

张艺帆, 席楠, 杨天青, 等. 2025. 基于 VGG16 卷积神经网络模型的川滇地区谱比法场地分类研究. 中国地震, 41(1): 46~57.

基于 VGG16 卷积神经网络模型的 川滇地区谱比法场地分类研究

张艺帆^{1,2)} 席楠³⁾ 杨天青³⁾ 姜立新³⁾

1) 中国地震局地震预测研究所, 北京 100036

2) 河南省地震局, 郑州 450016

3) 中国地震台网中心, 北京 100045

摘要 传统的谱比法在地震场地分类中存在一定的主观性和不确定性, 引入机器学习算法可以通过大量的样本数据和特征提取方法, 提高分类方法的客观性和准确性, 为地震工程设计和地震危险性评估提供更可靠的依据。本研究利用川滇地区历史地震的强震台事件波形数据, 计算不同场地类别强震动记录的水平和竖向速度反应谱值获取谱比曲线, 分析了其在卓越周期、峰值和整体幅值上的特征。引入深度学习算法, 整理并生成了川滇地区强震动台站谱比曲线样本集, 并将其用作为 VGG16 卷积神经网络模型的输入数据, 训练得到了 NEHRP 下三类标准场地的概率分布模型。通过特征提取和交叉验证算法, 提高了模型的精度和泛化能力。研究所建立的场地分类模型在验证结果中表现良好, 对于谱比法场地分类方法的改进和应用具有一定的参考价值。

关键词: 地震学 场地分类 谱比法 深度学习

[文章编号] 1001-4683(2025)01-0046-12 [中图分类号] P315 [文献标识码] A

0 引言

大量地震灾害研究表明, 场地条件对地震动有重大影响(郭宗斌等, 2021)。场地分类是工程抗震和地震危险性研究的重要内容, 地震发生时, 不同场地对来自地下深处的地震动会产生不同的影响, 场地分类的本质就是反应这种影响的差别。确定一个区域的场地类别, 对于确定该区域抗震设计参数, 提升地震危险性评估水平和提高烈度速报准确度等方面具有重要意义。

我国地震台网建设初期, 数字强震台站数量达数千个, 部分台站所处地区偏远, 通过钻孔勘测确定其场地分类比较困难(温瑞智等, 2015)。随着预警工程建设, 目前我国已形成由 18000 多个强震仪和烈度仪站点构成的强震动观测网络, 其中有超过 12000 多个站点并未通

[收稿日期] 2024-04-18 [修定日期] 2024-12-09

[项目类别] 地震科技星火计划项目(XH23052C)资助

[作者简介] 张艺帆, 女, 1998 年生, 硕士研究生, 主要从事地震场地分类研究。E-mail: 15838120611@163.com

席楠, 通讯作者, 男, 1986 年生, 高级工程师, 主要从事地震烈度速报与地震预警研究。E-mail: xinan314@seis.ac.cn

过钻孔勘测确定其场地类别,其中主要为烈度仪台站。确定台站场地条件,对于确定台站的地震动放大系数、减少地震动数据的应用误差、震前抗震设防和震后应急决策具有重要意义。对数量庞大的台站进行钻孔勘测来划分场地效率低且不经济,因此如何利用更加简便高效的方法来评估台站场地条件成为一个重要课题。

谱比法作为场地分类方法中的常用方法,最早由日本学者 Nakamura 等(1983)提出,通过假设地面运动的竖向分量比水平分量包含更多的地面运动信息,由此发现并论证了地脉动的竖向和横向傅里叶谱比值,其可以反应不同场地的软硬程度并估计场地特征。Yamazaki 等(1997)将该方法整体延伸至对强震动记录的研究,得到同一场地的水平和竖向傅里叶谱比值基本一致的结论。因傅里叶谱具有波动大,尖刺明显的特点,Zhao 等(2006)使用阻尼比为 5%的水平与竖向速度反应谱代替傅里叶谱进行了场地特征研究和场地类别划分,方法可行性和精度均得到较好的验证。Ji 等(2017)采用日本 KiK-net 台站数据,按照中国抗震设防规范分为三种场地类别,并得到这三类场地的 H/V 反应谱标准谱比曲线,将其与中国强震台站的谱比曲线进行匹配,得到中国国家强震动观测网络系统台站的站点分类。针对场地波速资料缺乏的情况,Xie 等(2022)提出一种基于岩性的剪切波速估算方法,将其应用于强震动台站 V_{s30} 的估计和场地分类。

机器学习在波形匹配和特征提取方面的改进中效果较为出色。在早期研究中,Gibbons 等(2006)、Peng 等(2009)多基于波形匹配的方法使用附近震源的波形相似性来探测未被识别的地震。近年来,根据地震波形间的相似性进行无监督学习,可实现波形特征提取和波形匹配功能(Wu 等,2019; 凌里杨等,2022)。Ji 等(2023)使用深度卷积神经网络(DCNN)对日本 1649 个台站进行分类,得到的 C、D 和 E 三类分类准确性高于经验谱比法。因此,本研究将机器学习方法应用于对 H/V 谱比法波形特征提取和模型建立过程,结合卷积神经网络和交叉检验机制,提高传统谱比法的精度和效率。

