结合台阵策略的震相拾取深度学习方法

刘芳¹, 蒋一然², 宁杰远^{2*}, 张建中¹, 赵艳红¹

1. 内蒙古自治区地震局, 呼和浩特 010010;

2. 北京大学地球与空间科学学院, 北京 100871

* 联系人, E-mail: njy@pku.edu.cn

2019-10-09 收稿, 2020-02-19 修回, 2020-02-20 接受, 2020-03-20 网络版发表

中国地震局地震科技星火计划(XH18012)、内蒙古自治区科技重大专项和中国地震科学实验场专项(2019CSES35)资助

摘要 深度学习的突破性发展及其在地震学领域的初步应用, 为有效处理和利用地震资料提供了可能. 震相拾取是地震资料处理中的基础性工作, 目前已经提出了很多基于深度学习的震相拾取方法, 取得了很好的效果. 但是, 为了满足地震学研究中处理连续地震记录的需要, 尚需对方法进行进一步的改进. 本研究针对性地设计了结合台阵策略, 单独识别P波和S波的长时窗震相拾取深度学习模型PP(phase picker)及其训练方式, 提出了具有实用性的震相拾取方法APP(array-assisted phase picker). 利用阿里余震AI捕捉大赛和Hi-net数据进行测试的结果表明, 模型能够有效地在连续波形上拾取体波震相并具有很好的泛化能力. 通过比较该模型与其他模型(较短时窗的模型和同时识别P波、S波的模型)的拾取效果, 验证了模型设计的合理性. 具体的测试样例显示, 该方法能够正确地处理地震密集的波形数据并能避免典型噪声的影响. 将该方法运用到内蒙古地区台网的观测数据中, 检测到了人工目录中98.1%的地震, 地震拾取总数为人工目录数的30倍, 进一步表明本研究方法具有很好的实用性.

关键词 深度学习, 震相拾取, 台阵策略, U型网络

地震资料的积累速度日益增长, 为地震学研究提供了海量资料. 以我国为例, 国民经济和社会发展第十个五年计划以来, 固定台站数目不断增长, 并开展了大量流动观测^[1]. 人工处理难以完成震相识别任务, 需要发展高精度、可实用的自动化资料处理方法.

近年来, 随着深度学习方法的突破性发展, 许多深度学习方法被应用到震相自动拾取领域, 如卷积神经网络CNN(convolutional neural network)^[2~7]、递归神经网络RNN(recurrent neural network)^[8]、全连接神经网络FCN(fully convolutional network)^[9]和U形网络U-Net^[10,11]. 这些深度学习方法, 借助于复杂的网络结构, 能够直接从波形中提取特征, 建立波形与拾取结果间的映射关系. 在使用大量的震相资料进行训练之后, 这

些方法表现出远超传统方法的拾取效果, 使得震相拾取工作自动化变为可能^[12].

Zhu等人^[4]基于卷积神经网络CNN提出了一种震相拾取方法, 利用某一时间点前后的波形信息对于该点所处的类别作出判断, 进行震相拾取. Zhou等人^[8]将波形看作有前后关联的时间序列, 利用RNN对时间序列的阶段变化作出预测, 并根据每个阶段开始的时间拾取震相. 基于U-Net的震相拾取方法^[10,11]则是进行图像分割, 识别出波形的不同类别, 并标注出不同类别间的分界点. 其中, U-Net^[13]是在FCN^[14]基础上发展起来的一种端对端的神经网络. 借助于包含收缩路径和扩展路径的网络结构, U-Net既能准确地识别目标对象又能对其边界进行精准定位, 能够很好地处理图像分割

引用格式: 刘芳, 蒋一然, 宁杰远, 等. 结合台阵策略的震相拾取深度学习方法. 科学通报, 2020, 65: 1016–1026

Liu F, Jiang Y R, Ning J Y, et al. An array-assisted deep learning approach to seismic phase-picking (in Chinese). Chin Sci Bull, 2020, 65: 1016–1026, doi: 10.1360/TB-2019-0608

问题. 这种网络结构非常适合震相拾取问题: 既可以准确地对波形分类, 又可以精确地确定震相的起始位置.

Zhu和Beroza^[10]首次将U-Net应用于震相拾取, 提出了PhaseNet方法, 并取得非常好的效果. 但是, PhaseNet在设计和训练的过程中都是以事件为核心进行的, 能够在有单一地震的波段出色地完成震相到时拾取工作, 但在直接应用到在连续波形上进行震相识别时面临困难. 赵明等人^[11]尝试将与PhaseNet类似的U-Net震相拾取模型用于连续波形, 取得了初步成功, 但仍然被实际数据中复杂的噪声所困扰. 为保证正确率, 他们设定了较高的阈值(95%), 但却因此显著地降低了召回率, 极大地影响了方法的实用性, 因此需要做针对性的改进.

