



地震学报

ACTA SEISMOLOGICA SINICA



ISSN 0253-3782

CN 11-2021/P

基于深度学习的地震震级分类

刘涛 戴志军 陈苏 傅磊

Earthquake magnitude classification based on deep learning

Liu Tao, Dai Zhijun, Chen Su, Fu Lei

引用本文:

刘涛, 戴志军, 陈苏, 傅磊. 2022. 基于深度学习的地震震级分类. *地震学报*[J], 44(4): 656–664. DOI: 10.11939/jass.20210046

Liu T, Dai Z J, Chen S, Fu L. 2022. Earthquake magnitude classification based on deep learning. *Acta Seismologica Sinica*[J], 44(4): 656–664. DOI: 10.11939/jass.20210046

在线阅读 View online: <https://doi.org/10.11939/jass.20210046>

您可能感兴趣的其他文章

Articles you may be interested in

基于深度学习残差网络模型的地震和爆破识别

Earthquake and explosion identification based on Deep Learning residual network model
地震学报. 2019, 41(5): 646–657 <https://doi.org/10.11939/jass.20190030>

基于深度卷积神经网络的地震数据随机噪声压制

Random noise suppression of seismic data based on deep convolution neural network
地震学报. 2021, 43(4): 474–482 <https://doi.org/10.11939/jass.20200135>

基于深度降噪自编码神经网络的中国大陆地壳厚度反演

Inversion of crustal thickness in Chinese mainland based on deep denoising autoencoder neural network
地震学报. 2021, 43(1): 34–47 <https://doi.org/10.11939/jass.20200043>

利用卷积神经网络检测地震的方法与优化

Earthquake detection using convolutional neural network and its optimization
地震学报. 2020, 42(6): 669–683 <https://doi.org/10.11939/jass.20200045>

基于反向传播神经网络的闽台 M_L 震级偏差分析与修正

Analysis and revision of magnitude M_L deviation between Fujian and Taiwan based on BP neural network
地震学报. 2019, 41(6): 723–734 <https://doi.org/10.11939/jass.20190038>

兼顾速度和精度的深度神经网络震相拾取

A new deep neural network for phase picking with balanced speed and accuracy
地震学报. 2020, 42(3): 269–282 <https://doi.org/10.11939/jass.20190154>

刘涛, 戴志军, 陈苏, 傅磊. 2022. 基于深度学习的地震震级分类. 地震学报, 44(4): 656-664. doi: 10.11939/jass.20210046.

Liu T, Dai Z J, Chen S, Fu L. 2022. Earthquake magnitude classification based on deep learning. *Acta Seismologica Sinica*, 44(4): 656-664. doi: 10.11939/jass.20210046.

基于深度学习的地震震级分类^{*}

刘 涛 戴志军[†] 陈 苏 傅 磊

(中国北京 100081 中国地震局地球物理研究所)

摘要 为了探索地震加速度时程记录的震级信息, 训练卷积神经网络基于地震震级大小对地震记录进行分类, 将 K-NET 和 KiK-net 中将近 12 万个地震记录作为样本, 对其进行信息筛选和归一化, 之后将地震加速度时程记录用作输入, 训练卷积神经网络模型以 $M5.5$ 为分类界限来区分大震和小震. 结果显示, 在训练集中基于该模型的分类准确率为 93.6%, 在测试集中的准确率为 92.3%, 具有良好的分类效果, 这表明大震记录与小震记录之间存在一些根本的区别, 即可通过地震动加速度时程记录获取一定的震级信息.

关键词 深度学习 地震记录分析 卷积神经网络 震级信息

doi: 10.11939/jass.20210046 中图分类号: P315.9 文献标识码: A

Earthquake magnitude classification based on deep learning

Liu Tao Dai Zhijun[†] Chen Su Fu Lei

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

Abstract: In order to explore the magnitude information of the seismic acceleration time history recordings, we train a convolutional neural network to classify the seismic recordings based on the magnitude of the earthquakes. Nearly 120 000 earthquake recordings in K-NET and KiK-net are used as samples, and these acceleration time history recordings are used as inputs for model training after information screening and normalization. Taking the magnitude $M5.5$ as the classification standard, we train a deep learning model of convolutional neural network to classify large and small earthquakes. The results show that the model has an accuracy rate of 93.6% on the training set and 92.3% on the test set, which has a good classification effect. This suggests there are some fundamental differences between large earthquake recordings and small ones. Thus, earthquake magnitude information may be revealed from acceleration time history recordings of earthquakes.

