

隗永刚, 杨千里, 王婷婷, 蒋长胜, 边银菊. 2019. 基于深度学习残差网络模型的地震和爆破识别. 地震学报, 41(5): 646–657. doi: 10.11939/jass.20190030.
Wei Y G, Yang Q L, Wang T T, Jiang C S, Bian Y J. 2019. Earthquake and explosion identification based on Deep Learning residual network model. *Acta Seismologica Sinica*, 41(5): 646–657. doi: 10.11939/jass.20190030.

基于深度学习残差网络模型的 地震和爆破识别*

隗永刚^{1), †} 杨千里 王婷婷 蒋长胜 边银菊

(中国北京 100081 中国地震局地球物理研究所)

摘要 为加强对地震台网记录的天然地震与人工爆破事件进行准确的性质识别, 本文基于深度学习技术中的残差网络模型, 提出了一种新的爆破识别方法, 并根据北京数字遥测地震台网及国家数字测震台网中心记录的波形数据及其发布的包含事件性质的地震报告, 选取河北三河采石场的 93 次爆破事件和 54 次周边地震事件的波形功率谱, 分别采用不同的训练样本比例进行了 100 次和 1 000 次独立的随机抽样子试验以及“留一交叉验证法”试验, 对人工爆破与天然地震进行了识别研究. 试验结果表明, 深度学习残差网络模型在天然地震与爆破事件的性质识别中具有很高的识别率且效果稳定, 具有较好的应用前景.

关键词 地震识别 爆破识别 深度学习 残差网络

doi: 10.11939/jass.20190030

中图分类号: P315.63

文献标识码: A

Earthquake and explosion identification based on Deep Learning residual network model

Wei Yonggang^{1), †} Yang Qianli Wang Tingting Jiang Changsheng Bian Yinju

(Institute of Geophysics, China Earthquake Administration, Beijing 100081, China)

Abstract: In order to enhance the property identification of earthquakes and explosions recorded by seismic network, this paper proposed a new technology of explosion discrimination based on the residual network model in Deep Learning technology, and utilized it to identifying explosion and surrounding earthquakes in Sanhe Quarry of Hebei Province. According to the waveform data recorded by the Beijing Digital Telemetry Seismic Network and China Center of Digital Seismic Network, and the released seismic phase reports, we analyzed the waveform power spectrum of 93 explosion events and 54 surrounding seismic events in Sanhe Quarry of Hebei Province. Moreover, 100 independent random sampling sub-tests, 1 000 independent random sampling sub-tests and leave-one-out-cross-validation test were conducted by adopting different training sample proportions, respectively. The test results show that the residual network model in Deep Learning has a high recognition rate and a stable effect in identifying the property of earthquakes and explosions, therefore it has a wonderfully potential application.

* 基金项目 中央级公益性科研院所基本科研专项(DQJB18B17, DQJB19B10)资助.

收稿日期 2019-02-14 收到初稿, 2019-04-02 决定采用修改稿.

† 通信作者 e-mail: weiyonggang@cea-igp.ac.cn



Key words: earthquake discrimination; explosion discrimination; Deep Learning; residual network

引言

天然地震与人工爆破事件的识别研究, 对于安全生产、国防安全和地球科学研究以及高质量地震数据的创建具有重要意义, 准确地、自动地区分天然地震与人工爆破事件始终是相关研究中的热点问题(Stump *et al.*, 2002; Rabin *et al.*, 2016). 区域地震目录中经常混杂着工人爆破事件, 导致地震危险性评估不准确(Kortström *et al.*, 2016), 因此在分析微震活动性时需要从地震目录中剔除人工爆破事件(Wiemer, 2000), 而人工爆破事件又会引发构造应力释放剪切分量, 增加了天然地震与爆破事件的判别难度.

自 20 世纪 50 年代起, 对天然地震与人工爆破事件的识别就开始了广泛深入的研究, 并提出了多种识别判据, 包括 P 波初动、震源深度、体波震级 m_b 与面波震级 M_S 之比、倒谱、P 波与 S 波震相幅值比以及从频率域中提取的判据等(王婷婷等, 2013). 这些单特征向量最初通常被单独使用, 但当爆炸当量较小时, 其识别效果并不理想(杨宏, 贾维敏, 2000), 因此利用多个判据进行综合识别成为主要的研究方向. 天然地震相较于人工爆破事件会产生更多的低频信息(Tiira, 1995), 因此从频率域中提取的判据常用于事件识别, 而且频率域判据较时域判据具有更好的判别效果(Falsaperla *et al.*, 1996). 随着神经网络研究的兴起, 频率域信息或频率域和时域混合信息与人工神经网络相结合的方法应用于天然地震与人工爆破的识别中, 有效地提高了事件的识别能力(Esposito *et al.*, 2006; Kuyuk *et al.*, 2011; AitLaasri *et al.*, 2013; Vallejos, McKinnon, 2013; Riggelsen, Ohrnberger, 2014; Reynen, Audet, 2017).

由于早期人工神经网络受限于算法和当时硬件的计算能力, 在面对海量的地震数据时, 未能将事件的全部频谱信息输入网络, 仅从频谱信息中预先提取某些特征(如 P/S 型谱比值)对数据进行降维使用, 不但增加了处理过程的复杂性, 而且损失了大量的事件信息, 模型极易陷入局部最优解. 在网络层数很深的情况下, 神经网络参数众多, 使得模型训练速度非常缓慢, 容易出现梯度消失现象, 进而导致模型的泛化能力不强.

近年来, 深度学习技术在大量智能计算任务中得以广泛应用, 在很多方面带来了超出预期的效果, 例如语音处理(Lee *et al.*, 2009; Dahl *et al.*, 2011)、视觉计算和自然语言理解、搜索、推荐, 尤其在图像分类、图像分割、目标检测(Ren *et al.*, 2015)、事件检测、动作识别和场景识别等视觉信息处理方面均取得了显著的进步(山世光等, 2016). 深度学习的概念源于人工神经网络研究, 是人工神经网络的延伸和发展, 它通过低层特征组合形成更加抽象的高层表示属性类别或特征, 以寻求数据的分布式特征表示(孙志军等, 2012), 两者本质上都是从大量训练样本中学习统计规律, 从而对未知事件作出预测, 但深度学习模型在处理大样本和复杂的函数关系时更为便捷.