在本研究中, 我们从网络结构、样本集构造等方面对U-Net型的震相拾取模型进行改进, 设计了震相拾取模型PP(phase picker), 以更好地处理连续记录. 进一步地, 将该网络与蒋一然和宁杰远^[15]所提出的台阵策略相结合, 利用多台约束, 将不同台的拾取结果匹配起来, 并剔除其中的误识别结果, 以进一步提升方法的准确率和处理复杂情况的能力, 形成了震相拾取方法APP(array-assisted phase picker). 使用余震捕捉大赛^[16]和Hinet^[17,18]所提供的两个实际数据集, 对模型进行训练和交叉验证, 以测试模型的拾取效果和泛化能力. 最后, 将模型运用到内蒙古地区的流动台阵观测资料分析中, 并将拾取结果与人工目录进行了比较, 进一步验证了方法的实用性.

1 震相自动拾取方法

1.1 PP的网络结构及其样本集的构造

以往, 基于U-Net的震相自动拾取方法的网络和训练都是以事件为核心设计的, 样本由标注好P波、S波的事件构成, 模型也需要同时给出某一时刻出现P波、S波和噪声的概率^[10,11]. 但这样做会面临3个问题. (1) 在训练时, 样本必须同时给出P波和S波的概率, 只含有P波标注的地震记录无法用于训练. (2) 在测试或在连续波形上进行震相识别时, 为了将同属于一个事件的P震相和S震相包含在同一时间窗内, 样本往往是P波在前、S波在后, 导致模型在判断时会倾向于认为在前的P波, 在后的是S波. 但在应用训练好的模型按某一长度的时间窗扫描连续记录进行震相识别时, 不能保证时间窗总将P和S震相同时包含进来, 与样本中P波在

前、S波在后的模式存在差异, 从而容易发生识别错误. (3) 如果将二者放在同一模型中进行识别, 则对于一类震相的误识别会干扰对于另一震相的识别, 从而降低整体的拾取效果(详见第2节).

为了使得模型能够更好地满足扫描连续波形的需求, 并能够适用于更多的地震资料, 我们设计的PP模型由两个网络组成, 分别对P波和S波进行识别. 这时, 每一个网络的输出维度变为1. 这样, 以上三方面的困难就自然地得到了克服.

赵明等人^[11]将U-Net震相拾取模型用于连续波形震相识别的研究中, 时间窗选择较短, 无法利用更长时间窗内的噪声和震相信息, 不能很好地处理更大震中距的地震, 不适用于台站分布和监测范围更广的台网.

为了克服这一问题, 本研究将模型设计的信号输入长度拓展为40 s, 便于模型从更长的时间窗的记录中获取噪音、震相信息, 从而更好地确定震相及其到时. 为了不增加模型规模和计算量, 本研究将模型输入信号的采样率设为50 Hz. 这样做, 同时也增加了模型的适用范围, 使其能够处理50 Hz及以上采样率的数据. 同时, 为了能够尽可能多地使用更长时间窗内的信息, 从而使得识别效果更好, 我们针对性地采用了更深的U型网络结构, 其中最深一层的数据在时间维度上有5个点.

PP网络结构如图1. 模型的直接输出是一个维度的, 分布在 $(-\infty, +\infty)$ 的区间内. 为了表示对象属于目标类别的概率 $q(x)$, 需要把网络直接输出 $y_0(x)$ 映射为 $q(x)$. 因此, 我们设置输出层的激活函数为sigmoid函数, 其形式如下:

$$q(x) = \frac{1}{1 + e^{-y_0(x)}}, \quad (1)$$

其中, x 为输入, $q(x) \in (0, 1)$. 当 $y_0(x)$ 趋于正无穷时, 概率 $q(x)$ 趋近于1, 表示为震相的概率高; 趋近于负无穷时, 概率 $q(x)$ 趋近于0, 表示为噪音的概率高.

类似于Zhu和Beroza^[10]的研究, 将P、S波的到时转化为时间上的高斯概率分布, 称作真实概率, 作为模型的拟合目标, 用 $p(x)$ 表示. 在本研究中, 真实概率设置为以人工拾取到时为中心的高斯函数, 如图2(a), (b)所示. 高斯函数的形态由拾取精度决定. 这里, P的拾取精度较高, 高斯函数的 σ 设为0.1 s; S的拾取精度较低, 高斯函数的 σ 设为0.2 s.

图1 PP模型的网络结构. 彩色的框和连接线表示模型内不同类型的层; 虚线框标识出模型中针对性设计的部分

Figure 1 The structure of the PP model. Boxes and lines with different colors represent different types of layers in the model; dashed boxes mark the parts of the model specifically designed in this paper

由于PP模型仅输出一个类别的概率, 我们需要定义与之相适应的损失函数. Zhu和Beroza^[10]的研究中, 使用交叉熵定义了模型的损失函数. 考虑到训练样本中有震相的部分所占的比例较低, 本研究在交叉熵的形式上, 增加了有震相部分的权重以均衡样本, 得到如下的损失函数 $F(x)$:

$$F(x) = -\sum_x w_0 \times p(x) \times \log(q(x)) + w_1 \times (1 - p(x)) \times \log(1 - q(x)), \quad (2)$$

其中, $p(x)$ 表示前述真实概率分布, $q(x)$ 表示模型输出的属于该类别的概率, w_0 和 w_1 分别为对震相和噪音两个类别设置的权重. 经过测试, 设置 w_0 为7.1, w_1 为1.1, 以使不产生大量误识别的情况下较好地均衡样本.