1 H/V 谱比法场地分类

1.1 场地分类标准

传统场地分类标准以地下 30m 深度的土层平均剪切波速(V_{s30})为分类依据进行划分,目前采用 V_{s30} 作为分类指标的国家包括美国、日本及欧洲大部分国家(袁吉,2021; 薄景山等,2021; 王竞等,2022)。美国抗震规范(NEHRP)(Building Seismic Safety Council, 2020)采用 V_{s30} 单一指标进行分类,基于该指标将场地分为 A、B、C、D 和 E 五类,具体分类如表 1 所示。

表 1 使用 NEHRP 标准的 V_{s30} 与场地类别关系(袁吉,2021)

场地类别	V_{s30} 范围/($\text{m} \cdot \text{s}^{-1}$)	描述
A	$V_{s30} > 1500$	坚硬岩石
B	$760 < V_{s30} \leq 1500$	岩石
C	$360 \leq V_{s30} \leq 760$	软岩石或致密的土
D	$180 \leq V_{s30} \leq 360$	硬土
E	$V_{s30} < 180$	软土

表 2 GB50011—2010《建筑抗震设计规范》中的场地分类方法

等效剪切波速 $V_s/(m \cdot s^{-1})$	土层覆盖厚度 H/m				
	I0 类	I1 类	II 类	III 类	IV 类
$V_s > 800$	0				
$500 < V_s \leq 800$		0			
$250 < V_s \leq 500$		< 5	≥ 5		
$150 < V_s \leq 2500$		< 3	$3 \sim 50$	> 50	
$V_s < 150$		< 3	$3 \sim 15$	$15 \sim 80$	> 80

我国按照 GB 50011—2010《建筑抗震设计规范》(2016 年版)局部修订条文(2024 年版)(住房和城乡建设部,2024)规定的场地分类标准进行分类,该场地分类方法将等效剪切波速和覆盖土层厚度两个指标共同作为划分标准。其中,等效剪切波速为 20m 和剪切波速达到 500m/s 时深度之间较小深度的等效剪切波速,依据该值将场地分为 I0、I1、II、III 和 IV 类,具体分类情况如表 2 所示。

由于缺乏部分台站的土层厚度数据,无法准确计算出台站的等效剪切波速,本文以 NEHRP 标准对台站场地类别进行判定,同时该标准也更易与国际接轨,便于未来将训练模型应用于其他地区。

本文收集了川滇地区 2013—2022 年 412 个强震动台站共计 5769 条地震波原始三分量数据,地震事件和台站分布情况如图 1 所示。对于强震动台站 V_{s30} 数据的获取参考 Xie 等(2022)研究中的川滇地区台站的场地钻孔和波速测试资料,对于钻孔深度不足 30m 的台站的剪切波速数据推演至 30m,根据得到的剪切波速资料按照 NEHRP 标准对台站进行分类,与 Xie 等(2022)的分类结果进行对照后,得到 B 类台站 45 个,C 类台站 178 个,D 类台站 189 个,如表 3 所示。分类标准中的 E 类场地为软弱场地,根据我国强震台站选址要求和工程实践,一般不在该类场地上建设强震动台站,因此本文中并未考虑该类场地。上述强震动台站均为基本站,由于 B 类台站的地震动数据数量与另外两类台站的地震动数据数量差距明显,为避免在后续数据训练过程中出现因数据分布不均而产生的误差,选取部分强震事件中的基准站数据加入 B 类数据集。基准站主要建立在地震活动性较强的断裂带附近,是预警台网的核心台站,我国的基准站均建立在基岩、页岩、花岗岩、风化岩等硬质土层上,因此可直接归为 B 类场地。

1.2 强震记录选取与分析

强震动数据中普遍存在噪声污染,为了降低高频和低频噪声对记录的影响,需要对数据进行滤波处理(周宝峰等,2012;姚鑫鑫等,2022)。在噪声处理中,选择适当的滤波器对强震记录进行滤波至关重要,本文选择了 Butterworth 四阶非因果滤波器,其在移除强震记录噪声方面效果良好。滤波的截止频率范围选择 0.2~50Hz,以确保本文的研究在 0.02~5s 周期范围内得到有效的结果。该滤波方法可以有效地降低高频和低频噪声的干扰,提高数据的质量和可靠性。

根据对土层非线性反应的研究结果,当地震动水平增大时,土体的剪切模量降低,阻尼增强,场地的放大效应减小。通过移除 PGA 超过 100gal 的强震记录数据,通常可以有效地

