

(主持人: 洪亮, 云南师范大学教授, 博士生导师, 云南师范大学地理学部副主任, 主要研究方向: 遥感大数据智能处理与时空分析.)

基于改进 ResNet50 网络的泥石流沟谷识别^{*}

刘秋雨¹, 王保云^{1,2**}

(1. 云南师范大学 数学学院, 云南 昆明 650500;

2. 云南师范大学 云南省现代分析数学及应用重点实验室, 云南 昆明 650500)

[摘 要] 针对传统神经网络进行泥石流沟谷图像分类时, 可能出现准确率不高、提取图像特征较差、边缘模糊等问题, 对 ResNet50 网络进行改进. 在 ResNet50 网络部分残差块前加入注意力机制模块, 使其具有更高的性能和准确性, 可以精确捕捉到泥石流沟谷图像中的地形地貌. 试验结果表明, 改进后的 ResNet50 网络在泥石流沟谷图像的分类准确率达到 83.02%, 其分类性能在 ResNet50 网络的基础上提升了 11.32 个百分点, 且准确率、召回率、精确率、 F_1 值和 AUC 值等各项指标均优于 ResNet50 网络和其他深度学习识别算法.

[关键词] 注意力机制; 卷积神经网络; 泥石流灾害; ResNet50

[中图分类号] TP391.41; TP183; X928.7 **[文献标志码]** A **[文章编号]** 1674-5639(2024)06-0114-06

DOI: 10.14091/j.cnki.kmxyxb.2024.06.016

我国疆域辽阔, 而位于青藏高原地区的部分四川省地形地貌多样且复杂多变. 由于其地势和气候的复杂多变, 四川发生的泥石流灾害具有来势猛、成灾快、数量多、伤亡大、损失大、灾后恢复治理困难的特点, 危害十分严重, 是四川严重的自然灾害之一^[1]. 根据自然资源部关于山区丘陵地质灾害的研究表明, 在 1999 年至 2008 年, 四川省泥石流灾害发生的数量高达 3 171 次, 在全国泥石流灾害总数中占比达到 11%, 比云南省泥石流灾害占比高出 2%, 在全国排名前五^[2]. 根据四川地质灾害主要类型、成因及影响研究表明, 在四川省, 由暴雨激发的泥石流灾害为多沟齐发, 甚至数十条沟同时暴发. 如 1984 年 8 月 11 日, 得荣县城后山出现局部性暴雨, 20 分钟便导致 4 条沟同时暴发泥石流, 6 万 m^3 泥石流体在约 5 min 内被冲到沟口, 冲入机关民房, 淤积县城主要街道, 造成重大经济损失和人员伤亡. 因此对泥石流灾害区域进行快速识别十分重要^[3].

目前, 对泥石流沟谷的识别方法大致有以下 4 类: 第 1 类是通过实地考察, 对沟谷的地形、地貌、冲积扇特征、植被情况等进行调查识别, 此类方法准确率高, 但需要投入较大的人力和时间成本^[4]. 第 2 类主要基于遥感技术, 通过构建沟谷的目视解译特征进行泥石流的识别^[5]. 第 3 类使用统计学的方法, 对泥石流流域面积、相对高差、松散物源面积等进行建模, 通过计算出的危险因子大小, 判别沟谷的危险程度^[6]. 第 2、3 类方法在进行特征构建或选取危险因子时, 主要依赖研究者的主观判断, 需要一定的专业知识, 且模型结果受参数选取的影响较大. 最后一类方法是基于深度学习的方法, 在遥感影像的基础上, 通过网络提取泥石流沟谷的特征, 可更加高效、快速地预测泥石流, 该类方法往往需要大量数据作为支持^[7].

卷积神经网络识别性能会随着其层数的增加而逐渐提升, 但模型层数增加会导致模型过拟合和梯度消失以及训练持续时间过长等问题, 使模型训练困难、网络识别错误率增加, 导致神经网络性能退化^[8]. 何凯明^[9]等科学家在图像识别领域提出了 ResNet 网络. 具体来说, ResNet 在每个残差模块中增加一个跨

^{*} [收稿日期] 2024-01-11

[作者简介] 刘秋雨, 女, 山东临沂人, 云南师范大学在读硕士研究生, 研究方向为图像处理和深度学习.

