

基于两阶段混合模型的地震数据到时拾取研究

付敏, 马中华*

天津职业技术师范大学理学院, 天津

收稿日期: 2023年8月21日; 录用日期: 2023年9月19日; 发布日期: 2023年9月26日

摘要

相位拾取是地震资料处理中一项基础而重要的工作。为解决在连续波形中, 地震事件及初至波拾取准确率不高的难题, 基于深度学习提出了一种两阶段混合拾取模型, 分两步进行震相拾取。首先构建CNGNet模型检测地震事件; 其次构建GPkNet模型并引入自注意力机制, 对检测到的地震事件进行P波到时的精确拾取。用于网络训练和验证测试的数据来源于2019年加利福尼亚州地震序列的20个台站的连续波形。在测试集上, 地震事件检测的精确率和召回率都达到99%以上, P波拾取到时的估计误差为0.019秒。

关键词

深度学习, 神经网络, P波到时拾取, 两阶段拾取

Seismic Data Arrival Pickup Based on a Two-Stage Hybrid Model

Min Fu, Zhonghua Ma*

School of Science, Tianjin University of Technology and Education, Tianjin

Received: Aug. 21st, 2023; accepted: Sep. 19th, 2023; published: Sep. 26th, 2023

Abstract

Phase-Picking is a basic and important task in seismic data processing. In order to solve the problem of low accuracy of seismic event and first-to-wave pickup in continuous waveforms, a two-phase hybrid pickup model is proposed based on deep learning to perform phase pickup in two steps. Firstly, a CNGNet model is constructed to detect seismic events; secondly, a GPkNet model is constructed to accurately pick up the P-arrivals time for the detected events by introducing the self-attention mechanism and the gated recurrent unit. The data used for network training and

*通讯作者。

validation testing are derived from continuous waveforms from 20 stations of the 2019 California earthquake sequence. The precision and the recall ratio of seismic event detection by our method are more than 98% for the test sets, and detection reaches more than 99%, and the estimated error of P-arrival is 0.019 s.

Keywords

Deep Learning, Neural Network, P-Arrival Picking, Two-Stage Picking

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

地震事件检测和相位拾取是确定地震位置和建立地震目录的关键一步, 对于地震数据的研究至关重要。目前地震波初至拾取算法通常基于振幅、能量或频率等确定初至时间, 人们已提出多种初至拾取方法, 如能量比法、相关法、分形维法, 这些方法通常仅考虑单一特征, 在精度上表现较差。基于神经网络拾取初至时间的算法诸如: 人工神经网络[1]、支持向量机[2]、模糊聚类[3]等, 此类算法是基于一维特征的时间序列进行训练, 拾取过程中易受噪声干扰。

深度学习技术在人工智能时代已经成为了各个领域研究热点, 其强大的自学习特征能力, 各个领域都有广泛的应用。近年来, 在地震学领域也得到一系列的成功应用, 都有比较明显的进展。初至数据的自动拾取已经有了一些深度学习方法的应用, 包括当今比较流行的卷积神经网络[4]。基于 U-net 网络设计的 PhaseNet 震相拾取算法[5]在近震事件震相挑取中有着出色的泛化能力。DetNet + PpkNet 的事件检测和相位选取算法[6]在地震波形的初至拾取上表现良好。LPPN 算法[7]是一种轻型震相拾取网络, 用于从连续地震记录中拾取 P/S 相位。EQtransformer 算法[8]是用于地震事件检测和相位拾取的全局深度学习模型, 使用分层注意机制组合相位信息和地震信号的全波形, 在单站的测试精度上优于以前的深度学习和传统的相位拾取和检测算法。

在地震数据研究中, 深度学习算法都涉及信号检测或者相位提取问题, 初至波到时拾取则需要将两个任务结合。本研究将两个任务分别进行训练, 用于在连续波形中进行地震事件检测及 P 波到时拾取。将地震事件检测视为图像分类问题, 与现有研究使用单个台站的地震数据[9]训练网络不同, 在本研究中, 我们使用多个台站的记录联合训练 CNN 模型。由于 CNN 自动综合了多个台站之间的信息, 因此可以更可靠地检测地震事件。时间序列网络模型则可以有效地捕获事件的时序特征。利用这两种网络模型的优势, 引入了自注意力模块, 提高模型的精度, 以期通过上述研究能够有效拾取地震事件, 并进行 P 波到时拾取。