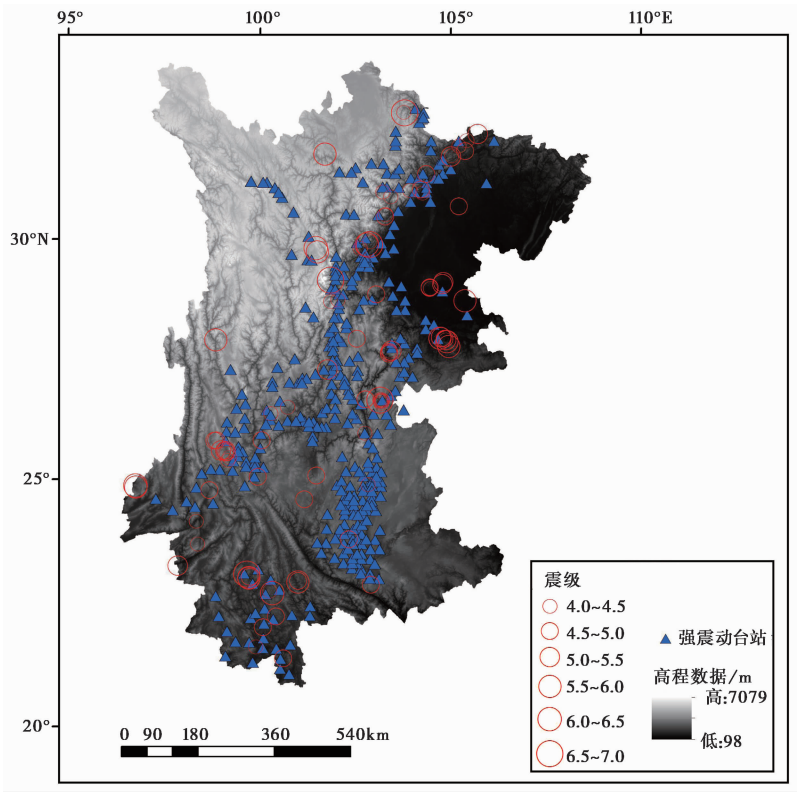


图 1 地震事件和台站分布情况

表 3 台站原始数据分类结果

场地分类	台站数量/个	记录数量/条	台站比例/%	记录比例/%
B	45	378	10.9	6.5
C	178	2853	43.2	49.4
D	189	2539	45.9	44.1

消除土层非线性效应对结果的干扰(温瑞智等,2015; 陈永新,2015)。本文对强震数据主动降低重采样率,由原始的 200 次/s 降低为 50 次/s。通过降低重采样率,可以有效减少数据中的细微波动,使得波形数据在保留绝大部分信息的情况下更加平滑和简化,提高数据处理效率,减少计算和存储的负担。

经过对强震记录的处理和筛选,最终确定了川滇地区共 503 个台站,4938 条强震动记录作为训练样本,台站分类情况如表 4 所示。

对于原始波形数据计算其三分量的 5% 阻尼比的速度反应谱值,采样率为 50 次/s,得到每组数据的 H/V 反应谱比

$$\frac{H}{V} = \frac{\sqrt{EW \times NS}}{UD} \tag{1}$$

其中, EW 、 NS 、 UD 分别为东西、南北、竖向方向记录的 5% 阻尼比速度反应谱比值; H/V 即水平/竖向方向谱比(温瑞智等,2015)。

表 4 预处理后台站数据分类结果

场地分类	台站数量/个	记录数量/条	台站比例/%	记录比例/%
B	212	1911	42.1	26.5
C	135	1914	26.8	38.8
D	156	1713	31.1	34.7

同时,依据场地类别对得到的谱比曲线进行分类,得到三类的平均谱比曲线图,如图 2 所示,图中的横轴以对数形式显示,横轴范围限制在 0.02~5s 之间。从图中可以看出,B 类场地的曲线较为平坦; C 类有一个较为明显的峰值,其卓越周期为 0.33~0.67s 左右, H/V 幅值为 1.6 左右; D 类场地曲线有较为明显的两个峰值,第一个峰值周期与 C 类周期大致重合,整体 H/V 幅值为 2.0 左右。三类谱比曲线的明显差异体现了不同类型场地对于地震动的反应: B 类场地较为坚硬,对地震动的放大效应较小, C 类和 D 类对地震动的放大效应依次增大,并且, D 类出现了明显的双波峰现象。

