**先导数据大模型：用于单台地震波形数据分析的双向神经网络预训练模型**

蔡育琦, 于子叶, 王伟涛, 安艳茹, 李 璐

Bi-directional Pre-trained Network for Single-station Seismic Waveform Analysis

CAI Yuqi, YU Ziye, WANG Weitao, AN Yanru, and LI Lu

在线阅读 View online: <https://doi.org/10.15953/j.ctta.2024.162>**您可能感兴趣的其他文章****Articles you may be interested in****基于单步神经网络的牙根与下颌管关系自动检测**

Automatic Identification of Relationship between Tooth Root and Mandibular Canal Based on One Step Deep Neural Network

CT理论与应用研究. 2023, 32(2): 198–208

基于卷积神经网络的断层预测方法

Fault Prediction Method Based on Convolutional Neural Network

CT理论与应用研究. 2020, 29(5): 522–533

基于EMD与EEMD的最优降噪平滑模型在地震波中的应用

Application of Optimal Noise Reduction Smooth Model Based on EMD and EEMD in Seismic Waves

CT理论与应用研究. 2020, 29(1): 11–21

基于波形特征向量的谱聚类地震相分析

Seismic Facies Analysis Based on Spectral Clustering with Waveform Characteristic Vector

CT理论与应用研究. 2024, 33(1): 13–23

探讨基于卷积神经网络对颅底骨折CT图像精准诊断的应用价值

The Application Value of Accurate Diagnosis of CT Image of Skull Base Fractures based on Convolutional Neural Network

CT理论与应用研究. 2021, 30(6): 769–776

基于对抗式残差密集深度神经网络的CT稀疏重建

Sparse CT Reconstruction Based on Adversarial Residual Dense Deep Neural Network

CT理论与应用研究. 2022, 31(2): 163–172



关注微信公众号，获得更多资讯信息

蔡育琦, 于子叶, 王伟涛, 等. 先导数据大模型: 用于单台地震波形数据分析的双向神经网络预训练模型[J]. CT 理论与应用研究 (中英文), 2025, 34(1): 111-116. DOI:10.15953/j.ctta.2024.162.
CAI Y Q, YU Z Y, WANG W T, et al. Bi-directional Pre-trained Network for Single-station Seismic Waveform Analysis[J]. CT Theory and Applications, 2025, 34(1): 111-116. DOI:10.15953/j.ctta.2024.162. (in Chinese).

先导数据大模型: 用于单台地震波形数据分析的双向神经网络预训练模型

蔡育琦¹, 于子叶^{1✉}, 王伟涛¹, 安艳茹², 李璐¹

1. 中国地震局地球物理研究所, 北京 100081

2. 中国地震台网中心, 北京 100045

摘要: 机器学习特别是深度学习方法在地震学中的应用越来越广泛, 其在震相检测、地震分类中都达到了接近人类的精度。但目前, 多数地震学神经网络模型专注于单一任务。我们基于中国地震台网中心发布的 CSNCD 数据集构建一个用于单台数据分析的双向神经网络预训练模型, 模型以原始波形数据为输入, 通过卷积神经网络和双向 Transformer 模型进行特征提取和处理, 不仅可以完成常规的 Pg、Sg、Pn 和 Sn 震相检测、P 波初动方向判定和事件类型判断工作, 还可以通过迁移学习将模型用于其他地震波形数据分析工作中。

关键词: 深度学习; 震相检测; 初动检测; 事件分类; 预训练模型

DOI:10.15953/j.ctta.2024.162 中图分类号: P 315; O 242 文献标识码: A

Bi-directional Pre-trained Network for Single-station Seismic Waveform Analysis

CAI Yuqi¹, YU Ziye^{1✉}, WANG Weitao¹, AN Yanru², LI Lu¹

1. Institute of Geophysics, China Earthquake Administration, Beijing 100081, China

2. China Earthquake Networks Center, Beijing 100045, China

Abstract: The application of machine learning, particularly deep learning methods, is becoming increasingly widespread in seismology, achieving near-human accuracy in tasks such as phase detection and event classification. However, most neural network models in seismology currently focus on single tasks. Based on the CSNCD dataset released by the China Earthquake Networks Center, we have developed a bi-directional neural network pre-trained model for single-station data analysis. This model uses three-component seismic waveform data as input and employs convolutional neural networks and bi-directional Transformer models for feature extraction and processing. It not only performs routine tasks such as Pg, Sg, Pn and Sn phase detection, P-wave first-motion direction determination, and event type classification but can also be adapted to other seismic waveform data analysis tasks through transfer learning.

