

深度学习在地震数据处理中的应用

李帅¹, 蒋策², 田平², 范莉苹¹, 夏登科¹, 房立华³

(1. 中国地震局地球物理研究所, 北京 100081;

2. 广东省地震局, 广州 510070;

3. 中国地震局地震预测研究所, 北京 100036)

摘要: 近年来, 以深度学习为代表的新一代人工智能技术在地震检测与震相拾取、地震事件分类、P波初动极性判别及震源机制反演等地震学领域取得了显著进展, 正逐步成为推动地震数据处理智能化转型的关键技术。笔者系统回顾了深度学习方法在地震学领域的主要应用进展, 重点介绍了近两年地震学大模型的发展和运用, 梳理了基于人工智能技术在地震科学发现方面的最新成果。在此基础上, 笔者指出当前深度学习在地震学应用中面临的数据质量、模型架构、评测体系以及AI赋能地震学发现等技术瓶颈与挑战, 并对未来发展趋势进行展望, 旨在为相关研究提供系统性参考和指导。

关键词: 深度学习; 人工智能; 地震数据处理; 地震学大模型

中图分类号: P631 文献标志码: A DOI: 10.12474/wthtjs.20250902-0001

Application of deep learning in seismic data processing

LI Shuai¹, JIANG Ce², TIAN Ping², FAN Liping¹, XIA Dengke¹, FANG Lihua³

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

2. Guangdong Earthquake Agency, Guangzhou 510070, China;

3. Institute of Earthquake Forecasting, China Earthquake Administration, Beijing 100036, China)

Abstract: In recent years, Deep Learning has made remarkable strides in various seismological domains, including earthquake detection and phase picking, event classification, P-wave first-motion polarity determination, and earthquake focal mechanism inversion. It is increasingly emerging as a key driver in the intelligent transformation of seismic data processing. This paper systematically reviews the major advancements in the application of deep learning within seismology, with a particular focus on the development and implementation of large-scale seismological models over the past two years. Furthermore, it summarizes the latest achievements in scientific discovery enabled by artificial intelligence in seismology. On this basis, the paper identifies current technical bottlenecks and challenges faced by deep learning in seismological applications, such as data quality, model architecture, evaluation systems, and the role of AI in facilitating scientific discovery. Finally, it provides an outlook on future trends, aiming to offer systematic reference and guidance for related research.

Keywords: deep learning; artificial intelligence; seismic data processing; large seismology models

收稿日期: 2025-09-02

基金项目: 国家自然科学基金(42374081)

第一作者: 李帅(1998—), 男, 博士, 主要从事微震监测与震源机制研究, E-mail: 11559293003s@gmail.com。

通信作者: 房立华(1981—), 男, 博士, 研究员, 主要从事人工智能地震学研究, E-mail: fanglihua@ief.ac.cn。

引用格式: 李帅, 蒋策, 田平, 等, 2025. 深度学习在地震数据处理中的应用[J]. 物探化探计算技术, 47(5): 623-640.

Li S, Jiang C, Tian P, et al., 2025. Application of deep learning in seismic data processing[J]. Computing Techniques for Geophysical and Geochemical Exploration, 47(5): 623-640.

0 引言

近年来,深度学习技术的高速发展正推动地震学研究范式的根本性转变(Wang et al., 2025; Zhu et al., 2025b)。密集台阵的高采样率与空间覆盖导致数据规模激增,传统人工处理方法已难以满足地震数据处理需求(廖诗荣等, 2021; Zor et al., 2025)。以中国大陆为例,近5年的地震台站数量由1 200多个快速增加至20 000多个(蒋策, 2025)。同时,随着页岩气开发、水库蓄水等活动的开展,微小地震的数量也呈现快速增加的趋势。传统的数据分析处理方法已无法满足地震数据自动处理的需求。

深度学习模型凭借其强大的特征自动提取能力与非线性映射能力,在地震检测(Ross et al., 2018a; Zhu and Beroza, 2018; Mousavi et al., 2020; Zhou et al., 2022; Li et al., 2023)、地震分类(Kong et al., 2022; Lu et al., 2025)、P波初动极性判别(Ross et al., 2018b; Hara et al., 2019; Zhao et al., 2023a; Han et al., 2025)以及震源机制反演(Li et al., 2023; Chen et al., 2024a;

Wang et al., 2024)等关键任务中得到广泛应用(图1)。

为系统评估深度学习在地震学中的研究和应用进展,笔者检索了Web of Science英文数据库近十年(2015—2024年)的文献,使用关键词组合“seismology”“geophysics”“artificial intelligence”“deep learning”与“machine learning”,一共有3 332篇相关文献。结果显示,相关研究结果随年份呈现出指数级增长(图2a),2019年发表量突破100篇,2024年达911篇,反映出该领域研究热度的持续攀升。

关键词作为反映研究核心的要素,其频次统计对把握研究焦点具有关键作用。笔者对每篇文献的关键词进行词频分析并生成词云图(图2b)。统计结果显示,“Deep Learning”(1 157次)、“machine learning”(1 078次)与“earthquake”(1 039次)位列前三。尽管检索阶段使用了“seismology”“geophysics”和“artificial intelligence”等广义术语,但在关键词频次中却显著低于更具体的技术与应用词汇。表明在人工智能地震学领域内,研究者更倾向于采用具体技术方法和明确应用对象来标注其研究,反映出人工智能技术正以高度专业化、场景化的方式深度融合

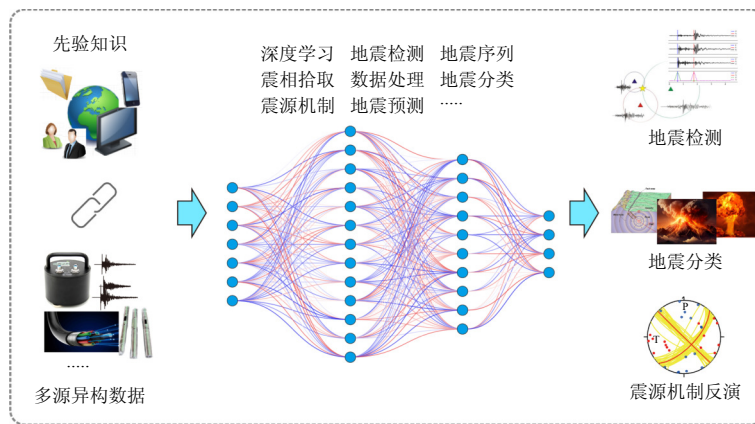


图1 深度学习方法在地震数据处理中的应用

Fig. 1 Application of deep learning in seismic data processing

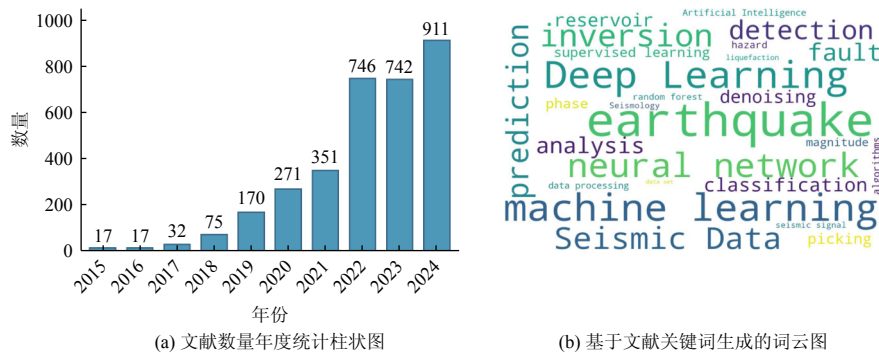


图2 2015—2024年人工智能地震学文献计量与关键词分析

Fig. 2 Bibliometric and keyword analysis of AI seismology literature from 2015 to 2024

于地震学研究之中。进一步地,“prediction”(955次)、“inversion”(692次)、“detection”(550次)、“fault”(536次)等关键词,突显了人工智能在地震预测、反演解释、事件检测以及断层识别等具体任务中的广泛应用。同时,“denoising”(339次)、“data processing”(250次)等技术性术语,体现了信号处理与数据质量控制等基础支撑环节仍受到持续关注。上述词频分布共同勾勒出以深度学习为代表的人工智能技术在地球物理学领域多层面、多场景的深入应用生态。

1 深度学习方法概述

深度学习依托深层神经网络架构,通过多层次非线性变换实现对原始数据的层级特征抽象,从而自主挖掘输入数据的分布式表征,无需依赖人工特

征工程即可完成模式识别与决策建模(LeCun et al., 2015)。深度学习中常用的网络结构主要包括卷积神经网络(Convolutional Neural Network, CNN)(LeCun et al., 1989)、循环神经网络(Recurrent Neural Network, RNN)(Zaremba et al., 2014)、自注意力机制(Transformer)(Vaswani et al., 2017)、自编码器(Auto-encoder, AE)(Bank et al., 2020)、生成对抗网络(Generative Adversarial Network, GAN)(Goodfellow et al., 2014)、图神经网络(Graph Neural Network, GNN)(Scarselli et al., 2009)等模型,以及在近几年在大语言模型(Large Language Models, LLMs),如 GPT(Radford et al., 2018)、DeepSeek(Bi et al., 2024; Zhao et al., 2025)等基础上发展的大规模预训练模型(Large-scale Pre-trained Models, LPMs)。这些模型及其变体在不同地震任务中展现出不同的优势(表 1):

表 1 不同人工智能模型在地震学中的应用
Tab. 1 Applications of different AI models in seismology

模型类型	主要特点	典型应用
CNN	局部感受野, 擅长空间特征提取	地震检测、信号分类
RNN/LSTM	处理时间序列, 捕获长短期依赖	震相识别、地震预测
Transformer	全局关系建模, 强大表达能力	多模态数据融合、预训练模型
AE	信号降噪和压缩编码, 特征提取	地震数据特征提取
GAN	生成模型, 数据增强	地震图像重建、数据扩增
GNN	处理非结构化数据	震相关联、端到端的地震检测
LPMs	大规模参数架构, 通用任务适应性	跨模态知识融合

卷积神经网络(CNN)是深度学习中应用最广泛的架构之一,最初主要面向图像和视频数据处理。其典型结构包含卷积层、池化层、激活函数及全连接层。CNN 利用卷积操作从输入数据中捕捉特定特征,池化操作则负责压缩数据维度。通过最小化损失函数并借助反向传播算法、梯度下降等优化方法学习模型参数,该网络能自动抽取数据中的特征与模式,实现分类、回归等任务(Lecun et al., 1998)。在地震监测领域, CNN 被有效地应用于地震波形分析,以提取波形特征(Hara et al., 2019; Ross et al., 2018a; Uchide, 2020; Zhu and Beroza, 2018)。其优势在于能够通过卷积层对高维地震数据进行降维和自动特征提取,显著提升处理这类任务的效率和精度。

循环神经网络(RNN)由 Zaremba et al.(2014)提出,专为处理序列数据而设计。其核心机制是在不同时间步间传递信息,将当前步的输出或隐藏状态作为下一步的输入。RNN 在语音识别、自然语言处

