

基于元学习Transformer-BiLSTM的测井曲线重构

何昌龙^{1*}, 张琳智^{1#}, 李亚兰², 房平超¹, 应文峰¹, 刘庆松¹, 毛子原¹

¹中国石油西南油气田公司勘探事业部, 四川 成都

²四川吉诺恩油气工程技术有限公司, 四川 成都

收稿日期: 2026年7月10日; 录用日期: 2026年8月24日; 发布日期: 2026年9月2日

摘要

随着页岩气资源勘探和开发工作的不断深入, 测井数据在储层评价中所起的作用也越来越大。但是由于设备故障、作业成本等各方面的原因, 在实际的测井过程中常常会出现曲线缺失或者异常的情况, 从而影响到地质解释的精度以及资源开发的决策。本文设计出一种将元学习、Transformer和双向长短期记忆神经网络(BiLSTM)结合在一起的深度学习模型, 简称MTB。元学习模块依靠跨任务元训练机制, 使得模型可以从很少的数据样本中迅速学会测井数据的内在规律, 进而明显改善小样本情况下的重构性能和泛化水平; Transformer编码器可以捕捉到测井序列里长距离的依赖关系, 加强了全局特征的表达能力; BiLSTM会继续对前后向时序特征加以提取, 从而加强对于测井曲线动态改变的建模精确度。三者共同作用可以使得模型具有很好的学习和重构能力, 提高测井曲线的重构精度和稳定性。在四川绵阳北部地区6口井的实测数据集中进行单目标测井曲线重构实验, 结果表明MTB模型在RMSE、MAE、MAPE、R²等各方面都比现有的主流模型有更高的预测精度和更好的泛化能力。

关键词

测井曲线重构, 元学习, 双向长短期记忆神经网络, Transformer编码器, 深度学习

Well-Log Reconstruction Based on Meta-Learning with Transformer-BiLSTM

Changlong He^{1*}, Linzhi Zhang^{1#}, Yalan Li², Pingchao Fang¹, Wenfeng Ying¹, Qingsong Liu¹, Ziyuan Mao¹

¹Exploration Division, PetroChina Southwest Oil and Gasfield Company, Chengdu Sichuan

²Sichuan Ginuo Oil & Gas Engineering Technology Service Co., Ltd., Chengdu Sichuan

*第一作者。

#通讯作者。

文章引用: 何昌龙, 张琳智, 李亚兰, 房平超, 应文峰, 刘庆松, 毛子原. 基于元学习 Transformer-BiLSTM 的测井曲线重构[J]. 石油天然气学报, 2026, 48(3): 380-392. DOI: 10.12677/jogt.2026.483043

Abstract

With the continuous development of shale gas exploration and production, more and more data from well logs are being used to assess the reservoirs. Due to various reasons such as equipment failure, operating costs, etc., there are often missing or irregular curves in the actual logging operation; as a result, the accuracy of geological interpretation and the reliability of resource development plans are greatly affected. To address the problem of missing or abnormal well logs, a new kind of deep learning model called MTB has been put forward in this paper, which combines meta-learning, Transformer and bidirectional long short-term memory (BiLSTM). A meta-learning module is a cross-task meta-training method that can learn the general characteristics of well-logging data from a small number of samples very quickly to improve reconstruction performance and generalisation ability in a few-shot scenario. The Transformer encoder can collect information from the far-away parts of the well log sequence. BiLSTM is used to obtain the forward and reverse temporal information, and then changes in the well log over time can be modelled. The three components together can achieve good learning and reconstruction by the model to improve the accuracy and stability of well log curve reconstruction. Experiments on the single-target well log reconstruction were carried out based on the actual data set of the six wells in the northern Mianyang area of the Sichuan Basin. From the above results, it can be seen that the MTB model is better than other typical models in terms of RMSE, MAE, MAPE and R^2 , and has relatively high prediction accuracy and generalisation ability.

Keywords

Well Log Reconstruction, Meta-Learning, Bidirectional Long Short-Term Memory (BiLSTM), Transformer Encoder, Deep Learning

Copyright © 2026 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

非常规油气资源开发速度不断加快, 页岩气在我国能源结构中所占的比重越来越大。绵阳北部位于四川盆地重点勘探区, 具有资源基础好、埋藏条件适宜商业开发的客观条件[1]。该区域地质构造比较复杂, 在生产过程中受到钻井作业成本、测井仪器性能和复杂井筒环境的限制, 数据丢失或者采样稀疏的现象比较普遍[2] [3]。直接影响到地层解释的可靠性, 也给后续储层评价及开发方案的制订带来不确定性, 因此高精度测井曲线重构的研究意义十分突出。页岩气储层一般都存在较强的非均质性、多尺度结构特点以及沿着深度方向复杂的序列联系, 传统的依靠地质经验或者简单插值的方法, 在表现地层内部非线性响应时存在着明显的缺陷[4]。测井数据一直被当作油气勘探开发过程中形成地下模型、评判资源潜力的基本依据。利用各种地球物理测井手段综合起来使用, 不但可以提高储层物性参数预测的准确性, 还可以提高地质解释的可信度, 为建立精细的地下模型、保证页岩气资源高效动用提供数据支持。现场作业时, 由于经济成本限制、仪器测量误差、地质条件不确定等各方面原因, 测井曲线不完整或者数据异常的记录较多[5] [6]。怎样有效地恢复缺失的测井信号, 已经成为测井解释和储层参数评价领域急需解决的技术难题。由于测井曲线具有非线性特点, 单纯用传统的测量方法不能很好地反映测井曲线之间复