Keywords: deep learning; phase detection; first-motion polarity; event classification; pre-trained model

近些年, 机器学习方法特别是深度学习方法在地震学中的应用愈发广泛。深度学习方法在地震数据处理中表现出了接近人工分析的精度, 并且处理效率高于人工处理^[1-2]。典型的应用包括地震震相拾取^[3-8]、地震检测^[9-10]、P 波初动分析^[11-14]、震源参数分析^[15-16]、地震波形去噪^[17-20]、震相关联^[21-24]、地震定位^[25-30]和地震分类^[31-34]等。

目前地震学中多数神经网络模型着眼于单一任务, 比如用于震相拾取的神经网络模型以地震波形为输入并以 P、S 震相作为输出。在一些研究中, 尝试将更多信息作为输入, 比如 Münchmeyer 等^[35]尝试将地震定位和震级估计同时进行处理, Seis-CLIP^[36]尝试将震相到时、震中距、方位角等信息输入到神经网络中, 并构建了一个能够处理更多信息

收稿日期: 2024-07-30。

第一作者: 蔡育琦, 男, 固体地球物理专业硕士研究生, 主要从事地震信号检测研究, E-mail: caiyuqiming@foxmail.com。

通信作者: 于子叶[✉], 男, 副研究员, 主要从事地震数据分析工作, E-mail: yuziye@cea-igp.ac.cn。

的神经网络模型。

2023 年, 中国地震台网中心发布了 CSNCD 数据集^[37], 其中包含了超过 4500 万条人工标注震相到时、震相初动、地震类型、地震震级等信息。在此基础上, 我们构建一个用于单台数据分析的双向神经网络预训练模型, 称为先导数据大模型。

作为预训练模型, 我们加入了 Pg、Sg、Pn、Sn 4 种震相、P 波初动方向、地震类型和原始波形 4 种输出, 其中原始波形输出为自监督训练, 其他为监督训练。测试表明, 我们的模型不仅可以完成常规的 Pg、Sg、Pn 震相检测、P 波初动方向判定和事件类型判断工作, 还可以通过迁移学习将模型用于其他波形数据分析工作中。

1 预训练模型和迁移学习方法简介

通过大量数据训练供不同机器学习任务使用的模型称为预训练模型, 目前广泛使用的预训练模型包括单向模型和双向模型。其中单向模型以 GPT 系列为代表, 在处理数据中通常只考虑前文, 其通常用于文本生成、文本图像音频生成等多模态工作中。双向模型则以 BERT 系列为代表, 其在处理文本过程中会考虑前后文数据。生成式模型在目前可训练参数数量已经可以达到千亿量级, 在文本处理中表现出了良好的性能。SeisCLIP 使用生成式模型来处理多台地震频谱并生成所需信息。但是地震学研究中通常需要对原始波形进行分析并从中提取更加准确的信息, 在此需求中双向模型通常能够获得

更高的精度^[38]。

迁移学习指的是通过海量数据训练后的模型在处理其他任务中可以通过少量样本训练即可迁移至新的工作中。迁移学习目前在图像处理、文本处理中均表现出良好的性能。地震学数据分析中需要处理大量的数据, 而分析数据的需求各不相同。如果针对每个需求训练深度学习模型, 那么在每个场景中都需要海量的人工标注数据, 但是在诸如滑坡微震等场景人工标注数据是偏少的。因此, 在地震学中同样有研究者使用迁移学习方法以减少对训练数据量的需求。

研究表明, 由于预训练模型使用海量数据进行训练, 在只使用少量数据进行迁移学习后, 神经网络模型在新的场景中表现出良好的效果, 比如地震事件分类和远震震相检测。

2 训练数据

预模型训练需要海量的各种类型数据。因此, 我们使用中国地震台网中心的 CSNCD 数据集来对模型进行训练。CSNCD 数据集是一个大规模的全球地震信号标记数据集, 其特点在于覆盖了多样化的地理区域和各种地质环境下的地震事件。我们使用的事件和台站如图 1 所示, 这些事件包含了丰富的震相类型和大量的初动数量。