理等领域应用广泛。作为其重要变体,长短时记忆网络(Hochreiter and Schmidhuber, 1997)(Long short-term memory, LSTM)能有效克服传统 RNN 在处理长序列时面临的梯度消失的问题,从而更好地保留长期依赖信息。在地震学研究中, RNN 可用于分析地震事件序列数据,建立时间上的关联模型(Aydin et al., 2025)。

Transformer(Vaswani et al., 2017)是另一种用于序列建模的架构。其创新之处在于引入了自注意力机制(Self-Attention Mechanism),能够直接建模输入序列内部元素间的长距离依赖关系,规避了 RNN 固有的梯度问题。近年来,Transformer 在机器翻译、文本生成及自然语言处理(Natural Language Processing, NLP)等领域成果斐然, GPT-2(Radford et al., 2019)是其最显著的例子之一。在地震学应用中,Transformer 模型已被证明可有效应用于地震事件检测和震相识别等任务(Mousavi et al., 2020)。

自编码器(AE)(Bank et al., 2020)是一种无监督学习模型。它通过编码(压缩)和解码(重建)数据的过程,学习数据中的核心特征表达,实现数据的降维。AE 常用于数据降噪、压缩编码,为后续的分类、回归任务提供低维特征输入(如图像、音频、视频信号处理)。在地震学中,AE 主要用于波形去噪、对高维地震数据进行降维与特征提取(Jiang et al., 2022)。

生成对抗网络(GAN)(Goodfellow et al., 2014)由生成器和判别器两部分组成。生成器基于低维随机向量生成样本,判别器则负责鉴别样本的真实性。GAN 最初用于生成图像、视频等合成数据,近年来也被扩展应用于特征提取,以服务于分类或目标检测等任务。在地震学领域,GAN 可被用于生成合成地震波形数据或增强其他类型样本数据的多样性(Oliveira et al., 2018)。

图神经网络(GNN)(Scarselli et al., 2009)通过节点特征与图结构的联合建模机制,以及基于邻域信息传递的特征聚合策略,在处理复杂网络结构数据方面展现出独特优势。在地震学领域,该方法通过构建节点-边的拓扑关系,将地震台网拓扑图、三分量地震波形及台站地理位置信息组成图特征向量,在地震定位、震相关联和震级估算等任务中取得显著进展(Van and Ampuero, 2020; Feng et al., 2022)。

相较于传统单一架构的神经网络模型,现代大规模预训练模型(Large-scale Pre-trained Models, LPMs)的参数空间已达到百亿至万亿量级(Brown et al., 2020)。这种参数规模的扩展显著增强了模型对复杂数据模式的表征能力与泛化性能。通过自监督预训练范式在海量无标注数据上的优化,模型能够捕获跨领域的通用特征表示,借助少量标注数据实现高效的迁移任务(Devlin et al., 2019)。此类模型通过多模态融合机制整合文本、图像、音频及时序数据等异构信息,在跨模态表示学习与多任务迁移场景中展现出显著优势(Radford et al., 2021)。

2 主要应用领域

2.1 地震检测和震相拾取

地震检测和震相拾取是地震学研究的重要方向,也是地震数据自动处理的两个关键环节。地震检测的核心任务是从连续波形数据中识别地震信号特征,有效区分地震信号与背景噪声(包括自然噪

声、仪器噪声以及人为干扰),从而实现地震事件的初步定位(Mousavi and Beroza, 2023)。该任务本质上是针对长时间序列的二元分类问题,需兼顾计算效率与模型泛化能力。震相拾取则聚焦于精确识别 P 波和 S 波的初至到时,本质上是一个面向长时间序列的语义分割任务。它要求从复杂的波形特征中剔除噪声干扰,并依赖波形前后时序关联性,以精准标定震相到时。

自 2018 年以来,基于深度学习的震相拾取模型通过自动学习波形多尺度特征,在低信噪比条件下显著提升了地震事件检出率与到时精度。早期模型如 GPD(Ross et al., 2018a)、DetNet&PpkNet(Zhou et al., 2019)主要依赖 CNN 提取局部波形特征,但其感受野受限于局部卷积核的大小,导致对跨站、跨时段的相关信息捕获不足;PickNet(Wang et al., 2019)通过多尺度卷积层融合高频细节与低频语义特征,初步提升了 Pg 和 Sg 震相的区分能力。随后,基于对称编码器-解码器结构的 PhaseNet(Zhu and Beroza, 2018)及其改进模型 CSESnet(Li et al., 2023),通过波形分割实现了高精度震相拾取。

Mousavi et al. (2020)基于注意力机制提出的 EqTransformer (Earthquake Transformer) 是一个重要的里程碑。该模型首次将 Transformer 架构引入地震学研究中,融合了 CNN、RNN 与注意力机制,包含 17 层编码层和 8~10 层解码层,实现了地震事件检测与震相拾取的同步处理(图 3a)。在 2000 年日本 Tottori 地震近 5 周的连续波形数据中,仅使用不到三分之一的台站,模型检测并定位的地震事件数量较 JMA(Japan Meteorological Agency)目录增加两倍,且新增事件中大部分为 1.0 级以下的微震,显著提升了微震识别能力(图 3b)。

进一步地,在 EqTransformer 的基础上,S-EqT (Siamese Earthquake Transformer)(Xiao et al., 2021)创新性地利用孪生网络建模台站对间的物理关联性,在洛杉矶区域台阵 14 天高噪声连续观测中,较 EqTransformer 额外检出 40% 的震相。为满足工程化需求,轻量化模型如 LPPN(Yu and Wang, 2022)通过压缩网络结构显著降低计算开销,为实时波形处理提供了可行性。PhaseNO 模型(Sun et al., 2023)采用神经算子(Neural Operator)范式,通过核积分聚合不规则台站的跨台站特征,在 Ridgecrest 地震序列中事件检出率较 PhaseNet 提升 20%,召回率达 94%,实现了台网几何结构隐含物理规律的学习。

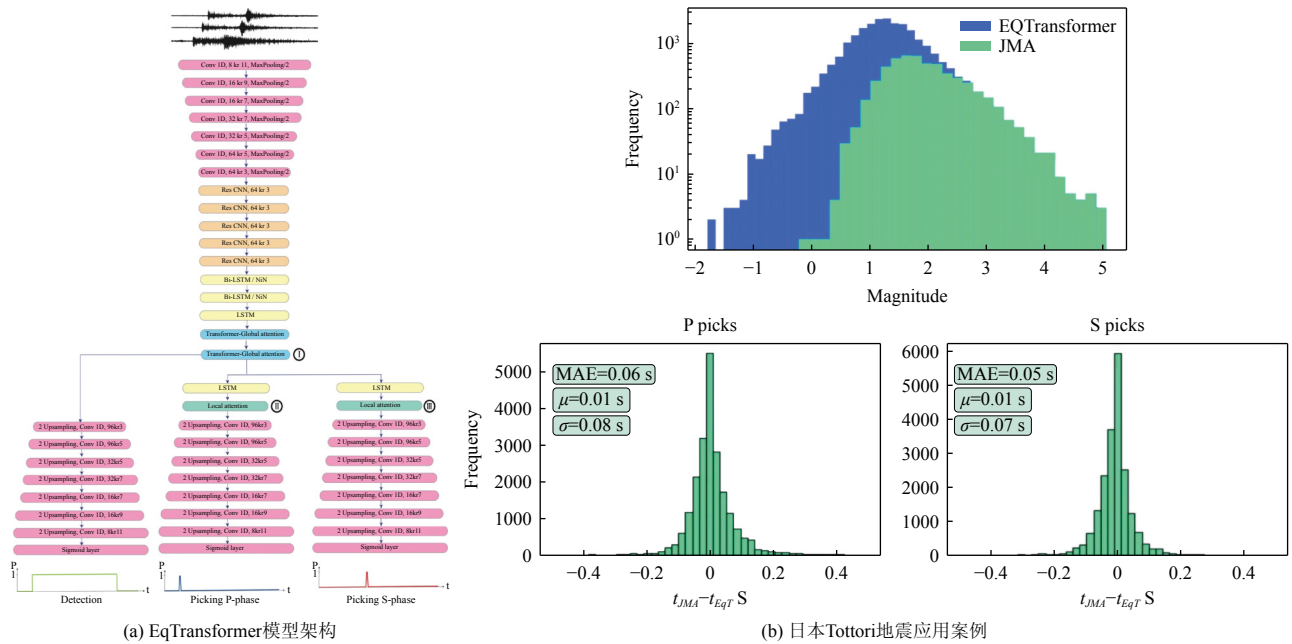


图3 EqTransformer 模型架构及其应用实例 (修改自 Mousavi et al., 2020, 图 1 和图 6)

Fig. 3 Architecture and Application of the EqTransformer Model

2.2 地震事件类型判别

地震事件类型判别是地震学中的一项经典分类任务,旨在区分天然地震与诸如工业爆破、矿山塌陷、核试验等非天然地震事件。此类事件震源深度一般不大,与震级较小的浅源地震在波形特征上存在一定共性,为准确判别事件类型带来挑战。将天然地震与非天然地震进行区分是利用深度学习进行分类的主要研究内容之一。深度学习方法凭借卓越的自动特征提取能力与非线性映射特性,突破了传统方法对人工特征工程的依赖。该方法直接以地震波形或其频域谱作为输入,使用监督学习或半监督学习框架从地震波形中提取与震源相关的参数特征完成分类任务。

最早由 Dowla et al.(1990)基于 RNN 对天然地震与地下核试验的区分进行研究,发现 RNN 在地震事件类型判别中相比于传统方法展现出的大量计算优势。Mousavi et al.(2016)基于人工神经网络和逻辑回归方法(Logistic Regression, LR),对美国路易斯安那州盐丘附近地震记录中的发震位置较浅的塌陷事件与较深的微震事件进行区分。

随着计算机性能的进一步提升,以 CNN 为代表的深度学习算法,由于其通过共享权重,可以构建更深的网络层数,提取特征的能力得到极大提高,同样被应用到事件类型判别领域。Linville et al.(2019)利用美国犹他州 13 313 起地震、爆破事件记录,将频谱作为输入进行训练,在该区域进行的地震事件

类型二分类测试准确率均达到 99% 以上; 隗永刚等(2019)研究了将波形功率谱作为训练和识别的样本进行深度模型训练,发现可以对天然地震和爆破进行准确地区分,三组试验结果的准确率均可以达到 97% 以上; Miao et al.(2020)基于神经网络结合人工提取的特征对美国肯塔基州的爆破识别进行研究; Liu et al.(2020)构建的 7 层 CNN 模型在 2009—2018 年中国大陆 4 842 个地震事件数据集上实现了 93% 的构造地震判别准确率。这些早期研究为后续方法改进奠定了基础。

后续研究进一步探索了输入特征与网络结构的优化组合。周少辉等(2021)分别利用 AlexNet、VGG16、VGG19 和 GoogLeNet 卷积神经网络结构进行学习训练,对各判别模型进行测试发现,各个卷积神经网络在训练过程中代价函数的走势曲线基本一致, AlexNet 网络结构的事件检测结果(98.5%)优于其他 3 种网络结构; Kong et al.(2022)在卷积神经网络基础上,加入事件中的 P/S 振幅比和地方震级与尾波持续震级差值这两种物理特征进行约束,对事件类型进行综合判别,具有较高的判别准确度和泛化能力; 田宵等(2022)基于波形、时频数据、多道波形的单输入网络和基于时频数据和多道波形的多输入网络的四种 CNN 分类模型进行天然地震和爆破事件的区分研究,发现以时频数据作为输入的单输入 CNN 模型精度高于以波形数据作为输入的结果。Zhu et al.(2023)基于卷积神经网络(CNN)训练了地震事