Key words: deep learning; seismic recording analysis; convolution neural network; magnitude information

* 基金项目 国家重点研究发展计划(2018YFE0109800)和国家自然科学基金(51738001, U1839202)共同资助.

收稿日期 2021-04-01 收到初稿, 2021-05-31 决定采用修改稿.

作者简介 刘涛, 在读硕士研究生, 主要从事地震动记录特征分析的相关研究, e-mail: 844613078@qq.com; [†]戴志军, 博士, 研究员, 主要从事强地震动特性分析与合成、地震动过程数值模拟、计算机视觉方法在工程地震方面的应用和深度学习相关应用研究, e-mail: dzj@cea-igp.ac.cn



引言

人工神经网络(artificial neural networks, 缩写为 ANN)在过去的三十年中取得了长足的发展, 从最开始模仿神经元而建立数学模型发展到如今已经成为广泛应用于众多领域的实用技术(Murphy, 2012; Jordan, Mitchell, 2015), 特别是在地震学领域也有很多应用, 例如地震识别和分类(Dysart, Pulli, 1990; Ursino *et al*, 2001; 周本伟等, 2020)、地震相位拾取(Tiira, 1999; Wiszniowski *et al*, 2014; 李安等, 2020)等. 深度神经网络作为人工神经网络的一个分支, 由于需要大量的训练数据和其它约束条件, 并未得到广泛使用, 但最近十多年里该技术在数据收集、存储、传输和分析等方面的应用得到了突破性的发展. 数据的爆炸性增长迫切需要能够对其进行有效分析的方法, 而深度神经网络正好可以满足这一迫切需求, 而且由于该方法具有对事物或抽象概念建立更复杂模型的能力, 因而在处理大样本和复杂函数关系时更为便捷(隗永刚等, 2019). 对于运用深度学习方法的震学研究, 其核心是利用深度学习模型分析数据以获取、使用有效的信息. 经过地震学研究人员近年来的努力, 深度学习技术已成功用于许多挑战性的研究中, 例如地震岩性预测(Zhang *et al*, 2018)、地震事件检测与定位(Huang *et al*, 2018)、地震相位检测与拾取(Zhu *et al*, 2019)、相位关联(Ross *et al*, 2019)等. 而现阶段, 地震的准确预测作为公认的世界性科学难题, 还很难实现(张肇诚, 张炜, 2016). 为了预防地震带来较大的危害和损失, 须根据当地的抗震设防标准进行抗震设计. 在工程抗震设计、研究和分析中, 往往需要选择实际的地震动记录来代表地震对结构的作用, 亦或是代表施加于该结构的一种地震荷载(谢礼立, 翟长海, 2003). 对于工程而言, 考虑到路径和场地的影响, 实际的地震动应该是当地的大震记录, 显然满足此要求的地震动记录很少, 甚至很多地方都无大的地震动记录, 这就需要对当地的地震动记录进行一定调整以使其满足作为地震动输入的要求. 国外对这方面的研究开展较早, 美国太平洋地震研究中心(Pacific Earthquake Engineering Research, 缩写为 PEER)将地震动记录的缩放方法分成五类: ①通过震级、断层距将记录在已知结构基本周期处的加速度谱值缩放至目标谱在该周期处的值; ②通过缩放记录幅值, 使所选记录的反应谱与危险谱拟合一致; ③选择的记录谱应很好地拟合条件均值谱(Baker, Cornell, 2005; Baker, Cornell, 2006); ④所选记录的 ε 应与预测地震的 ε 值相接近, ε 是给定周期点的记录谱值与地震动预测方程平均值的差值(Goulet *et al*, 2004), ε 在预测结构反应上有显著表现, 是一个谱形的指标(Goulet, 2005; Goulet *et al*, 2006); ⑤所选记录的位移谱应较好地拟合非线性目标位移谱(Baker, Cornell, 2006b). 而在国内, 对于地震动输入多是选择比较常用的大震记录, 对于地震动记录的处理往往是将加速度记录调整到目标场地、目标设防的规范加速度值(朱晓炜, 2011), 具有明显的局限性. 为了选择一个合适的地震动记录调整方法, 本文拟建立一个卷积神经网络(convolutional neural networks, 缩写为 CNN)来分析地震加速度时程记录的特征, 并选择归一化的加速度记录作为样本输入以训练模型对大、小地震进行分类, 并基于单方向地震加速度记录来判断地震的震级大小, 由此判断小震记录经过调整是否具有一定的大震特性, 以提高抗震分析的有效性.

