

# 基于双向长短时循环神经网络的测井曲线重构方法研究

张 茜<sup>1</sup> 丁 磊<sup>1</sup> 王子国<sup>1</sup> 武凌峰<sup>1</sup> 吴 妍<sup>1</sup>

ZHANG Xi DING Lei WANG Ziguo WU Lingfeng WU Yan

## 摘 要

南海西部海域油气田勘探开发过程中, 广泛采集测井资料用于油气层精确解释评价。然而随着钻井过程中井况变得日趋复杂, 经常会出现测井曲线缺失或失真的情况。针对该问题, 比较经济便捷的方法是利用机器学习算法基于现有测井曲线进行曲线重构, 但是传统技术无法综合测井曲线上下文信息, 影响了重构精度。文章基于 Bi-LSTM 双向长短时循环神经网络进行测井曲线重构方法研究, 首先对比了不同类型循环神经网络的测井曲线重构效果, 然后对神经网络结构进行优化。研究表明, 基于多层 Bi-LSTM 神经网络进行测井曲线重构的精度更高。

## 关键词

测井; 机器学习; 曲线重构; 双向长短时循环神经网络

doi: 10.3969/j.issn.1672-9528.2025.11.018

## 0 引言

在南海西部海域油气田勘探开发过程中, 广泛采用测井技术来测量目的层的声学、电学、化学以及放射性等特征。利用种类丰富的测井资料, 研究人员可以分析出目的层的岩性、物性以及含油性等多项物理性质, 为勘探开发的下一步工作提供精确指导。然而, 随着勘探开发领域由常规油气藏向低孔低渗、高温高压等复杂油气藏转变, 钻井过程中井况也变得日趋复杂, 经常出现测井曲线缺失或失真的情况。虽然可以通过复测的方式重新采集测井资料, 但耗时长、费用高。因此, 基于现有测井资料重构缺失或失真的测井曲线, 是更经济便捷的方法。

针对测井曲线重构, 国内外研究人员首先是采用岩石物理建模以及多元线性回归的方式进行重构, 这些方法在局部区块应用效果较好, 但是泛化性较差。近年来, 随着机器学习技术的不断发展与普及, 研究人员开始尝试应用各种机器学习技术进行测井曲线重构研究。张家臣等人<sup>[1]</sup>和齐春生等人<sup>[2]</sup>应用 XGBoost 极端梯度提升树算法进行测井曲线重构, 该算法支持对稀疏矩阵的处理能力, 可以有效节约数据准备时间。石玉江等人<sup>[3]</sup>应用 FCNN 全连接神经网络进行电阻率曲线重构并应用于水淹层定性识别中。虽然全连接神经网络具有较强的表示能力, 但是也存在训练时间过长、计算成本过高的问题, 所以对大规模数据集的应用效果不佳。为解决该问题, 研究人员开始广泛使用深度学习神经网络, 在有效降低计算成本的同时, 也具备了大规模数据的处理能力、强大的泛化能力与抗噪能力。李枫林等人<sup>[4]</sup>应用 U-Net 神经网络

进行声波测井曲线重构, 翟晓岩等人<sup>[5]</sup>应用基于时空注意力机制的 CNN 网络进行声波和密度测井曲线, 取得了良好的应用效果。但是, 上述模型均属于点对点预测的机器学习模型, 只能分析同一深度点不同测井曲线数据之间的相关关系, 而忽略了测井曲线数据随地层深度的变化趋势以及数据间的前后关联<sup>[6]</sup>。

针对测井数据这类序列数据, 循环神经网络作为一种功能强大的序列处理模型被广泛应用, 可以综合考虑不同深度、不同测井曲线之间的综合变化趋势, 更符合岩石物理学思想。杨满玉等人<sup>[7]</sup>应用 RNN 神经网络对岩性密度曲线和声波测井曲线进行了重构, 重构精度高于传统机器学习算法, 但由于 RNN 循环神经网络需要为每一步保持一个激活向量, 容易在训练阶段出现梯度爆炸问题<sup>[8]</sup>。尚福华等人<sup>[9]</sup>应用 LSTM 长短时循环神经网络, 王俊等人<sup>[10]</sup>应用 GRU 门控制循环神经网络进行处理, 避免了梯度爆炸问题。但传统的长短时循环神经网络以及门控制循环神经网络都只是针对前向序列数据进行处理, 而忽略了后向序列数据的影响, 不利于测井曲线的精确重构。