件类型判别模型(图 4),该模型以原始波形数据作为输入,通过多层级卷积与池化操作自动提取波形的时频特征,并最终通过全连接层实现分类。该模型对美国南加州爆破事件的识别准确率达到 98.1%,显著降低了人工判别的工作量。此外,该研究成功验证了跨区域迁移学习的有效性,将在南加州训练得到的模型迁移至美国肯塔基州东部后,仍能保持优异的判别性能(图 5),为解决深度学习模型在实际应用中面临的区域泛化难题提供了关键实证。

Hourcade et al.(2023)基于卷积神经网络将频谱作为输入,训练得到的模型表现出优异的泛化能力,

在法国和美国犹他州的判别结果准确率均达到 95% 以上;Maguire et al.(2024)分别基于 8 个不同区域的波形数据在同一卷积神经网络构架下进行训练,发现不同数据集的使用对模型的泛化性产生比较大的影响;Lu et al.(2025)以波形、频谱和振幅比作为三分支 CNN 模型的输入,进行对地震、爆破的二分类以及地震、爆破和塌陷的三分类研究,其中跨区域的二分类表现出良好的泛化性能;Hong et al.(2025)利用福建测震台网记录的爆破事件,基于 AlexNet、VGG 和 ResNet 构架的三种模型均取得 95% 以上的判别准确率。这些研究共同揭示了深度学习方法在

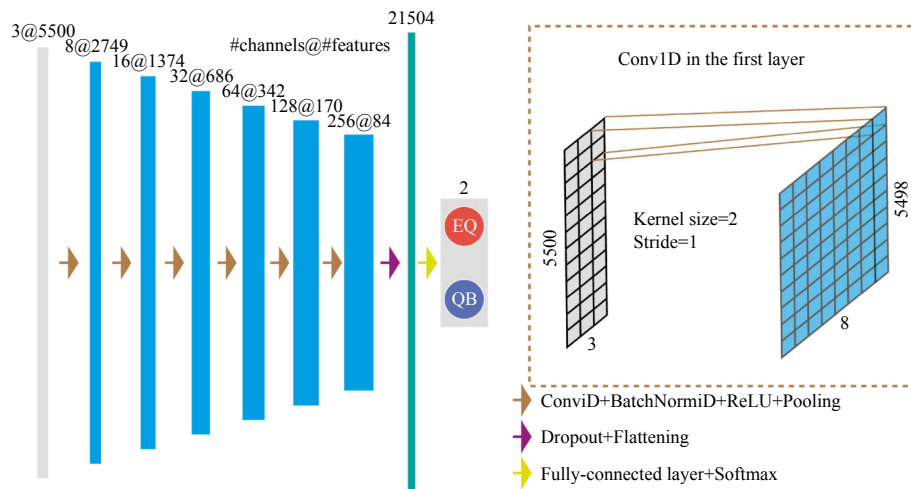


图 4 用于地震分类的 CNN 模型架构 (引自 Zhu et al., 2023, 图 4)

Fig. 4 CNN model architecture for earthquake discrimination

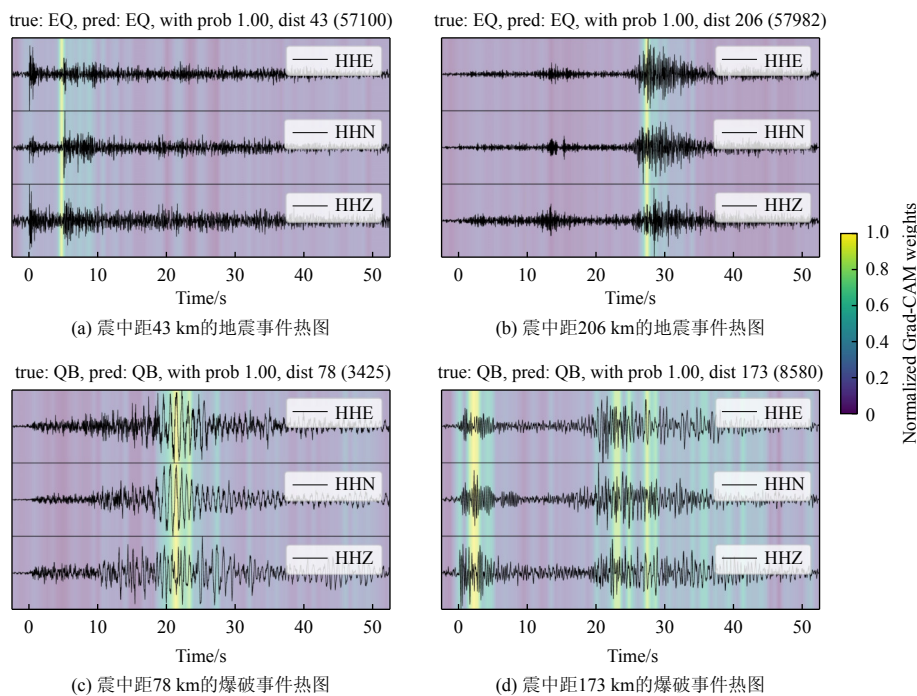


图 5 肯塔基州东部 2 次典型地震和 2 次爆破事件的热图及对应波形 (引自 Zhu et al., 2023, 图 9)

Fig. 5 Heat maps and corresponding waveforms of of four typical EQs and blasts in eastern Kentucky

地震事件类型判别中的技术演进路径与应用潜力。

2.3 利用 P 波初动极性反演震源机制

P 波初动极性是地震发生后 P 波在地震仪上记录的初始振动方向。准确判别 P 波初动极性对震源机制反演 (Hardebeck, 2002; Hardebeck, 2003; Nakamura, 2004; Hardebeck and Shearer, 2008; 俞春泉等, 2009)、震源区应力场分析 (许忠淮等, 1979; 许忠淮等, 1983; Uchide, 2020)、微震监测 (Skoumal et al., 2022; Cheng et al., 2023; 任义等, 2025) 与地下核爆识别 (Chi-Durán et al., 2021) 等具有重要意义。随着地震台站密度的快速增加, 传统人工判别方法已难以满足海量数据处理需求, 亟需高效自动化解决方案 (Li et al., 2023)。

深度学习方法在地震 P 波初动极性判别中的研究经历了从单一任务模型到多源数据联合训练, 再到跨模态迁移策略的逐步演进。为提高 P 波初动极性判别效率和一致性, Ross et al. (2018a) 率先提出基于卷积神经网络的 P 波初动自动化判别模型, 并在南加州地震台网 2000—2017 年数据中验证其有效性。此后, Hara et al. (2019) 和 Uchide (2020) 利用日本区域台网数据训练模型, 显著提升了极性判别效率; Chakraborty et al. (2022) 通过意大利 INSTANCE 数据集 (Michellini et al., 2021) 验证了跨区域应用的可行性。为提升低信噪比地震波形中 P 波初动极性判别的准确性, Tian et al. (2020) 提出基于多通道波形特征融合的卷积神经网络 (Multitrace-based CNN, MT-CNN) 采用相邻台站的多道地震记录作为训练样本, 通过构建空间极性分布特征的多通道卷积架构, 有效增强了微震数据的极性判别稳定性, 提高了微震震源机制反演的准确性。

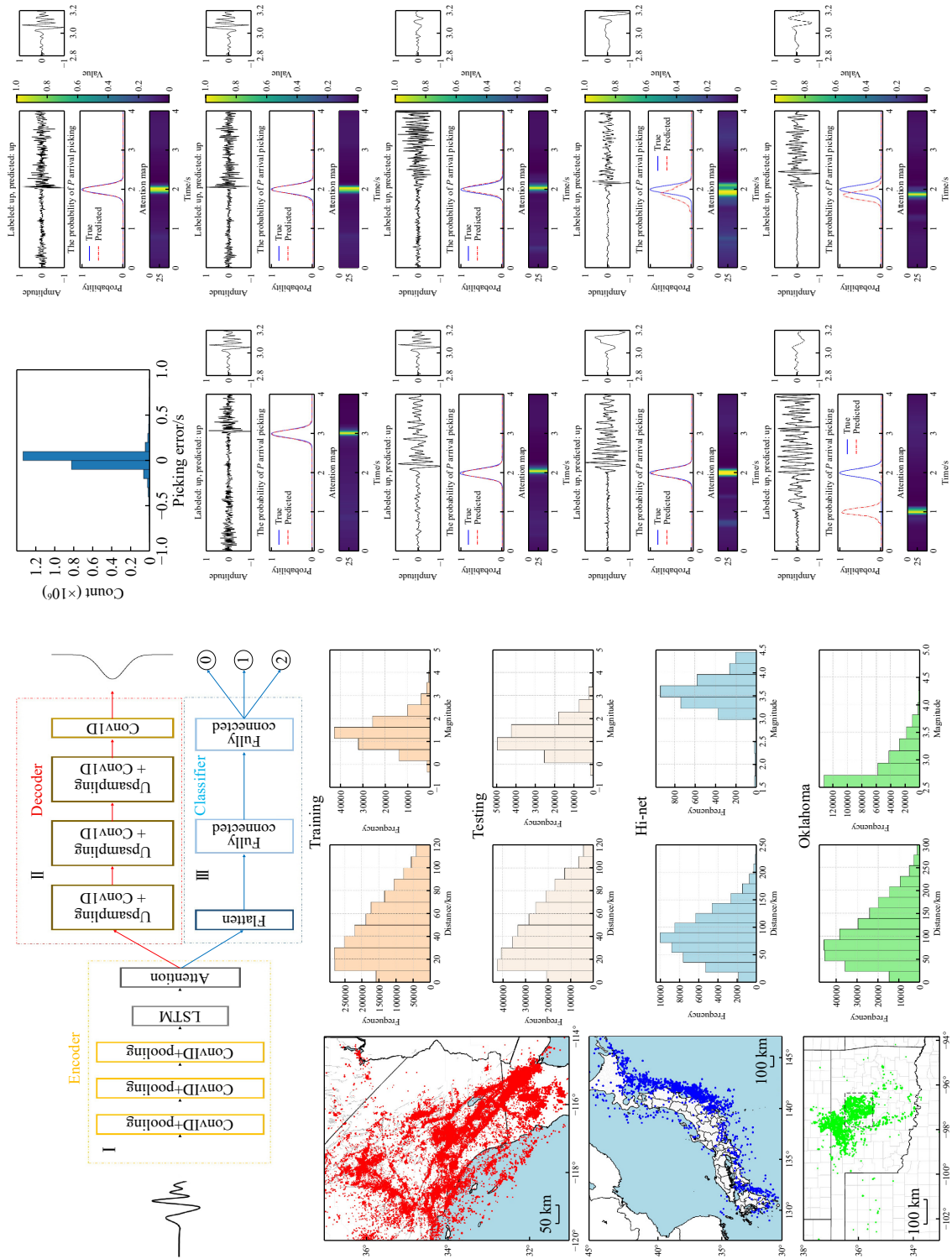
为解决单一区域数据训练的模型泛化性不足问题, Zhao et al. (2023a) 提出了 DiTingMotion 模型, 融合南加州与中国大陆“谛听”数据集 (Zhao et al., 2023b), 结合时频域多尺度特征, 在低信噪比条件下实现 98.49% 的判别准确率, 并在 2021 年漾濞地震序列中验证了其可靠性 (Li et al., 2023)。Zhai (2022) 提出 NPC (Network for Polarity Classification) 模型, 结合卷积神经网络与注意力机制, 训练数据涵盖多区域波形, 显著提升了模型在不同构造环境中的适应性。Chen et al. (2024a) 提出 Eqpolarity 模型, 采用 CCT 架构与迁移学习, 在得州 TexNet 数据集中获得 98.82% 的准确率, 证明小样本微调在新台网部署中的有效性。针对实际观测中极性数据常存在缺失