1 数据预处理

在训练过程中, 为了提高模型的有效性, 确保每个样本数据的规模相当, 需先对数据进行归一化处理, 使每个地面运动记录输入具有相同的峰值加速度. 将每个地震加速度记录定

义为一组向量,即

$$\mathbf{x}=[x_1, x_2, \cdots, x_n], \tag{1}$$

取每个地震记录的绝对加速度最大值为

$$x_{\max}=\max(|x_1|, |x_2|, \cdots, |x_n|), \tag{2}$$

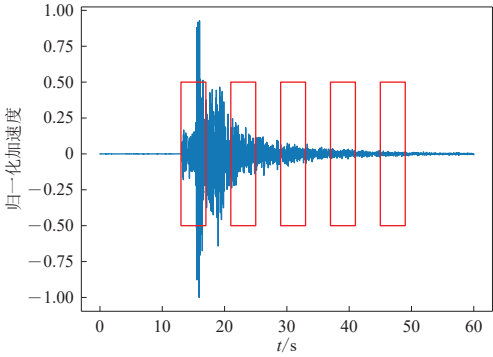


图 1 预处理阶段采样图

红线部分是采样频率为 100 Hz 的五个采样位置,
每个位置采样时长为 4 s, 五个部分共 20 s

Fig. 1 Pre-processing sampling

The red boxes delineate the five sampling positions with
a sampling frequency of 100 Hz. The sampling time of
each position is 4 s, and the five parts are 20 s in total

依次将每个加速度值除以加速度最大绝对值 x_{\max} , 相应的归一化公式为

$$x_i^*=\frac{x_i}{x_{\max}}, \quad i=1, 2, \cdots, n, \tag{3}$$

式中, x_i^* 为地震记录的归一化值, x_i 为地震记录的加速度值.

卷积神经网络要求每个地震记录样本的输入形状必须一致, 但地震记录具有不同的持时和采样频率, 因此, 在数据预处理阶段, 本文以 20 s 的采样时间和 100 Hz 的采样频率对每个地震记录进行均匀采样. 在每个地震记录截取五段共获取 20 s 长的输入数据, 前 0.05% 的阿里亚斯强度 (Arias, 1970) 是采样的起点, 后 0.05% 的阿里亚斯强度是采样的终点, 起点与终点之间平均取五段, 每段的采样时间为 4 s, 总计 20 s, 如图 1 所示.

2 卷积神经网络模型

近年来, 深度学习的飞速发展使其成为智能数据分析的有力工具, 而地震学是一门以数据为驱动力的学科, 因此构建深度学习模型成为我们的首选, 其中卷积神经网络因其特征提取的鲁棒性而被广泛应用. 特别是在语音识别领域中 (Sainath *et al*, 2013; Tóth, 2013; Qian *et al*, 2016; Sercu *et al*, 2016; Yu *et al*, 2016), 每个人的发音大不相同, 卷积神经网络因为有局部滤波和最大池化技术可以有效地消除这种差异, 有利于语音的声学建模, 并且可以提高训练效果. 考虑到地震动数据在很多方面与语音数据相似, 例如都需要介质, 都是通过振动传播, 都是非平稳时间序列信号等, 本文拟采用卷积神经网络模型基于地震动记录识别地震的大致震级.

如图 2b 所示, 模型中卷积层共有三层, 每一层只有一个卷积层, 没有池化层. 每层分为五个部分, 分别对应于每个样本的五段输入, 也就是采样过程中的五段数据. 第一层每段输入数据形状为 1×400 , 核心数为 4, 卷积窗尺寸为 1×5 , 步长为 5; 第二层卷积层将第一层卷积层的输出作为输入, 核心数为 8, 卷积窗尺寸为 1×4 , 步长为 4; 第三层卷积层将第二层输出作为输入, 核心数为 16, 卷积窗尺寸为 1×2 , 步长为 2. 上述卷积层的所有激活函数都使用 ReLu 激活函数 (Krizhevsky *et al*, 2017), 这是因为 ReLu 激活函数可以有效地避免梯度消失和过度拟合的问题, 并且具有样本导数形式, 可以加快训练速度. ReLu 激活函数如下:

$$\text{ReLu}(x)=\max(0, x). \tag{4}$$

样本每次经过卷积层时, 数据长度都会以卷积步长为倍数而减少, 数据宽度会以核心数

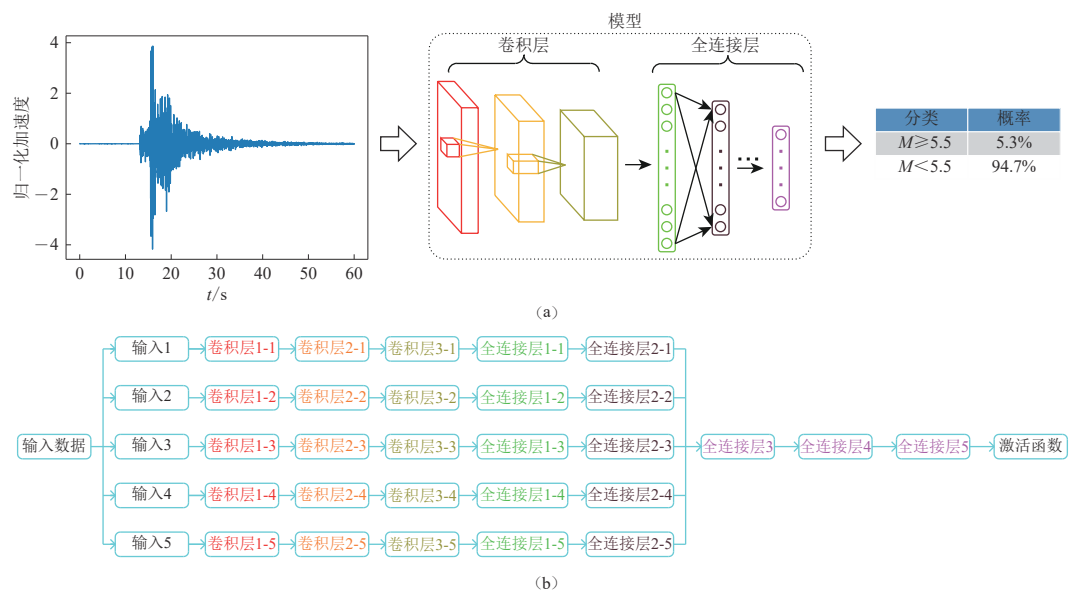


图2 神经网络模型从输入数据到获取分类结果的流程图

(a) 模型的流程说明; (b) 模型的架构

Fig. 2 Flow chart of neural network model from data inputting to classification result acquisition

(a) Process description of the model; (b) Model architecture

为倍数而增加。在三层卷积层之后，样本的五段数据形状为 16×10 。

因为全连接神经网络的输入是线性的，所以需要 对卷积层的输出结果进行线性处理。将卷积层输出的结果经过两个全连接层之后合并为一条线性数据，然后在三个全连接层之后输出两个结果，它们分别代表模型判断样本为大震或小震的概率。除最后一层之外，所有层的激活函数均使用 ReLu 激活函数。最后一层不使用激活函数，但是在计算损失函数时会添加一个 softmax[式(5)]层。Softmax 激活函数在规范化过程中使用指数形式，这意味着较大的值更大而较小的值会更小，增加了区分的对比度，可使模型训练更加有效，这对于分类问题尤其重要。

$$\text{softmax}(x_i) = \frac{e^{x_i}}{\sum_{j=1}^n e^{x_j}}, \quad i = 1, 2, \dots, n.$$

(5)

本文使用自适应矩估计优化器(adaptive moment estimation optimizer, 缩写为 Adam)来训练模型。该优化器针对 AdaGrad 和 RMSProp 的缺点弥补而来，具有以下优点：实现简单且计算效率较高，几乎无需调整超参数，可以自动调节学习速度，非常适合大型数据和参数模型训练。

本文模型所用训练集和验证集数据来自 K-NET 和 KiK-net，共有 11 万 9 760 个地震记录，涉及 1 698 个台站。按时间顺序排列，将 9 万 1 488 个记录作为训练集，2 万 8 272 个记录作为验证集，共 6 万 8 580 个大震记录和 5 万 1 180 个小震记录。

3 超参数调整

模型训练过程中，超参数的调整对训练效率和训练结果均会产生影响，本文涉及两个超参数：学习率和批量(batch size)。

1) 学习率. 学习率是模型每次减小损失函数值的程度, 学习率参数值设置得较大, 则模型前几次可能会很快收敛, 但学习率设置得过大, 可能使模型无法达到全局最优, 学习率参数值较小对训练效率也会产生较大的影响. 在保证其它变量相同的情况下用控制变量法测得模型训练过程中的准确率, 结果如图 3 所示. 可见: 当学习率为 0.001 (橙线) 时, 随着训练次数的增加, 训练集的准确率逐步提高, 测试集的准确率反而下降, 出现了过拟合现象, 也就是模型过度拟合训练集, 因而导致该模型在其它数据集上的表现不佳; 当学习率为 0.01 (蓝线) 和 0.0001 (绿线) 时, 训练集的准确率表现基本持平, 而学习率为 0.01 (蓝线) 时模型在测试集更胜一筹, 所以面对其它数据学习率为 0.01 (蓝线) 的模型可能会有更好的效果. 因此本文将学习率设置为 0.01, 每迭代一周学习率乘以 0.99, 随着训练次数增加, 学习率逐渐降低, 使模型损失更接近全局最小值.