除此之外, 这些数据还覆盖了从 -2.0 到 9.0 级不同震级的地震事件。数据集中还包含的广泛的信噪比 (SNR) 范围, 这一特点使得模型能够在不同

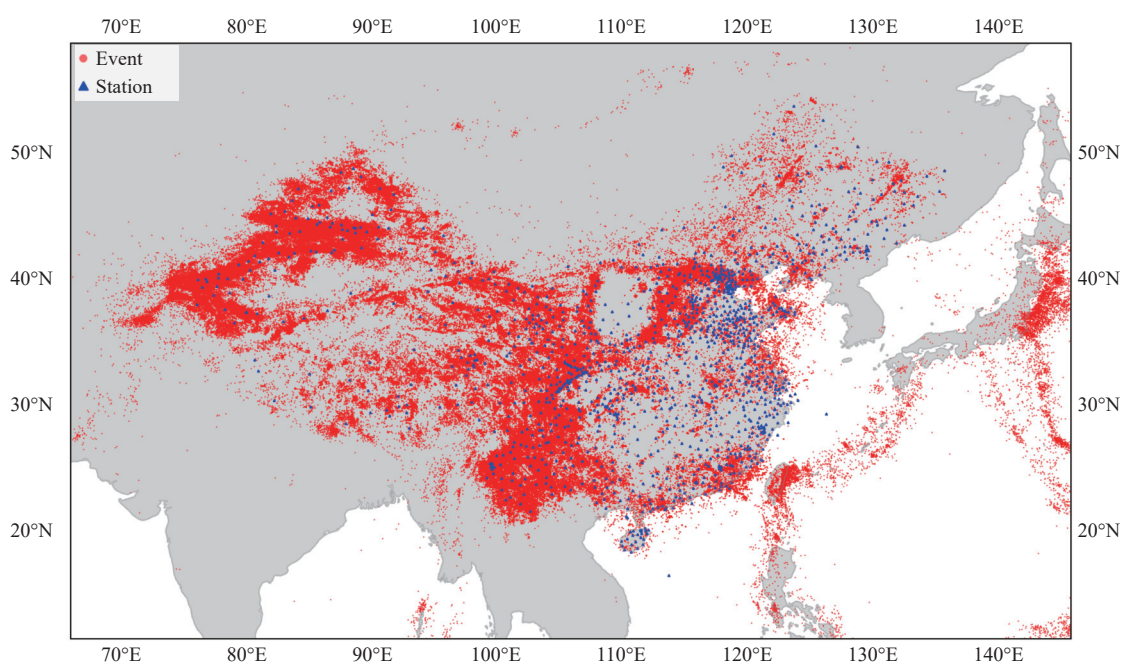


图 1 使用的台站 (蓝色三角形) 和事件 (红色点) 的分布

Fig.1 Distribution of stations (blue triangles) and events (red points)

质量的数据上进行有效的学习，从而提高模型的适应性和泛化能力。

CSNCD 数据集的多样性和复杂性为先导数据大模型提供了坚实的基础。

3 模型结构

在处理更多地震波波形信息时需要较大的模型容量，即需要更多可训练参数。因此我们在构建神经网络过程中以 Transformer 模型为核心，同时为了处理波形输入和输出，我们加入卷积神经网络，具体网络结构如图 2 所示。

因为地震波形数据中包含较多高频信息，因此我们使用卷积神经网络构建编码器来处理原始波形，得到波形特征。但波形特征中缺少位置相关的信息，为此我们加入双向循环神经网络构建的位置编码器来对特征位置信息进行编码。在此之上我们

加入多层双向 Transformer，这可以综合考虑全部波形的特征，并输出特征向量。由于所得特征需要输出事件类型等信息，我们加入了多任务编码 (Task ID)，这可以提供额外特征供后续分析。

4 先导大模型的应用

4.1 震相检测和 P 波初动检测

目前大多数地震学震相检测神经网络模型只能检测 Pg 和 Sg 震相，如 PhaseNet 和 EQTransformer 等^[3-8]。除了 Pg、Sg 震相拾取模型外，一些研究者还构建了其他的地震学神经网络模型。例如，Yuan 等^[39]提出的远震震相拾取方案，可以检测 PcP 震相和 PKiKP 震相；Ding 等^[40]构建的 PmPNet 可以检测 PmP 震相；Ross 等^[11]和 Zhao 等^[14]分别构建了用于确定 P 波初动极性的神经网络模型。

