

风电变频器功率的多网络融合可解释预测模型

罗 威¹
LUO Wei

摘 要

为应对“双碳”目标下风电高效智能运维需求,解决风电机组变频器输出功率因多变量时序特征引发的非线性强相关建模难题,文章提出一种融合 CNN、LSTM 与 Transformer 的可解释性预测模型。该模型利用 CNN 压缩输入维度并提取局部特征,通过 LSTM 挖掘时序依赖关系,借助 Transformer 多头注意力机制强化全局关联建模;同时引入 SHAP 方法,定量分析各输入变量对功率输出的影响以增强可解释性。基于真实风电 SCADA 数据的实验结果表明,该模型在 SMAPE、MAE、RMSE 等评估指标上均优于传统机器学习及单一深度学习方法;SHAP 可视化进一步验证了模型对关键变量的有效识别能力。研究为风电场智能化管理与运维优化提供了可行方案及理论支撑。

关键词

风电变频器; 功率预测; 可解释性; 机器学习

doi: 10.3969/j.issn.1672-9528.2026.01.040

0 引言

随着“双碳”目标的深入推进,风电作为我国最具潜力的可再生能源之一,其在能源结构中的占比持续提升^[1]。风电机组变频器是风电并网与功率调节的核心部件,其运行状态直接影响风电场的能量输出及设备可靠性^[2]。因此,基于多源时序数据的变频器输出功率精确预测,不仅可提升风电场运行效率,还能故障预警与运维优化提供决策支持,对实现风电高效、安全、智能化管理具有重要意义^[3]。

现有研究中,传统机器学习方法在风电监测领域已得到广泛应用:决策树^[4]、随机森林^[5]等树模型凭借可解释性优势用于故障检测,但难以捕捉功率时序数据的动态依赖;支持向量回归^[6]、XGBoost^[7]、LightGBM^[8]等通过非线性拟合提升预测精度,却受限于手动特征工程,难以处理高维耦合特征。例如,刘冰冰等^[9]基于随机森林构建的变频器故障检测模型虽能识别异常状态,但对功率连续波动的预测精度有限;李森娟等^[10]采用支持向量机实现特定故障的高精度预测,却难以推广至复杂工况下的功率时序建模。Ren 等^[11]引入主动学习策略,提升了跨风电场故障预测的泛化能力。

近年来,深度学习方法逐步突破传统方法的局限:张西晨等^[12]将边缘智能引入风电监测,验证了端侧实时建模的工程可行性;徐武等^[13]通过改进 Transformer 模型增强了对 SCADA 数据长程依赖的捕捉能力。然而,现有深度模型仍存在三方面瓶颈:一是长距离时序依赖建模不足,普通

卷积网络(CNN)仅能提取局部特征,传统循环网络(如 LSTM)在长序列中易出现梯度衰减;二是非线性耦合拟合能力有限,单一模型难以兼顾风速、温度等多因素的复杂交互作用;三是可解释性缺失,深度模型的“黑箱”特性导致预测结果难以被运维人员理解,制约了工程落地^[14]。

针对上述挑战,本文提出融合 CNN、LSTM 与 Transformer 的可解释性功率预测模型,多模块协同实现“局部特征-时序依赖-全局关联”多层次建模:CNN 提取功率波动局部突变特征,LSTM 捕捉长周期时序依赖,Transformer 以多头自注意力强化多因素全局关联;同时植入可学习特征选择模块,结合注意力权重动态调整输入特征重要性,通过 SHAP 值量化分析参数对预测结果的影响,实现模型决策可视化。

实验验证表明,基于真实风电场 SCADA 数据的测试中,该模型测试集 R^2 达 0.964 6,在 MAE、RMSE 等关键指标上显著优于传统机器学习方法及单一深度学习模型(如 BiLSTM、CNN-GRU),充分验证了多模块融合在捕捉复杂时序特征上的优势。更重要的是,通过 SHAP 可视化分析,为运维人员提供了可信赖的决策依据,为风电系统的智能控制与运行优化提供可靠的数据支撑。

1 模型构建

1.1 卷积神经网络

在风电机组变频器功率预测模型中,卷积神经网络(CNN)发挥着关键作用。CNN 通过一维卷积提取功率波动的局部突变特征,池化操作降低维度、抑制噪声,为后续模块提供鲁棒特征表示^[15]。

1. 武汉轻工大学 湖北武汉 430048

通过一维卷积操作, CNN 能够有效提取风电功率时间序列中的局部特征, 捕捉其潜在模式与趋势, 为后续的 LSTM 和 Transformer 层提供丰富的特征表示。同时, 池化操作可降低数据维度, 抑制噪声, 保留关键信息, 从而提升模型的鲁棒性与泛化能力。

1.2 长短时记忆神经网络

LSTM 采用 128、64 单元双层结构, 通过门控机制捕捉长周期时序依赖, 向 Transformer 层输出完整序列信息。

风电功率数据具有显著的时间序列特性, 输入为经过卷积层提取的序列特征:

$$\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T] \in \mathbf{R}^{T \times d} \quad (1)$$