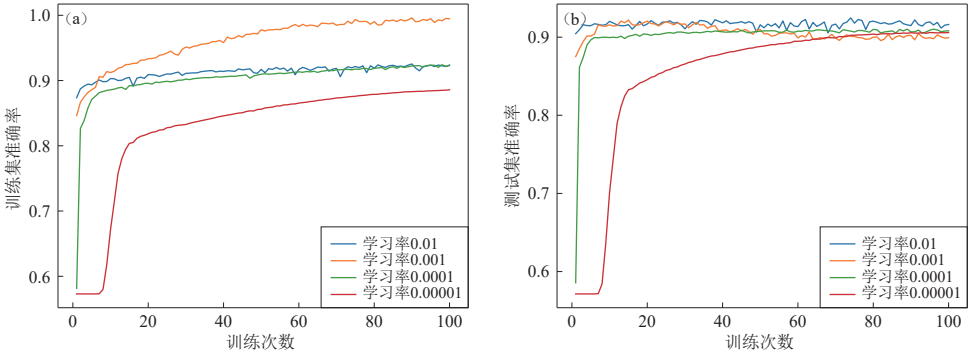


图 3 不同学习率下训练集(a)和测试集(b)的准确率随训练次数增加的变化
Fig. 3 The change in the correct rates of the trainings set (a) and the test set (b)
with the training time increasing on the condition of different learning rate

2) 批量(batch size). 批量是指每次输入模型的样本数量, 批量太小容易使模型收敛方向出现偏差, 太大容易使模型困在局部最优而无法达到全局最优. 根据训练集和测试集准确率的变化(图 4), 可以看到不同批量情况下均出现了程度不同的过拟合现象. 考虑到模型以较强的泛化能力和最低过拟合现象为佳, 我们选取批量为 400, 这种情况下模型在训练过程中

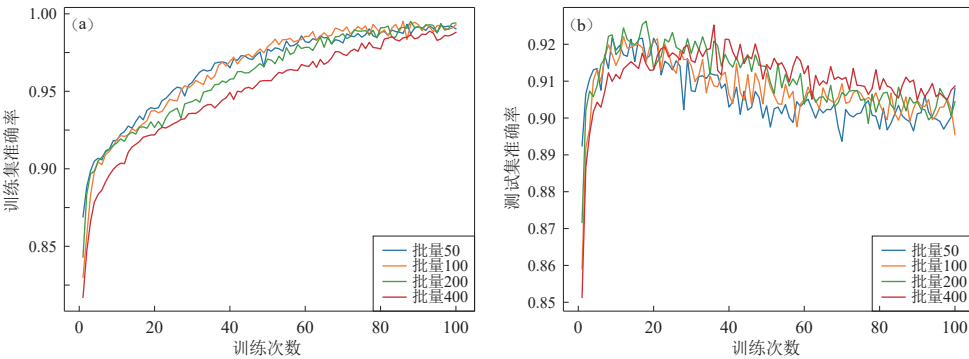


图 4 不同批量大小下训练集(a)和测试集(b)的准确率随训练次数增加的变化
Fig. 4 The change in the correct rates of the trainings set (a) and the test set (b)
with the training time increasing on the condition of different batch size

过拟合现象最轻且在测试集效果最好, 有较强的泛化能力. 这样我们就可以选择出现过拟合现象前的模型作为最终结果.

4 训练结果

本文提出使用 CNN 识别地震震级大小的方法. 在对原始地震数据进行筛选和归一化之后, 使用 CNN 模型对预处理后的地震记录进行识别和分类. 经过 100 次训练后, 将验证数据集上精度最高的模型用于分析. 模型准确率统计流程如图 5a 所示, 训练准确率随训练次数