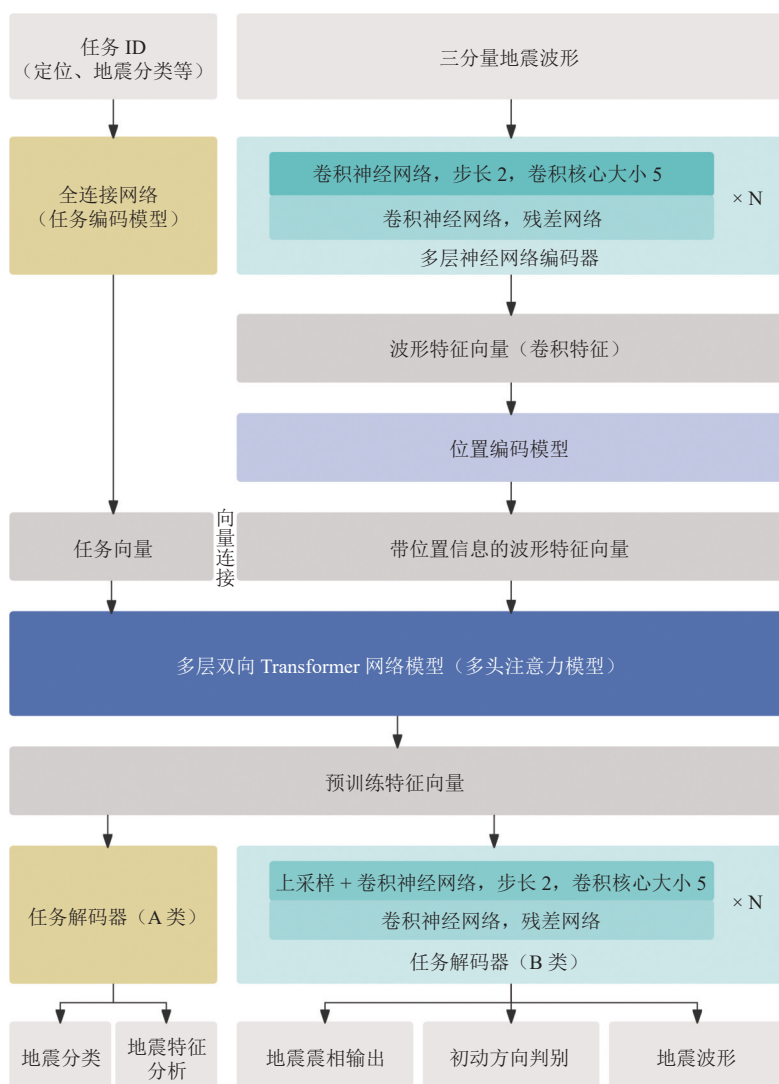


图 2 双向神经网络结构

Fig.2 Bidirectional neural network architecture

基于 CSNCD 数据的特点, 我们构建的先导大模型通过多个解码器同时实现了震相检测和 P 波初动检测工作。而且在震相检测方面, 不仅可以检测

常规的 Pg 和 Sg 震相, 还可以检测传统上认为难以检测的 Pn 和 Sn 震相。图 3 展示了一个模型震相检测的一个例子。

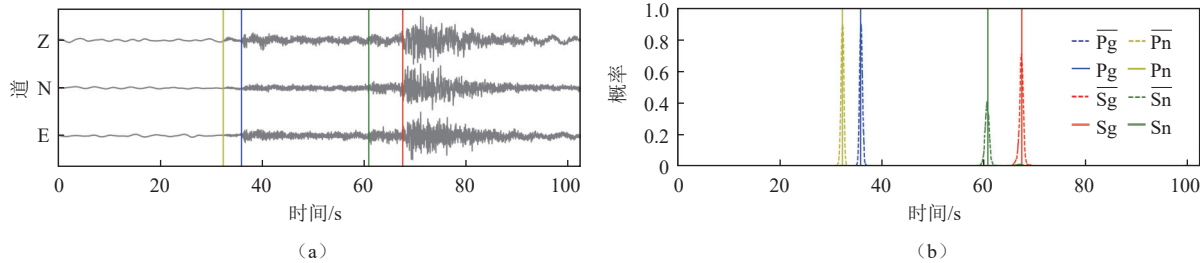


图 3 模型拾取结果示例, 实线为人工标注震相到时, 虚线为模型预测震相到时

Fig.3 A example of picks in the test dataset. The solid vertical lines represent the arrival times of phases picked by analysts. The dashed lines represent the probability distributions of phases predicted by the model