^{**} [通信作者] 王保云, 男, 云南玉溪人, 云南师范大学副教授, 博士, 研究方向为机器学习及图像处理, E-mail: wspbmly@163.com.

[基金项目] 国家自然科学基金项目 (61966040).

层连接, 让特征可以直接传递到后面的层次, 从而保留原始特征, 避免特征逐层消失. 因此本文选择在具有 50 层网络结构的残差网络 ResNet50 上进行改进.

本文采用四川省泥石流沟谷数据集, 每个样本中包含了大量沟谷的地形地势、地貌特征等信息, 但在样本量较少的情况下, 简单的神经网络对泥石流灾害沟谷的特征捕捉不够完整. 因此, 提出基于注意力机制^[10-12]的泥石流灾害沟谷分类模型, 对泥石流沟谷图像的正负样本进行二分类训练. 注意力机制模块 CBAM (Convolutional Block Attention Module) 倾向于关注图像中比较感兴趣的部分, 因此基于注意力机制的网络模型会更加关注图像中的重点信息. 本文通过泥石流沟谷地形地貌特征, 实现对泥石流发生与否的概率预测, 并将改进的 ResNet50 网络模型与多个经典网络模型对比.

1 试验数据

1.1 数据收集

本文采用的数据为四川省没有发生过泥石流的沟谷和发生过泥石流的沟谷图像. 通过 Arc GIS 软件, 以沟谷流域为单元提取四川沟谷的 DEM (Digital Elevation Model) 和高分一号遥感构建泥石流灾害数据集, 具体信息如表 1 所示. DEM 可以提供泥石流沟谷的地形地势信息, 如沟谷面积、沟谷长度、坡度、地势高差等信息; 高分一号遥感可以更好反映泥石流沟谷的地貌特征, 如植被覆盖、冲刷面积等特征. 本文的目的是通过泥石流沟谷地形地貌特征, 实现对泥石流发生与否的概率预测, 因此需要将提取的泥石流沟谷按照灾害是否发生进行分类, 分为 0 类和 1 类. 0 类代表该沟谷处记录发生过泥石流, 1 类代表该沟谷处未记录发生过泥石流.

表 1 数据信息

数据源	数据来源	数据类型	空间分辨率	通道数
DEM	美国地质勘察局 https://glovis.usgs.gov	数字高程模型	空间分辨率为 30 m	1
高分一号	地理空间数据云 http://www.gscloud.cn/home	遥感影像	空间分辨率为 16 m 按照 DEM 采样为 30 m	4

通过查阅四川省相关监测站的记录及网络新闻报告, 按记载信息并结合 Google Earth 和高德地图, 确定具体发生的村和沟谷, 将发生过泥石流灾害的沟谷图像归为正样本 (为 0). 同理将历史记载中没有发生过泥石流灾害, 且周边村庄没有明显冲刷痕迹和冲积扇的沟谷图像归为负样本 (为 1). 最终提取的四川省所有沟谷样本中, 记录发生过泥石流的沟谷 267 条作为正样本, 未记录发生过泥石流的沟谷 263 条作为负样本, 即 0 类中有 267 条沟谷, 1 类中有 263 条沟谷.

1.2 数据预处理

由于泥石流沟谷图像大小尺寸不同, 为确保图像中沟谷的比例不变且图像大小统一, 需要对每张图像, 根据自身的长宽比例进行缩放, 确保所有图像的宽度一致. 随后, 对长度进行填充处理, 从而确保每个泥石流沟谷图像的长宽尺寸相同.