2. 方法

2.1. 基本原理

本研究主要解决在连续地震波形中的地震事件检测和到时拾取问题, 如图 1 所示, 输入三分量地震数据来进行初至波拾取的训练。首先将输入的地震数据进行预处理, 如: 数据增强等; 将地震数据输入卷积神经网络模型(CNGNet)进行事件的检测; 最后将地震波形事件作为时间序列输入到到时拾取模型(GPpkNet), 采用 GRU 模块结合自注意力机制拾取 P 波与 S 波。

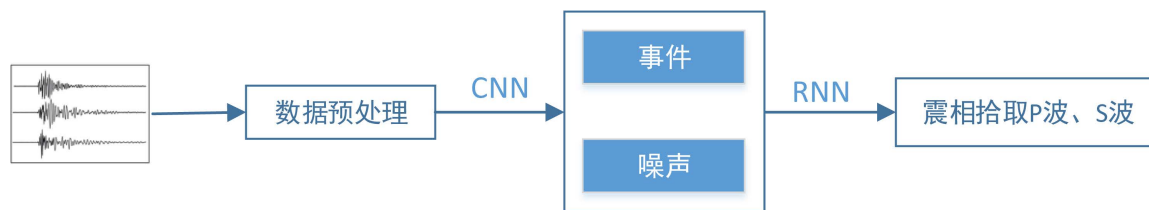


Figure 1. Fundamentals of network design
图 1. 网络设计的基本原理

2.2. 事件检测模型(图 2)

在本文中地震事件检测模型是由噪声和事件组成的一个二分类模型。常见的 CNN 卷积层通常由卷积操作、激活函数和最大池化层三种部分组成。其中卷积操作实现了输入数据的滤波和线性组合，主要是为了对地震事件进行特征的提取。它可以通过滑动窗口的方式对输入的地震数据进行扫描，并通过局部感知野提取不同位置的特征。其中卷积核的权重共享机制也使得参数量相对较小，减少了过拟合的风险。使用 ReLU 激活函数，能够引入更丰富的表达能力，并帮助网络学习更复杂的特征表示。池化操作例如最大池化，可以保留重要的特征，在一定程度上提升了网络对平移不变性的建模能力，并减少了计算量和参数数量。

在设计地震事件检测模型时，文献[10]中借鉴了 LeNet 以及 VGG 模型的思路，设计了简单的 CNN 模型。而在本文中把整个网络模型分为了三个部分，先是前 4 个卷积层，使用最经典的卷积层即卷积操作、ReLU 激活函数以及最大池化操作，其中滤波器的数量倍数增长；5 到 6 层滤波器数量不变；最后是两个类似于 VGG 块的两个卷积层。