杂的响应关系。近些年来,有关的研究也开始使用机器学习、深度神经网络这些手段来探究更为具有适应性与推广性的建模方法[2] [4]。

Park 等人[7]对多种基于机器学习的测井数据重构算法性能进行了比较分析,结果表明集成学习方法对于缺失层速度数据的预测任务来说准确度比较高。Haritha 等[5]提出了一种结合卷积神经网络和双向长短期记忆网络的深度学习模型(CNN-Bi-LSTM),在 Krishna Godavari 盆地两个站点上进行实验验证,得到较好的预测结果,均方误差低。Kang 等人[8]为了解决深层碳酸盐岩微相识别难的问题,提出了一种残差长短期记忆网络(ResLSTM)智能预测模型,该模型对深层碳酸盐岩微相预测的准确率比传统的 LSTM 高很多,在薄互层识别上也有较高的准确率。张俊华等人[9]使用门控循环单元(GRU)结合分位数回归法,将均方根误差降低了很多,在储层孔隙度预测上比 BP 神经网络更优。郭滨等[10]提出了一种将多模块结合在一起的深度学习测井曲线重构方法,对复杂储层的特征表征有较好的应用前景。曹志民等人[11]提出了一个 CCD-WLR (CNN-Bi-LSTM 和 CNN-Transformer 两个通道的缺失测井重构)框架,用小波包分解增强可以很好地捕捉到测井数据的多尺度特征以及长程时序依赖。邹春亮等[12]提出了 meta-ViT 方法,把元学习的参数更新机制和视觉 Transformer 的自注意力结构结合起来,用 Transformer 的自注意力机制捕捉传统 LSTM 不能处理的远距离依赖,解决了小样本钻井数据解释中模型适应性差、预测精度低的问题。双向长短期记忆网络在测井重构任务上也有较多的应用。以上研究为建立元学习、Transformer、BiLSTM 相结合的测井重构模型奠定了较好的基础。

本文以绵阳北部页岩气区块测井曲线高精度重建为研究对象,提出了一种名为 MTB 的深度网络模型,该模型将元学习、Transformer 编码器和双向长短期记忆网络三种技术结合在一起。在模型结构里,元学习模块赋予网络跨任务快速适应的本领,在不同任务和数据分布的环境下,模仿元学习的过程,从而大幅加强模型对于样本分布异质性以及小样本状况的稳健水平以及推广性能。Transformer 编码器用多头自注意力机制沿井深方向挖掘测井序列的全局依赖关系,可以很好地建模长距离特征。BiLSTM 模块从正反两个方面来捕捉测井曲线的时序联系,并且使得模型对于地层动态发展规律有更强的表现力。三者有机结合,可以全面理解测井数据复杂的结构,从全局上下文、双向时序特征、跨任务自适应性这三个方面来提高曲线重建的精度和稳定性。通过绵阳北部地区实际测井数据集的验证可知,本模型重建精度和稳定性比目前主流的方法要好,为页岩气储层精细描述、智能化测井数据补全提供了一种高效可靠的途径。

2. MTB 模型结构设计与原理分析

2.1. MTB 模型原理及结构

测井曲线重构遇到局部变化剧烈、跨尺度非线性复杂以及长距离依赖等诸多难题,而且实际勘探中标注样本数量较少,井间分布情况也不一致,小样本问题比较明显。为此设计了一个 MTB 深度网络模型,即先用元学习从多个任务中提取出可以迁移的先验知识,使得模型能够快速地适应新的井;再用 BiLSTM 来捕捉到短时的局部模式;最后用 Transformer 来建立全局依赖,弥补了 BiLSTM 长距离建模的不足。三者共同完成高质量的重建。

输入为二维矩阵,行深度采样点,列测井参数。序列先经过元学习模块,用跨任务元训练,内循环用少量支持集样本做几步梯度下降,使模型适应任务分布;外循环用查询集损失更新全局初始参数。内外交替优化,从多个标注任务中得到共性的先验知识,存入初始的权重。新井只需要少量的实测点完成几轮内循环更新就可以达到稳定的重构效果,缓解了小样本过拟合的问题。接着就是 BiLSTM 层,前向 LSTM 沿深度增加提取局部演化特征,后向 LSTM 反向扫描捕捉未来深度的影响。双向输出在每一个深度上都进行处理,即模型可以同时得到过去的信息和未来的信息。结构对测井来说很重要,因为响应常

常受上下邻层的共同影响。BiLSTM 的输出有很强的局部上下文，对短程波动、渐变层理都很敏感。接着就是 Transformer 编码器，先对 BiLSTM 的输出加上位置编码，区分深度顺序。经过多头自注意力之后，每一个深度点都可以和任意一个点进行直接的交互，以此建立全局依赖图，捕获几千米外相似岩性段之间的联系，克服了 BiLSTM 长距的不足。自注意力输出经过前馈网络来加强非线性。每层后面加上残差连接和层归一化，保证梯度可以跨层传播，稳定收敛。

在 BiLSTM 和 Transformer 之间加入直接的残差连接，把两者输出逐元素相加，实现多层融合。最后得到的特征是带有 Transformer 全局抽象但又保留了 BiLSTM 局部精细结构的一种特征。最后用线性全连接层把目标测井曲线映射到一维空间中，对每个预测点进行预测，得到完整的缺失段。MTB 模型把元学习、BiLSTM 和 Transformer 结合起来，分别发挥出小样本适应、局部时序刻画和全局依赖建模的优势，给复杂测井曲线高质量重构提供一条可行的技术路线。图 1 为结构示意图。