在模型训练过程中, 为防止过拟合以及得到泛化能力较强的神经网络模型, 通常需要庞大的数据集作为支撑. 然而, 在泥石流灾害研究中, 难以提取到大量复杂的、有标注的泥石流沟谷图像作为训练的数据集. 数据增强可以让有限数据产生更多数据价值, 提高神经网络的学习效果和性能, 因此选择通过数据增强的方式扩充原数据集. 值得注意的是, 并非所有的数据增强方式均可用于 DEM 图像和遥感图像, 由于几何结构特征是本研究的重点因素, 因此所采用的数据增强方式必须确保不破坏原始图像中沟谷的形状特征. 故最终采用的增强方式为: 水平翻转、垂直翻转和水平垂直翻转, 解决原始样本不足的问题.

通过数据增强, 泥石流沟谷图像共有 1 060 张, 其中 0 类有泥石流沟谷图像 534 张, 1 类有泥石流沟谷图像 526 张. 在泥石流灾害研究中, 通常按照 9:1 的比例来划分数据集. 为确保公正性和随机性, 采用随机选取方式将数据集划分为训练集和测试集, 每类图像抽取样本的 10% 作为测试集, 余下样本的 90% 作为训练集, 数据集划分结果如表 2 所示.

表 2 图像数据集数量划分

类别	训练集	测试集	总计
记录发生过泥石流 (0: 正样本)	487	47	534
未记录发生过泥石流 (1: 负样本)	467	59	526
总计	954	106	1 060

2 基于注意力机制的 ResNet50 泥石流沟谷识别网络模型

2.1 ResNet50 网络模型

当层数少时，一般的卷积神经网络学习图像特征不够充分，而当层数多时又容易导致模型过拟合和梯度消失以及训练持续时间过长等问题。残差神经网络（ResNet）于 2015 年被 Kaiming He 等^[9]提出，其特有的网络结构可以克服上述问题。本文采用 ResNet 网络中最具代表性的 ResNet50 网络作为主干网络，网络层数适中，收敛速度快，在分类等任务中应用广泛，其残差结构如图 1 所示。

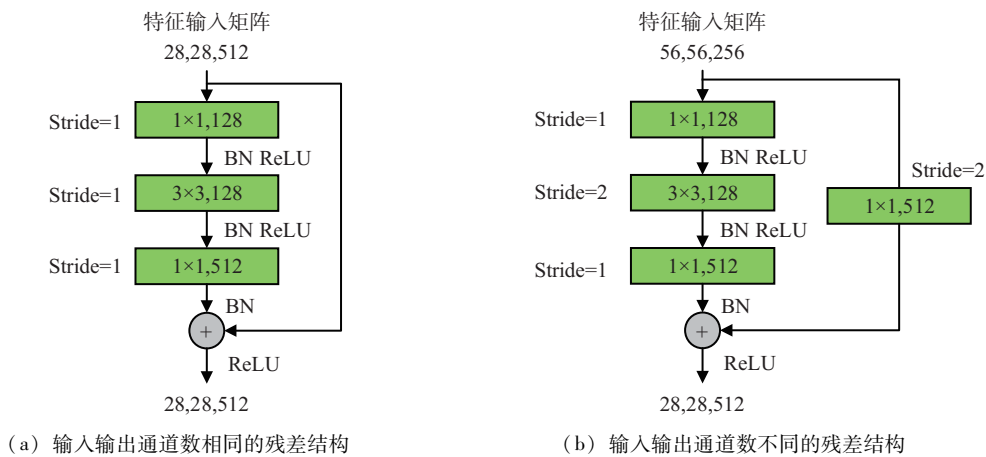


图 1 残差结构模块示意图

ResNet50 网络模型包括 6 个部分。输入样本先经过第 1 部分，包括卷积核尺寸为 7×7 ，步长为 2，padding 为 3，out_channels 为 64 的卷积层、标准化层 BN（Batch Normalization）、激活层 ReLU、卷积核尺寸为 3×3 ，步长为 2，padding 为 1 的最大池化层。然后依次进入第 2 部分到第 5 部分，这 4 个部分主要为残差块，残差块包括恒等残差块（残差块 I）与卷积残差块（残差块 II）。在每个部分，卷积残差块数量相同，个数均为 1，恒等残差块数量不同，个数依次为 2、3、5、2。恒等残差块与卷积残差块均是由卷积核尺寸为 1×1 、 3×3 、 1×1 的卷积层构成，每个卷积层后添加标准化层 BN、激活层 ReLU，增加网络的稳定性。最后进入第 5 部分，包括平均池化层，全连接层，对提取到的特征进行分类。