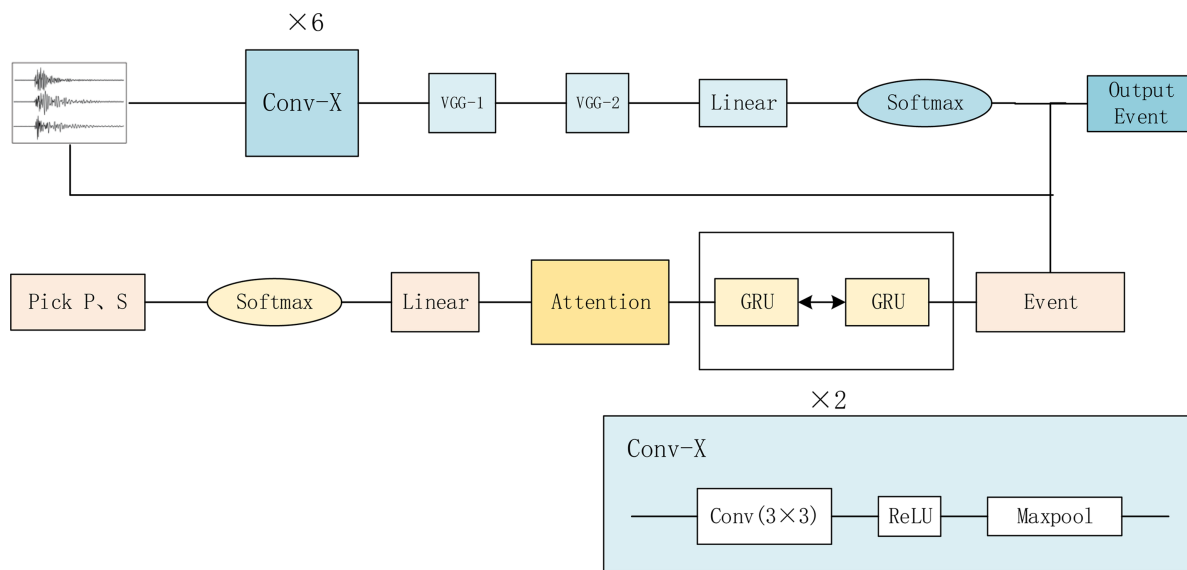


Figure 2. CNGNet for event detection and GPKNet model for waveform picking
图 2. 进行事件检测的 CNGNet 与进行波形拾取的 GPKNet 模型

这种设计可以防止梯度消失和梯度爆炸抑制，激活函数引入非线性变换，使得梯度能够在反向传播过程中更好地流动，从而保证梯度的有效传递和更新。在前 4 个卷积层中，滤波器的数量倍数增长，这意味着网络可以在较浅的层级上捕获到更多的特征表示，较少的滤波器数量能够处理局部细节信息，而更多的滤波器数量则可以关注更大范围的全局结构。第 5 到第 6 层保持滤波器数量不变，可以控制网络

的参数量和复杂度。最后两个类似于 VGG 块的卷积层通常具有较大的滤波器数量和深度, 可以进一步提取更高级别的抽象特征, 增强网络的表达能力。

由于我们将地震检测任务看作为一个二分类任务[5], 因此在模型的末端使用 softmax 函数进行概率预测。Softmax 函数可以使得所有类别的输出值都落在 0 到 1 之间, 并且各个类别的输出和为 1, 这样可以将模型的输出解释为对不同类别的预测概率。

2.3. 到时拾取模型

GRU 是一种循环神经网络(RNN)的变体, 它具有门控机制, 可以更好地捕捉长期依赖关系。相比于传统的 RNN, GRU 引入了更新门和重置门, 通过动态调整当前时刻的信息对过去的记忆进行遗忘或者记忆的操作。假设隐藏状态的维度为 h_t , 输入特征的维度为 x_t , 图 3 为 GRU 模块的基本模型结构。

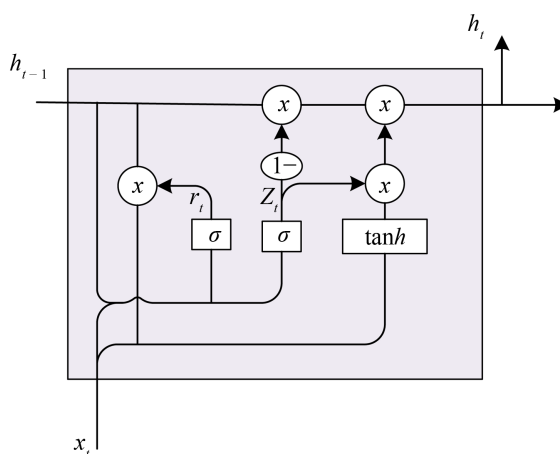


Figure 3. Basic schematic diagram of the GRU module
图 3. GRU 模块的基本原理图

自注意力机制(Self-Attention)是一种允许模型根据输入的不同位置进行自适应权重分配的机制。通过计算输入序列中各个位置之间的相关性, 自注意力机制能够捕捉到不同单词或者元素之间的重要关系, 并且分配给它们不同的注意力权重。这使得模型更加关注输入序列中最相关的信息, 并减少了传统固定窗口大小的局限性。在注意力机制中, 主要涉及三个要素: 查询(query)、键(key)和值(value)。查询是用来获取注意力权重的向量, 键和值则表示输入序列中每个位置的特征向量。注意力机制根据查询、键和值之间的相似度来计算注意力权重。注意力机制的公式表示:

$$\text{Attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{softmax}\left(\frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{d_k}}\right)\mathbf{V} \quad (1)$$