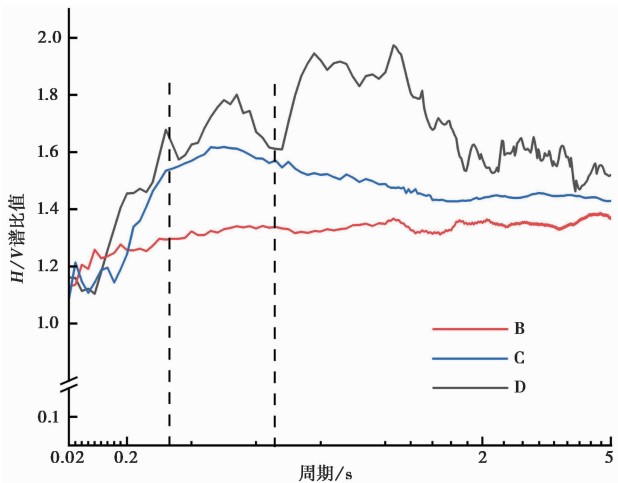


图 2 三类场地平均谱比

2 基于机器学习算法改进分类结果

2.1 图像预处理

对于原始的地震波信号进行滤波处理、加速度反应谱计算和 H/V 谱计算后得到一系列的谱图。这些谱图作为模型训练的样本,输入时要将图的大小和通道统一至神经网络所兼容的状态。在模型的数据生成器模块中,将输入的图像进行归一化处理,像素缩放至 0~1 范围内,尺寸作相应调整。

由于数据集本身的数量不多,文中使用图像增强技术来增加数据的多样性,对图像进行旋转、平移、缩放和翻转等处理。同时,将原始分类设置转化为多分类的形式,使用 one-hot 编码表示。将处理好的数据和标签分批输入到模型中进行训练,这样可以避免过拟合问题,并提高模型性能和泛化能力。

2.2 模型构建

卷积神经网络是一种具有深度结构的前馈神经网络,其基本组成包括输入层、隐含层、输出层。其中,隐含层包括卷积层、池化层和全连接层等,部分网络还会加入激活函数。VGG 系列网络在卷积神经网络中是较为成熟的网络结构,由 Simonyan 等(2015)在 Alex Net 的基础上进一步开发出来。VGG 模型结构简洁,由连续的 3×3 小卷积核和池化层构建而成,对于给定的输入图片,使用更少的卷积核多次扫描图片并组合特征的形式优于单纯使用一个大卷积核。因此,VGG 模型不仅可减少参数的数量,且能够使决策函数更具判别性,在一定程度上提升了神经网络的效果。在图像分类、目标检测和语义分割上都取得较好的效果(周少辉等,2021; 韩红,2022)。实验表明,VGG 模型层数在 16~19 的范围内效果最好,本文采用网络层数为 16 的 VGG16 模型进行迁移学习。

引入 VGG16 预训练模型进行迁移学习,冻结预训练权重,之后根据数据特性设置相应的模型结构,具体结构如图 3 所示。

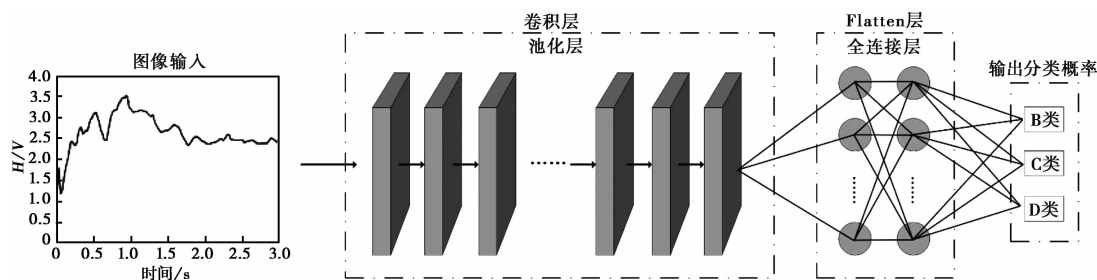


图 3 卷积神经网络模型结构

卷积层对输入的数据进行特征分析和提取,VGG16 网络使用连续的 3×3 小卷积核串联为一个感受野更大的卷积核,其非线性能力更强,可以在不增加计算量的情况下增加感受野,提高网络的性能,增强特征提取的能力;池化层对数据特征进行压缩,剔除对分类贡献较少的部分,降低模型的过拟合率。在每个卷积层之后加入 Relu 激活函数,为模型训练加入非线性特征,同时缓解梯度消失问题,提高模型分类性能。

Flatten 层将图像特征张开平铺为一维数据,输入到全连接层中,全连接层中的神经元通过训练过程中的正向和反向传播,学习到最优的权重和偏置值,将输入数据转化为更高级别的特征表示。全连接层后加入 softmax 激活函数,将输出转化为 B 类、C 类和 D 类的概率分布,用于分类判别。

2.3 模型性能

对模型使用 k 折交叉验证方法进行训练,将经过数据增强的完整数据集划分为 k 个子集,每次使用不同的 $k-1$ 个子集作为训练集,1 个子集作为验证集进行循环训练和验证,最终将 k 次结果的平均值作为模型的性能评估指标。在本研究中, k 值设置为 10,每次循环中数据集通过神经网络进行前向传播和反向传播次数为 40 次,共 400 次训练。