本次研究以南海西部海域 A 油田为靶区, 基于 Bi-LSTM 双向长短时循环神经网络进行测井曲线重构, 首先对比 RNN、LSTM 及 Bi-LSTM 神经网络的重构效果, 然后分析并优选出最佳的神经网络层结构。实验结果表明, 基于多层 Bi-LSTM 神经网络进行测井曲线重构的精度更高。

## 1 原理介绍

### 1.1 RNN 神经网络

RNN 神经网络是一种专门处理序列数据建模问题的神

1. 中海石油(中国)有限公司海南分公司 海南海口 570100

神经网络,如图1所示为网络结构示意图,它通过循环连接来传递信息,使得网络能够持有一定的“记忆”。在处理序列的时间步 $t$ 时,网络会同时接收时间步 $t$ 的输入 $x_t$ 和时间步 $t-1$ 的隐藏状态 $S_{t-1}$ ,然后生成时间步 $t$ 的输出 $O_t$ 和隐藏状态 $S_t$ ,其中 $S_t$ 会进一步传到时间步 $t+1$ 参与下一步计算。通过这种方式,隐藏状态就起到了记忆的作用,携带着序列之前时间步的信息,并影响后续时间步的操作<sup>[11]</sup>。

RNN神经网络反向传播使用的是(back propagation through time, BPTT)算法,其损失Loss对参数 $w$ 的梯度等于Loss在各个时间步对 $w$ 求导之和,如果时间步之间距离较远,可能会产生梯度消失或梯度爆炸的问题,所以只适用于短时间序列。

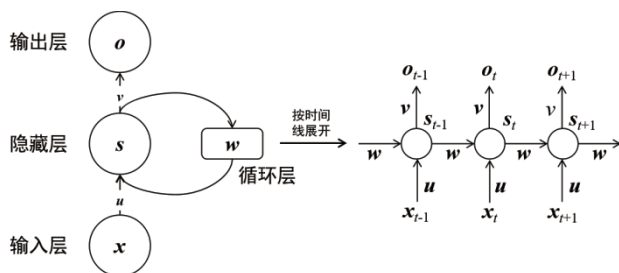


图1 RNN循环神经网络结构示意图

## 1.2 LSTM 神经网络

LSTM神经网络可解决RNN神经网络在长时间序列建模过程中的梯度消失或梯度爆炸问题(结构如图2所示)。该网络通过引入3个门结构(遗忘门、输入门、输出门)和一个称为细胞状态的内部记忆单元来控制信息的流动。在处理序列的时间步 $t$ 时,遗忘门决定了细胞状态 $S_{t-1}$ 中需要丢弃的信息,输入门决定了需要添加到细胞状态 $S_t$ 中的新信息,而输出门则控制了下一个隐藏状态 $h_t$ 应该包含的信息。细胞状态允许信息在序列中长时间传播,从而使LSTM神经网络能够学习到长距离的时间依赖关系<sup>[12]</sup>。

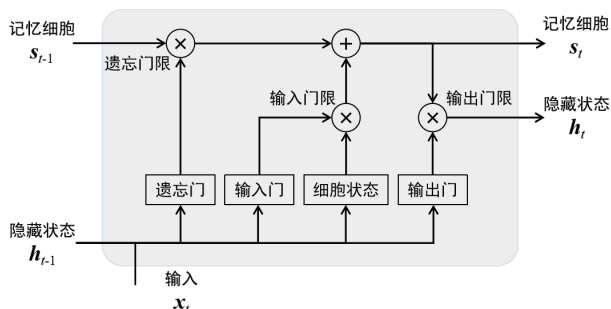


图2 LSTM长短时循环神经网络结构图

## 1.3 Bi-LSTM 神经网络

LSTM神经网络可以避免RNN神经网络存在的梯度消失或梯度爆炸问题,但是只能考虑前向数据,而忽略了后向序列数据的影响。为此研究人员提出了Bi-LSTM神经网络(结

构如图3所示),在LSTM神经网络的基础上增加了反向传递的循环层,使得网络在时间步 $t$ 的输出不仅依赖于时间步 $t-1$ 的信息,也依赖于时间步 $t+1$ 的信息。Bi-LSTM神经网络有两个LSTM层,一个处理正向序列,另一个处理反向序列。两个LSTM层的输出在每个时间步被合并,以提供更全面的上下文信息。