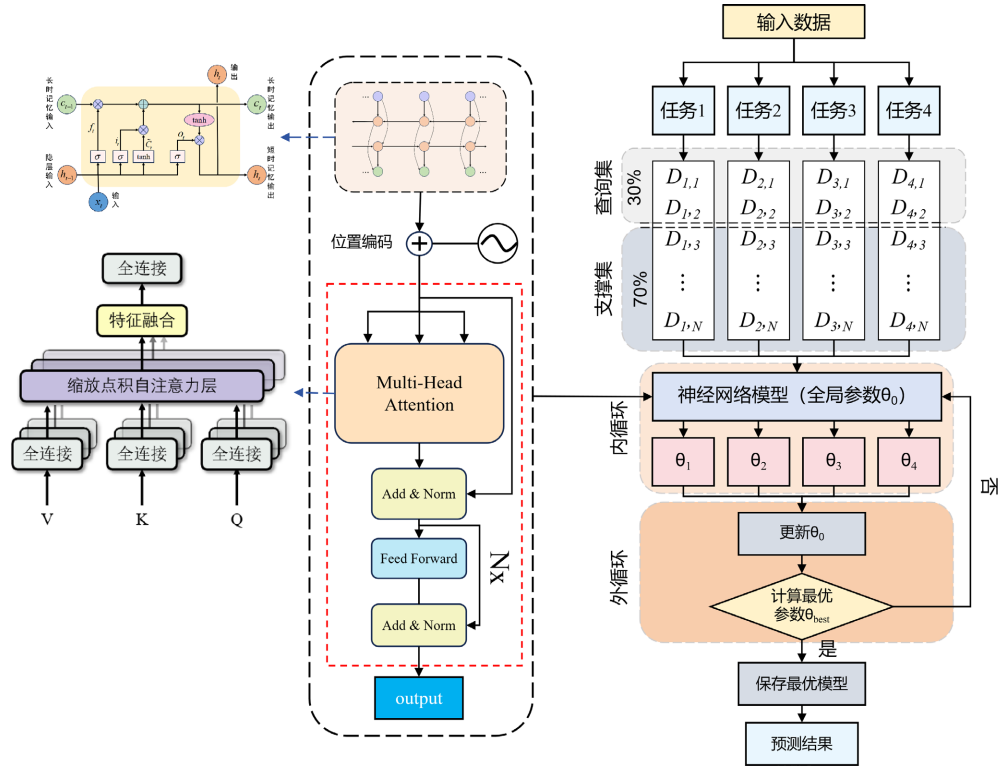


Figure 1. MTB model structure
图 1. MTB 模型结构

2.2. 关键模块原理

元学习就是从多个相关任务中学习通用的先验知识，使模型在小样本情况下也能很快适应新的任务。本文使用模型 Reptile 元学习算法，它的主要思想就是通过内循环和外循环的交替优化来得到一组对任务变化敏感的初始参数[13]。元学习的过程分为内循环和外循环两部分，在内循环里，模型在一个元任务的支持集中做若干步的梯度下降来适应该任务的数据分布，在外循环里，模型用所有任务内循环更新之后在查询集上的表现来优化初始参数，使它更容易通过少量的梯度更新达到较好的性能，其结构如图 2 所示。

设元任务分布为 $p(T)$ ，每个任务 T_i 包含支持集 D_i^{spr} 和查询集 D_i^{qry} 。模型参数初始化为 θ 。在内循环中，对每个任务 T_i ，模型在 D_i^{spr} 上计算损失 $L_{T_i}(f_\theta)$ ，并通过梯度下降更新参数：

$$\theta'_i = \theta - \alpha \nabla_{\theta} \mathcal{L}_{\mathcal{T}_i}(f_{\theta}) \quad (1)$$

其中 α 为内循环学习率, θ'_i 为任务 $\mathcal{T}_i$ 的快速适应参数。通常该更新步数较少(如 1~5 步), 以保证快速适应能力。

在外循环中, 模型利用所有任务在查询集上的损失来更新初始参数 θ :

$$\theta = \theta - \beta \nabla_{\theta} \sum_{\mathcal{T}_i \sim p(\mathcal{T})} \mathcal{L}_{\mathcal{T}_i}(f_{\theta'_i}) \quad (2)$$

其中 β 为外循环学习率, $f_{\theta'_i}$ 表示以内循环更新后的 θ'_i 为参数的模型。通过内、外循环的交替优化, 模型能够从多个相关任务中提炼出可快速迁移的共性知识, 并在面对新任务时仅需少量样本即可获得稳定可靠的重构结果, 从而有效缓解小样本场景下的过拟合风险。

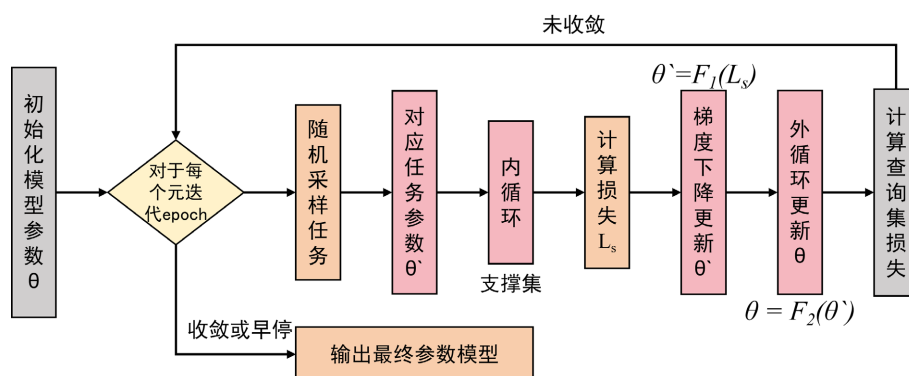


Figure 2. Meta-learning module structure

图 2. 元学习模块结构

长短期记忆神经网络是 Hochreiter 和 Schmidhuber 于 1997 年提出的[14], 它能够克服传统 RNN 对于长序列数据的梯度消失、梯度爆炸等问题。LSTM 用门控机制和特殊单元结构来解决这个问题, 从而使网络可以更好地学习长期依赖关系。LSTM 最核心的创新之处就是它的记忆单元, 它包含了一条可以长时间保存信息的“细胞状态”, 这条信息高速通道使得信息可以在多个时间步之间传递, 不会因为梯度传播得太快而消失。此外, LSTM 设计了三个关键的门结构, 即输入门(Input Gate)、遗忘门(Forget Gate)和输出门(Output Gate), 用于对信息流动进行控制其核心示意图如图 3 所示。

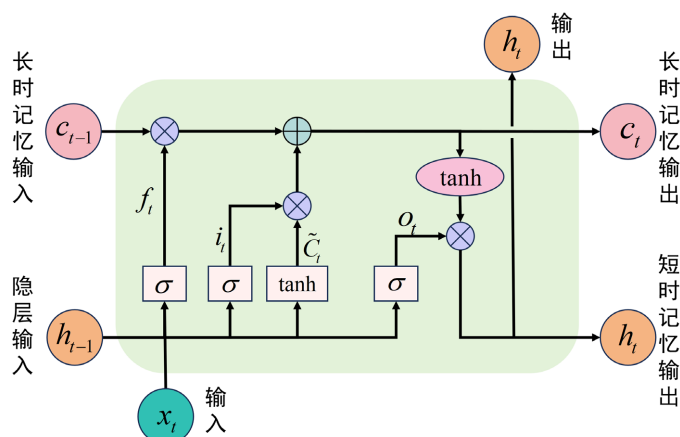


Figure 3. LSTM neural network state diagram

图 3. LSTM 神经网络状态图

输入门决定当前时间步的新信息应该存入细胞状态中的多少, 其计算式为

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i); \tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (3)$$

式中, i_t 为当前时间步信息的存储比例, W_i, b_i 为输入门的权重矩阵和偏置项, W_c, b_c 为候选信息的权重矩阵和偏置项, h_{t-1} 为上一时刻的隐藏状态, x_t 为当前时间步的输入, $\tilde{C}_t$ 为新的候选信息, $\tanh$ 为双曲正切函数。当 i_t 接近“1”时则代表新信息能较大程度的存储, i_t 接近“0”时, 则会舍去新信息。遗忘门决定哪些历史信息需要被丢弃, 其计算式如下:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (4)$$