为提高天然地震事件与人工爆破事件的识别率, 本文拟应用深度学习技术, 在该领域尝试使用可以训练超深网络的残差网络模型(residual network), 探索使用波形的全部功率谱作为模型训练与测试的输入数据, 而非在频率域中预先提取若干个判据或者选取部分波形的频域数据进行降维使用, 以期简化数据的预处理流程, 在模型训练过程中保留事件更多的原始信息, 进一步提高事件的识别能力.

1 方法

1.1 深度残差网络模型基本原理

模型深度是图像识别和分类任务中非常重要的因素(Simonyan, Zisserman, 2015; Szegedy et al, 2015), 模型深度越深, 对于图像的抽象刻画能力越强, 但同时会出现梯度消失或梯度爆炸现象(Bengio et al, 2002), 从而影响模型的收敛, 不容易找到全局最优解. 为解决该问题, He 等(2016)提出了深度残差网络模型. 理论上, 只要硬件能力允许, 该模型可以训练多达上千层深的网络. 受限于计算机运算能力, 本文仅采用 14 层网络. 该模型的基本结构如图 1 所示.

在卷积神经网络(Krizhevsky et al, 2012)的基础上, 本文模型的结构增加了快捷链接, 该链接可以跨越多个卷积层(也可称为映射层). 以跨越两个卷积层为例(图 1), 输入数据 x 经过快捷链接与残差函数 $\mathcal{F}(x)$ 相加构成了一个映射层, 其输出值为 y , 即

$$y = \mathcal{F}(x) + x, \tag{1}$$

其目的是使输入数据的信息随着卷积层数的增加而不会减弱, 有利于随后的映射层拟合输入数据. 式(1)中,

$$\mathcal{F}(x) = W_2 \delta(W_1 \cdot x), \tag{2}$$

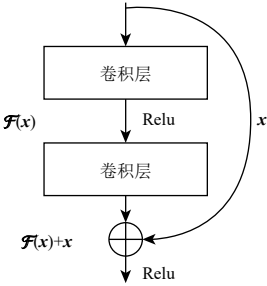


图 1 深度残差网络模型的基本结构

Fig. 1 Basic structure of deep residual network model

式中, W_1 和 W_2 为权重, x 为输入数据, δ 代表非线性激活函数, 其作用是将 $W_1 \cdot x$ 加入非线性因素, 解决线性模型所不能解决分类问题. 采用的激活函数为线性整流(rectified linear unit, 缩写为 Relu)函数, 使用 Relu 函数得到的随机梯度的下降收敛速度会较 Sigmoid 函数或 Tanh 函数快很多(Krizhevsky et al, 2012). 另外, x 和 $\mathcal{F}(x)$ 的维度必须保持一致, 才可以合并计算, 因此需要将 x 乘以一个线性映射矩阵 W_S 或将 x 的二维形状周边补 0, 若 x 与 $\mathcal{F}(x)$ 的维度一致, 则可以直接相加.

若跨越多于两个卷积层, 则式(1)和式(2)可合并为

$$y = \mathcal{F}(x, \{W_i\}) + W_S \cdot x. \tag{3}$$

拟合输入数据, 就是使 y 的分布与输入数据 x 的分布相一致, 从而求得模型的各个参数. 该模型依然可以使用结合反向传播算法的随机梯度下降法来求解. 随着该模型基本结构的堆叠, 层深会越来越深, 但由于快捷链接的存在, 不会出现梯度消失或者爆炸现象, 模型收敛效果依然很好.

1.2 模型的构建

实际应用中, 残差网络模型的结构多样, 其中卷积层中的卷积核大小、池化层大小、输入数据形状、通道数量以及快捷链接跨越层数等均根据经验设置, 并非固定不变, 细微的差别并不会影响模型的收敛效果.

本文应用的残差网络模型的具体结构如图 2 所示. 将 x 输入模型时, 需将 x 的形状调整为二维平面形式, 以方便使用卷积核进行卷积计算. 本文设置 x 的形状为 20×10 , 各卷积层的输入、输出通道数列于表 1. 第 14 层为全连接层及分类器.

需要注意的是, 在本模型中, 连接快捷链接的输入数据与相应卷积层的输出数据维度不



图 2 本文应用的残差网络模型

图中省略了 4 层卷积核为 64 和 128 的卷积层, 加上全连接层及分类器层, 共计 14 层,
图中上方数字 2×2 为卷积核大小, 下方数字为卷积核的数量

Fig. 2 The residual network model applied in this paper

The four convolutional layers with 64 and 128 convolution kernels was omitted; coupled with the fully connected layer and the classifier layer, there are 14 layers in total. 2×2 on the upper in the figures is the size of the convolution kernel, and on the bottom is the number of the convolution kernel

表 1 各卷积层输入及输出的通道数量

Table 1 The number of channels for input and output of each convolution layer

卷积层	输入通道数	输出通道数	卷积层	输入通道数	输出通道数	卷积层	输入通道数	输出通道数
1	1	8	6	32	64	11	256	256
2	8	16	7	64	64	12	256	512
3	16	16	8	64	128	13	512	512
4	16	32	9	128	128	14	全连接	
5	32	32	10	128	256			

一致, 不能直接相加, 故将输入数据二维平面的边缘补 0 来保持与输出数据维度的一致性. 此外, 当输入数据维度较高时, 卷积层与卷积层之间需设有池化层以对数据进行降维, 本文的输入数据维度仅为 200, 故未进行池化.