4.2 迁移学习应用

我们分别使用内蒙地区数据和远震数据测试模型的迁移学习能力。其中, 内蒙地区的数据用于事件分类测试, 在预训练过程中也有不同事件类型的标签, 但是由于预训练样本极度不均衡 (天然地震占比 99.5%), 这使得原始的预训练模型对事件类

型的预测均为天然地震, 因此需要进行迁移学习。事件分类测试数量总计 419 个, 其中 251 个地震事件用于训练, 168 个地震事件用于测试。主要包含 3 种事件类型: 天然地震 (EQ, 103 个事件), 爆破 (EP, 244 个事件) 和塌陷 (SS, 72 个事件)。在经过迁移学习后, 模型的分类准确度达到了 92.2% (图 4)。

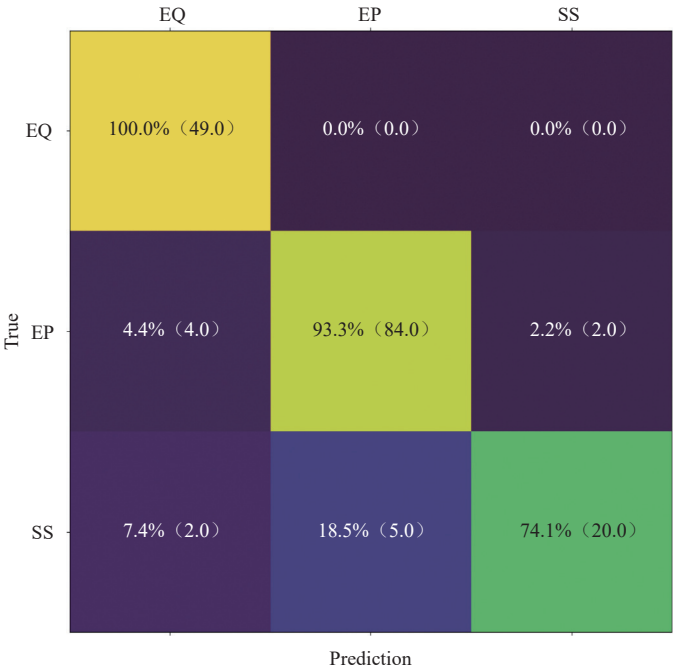


图 4 地震分类混淆矩阵

Fig.4 Confusion matrix of event classification

我们整理全球的远震数据, 挑选 1000 条样本进行远震迁移学习训练和测试。测试中, 原始的预训练模型的查准率较高, 但查全率偏低, 这代表大量样本未被检测到, 这说明远震波形特征和近震特征之间存在差异。经过迁移学习后, 模型对远震的

检测精度明显提高, 查准率由原来的 0.270 提高到了 0.708。

5 结论

先导数据大模型在地震学中具有广泛的应用前

景, 通过利用先导数据大模型, 地震台站可以实时处理和分析地震波形数据, 快速识别地震事件、震相和初动方向, 这可以显著提高地震监测的效率和准确性, 尤其在地震多发地区和紧急情况下, 能够提供更及时的预警信息。此外, 先导数据大模型可以通过迁移学习, 适应不同地理区域和不同类型的地震数据。这样可以减少数据集的依赖和节约训练时间成本, 实现在不同的地区和任务中快速部署。

先导数据大模型在地震监测、预警、科学研究和数据管理等方面具有重要的应用前景, 能够显著提升地震数据处理的自动化和智能化水平, 为地震学的发展和地震灾害的防范提供有力支持。

感兴趣的专家读者可以通过数据大模型网址 <https://github.com/cangyeone/seismological-ai-tools> 来更好的了解和使用模型。