2.2 CBAM 注意力机制模块

注意力机制是一种模仿人类视觉和认知系统的方法，它允许神经网络在处理输入数据时集中注意力于相关的部分。在卷积神经网络中加入该机制，能够让网络模型着重关注重点区域，而忽略非重点区域。提高模型的性能和泛化能力。本文使用的注意力机制模块是 CBAM，与 SENet（Squeeze-and-Excitation Networks）相比，CBAM 注意机制既关注通道信息又关注空间信息，从而可以取得更好的效果。此外，CBAM 注意机制中添加了最大池化，在一定程度上弥补了平均池化丢失的信息。

它由通道注意力模块（Channel Attention Module）和空间注意力模块（Spatial Attention Module）组成，将二者结合在一起，可以同时获得特征映射的空间域和通道域信息，结构如图 2 所示。特征首先通过通道注意力模块取得新的特征，然后将新的特征作为输入，进入空间注意力模块，最后取得新的特征。

2.3 基于注意力机制的 ResNet50 网络模型

卷积神经网络参数量庞大，在样本量足够的情况下，通常能达到比较理想的性能。但文章现有的泥石流灾害沟谷图像数据集较小，每类样本量只有 500 多个。因此，文章基于注意力机制设计了改进的 Res-

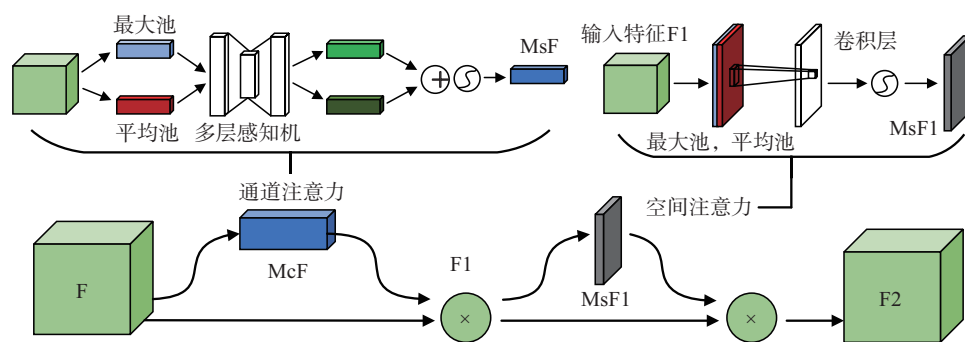


图2 CBAM 注意力机制

Net50 网络模型,用于加强 ResNet50 网络对泥石流灾害沟谷的特征的捕捉,改善 ResNet50 网络在泥石流灾害沟谷图像数据集上的分类性能。

注意力机制的加入可以增强网络提取任务相关的区域的能力,提高关联度高的通道数据特征权重,降低关联度低的通道数据特征权重,从而获取最优的识别能力,加大网络的识别精度。因此本文分别在 ResNet50 网络模型第1部分、第3部分、第5部分的后面加入 CBAM 注意力机制模块,加强网络提取有效信息的力度,提高 ResNet50 网络模型表达能力。改进的 ResNet50 网络结构如图3所示。