式中, f_t 为每个单元应该保留的比例, W_f, b_f 为遗忘门的权重矩阵和偏置项, σ 为 Sigmoid 激活函数。当 f_t 接近“1”时, 表示该信息需要保留, 当 f_t 接近“0”时, 表示该信息需要遗忘。输出门则控制当前时间步的输出, 其计算式如下:

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o); h_t = o_t \odot \tanh(C_t) \quad (5)$$

式中, o_t 为输出门的输出比例, W_o, b_o 为输出门的权重矩阵和偏置项, h_t 为当前时间步的输出量, $\odot$ 为逐元素乘法。当 o_t 接近“1”时则表示输出信息, 当 o_t 接近“0”时, 则表示当前时间步不输出信息。

相比于传统的单向 LSTM 来说, Bi-LSTM 可以更好地利用数据的时序结构, 在学习长依赖关系的时候更有优势。双向传播机制能同时考虑历史信息和未来信息, Bi-LSTM 得到的特征表示比单向的 LSTM 更丰富, 可以提高序列建模的效果[15], 结构如图 4 所示。

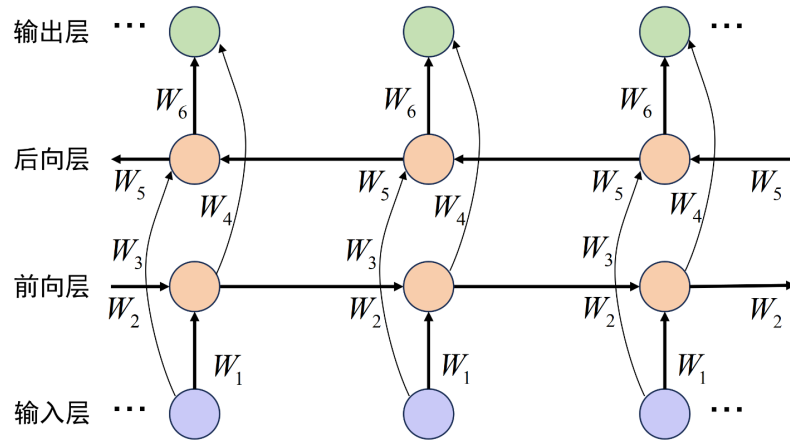


Figure 4. Bi-LSTM structure diagram

图 4. Bi-LSTM 结构图

Transformer 是一种基于自注意力机制的深度学习架构, 由 Vaswani 等(2017)提出[16]。Transformer 完全依靠自注意力机制, 克服了 RNN 在处理长序列时出现的梯度消失问题[17], 并且可以进行更高效的并行计算, 使训练速度更快、性能更好。

假设输入序列为 $X = [x_1, \dots, x_N] \in R^{D_x \times N}$, 输出序列为 $H = [h_1, \dots, h_N] \in R^{D_h \times N}$, 将 X 输入自注意力机制时, 先将其映射为 Q 、 K 、 V 三个向量矩阵, 其表达形式如下所示:

$$\begin{aligned} Q &= W_q X \in R^{D_q \times N} \\ K &= W_k X \in R^{D_k \times N} \\ V &= W_v X \in R^{D_v \times N} \end{aligned} \quad (6)$$

式中, Q 、 K 、 V 分别是查询向量、键向量、值向量, W_q 、 W_k 、 W_v 是对应向量的权重矩阵。将查询向量 Q 与键向量 K 进行点积运算后, 除以比例因子 $\sqrt{D_k}$, 再经过 softmax 函数后与值向量 V 进行点积运算后得到注意力分布, 其表达式如下所示:

$$H = V \text{softmax} \left(\frac{QK^T}{\sqrt{D_k}} \right) \tag{7}$$

式中, K^T 是键向量 K 的转置。其具体计算过程如图 5 所示:

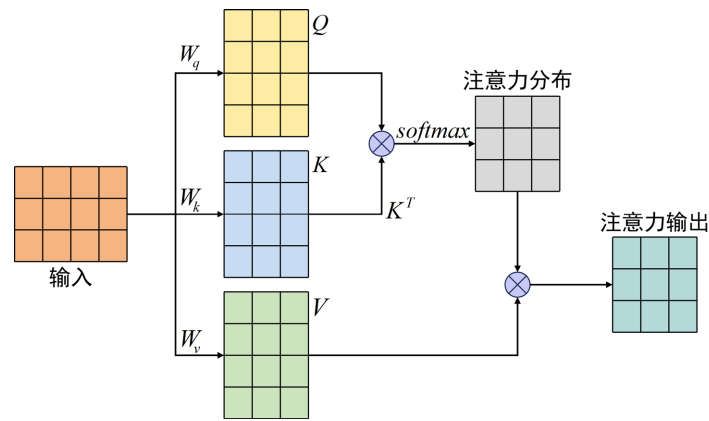


Figure 5. Computational graph of the self-attention mechanism
图 5. 自注意力机制计算图

在 Transformer 编码器里, 把多个子注意力头并行地构造成多头注意力机制, 每一个注意力头, 在不同的子空间向量里提取到的信息是不一样的。由于 Transformer 没有循环结构, 不能像 RNN 一样学习序列信息, 所以引入位置编码来表示每个数据在序列中的位置, 其计算式如下所示:

$$PE_{\text{pos}, 2i} = \sin \left(\frac{\text{pos}}{10000^{2i/d_{\text{model}}}} \right)$$
$$PE_{\text{pos}, 2i+1} = \cos \left(\frac{\text{pos}}{10000^{2i/d_{\text{model}}}} \right) \tag{8}$$

式中, pos 是数据的位置索引, i 表示维度索引, d_{model} 为嵌入维度。在自注意力计算之后, 每个 Transformer 层都会通过一个两层的前馈神经网络进行非线性变换, 然后进行残差连接和层归一化后, 将数据输出。Transformer 编码器结构如图 6 所示。