同时, 本研究对样本集的构造作了如下设计: (1) 将样本的水平分量进行随机旋转, 并随机地选取时间窗. 这样做, 一方面, 可以产生更多的样本; 另一方面, 随机选取时间窗意味着会产生一部分完全不包含震相的样

本. 在实际的连续波形中, 无震记录的比重远大于震相, Zhu和Beroza^[10]方法的网络训练时以事件为中心, 未对这部分充分训练, 容易发生误识别. 使用新的构造方法, 能够使模型更好地适应震相所占比例更少的连续数据. (2) 有一些台站建在了环境噪声较大的区域. 我们使用的含有人工拾取到时的地震数据, 往往来自于质量较好的记录, 几乎很少采样这类噪声. 为了避免模型因缺乏这类样本而在实用时产生大量误识别, 我们人工地增加了一些噪声高的数据到样本中. 需要特别指出的是, 因为其他模型在震相拾取时也面临类似的需求和问题, 这种样本集的构造设计对其他模型也同样适用.

1.2 APP:PP模型与台阵策略的结合

如前所述, 在震相拾取工作中, 很多研究首先从连续波形中对地震进行检测, 挑选出具有地震的波段, 然后在这个波段上拾取P波和S波^[4,8]. 这样处理的问题在于, 难以选择检测地震所用时间窗长: 过短的时间窗会

图 2 波形及其输入、输出示例. (a), (b) 分别展示P波和S波的真实概率分布; (c) 在JP测试集上拾取P波; (d) 在SC测试集上拾取P波; (e) 在JP测试集上拾取S波; (f) 在SC测试集上拾取S波; (g), (h) 分别为S波识别干扰P波识别和P波识别干扰S波识别的两个例子; (i) 处理地震密集的波形数据; (j) 对典型噪声的免疫力测试. E、N、Z为波形的三分量; $p_p(x)$ 和 $p_s(x)$ 分别为P波和S波的真实概率; $q_p(x)$ 和 $q_s(x)$ 分别是PP模型输出的P波和S波的概率; $q_p'(x)$ 和 $q_s'(x)$ 分别是P波和S波一起识别的模型输出的P波和S波的概率; q_{pn} 和 q_{sn} 表示未加入典型噪声的模型输出的P波和S波概率; q_{pw} 和 q_{sw} 表示加入典型噪声的模型输出的P波和S波概率; 在(g)和(h)中用框分别标注了P波和S波震相出现的部分

Figure 2 Samples of waveforms with their inputs and outputs. (a), (b) The true probability distribution of P phase and S phase; (c) picking P-waves on the JP test set; (d) picking P-waves on the SC test set; (e) picking S-waves on the JP test set; (f) picking S-waves on the SC test set; (g), (h) examples of S-picking interfering P-picking, and P-picking disturbing S-picking; (i) an example showing event-crowded waveform data processing; (j) an example showing anti-noise-interference ability. E, N and Z are three-component waveforms; $p_p(x)$ and $p_s(x)$ show the true probability of P and S phases; $q_p(x)$ and $q_s(x)$ show the probabilities of P and S phases from model PP; $q_p'(x)$ and $q_s'(x)$ show the probabilities of P and S phases from the model which simultaneously picks P and S phases; q_{pn} and q_{sn} represent the probabilities of P and S phases from the model without typical noise; q_{pw} and q_{sw} represent the probabilities of P and S phases from the model with typical noise; in (g), (h), the boxes are used to mark the parts of the P- and S-phases, respectively

使得方法只能处理震中距较小、波场持续时间较短的地震; 较长的时间窗又会在处理到时较为接近的多个地震时遇到困难. 我们在1.1中提出的使用两个模型分别对P波和S波进行识别的PP模型, 不依赖于时间窗的准确选取, 可以给出一系列P波和S波的到时, 但并未将同一台站上属于同一地震的P波和S波匹配起来, 不同台站上的震相也未被联系到具体对应的地震上.

蒋一然和宁杰远^[15]在Chen和Holland^[19]的基础上, 发展出一种以P、S震相到时差作为时空约束的台阵策略. 先由自动拾取算法给出连续记录里不同台站上的P、S震相的到时信息; 再利用P波、S波在同一台站上的到时差计算地震的发震时刻和位置, 根据不同台站