参考文献

- [1] BEROZA G C, SEGOU M, Mostafa Mousavi S. Machine learning and earthquake forecasting: Next steps[J]. *Nature communications*, 2021, 12(1): 4761.
- [2] MOUSAVI S M, BEROZA G C. Deep-learning seismology[J]. *Science*, 2022, 377(6607): eabm4470.
- [3] ZHU W, BEROZA G C. PhaseNet: A deep-neural-network-based seismic arrival-time picking method[J]. *Geophysical Journal International*, 2019, 216(1): 261-273.
- [4] ROSS Z E, MEIER M A, HAUSSON E, et al. Generalized seismic phase detection with deep learning[J]. *Bulletin of the Seismological Society of America*, 2018, 108(5A): 2894-2901.
- [5] WANG J, XIAO Z, LIU C, et al. Deep learning for picking seismic arrival times[J]. *Journal of Geophysical Research: Solid Earth*, 2019, 124(7): 6612-6624.
- [6] MOUSAVI S M, ELLSWORTH W L, ZHU W, et al. Earthquake transformer: An attentive deep-learning model for simultaneous earthquake detection and phase picking[J]. *Nature Communications*, 2020, 11(1): 3952.
- [7] XIAO Z, WANG J, LIU C, et al. Siamese earthquake transformer: A pair-input deep-learning model for earthquake detection and phase picking on a seismic array[J]. *Journal of Geophysical Research: Solid Earth*, 2021, 126(5): e2020JB021444.
- [8] YU Z, WANG W. LPPN: A lightweight network for fast phase picking[J]. *Seismological Research Letters*, 2022, 93(5): 2834-2846.
- [9] YANG S, HU J, ZHANG H, et al. Simultaneous earthquake detection on multiple stations via a convolutional neural network[J]. *Seismological Research Letters*, 2021, 92(1): 246-260.
- [10] YANO K, SHIINA T, KURATA S, et al. Graph-partitioning based convolutional neural network for earthquake detection using a seismic array[J]. *Journal of Geophysical Research: Solid Earth*, 2021, 126(5): e2020JB020269.
- [11] ROSS Z E, MEIER M A, HAUSSON E. P wave arrival picking and first-motion polarity determination with deep learning[J]. *Journal of Geophysical Research: Solid Earth*, 2018, 123(6): 5120-5129.
- [12] HARA S, FUKAHATA Y, IIO Y. P-wave first-motion polarity determination of waveform data in western Japan using deep learning[J]. *Earth, Planets and Space*, 2019, 71(1): 127.
- [13] TIAN X, ZHANG W, ZHANG X, et al. Comparison of single-trace and multiple-trace polarity determination for surface microseismic data using deep learning[J]. *Seismological Research Letters*, 2020, 91(3): 1794-1803.
- [14] ZHAO M, XIAO Z, ZHANG M, et al. DiTing Motion: A deep-learning first-motion-polarity classifier and its application to focal mechanism inversion[J]. *Frontiers in Earth Science*, 2023, 11: 1103914.
- [15] KUANG W, YUAN C, ZHANG J. Real-time determination of earthquake focal mechanism via deep learning[J]. *Nature Communications*, 2021, 12(1): 1432.
- [16] LI S, FANG L, XIAO Z, et al. Focmech-flow: Automatic determination of P-wave first-motion polarity and focal mechanism inversion and application to the 2021 yangbi earthquake sequence[J]. *Applied Sciences*, 2023, 13(4): 2233.
- [17] ZHU W, MOUSAVI S M, BEROZA G C. Seismic signal denoising and decomposition using deep neural networks[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2019, 57(11): 9476-9488.
- [18] WANG T, TRUGMAN D, LIN Y. SeismoGen: Seismic waveform synthesis using GAN with application to seismic data augmentation[J]. *Journal of Geophysical Research: Solid Earth*, 2021, 126(4): e2020JB020077.
- [19] YANG L, LIU X, ZHU W, et al. Toward improved urban earthquake monitoring through deep-learning-based noise suppression[J]. *Science advances*, 2022, 8(15): eabl3564.
- [20] NOVOSELOV A, BALAZS P, BOKELMANN G. SEDENOSS: SEparating and DENOising Seismic Signals with dual-path recurrent neural network architecture[J]. *Journal of Geophysical Research: Solid Earth*, 2022, 127(3): e2021JB023183.
- [21] ROSS Z E, YUE Y, MEIER M A, et al. PhaseLink: A deep learning approach to seismic phase association[J]. *Journal of Geophysical Research: Solid Earth*, 2019, 124(1): 856-869.
- [22] MCBREARTY I W, DELOREY A A, JOHNSON P A. Pairwise association of seismic arrivals with convolutional neural networks[J]. *Seismological Research*