2 数据预处理

根据北京遥测数字地震台网及国家数字测震台网中心提供的波形数据和事件目录, 事件性质由国家数字测震台网中心判定. 试验数据选自 2010—2016 年期间发生在河北三河地区(39.8°N — 40.3°N , 116.8°E — 117.4°E)有明确分类的天然地震事件 54 次, 人工爆破事件 93 次, 震级范围为 $M_L 1.5$ — 3.0 , 震中距范围为 5 — 260 km . 研究区内分布 100 个观测台站, 每次事件根据理论到时截取这 100 个台站记录到的三分向记录, 后续的波形筛选在此基础上进行, 事件和台站分布如图 3 所示.

由于本文采用的事件发生在城市边缘地区, 部分台站受震中距和自身监测能力的影响, 事件信息中所含各类噪声的影响较大. 因此, 在开展研究之前需要进行筛选, 剔除被噪声淹没的台站波形, 之后对剩余波形进行功率谱估计.

2.1 波形筛选

本文选用短长时窗平均比(short-term average/long-term average, 缩写为 STA/LTA)方法(Stevenson, 1976; Baer, Kradolfer, 1987; 刘晗, 张建中, 2014)并结合能量算子(Teager-Kaiser energy operator, 缩写为 TKEO)对地震波形数据进行预处理, 挑选出有效信号. STA/LTA 的基本思想是利用信号短时窗平均值(STA)与长时窗平均值(LTA)之比来研究能量或振幅的变化, 当某一点比值大于设定阈值时, 视为事件发生. TKEO 能够准确地追踪信号的瞬时能量变化, 对微小事件有很好的识别效果. 本文针对三分向波形中的垂直分量进行研究, 整个筛选流程如下:

- 1) 对波形数据进行预处理, 去除仪器响应和均值倾向;

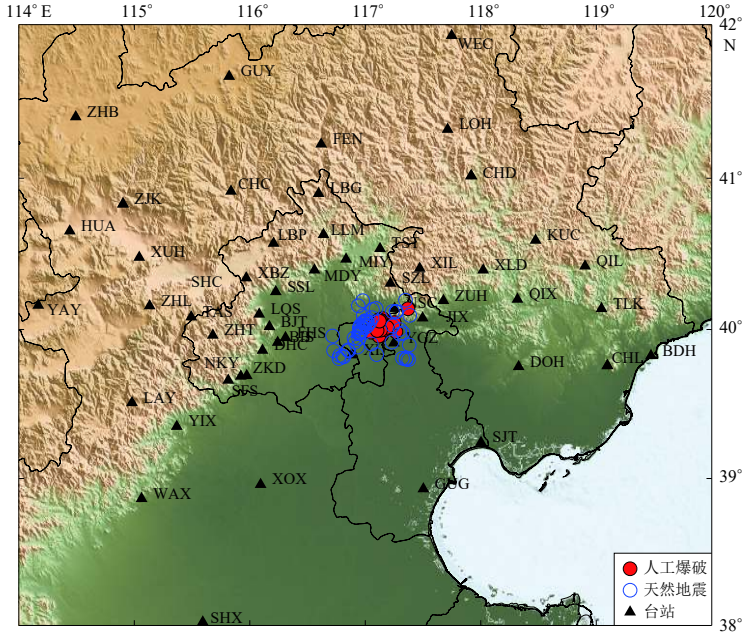


图 3 本文所用台站和事件的分布

Fig. 3 Distribution of stations and events used in this study

- 2) 根据事件发震时刻和震中距, 选取合理的速度窗口截取包含理论 P 波和 S 波在内的全波段波形;
- 3) 对数据进行 1—25 Hz 带通滤波, 消除长周期噪声干扰;
- 4) 设定阈值 ε 、长短时窗长度 N_{STA} 和 N_{LTA} ;
- 5) 进行波形扫描, 若 $\exists STA_i/LTA_i \geq \varepsilon$, 视为事件发生, 保存数据; 若 $\forall STA_i/LTA_i < \varepsilon$, 表明事件被完全淹没在噪声之中, 数据删除. 此处令 $\varepsilon=10$, $N_{STA}=0.2\text{ s}$, $N_{LTA}=2\text{ s}$, 波形筛选如图 4 所示.

由图 4 可以看出, 原始数据受长周期噪声影响严重(图 4a), 但经滤波之后(图 4b)P 波和 S 波发育清晰, 信噪比较高. 图 4c 中波起始位置的 STA/LTA 值高于红色虚线, 与人工判定结果一致. 以上示例说明, STA/LTA 方法在信噪比较低的情况下依然能对训练样本进行准确的筛选, 具有较强的实用性.

按照上述步骤筛选天然地震和人工爆破的有效波形, 每次事件最终保留样本数据数十份不等, 这样形成了 3 995 份波形数据, 较传统人工挑选波形节约时间, 效率更高.

2.2 功率谱估计

理论研究表明, 由于震源性质不同, 人工爆破和天然地震具有不同的主频、拐角频率和频带宽度(Bennett, Murphy, 1986; Hong, 2013), 这些差异或多或少能够呈现在频谱上. 台站接收到的天然地震和人工爆破可以看作随机信号, 频率分布服从一定的统计规律. 功率谱密度作为频谱分析的重要手段, 在数学计算上等于振幅谱的平方, 将其作为识别判据可以放大地震和爆炸的频域差异. 本文采用 Welch 平均周期法(郝春月等, 2012)对筛选出的波形数据进行谱估计.