其中: \mathbf{Q} 表示查询向量, \mathbf{K} 表示键向量, \mathbf{V} 表示值向量, 经过 softmax 函数处理得到注意力权重, 然后其应用于值向量上, 最后得到加权后的表示。

GRU 可以有效地捕获波形信号的时序特征, 自注意力机制能够学习到波形数据中不同位置之间的关联性, 并根据全局上下文信息进行权重调整。这使得模型能够在进行初至拾取时同时考虑到整个波形数据的信息, 而不仅仅是局部窗口内的信息, 同时可以提高模型的鲁棒性和可靠性。GRU 可以处理不同长度的输入波形序列, 并能够适应不同的时序特征, 自注意力机制能够实现对不同位置的自适应关注, 使得模型对于不同类型的波形数据都能具有较好的适应性和准确性。模型可以支持并行计算, 这意味着可以在大规模波形数据上高效训练和推理。

通过结合 GRU 和自注意力机制, 可以充分利用它们各自的优势, 提高波形初至拾取任务的性能和效果, 并获得了很好的泛化能力。

在模型中我们采用的是 30 秒长的三分量地震图, 把它切割成 59 个连续帧, 每帧持续时间为 1 秒, 步幅为 0.5 秒, 将这些三分量段连接成具有 300 个数据点的一维矢量。使用了两个结构相同的 GRU 模块, 包含 64 个特征向量, 并发送到自注意力机制中获取更多局部信息, 最后输入到完全连接层。根据 softmax 归一化给出的分类分数, 将每个时间步长分为噪声、P 波或 S 波三类之一。

由于波形有 50% 是重叠的, 包含输入到达的两个段都被标记为到达后, 真正的到达位于第一个标记段的后半部分。因此, 我们选择的到达时间在第一个标记段的初始时间后 0.75 s, 这样可以保证理论最大拾取精度小于 0.25 s。

3. 数据集

3.1. 数据集预处理

使用了 2019 年美国加利福尼亚州的东部的多个相近站台三天的连续地震波形(<https://www.iris.edu>), 台站设备由九个宽频带三分量微震台站组成。其中包括了 20 个相邻震相台站的三分量地震波形数据, 原始采样频率为 1 kHz, 其中主要检测的震相分别为直达的 P 波、S 波。所选的地震数据集按照 9:1 的比例分为训练集和测试集。

3.2. 数据预处理及数据增强

使用的是 100 Hz 的低通滤波器, 滤除频率高于 100 Hz 的信号成分, 这样可以降低或消除高频噪声的影响, 对于提高信号质量和减少噪声干扰非常有益; 接着进行数据的归一化, 如图 4 所示。

地震数据集中可能存在类别不平衡问题, 通过数据增强方法, 可以生成更多少数类别的样本, 以平衡各个类别之间的样本数量差异, 也可以添加额外的信息, 如地震波形的振幅、频率谱、信号延迟等, 帮助改善模型对这些信息的学习能力。本文中采用数据增强主要为时间偏移使其增加为原来的两倍, 由此增强模型的泛化能力平衡数据的分布。

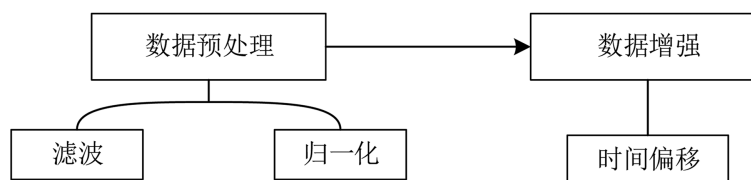


Figure 4. Data preprocessing and data augmentation
图 4. 数据预处理以及数据增强

4. 性能测试

两个阶段都是通过小批随机梯度下降(Mini-batch Stochastic Gradient Descent, 简称 Mini-batch SGD) 算法进行训练的, 它是梯度下降算法的一种变体, 在每次更新模型参数时利用部分训练数据样本的梯度来进行优化。通过使用小批量样本进行参数更新, 可以减少训练过程的计算开销, 并且还可以更好地逼近全局最优解。