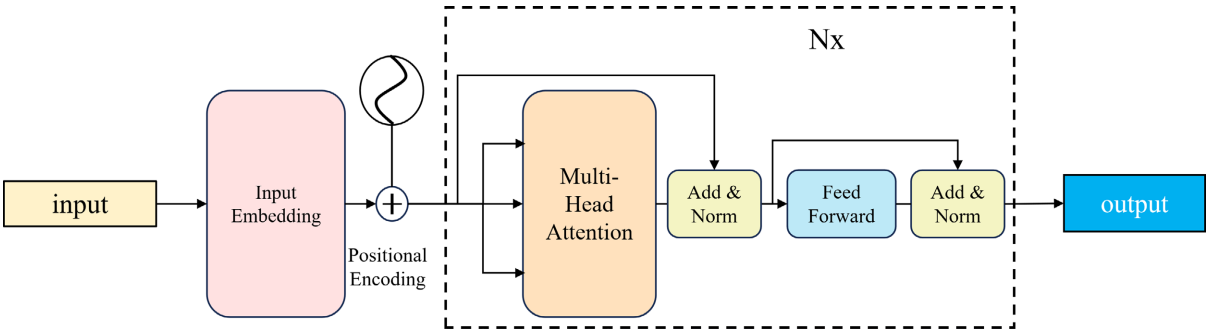


Figure 6. Transformer encoder structure diagram
图 6. Transformer 编码器结构图

3. MTB 测井曲线重构流程

根据 MTB 测井曲线重构流程图 7 可知, 整个操作可以分成三大部分。第一阶段, 先对输入的测井数据做相关性分析, 得到目标重构曲线与其它各种曲线的相关程度, 从中选出和待重建参数相关性较高的测井信息作为模型的特征输入。接着, 对选择的特征进行归一化处理, 用 4 口井做训练样本、1 口井做验证样本、1 口井做测试样本。第二阶段是给 MTB 模型的网络权重进行初始化设置, 然后用训练样本集进行模型的学习。训练阶段用均方误差(MSE)做代价函数, 具体的计算公式见式(13)。根据代价函数数值的变化来评价模型学习情况, 利用误差反向传播法更新 MTB 模型各个参数, 并采用 Adam 优化算法对参数进行循环调整。当价函数收敛到设定的目标值之后, 就将已经完成训练的模型保存下来。第三阶段调用之前存储好的模型对测试井进行曲线重建预测, 然后对预测结果进行反归一化变换, 得到重建后的测井曲线数值序列。

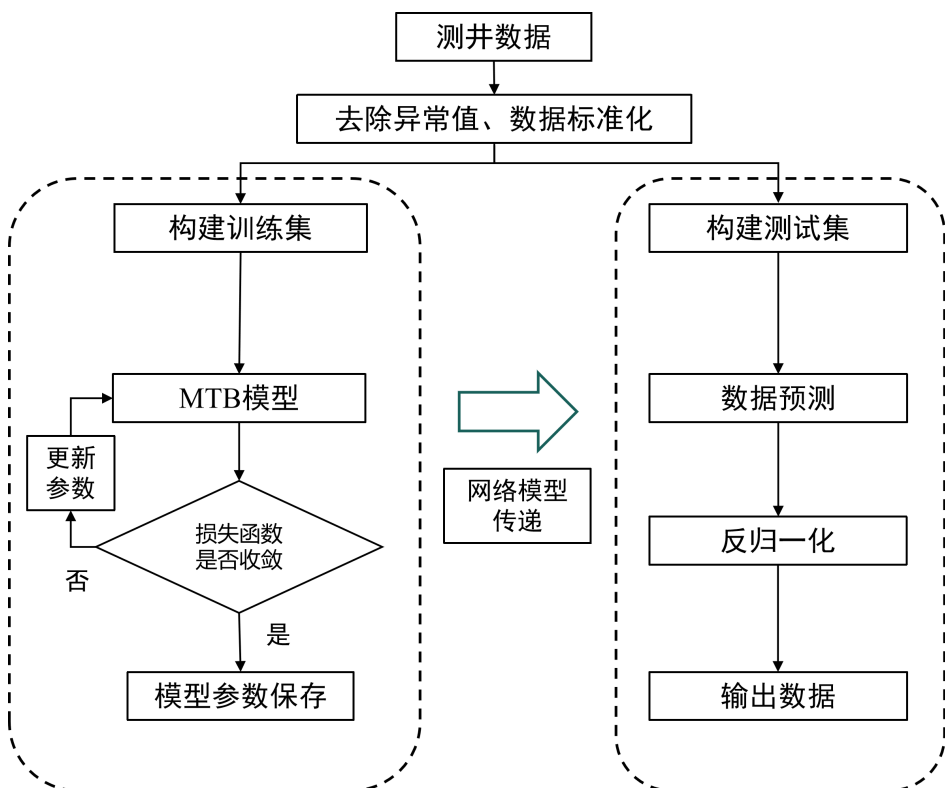


Figure 7. MTB workflow for well-log reconstruction
图 7. MTB 测井曲线重构流程

4. 实验及结果分析

4.1. 实验数据集

本文所用测井数据来源于绵阳北部页岩气区块, 共 6 口井的测井资料。为了方便区分, 将它们分别用 W1 到 W6 表示, 整体称为数据集 W。模型输入数据为井内所记录下来的声波时差、自然伽马、深侧向电阻率这三种主要测井参数, 单位分别为 $\mu\text{s}/\text{ft}$ 、API、 $\Omega\cdot\text{m}$ 。所有的井的采样间距都是 0.125 m, 数据集 W 的详细指标如表 1 所示。为了对 MTB 模型有更深入的了解, 用现有的数据集做单目标缺失重构实验来检验 MTB 模型的性能。

Table 1. Information related to well-log dataset W
表 1. 测井曲线数据集 W 相关信息

井名	深度(m)	采样间隔(m)	样本数
W1	3567.75~3640.25	0.125	580
W2	3490.25~3570.75	0.125	644
W3	3558~3633	0.125	600
W4	3482.25~3556.5	0.125	594
W5	3567.75~3640.25	0.125	580
W6	3560.5~3635.75	0.125	602

4.2. 相关性分析

为了考察 AC、GR、LLD 三者之间线性相关程度，也为输入特征的选择和预测目标的合理性打下基础，在训练 MTB 模型之前对样本来做皮尔逊相关性分析，并用热力图形式展示出来。皮尔逊相关系数的表达式为：

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (9)$$

其中， $\bar{x}$ 为样本预测值的平均值， $\bar{y}$ 为样本真实值的平均值。针对 AC 曲线的单目标重构任务，其相关性分析结果如图 8 所示。从相关系数矩阵可以看出，RT 与 AC 之间相关系数较低，GR 与 AC 相关系数较高。由于工区内可用的测井样本数量较少，因此在 AC 曲线重构时，最终选取 GR 和 RT 作为输入特征。