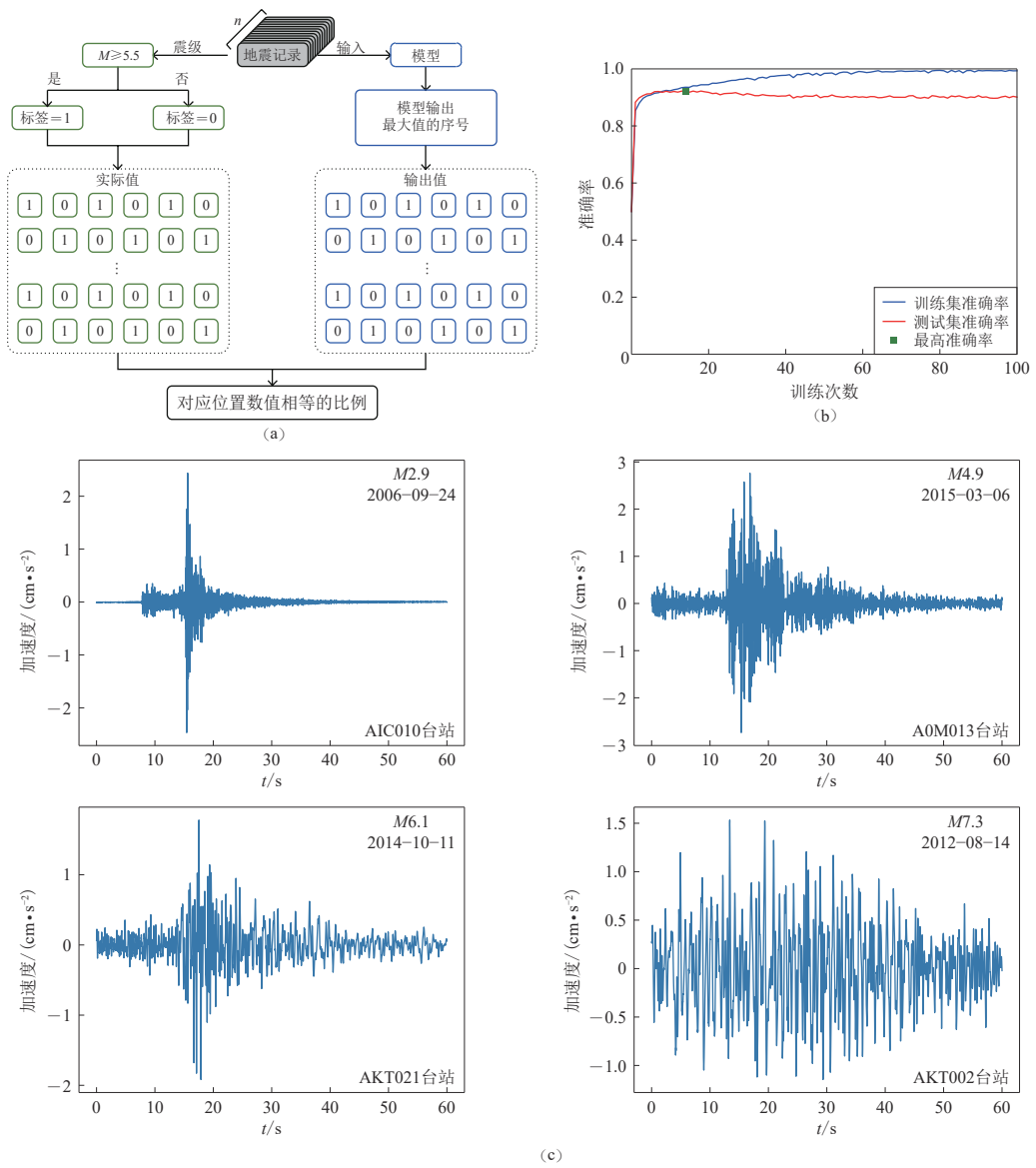


图 5 (a) 计算模型精度的流程; (b) 准确率; (c) 模型对于部分记录的识别结果
Fig. 5 (a) The flowchart of calculating model accuracy; (b) Training accuracy;
(c) The recognition results of the model for some recordings

的变化如图 5b 所示,部分记录的识别结果展示在图 5c 中,图中的四个记录分别来自 AIC010, AOM013, AKT021 和 AKT002 台站,地震分别发生于 2006 年 9 月 24 日,2015 年 3 月 6 日,2014 年 10 月 11 日和 2012 年 8 月 14 日。结果显示:基于 11 万 9 760 个记录进行训练,以 $M5.5$ 作为分界线进行分类,该模型在训练集上的平均准确率达到 93.6%,在测试集上的平均准确率达到 92.3%。根据统计,当地震震级处于 5.0—5.9 范围内即在 $M5.5$ 附近时,模型的准确率仅为 79.7%,这与平均准确率之间的差距较大,因此以 $M5.5$ 作为地震震级分类界限是可行的。

