, 等. 山西晋中地区“2021.10”暴雨特征与洪灾分析 [J]. 中国防汛抗旱, 2022, 32(12): 82-88.
WU W W, XU T, GAO Z Q, et al. Rainstorm characteristics and flood disaster analysis of “2021.10” in Jinzhong area, Shanxi Province[J]. China Flood & Drought Management, 2022, 32(12): 82-88. (in Chinese with English abstract)

[5] 李天毅, 戴宁, 丛佩娟, 等. “23·7”京津冀暴雨洪水特性分析: 以河北省临城县为例 [J/OL]. 水土保持通报: 1-13 [2026-01-14]. https://link.cnki.net/doi/10.13961/j.cnki.stbctb.2024.02.017.
LI T Y, DAI N, CONG P J, et al. Analysis of rainstorm flood characteristics in “23·7” Beijing-Tianjin-Hebei: A case study of Lincheng County, Hebei Province[J/OL]. Bulletin of Soil and Water Conservation: 1-13 [2026-01-14]. https://link.cnki.net/doi/10.13961/j.cnki.stbctb.2024.02.017.(in Chinese with English abstract)

[6] LI G E, HU B, LI H, et al. Early identifying and monitoring landslides in Guizhou Province with InSAR and optical remote sensing[J]. Journal of Sensors, 2021, 2021(1): 6616745.

[7] WANG Y A, LIU D L, DONG J, et al. On the applicability of satellite SAR interferometry to landslide hazards detection in hilly areas: A case study of Shuicheng, Guizhou in Southwest China[J]. Landslides, 2021, 18(7): 2609-2619.

[8] DAI K R, FENG Y, ZHUO G C, et al. Applicability analysis of potential landslide identification by InSAR in Alpine-canyon terrain: Case study on Yalong River[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2018, 15: 2110-2118.

[9] LIANG J T, DONG J H, ZHANG S, et al. Discussion on InSAR identification effectivity of potential landslides and factors that influence the effectivity[J]. Remote Sensing, 2022, 14(8): 1952.

[10] 许强, 陆会燕, 李为乐, 等. 滑坡隐患类型与对应识别方法 [J]. 武汉大学学报(信息科学版), 2022, 47(3): 377-387.
XU Q, LU H Y, LI W L, et al. Types of potential landslide and corresponding identification technologies[J]. Geomatics and Information Science of Wuhan University, 2022, 47(3): 377-387. (in Chinese with English abstract)

[11] LIU Y, WU L Z. Geological disaster recognition on optical remote sensing images using deep learning[J]. Procedia Computer Science, 2016, 91: 566-575.

[12] SAMEEN M I, PRADHAN B, LEE S. Application of convolutional neural networks featuring Bayesian optimization for landslide susceptibility assessment[J]. Catena, 2020, 186: 104249.

[13] GUARNIERI A, MASIERO A, VETTORE A, et al. Evaluation of the dynamic processes of a landslide with laser scanners and Bayesian methods[J]. Geomatics, Natural Hazards and Risk, 2015, 6(5/6/7): 614-634.

[14] 邬礼扬, 殷坤龙, 曾韬睿, 等. 不同栅格尺寸下输电线路地质灾害易发性评价 [J]. 地质科技通报, 2024, 43(1): 241-252.
WU L Y, YIN K L, ZENG T R, et al. Evaluation of geological disaster susceptibility of transmission lines under different grid resolutions[J]. Bulletin of Geological Science and Technology, 2024, 43(1): 241-252. (in Chinese with English abstract)

[15] NHU V H, MOHAMMADI A, SHAHABI H, et al. Landslide detection and susceptibility modeling on Cameron Highlands (Malaysia): A comparison between random forest, logistic regression and logistic model tree algorithms[J]. Forests, 2020, 11(8): 830.

[16] DOU J, YUNUS A P, BUI D T, et al. Improved landslide assessment using support vector machine with bagging, boosting, and stacking ensemble machine learning framework in a mountainous watershed, Japan[J]. Landslides, 2020, 17(3): 641-658.

[17] 吴季寰, 张春山, 孟华君, 等. 抚顺西露天矿区滑坡易发性评价与时空特征分析 [J]. 地质力学学报, 2021, 27(3): 409-417.
WU J H, ZHANG C S, MENG H J, et al. Temporal and spatial characteristics of landslide susceptibility in the west open-pit mining area, Fushun, China[J]. Journal of Geomechanics, 2021, 27(3): 409-417. (in Chinese with English abstract)

[18] CHEN F, YU B, LI B. A practical trial of landslide detection from single-temporal Landsat8 images using contour-based proposals and random forest: A case study of national Nepal[J]. Landslides, 2018, 15(3): 453-464.