采用训练精度、验证精度、训练损失和验证损失来评估模型性能。模型在训练集上的平均精度为 0.736,在验证集上的平均精度为 0.756,在训练集上的平均损失为 0.58,在验证集上的平均损失为 0.49,上述模型性能参数波动曲线图如图 4 所示。

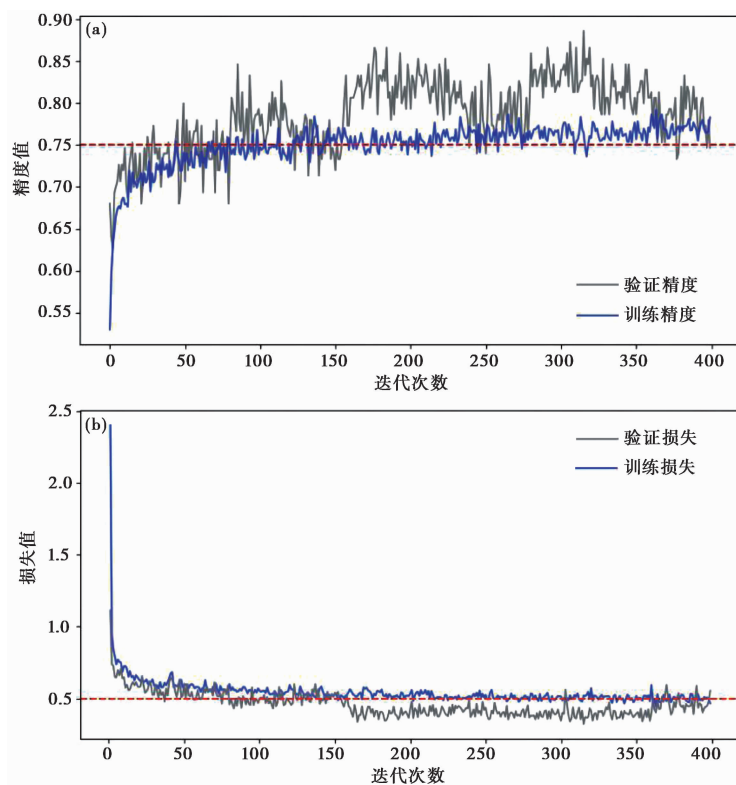


图4 精度函数(a)和损失函数(b)变化曲线

由图4可以看出,模型的训练精度在训练中稳步提升,在350次之后趋于平缓;模型的验证精度在0.7~0.8内存在波动,在350次之后开始出现波动下降的趋势,说明可能出现过拟合现象,验证精度的均值略高于训练精度,这可能是由于验证集的样本数量小于训练集,故在验证集上的泛化能力较低,精度略高。模型的训练损失在训练初期都下降明显,且随着迭代次数增加,逐步稳定在0.5左右;模型验证损失也有所下降,但没有训练损失明显,且在300次迭代之后,开始出现小幅波动,在350次之后有上升趋势,说明迭代400次是合理的取值,再增加迭代次数可能会出现损失值增加,模型过拟合等现象。模型整体的训练集和验证集的准确率高,损失精度低,并且能够在训练中得到收敛,说明该模型的性能较好,可以使数据集得到较好的分类。

对模型训练集的预测结果与 NEHRP 标准结果进行混淆矩阵分析(图5),对混淆矩阵进行精度评价,选择的指标为正确率、精确率、召回率和 F1 值。混淆矩阵将预测值和实际值之间的结果关系分为真正值(true positive,简称 TP)、真负值(true negative,简称 TN)、假正值(false positive,简称 FP)和假负值(false negative,简称 FN)四种结果。正确率(Precision)为模型识别为正例的样本中,真正为正类的样本所占的比例,一般情况下,正确率越高,说明模型的效果越好。精确率(Accuracy)为模型识别正确的个数占样本总个数的比值,即传统意义上模型的精度。召回率(Recall)表示模型正确识别出为正类的样本的数量占总的正类样本数量的比值,一般情况下,召回率越高,说明有更多的正类样本被模型预测正确,模型的效

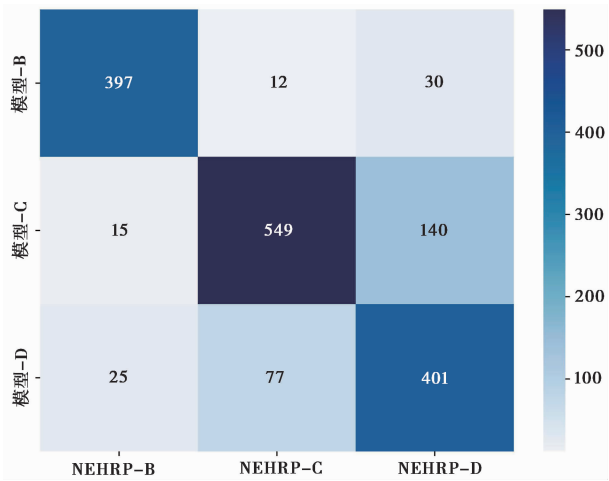


图 5 模型训练集混淆矩阵

果越好。F1 值是模型准确率和召回率的加权平均,值越大意味着模型的效果越好。正确率、精确率、召回率和 F1 值的计算公式见式(2)~(5)