估计的发震时刻和位置的一致性将不同台站上属于同一地震的震相匹配起来, 最终完成对地震的识别和对震相到时的拾取. 这样可以不先在单台上把属于同一地震的P、S波匹配起来, 而是借助台阵约束来完成这种匹配, 通过更强的约束, 让方法具有了处理复杂情形的能力. 在单台上, 以现有方法和样本, 在很多情况下对震相、噪声几乎不可能做出准确的判断, 始终会有大量的误识别产生. 借助于台阵方法, 可以去除掉绝大部分孤立出现的误识别信号, 能够显著提高拾取的准确率. 本研究因此进一步借鉴蒋一然和宁杰远^[15]的思路, 提出将本研究的PP模型与台阵策略相结合, 减少误识别的比例, 形成震相拾取方法APP(array-assisted

phase picking), 具体拾取流程如下: (1) 使用训练好的PP模型的两个网络, 对三分量数据进行滑动扫描, 分别识别P波和S波, 取输出结果大于0.5的峰出现的位置作为P波、S波的到时. 时间窗每次滑动10 s, P波取每个时间窗中15~25 s时间段的输出结果, S波取每个时间窗中20~30 s时间段的输出结果. (2) 寻找同一台站上到时差小于80 s的P、S震相, 并根据到时差估算发震时刻和台站与震中的距离. (3) 根据蒋一然和宁杰远^[15]的台阵策略, 将研究区域划分为多个子区域, 联合不同台站的估计结果, 取发震时刻和震中位置一致性最好的时刻和子区域作为发震时刻和震中位置. (4) 将不同台站上满足发震时刻和震中位置的拾取结果和相应的地震匹配起来, 剔除余下未被匹配的震相, 得到最终的地震和震相到时目录.

2 PP模型的训练与测试

本研究使用的数据由两部分构成: 余震捕捉AI大赛^[16]主办方中国地震局地球物理研究所和阿里云计算有限公司在2017年8月公布的复赛阶段使用的人工拾取震相20647条, 用SC表示; 在日本的Hi-net^[17,18]网站上下载的125°~150°E、25°~50°N范围内的29234个地震的波形数据, 并截取其中人工拾取的P、S震相数据各331000条, 用JP表示. 对所有样本都进行2~15 Hz滤波、去均值, 按标准差进行归一化. 这两部分数据来源于不同地区的不同台网, 在地质背景、地震类型、建台环境、仪器型号、数采响应等方面都存在一定的差别. 使用这两部分数据, 进行交叉训练与测试, 可以帮助我们更好地衡量模型的泛化能力.

图2(c)~(f)展示了模型分别在JP和SC测试集上的输出样例, 其中的(d)、(f)图展示了两个时间窗选取十分不好的情况下, 模型的拾取结果. 在图(d)中, P波出现在窗的最末端, 在图(f)中, S波出现在窗的前半部分, 且

窗内没有P波. 在这两种情况下, 模型都成功地对震相予以识别. 当然, 为了克服在时窗前端拾取效果变差的问题, 需要在APP方法拾取流程中有重叠地在连续波形上滑动时窗, 并保留每个时窗内5 s以后的拾取结果, 舍去前端的拾取结果.

为了分析用P波和S波两个模型分别进行拾取、拓宽时窗等设计对震相拾取效果的影响, 将PP模型(40 s时窗)与之前P波、S波一起识别的模型(40 s时窗), 以及30 s时窗模型和24 s时窗模型的拾取效果进行了比较, 其结果如表1. 这里, 使用1/1 JP训练集训练4个模型, 并在JP测试集上测试, 最大误差设置与前文一致. 为了避免短时窗模型在边界上拾取效果下降的问题, 仅取时窗内震相到时在5~20 s内的样例进行统计.

可以看出, 本研究单独识别P波和S波的模型, 在拾取效果上明显好于过去一起识别的模型: 正确率 P 相当的情况下, 召回率更高、拾取误差更小. 我们认为这种表现的差异一部分是因为将P波、S波一起识别时, 对其中一类震相的误识别会影响对另一类震相的识别. 图2(g)、(h)展示的就是震相识别相互干扰的两个例子. 单独识别的模型给出了较为准确的结果, 而一起识别的模型则因为对另一震相的误识别而做出了错误的判断. 这说明将二者分开识别对提高震相拾取的正确率是有效的.

表1中, 40 s时窗的模型略优于30 s时窗和24 s时窗的模型. 实际上, 当P-S到时差大于20 s时, 40 s时窗模型对P波和S波的召回率 R 分别为: 0.926、0.766; 30 s时窗模型则为0.911、0.755; 24 s时窗模型为: 0.895、0.766. 三者中, 40 s时窗的模型能更好地处理远震中距时的P震相拾取; 而对S波的拾取, 三者效率相当. 此外, 在扫描连续波形时, 需要有重叠的移动时窗以减少边缘拾取率下降的影响. 而长时窗模型需要舍去的部分所占的比例比短时窗的模型更低. 在保证拾取效率不

表1 4种不同模型在JP测试集的P波、S波拾取效果(P , R , F_1 , 均值, 标准差)

Table 1 P and S phase pickup performance (P , R , F_1 , mean, std) of four different models in the JP test set of four models with different structure on P and S phases in JP test set

震相指标	P波					S波				
	P	R	F_1	均值(s)	标准差(s)	P	R	F_1	均值(s)	标准差(s)
单独识别P/S波(40 s时窗)	0.984	0.974	0.979	0.015	0.092	0.898	0.891	0.894	0.031	0.138
一起识别P/S波(40 s时窗)	0.981	0.958	0.969	0.037	0.102	0.897	0.873	0.885	0.041	0.139
30 s时窗	0.982	0.970	0.976	0.023	0.090	0.895	0.884	0.889	0.037	0.141
24 s时窗	0.984	0.965	0.975	0.014	0.086	0.898	0.887	0.892	0.029	0.136