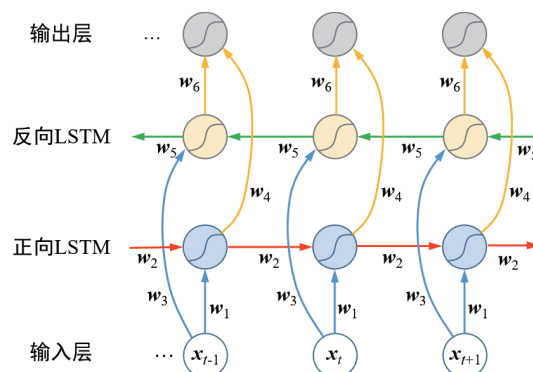


图3 Bi-LSTM双向长短时循环神经网络结构图

对于时间步 $t$ 的输入 $x_t$ ,  $w_1 \sim w_6$ 为节点权值,正向LSTM层输出 $h_{ft}$ 、反向LSTM层输出 $h_{bt}$ 表示为<sup>[13]</sup>:

$$h_{ft} = f(w_1 x_t + w_2 h_{ft-1}) \quad (1)$$

$$h_{bt} = f(w_3 x_t + w_5 h_{bt+1}) \quad (2)$$

此时,输出层输出 $y_t$ 可以表示为:

$$y_t = f(w_4 h_{ft} + w_6 h_{bt}) \quad (3)$$

## 1.4 优化算法

优化算法主要用于调整神经网络中的超参数,使得训练数据集上的损失函数尽可能小,从而使神经网络运行得更快。常见的优化算法有Momentum、SGDM、Adam、RMSprop、LBFGS算法等,本次研究中使用Adam优化算法,该算法集成了Momentum和RMSprop算法的理论,既会跟踪过去梯度的指数衰减平均值,也会跟踪过去梯度平方的指数衰减平均值。

## 1.5 防止过拟合

在机器学习过程中,若训练样本数量不足、噪声干扰过多或模型过于复杂,可能会出现过拟合的现象。为防止过拟合现象发生,本文通过添加DropOut层的方式,在每次迭代过程中随机丢弃一定比例的神经元节点,相当于训练了多个神经网络,增强了模型的泛化能力。

## 2 实验分析

本次研究选择南海西部海域莺歌海盆地A油田作为靶区,取8口井测井数据开展实验,包括:PEF光电吸收截面指数、DCAV井径、GR自然伽马、P16H浅相位电阻率、P28H中相位电阻率、P40H深相位电阻率、A40H深衰减电

阻率、RHOB 密度以及 TNPH 中子九条曲线。将 1 井至 7 井作为训练样本, 8 井作为预测样本, 在此基础上进行两个实验: 一是对比不同类型循环神经网络重构测井曲线的效果; 二是对比不同层数循环神经网络重构测井曲线的效果。由此优选出最佳的神经网络结构。

## 2.1 RNN、LSTM 及 Bi-LSTM 神经网络重构效果对比

本实验选择单层 RNN、LSTM 以及 Bi-LSTM 神经网络进行训练, 假设 TNPH 中子曲线缺失, 利用其他测井曲线重构该曲线。图 4 为各模型重构效果图, 图中绿线、红线分别为 TNPH 中子曲线原始值和预测值。在 3 169~3 218 m 井段, 3 种神经网络的重构效果均与真实值有一定偏差, 但整体效果 Bi-LSTM 最好, LSTM 次之, RNN 较差; 在 3 218~3 280 m 井段, 3 种神经网络的重构效果均与真实值较为接近, 但整体效果 Bi-LSTM 最好, RNN 次之, LSTM 较差。综合而言, Bi-LSTM 神经网络可以综合反映测井曲线上下文的信息, 重构精度最高。

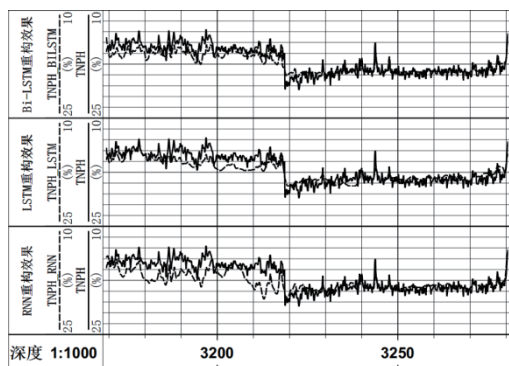


图 4 RNN、LSTM 及 Bi-LSTM 神经网络重构效果对比

## 2.2 不同隐含层层数 Bi-LSTM 神经网络重构效果对比

由前述可知, Bi-LSTM 神经网络相比其他循环神经网络重构精度更高。因此, 本实验在单层 Bi-LSTM 神经网络基础上, 通过对比不同 Bi-LSTM 层数的重构效果, 进一步对神经网络层结构进行优化。