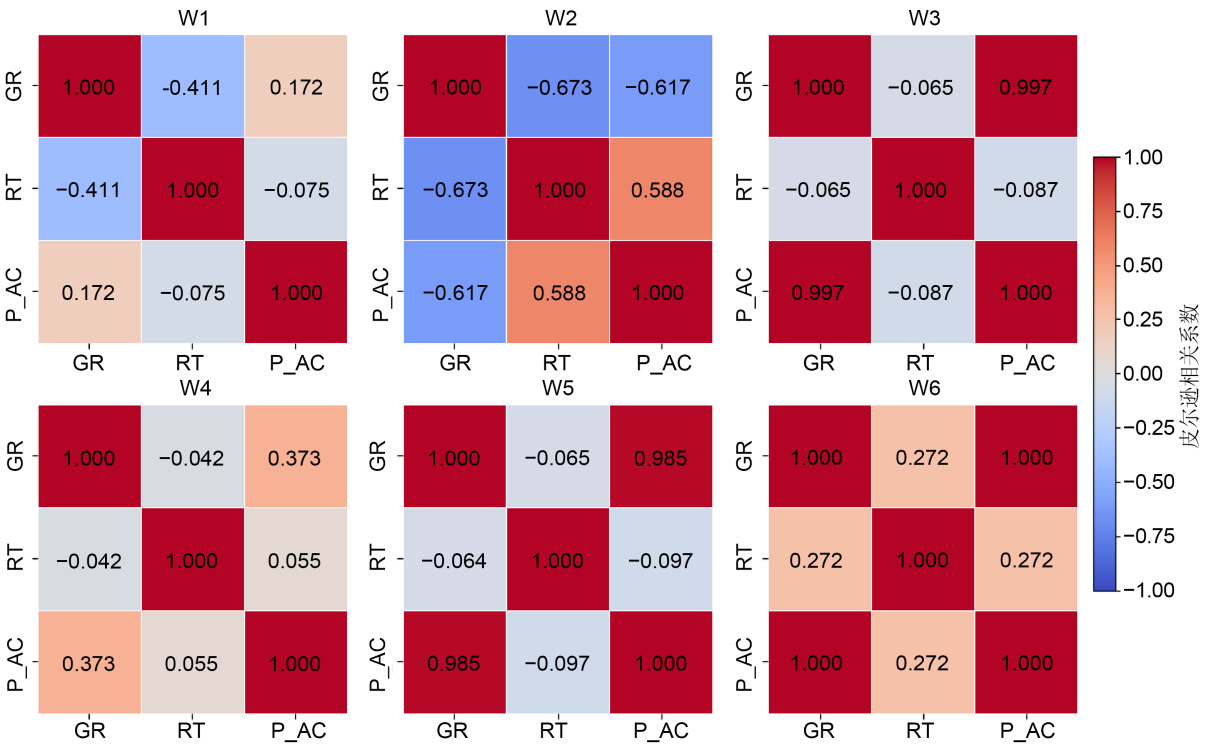


Figure 8. Pearson correlation coefficient matrix between AC and other well logs
图 8. AC 与其他测井曲线皮尔逊相关系数矩阵

4.3. 实验设置与评价指标

本文所提出的测井曲线重构模型采用 MTB (Matelearning-Transformer-BiLSTM)网络结构。为了充分发挥各个模块的特征提取能力以及序列建模能力,在元学习训练过程中保证模型具有较好的收敛性和泛化性,经过多次实验最终得到模型结构参数和训练超参数如下表所示。

因为实验中要进行单目标预测和多目标预测,所以输入维度根据所选测井曲线特征数量动态赋值,在模型运行时自动确定,保证不同的实验方案之间的灵活性以及兼容性。模型内部结构为 BiLSTM 隐藏层维度设为 64, BiLSTM 层数设为 2 层,双向结构有利于捕捉测井曲线上下文信息,提高模型对测井曲线上下文信息的建模能力。在 BiLSTM 之后加入多头自注意力机制,头数设为 4,用多头机制来实现不同子空间的特征交互,提高序列内部不同位置之间全局相关性的建模效果。模型用残差连接、层归一化解决了深层网络中梯度消失的问题,加快了收敛速度。dropout 比例设为 0.5,可以使用配置文件进行调节,随机失活方式有助于防止过拟合,提高模型的泛化能力。

训练阶段用 Reptile 元学习算法。内循环学习率为 0.01,内循环更新步数为 5;外循环学习率为 0.001。序列长度设为 60,可以涵盖测井曲线里常见的地质变动范围,从而使得模型对于各个深度区间都有较好的感受效果。元训练的总迭代轮数设为 1000,用早停机制根据验证井的均方误差连续 30 次更新幅度较小的时候就停止训练,实验中发现模型在这个设置下收敛比较稳定,可以得到较好的预测性能。需要说明的是,由于 Reptile 内循环使用支撑集中所有的样本进行批量梯度下降,批量大小由支撑集样本数决定,不需要事先固定。

测井数据存在不同的量纲和数值范围,为了消除测井数据对模型训练的影响,本文用最大最小归一化法对输入特征进行标准化处理。归一化公式如式(10)所示:

$$Y_{pre} = \frac{Y - Y_{\min}}{Y_{\max} - Y_{\min}} \quad (10)$$

其中, Y_{pre} 为归一化后的模型输入值, Y 为原始测井值, $Y_{\max}$ 、 $Y_{\min}$ 分别为目标测井曲线的最大值与最小值。预测完成后需按相同公式进行反归一化,恢复至原始量纲。同时,为评估测井曲线重构的准确性及其在不同模型中的表现,本文引入均方根误差(RMSE)、平均绝对误差(MAE)、平均绝对百分比误差(MAPE)和决定系数(R^2)这四个评价指标。能够较全面地反映模型在重构任务中的性能优劣。四个评估指标的公式如下:

$$\begin{aligned} RMSE &= \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \\ MAE &= \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \\ MAPE &= \frac{100\%}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \\ R^2 &= 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \end{aligned} \quad (11)$$

其中, y_i 为真实值, $\hat{y}_i$ 为预测值, $\bar{y}$ 为真实值的均值。

4.4. AC 测井曲线重构

本部分就 AC 测井曲线设计了单目标回归重构实验,用来比较不同深度的神经网络模型在测井预测任务中表现。实验使用数据集 W,用 W5 井做测试集、W4 井做验证集,剩下的 4 口井作为训练集。训

练过程里所有的模型都会把 GR 和 LLD 当作输入特征，从而创建起 AC 曲线预测的联系。各个模型在测试井上所做预测的误差以及拟合结果见表 2。

Table 2. Comparison of model predictions for single-target CNL well-log reconstruction
表 2. 单目标 CNL 测井曲线重构各模型预测对比