优化器根据损失函数的梯度更新神经网络的权值。每次迭代后, 对训练集中的样本进行打乱, 重复整个过程。在事件检测模型训练中的损失函数为 L2 正则化的交叉熵损失。在初至拾取波形训练中损失函数为所有时间步长的平均交叉熵。所有权重都由 Xavier 初始化器初始化, 并由 Adam 优化器优化。

我们使用 Tensorflow 构建模型, 在事件检测训练中采用正则化来防止过拟合, 其中包括: 基于验证集的提前停止, 停止训练过程; L2 正则化来惩罚损失函数和加权数的绝对值; 批归一化, 通过对每个卷积输出进行归一化, 减少从浅层到深层的分布转移。

使用图 2 中的网络模型架构进行模型训练, 用准确率和召回率来评估训练过程和训练模型在测试数据集集中的性能, 这两个指标的定义如下:

$$\text{Precision} = \frac{N_{\text{TP}}}{N_{\text{TP}} + N_{\text{FP}}} \quad (2)$$

$$\text{Recall} = \frac{N_{\text{TP}}}{N_{\text{TP}} + N_{\text{FN}}} \quad (3)$$

其中, 式中 TP 为真阳性数, FP 为假阳性数, TN 为真阴性数, FN 为假阴性数。在本文中, 地震被视为积极的, 噪声被视为消极的。

训练和验证过程中的分类精度以及召回率, 如图 5 所示。经过约 11,000 次迭代后, DeGNet 的训练精度及召回率收敛到 99.5% 以上, 并且峰值多次达到 100%。对于地震事件和噪声, 其在测试集的精确度和召回率同样达到约 99%, 峰值在 100%。经过 100,000 次迭代后, RNN 阶段的训练精度和验证精度均收敛到 98% 以上, 并且没有出现过拟合的状况。