的问题, Skoumal et al. (2022) 提出基于随机森林的无监督学习方法, 通过迭代算法填补初动极性数据, 结合 k-modes 聚类显著提升了复合震源机制解精度。该方法通过对缺失极性进行概率填补, 降低了因台网稀疏导致的极性缺失率, 提升了整体震源机制求解的稳定性。

Zhang et al. (2023) 提出一种基于注意力机制的多任务神经网络模型 (Attention mechanism for Picks and Polarity, APP), 由编码器、解码器和分类器组成, 其架构如图 6a 所示。该模型结合事件检测算法确定分析窗口位置, 输入包含 P 波的 4 s 垂直分量波形, 输出 P 波到时和初动极性的概率。在日本、美国南加州和俄克拉荷马数据中均表现优异 (图 6b), 显著提升了模型在低信噪比和跨区域条件下的性能。

Han et al. (2025) 提出 RPNNet (Robust P-wave First-motion Polarity Determination Network) 模型, 结合 Inception 模块与注意力机制, 通过多尺度卷积捕捉波形细节, 并引入蒙特卡洛 Dropout 技术量化不确定性, 在美国西部和日本 Hi-net 数据集 P 波到时随机偏移 0.5 s 的情况下证实了其可靠性与实用价值。PhaseNet+ 模型 (Zhu et al., 2025b) 通过多任务框架整合震相拾取、初动极性判别与震相关联, 在 2019 年 Ridgecrest 序列中生成的地震目录比传统方法多检测 73%~81% 的事件, 结合极性数据后反演震源机制解数量提升至 SCSN 目录的 4 倍。

除 P 波初动极性判别外, 深度学习在地震波形反演中的应用已成为震源机制研究的重要方向。Kuang et al. (2021) 提出的 FMNet (Focal Mechanism Network) 模型通过理论地震图训练, 构建了中强地震震源机制的高效反演框架。Wang et al. (2024) 提出的多任务学习震源机制网络 MTFMN (Multi-Task Focal Mechanism Network) 通过构建卷积神经网络 (CNN)、双向长短时记忆网络 (Bi-LSTM) 与自注意力机制 (Self-Attention) 的复合架构, 使用多门控混合专家模型 (MMoE) 实现动态任务权重分配, 在保留地震学先验知识物理机制约束的同时, 通过数据驱动优化提升模型泛化能力, 在工程地震学领域展现出重要应用潜力。类似地, 针对传统网格搜索算法数据利用不充分、固定权重以及计算效率等问题, Song et al. (2025) 将 Transformer 架构引入震源机制反演过程, 提出 FocoNet 模型, 通过自注意力机制学习不同台站间的复杂关系, 突破了传统方法单台独立反演的局限, 通过将 P 波初动极性、多通道 S/P 振幅比、



(b) P波震相检测和极性判别结果

(a) APP模型架构与SCSN、Hi-net和Oklahoma数据集事件分布

图 6 APP 模型架构、不同数据集分布及 P 波检测结果 (修改自 Zhang et al., 2023, 图 1 和图 2)

Fig. 6 APP model architecture, distribution of different data sets and P-wave detection results

波形信噪比以及台站相对位置坐标构建了融合多元信息的 token 向量为准确确定震源机制提供了更全面的约束。

2.4 地震学大模型的应用

在人工智能领域,大型预训练模型(Large Pre-trained Models, LPMs)通常指基于海量多模态数据构建、具备数十亿至万亿级可训练参数的深度学习架构。这类模型通过无监督或自监督学习范式,在自然语言处理、计算机视觉等领域展现出强大的特征提取与迁移学习能力。聚焦地震学领域,近年涌现出一系列专门针对地震波形、文本、地震元数据等分析的大模型架构,包括 SeisT、SeisCLIP、SeisLM、Prime-DP、SeisMoLLM 及“谛听”地震波等大模型。

传统深度学习框架在事件分类和震源机制反演等特定任务模型训练中面临标注数据稀缺和区域迁移时的泛化能力受限的困境(Si et al., 2024)。近年来,地震学领域通过跨模态迁移、多任务协同训练及分布式计算等技术创新,发展出一系列基于深度学习的大型预训练模型,显著提升了多模态数据处理能力与跨任务适配性。

SeisT(Li et al., 2024)模型构建了多任务联合学习架构,通过共享特征提取层实现参数复用策略,整合震相识别、震源定位与震级估计三大核心任务。基于川滇地区 2 000 个地震事件的实证研究表明,该模型将震中定位误差中位数降至 2.1 km,震级预测绝对误差控制在 0.3 以内。

SeisCLIP(Si et al., 2024)模型受音频基础模型 AudioCLIP(Guzhov et al., 2021)启发,基于音频信号与地震波在时序特性上的共性特征,通过融合地震频谱特征与事件震相及震源信息进行对比训练,构建了包含 Transformer 编码器与多层感知机(MLP)模块双分支架构的基础地震学模型。在事件分类、定位及震源机制分析等任务中,该模型通过微调、特征冻结及从头训练三种策略均取得优于基准模型的性能,尤其在定位任务中展现出显著优势。

SeisLM(Liu et al., 2024)模型构建了地震波形分析的自监督学习框架,采用卷积网络(ConvNet)编码器、量化模块与 Transformer 编码器的复合架构,通过对比损失函数与掩码信号重建目标实现通用特征学习。该模型在大规模未标注地震数据上的预训练成果,在事件检测、相位识别、震相到时回归及前震-余震分类等任务中表现卓越,尤其在数据稀缺场景下验证了自监督方法的泛化潜力,为地震预警与风

险评估提供了新的技术路径。

Prime-DP(蔡育琦等, 2025)模型针对滑坡微地震、爆破、塌陷等非天然地震事件构建多分类框架,通过迁移学习策略结合 262 条微地震数据实现模型优化。实验结果显示,该方法将分类精度从 79.4% 提升至 94.8%,性能超越传统短时傅里叶变换(STFT)特征工程方法,揭示了预训练模型在复杂地震信号分类中的泛化优势。

SeisMoLLM(Wang et al., 2025)模型开创了地震监测领域的跨模态迁移新范式,采用 LoRA(Low-Rank Adaptation)策略对 GPT-2 进行参数微调(图 7),仅用 10% 的训练参数在 DiTing 与 STEAD 两大数据集上实现跨模态特征学习;在 DiTing 和 STEAD 数据集的震中距离估计、震级预测等五项任务中取得 SOTA (state-of-the-art)性能,43 项评估指标中 36 项达最优,少样本学习场景下的泛化能力提升 10%~50%。该模型在 4 张 GPU 卡上训练 2 天,将大模型训练的硬件需求、时间成本和经济成本降低 90% 以上。

“谛听”地震波大模型(赵明, 2025)采用 MAE (Masked Auto Encoder)技术架构,基于“谛听”数据集进行预训练,具备地震识别、震级估计、地震预警预测等功能。模型通过 0.1 B 参数规模的深度神经网络对海量地震波形数据进行学习,实现了高精度震相识别与地震序列自动化处理能力。

2.5 人工智能驱动的科学发现

人工智能驱动的科学发现(AI for Science)通过融合理论模型、数据科学和深度学习技术,在经验观察、理论推演、计算模拟与数据驱动四大研究范式的基础上构建了“人机协同、跨界融合”的第五研究范式(AI4Science)。这一范式通过算法模型与物理规律的协同优化,已在全球生物医学、自动驾驶等前沿领域取得突破性进展,其在地震学领域的应用更催生了多项研究成果。

AI 驱动的高分辨率地震目录不仅提升了数据处理效率,更揭示了传统方法难以捕捉的深部动力学特征,为地震学研究开辟了新的技术路径。三维断层几何形态是解析震群(swarm)时空演化机制的关键基础,对提升地震预测精度和危险性评估的可靠性具有重要意义。Ross et al.(2020)基于深度学习模型构建的高分辨率地震目录对美国加州南部四年震群活动的系统性分析,表明断层区的通道结构、屏障特征及几何弯曲等三维构造要素是调控震群动态演化的核心控制因素。研究进一步揭示,断层带渗透

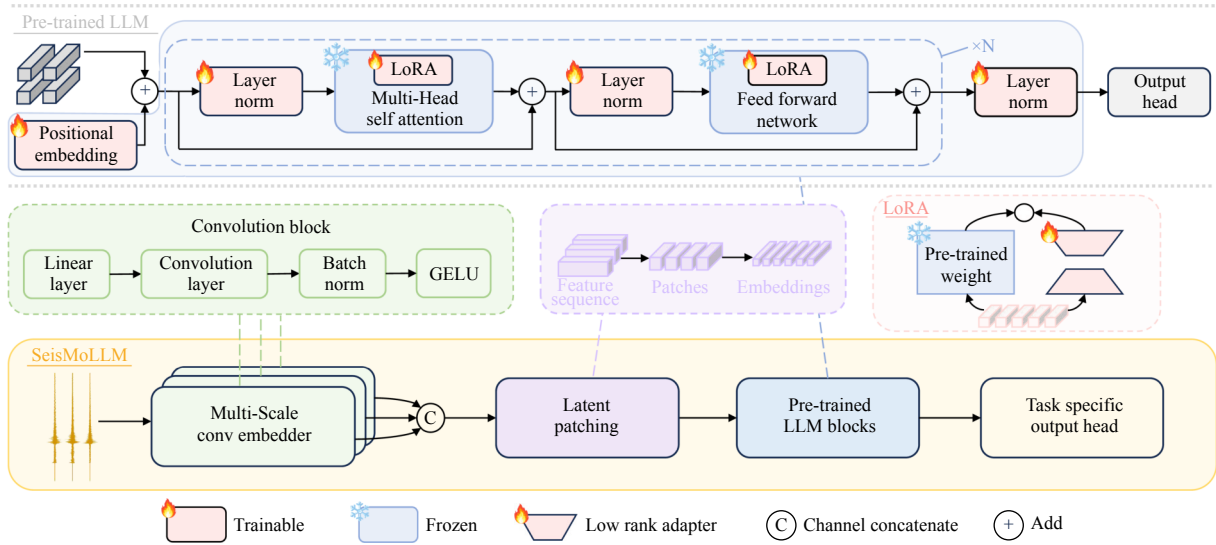


图 7 SeisMoLLM 模型架构 (引自 Wang et al., 2025, 图 1)