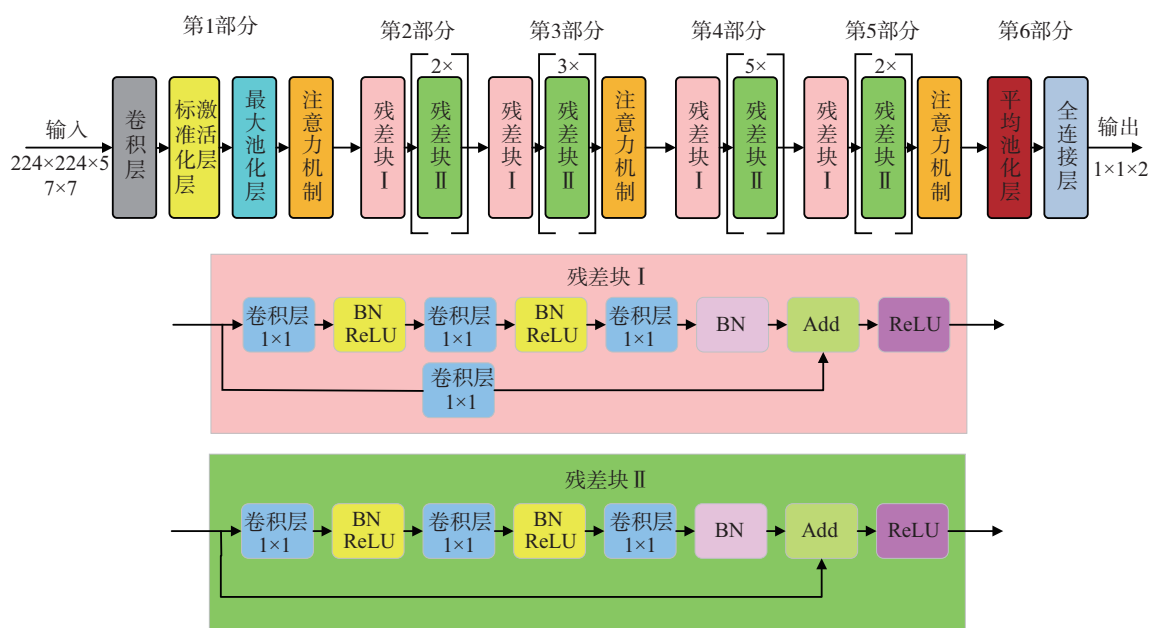


图3 改进的 ResNet50 网络结构

3 试验结果与分析

3.1 训练环境与参数设置

试验设备为 64 位 Microsoft Windows11 操作系统,英特尔 Core (TM) i5-8265U CPU@1.60 GHz 4 核处理器, SODIMM 2 400 MHz 4 GB 内存, NVIDIA GeForce MX130 显卡。本次试验的训练和测试在 Pytorch 上进行,编程语言为 Python。模型训练时采用 12 G 内存 NVIDIA RTX A 2 000 图形处理器 GPU 加速,训练设定如下:训练轮数 Epochs 为 200,迭代次数 Batch_size 为 48,学习率 lr 为 0.001,优化方法为随机梯度下降法 (SGD),分类函数为 softmax 函数,损失函数为 CrossEntropyLoss 函数 (交叉熵损失函数)。

3.2 训练结果

为了验证改进的 ResNet50 网络准确性,在相同的训练环境和参数设置下,对改进的 ResNet50, ResNet50, ResNet18, VGG16 网络分别进行 20 组平行试验,各组试验结果比较相近,故选取其中一组试验结果。各种网络的试验结果对比如表3所示,其中测试集准确率对比如图4所示。从表3可以看出,改进的

ResNet50 网络测试集准确率比 ResNet50、ResNet18、VGG16 网络分别提升了 11.32、19.82 和 10.38 个百分点, 分类准确度达到 83.02%。说明改进的网络可以更好地提取泥石流沟谷图像特征, 获得更优异的准确率。改进的 ResNet50 网络与其他两种网络相比, 整体训练时间有所提升, 训练速度变快, 同时损失有所降低。故改进的 ResNet50 网络在泥石流灾害沟谷图像数据集分类上具有更好的性能。

图 5 是改进的 ResNet50 训练集与测试集准确率变化曲线图。从图中可以看出, 在训练 50 轮后, 准确率达到相对稳定, 上下波动减小, 测试集准确率与训练集准确率变化趋势基本一致, 并收敛到 83.02%。