[19] BUI D T, TSANGARATOS P, NGUYEN V T, et al. Comparing the prediction performance of a deep learning neural network model with conventional machine learning models in landslide susceptibility assessment[J]. Catena, 2020, 188: 104426.

[20] PANG D D, LIU G, HE J, et al. Automatic remote sensing identification of co-seismic landslides using deep learning methods[J]. Forests, 2022, 13(8): 1213.

[21] WANG Y, WANG X F, JIAN J F. Remote sensing landslide recognition based on convolutional neural network[J]. Mathematical Problems in Engineering, 2019, 2019(1): 8389368.

[22] 高瑞源. 基于机器学习方法的区域泥石流易发性评价研究 [D]. 长春: 吉林大学, 2023.
GAO R Y. Study on regional debris flow susceptibility evaluation based on machine learning methods[D]. Changchun: Jilin University, 2023. (in Chinese with English abstract)

[23] 冀全伟, 王文磊, 刘治博, 等. 一种基于机器学习算法的岩性填图方法 [J]. 地质力学学报, 2021, 27(3): 339-349.
JI Q W, WANG W L, LIU Z B, et al. A machine learning-based lithologic mapping method[J]. Journal of Geomechanics, 2021, 27(3): 339-349. (in Chinese with English abstract)

[24] DING A, ZHANG Q, ZHOU X, et al. Automatic identification of landslide based on CNN and texture change detection[C]//Anon. Proceedings of the Chinese Association of Automation (YAC), Youth Academic Annual Conference. [S. l.], IEEE, 2016: 444-448.

[25] WANG H J, ZHANG L M, YIN K S, et al. Landslide identification using machine learning[J]. Geoscience Frontiers, 2021, 12(1): 351-364.

[26] KEYPORT R N, OOMMEN T, MARTHA T R, et al. A comparative analysis of pixel- and object-based detection of landslides from very high-resolution images[J]. International Journal of Applied Earth Observation and Geoinformation, 2018, 64: 1-11.

[27] GHORBANZADEH O, BLASCHKE T, GHOLAMNIA K, et al. Evaluation of different machine learning methods and deep-learning convolutional neural networks for landslide detection[J]. Remote Sensing, 2019, 11(2): 196.

[28] 福建省煤田地质局, 福建省 121 地质大队. 武平县 1∶5 万地质灾害调查与风险评价成果报告 [R]. 福州: 福建省煤田地质局, 福建省 121 地质大队, 2021.
Fujian 121 Geological Brigade, Fujian Provincial Coal Geology Bureau. Report on the results of 1∶50 000 geological hazard investigation and risk assessment in Wuping County[R]. Fujian 121 Geological Brigade, Fujian Provincial Coal Geology Bureau, 2021. (in Chinese)

[29] LIU Y H, HUANG J B, XIAO R H, et al. Research on a regional landslide early-warning model based on machine learning: A case study of Fujian Province, China[J]. Forests, 2022, 13(12): 2182.

[30] WANG T, LI R, CHEN C, et al. Kinetic analysis of rainfall-induced landslides in May 2022 in Wuping, Fujian, SE China[J]. Water, 2024, 16(21): 3018.

[31] 罗常伟, 王双双, 尹峻松, 等. 集成学习研究现状及展望 [J]. 指挥与控制学报, 2023, 9(1): 1-8.
LUO C W, WANG S S, YIN J S, et al. Research status and prospect of ensemble learning[J]. Journal of Command and Control, 2023, 9(1): 1-8. (in Chinese with English abstract)

[32] LEO B. Bagging predictors[J]. Machine Learning, 1996, 24(2): 123-140.

[33] 徐继伟, 杨云. 集成学习方法: 研究综述 [J]. 云南大学学报(自然科学版), 2018, 40(6): 1082-1092.
XU J W, YANG Y. A survey of ensemble learning approaches[J]. Journal of Yunnan University (Natural Sciences Edition), 2018, 40(6): 1082-1092. (in Chinese with English abstract)

[34] 李航. 统计学习方法 [M]. 北京: 清华大学出版社, 2012.
LI H. Statistical learning method[M]. Beijing: Tsinghua University Press, 2012. (in Chinese)

[35] XIE M W, HUANG J X, WANG L W, et al. Early landslide detection based on D-InSAR technique at the Wudongde hydropower reservoir[J]. Environmental Earth Sciences, 2016, 75(8): 717.

[36] REN T H, GONG W P, GAO L, et al. An interpretation approach of ascending-descending SAR data for landslide identification[J]. Remote Sensing, 2022, 14(5): 1299.

[37] VAN NATIJNE A L, BOGAARD T A, VAN LEIJEN F J, et al. World-wide InSAR sensitivity index for landslide deformation tracking[J]. International Journal of Applied Earth Observation and Geoinformation, 2022, 111: 102829.