多层 Bi-LSTM 神经网络的输入层为 1 个 `sequenceInputLayer`, 隐藏层为多个 `biLstmLayer`, 每个 `biLstmLayer` 后添加一个 `dropoutLayer` 以防止过拟合, 输出层为一个 `fullyConnectedLayer` 和一个 `regressionLayer`。在此基础上再次进行 TNPH 中子曲线重构实验, 得到不同 Bi-LSTM 层数的重构效果对比 (如图 5 所示), 可见随着堆叠的隐含层数量增加, 训练集和预测集的重构 RMSE 首先会下降, 在 Bi-LSTM 层数为 3 时达到最低, 当层数继续增加时, 层与层之间的梯度消失情况变得更加明显, 收敛效果与计算效率会快速下降, 重构 RMSE 会上升, 甚至进入局部最小的情况<sup>[14]</sup>, 最终确定多层 Bi-LSTM 神经网络模型的层数为 3。

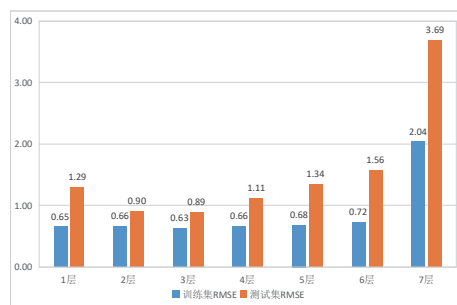


图 5 多层 Bi-LSTM 神经网络重构 RMSE 对比

利用 3 层 Bi-LSTM 神经网络重构 8 井 TNPH 中子曲线, 并与单层 Bi-LSTM 重构效果进行对比 (如图 6 所示), 可见 3 层 Bi-LSTM 神经网络重构精度更高, 尤其在 3 169~3 218 m 井段有明显改善。综上所述, 利用 3 层 Bi-LSTM 神经网络进行测井曲线重构, 可取得相比其他循环神经网络更高的重构精度。

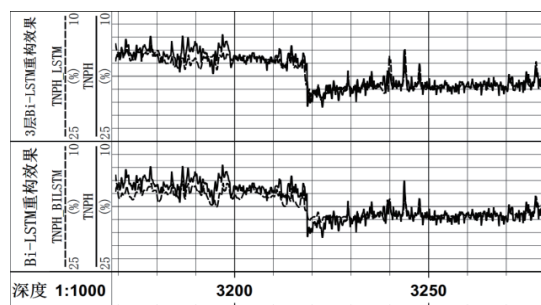


图 6 多层 Bi-LSTM 神经网络重构 RMSE 对比

## 3 结论

本文从循环神经网络机理出发, 对比了多种类型循环神经网络在测井曲线重构任务中的应用效果, 并进行了神经网络层结构的优化分析。结果表明, 由于 Bi-LSTM 双向长短时循环神经网络可同时分析测井曲线上下文信息, 因此测井曲线重构精度明显优于其他循环神经网络。增加 Bi-LSTM 层数可以进一步提高神经网络重构精度, 但并不是越多越好, 需要针对实际情况进行分析, 以确定最优的 Bi-LSTM 层数。

## 参考文献:

- [1] 张家臣, 邓金根, 谭强, 等. 基于 XGBoost 的测井曲线重构方法 [J]. 石油地球物理勘探, 2022, 57(3): 697-705.
- [2] 齐春生, 丁磊, 王一, 等. 基于极端梯度提升树算法的测井曲线重构方法 [J]. 科技和产业, 2022, 22(10): 405-410.
- [3] 石玉江, 周金显, 钟吉彬, 等. 重构电阻率曲线识别水淹层的方法及应用 [J]. 测井技术, 2018, 42(1): 42-48.
- [4] 李枫林, 刘怀山, 杨熙镭, 等. 基于 U-Net 神经网络的声波测井曲线重构 [J]. 中国海洋大学学报 (自然科学版), 2023, 53(8): 86-92, 103.