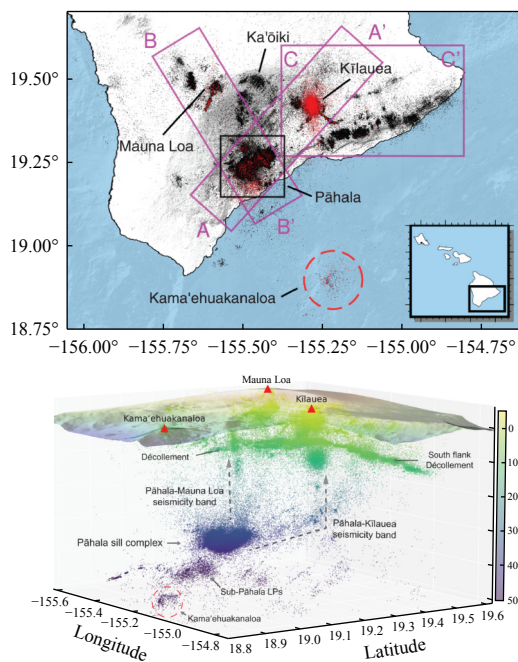
Fig. 7 Architecture of SeisMoLLM

性结构的空间异质性随时间演化将显著改变流体迁移路径与地震触发模式的时空分布特征。该发现为地震物理过程的机理解释、流体-断层相互作用机制的量化研究以及灾害评估模型的优化提供了新的观测依据,同时凸显了传统静态二维构造模型在刻画流体动力学与断层响应动态关系方面的理论局限性,为后续研究指明了三维结构时变特征在地震灾害预测中的关键作用。

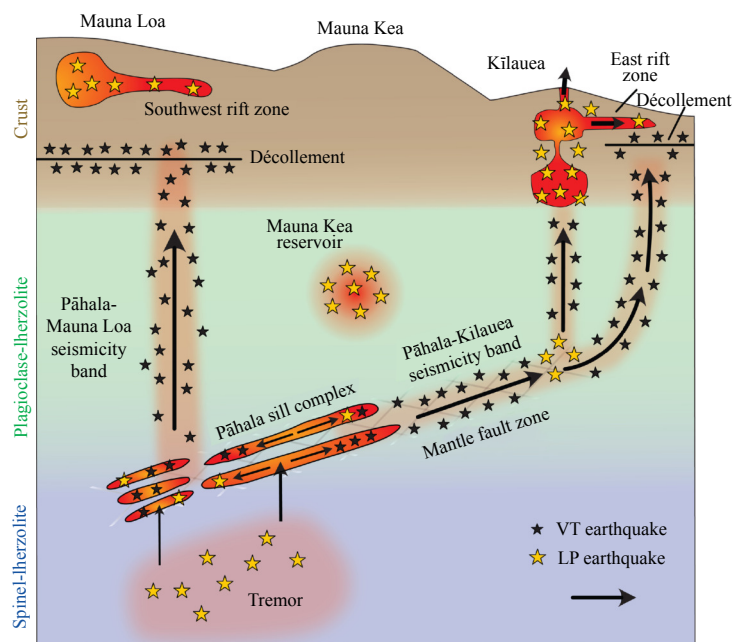
另一项基于深度学习的实证研究(Wilding et al., 2023)通过构建包含 20 万余次地震的高分辨率目录,

成功解析了夏威夷火山系统底部 15 km 深度的离散层状结构(图 8a)。该研究证实近水平岩床作为岩浆运移枢纽,其弹性波速异常与电阻率特征共同指向部分熔融态物质的存在。这一成果首次将地震层析成像与岩浆动力学模型相结合,为深部岩浆房与浅部火山系统的连通性机制提供了直接证据(图 8b),拓展了人工智能在地质结构解释中的应用边界。

在实验室尺度的地震预测领域,人工智能同样展现出突破性潜力。Borate et al.(2023)构建的物理信息神经网络(Physics-Informed Neural Network, PINN)



(a) Hawai'i地震活动概述



(b) 观测结论卡通图

图 8 Hawai'i 地震活动概述与观测结论卡通图 (修改自 Wilding et al., 2023, 图 1 和图 6)

Fig. 8 Cartoon of Hawai'i seismic activity overview and observation conclusions

框架通过融合断层带弹性波速(V_p/V_s)与地震波振幅的时序演化特征,实现剪切应力与滑移率历史的准确预测。该模型引入两个物理约束:弹性耦合方程与刚度-传输系数关系式。在数据稀缺场景下,PINN模型预测精度较纯数据驱动方法提升 10%~15%,且在跨数据集迁移学习中展现出显著优势。该研究为实验室地震向实际断层预测的转化提供了理论桥梁。

Cui et al. (2023) 基于全球震源时间函数(Earthquake source time functions, STFs)数据,结合机器学习

算法对全球 3 000 余次 $M_w > 5.5$ 级地震破裂动力学特征与震源深度相关性展开系统性研究。发现深源地震(300 km~700 km)与浅源地震(0~60 km)的破裂特性差异主要由地球介质的深度依赖性刚性控制,而非传统认知中的触发机制差异。基于观测数据分析,提出“常应变降假说”:随着深度增加,地震破裂尺度和滑动量减小,破裂持续时间缩短,而应力降显著升高(图 9)。在应变降恒定的理论框架下,不同深度地震的破裂尺度与滑移量之比维持相似关系,该发现为修正地震自相似性理论提供了实证支持。

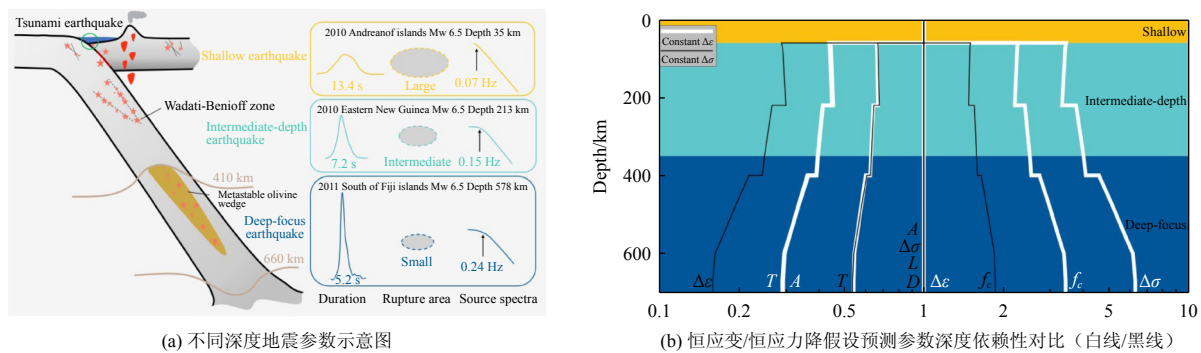


图 9 常应变降假说下地震参数的系统变化(修改自 Cui et al., 2023, 图 4)

Fig. 9 Systematic variations in earthquake parameters under the constant strain drop hypothesis

3 当前面临的挑战与未来发展趋势

3.1 数据集分散与标注问题

当前主流地震数据集主要聚焦于特定区域或任务,其空间分布和数据规模存在明显不均衡性,难以满足全球多场景地震研究的系统性需求。日本 Hi-Net、F-Net 和 K-Net 地震数据集(Okada et al., 2014)以高密度台网数据为特点,广泛应用于 P 波初动极性判别(Uchide, 2020)和微震检测(Kaneko et al., 2023)等任务,有效提升了复杂介质环境下震相识别的效率与精度。“谛听(DiTing)”数据集(Zhao et al., 2023b)和“震典(CSNCD)”数据集(安艳茹等, 2023)则聚焦中国大陆区域,通过高密度观测数据为深度学习模型的训练提供了关键支撑,而美国南加州数据集和北加州数据集则主要覆盖南加州与北加州地区,被广泛用于地震事件检测模型的开发(Ross et al., 2018a)。尽管 STEAD(Mousavi et al., 2019)数据集包含全球范围的事件波形,但已标注的事件主要集中在美国与欧洲,呈现出显著的空间分布不均匀特征。INSTANCE(Michelini et al., 2021)数据集主要来自于意大利,整合多模态数据用于地震分类研究;

TXED(Chen et al., 2024b)和 CEED(Zhu et al., 2025a)则集中于美国德克萨斯州和加州,强调实时处理与轻量化模型的适配性。

上述区域性数据集在特定应用场景中已取得显著成效,通过深度学习方法在高分辨率地震目录生成和震源机制反演等领域实现技术突破。然而,现有数据集空间分布高度不均,北美、欧洲等地震数据集占比较高,制约了全球化标准数据集的发展;数据采样率、格式、长度、预处理过程等不统一,导致各数据集之间融合较差;数据集标注依赖专家主观判断与不同标注标准,台网密度和仪器响应等观测层面的差别则影响波形质量,从而间接导致特征提取误差(Wu et al., 2023; 易继东等, 2023)。

针对目前数据集空间分布不均的问题,未来可在 SeisBench(Woollam et al., 2022)平台的基础上,进一步实现统一元数据标准,整合全球范围内的地震数据资源,推动不同区域数据集之间的协作与共享,提升数据利用效率。同时,地震学领域特有的连续波形数据为自监督任务设计提供了天然场景:通过对比学习(Contrastive Learning)或时序预测任务自动生成伪标签,再使用交叉验证或少量人工标签进行校准,可显著减少对人工标注的依赖。在 STEAD

数据集中已展现出潜力,通过结合未标注数据与少量标注样本,缓解人工标注成本高、周期长的问题,提高数据集构建的自动化水平和可扩展性。此外,地震学作为一门高度依赖先验知识的学科,其数据集建设应不同于计算机视觉领域已成熟的 ImageNet (Russakovsky et al., 2015)、COCO(Lin et al., 2014)等标准化基准体系。

Wu et al.(2023)在深度神经网络(Deep Neural Network, DNN)中引入地质和地球物理的先验约束,显著提高了 DNN 在处理复杂地球物理数据时的性能,表明将领域知识融入模型架构与训练过程,是提升模型泛化能力与物理一致性的关键路径。然而,此类方法的有效性高度依赖于高质量、具有明确地质构造含义的训练数据。为此,需建立融合区域地质构造参数与地球物理约束的训练数据集,并配套构建定期更新的评测数据集。定期更新的评测数据集有助于持续评估模型性能,并反映地震学研究的最新进展。

3.2 模型架构创新瓶颈

在医学图像分割领域, Ronneberger et al.(2015)提出的 U-Net 神经网络因其优异的特征提取与上下文感知能力,为地震学领域提供了重要的技术启发。Zhu and Beroza(2018)基于 U-Net 框架开发了 PhaseNet 模型,用于地震波形的震相拾取任务,通过引入多尺度特征融合和动态权重机制,显著提升了微震事件检测的鲁棒性,尤其在低信噪比环境中表现突出。随着自然语言处理领域 Transformer 模型(Vaswani et al., 2017)的兴起, Mousavi et al.(2020)将其扩展至地震学,提出了 EQTransformer,利用自注意力机制对地震波形的时空特征进行建模,实现了跨区域地震事件检测的精度提升。