[38] NIU C, YIN W P, XUE W, et al. Multi-window identification of landslide hazards based on InSAR technology and factors predisposing to disasters[J]. Land, 2023, 12(1): 173.

[39] 王灿星, 朱杰勇, 喻聪骏, 等. 基于皮尔逊Ⅲ型曲线的不同降雨工况下的崩滑地质灾害危险性评价 [J]. 地质科技通报, 2025, 44(1): 194-204.
WANG C X, ZHU J Y, YU C J, et al. Risk assessment of landslide geological hazards under different rainfall conditions based on the Pearson Ⅲ curves[J]. Bulletin of Geological Science and Technology, 2025, 44(1): 194-204. (in Chinese with English abstract)

[40] 李星, 杨赛, 李远耀, 等. 面向区域滑坡易发性精细化评价的改进斜坡单元法 [J]. 地质科技通报, 2023, 42(3): 81-92.
LI X, YANG S, LI Y Y, et al. Improved slope unit method for fine evaluation of regional landslide susceptibility[J]. Bulletin of Geological Science and Technology, 2023, 42(3): 81-92. (in Chinese with English abstract)

[41] CHEN B, LI Z H, SONG C, et al. Automatic detection of active geohazards with millimeter-to-meter-scale deformation and quantitative analysis of factors influencing spatial distribution: A case study in the Hexi Corridor, China[J]. International Journal of Applied Earth Observation and Geoinformation, 2024, 131: 103995.

Summary for “基于机器学习的高植被覆盖区降雨滑坡识别”

Recognition of rainfall-induced landslide in high vegetation coverage area based on machine learning

QI Lining[1,2,3,4], GUO Chaoxu[2,3], MENG Huajun[1,5*], CHU Kelei[2,3,6], WANG Yuanyuan[1,4]

1. Institute of Geomechanics, Chinese Academy of Geological Sciences, Beijing 100081, China;
2. Key Laboratory of Geohazard Prevention of Hilly Mountains, Ministry of Natural Resources, Fuzhou 350002, China;
3. Key Laboratory of Geohazard of Fujian Province, Fuzhou 350002, China;
4. School of Earth Sciences and Resources, China University of Geosciences (Beijing), Beijing 100083, China;
5. Observation and Research Station of Geological Disaster in Baoji, Shaanxi Province, Ministry of Natural Resources, Baoji Shaanxi 721000, China;
6. Chinese Academy of Geological Sciences, Beijing 100037, China

First author, E-mail: qilining19990228@163.com

* Corresponding author, E-mail: mhjun-521@foxmail.com / mhjun521@cags.ac.cn

Abstract: In recent years, influenced by global warming, the southeast coastal provinces of China have frequently experienced super typhoons, heavy rainfall, and prolonged monsoon rains, triggering numerous shallow landslides the pose serious threats to local lives and property. [**Objective**] To rapidly obtain more accurate disaster information for support emergency response and scientific decision-making, [**Methods**] this study developed an automatic recognition model for shallow landslides using a machine learning approach, the LightGBM decision tree algorithm. The model targets clustered shallow landslides developed in highly weathered granite layers induced by heavy rainfall in high vegetation coverage area. High-resolution aerial images and DOM data from Wuping County, Fujian Province, were used, with feature parameters including slope, elevation, plane curvature (*SOA*), profile curvature (*SOS*), and R, G, and B bands selected to construct, train and test the recognition model. [**Results**] The results indicate that: ①The multi-feature combination model achieves higher accuracy in landslide recognition compared to single-feature models; ②The rapid recognition model incorporating slope, elevation, plane curvature, profile curvature, and R, G, and B bands accurately identifies existing landslides, with both *Accuracy* and *Recall* exceeding 0.99; ③Even after removing the visual layer, the model maintains an *Accuracy* and *Recall* of 0.86 for early landslide recognition in high vegetation coverage area. When extended to the entire Wuping County, *Accuracy* reached 0.80, demonstrating high recognition performance; ④The model achieves and *F*1*score* of 0.86, and when generalized to the entire county, the *F*1*score* remains as high as 0.84, indicating strong generalization capability and reliable performance in automatically recognizing landslide hazards in similar geological environments. [**Conclusion**] This research provides a novel approach for early recognition of rainfall-induced landslides in high vegetation coverage area and offers scientific and technological support for risk prevention and disaster response decision-making in typhoon-related and rainstorm-affected areas such as Haixi.

Key words: rainfall-induced; landslide; machine learning; high vegetation coverage area

Received: 2024-06-03; Revised: 2024-10-18; Accepted: 2024-10-21

doi: 10.19509/j.cnki.dzkq.tb20240298