频谱分析的优势在于不受震源和震中距影响, 不同台站记录的天然地震和人工爆破事

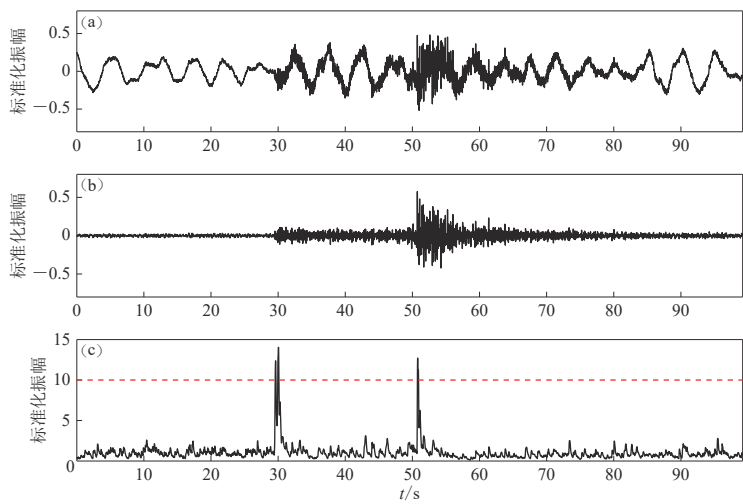


图4 波形筛选示意图
(a) 原始波形; (b) 带通滤波去噪结果; (c) TKEO 短长时窗比

Fig. 4 Schematic diagram of waveform selection

(a) Original waveform; (b) De-noising result by band-pass filtering; (c) TKEO short-length time-to-window ratio

件的总长度各不相同,但在频域的功率谱密度长度一致,不影响后续的训练和识别,选取2012-01-18 23:03:54.6和2010-01-22 15:40:26.6震级同为 $M_S2.0$ 的天然地震和人工爆破事件,震中位置分别为(39.803°N, 117.352°E)和(40.049°N, 117.139°E);震中距分别为 $\Delta=67.52\text{ km}$ 和 $\Delta=65.07\text{ km}$,对两事件进行信号功率谱对比,结果如图5所示。

由图5可见,爆炸信号功率谱在1—5 Hz低频部分明显高于地震信号,高频区域在分布规律上与地震相似。由于单一样本不能反映整体类别的频谱特征,本文将功率谱作为训练和测试数据,对筛选出的每一条数据进行功率谱估计,为后续的训练与判别奠定基础。

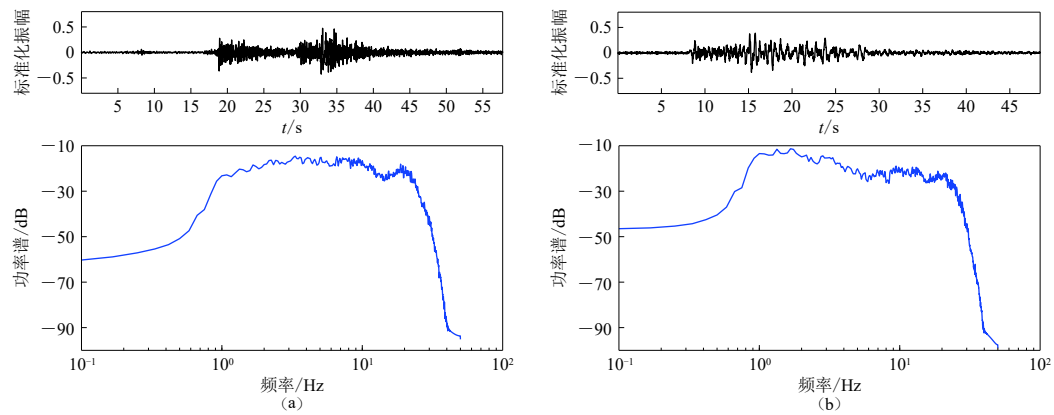


图5 $M_S2.0$ 天然地震(a)和人工爆破(b)经1—25 Hz带通滤波的波形及功率谱

Fig. 5 Waveforms with 1-25 Hz bandpass filter and power spectra of $M_S2.0$ earthquake (a) and explosion (b)

3 训练集和测试集的创建及测试过程

训练集和测试集分别用于模型的训练和测试,两个集合分别包含若干个天然地震样本

和人工爆破样本,且两个集合之间的样本互不重叠.集合中每个样本均由事件波形的功率谱和标签组成.每个样本的功率谱由 200 个谱值构成,以其为模型的输入;将天然地震样本标签标注为 0,人工爆破样本标签标注为 1,之后对标签进行 one-hot 编码,将其作为观测值,而将模型的输出作为模型的预测值,模型训练的目的为使训练集的输出值分布与观测值的分布尽量一致,即输出值的分布与预测值的分布之间的差值尽可能小,其差值定义为损失函数,本文采用的损失函数为交叉熵,优化函数采用自适应矩估计(adaptive moment estimation,缩写为 Adam)(Kingma, Ba, 2015).自适应矩估计是一种计算每个参数不同更新速率的方法,它利用梯度的一阶矩估计和二阶矩估计动态地调整每个参数的学习率,其优点主要在于经过偏置校正后,每一次迭代学习率均有个确定范围,使得参数比较平稳.

为了验证模型的准确率和稳定性,本文进行两组随机抽样试验和一组留一交叉验证试验.前两组随机抽样试验中,每组试验各进行 100 次和 1 000 次独立的随机抽样子试验.第一组试验中,每次子试验从天然地震和人工爆破中各随机抽取 20 次事件构成训练集进行训练,余下的样本作为测试集.第二组试验中,每次子试验从天然地震和人工爆破中分别随机抽取 27 次和 46 次事件构成训练集进行训练,余下的样本作为测试集.第三组试验中,采用留一交叉验证法(Kohavi, 1995)进行试验,即每次子试验从 147 次事件中留下一个事件作为测试,其它 146 次事件全部用于模型训练,该步骤将一直持续到每次事件都被作为一次测试事件时结束,这样最终可得到 147 次测试结果,其平均值用于计算模型的准确率,因此需要进行 147 次子试验.