表 3 各模型性能比较

模型	训练集准确率/%	测试集准确率/%	测试集损失值	训练时间/s
VGG16	100	72.64	1.57	0.86
ResNet18	100	63.20	1.40	0.46
ResNet50	100	71.70	1.92	0.65
改进的 ResNet50	100	83.02	0.98	0.59

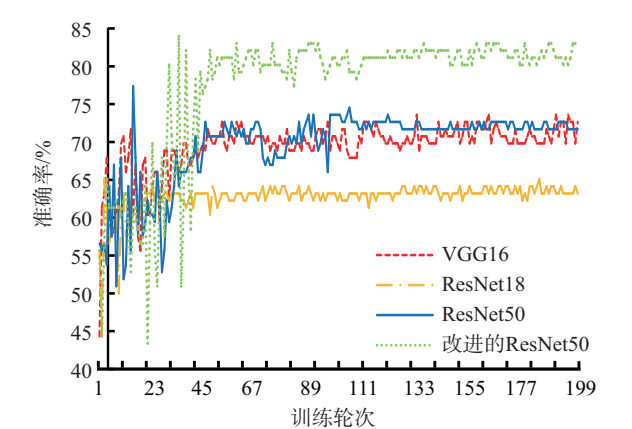


图 4 各个模型准确率变化

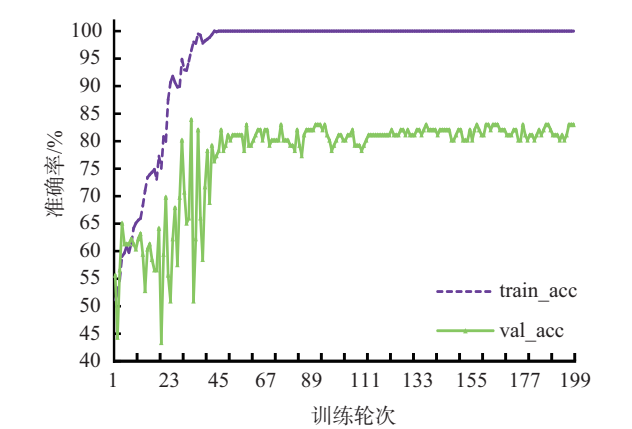


图 5 训练与测试准确率变化

3.3 性能评估

在分类任务中, 常用评价指标有准确率 (Accuracy)、精确率 (Precision)、召回率 (Recall)、 F_1 值 (F_1 -Score)、ACU 值 (Area Under ROC Curve) 等。本文在评估泥石流沟谷 2 分类任务中, 采用以上 5 个评价指标对网络模型进行评估。首先, 给出网络模型在测试集上的混淆矩阵, 如表 4 所示, 其中 TP 表示正样本被预测正确的数量; TN 表示负样本被预测正确的数量; FP 表示负样本被预测为正样本的数量; FN 表示正样本被预测为负样本的数量。

接着, 基于表 4 的混淆矩阵, 可计算出各个模型的评价指标, 结果如表 5 所示。从表 5 可以看出, 改进的 ResNet50 网络模型召回率是 85.11%, 说明该模型识别沟谷发生泥石流的能力要强于其他网络模型。改进的 ResNet50 网络模型精确率是 78.43%, 说明该模型预测沟谷发生泥石流的准确度高于其他网络模型。

表 4 混淆矩阵

模型	TP	FN	FP	TN
VGG16	39	8	21	38
ResNet18	29	18	21	38
ResNet50	31	16	14	45
改进的 ResNet50	40	7	11	48

表 5 各模型评价指标

模型	准确率/%	精确率/%	召回率/%	F_1 值	ACU 值
VGG16	72.64	65.00	82.98	0.729	0.745
ResNet18	63.20	58.00	61.70	0.598	0.691
ResNet50	71.70	68.89	65.96	0.674	0.715
改进的 ResNet50	83.02	78.43	85.11	0.816	0.889

为了进一步验证模型的可靠性,分别绘制出 4 种模型的 ROC (Receiver Operating Characteristic) 曲线,又称接受者操作特征曲线,如图 6 所示.改进的 ResNet50 网络模型与 ResNet50、ResNet18、VGG16 网络模型的 AUC 值分别为 0.889、0.715、0.691、0.745,这说明与其他 3 种网络模型相比,改进后的 ResNet50 网络模型具有更好的识别能力.