下降的前提下, 40 s长时窗的模型也在一定程度上提高计算效率。

SC数据包含了大量的汶川地震的余震, 发震频繁, 不同地震震相交错在一起。我们将模型直接用于SC数据的连续记录里, 并发现了大量如图2(i)例子所示的情况。3个地震的到时非常接近, 后到地震的P波处在先到地震的尾波中。如使用先至地震, 再在其附近进行拾取的方法则难免会漏掉后面两个地震。而本研究的模型仍然能够准确地给出各个震相的到时。然后, 结合台阵策略就可以对其进行分别拾取。

实际的观测数据有时候是由不同小段记录拼接而成, 数据在拼接前后可能产生较大的突变, 在滤波后, 信号呈现的模式与地震信号很相似, 易被误识别, 是前面提到的典型噪声中的一种。比较了训练集中加入与不加入典型噪声得到的模型在这类噪声上表现的差异, 如图2(j)。在没有加入典型噪声时, 模型对这类信号产生了大量的误识别, 而训练集中加入部分典型噪声时, 效果得到改善。这说明收集多种类的典型噪声样本, 并加入训练集中, 将减少模型的误识别, 对于模型应用到数据情况较复杂的流动台网上是非常必要的。

为了测试在不同训练集和训练规模下, PP模型在不同数据集上的识别效果, 同时研究其泛化能力, 本研究针对P、S分别设计了8组具有不同规模和数据来源的训练集。P波这8组训练集为: SC(8633条), 1/128 JP(2500条), 1/64 JP(5000条), 1/32 JP(10000条), 1/16 JP(20000条), 1/4 JP(80000条), 1/1 JP(320000条), SC(8633条)+1/1 JP(320000条)。对于S波, 这8组训练集为: SC(8576条), 1/128 JP(2500条), 1/64 JP(5000条), 1/32 JP(10000条), 1/16 JP(20000条), 1/4 JP(80000条), 1/1 JP(320000条), SC(8576条)+1/1 JP(320000条)。测试集为: 10000条JP样本; 1000条SC样本。

本研究统计了模型拾取与人工拾取差在1.5 s内的结果, 误差的分布如图3(a)~(d)。可以看出, 由SC训练集得到模型的拾取结果早于由JP训练集得到的模型: 在JP测试集上, SC训练集得到模型的拾取结果偏早; 在SC测试集上, JP训练集得到模型的拾取结果偏晚。经过对具体波形的查看, 发现这里的差异是由于两个台网人工拾取习惯造成的, SC数据集的人工拾取会偏早一些。对于震相拾取而言, 一个稳定的小的到时偏差, 对于定位等方法没有影响, 也可以通过标定很容易地消除这种偏差。我们对同一组拾取结果都采用去均值处理, 再评价其拾取结果。这里的均值取到时差在2 s以

图3 不同训练集得到模型在不同测试集下的测试结果。(a)~(d) 拾取误差分布: (a) 在JP测试集上拾取P波; (b) 在SC测试集上拾取P波; (c) 在JP测试集上拾取S波; (d) 在SC测试集上拾取S波。(e)~(i) 不同训练集得到模型在不同测试集下的拾取效果(P , R , F_1 , 均值和标准差数值随训练集的变化); 横坐标为前文提到的8个训练集

Figure 3 Results of the models trained by different training sets on different test sets. (a)~(d) Picking error distributions: (a) P-wave picking on the JP test set; (b) P-wave picking on the SC test set; (c) S-wave picking on the JP test set; (d) S-wave picking on the SC test set. (e)~(i) Performance (P , R , F_1 , mean and std) of the models trained by different training sets on different test sets; the abscissa shows the eight training sets mentioned above

内结果的平均值。

为了定量地对比模型的拾取效率, 采用与Zhu和Beroza^[10]相同的评价标准, 下面为正确率 P 、召回率 R 和综合正确率 F_1 的表达式:

$$P = \frac{T_p}{T_p + F_p}, \quad (3)$$

$$R = \frac{T_p}{T_p + F_n}, \quad (4)$$

$$F_1 = \frac{2 \times P \times R}{P + R}. \quad (5)$$

当预测值 >0.5 , 且其拾取的到时与人工标注差小于等于最大到时差(max dt)时, 视为正确的正类预测, T_p+1 ; 当预测值 >0.5 , 且其拾取的到时与人工标注差大于最大到时差(max dt)时, 视为错误的正类预测, F_p+1 ; 当预测值 <0.5 , 但实际存在震相时, 视为错误的负类预测, F_n+1 .

我们对 P 、 R 、 F_1 、均值、标准差设置最大拾取误差为 0.5 s, 统计拾取结果所对应的 P 、 R 、 F_1 、均值、标准差, 结果如图3(e)~(i). 研究评价参数的变化, 我们发现如下结果.