后续深度学习技术在自然语言处理(如 GPT-2、LLaMA 系列)(Radford et al., 2019; Grattafiori et al., 2024)、语音识别(Wav2Vec2)(Baevski et al., 2020; Hsu et al., 2021)及计算机视觉(Vision Transformer)(Radford et al., 2021)等领域取得突破性进展。以 OpenAI 的 GPT 系列为例,其从 GPT-1(2018 年, 1.17 亿参数)到 GPT-5(2025 年, 万亿级参数)的迭代路径清晰展现了架构优化、数据规模扩展与多模态融合的系统性演进特征。然而,当前地震学领域仍存在明显的模型迭代滞后现象:现有算法版本多处于 1.0 阶段,仅在特定应用场景中进行局部参数微调(fine-tuning),模型在跨区域迁移应用时表现出泛化

问题,缺少针对地震波形特征的结构化优化策略。

针对现有深度学习模型在地震学应用中面临的创新性瓶颈,可从以下两个维度构建突破路径:首先,针对区域化定制模型的动态适应需求,建立基于数据演进的持续迭代机制——通过逐年累积的高质量观测数据进行模型重训练与版本迭代,可有效应对台网部署变化、地震序列前后地质构造时变特性,从而增强模型对区域地质特征的适应能力。其次,需构建融合物理约束的多任务协同学习框架,通过整合同一区域的多模态地震任务(如检测、到时拾取、极性判别、事件分类等),结合无监督学习与多任务学习策略,将先验地质信息(如断层网络结构、介质属性)和地球物理定律(如波动方程、震源机制约束)嵌入网络架构设计,实现任务间知识迁移与特征共享(Wu et al., 2023)。最终构建具备物理可解释性的跨任务协同分析系统。

3.3 评测体系规范化不足

当前已有多种架构和不同数据集训练的深度学习模型用于地震检测和震相拾取任务。PhaseNet 模型在 2019 年 Ridgecrest 地震序列取得不错的效果(Cianetti et al., 2025),但是在四川盆地应用时的拾取准确率都较在南加州地区应用时要低(林学楷和许才军, 2022)。同样, EQTransformer 在 STEAD 数据集上表现结果优秀,在南加州 14 天连续波形应用中对低信噪比的波形记录处理效果欠佳,存在较高的漏检率(Xiao et al., 2021)。因此,有必要建立统一的评测标准,对数据集和深度学习模型进行评测。

尽管研究者已尝试构建地震检测模型的评价框架(Cianetti et al., 2021; Jiang et al., 2021; Mousavi et al., 2020; Münchmeyer et al., 2022; Yu et al., 2023),但现有评估方法仍存在评测数据集的公开性不足和数据泄露、评估规模受限于区域数据分布、评测规模小以及评测维度单一等问题。

近年来人工智能在地震监测中的实践为评测体系建设提供了重要参考。以 2017 年由 中国地震局地球物理研究所与阿里云联合发起的“余震捕捉 AI 大赛”为例,该赛事首次将大数据与深度学习技术引入余震监测领域。通过汶川地震与美国纳帕地震的实测数据对比,系统评估了传统人工拾取方法与 AI 算法在处理短时密集余震序列时的效能差异。赛事历时 5 个月,吸引全球 1 143 支队伍参与,开创了地震检测算法量化评测模式(Fang et al., 2017; Bergen et al., 2019)。2025 年,中国地震局地震预测研究

所联合上海人工智能实验室和上海市地震局发布的 PrismaX Domain Leaderboard(<https://primax.opencompass.org.cn/domainlb>)通过整合全球地震数据,构建了面向主流地震检测模型的系统性评估体系。该评估框架采用多维度指标体系,涵盖精度(Precision)、召回率(Recall)、F1 分数、均值误差(μ)及标准差(σ)等关键性能参数,对 14 个主流地震检测模型(含 5 个地震学大模型)进行全面量化分析。研究结果表明:①大模型与专用模型在不同任务场景下呈现差异化优势,其中专业模型在推理效率方面表现更优;②现有模型对 S 波拾取精度仍存在显著提升空间,P 波拾取性能相对更优;③同一评估指标下,不同模型的性能差异显著。

未来,应在现有基础上,进一步完善评测体系,建立统一的评测标准。集成多模型评价机制,设计覆盖震相拾取、震源定位及震源机制反演等多任务的综合性能评价指标体系,为不同模型的横向比较与持续改进提供科学依据。

3.4 人工智能驱动科学发现较少

当前深度学习在地震学中的应用呈现显著的任务偏向性,其研究重心仍集中于数据密集型任务(如事件检测与震相拾取),而在涉及地震物理机制的地震成核过程、断层动力学、地震预测等核心科学问题中缺乏系统性突破。现有模型对地震波传播的非线性特征、多尺度时空关联及多物理场耦合机制缺乏显式建模,导致其在复杂地质环境下的泛化能力受限;物理规律与数据驱动方法的融合仍处于浅层耦合阶段,深度学习算法在处理地震波传播的弹性波方程约束时,尚未实现物理模型与神经网络参数的有效协同优化(Mousavi and Beroza, 2023);科学发现能力的不足还体现在模型可解释性缺失,现有方法多聚焦于表层特征提取,而难以揭示地震活动的深层次动力学规律。

未来需推动跨学科协作机制建设,建立地震学专家与 AI 研究者的联合实验室,共同设计符合物理规律的损失函数与网络架构,为模型训练提供物理约束的显式监督信号,从而有效缩小 AI 方法与地震科学本质规律的距离,推动深度学习从工具性应用向机理发现型研究转型。

4 结论与展望

笔者系统回顾了深度学习在地震事件检测、震

相拾取、地震分类、P 波初动极性判别震源机制反演等任务中的应用。人工智能技术正快速渗透至地震学各个分支方向,但其发展仍面临数据集规模小、制作规范不统一、标注质量良莠不齐、网络架构缺乏针对性、模型缺乏迭代更新、模型评测体系尚未建立,以及人工智能在地震科学发现中赋能不足等诸多瓶颈。

未来的发展路径应首先着眼于解决数据集分散与标注难题。因此有必要在 SeisBench 等现有平台的基础上,推动建立统一的数据共享协议与接口,整合全球数据资源。同时,连续波形数据为自监督学习提供了理想场景,利用对比学习或掩码信号重建等方法,从海量未标注数据中提取有效特征,有望大幅降低对人工标注的依赖。

其次,需突破模型架构的创新瓶颈。目前地震学领域的模型多借鉴自计算机视觉或自然语言处理,缺乏针对地震波物理特性的深度定制化模型。未来的研究重点应转向发展物理规律驱动的深度学习框架,将弹性波方程等物理规律作为约束项融入神经网络的训练过程,或设计能够融合地震学先验知识的混合模型架构。这不仅能提升模型在复杂地质环境下的泛化能力,也是增强其物理可解释性的关键一步。

再次,建立统一的评测标准。由于缺乏公认的基准数据集和评估体系,目前很多模型的指标存在虚高的问题,不同模型之间的性能难以进行公平比较。因此,应依托 PrismaX 等新兴评测平台,围绕地震检测、地震预警、地震预测等领域的关键技术进行评测,构建涵盖精度、召回率、鲁棒性及不确定性量化等多维度的综合评估指标体系。通过设立持续更新的公开排行榜,为算法迭代提供明确的优化目标,并促进整个社区的良性竞争与合作。

最后,推动人工智能从“数据处理工具”向“科学发现引擎”的转型,是人工智能地震学领域的终极目标。当前应用多集中于重复性任务自动化,而在地震成核机理、断层动力学等核心科学问题的探索上贡献有限。这需要通过建立跨学科协同机制,整合地震学专家、数据科学家与高性能计算工程师的集体智慧,共同设计能够揭示深层物理规律的新范式。通过上述努力,深度学习将不仅是提升数据处理效率的工具,更有望成为驱动地震科学理论突破、赋能新一代地震预警与风险评估体系的关键力量,为全球防震减灾事业提供更为可靠、透明且可持续的技术支撑。

参考文献:

- 安艳茹, 张莹莹, 苗春兰等, 2023. 中国测震站网完备数据集 (CSNCD) [EB/OL]. <https://data.earthquake.cn>.
- An Y R, Zhang Y Y., Miao C L, et al., 2023. CSNCD: A Comprehensive Dataset of Chinese Seismic Network [EB/OL]. <https://data.earthquake.cn>. (In Chinese).
- 蔡育琦, 邵博, 于子叶, 等, 2025. 基于 PRIME-DP 预训练模型的滑坡微地震等非天然地震分类研究 [J]. 华南地震, 45(1): 12—18.
- Cai Y Q, Shao B, Yu Z Y, et al., 2025. Classification of Non-Natural seismic events with landslide microseismicity as representative based on PRIME-DP Pre-Training model [J]. South China Journal of Seismology, 45(1): 12—18. (In Chinese)
- 蒋策, 2025. 区域台网地震数据智能处理方法研究与应用 [D]. 北京: 中国地震局地球物理研究所.
- Jiang C. 2025. Research and application of intelligent processing methods for regional seismic network data [D]. Beijing: Institute of Geophysics, China Earthquake Administration. (In Chinese).
- 林学楷, 许才军, 2022. 深度学习驱动的地震目录构建: PhaseNet 和 EqT 模型的对比与评估 [J]. 武汉大学学报 (信息科学版), 47(6): 855—865.
- Lin X K, Xu C J. 2022. Deep-Learning-Empowered earthquake catalog building: comparison and evaluation of PhaseNet and EqT Models [J]. Geomatics and Information Science of Wuhan University, 47(6): 855—865. (In Chinese)
- 廖诗荣, 张红才, 范莉苹, 等, 2021. 实时智能地震处理系统研发及其在 2021 年云南漾濞 Ms6.4 地震中的应用 [J]. 地球物理学报, 64(10): 3632—3645.
- Liao S R, Zhang H C, Fan L P, et al., 2021. Development of a real-time intelligent seismic processing system and its application to the 2021 Yangbi Ms6.4 earthquake [J]. Chinese Journal of Geophysics, 64(10): 3632—3645. (In Chinese)
- 任义, 姚志祥, 吴庆举, 等, 2025. 利用 P 波初动振幅反演小微地震的震源机制 [J]. 地球物理学报, 68(5): 1808—1825.
- Ren Y, Yao Z X, Wu Q J, et al. 2025. Focal mechanism inversion of small and micro earthquakes using P-wave initial amplitude [J]. Chinese Journal of Geophysics, 68(5): 1808—1825. (In Chinese)
- 田宵, 汪明军, 张雄, 等, 2022. 基于多输入卷积神经网络的天然地震和爆破事件识别 [J]. 地球物理学报, 65(5): 1802—1812.
- Tian X, Wang M J, Zhang X, et al. 2022. Discrimination of natural earthquakes and explosion events based on multi-input convolutional neural network [J]. Chinese Journal of Geophysics, 65(5): 1802—1812. (In Chinese)
- 隗永刚, 杨千里, 王婷婷, 等, 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(6): 646—657.
- Wei Y G, Yang Q L, Wang T T, et al. 2019. Discrimination of earthquakes and explosions based on deep learning residual network model [J]. Acta Seismologica Sinica, 41(6): 646—657. (In Chinese)
- 许忠淮, 刘玉芬, 张郢珍, 1979. 京、津、唐、张地区地震应力场的方向特征 [J]. 地震学报, 1(2): 122—132.
- Xu Z H, Liu Y F, Zhang Y Z. 1979. The characteristics of the direction of seismic stress field in the Beijing-Tianjin-Tangshan-Zhangjiakou region [J]. Acta Seismologica Sinica, 1(2): 122—132. (In Chinese)
- 许忠淮, 阎明, 赵仲和, 1983. 由多个小地震推断的华北地区构造应力场的方向 [J]. 地震学报, 5(3): 268—279.
- Xu Z H, Yan M, Zhao Z H. 1983. The direction of tectonic stress field in North China inferred from a number of small earthquakes [J]. Acta Seismologica Sinica, 5(3): 268—279. (In Chinese)
- 易继东, 张敏, 李振春, 等, 2023. 深度学习地震数据重建方法研究综述 [J]. 地球物理学进展, 38(2): 361—381.
- Yi J D, Zhang M, Li Z C, et al. 2023. A review of seismic data reconstruction methods based on deep learning [J]. Progress in Geophysics, 38(2): 361—381. (In Chinese)
- 俞春泉, 陶开, 崔效锋, 等, 2009. 用格点尝试法求解 P 波初动震源机制解及解的质量评价 [J]. 地球物理学报, 52(5): 1402—1411.
- Yu C Q, Tao K, Cui X F, et al. 2009. A grid search method for focal mechanism solution and its evaluation using P wave first motion [J]. Chinese Journal of Geophysics, 52(5): 1402—1411. (In Chinese)
- 赵明, 2025. 谛听十亿级参数地震波大模型泛化能力测评 [EB/OL]. 第六届人工智能地震学研讨会, 2025-07-20. <https://www.koushare.com/live/details/44536?vid=182313>.
- Zhao M. 2025. Evaluation of the generalization ability of the Di Ting Billion-Parameter Seismic wave large model [EB/OL]. The 6th Artificial Intelligence Seismology Workshop, 2025-07-20. <https://www.koushare.com/live/details/44536?vid=182313>. (In Chinese)
- 周少辉, 蒋长胜, 曲均浩, 等, 2021. 基于深度学习的地震事件分类识别——以山东地震台网记录为例 [J]. 地震地质, 43(3): 663—676.
- Zhou S H, Jiang C S, Qu J H, et al., 2021. Classification of seismic events based on deep learning: A case study of Shandong Seismic Network records [J]. Seismology and Geology, 43(3): 663—676. (In Chinese)
- Aydin O, Ardiç S, Kilar H, et al., 2025. Modelling the seismic activity of Kahramanmaraş, Türkiye with recurrent neural network (RNN) and long short-term memory (LSTM)