井名	模型名称	RMSE	MAE	MAPE (%)	R ²
W5	BiLSTM	6.1468	4.5964	2.16	0.8482
	Transformer	6.8988	5.9543	2.87	0.8088
	1D-CNN	6.7890	4.9431	2.30	0.8148
	BiGRU	6.4194	4.7598	2.24	0.8344
	MTB	3.9531	3.2538	1.54	0.9372

根据表 2 中 W5 井各模型的评估结果，BiLSTM 的 R² 为 0.8482、RMSE 为 6.1468，展现出较强的整体拟合能力；BiGRU 的 R² 为 0.8344、RMSE 为 6.4194，Transformer 的 R² 为 0.8088、RMSE 为 6.8988，两者在趋势捕捉上相对稳定，但在局部波动较大区段仍存在较大误差；1D-CNN 的 R² 为 0.8148、RMSE 为 6.7890，对复杂测井特征的建模能力有限。综合全部指标，MTB 模型取得了最优的预测性能，其 RMSE 仅为 3.9531，MAE 为 3.2538，MAPE 为 1.54%，R² 高达 0.9372，重构曲线与真实测量值高度吻合，能够精确还原不同井深区间的非线性波动及整体变化趋势，显著优于其他对比模型。

为了清楚地比较不同的方法在 AC 曲线预测上所表现出的差别，图 9 给出了 W5 井实测数据和各个模型的重建结果。图中蓝色是真实的 AC 值，红色是模型的预测值。从结果可知，MTB 的重构曲线最接近实际值，尤其是 3580~3610 m 井段局部波动和整体走势都得到了准确的再现。图 10 为 W5 井 AC 测井曲线真实值和预测值散点图，从图中可以看出 MTB 预测结果与真实测井结果更接近线性关系。因此从以上分析可知，MTB 在 W5 井上取得最好的预测效果，可以很好地学习到 AC 曲线的非线性动态和深层时序特征，具有较好的推广性和地质适应性。