$$\text{Precision} = TP / (TP + FP) \tag{2}$$

$$\text{Accuracy} = (TP + TN) / (TP + FN + FP + TN) \tag{3}$$

$$\text{Recall} = TP / (TP + FN) \tag{4}$$

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{5}$$

分析结果(表 5 和表 6)可以看到,模型的整体训练效果较好,精确率和正确率均在 0.7 以上。B 类的整体精度最高,主要是因为数据集制作过程中加入了具有明确场地特征的基准台站的地震波数据,且通过各类平均曲线图(图 2)所示,B 类的整体曲线与 C 类和 D 类具有明显的差别,因此在模型训练过程中可以被较好地提取和分类出来。

表 5 模型训练集预测分类结果分布

	NEHRP 标准 B 类	NEHRP 标准 C 类	NEHRP 标准 D 类
模型预测 B 类	397	12	30
模型预测 C 类	15	549	140
模型预测 D 类	25	77	401

表 6 模型训练集预测结果分析

场地类别	精确率	正确率	召回率	F1 值
B	0.909	0.943	0.904	0.907
C	0.861	0.847	0.780	0.818
D	0.702	0.832	0.797	0.747

分类过程中 C 类和 D 类存在一定的混淆情况,针对这一问题进行详细分析,发现存在部分 NEHRP 标准 C 类的同一站点的多条数据均被错误分到了 D 类,同样的,对于 D 类站点也存在这样的现象。图 6 为部分混淆的站点的分布情况,其中图 6(b)和图 6(c)为局部放大图。图 6(b)中根据地形数据可以看出,部分台站位于河谷和谷底等地势较低地带,部分位于东南的四川盆地地区,这些台站容易受到盆地效应的影响,使得台站记录的地震动被放大,地震动中的中低频成分强度增大,地震动明显高于盆地周边地区。图 6(c)中的台站同样位于河谷地带,地震动经河谷的松软土层会放大明显,计算得到的谱比曲线对应的场地土层会比实际要偏软。地形的复杂性和盆地效应的影响是造成这些区域进行场地分类训练时,C 类和 D 类分类混淆的一个重要原因。

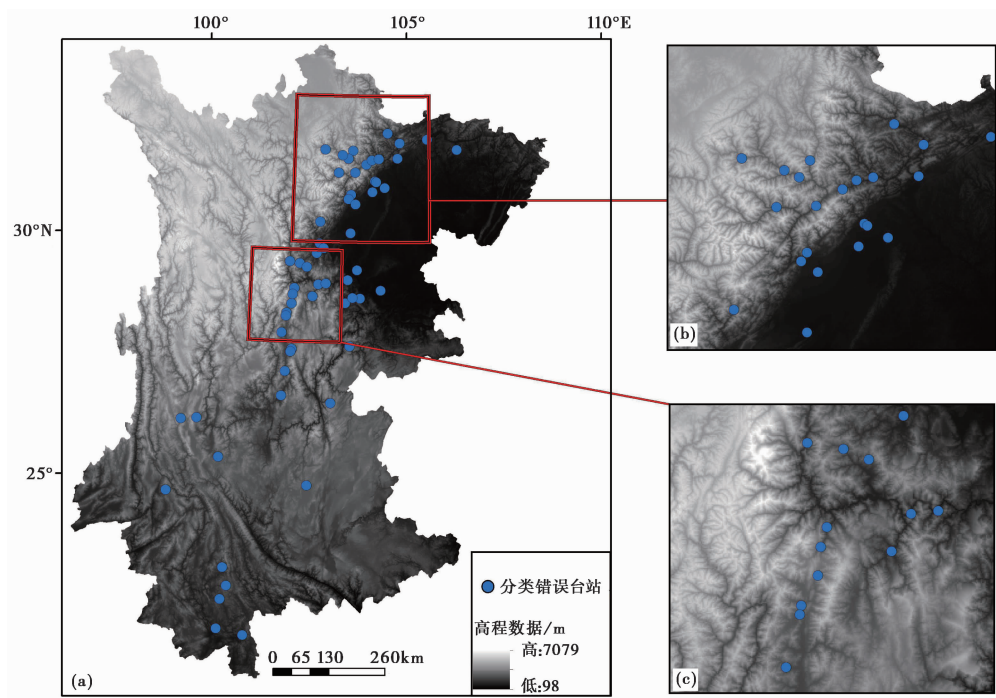


图 6 错误分类台站分布情况

2.4 方法验证

为了验证模型分类的有效性,选取未参与训练的 2008 年汶川地震中部分强震动台站的地震动记录作为验证数据集,经过相同的数据处理之后输入到模型中。以台站实际 V_{s30} 数据得到的分类结果为准,与模型的分类结果进行对比分析,得到模型预测的分类精确度为 0.746,混淆矩阵结果如表 7 所示。