在前两组 100 次和 1 000 次随机抽样的子试验中,将每次输入模型的样本数量即块尺寸(batch_size)设置为训练集的全部样本,由于未采用分批输入的方式,所以每次子试验只进行 100 次或 1 000 次父循环(训练).而在第三次试验中,块尺寸设置为 500 个样本,因每次输入模型的样本数量并非训练集的全部样本数,一次父循环(训练)过程中会将全部样本分为若干批次输入模型进行子循环(训练),所以每次子试验的父循环(训练)次数可相应减少,本文选取 3 次父循环(训练).子循环(训练)次数即批次为:训练样本数/块尺寸-1.需要指出的是,父循环(训练)次数和块尺寸大小,可根据 GPU 的内存大小以及模型收敛的速度进行调整.一般来说,块尺寸越大,批次越小,子循环(训练)次数就越少,因此需要更多的父循环(训练)才能使模型收敛.另外 GPU 内存的大小,也决定块尺寸并不能设置为无限大.

在计算速度方面,模型训练时间受样本量大小、模型复杂程度、迭代次数以及计算机运算能力的影响而长短不一.以本文为例,单次训练过程(1 000 次迭代)的时间在 2 分钟左右.而测试过程是在模型参数都训练完毕后进行,只是输入和输出的过程,所以计算相当迅速,仅需毫秒即可完成.

4 研究结果

若以每个测试样本(台站记录)是否被正确识别作为评判标准(基于记录),则第一组试验的 100 次子试验的平均识别率、地震平均识别率和爆破平均识别率分别为 90.3%, 93.6% 和 88.3%; 1 000 次子试验的平均识别率、地震平均识别率和爆破平均识别率分别为 90.2%, 93.6% 和 88.3%. 第二组试验的 100 次子试验的平均识别率、地震平均识别率和爆破平均识别率分别为 91.7%, 91.0% 和 92.3%; 1 000 次子试验的平均识别率、地震平均识别率和爆破平均识别率分别为 91.6%, 90.6% 和 92.1%. 详细结果见图 6.

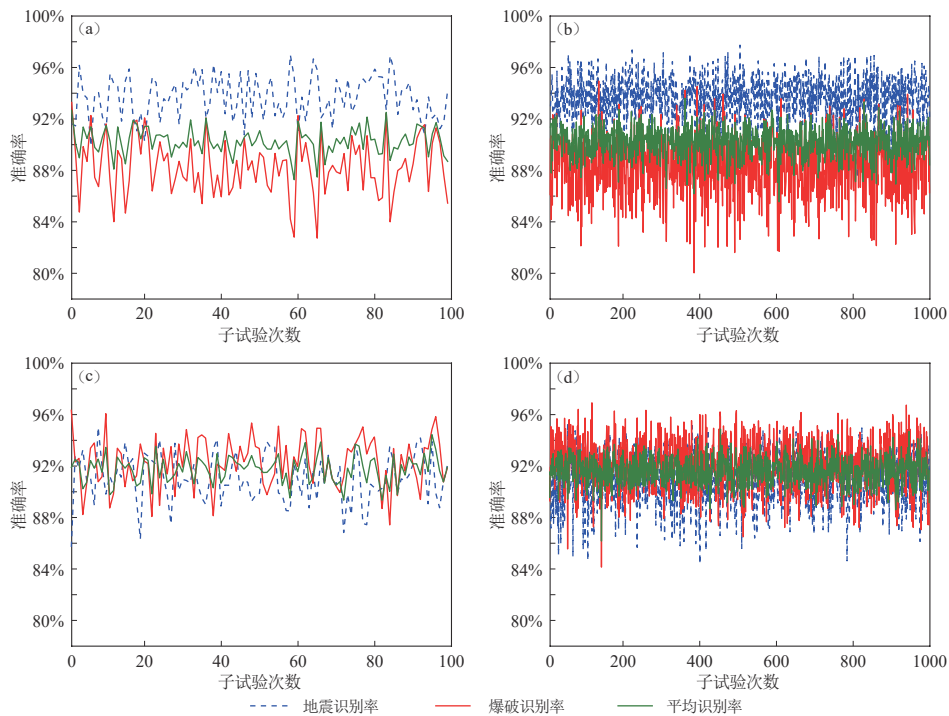


图 6 基于台站记录的试验结果

图(a)和(b)为第一组试验中的 100 次随机子试验和 1 000 次随机子试验, 每次子试验从地震和爆破中分别抽取 20 次事件进行训练, 余下的事件作为预测; 图(c)和(d)分别表示第二组试验中 100 次随机子试验和 1 000 次随机子试验, 每次子试验从地震和爆破中分别抽取 27 和 46 次事件进行训练, 余下的事件作为预测

Fig. 6 Test results based on station records

Figs. (a) and (b) are the results for the 100 random subtests and 1 000 random subtests in the first group, each subtest selects 20 events from earthquakes and explosions for training, and the remaining events are used as predictions; Figs. (c) and (d) are the results for 100 random subtests and 1 000 random subtests in the second group, each subtest selects 27 and 46 events from earthquakes and explosions for training, and the remaining events are used as predictions

若以每次事件基于 50%, 60% 和 70% 的台站记录被正确识别作为评判标准(基于事件)统计, 准确率会进一步提高. 为保持图表简洁, 仅将基于 50% 的台站记录被正确识别作为评判标准(基于事件)的结果列于表 2. 同时, 给出基于 50% 和 60% 台站记录被正确判别为识别标准(基于事件)的 100 次随机子试验进程及结果(图 7), 同时给出基于 50% 台站记录被正确判

表 2 基于事件的 3 组试验结果对比

Table 2 Comparison of the three groups of event-based test results										
试验组	随机子 试验次数	基于50%记录被正确 判别的平均识别率			基于50%记录被正确 判别的地震识别率			基于50%记录被正确 判别的爆破识别率		
		平均值	最大值	最小值	平均值	最大值	最小值	平均值	最大值	最小值
第一组	100	97.7%	100%	93.5%	99.4%	100%	97.1%	96.9%	100%	91.8%
	1 000	97.6%	100%	92.5%	99.3%	100%	97.1%	96.8%	100%	90.4%
第二组	100	98.4%	100%	95.9%	98.7%	100%	96.3%	98.3%	100%	93.6%
	1 000	98.3%	100%	94.6%	98.8%	100%	90.5%	98.1%	100%	93.6%
留一交叉验证		97.3%			96.3%			97.8%		