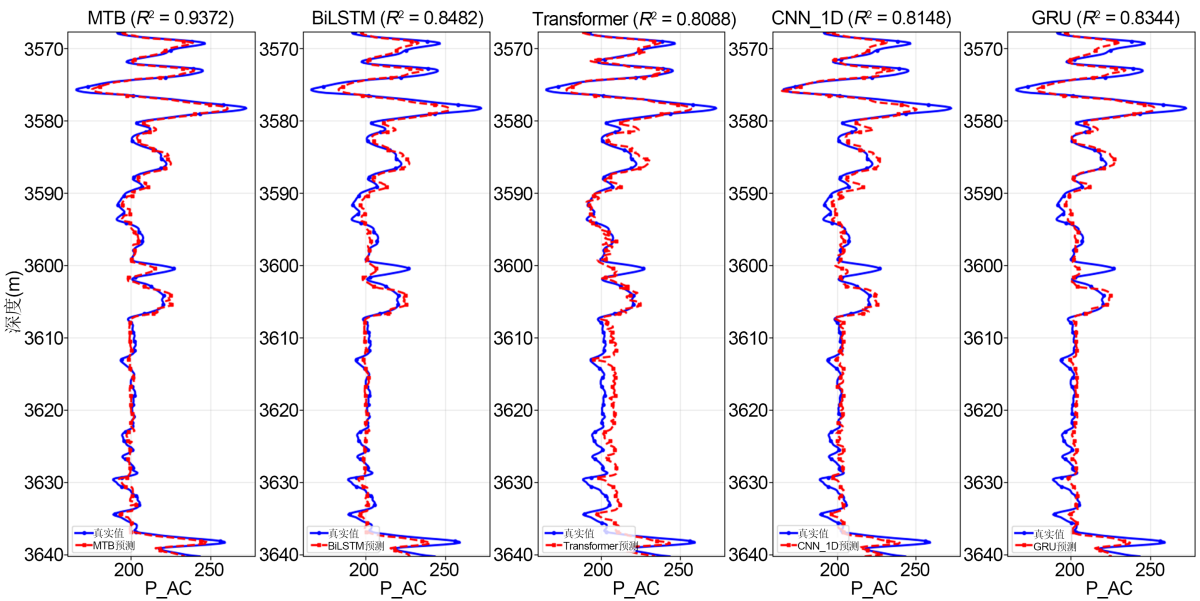


Figure 9. Prediction of well log W5
图 9. W5 测井曲线预测

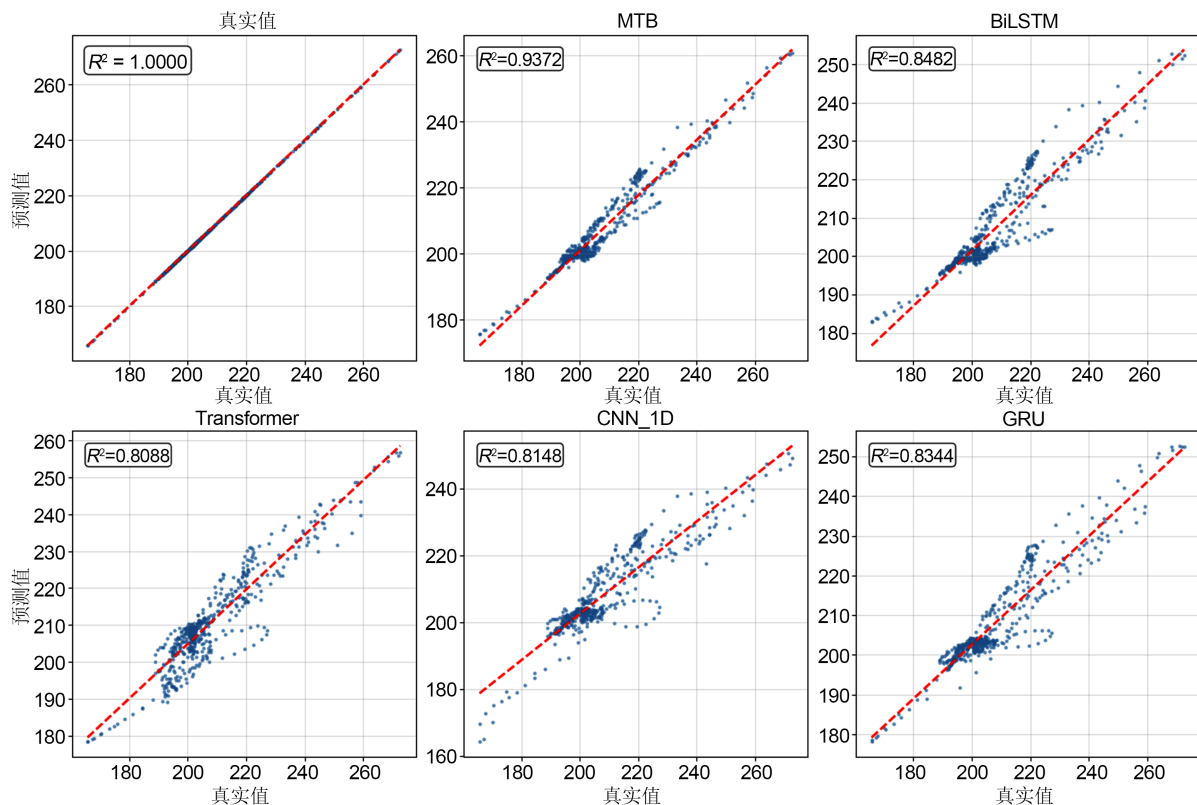


Figure 10. Scatter plot of true and predicted values of the AC well log from well W5

图 10. W5 井 AC 测井曲线的真实值和预测值散点图

5. 结论与展望

针对页岩气勘探过程中测井曲线经常出现数据缺失的问题, 本文设计出一个叫 MTB 的深度学习神经网络模型来重建测井曲线。该模型把 Reptile 元学习策略、BiLSTM 双向时序处理能力、Transformer 编码器全局注意力机制这三者结合起来, 可以比较全面地捕捉到测井数据里存在的非线性关系和序列依赖特征。在绵阳北部页岩气区块实际数据上, 以 AC 曲线为目标进行重构实验时, MTB 模型在 RMSE、MAE、MAPE 以及 R^2 等各项评价指标上都比参与对比的其它模型好。下一步可以把它应用到多类测井参数的联合预测任务上。

参考文献

- [1] 杨跃明, 陈玉龙, 刘燊阳, 等. 四川盆地及其周缘页岩气勘探开发现状、潜力与展望[J]. 天然气工业, 2021, 41(1): 42-58.
- [2] Wang, Z. and Wang, Y. (2025) Integrating Unsupervised Learning and Transformer for Missing Log Prediction. *Geophysics*, **90**, D85-D100. <https://doi.org/10.1190/geo2024-0558.1>
- [3] Hao, L., Cai, B., Cao, Z., Han, J., Li, J. and Qie, D. (2024) Well Logging Restoration Based on Spatial-Temporal Information Mining from Multi-Scale Correlation Graph Representation. *Petroleum Science and Technology*, **43**, 2080-2103. <https://doi.org/10.1080/10916466.2024.2378205>
- [4] 罗功伟, 安小平, 姚卫华, 等. 非常规油气储层测井智能解释应用现状与发展趋势[J]. 石油科学通报, 2025, 10(5): 908-925.
- [5] Haritha, D., Satyavani, N. and Ramesh, A. (2025) Generation of Missing Well Log Data with Deep Learning: CNN-Bi-LSTM Approach. *Journal of Applied Geophysics*, **233**, Article ID: 105628. <https://doi.org/10.1016/j.jappgeo.2025.105628>

-
- [6] Garini, S.A., Shiddiqi, A.M., Utama, W. and Abduh, M.U.N. (2025). Reconstructing Missing Well-Log Data with Lightgbm and Birnn in Oil Fields. 2025 International Conference on Smart Computing, IoT and Machine Learning (SIML), Surakarta, 3-4 June 2025, 1-6. <https://doi.org/10.1109/siml65326.2025.11081019>
 - [7] Park, G., Jeong, J., Park, E., Jeong, J., Lee, B. S., and Song, S. H. (2019) Comparative Analysis on the Performance of Machine Learning-Based Algorithms to Reconstruct the Missing Geophysical Well Log Data. *AGU Fall Meeting Abstracts*, **2019**, Artical ID: IN41D-0888.
 - [8] Kang, Z.Z., Guo, J.H., Gu, B.X., *et al.* (2025) Carbonate Microfacies Identification Using Residual LSTM Network. *Progress in Geophysics*, **40**, 2736-2749.
 - [9] Yu, Z., Sun, Y., Zhang, J., Zhang, Y. and Liu, Z. (2023) Gated Recurrent Unit Neural Network (GRU) Based on Quantile Regression (QR) Predicts Reservoir Parameters through Well Logging Data. *Frontiers in Earth Science*, **11**, 1087385. <https://doi.org/10.3389/feart.2023.1087385>
 - [10] Zhang, L. and Gong, A. (2025) A Deep Learning Multi-Module Fusion Method for Well Logging Curve Reconstruction. *Physics of Fluids*, **37**, Article ID: 076627. <https://doi.org/10.1063/5.0276398>
 - [11] Cao, Z., Ye, L., Zheng, B., Han, J., Liu, X. and Xia, S. (2025) CCD-WLR: CNN-Bi-LSTM and CNN-Transformer-Based Dual-Channel for Missing Well Logs Reconstruction. *Engineering Research Express*, **7**, Article ID: 0452b5. <https://doi.org/10.1088/2631-8695/ae1883>
 - [12] Zou, C.L., Zhang, J.H., Sun, Y.Z., Pang, S. and Zhang, Y. (2024) Enhancing Fluid Classification Using Meta-Learning and Transformer through Small-Sample Drilling Data to Interpret Well Logging Data. *Physics of Fluids*, **36**, Article ID: 076608. <https://doi.org/10.1063/5.0211187>
 - [13] Nichol, A., Achiam, J. and Schulman, J. (2018) On First-Order Meta-Learning Algorithms. arXiv: 1803.02999.
 - [14] Hochreiter, S. and Schmidhuber, J. (1997) Long Short-Term Memory. *Neural Computation*, **9**, 1735-1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
 - [15] Graves, A. and Schmidhuber, J. (2005) Framewise Phoneme Classification with Bidirectional LSTM and Other Neural Network Architectures. *Neural Networks*, **18**, 602-610. <https://doi.org/10.1016/j.neunet.2005.06.042>
 - [16] Vaswani, A., Shazeer, N., Parmar, N., *et al.* (2017) Attention Is All You Need. arXiv: 1706.03762.
 - [17] Pascanu, R., Mikolov, T. and Bengio, Y. (2013) On the Difficulty of Training Recurrent Neural Networks. *Proceedings of the 30th International Conference on Machine Learning*, Atlanta, Georgia, 16-21 June 2013, 1310-1318.