4 结 论

鉴于目前泥石流灾害沟谷数据集原始样本不足、网络模型训练效果差等问题,本文通过网络提取沟谷图像,采用数据增强的方式扩充图像数量,并提出在 ResNet50 网络模型的基础上加入 CBAM

注意力机制模块,对泥石流灾害沟谷进行识别.试验结果表明,CBAM 注意力机制运用到 ResNet50 网络模型中,可以更有效获取模型的位置信息和空间信息,也可以更好地实现泥石流灾害沟谷数据集识别. ResNet50 网络模型在训练集中的分类准确率为 83.02%,与 VGG16、ResNet18、ResNet50 网络模型相比,其分类准确率分别提升了 10.38、19.82 和 11.32 个百分点.在以后的工作中,将会增加泥石流沟谷数据,在 DEM 数据和高分一号遥感数据的基础上,加入一通道的岩性、土壤、植被数据,将这些特征信息充分运用至泥石流灾害沟谷分类工作中,提升模型分类准确率.

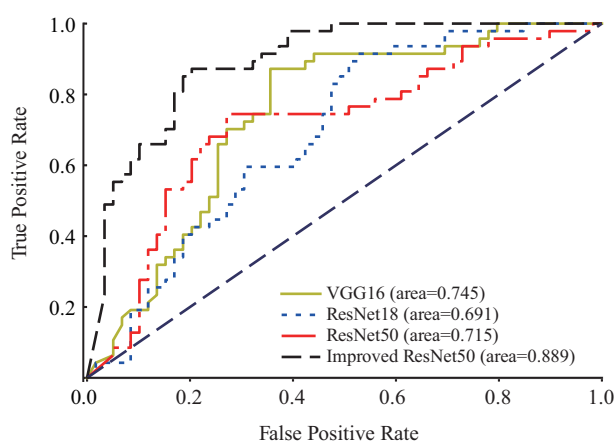


图 6 ROC 曲线图

[参考文献]

- [1] 张楠,方志伟,韩笑,等.近年来我国泥石流灾害时空分布规律及成因分析[J].地学前缘,2018,25(2):299-308.
- [2] 槐永波,叶胜华,赵其苏.汶川地震10年来北川县地质灾害发育规律分析[J].人民长江,2021,52(增刊2):83-87.
- [3] 四川省情网.四川地质灾害主要类型、成因及影响[EB/OL].(2018-10-17)[2023-12-01].http://scdfz.sc.gov.cn/whzh/slzc1/content_16788.
- [4] 李婧玥,孙永彬,王洗,等.北京市延庆区石窑村南地沟泥石流特征及发展趋势预测[J].矿产勘查,2021,12(6):1462-1471.
- [5] 魏永明,郭华东,陈玉,等.甘肃省舟曲特大泥石流的遥感影像特征及古泥石流遥感识别的地质意义[J].第四纪研究,2014,34(2):325-335.
- [6] 苏鹏程,倪长健,孔纪名,等.区域泥石流危险度评价的影响因子识别[J].水土保持通报,2009,29(1):128-132.
- [7] 李秀珍,孔纪名,李朝凤.多分类支持向量机在泥石流危险性区划中的应用[J].水土保持通报,2010,30(5):128-133.
- [8] 周飞燕,金林鹏,董军.卷积神经网络研究综述[J].计算机学报,2017,40(6):1229-1251.
- [9] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C]//IEEE. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) CoRR. Las Vegas: IEEE, 2016: 770-778.
- [10] 刘坤香,王保云,徐繁树,等.基于残差注意力机制的泥石流沟谷识别[J].中国地质灾害与防治学报,2022,33(6):134-141.
- [11] 刘源洞,何茂征,黄益斌,等.基于 ResNet50 和改进注意力机制的船舶识别算法[EB/OL].(2023-09-05)[2023-12-01].<https://link.cnki.net/urlid/51.1307.TP.20230904.1509.012>.
- [12] 王德雨,石伟,张元良.一种融合 CBAM 注意力机制和金字塔卷积的经济鱼类识别方法[J].江苏海洋大学学报(自然科学版),2023,32(1):73-80.