- Letters, 2019, 90(2A): 503-509.
- [23] MCBREARTY I W, GOMBERG J, DELOREY A A, et al. Earthquake arrival association with backprojection and graph theory[J]. Bulletin of the Seismological Society of America, 2019, 109(6): 2510-2531.
 - [24] YU Z, WANG W. FastLink: A machine learning and GPU-based fast phase association method and its application to Yangbi Ms6.4 aftershock sequences[J]. Geophysical Journal International, 2022, 230(1): 673-683.
 - [25] DEVRIES P M R, VIÉGAS F, WATTENBERG M, et al. Deep learning of aftershock patterns following large earthquakes[J]. Nature, 2018, 560(7720): 632-634.
 - [26] LOMAX A, MICHELINI A, JOZINOVIĆ D. An investigation of rapid earthquake characterization using single-station waveforms and a convolutional neural network[J]. Seismological Research Letters, 90(2A): 517-529.
 - [27] MOUSAVI S M, BEROZA G C. Bayesian-deep-learning estimation of earthquake location from single-station observations[J]. arXiv preprint arXiv: 1912.01144, 2019.
 - [28] ZHANG X, ZHANG J, YUAN C, et al. Locating induced earthquakes with a network of seismic stations in Oklahoma via a deep learning method[J]. Scientific reports, 2020, 10(1): 1941.
 - [29] ZHANG X, REICHARD-FLYNN W, ZHANG M, et al. Spatiotemporal graph convolutional networks for earthquake source characterization[J]. Journal of Geophysical Research: Solid Earth, 2022, 127(11): e2022JB024401.
 - [30] MCBREARTY I W, BEROZA G C. Earthquake phase association with graph neural networks[J]. Bulletin of the Seismological Society of America, 2023, 113(2): 524-547.
 - [31] LINVILLE L, PANKOW K, DRAELOS T. Deep learning models augment analyst decisions for event discrimination[J]. Geophysical Research Letters, 2019, 46(7): 3643-3651.
 - [32] KIM G, KU B, AHN J K, et al. Graph convolution networks for seismic events classification using raw waveform data from multiple stations[J]. IEEE Geoscience and Remote Sensing Letters, 2021, 19: 1-5.
 - [33] BREGMAN Y, LINDENBAUM O, RABIN N. Array based earthquakes-explosion discrimination using diffusion maps[J]. Pure and Applied Geophysics, 2021, 178: 2403-2418.
 - [34] KONG Q, WANG R, WALTER W R, et al. Combining deep learning with physics based features in explosion-earthquake discrimination[J]. Geophysical Research Letters, 2022, 49(13): e2022GL098645.
 - [35] MÜNCHMEYER J, BINDI D, LESER U, et al. Earthquake magnitude and location estimation from real time seismic waveforms with a transformer network[J]. Geophysical Journal International, 2021, 226(2): 1086-1104.
 - [36] SI X, WU X, SHENG H, et al. SeisCLIP: A seismology foundation model pre-trained by multi-modal data for multi-purpose seismic feature extraction[J]. IEEE Transactions on Geoscience and Remote Sensing, 2024.
 - [37] AN Y. Introduction to a recently released dataset entitled CSNCD: A comprehensive dataset of chinese seismic network[J]. Earthquake Research Advances, 2024, 4: 100255.
 - [38] YU Z, WANG W, Chen Y. Benchmark on the accuracy and efficiency of several neural network based phase pickers using datasets from China Seismic Network[J]. Earthquake Science, 2023, 36(2): 113-131.
 - [39] YUAN C, ZHANG J. Applying deep learning to teleseismic phase detection and picking: PcP and PKiKP cases[J]. arXiv preprint arXiv: 1910.09049, 2019.
 - [40] DING W, LI T, YANG X, et al. Deep neural networks for creating reliable PmP database with a case study in southern California[J]. Journal of Geophysical Research: Solid Earth, 2022, 127(4): e2021JB023830.