(下转第 84 页)

- org/10.48550/arXiv.2306.00890.
- [3]GRATTAFIORI A, DUBEY A, JAUHRI A, et al. The Llama 3 herd of models[EB/OL]. (2024-11-23)[2025-09-25]. <https://doi.org/10.48550/arXiv.2407.217>.
- [4]SHEN S, YAO Z W, LI C Y, et al. Scaling vision-language models with sparse mixture of experts[C]//Findings of the Association for Computational Linguistics: EMNLP 2023. Brussels: ACL, 2023:11329-11344.
- [5]HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]// 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2016: 770-778.
- [6]FAN Z H, TANG J L, CHEN W, et al. AI hospital: interactive evaluation and collaboration of LLMs as intern doctors for clinical diagnosis[EB/OL].(2024-02-21)[2025-09-12]. <https://arxiv.org/html/2402.09742v2>.
- [7]YOU H X, SUN R, WANG Z C, et al. IdealGPT: iteratively decomposing vision and language reasoning via large language models[EB/OL].(2025-04-11)[2025-09-12]. <https://doi.org/10.48550/arXiv.2305.14985>.
- [8]ZHANG Z S, ZHANG A, LI M, et al. Multimodal chain-of-thought reasoning in language models[EB/OL]. (2024-05-20)[2025-09-25].<https://doi.org/10.48550/arXiv.2302.00923>.
- [9]LIN B, TANG Z Y, YE Y, et al. MoE-LLaVA: mixture of experts for large vision-language models[EB/OL].(2024-12-23)[2025-09-25].<https://doi.org/10.48550/arXiv.2401.15947>.
- [10]LAU J J, GAYEN S, ABACHA A B, et al. A dataset of clinically generated visual questions and answers about radiology images[J/OL]. Scientific data, 2018[2025-09-11].<https://www.nature.com/articles/sdata2018251>.
- [11]HE X H, ZHANG Y C, MOU L T, et al. PathVQA: 30000+ questions for medical visual question answering[EB/OL]. (2020-03-07)[2025-09-25].<https://doi.org/10.48550/arXiv.2003.10286>.

#### 【作者简介】

崔芸山(1999—), 男, 四川绵阳人, 硕士研究生, 研究方向: 多模态、计算机视觉等。

孙希选(1999—), 男, 山东聊城人, 硕士研究生, 研究方向: 自然语言处理、深度学习等。

陈育德(1981—), 男, 黑龙江佳木斯人, 硕士研究生, 讲师, 研究方向: 医学信息学、深度学习。

李殿奎(1964—), 通信作者(email:datutu2002@sina.com), 男, 黑龙江齐齐哈尔人, 硕士研究生, 教授、硕士研究生指导教师, 研究方向: 图像处理、深度学习。

(收稿日期: 2025-05-30 修回日期: 2025-11-10)

(上接第78页)

- [5] 翟晓岩, 高刚, 李勇根, 等. 融合注意力机制的二维卷积神经网络测井曲线重构方法[J]. 石油地球物理勘探, 2023, 58(5): 1031-1041.
- [6] 周欣, 曹俊兴, 王兴建, 等. 基于双向门控循环单元神经网络的声波测井曲线重构技术[J]. 地球物理学进展, 2022, 37(1): 357-366.
- [7] 杨满玉, 李晶, 王锦鹏, 等. 基于循环神经网络的测井曲线重构技术研究[J]. 当代化工研究, 2023(8): 158-160.
- [8] 张海涛, 杨小明, 陈阵, 等. 基于增强双向长短时记忆神经网络的测井数据重构[J]. 地球物理学进展, 2022, 37(3): 1214-1222.
- [9] 尚福华, 卢玉莹, 曹茂俊. 基于改进 LSTM 神经网络的测井曲线重构方法[J]. 计算机技术与发展, 2022, 32(6): 198-202.
- [10] 王俊, 曹俊兴, 尤加春. 基于 GRU 神经网络的测井曲线重构[J]. 石油地球物理勘探, 2020, 55(3): 510-520.
- [11] 赵晖光. 深度学习与神经网络[M]. 北京: 电子工业出版社, 2023.
- [12] AURELIEN G. 机器学习实战: 基于 Scikit-Learn 和 TensorFlow[M]. 北京: 机械工业出版社, 2018.
- [13] 齐春生, 丁磊, 焦祥燕, 等. 基于双向长短时循环神经网络的沉积微相自动识别方法: 以莺歌海盆地东方 B 气田为例[J]. 科技和产业, 2023, 23(5): 217-221.
- [14] 付强, 王华伟. 基于多层 LSTM 的复杂系统剩余寿命智能预测[J]. 兵器装备工程学报, 2022, 43(1): 161-169.

#### 【作者简介】

张茜(1987—), 女, 山东潍坊人, 硕士, 工程师, 研究方向: 油气勘探开发数字化和智能化。

(收稿日期: 2025-05-28 修回日期: 2025-11-14)