(下转第 128 页)

- differential SAR interferograms [J]. IEEE Transactions on geoscience and remote sensing, 2002, 40 (11): 2375-2383.
- [14] 李凡, 李素敏, 杨渊, 等. 基于时序 InSAR 的沙湾大沟滑坡型泥石流发育特征研究 [J]. 地球物理学进展, 2023, 38 (2): 32-541.

Study on the Applicability of Ascending and Descending Orbit Sentinel-1A Data in Geological Hazard Monitoring in High Mountainous Canyon Areas

CHEN Kuayue¹, WANG Baoyun^{1,2}, WANG Ting³

- (1. School of Mathematics, Yunnan Normal University, Kunming, Yunnan, China 650500;
2. Yunnan Key Laboratory of Modern Analytical Mathematics and Applications, Kunming, Yunnan, China 650500;
3. School of Foreign Languages & Literature, Yunnan Normal University, Kunming, Yunnan, China 650500)

Abstract: In response to the difficulty in selecting the appropriate ascending or descending orbit Sentinel-1A data for geological hazard monitoring in high mountainous canyon areas using InSAR technology, this article uses SBAS-InSAR technology and combines slope and aspect data to calculate the deformation values of the vertical slope direction and along the slope direction in the study area. The applicability of the lifting rail Sentinel-1A data source for different slope directions in geological hazard deformation monitoring is studied. The results indicate that: 1) Sentinel-1A ascending orbit data are more suitable for monitoring deformation information on north-facing and east-facing slopes, while descending orbit data are more suitable for monitoring deformation information on south-facing and west-facing slopes. 2) Deformation rates along the slope are greater than vertical slope deformation rates on north-facing slopes; the difference between along-slope and vertical slope deformation rates is not significant on south-facing slopes; vertical slope deformation rates are greater than along-slope deformation rates on east-facing slopes; both vertical and along-slope deformations are significant on west-facing slopes. Therefore, combining ascending and descending orbit data can address the issue of insufficient accuracy of monitoring results for certain areas or specific angles using a single orbit, and can provide more comprehensive and accurate deformation information.

Key words: small baseline subset interferometric synthetic aperture radar (SBAS-InSAR); Sentinel-1A data source; Two-dimensional deformation decomposition; deformation monitoring; geological hazard

(责任编辑: 陈伟超)

(上接第 119 页)

Debris Flow Gully Identification Based on Improved ResNet50 Network

LIU Qiuyu¹, WANG Baoyun^{1,2}

- (1. School of Mathematics, Yunnan Normal University, Kunming, Yunnan, China 650500;
2. Yunnan Key Laboratory of Modern Analytical Mathematics and Applications,
Yunnan Normal University, Kunming, Yunnan, China 650500)

Abstract: Traditional convolutional neural networks have problems such as low accuracy, poor image feature extraction and blurred edges when used for landslide disaster valley image classification. This paper improves the ResNet50 network by adding a CBAM attention mechanism module before some residual blocks of the ResNet50 network, which enables it to have higher performance and accuracy and accurately capture the terrain and landforms in landslide disaster valley images. The experimental results show that the improved ResNet50 network achieves a classification accuracy of 83.02% for landslide disaster valley images, which improves its classification performance by 11.32 percentage points compared to the ResNet50 network. Moreover, its accuracy, recall rate, precision rate, F1 value, and AUC value are better than those of the ResNet50 network and other deep learning recognition algorithms.

Key words: attention mechanism; convolutional neural network; debris flow disaster; ResNet50

(责任编辑: 陈伟超)