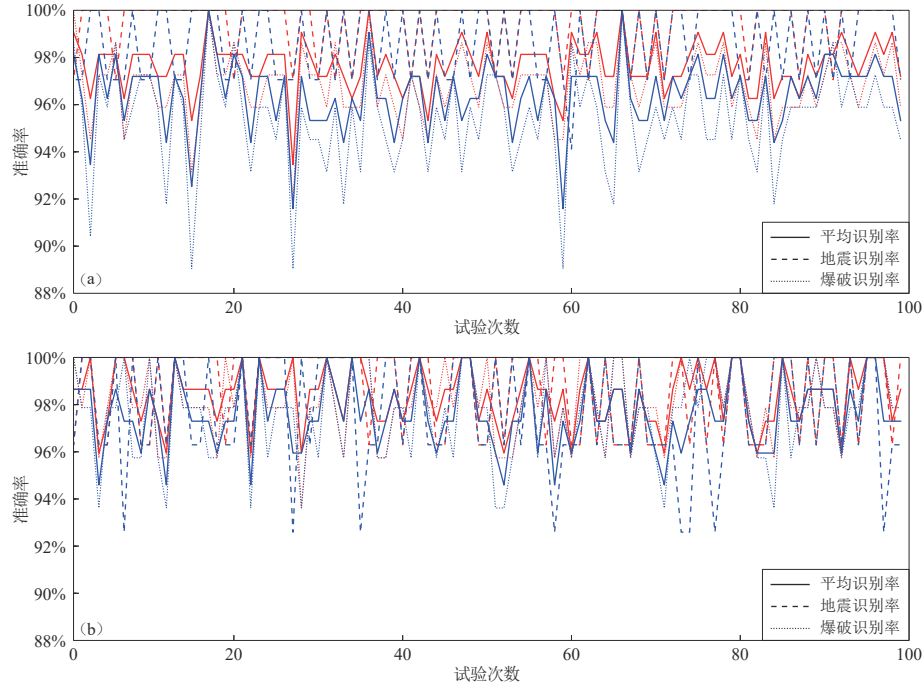


图 7 基于事件的 100 次随机子试验结果

图(a)为第一组试验中的 100 次随机子试验,每次的子试验从地震和爆破中分别抽取 20 次事件进行训练,余下的事件作为预测;图(b)为第二组试验中的 100 次随机子试验,每次子试验从地震和爆破中分别抽取 27 和 46 次事件进行训练,余下的事件作为预测.图中红、蓝线条分别代表基于 50%, 60% 台站记录被正确判别为识别标准(基于事件)的识别率

Fig. 7 Results of 100 random subtests based on the events

Figs. (a) and (b) represent 100 random subtests in the first group and the second group, respectively. Fig. (a) indicates that each subtest takes 20 events from the blasting and earthquakes for training, and the remaining events are used as predictions; Fig. (b) indicates that each subtest takes 46 and 27 events from the blasting and earthquakes for training, and the remaining events are used as predictions.

别为识别标准(基于事件)的 1 000 次随机子试验进程及结果(图 8). 此外,采用留一交叉验证对模型的准确率和稳定性进行验证,因留一交叉验证试验中每次测试(子试验)仅有一次事件,基于事件的识别率仅当全部子试验完毕后才能给出,因此仅存在最终识别率,不存在平均识别率、最高识别率和最低识别率.结果显示:若以每次事件超过 50%, 60% 和 70% 的台站记录被正确识别作为评判标准(基于事件),最终识别率分别为 97.3%, 95.2% 和 93.9%. 为进行方法对比,基于 50% 台站记录被正确识别作为评判标准(基于事件)的结果列于表 2.

5 讨论与结论

为探索将深度学习技术在天然地震与人工爆破的识别中的应用,本文尝试在该领域应用可以训练超深网络的残差网络模型. 研究中,将波形的全部功率谱作为模型训练与测试的输入数据. 为了测试模型的识别率和稳定性,共进行了 3 组对比试验,前两组试验中分别依据训练样本比例不同各进行了 100 次和 1 000 次随机抽样子试验,第 3 组试验采取留一交叉验证法进行了测试. 根据以上试验的结果,得到如下认识:

- 1) 以超过 50% 的台站记录被正确识别(基于事件)作为判别标准较为客观. 结果显示:

按照 50% 的判别标准, 前两组试验的 100 次随机试验平均识别率分别为 97.7% 和 98.4%, 1 000 次随机试验平均识别率分别为 97.6% 和 98.3%; 第 3 组试验的识别率为 97.3%.

2) 3 组比较接近的试验结果表明, 在训练集达到总样本的一定比例后, 训练集的大小对结果影响并不大, 模型可从有限的样本中统计出样本的规律, 展现出模型强大的学习能力和泛化能力. 另外, 分析认为, 事件的波形功率谱除了震中距的因素, 一定程度上消除了传播路径的影响, 更能反映出震源的信息, 为该模型较高的识别率奠定了基础.

前两组试验中, 相同随机抽样数量下的 1 000 次与 100 次随机试验的平均识别率仅相差 0.1%, 且在 1 000 次随机试验(基于事件)的结果(图 8)中, 可以清晰地看到, 平均识别率、地震识别率、爆破识别率的分布均呈现条带状, 仅有少数值偏离条带, 这些结果充分展示出模型对于地震和爆破的识别不受试验次数的影响, 具有非常高的识别率和稳定性.

本文中的天然地震事件和人工爆破事件在空间上呈现包围分布, 如图 3 所示, 具有明显的规律性. 对于这种有规律性的样本分布, 模型采用较少的训练样本就可找到样本的统计规律. 而对于复杂的空间分布, 则需要更多的样本进行训练.