(1) 在SC测试集上, 与人工结果相比, 多个模型拾取P震相的标准差在 0.15 s左右(设置最大拾取误差为 1 s时, 为 0.2 s左右), 拾取S震相的标准差在 0.2 s左右(设置拾取最大误差为 1 s时, 为 0.3 s左右), 对比使用同样数据集的研究, 效果好于Zhou等人^[8], 与Zhu等人^[4]

持平.

(2) 模型识别效果随训练集规模的增长而缓慢增长, 128 倍的规模增长带来的拾取效果(P , R , F_1)的提高在 2% 以内. 如果对拾取精度要求不高, 使用小数据量的样本训练就能够达到可接受的水平; 如果对拾取精度要求较高, 可使用更大规模的样本, 以取得更好的拾取效果.

(3) 统计中, 同一组训练集得到的模型在JP测试集上的表现(P , R , F_1 , 标准差)优于在SC测试集上的表现. 我们人工地对数据集中部分波形和标注进行查看, 发现JP数据集在信噪比、人工拾取精度上好于SC数据集.

(4) 在SC测试集上, 仅使用JP训练集的模型表现(P , R , F_1)比仅使用SC训练集的模型对P波的拾取效果更好, 对S波的拾取效果二者相当; SC训练集得到的模型在人工拾取质量更好的JP测试集上比其在SC测试集上表现(P , R , F_1 , 标准差)更好.

(5) SC+JP训练集在SC测试集上的表现(P , R , F_1)略优于仅使用JP数据. 这说明适当使用当地数据, 有助于

图4 模型对于P波(a, c, e)和S波(b, d, f)的拾取效果随P-S波到时差($t_s - t_p$)、信噪比(SNR)、震相在时窗中位置(arrival time)的变化. 深色柱体表示测试集中样本的总数量; 浅色柱体表示被模型检测出的样本数量; 黑色折线为反映模型召回率的检出数量与总数量比值

Figure 4 Variation of the model's performance on picking P phases (a, c, e) and S phases (b, d, f) for different P-S arrival time difference ($t_s - t_p$), signal noise ratio (SNR) and phase arrival time in the time window (arrival time). Dark bars indicate the total number of samples in the test set; light bars indicate the number of samples detected by the model; black polylines represent the ratio of the detected number to the total number, which reflects the recall rate of the model

提高拾取的效果.但这种效果的提升非常有限,即便仅仅使用JP数据也可以达到理想的拾取效果.得益于更高的信噪比和人工拾取精度,在更小的训练集规模下1/128JP得到的模型比SC训练集得到的模型在两组测试集上表现(P, R, F_1)总体更优.这说明采用更高精度的数据集,是提高U-Net这类网络结构识别效果的有力手段.也说明,因为U-Net网络较少的参数和端对端充分利用样本信息的优势,在小样本下就能够对参数进行很好的训练.

图4展示了使用1/1 JP训练集得到的模型在JP测试集上震相拾取效果随P波和S波到时差(震中距)、信噪比、震相在时间窗内所处位置的变化,这里最大到时差设置为0.5 s.可以发现模型对于震中距较近的记录拾取效果较好;较远的记录拾取效果变差.对误识别的

样例进行查看,发现距离较远时,样例中存在首波等复杂震相,对震相拾取造成了影响.图4(e), (f)反映出,在时窗内的大部分时间,震相拾取效果都很好,只有当震相十分靠近时间窗前端的时候,拾取效果显著变差.

经过上述的比较和分析,我们认为用高精度JP数据训练得到的PP模型,能够处理复杂的连续数据,并具有直接异地使用的可能.

3 应用和结果分析

将上述流程初步应用到内蒙古地区的台阵资料上.数据由2014年1月1日~12月31日,96个喜马拉雅二期的流动台、11个甘肃地震局的固定台和2个内蒙古自治区测震台网的固定台的三分量记录构成.台阵分布如图5中红色三角形所示.该台阵的资料较复杂,有的台

图5 内蒙古台阵拾取结果.蓝色小点代表拾取到的地震;黑色和红色圆圈表示人工目录中地震(黑色圆圈:被模型检测出地震,红色圆圈:未被模型检测);红色三角形表示台站,黑色线条表示断层分布,绿色方框表示用于比较自动拾取结果和人工目录的区域

Figure 5 Picking results by the Inner Mongolia array. The blue dots represent the picked earthquakes; the black or red circles represent the earthquakes in the manual catalog (black circle, detected by model; red circle, not detected by model); the red triangles represent the stations, the black lines indicate the fault distribution, and the green squares represent the areas used to compare the automatic picking results with the manual catalog

站会出现大量的环境噪声和不连续记录,对自动拾取算法要求较高。

从109个台站上365天的记录中一共检测到31862个地震,拾取到440142个P震相,313050个S震相。地震位置分布如图5中蓝色小点所示(剔除了一些台站数较少,不能定位的地震),图6(a), (b)是拾取结果中的一个例子。

在研究区域的南部分布有较多断层,检测到的地震多沿断层展布。北部已知断层较少,但也观测到了不少地震,与地形有一定关联。总体而言,拾取结果反映了该地区的地震活动性随空间的变化。