- methods[J]. *Natural Hazards*, 1(121): 18361—18390.
- Baevski A, Zhou H, Mohamed A, et al., 2020. wav2vec 2.0: A Framework for Self-Supervised Learning of Speech Representations[J]. *Advances in Neural Information Processing Systems*, 33(1): 12449—12460.
- Bank D, Koenigstein N and Giryas R, 2020. Autoencoders[M]. *Machine learning*. 193—208.
- Bergen K J, Johnson P A, de Hoop M V, et al., 2019. Machine learning for data-driven discovery in solid Earth geoscience[J]. *Science*, 363(6433): eaau0323.
- Bi X, Chen D, Chen G, et al., 2024. DeepSeek LLM: Scaling Open-Source Language Models with Longtermism[J]. *arXiv preprint arXiv: 2401.02954*
- Borate P, Rivière J, Marone C, et al., 2023. Using a physics-informed neural network and fault zone acoustic monitoring to predict lab earthquakes[J]. *Nature Communications*, 14(1): 3693.
- Brown T B, Mann B, Ryder N, et al. 2020. Language Models are Few-Shot Learners[J]. *Advances in Neural Information Processing Systems*, 33(1): 1877—1901.
- Chakraborty M, Cartaya C Q, Li W, et al., 2022. PolarCAP – A deep learning approach for first motion polarity classification of earthquake waveforms[J]. *Artificial Intelligence in Geosciences*, 3(1): 46—52.
- Chen Y K, Saad O M, Savvaidis A, et al, 2024 a. Deep Learning for P-Wave First-Motion Polarity Determination and Its Application in Focal Mechanism Inversion[J]. *IEEE Trans Geosci Remote Sensing*, 62(1): 1—11.
- Chen Y K, Savvaidis A, Saad O M, et al., 2024 b. TXED: The Texas Earthquake Dataset for AI[J]. *Seismological Research Letters*, 95(3): 2013—2022.
- Cheng Y F, Richard M A and Taka'aki T, 2023. A New Focal Mechanism Calculation Algorithm (REFOC) Using Inter-Event Relative Radiation Patterns: Application to the Earthquakes in the Parkfield Area[J]. *Journal of Geophysical Research: Solid Earth*, 128(3): e2022JB025006.
- Chi-Durán R, Dreger D S, Rodgers A J, et al., 2021. Joint Regional Waveform, First-Motion Polarity, and Surface Displacement Moment Tensor Inversion of the 3 September 2017 North Korean Nuclear Test[J]. *The Seismic Record*, 1(2): 107—116.
- Cianetti S, Bruni R, Gaviano S, et al., 2021. Comparison of Deep Learning Techniques for the Investigation of a Seismic Sequence: An Application to the 2019, Mw 4.5 Mugello (Italy) Earthquake[J]. *JGR Solid Earth*, 126(12): e2021JB023405.
- Cui X, Li Z F, Hu Y, 2023. Similar seismic moment release process for shallow and deep earthquakes[J]. *Nature Geoscience*, 16(5): 454—460.
- Devlin J, Chang M W, Lee K, et al., 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding[C]. *Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies*, 1(1): 4171—4186.
- Dowla F, Taylor S, Anderson R, 1990. Seismic discrimination with artificial neural networks: Preliminary results with regional spectral data[J]. *Bulletin of the Seismological Society of America*, 80(5): 1346—1373.
- Fang L H, Wu Z L and Song K, 2017. SeismOlympics[J]. *Seismological Research Letters*, 88(6): 1429—1430.
- Feng T, Mohanna S, Meng L, 2022. EdgePhase: A Deep Learning Model for Multi-Station Seismic Phase Picking[J]. *Geochem Geophys Geosyst*, 23(11): e2022GC010453.
- Goodfellow I J, Pouget-Abadie J, Mirza M, et al., 2014. Generative Adversarial Networks[J]. *Communications of the ACM*, 63(11): 139—144.
- Grattafiori A, Dubey A, Jauhri A, et al., 2024. The LLaMA 3 Herd of Models. *arXiv: 2407.21783*.
- Guzhov A, Raue F, Hees J, et al., 2021. AudioCLIP: Extending CLIP to Image, Text and Audio[C]. *CASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing*, 976—980.
- Han J, Kim S, Sheen D-H. 2025. RPNNet: Robust P-Wave First-Motion Polarity Determination Using Deep Learning[J]. *Seismological Research Letters*, 96(4): 2405—2417.
- Hara S, Fukahata Y, Iio Y. 2019. P-wave first-motion polarity determination of waveform data in western Japan using deep learning[J]. *Earth, Planets and Space*, 71(1): 117—127.
- Hardebeck J L. 2002. A New Method for Determining First-Motion Focal Mechanisms[J]. *Bulletin of the Seismological Society of America*, 92(6): 2264—2276.
- Hardebeck J L. 2003. Using S/P Amplitude Ratios to Constrain the Focal Mechanisms of Small Earthquakes[J]. *Bulletin of the Seismological Society of America*, 93(6): 2434—2444.
- Hardebeck J L, Shearer P M. 2008. HASH: A FORTRAN Program for Computing Earthquake First-Motion Focal Mechanisms – v1.2 – January 31, 2008[OL]. U. S. Geological Survey Open-File Report.
- Hochreiter S, Schmidhuber J. 1997. Long Short-Term Memory[J]. *Neural Computation*, 9(8): 1735—1780.
- Hong M Q, Zhang H C, Wu L H, et al. 2025. Using Deep Convolutional Neural Networks for Earthquake and Explosion Classification[J]. *IEEE Access*, 13(1): 56144—56159.
- Hourcade C, Bonnin M, Beucler É. 2023. New CNN-based tool to discriminate anthropogenic from natural low magni-