虽然本文研究的区域仅限于河北三河采石场及周边地区, 模型经过训练后得到的模型参数不能完全适用到其它地区, 但就模型和方法而言具有很好的通用性和稳定性. 此外, 可以将多个地区的天然地震和人工爆破统一进行训练和测试, 这样得到的模型参数会更具有

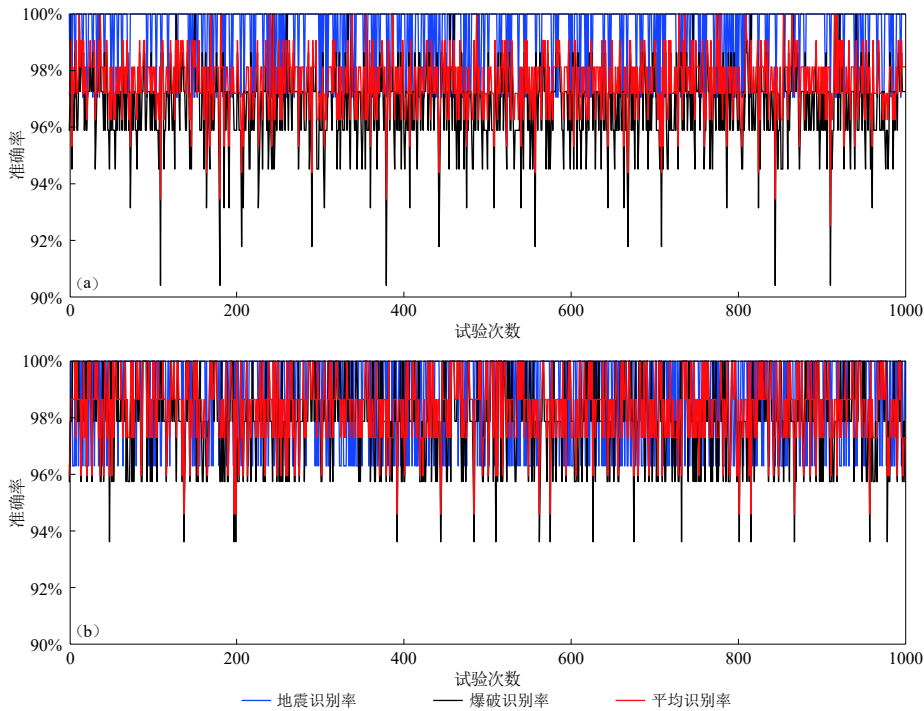


图 8 基于事件的 50% 地震记录的 1 000 次随机子试验结果
图(a)和(b)分别代表第一组试验和第二组试验中的 1000 次随机子试验, 试验说明同图 7
Fig. 8 Results of 1 000 random subtests based on the event
Figs. (a) and (b) represent 1 000 random subtests in the first group and the second group, respectively. Other explanations are the same as Fig. 7

普适性. 另外, 本文采用的模型深度仅为 14 层, 还可将模型深度进一步加深, 预计可以进一步提高模型的识别率.

波形功率谱作为训练和识别的样本是否有物理机制可以支撑, 能否作为“判据”使用是地震学界比较关心的问题. 在深度学习中使用的波形功率谱并不是直接作为判据使用, 而是作为训练(学习)和测试样本使用, 应用深度学习技术从波形功率谱中学习出爆破和地震的频谱特征, 这些被学习到的特征是一种概率分布, 而这些学习到的概率分布才是本文中应用到的判据, 但是深度学习作为一种黑箱算法很难从物理机制上阐述出这些特征(概率分布)的物理意义, 这确实是深度学习解释中的一个短板. 但是这并不影响波形功率谱可以作为训练(学习)和测试样本使用, 只要应用测试样本进行测试的准确率比较高, 就可以证明从训练样本中是可以学习出爆破和地震的特征规律的. 本文试验结果也表明, 应用波形功率谱进行训练和测试, 可以准确区分爆破和地震事件, 因而将波形功率谱作为训练和测试样本是可行的.

参 考 文 献

- 郝春月, 郑重, 张爽. 2012. 玉树地震前后当地的噪声变化研究[J]. *地球物理学进展*, **27**(6): 2418–2428.
- Hao C Y, Zheng Z, Zhang S. 2012. Research about noise variation around YUS station before and after the Yushu earthquake[J]. *Progress in Geophysics*, **27**(6): 2418–2428 (in Chinese).
- 刘晗, 张建中. 2014. 微震信号自动检测的 STA/LTA 算法及其改进分析[J]. *地球物理学进展*, **29**(4): 1708–1714.
- Liu H, Zhang J Z. 2014. STA/LTA algorithm analysis and improvement of microseismic signal automatic detection[J]. *Progress in Geophysics*, **29**(4): 1708–1714 (in Chinese).
- 山世光, 阚美娜, 刘昕, 刘梦怡, 邬书哲. 2016. 深度学习: 多层神经网络的复兴与变革[J]. *科技导报*, **34**(14): 60–70.
- Shan S G, Kan M N, Liu X, Liu M Y, Wu S Z. 2016. Deep Learning: The revival and transformation of multi layer neural networks[J]. *Science & Technology Review*, **34**(14): 60–70 (in Chinese).
- 孙志军, 薛磊, 许阳明, 王正. 2012. 深度学习研究综述[J]. *计算机应用研究*, **29**(8): 2806–2810.
- Sun Z J, Xue L, Xu Y M, Wang Z. 2012. Overview of Deep Learning[J]. *Application Research of Computers*, **29**(8): 2806–2810 (in Chinese).
- 王婷婷, 边银菊, 张博. 2013. 地震和爆破的综合识别方法研究[J]. *地球物理学进展*, **28**(5): 2433–2443.
- Wang T T, Bian Y J, Zhang B. 2013. The comprehensive identification methods between earthquakes and explosions[J]. *Progress in Geophysics*, **28**(5): 2433–2443 (in Chinese).
- 杨宏, 贾维敏. 2000. 基于神经网络的综合评判在核爆模式识别中的应用[J]. *核电子学与探测技术*, **20**(4): 279–283.
- Yang H, Jia W M. 2000. Recognition of underground nuclear explosion and natural earthquake based on neural network[J]. *Nuclear Electronics & Detection Technology*, **20**(4): 279–283 (in Chinese).
- AitLaasri E H, Akhouayri E S, Agliz D, Atmani A. 2013. Seismic signal classification using multi-layer perceptron neural network[J]. *Int J Comput Appl*, **79**(15): 35–43.
- Baer M, Kradolfer U. 1987. An automatic phase picker for local and teleseismic events[J]. *Bull Seismol Soc Am*, **77**(4): 1437–1445.
- Bengio Y, Simard P, Frasconi P. 2002. Learning long-term dependencies with gradient descent is difficult[J]. *IEEE Trans Neural Netw*, **5**(2): 157–166.
- Bennett T J, Murphy J R. 1986. Analysis of seismic discrimination capabilities using regional data from western United States events[J]. *Bull Seismol Soc Am*, **76**(4): 1069–1086.
- Dahl G E, Yu D, Deng L, Acero A. 2011. Large vocabulary continuous speech recognition with context-dependent DBN-HMMs[C]//*IEEE International Conference on Acoustics Speech and Signal Processing*. Prague: IEEE: 4688–4691.
- Esposito A M, Giudicepietro F, Scarpetta S, D'Auria L, Marinaro M, Martini, M. 2006. Automatic discrimination among landslide, explosion-quake, and microtremor seismic signals at Stromboli Volcano using neural networks[J]. *Bull Seismol Soc Am*, **96**(4A): 1230–1240.