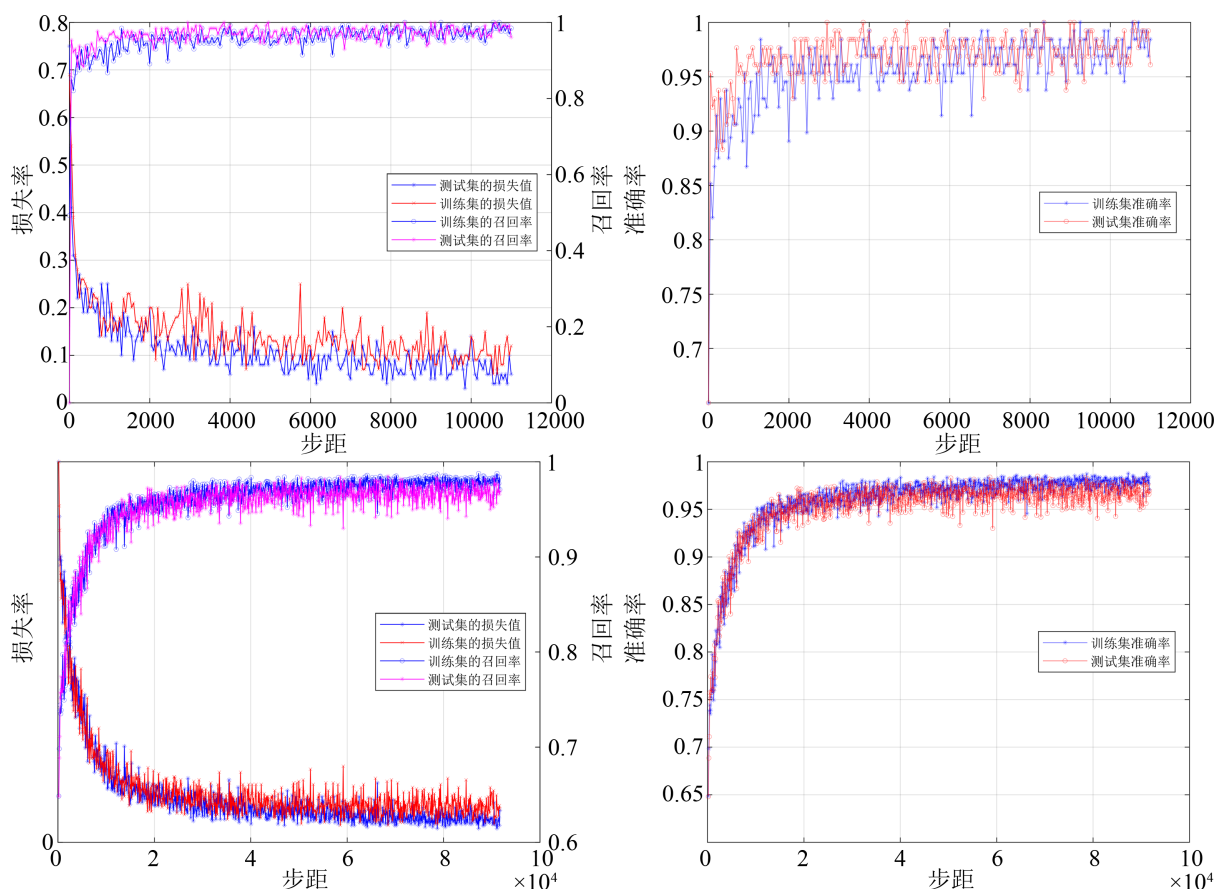


Figure 5. CNGNet (top) and GPKNet (bottom) training iterative processes

图 5. CNGNet (上)和 GPKNet (下)训练迭代过程

为了计算其到时误差, 我们测量了平均绝对误差(Mean Absolute Error, 简称 MAE), MAE 的计算公

式如下:

$$\text{MAE} = \left(\frac{1}{n}\right) \times \sum |y_i - \hat{y}_i| \quad (4)$$

其中, n 表示样本数量, y_i 表示实际值, \hat{y}_i 表示预测值。MAE 的值越小表示预测结果与真实值之间的绝对差异越小。通过计算 MAE 可以很明显的看出模型在性能上是非常优秀的。

Table 1. Performance test results for each method

表 1. 各个方法的性能测试结果

方法	精确度	召回率	平均到时误差(MAE)	误差 > 0.2 s 占比
PhaseNet	0.9505	0.9493	0.023491	2.97%
CNNNet + CGANet	0.9039	0.9017	0.026061	3.32%
DetNet + PpkNet	0.9782	0.9783	0.025341	2.82%
CNGNet + LpkNet	0.9804	0.9806	0.020136	2.53%
CNGNet + GpkNet	0.9800	0.9841	0.019916	2.17%

为更加客观的评估我们模型的优越性,进行了几种对比试验。DetNet + PpkNet 算法[5]采用了双向的 GRU 模型,可以很好地捕获时序信息,但是对于局部信息很难关注到。CNNNet + CGANet 算法(Mingruo, J. 2023) [11]中引入了注意力机制和卷积模块,拾取效果良好,所以在设计网络时,我们在原有的 GRU 模块的基础上引入了注意力机制,通过对表 1 中的结果分析,可以很明显的发现,本方法在事件检测和初至拾取的性能均表现的很好。在此基础上,我们又想到了将双向的 GRU 替换为 LSTM (LpkNet),结果发现精确度虽然很高但是召回率和 MAE 偏低,性能并不算太好。GRU 算法相对于 LSTM 算法具有更简单的结构和更少的门控单元更容易训练和调整,所以我们的模型选择 GRU 模块作为基础。

5. 结果

在地震频发区,需要短时间对大量的地震数据进行处理,人工的方式耗费了大量的人力物力,本文针对美国加利福尼亚州东部在 2019 年 7 月 4 日某 4 个台站获取的地震事件进行检测并进行 P 波与 S 波的拾取,表 2 是在事件检测阶段所获得的事件数量,由于有些地震事件置信度较低,在进行波形拾取之前对其选择丢弃。