5 讨论与结论

本文以归一化的地震动记录为数据集来构建卷积神经网络,分类效果良好,初步得到以下结论:①模型进行了多次对比训练,均出现了不同程度的过拟合现象,可采用文中提到的早停法等策略解决;②超参数中的学习率大小宜适中,学习率过大,模型不易收敛,过小则训练较慢,批量过大容易造成局部最小,过小则训练过程中波动太大;③模型能够识别经过归一化的地震动记录的震级大小,说明地震的加速度时程记录带有一定的地震震级信息。

尽管该模型具有比较好的训练效果,但仍有一些问题需要解决,还有需要优化的方面:

1) 如何选择分类分界线来区分大地震与小地震。我们使用 $M5.5$ 作为模型的分类标准,仅基于简单的统计信息,且 $M5.0$ — 5.9 地震记录在模型中的平均准确率为 79.7%,远低于整个数据集的平均准确率,所以我们初步判断 $M5.5$ 作为分类界限有一定可行性,但可能会有更合适的震级界线来区分大小地震,应该也在 $M5.5$ 左右,未来我们会不断尝试优化模型,找到一个最优震级作为分类界限;

2) 本文使用的数据来自 K-NET 和 Kik-net,尚需验证该模型是否适用于其它地区。下一步会将来自其它国家地区的数据添加到训练数据中,以提高模型的泛化能力;

3) 未来我们会不断优化模型,并基于该模型测试常见的地震动模拟以及调整方法所得的模拟地震动或者经调整的小震记录能否被模型识别为大震,为这些方法的选择及模拟或调整效果提供参考。另一方面,基于该模型深入研究,不拘泥于二分类,可以做成大、中、小地震的三分类问题,甚至可以直接识别出大致震级等延展性工作。

参 考 文 献

- 李安, 杨建思, 彭朝勇, 郑钰, 刘莎. 2020. 基于样本增强的卷积神经网络震相拾取方法[J]. 地震学报, 42(2): 163–176.
- Li A, Yang J S, Peng C Y, Zheng Y, Liu S. 2020. Seismic phase identification using the convolutional neural networks based on sample enhancement[J]. *Acta Seismologica Sinica*, 42(2): 163–176 (in Chinese).
- 魏永刚, 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[J]. 地震学报, 41(5): 646–657.
- 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[J]. *Acta Seismologica Sinica*, 41(5): 646–657 (in Chinese).
- 谢礼立, 翟长海. 2003. 最不利设计地震动研究[J]. 地震学报, 25(3): 250–261.
- Xie L L, Zhai C H. 2003. Study on the severest real ground motion for seismic design and analysis[J]. *Acta Seismologica Sinica*, 25(3): 250–261 (in Chinese).
- 张肇诚, 张炜. 2016. 地震预报可行性的科学与实践问题讨论[J]. 地震学报, 38(4): 564–579.
- Zhang Z C, Zhang W. 2016. Discussion on scientific and practical problems of feasibility of earthquake prediction[J]. *Acta Seismologica Sinica*, 38(4): 564–579 (in Chinese).