- Falsaperla S, Graziani S, Nunnari G, Spampinato S. 1996. Automatic classification of volcanic earthquakes by using multi-layered neural networks[J]. *Nat Hazards*, **13**(3): 205–228.
- He K M, Zhang X Y, Ren S Q, Sun J. 2016. Deep residual learning for image recognition[C]//*IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas: IEEE: 770–778.
- Hong T K. 2013. Seismic discrimination of the 2009 North Korean nuclear explosion based on regional source spectra[J]. *J Seismol*, **17**(2): 753–769.
- Kingma D P, Ba J. 2015. Adam: A method for stochastic optimization[C]//*Proceedings of the 3rd International Conference for Learning Representations*. San Diego: The Institute for Catastrophic Loss Reduction.
- Kohavi R. 1995. A study of cross-validation and bootstrap for accuracy estimation and model selection[C]//*Proceedings of the 14th International Joint Conference on Artificial Intelligence*. Montreal: Morgan Kaufmann Publishers Inc, **14**: 1137–1143.
- Kortström J, Uski M, Tiira T. 2016. Automatic classification of seismic events within a regional seismograph network[J]. *Comput Geosci*, **87**: 22–30.
- Krizhevsky A, Sutskever I, Hinton G E. 2012. ImageNet classification with deep convolutional neural networks[C]//*Proceedings of the 25th International Conference on Neural Information Processing Systems*. Lake Tahoe: Curran Associates Inc, 1097–1105.
- Kuyuk H S, Yildirim E, Dogan E, Horasan G. 2011. An unsupervised learning algorithm: Application to the discrimination of seismic events and quarry blasts in the vicinity of Istanbul[J]. *Nat Hazards Earth Syst Sci*, **11**(1): 93–100.
- Lee H, Largman Y, Pham P, Ng A Y. 2009. Unsupervised feature learning for audio classification using convolutional deep belief networks[C]//*Proceedings of the 22nd International Conference on Neural Information Processing Systems*. Vancouver: Curran Associates Inc: 1096–1104.
- Rabin N, Bregman Y, Lindenbaum O, Ben-Horin Y, Averbuch A. 2016. Earthquake-explosion discrimination using diffusion maps[J]. *Geophys J Int*, **207**(3): 1484–1492.
- Ren S Q, He K M, Girshick R, Sun J. 2015. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. *IEEE Trans Pattern Anal Mach Intell*, **39**(6): 1137–1149.
- Reynen A, Audet P. 2017. Supervised machine learning on a network scale: Application to seismic event classification and detection[J]. *Geophys J Int*, **210**(3): 1394–1409.
- Riggelsen C, Ohrnberger M. 2014. A machine learning approach for improving the detection capabilities at 3C seismic stations[J]. *Pure Appl Geophys*, **171**(3/4/5): 395–411.
- Simonyan K, Zisserman A. 2015. Very deep convolutional networks for large-scale image recognition[C]//*Proceedings of the 3rd International Conference for Learning Representations*. San Diego: The Institute for Catastrophic Loss Reduction.
- Stevenson P R. 1976. Microearthquakes at Flathead Lake, Montana: A study using automatic earthquake processing[J]. *Bull Seismol Soc Am*, **66**(1): 61–80.
- Stump B W, Hedlin M A H, Pearson D C, Hsu V. 2002. Characterization of mining explosions at regional distances[J]. *Rev Geophys*, **40**(4): 2–1.
- Szegedy C, Liu W, Jia Y Q, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A. 2015. Going deeper with convolutions[C]//*IEEE Conference on Computer Vision and Pattern Recognition*. Boston: IEEE: 1–9.
- Tiira T. 1995. Discrimination of nuclear explosions and earthquakes from teleseismic distances with a local network of short period seismic stations using artificial neural networks[J]. *Phys Earth Planet Inter*, **97**(1/2/3/4): 247–268.
- Vallejos J A, McKinnon S D. 2013. Logistic regression and neural network classification of seismic records[J]. *Int J Rock Mech Min Sci*, **62**: 86–95.
- Wiemer S. 2000. Mapping and removing quarry blast events from seismicity catalogs[J]. *Bull Seismol Soc Am*, **90**(2): 525–530.