Table 2. CNGNet test results and number of GpkNet discards

表 2. CNGNet 的检测结果及 GpkNet 丢弃数量

台站	检测到事件	丢弃事件
CI.WBM	6622	699
CI.CLC	3220	963
CI.WBF	2705	500
CI.SLA	4928	600

接下来是对所选择的事件进行波形拾取,图 6 展示每个台站的 P 波、S 波的初至拾取结果。结果显示,我们的方法在地震方面的应用上依旧存在较小的误差,需要进一步的进行研究。

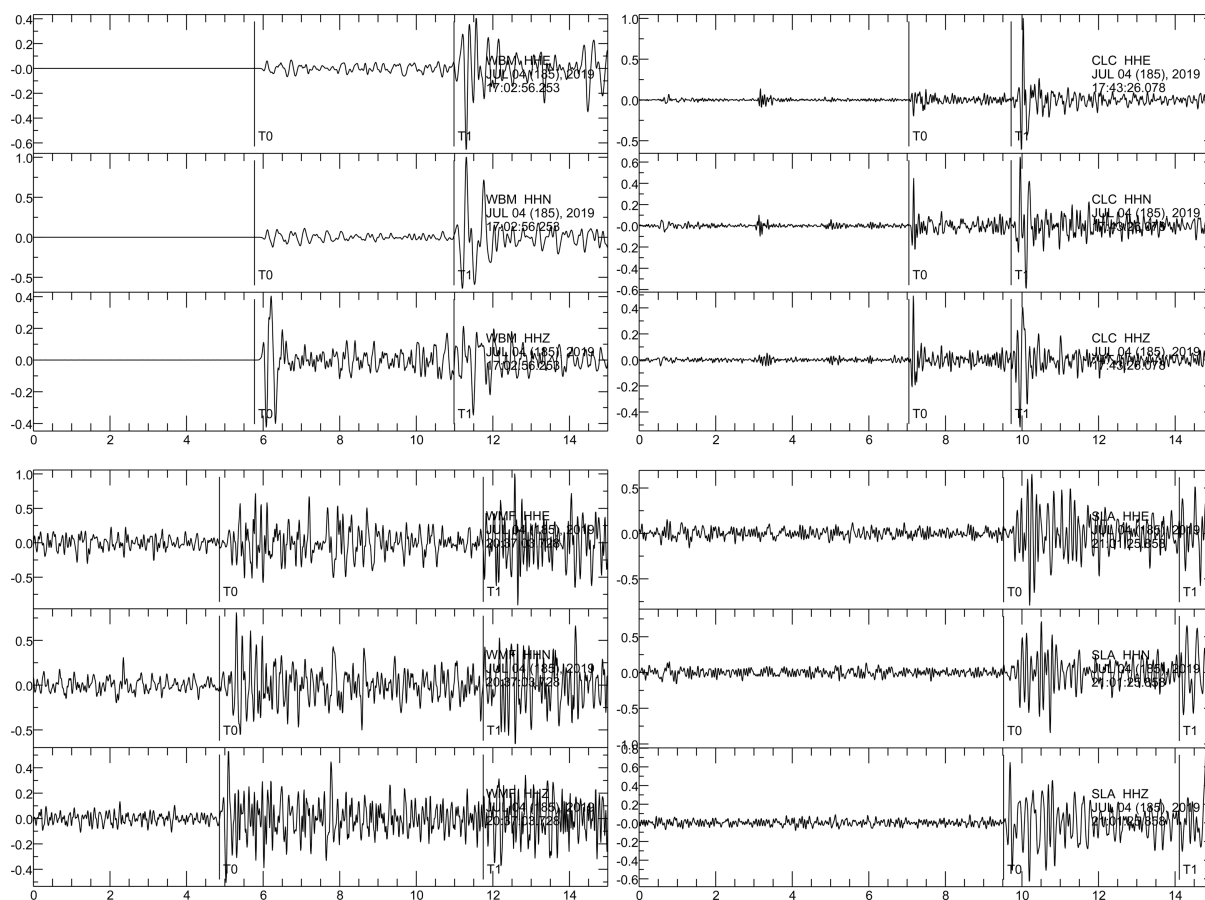


Figure 6. Example of station pickup results

图 6. 台站拾取结果示例

6. 结论

传统的地震波初至拾取方法存在了精度较低、拾取效率不高的问题, 在本文中使用深度学习方法, 引入了一种使用 CNGNet 和 GpkNet 的混合算法模型来检测事件并从连续波形中选择相位。CNGDet 在测试集上的事件检测精确率和召回率都达到 99% 以上, GpkNet 对于 P 波到时估计的误差均值为 0.019 s。对比结果表明, 在性能、精度上均优于其他算法。混合算法在采摘稳定性(低误检)和采摘精度方面优于其他算法, 在未来可以考虑使用不同类型的数据集进行测试, 不断优化本文中提到的混合模型, 在地震数据波形初至拾取方面发挥越来越重要的地位。

基金项目

天津市教委科研计划(2020KJ115)。

参考文献

- [1] Maity, D., Aminzadeh, F. and Karrenbach, M. (2014) Novel Hybrid Artificial Neural Network Based Autopicking Workflow for Passive Seismic Data. *Geophysical Prospecting*, **62** 834-847. <https://doi.org/10.1111/1365-2478.12125>
- [2] Jiang, Y. and Ning, J. (2019) Automatic Detection of Seismic Body-Wave Phases and Determination of Their Arrival Times Based on Support Vector Machine. *Chinese Journal of Geophysics*, **62**, 361-373.
- [3] 李帛珊, 唐丽华, 孙燕萍, 吕春来. 通过无监督机器学习自动拾取微震事件[J]. 世界地震译丛, 2019, 50(5): 411-432.

- [4] Yu, Z., Chu, R., Wang, W. and Sheng, M. (2020) CRPN: A Cascaded Classification and Regression DNN Framework for Seismic Phase Picking. *Earthquake Science*, **33**, 53-61. <https://doi.org/10.29382/eqs-2020-0053-01>