内蒙古自治区地震局该区域内2014年的地震目录,其位置分布如图5黑色圆圈和红色圆圈所示。地震目录

记录的地震主要分布于南部断层较多的区域。地震检测能力受台站分布的影响,为了更合理地分析模型拾取地震的能力,选取图5中绿色方框内的研究区域,比较在此区域内,自动拾取结果与人工地震目录异同。

(1) 模型自动拾取到13444个地震,人工目录共记录到424个地震。将两个目录进行匹配,模型一共找到了人工目录中的416个地震,召回率为98.1%。8个未被模型检测出的地震(如图5红色圆圈所示)都处于台站覆盖较差的区域。其中,有2个地震的波形被模型拾取到了,但由于射线覆盖不好,对震中和发震时刻与目录偏差较大,因而未匹配上目录中相应的地震。其余6个地震的震级最大为1.2,能量较小,信噪比较低。

(2) 图6(c), (d)展示自动拾取到的地震和人工目录

图6 内蒙古台阵拾取结果示例和统计。P波(a)和S波(b)按拾取到的到时对齐,并以台站震中距排列;(c)机器拾取结果的震级分布;(d)人工目录的震级分布;(e)拾取结果的台站记录数分布。在(c)和(d)中ml表示震级;在(e)中,n表示记录到地震的台站的数目

Figure 6 Samples and statistics of the picking results in the Inner Mongolia array. The P waves (a) and S waves (b) are aligned by the arrival times and lined by the station's epicentral distances; (c) the magnitude distribution of the picked-up results; (d) the magnitude distribution of the manual catalog; (e) the distribution of the number of stations in the picking results; in (c) and (d), ml indicates the magnitude; in (e), n indicates the number of stations which record the earthquake

中地震的震级分布. 从图中可大致估计, 模型在内蒙古台阵中地震检测能力大概在0~1级之间, 而人工目录所展示的地震监测能力在1~2级之间. 自动拾取模型将台网的地震监测能力提高了1级左右.

(3) 用 n 表示记录到地震的台站的数目, 其分布如图6(e). 可以看到, 存在大量被10个以上台站记录到的地震. 这为浅层速度结构反演、断层形态研究等提供了丰富的地震资料.

使用其他台网的数据训练得到的模型在该台阵具有良好的拾取效果, 进一步证明了该模型具有较强的

泛化能力, 也体现了震相拾取算法的实用价值.

4 结论

结合PP模型和台阵策略的震相拾取方法具有在实际连续波形上准确识别地震和精确提取震相到时的能力, 并在中国四川省、日本及其周边、中国内蒙古自治区3个区域的实际地震资料应用中展示出较强的泛化能力. 该方法可直接用于台阵资料, 并提供高精度的地震目录和震相到时数据. 相关的程序已发布到如下网址: <https://github.com/baogegeJiang/accuratePicker>.

致谢 感谢中国地震局地球物理研究所、NIED台网(Hi-net)和中国地震科学台阵探测计划提供的地震资料, 以及密歇根州立大学的田冬冬提供的地震数据下载软件HinetPy.