根据式(2)~(5),计算该混淆矩阵的精确率、正确率、召回率和 F1 值,结果如表 8 所示。可以看出,按照 NEHRP 标准,分类正确率和精确率均在 65%~75%之间。其中 B 类场地的预测精度最高,由于部分 C 类划分到了 D 类,使得 C 类的正确率较低,说明验证集中同样存在 C 类和 D 类之间的判别重合现象。从实际情况来看,部分分类错误的台站同样位于河谷、盆地等地形内,受到盆地和地形复杂性的影响,C 类和 D 类场地的界限相对模糊,数据的离

表 7 混淆矩阵分析结果

NEHRP 分类	模型预测 B 类	模型预测 C 类	模型预测 D 类
NEHRP 标准 B 类	5	2	0
NEHRP 标准 C 类	2	17	11
NEHRP 标准 D 类	1	6	22

表 8 预测结果分析

场地类别	精确率	正确率	召回率	F1 值
B	0.897	0.714	0.625	0.667
C	0.677	0.567	0.680	0.618
D	0.667	0.758	0.709	0.709

散性较高,依靠曲线判断容易产生误判,可以采用地形地质数据进行辅助判断。整体验证结果说明该模型具有一定的泛化性能,可以在陌生数据集上得到较好的分类结果。

3 结论

本文采用川滇地区 2013—2022 年强震台地震动数据,对不同场地类别的强震动记录计算谱比曲线,得到三种场地类别的平均谱比曲线,分析其在卓越周期、峰值和整体幅值上的特征。

通过将深度学习算法引入谱比法的场地分类方法中,本研究生成了川滇地区强震动台站谱比曲线样本集,并将其作为 VGG16 网络模型的输入数据。通过卷积层、池化层、Flatten 层和全连接层等结构对输入数据进行特征提取,通过交叉验证算法对模型进行迭代训练,从而提高了模型的整体精度并降低了损失。最终得到的模型精度达到 0.722,损失降至 0.58,在陌生数据集上的验证结果也表现良好,精度达到 0.746,这表明模型成功地提取了谱比曲线的特征,并且具有较好的分类效果和泛化能力。

传统谱比法主要依靠谱型匹配和曲线特征参数如卓越周期、峰值等进行分类,分类指标较为宏观。目前学者的研究中,有代表性的如 Ghasemi 等(2009)提出的基于斯皮尔曼秩相关系数的分类方法,依靠曲线首尾和卓越周期的匹配程度进行分类;温瑞智等(2015)结合卓越周期、谱比峰值和斯皮尔曼秩相关系数提出一种经验方法,对日本 Kik-net 台站进行分类。将本文的分类精度与结果进行对比,本研究得到的精度结果高于该经验方法的分类精度,与使用斯皮尔曼秩相关系数的分类结果持平,分类的稳定性高于斯皮尔曼秩相关系数方法,说明本文的研究方法对于精度提高是有效的。将机器学习算法引入谱比法的判别过程,通过大样本量的训练和分析,可以获取到曲线中隐藏的数据特征,有助于提升分类准确率,这对于谱比法场地分类方法的改进和应用具有参考价值。同时,由于谱比法自身存在缺陷,如竖向地震动不被放大的假设在一些场地条件下不成立,场地条件的复杂性和场地效应的非线性等,这些均会影响分类的准确率。

参考文献

薄景山,李琪,孙强强,等. 2021. 场地分类研究现状及有关问题的讨论. 自然灾害学报,30(3):1~13.