- [5] Zhu, W. and Beroza, G.C. (2019) PhaseNet: A Deep-Neural-Network-Based Seismic Arrival-Time Picking Method. *Geophysical Journal International*, **216**, 261-273. <https://doi.org/10.1093/gji/ggy423>
- [6] Zhou, Y., Yue, H., Kong, Q. and Zhou, S. (2019) Hybrid Event Detection and Phase-Picking Algorithm Using Convolutional and Recurrent Neural Networks. *Seismological Research Letters*, **90**, 1079-1087. <https://doi.org/10.1785/0220180319>
- [7] Yu, Z. and Wang, W. (2022) Lppn: A Lightweight Network for Fast Phase Picking. *Seismological Research Letters*, **93**, 2834-2846. <https://doi.org/10.1785/0220210309>
- [8] Mousavi, S.M., Ellsworth, W.L., Zhu, W., Chuang, L.Y. and Beroza, G.C. (2020) Earthquake Transformer: An Attentive Deep-Learning Model for Simultaneous Earthquake Detection and Phase Picking. *Nature Communications*, **11**, Article Number: 3952. <https://doi.org/10.1038/s41467-020-17591-w>
- [9] Perol, T., Gharbi, M. and Denolle, M. (2018) Convolutional Neural Network for Earthquake Detection and Location. *Science Advances*, **4**, e1700578. <https://doi.org/10.1126/sciadv.1700578>
- [10] Jiao, M.G., Dong, F.J., Luo, H., Yu, J.K. and Ma, L. (2023) P-Arrival Picking Method of Mine Microseisms by Fusing of GRU and Self-Attention Mechanism. *Acta Seismologica Sinica*, **45**, 234-245.
- [11] Yang, S., Hu, J., Zhang, H. and Liu, G. (2021) Simultaneous Earthquake Detection on Multiple Stations via a Convolutional Neural Network. *Seismological Research Letters*, **92**, 246-260. <https://doi.org/10.1785/0220200137>