参考文献

- 1 Song L L, Yang W, Ge H K, et al. The current status and progress of portable seismic array observation in China (in Chinese). *Recent Dev World Seismol*, 2012, (3): 16–21 [宋丽莉, 杨微, 葛洪魁, 等. 中国地震科学台阵流动观测现状及进展. *国际地震动态*, 2012, (3): 16–21]
- 2 Perol T, Gharbi M, Denolle M. Convolutional neural network for earthquake detection and location. *Sci Adv*, 2018, 4: e1700578
- 3 Ross Z E, Meier M A, Hauksson E. P-wave arrival picking and first-motion polarity determination with deep learning. *J Geophys Res Solid Earth*, 2018, 123: 5120–5129
- 4 Zhu L J, Peng Z G, McClellan J, et al. Deep learning for seismic phase detection and picking in the aftershock zone of 2008 M_w 7.9 Wenchuan Earthquake. *Phys Earth Planet Inter*, 2019, 293: 106261
- 5 Zhao M, Chen S, Dave Y. Waveform classification and seismic recognition by convolution neural network (in Chinese). *Chin J Phys*, 2019, 62: 374–382 [赵明, 陈石, Dave Y. 基于深度学习卷积神经网络的地震波形自动分类与识别. *地球物理学报*, 2019, 62: 374–382]
- 6 Yu Z Y, Chu R S, Sheng M H. Pick onset time of P and S phase by deep neural network (in Chinese). *Chin J Geophys*, 2018, 61: 4873–4886 [于子叶, 储日升, 盛敏汉. 深度神经网络拾取地震P和S波到时. *地球物理学报*, 2018, 61: 4873–4886]
- 7 Dokht R M, Kao H, Visser R, et al. Seismic event and phase detection using time-frequency representation and convolutional neural networks. *Seismol Res Lett*, 2019: 481–490
- 8 Zhou Y, Yue H, Kong Q, et al. Hybrid event detection and phase-picking algorithm using convolutional and recurrent neural networks. *Seismol Res Lett*, 2019, 90: 1079–1087
- 9 Woollam J, Rietbrock A, Bueno A, et al. Convolutional neural network for seismic phase classification, performance demonstration over a local seismic network. *Seismol Res Lett*, 2019, 90: 491–502
- 10 Zhu W, Beroza G C. PhaseNet: A deep-neural-network-based seismic arrival-time picking method. *Geophys J Int*, 2018, 216: 261–273
- 11 Zhao M, Chen S, Fang L H, et al. Earthquake phase arrival auto-picking based on U-shaped convolutional neural network (in Chinese). *Chin J Geophys*, 2019, 62: 3034–3042 [赵明, 陈石, 房立华, 等. 基于U形卷积神经网络的震相识别与到时拾取方法研究. *地球物理学报*, 2019, 62: 3034–3042]
- 12 Kong Q, Trugman D T, Ross Z E, et al. Machine learning in seismology: Turning data into insights. *Seismol Res Lett*, 2018, 90: 3–14
- 13 Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image segmentation. In: Navab N, Hornegger J, Wells W, et al., eds. *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Cham: Springer, 2015. 234–241
- 14 Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Boston: IEEE, 2015. 3431–3440
- 15 Jiang Y R, Ning J Y. Automatic detection of seismic body-wave phases and determination of their arrival times based on support vector machine (in Chinese). *Chin J Geophys*, 2019, 62: 361–373 [蒋一然, 宁杰远. 基于支持向量机的地震体波震相自动识别及到时自动拾取. *地球物理学报*, 2019, 62: 361–373]
- 16 Fang L, Wu Z, Song K. SeismOlympics. *Seismol Res Lett*, 2017, 88: 1429–1430
- 17 Okada Y, Kasahara K, Hori S, et al. Recent progress of seismic observation networks in Japan—Hi-net, F-net, K-NET and KiK-net—. *Earth Planet Sp*, 2004, 56: xv–xxviii
- 18 Obara K, Kasahara K, Hori S, et al. A densely distributed high-sensitivity seismograph network in Japan: Hi-net by National Research Institute for Earth Science and Disaster Prevention. *Rev Sci Instrum*, 2005, 76: 021301
- 19 Chen C, Holland A A. PhasePApy: A robust pure Python package for automatic identification of seismic phases. *Seismol Res Lett*, 2016, 87: 1384–1396

Summary for “结合台阵策略的震相拾取深度学习方法”

An array-assisted deep learning approach to seismic phase-picking

Fang Liu¹, Yiran Jiang², Jieyuan Ning^{2*}, Jianzhong Zhang¹ & Yanhong Zhao¹¹ Inner Mongolia Autonomous Region Earthquake Administration, Hohhot 010010, China;² School of Earth and Space Sciences, Peking University, Beijing 100871, China* Corresponding author, E-mail: njy@pku.edu.cn

Rapid increase of seismic data poses challenges to seismic data processing, in which phase-picking is a crucial procedure and accordingly attracts attention from seismologists. Among the breakthroughs in deep learning methods for seismic data processing, seismologists have especially proposed many seismic phase-picking methods based on deep learning, which have shown much better accuracy than traditional automatic phase-picking methods. However, further improvements are needed to meet the requirements for processing continuous seismic records. Here we propose a new approach combining seismic array scheme and U-Net, an end-to-end neural network developed in image segmentation. Previous studies have shown that U-Net can classify waveforms and determine the onset of seismic phases accurately with the help of contracting and expanding paths. Based on U-Net, we first design a deep learning phase picker (PP) model with a long (40 s) time window. To avoid difficulties in time window selection and interference between P- and S-phase identifications, the model involves two networks with the same structure, which pick the P and S phases separately. To improve model training, we employ data transformation to expand the sample set, add typical noises to the sample set for better resemblance of real data, and slightly increase the weight of the seismic signal containing seismic records for a better balance between precision and recall. We also adopt a more effective array-assisted phase picker (APP) model which correlates the picking results from different stations to minimize misidentification. Using data from Hi-net and Alibaba Aftershock AI Capture Contest, we design training and test sets of different sizes to examine the performance and generalization ability of the model. Results show that the APP can effectively pick the body-wave phases on continuous records with a good generalization ability. The relationship between the model performance and training set is such that the model performance is directly related to the data quality rather than the size of the training set, and there is no serious overfitting even for a small training set. Statistics further show that the model has no special requirements for the locations of time windows and can handle seismic records with large epicentral distances. By comparison with other models (model with short time windows or the one of simultaneous picking of P- and S-phases), we find that APP can successfully process event-crowded waveform data and effectively reduce the influence of noise. Finally, we apply the trained model to a seismic network in the Inner Mongolia. Results show that the number of earthquakes detected by the APP is 30 times the manually-detected one, and that 98.1% of the events in the catalog are detected by the APP. The successful application to a seismic network different from the training set suggests that the APP is robust and has strong generalization ability.

deep learning, phase-picking, array technique, U-Netdoi: [10.1360/TB-2019-0608](https://engine.scichina.com/doi/10.1360/TB-2019-0608)