- 陈永新. 2015. 基于 KiK-net 强震动记录对不同土层条件下场地非线性反应的研究. 硕士学位论文. 北京: 中国地震局地球物理研究所.
- 郭宗斌, 陈其峰, 冯恩国, 等. 2021. 工程地震场地分类方法研究进展. 内陆地震, **35**(4): 346~353.
- 韩红. 2022. 基于深度卷积神经网络的地震相识别. 硕士学位论文. 石家庄: 河北地质大学.
- 凌里杨, 徐天吉, 冯博, 等. 2022. 基于深度学习的多波地震信号智能匹配方法与应用. 石油地球物理勘探, **57**(4): 768~776.
- 王竞, 王世元, 宴金旭, 等. 2022. 中外场地分类研究现状及其应用讨论. 世界地震工程, **38**(3): 236~245.
- 温瑞智, 冀昆, 任叶飞, 等. 2015. 基于谱比法的我国强震台站场地分类. 岩石力学与工程学报, **34**(6): 1236~1241.
- 姚鑫鑫, 任叶飞, 岸田忠大, 等. 2022. 强震动记录的数据处理流程: 去噪滤波. 工程力学, **39**(增刊 1): 320~329.
- 袁吉. 2021. 基于日本 KiK-net 强震动数据的场地效应研究. 硕士学位论文. 哈尔滨: 中国地震局工程力学研究所.
- 周宝峰, 温瑞智, 谢礼立. 2012. 非因果滤波器在强震数据处理中的应用. 地震工程与工程振动, **32**(2): 25~34.
- 周少辉, 蒋海昆, 李健, 等. 2021. 基于深度学习的地震事件分类识别——以山东地震台网记录为例. 地震地质, **43**(3): 663~676.
- 住房和城乡建设部. 2024. GB 50011—2010 建筑抗震设计规范(2016 年版)局部修订条文(2024 年版). 工程建设标准化, **2024**(9): 72~79.
- Building Seismic Safety Council(BSSC). 2020. FEMA P-2082 NEHRP Recommended Seismic Provisions for New Buildings and Other Structures. Washington: Building Seismic Safety Council, Federal Emergency Management Agency.
- Ghasemi H, Zare M, Fukushima Y, et al. 2009. Applying empirical methods in site classification, using response spectral ratio (H/V): A case study on Iranian strong motion network (ISMN). Soil Dyn Earthq Eng, **29**(1): 121~132.
- Gibbons S J, Ringdal F. 2006. The detection of low magnitude seismic events using array-based waveform correlation. Geophys J Int, **165**(1): 149~166.
- Ji K, Ren Y F, Wen R Z. 2017. Site classification for National Strong Motion Observation Network System (NSMONS) stations in China using an empirical H/V spectral ratio method. J Asian Earth Sci, **147**: 79~94.
- Ji K, Zhu C B, Yaghmaei-Sabegh S, et al. 2023. Site classification using deep-learning-based image recognition techniques. Earthq Eng Struct Dyn, **52**(8): 2323~2338.
- Nakamura Y, Saito A. 1983. Estimations of amplification characteristics of surface ground and PGA using strong motion records[J]. Proc. 17th JSCE.
- Peng Z G, Zhao P. 2009. Migration of early aftershocks following the 2004 Parkfield earthquake. Nat Geosci, **2**(12): 877~881.
- Simonyan K, Zisserman A. 2015. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv: 1409.1556.
- Wu Y, Lin Y Z, Zhou Z, et al. 2019. DeepDetect: A cascaded region-based densely connected network for seismic event detection. IEEE Trans Geosci Remote Sens, **57**(1): 62~75.
- Xie J J, Li X J, Wen Z P, et al. 2022. Soil profile database and site classification for national strong-motion stations in Western China. Seismol Res Lett, **93**(3): 1930~1942.
- Yamazaki F, Ansary M A. 1997. Horizontal-to-vertical spectrum ratio of earthquake ground motion for site characterization. Earthq Eng Struct Dyn, **26**(7): 671~689.
- Zhao J X, Irikura K, Zhang J, et al. 2006. An empirical site-classification method for strong-motion stations in Japan using H/V response spectral ratio. Bull Seismol Soc Am, **96**(3): 914~925.

Site Classification by Spectral Ratio Method Based on VGG16 Convolutional Neural Network Model in Sichuan-Yunnan Region

Zhang Yifan^{1,2)}, Xi Nan³⁾, Yang Tianqing³⁾, Jiang Lixin³⁾

1) Institute of Earthquake Forecasting, China Earthquake Administration, Beijing 100036, China

2) Henan Earthquake Agency, Zhengzhou 450016, China

3) China Earthquake Networks Center, Beijing 100045, China

Abstract The traditional spectral ratio method for seismic site classification often involves subjectivity and uncertainty. By integrating machine learning algorithms and leveraging large datasets along with feature extraction techniques, the objectivity and accuracy of the classification process can be enhanced, providing a more reliable foundation for seismic engineering design and hazard assessment. In this study, strong-motion waveform data from historical earthquakes in the Sichuan-Yunnan region were used to generate spectral ratio curves by calculating the horizontal and vertical velocity response spectrum values for various site categories. The characteristics of these curves, including predominant period, peak, and overall values, were analyzed. By incorporating deep learning algorithms, a dataset of spectral ratio curves from strong-motion stations in the Sichuan-Yunnan region was compiled and used as input for the VGG16 convolutional neural network (CNN) model. This model was trained to predict the probability distribution of three site categories based on NEHRP standards. The accuracy and generalization ability of the model were further enhanced through feature extraction and cross-validation techniques. The resulting site classification model demonstrated strong performance in validation, offering significant reference value for improving and applying the spectral ratio method in seismic site classification.

Keywords: Seismology; Site classification; Spectral ratio method; Deep learning