- tude seismic events[J]. *Geophysical Journal International*, 232(3): 2119—2132.
- Hsu W-N, Bolte B, Tsai Y-H H, et al. 2021. HuBERT: Self-Supervised Speech Representation Learning by Masked Prediction of Hidden Units[OL]. arXiv preprint arXiv: 2106.07447.
- Jiang C, Fang L H, Fan L P, et al. 2021. Comparison of the earthquake detection abilities of PhaseNet and EQTransformer with the Yangbi and Maduo earthquakes[J]. *Earthquake Science*, 34(5): 425—435.
- Jiang J S, Ren H R, Zhang M. 2022. A Convolutional Autoencoder Method for Simultaneous Seismic Data Reconstruction and Denoising[J]. *IEEE Geoscience and Remote Sensing Letters*, 19(1): 1—5.
- Kaneko R, Nagao H, Ito S, et al. 2023. Detection of Deep Low-Frequency Tremors From Continuous Paper Records at a Station in Southwest Japan About 50 Years Ago Based on Convolutional Neural Network[J]. *Journal of Geophysical Research: Solid Earth*, 128(3): e2022JB024842.
- Kong Q K, Wang R J, Walter W R, et al. 2022. Combining Deep Learning With Physics Based Features in Explosion-Earthquake Discrimination[J]. *Geophysical Research Letters*, 49(3): e2021GL096577.
- Kuang W H, Yuan C C, Zhang J, 2021. Real-time determination of earthquake focal mechanism via deep learning[J]. *Nature Communications*, 12(1): 1432
- LeCun Y, Boser B, Denker J S, et al. 1989. Backpropagation Applied to Handwritten Zip Code Recognition[J]. *Neural Computation*, 1(4): 541—551.
- LeCun Y, Bottou L, Bengio Y, et al. 1998. Gradient-based learning applied to document recognition[J]. *Proceedings of the IEEE*, 86(11): 2278—2324.
- LeCun Y, Bengio Y, Hinton G. 2015. Deep learning[J]. *Nature*, 521(7553): 436—444.
- Li B R, Fan L P, Jiang C, et al., 2023. CSESnet: A deep learning P-wave detection model based on UNet++ designed for China Seismic Experimental Site[J]. *Frontiers in Earth Science*, 10: 1032839.
- Lin T Y, Maire M, Belongie S, et al., 2014. Microsoft COCO: Common Objects in Context[M]. Springer International Publishing.
- Li S, Fang L H, Xiao Z W, et al., 2023. 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*, 13(4): 2233—2240.
- Li S, Yang X, Cao A Y, et al., 2024. SeisT: A Foundational Deep-Learning Model for Earthquake Monitoring Tasks[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 62(1): 1—15.
- Linville L, Pankow K, Draelos T, 2019. Deep Learning Models Augment Analyst Decisions for Event Discrimination[J]. *Geophysical Research Letters*, 46(7): 3643—3651.
- Liu T, Münchmeyer J, Laurenti L, et al., 2024. SeisLM: a Foundation Model for Seismic Waveforms[J]. arXiv preprint arXiv: 2410.15765
- Liu X L, Ren T, Chen H F, et al., 2020. Classification of tectonic and non-tectonic seismicity based on convolutional neural network[J]. *Geophysical Journal International*, 224(1): 191—198.
- Lu T R, Zhou L Q, Zhang N, et al., 2025. EQTypeNet: deep learning tri-branch earthquake automatic classification model and its application in China[J]. *Geophysical Journal International*, 242(1): 242—259.
- Maguire R, Schmandt B, Wang R F, et al., 2024. Generalization of Deep-Learning Models for Classification of Local Distance Earthquakes and Explosions across Various Geologic Settings[J]. *Seismological Research Letters*, 95(4): 2229—2238.
- Miao F J, Carpenter N S, Wang Z, et al., 2020. High-Accuracy Discrimination of Blasts and Earthquakes Using Neural Networks With Multiwindow Spectral Data[J]. *Seismological Research Letters*, 91(3): 1646—1659.
- Michellini A, Cianetti S, Gaviano S, et al., 2021. INSTANCE – the Italian seismic dataset for machine learning[J]. *Earth System Science Data*, 13(11): 5509—5544.
- Mousavi S M, Horton S P, Langston C A, et al., 2016. Seismic features and automatic discrimination of deep and shallow induced-microearthquakes using neural network and logistic regression[J]. *Geophysical Journal International*, 207(1): 29—46.
- Mousavi S M, Sheng Y, Zhu W Q, et al., 2019. STanford EArthquake Dataset (STEAD): A Global Data Set of Seismic Signals for AI[J]. *IEEE Access*, 7: 179464—179476.
- Mousavi S M, Ellsworth W L, Zhu W Q, et al., 2020. Earthquake transformer—an attentive deep-learning model for simultaneous earthquake detection and phase picking[J]. *Nature Communications*, 11(1): 3952.
- Münchmeyer J, Woollam J, Rietbrock A, et al., 2022. Which Picker Fits My Data? A Quantitative Evaluation of Deep Learning Based Seismic Pickers[J]. *Journal of Geophysical Research: Solid Earth*, 127(3): e2021JB023499.
- Nakamura M, 2004. Automatic determination of focal mechanism solutions using initial motion polarities of P and S waves[J]. *Physics of the Earth and Planetary Interiors*, 146(3-4): 531—549.

- Okada Y, Kasahara K, Hori S, et al., 2014. Recent progress of seismic observation networks in Japan —Hi-net, F-net, K-NET and KiK-net—[J]. *Earth, Planets and Space*, 56: xv-xxviii.
- Oliveira D A B, Ferreira R S, Silva R, et al., 2018. Interpolating Seismic Data With Conditional Generative Adversarial Networks[J]. *IEEE Geoscience and Remote Sensing Letters*, 15(12): 1952—1956.
- Radford A, Narasimhan K. 2018. Improving Language Understanding by Generative Pre-Training[OL]. OpenAI blog.
- Radford A, Wu J, Child R, et al., 2019. Language Models are Unsupervised Multitask Learners [OL]. OpenAI blog.
- Radford A, Kim J W, Hallacy C, et al., 2021. Learning Transferable Visual Models From Natural Language Supervision[C]. *Proceedings of the 38th International Conference on Machine Learning*, PMLR 139: 8748—8763.
- Ronneberger O, Fischer P, Brox T, 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation[C]. Navab N, Hornegger J, Wells W M, et al. *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. *Lecture Notes in Computer Science*, vol 9351. Springer, Cham: 234—241.
- Ross Z E, Meier M, Hauksson E, et al., 2018a. Generalized Seismic Phase Detection with Deep Learning[J]. *Bulletin of the Seismological Society of America*, 108(5A): 2894—2901.
- Ross Z E, Meier M, Hauksson E, 2018b. P Wave Arrival Picking and First-Motion Polarity Determination With Deep Learning[J]. *Journal of Geophysical Research: Solid Earth*, 123(11): 5120—5129.
- Ross Z E, Cochran E S, Trugman D T, et al., 2020. 3D fault architecture controls the dynamism of earthquake swarms[J]. *Science*, 368(6497): 1357—1361.
- Russakovsky O, Deng J, Su H, et al., 2015. ImageNet Large Scale Visual Recognition Challenge[J]. *International Journal of Computer Vision*, 115(3): 211—252.
- Mousavi S M, Beroza G C, 2023. Machine Learning in Earthquake Seismology[J]. *Annual Review of Earth and Planetary Sciences*, 51: 391—418.
- Scarselli F, Gori M, Tsoi A C, et al., 2009. The Graph Neural Network Model[J]. *IEEE Transactions on Neural Networks*, 20(1): 61—80.
- Si X, Wu X M, Sheng H L, et al., 2024. SeisCLIP: A Seismology Foundation Model Pre-Trained by Multimodal Data for Multipurpose Seismic Feature Extraction[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 62: 1—13.
- Skoumal R J, Shelly D R, Hardebeck J L, 2022. Using Machine Learning Techniques with Incomplete Polarity Datasets to Improve Earthquake Focal Mechanism Determination[J]. *Seismological Research Letters*, 94(1): 294—304.
- Song X H, Men-Andrin M, Ellsworth W L, et al., 2025. FocoNet: transformer-based focal-mechanism determination[J]. Authorea Preprints.
- Sun H Y, Ross Z E, Zhu W Q, et al., 2023. Phase Neural Operator for Multi-Station Picking of Seismic Arrivals[J]. *Geophysical Research Letters*, 50(15): e2023GL106434.
- Tian X, Zhang W, Zhang X, et al., 2020. Comparison of Single-Trace and Multiple-Trace Polarity Determination for Surface Microseismic Data Using Deep Learning[J]. *Seismological Research Letters*, 91(3): 1794—1803.
- Uchide T, 2020. Focal mechanisms of small earthquakes beneath the Japanese islands based on first-motion polarities picked using deep learning[J]. *Geophysical Journal International*, 223(3): 1658—1671.
- Vaswani A, Shazeer N, Parmar N, et al., 2017. Attention Is All You Need[C]. *Advances in Neural Information Processing Systems 30 (NIPS 2017)*.
- Van den Ende M P A, Ampuero J P, 2020. Automated Seismic Source Characterization Using Deep Graph Neural Networks[J]. *Geophysical Research Letters*, 47(XX): e2020GL088690.
- Wang J, Xiao Z W, Liu C, et al., 2019. Deep Learning for Picking Seismic Arrival Times[J]. *Journal of Geophysical Research: Solid Earth*, 124(7): 6612—6624.
- Wang P Y, Ren T, Shen R, et al., 2024. Determination of earthquake focal mechanism via multi-task learning[J]. *Computers & Geosciences*, 184(1): 105513.
- Wang X, Liu F, Su R, et al., 2025. SeisMoLLM: Advancing Seismic Monitoring via Cross-modal Transfer with Pre-trained Large Language Model[J]. arXiv preprint arXiv: 2502.19960
- Wilding J D, Zhu W Q, Ross Z E, et al., 2023. The magmatic web beneath Hawai'i[J]. *Science*, 379(6633): 462—468.
- Woollam J, Münchmeyer J, Tilmann F, et al., 2022. SeisBench—A Toolbox for Machine Learning in Seismology[J]. *Seismological Research Letters*, 93(3): 1695—1709.
- Wu X M, Ma J W, Si X, et al., 2023. Sensing prior constraints in deep neural networks for solving exploration geophysical problems[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 120(23): e2219573120.
- Xiao Z W, Wang J, Liu C, et al., 2021. 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*, 126(6): e2020JB021444.
- Yu Z Y, Wang W T, 2022. LPPN: A Lightweight Network for

- Fast Phase Picking[J]. *Seismological Research Letters*, 93(5): 2834—2846.
- Yu Z Y, Wang W T, Chen Y N, 2023. Benchmark on the accuracy and efficiency of several neural network based phase pickers using datasets from China Seismic Network[J]. *Earthquake Science*, 36(2): 113—131.
- Zaremba W, Sutskever I, Vinyals O, 2014. Recurrent Neural Network Regularization[OL]. arXiv preprint arXiv: 1409.2329.
- Zhai Q S, 2022. Understanding microseismicity behavior and their response to earth processes by improving earthquake catalogs[D]. Georgia Institute of Technology.
- Zhang J, Li Z F, Zhang J, 2023. Simultaneous Seismic Phase Picking and Polarity Determination with an Attention-Based Neural Network[J]. *Seismological Research Letters*, 94(2): 813—828.
- Zhao C, Deng C, Ruan C, et al., 2025. Insights into DeepSeek-V3: Scaling Challenges and Reflections on Hardware for AI Architectures[C]. *Proceedings of the 52nd Annual International Symposium on Computer Architecture*. 1731—1745.
- Zhao M, Xiao Z W, Zhang M, et al., 2023a. DiTingMotion: A deep-learning first-motion-polarity classifier and its application to focal mechanism inversion[J]. *Frontiers in Earth Science*, 11(1): 1103914.
- Zhao M, Xiao Z W, Chen S, et al., 2023b. DiTing: A large-scale Chinese seismic benchmark dataset for artificial intelligence in seismology[J]. *Earthquake Science*, 36(1): 84—94.
- Zhou Y J, Yue H, Kong Q K, et al., 2019. Hybrid Event Detection and Phase-Picking Algorithm Using Convolutional and Recurrent Neural Networks[J]. *Seismological Research Letters*, 90(3): 1079—1087.
- Zhou Y J, Yue H, Fang L H, et al., 2022. An Earthquake Detection and Location Architecture for Continuous Seismograms: Phase Picking, Association, Location, and Matched Filter (PALM)[J]. *Seismological Research Letters*, 93(1): 413—425.
- Zhu J, Fang L H, Miao F J, et al., 2023. Deep learning and transfer learning of earthquake and quarry-blast discrimination: applications to southern California and eastern Kentucky[J]. *Geophysical Journal International*, 236(2): 979—993.
- Zhu W Q, Beroza G C, 2018. PhaseNet: A Deep-Neural-Network-Based Seismic Arrival Time Picking Method[J]. *Geophysical Journal International*, 217(2): 1386—1393.
- Zhu W Q, Wang H, Rong B, et al., 2025a. California Earthquake Dataset for Machine Learning and Cloud Computing[J]. arXiv preprint arXiv: 2502.11500
- Zhu W Q, Song J H, Wang H, et al., 2025b. Towards End-to-End Earthquake Monitoring Using a Multitask Deep Learning Model[J]. arXiv preprint arXiv: 2506.06939
- Zor E, Peng Z G, Ergin M, et al., 2025. Dense Seismic Recordings of the 2023 Kahramanmaraş Earthquake Sequence in Southeastern Türkiye[J]. *Seismological Research Letters*, 96(6): 2683—2697.