- 周本伟, 范莉苹, 张龙, 李珀任, 房立华. 2020. 利用卷积神经网络检测地震的方法与优化[J]. 地震学报, **42**(6): 669–683.
- Zhou B W, Fan L P, Zhang L, Li P R, Fang L H. 2020. Earthquake detection using convolutional neural network and its optimization[J]. *Acta Seismologica Sinica*, **42**(6): 669–683 (in Chinese).
- 朱晓伟. 2011. 强震记录的选择与缩放方法研究[D]. 哈尔滨: 中国地震局工程力学研究所: 4–5.
- Zhu X W. 2011. *The Selection and Scaling of Strong Motion Records*[D]. Harbin: Institute of Engineering Mechanics, China Earthquake Administration: 4–5 (in Chinese).
- Arias A. 1970. A measure of earthquake intensity[G]//*Seismic Design for Nuclear Power Plants*. Cambridge: Massachusetts Institute of Technology Press: 438–483.
- Baker J W, Cornell C A. 2005. A vector-valued ground motion intensity measure consisting of spectral acceleration and epsilon[J]. *Earthq Eng Struct Dyn*, **34**(10): 1193–1217.
- Baker J W, Cornell C A. 2006a. Correlation of response spectral values for multi-component ground motions[J]. *Bull Seismol Soc Am*, **96**(1): 215–227.
- Baker J W, Cornell C A. 2006b. Spectral shape, epsilon and record selection[J]. *Earthq Eng Struct Dyn*, **35**(9): 1077–1095.
- Dysart P S, Pulli J J. 1990. Regional seismic event classification at the NORESS array: Seismological measurements and the use of trained neural networks[J]. *Bull Seismol Soc Am*, **80**(6B): 1910–1933.
- Goulet C A, Haselton C B, Mitrani-Reiser J, Beck J L, Deierlein G, Porter K A, Stewart J P. 2004. Evaluation of the seismic performance of a code-conforming reinforced-concrete frame building: From seismic hazard to collapse safety and economic losses[J]. *Earthq Eng Struct Dyn*, **36**(13): 1973–1997.
- Goulet C. 2005. *Improving the Characterization of Seismic Hazard for Performance-Based Earthquake Engineering Design*[D]. Los Angeles, California: Environmental and Civil Engineering, University of California: 246–268.
- Goulet C A, Haselton C B, Mitrani-Reiser J, Deierlein G, Stewart J P, Taciroglu E. 2006. Evaluation of the seismic performance of a code-conforming reinforced-concrete frame building: Part I: Ground motion selection and structural collapse simulation[C]//*8th National Conference on Earthquake Engineering (8NCEE)*. San Francisco, California: Earthquake Engineering Research Institute: 456–489.
- Huang L Q, Li J, Hao H, Li X B. 2018. Micro-seismic event detection and location in underground mines by using convolutional neural networks (CNN) and deep learning[J]. *Tunnel Underground Space Technol*, **81**: 265–276.
- Jordan M I, Mitchell T M. 2015. Machine learning: Trends, perspectives, and prospects[J]. *Science*, **349**(6245): 255–260.
- Krizhevsky A, Sutskever I, Hinton G E. 2017. ImageNet classification with deep convolutional neural networks[J]. *Commun ACM*, **60**(6): 84–90.
- Murphy K P. 2012. *Machine Learning: A Probabilistic Perspective*[M]. Cambridge: Massachusetts Institute of Technology Press: 1–16.
- Qian Y M, Bi M X, Tan T, Yu K. 2016. Very deep convolutional neural networks for noise robust speech recognition[J]. *IEEE/ACM Trans Audio Speech Language Process*, **24**(12): 2263–2276.
- Ross Z E, Yue Y S, Meier M A, Hauksson E, Heaton T H. 2019. Phaselink: A deep learning approach to seismic phase association[J]. *J Geophys Res: Solid Earth*, **124**(1): 856–869.
- Sainath T N, Mohamed A R, Kingsbury B, Ramabhadran B. 2013. Deep convolutional neural networks for LVCSR[C]//2013 *IEEE International Conference on Acoustics, Speech and Signal Processing*. Vancouver B C: IEEE: 8614–8618.
- Sercu T, Puhrsch C, Kingsbury B, LeCun Y. 2016. Very deep multilingual convolutional neural networks for LVCSR[C]//2016 *IEEE International Conference on Acoustics, Speech and Signal Processing*. Shanghai: IEEE: 4955–4959.
- Tiira T. 1999. Detecting teleseismic events using artificial neural networks[J]. *Comput Geosci*, **25**(8): 929–938.
- Tóth L. 2013. Convolutional deep rectifier neural nets for phone recognition[C]//*Proceedings of the Annual Conference of the International Speech Communication Association, Interspeech*. Lyon: International Speech Communication Association (ISCA): 1722–1726.
- Ursino A, Langer H, Scarfi L, Di Grazia G, Gresta S. 2001. Discrimination of quarry blasts from tectonic microearthquakes in the Hyblean Plateau (Southeastern Sicily)[J]. *Ann Geophys*, **44**(4): 703–722.

- Wiszniowski J, Plesiewicz B M, Trojanowski J. 2014. Application of real time recurrent neural network for detection of small natural earthquakes in Poland[J]. *Acta Geophys*, **62**(3): 469–485.
- Yu D, Xiong W, Droppo J, Stolcke A, Ye G L, Li J Y, Zweig G. 2016. Deep convolutional neural networks with layer-wise context expansion and attention[C]//*Proceedings of the Annual Conference of the International Speech Communication Association, Interspeech*. Lyon: International Speech Communication Association (ISCA): 17–21.
- Zhang G Y, Wang Z Z, Chen Y K. 2018. Deep learning for seismic lithology prediction[J]. *Geophys J Int*, **215**(2): 1368–1387.
- Zhu L J, Peng Z G, McClellan J, Li C Y, Yao D D, Li Z F, Fang L H. 2019. Deep learning for seismic phase detection and picking in the aftershock zone of 2008 M_W 7.9 Wenchuan earthquake[J]. *Phys Earth Planet Inter*, **293**: 106261.