式中: T 表示时间步长度; d 表示每个时间步的特征维度。

卷积神经网络 (CNN) 首先提取局部特征, 随后两层 LSTM 分别以 128 和 64 个单元处理序列, 均返回完整序列, 以确保 Transformer 层接收到充分的时序信息。这种结构有助于更好地建模风电功率的时序特性, 提高预测精度。

1.3 Transformer 神经网络

Transformer 模型在处理风电功率数据中的长期依赖关系和复杂时序特征方面具有显著优势^[16]。Transformer 编码器结合位置编码与 8 头自注意力机制, 强化多因素全局关联建模, 缓解长序列梯度问题。

本研究采用 Transformer 编码器结构, 通过将输入特征与位置编码结合, 保留时间顺序信息, 位置编码表达式:

$$\text{PE}_{(\text{pos}, 2i)} = \sin\left(\frac{\text{pos}}{10000^{\frac{2i}{d_{\text{model}}}}}\right) \quad (2)$$

$$\text{PE}_{(\text{pos}, 2i+1)} = \cos\left(\frac{\text{pos}}{10000^{\frac{2i}{d_{\text{model}}}}}\right) \quad (3)$$

式中: pos 为当前序列的位置; i 为维度; d_{model} 为输入特征的维度。

再通过多头自注意力机制在多个子空间并行建模, 提取序列中丰富的上下文依赖。如图 1 所示, 该机制强化了模型对远距离依赖关系的学习能力, 显著提升了功率预测的准确性与鲁棒性。

Transformer 中包含自注意力机制使模型在处理序列数据时动态地关注输入序列中不同位置的信息, 以提高模型对输入序列数据的理解能力, 其表达式为:

$$\text{Attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{Softmax}\left(\frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{d_k}}\right)\mathbf{V} \quad (4)$$

式中: d_k 为键向量的维度; $\sqrt{d_k}$ 为缩放因子, 用于稳定梯度计算。

1.4 CNN-LSTM-Transformer 可解释预测模型

本文提出的模型分为两个主要部分: 左半部分为风机编号嵌入模块, 右半部分为主干网络与全连接预测层。

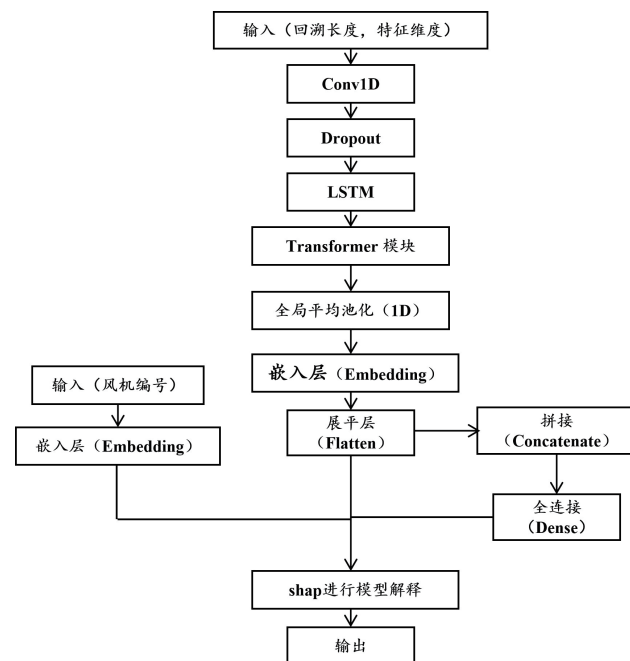


图 1 CNN-LSTM-Transformer 模型结构

左侧风机嵌入分支并行接收风机编号输入, 借助 Embedding 层将其映射为 10 维稠密向量, 以表达风机间的结构差异与性能差异。

主干网络首先对回溯时间窗口内的输入序列进行一维卷积运算, 提取局部时序特征并增强鲁棒性。随后, 卷积结果输入双层堆叠的 LSTM 模块, 通过门控机制建模风电功率的长期依赖关系。为进一步整合全局信息, LSTM 输出接入定制化 Transformer 编码器, 引入多头自注意力机制并行建模不同时间步之间的依赖, 强化对全局模式的感知与表达。该向量展平后与主干网络输出通过 Concatenate 融合, 再经全连接与 Dropout 层处理, 输出最终的功率预测值。

整体模型流程包括五个阶段: 数据预处理、局部特征提取、时序建模、全局融合、预测输出。数据层面采用正余弦时间编码与归一化处理构造标准输入; 模型层面采用串联 CNN-LSTM-Transformer 提取多层次特征, 并融合风机编号信息; 训练过程中结合 EarlyStopping 与动态学习率策略优化收敛性能^[17]。

为进一步提升模型的可解释性, 本文在模型训练完成后引入 SHAP (shapley additive explanations) 值分析方法, 从全局和局部两个维度对输入特征的重要性进行量化解释。可以更加直观展示各特征的影响程度, 为风电变频器功率预测提供了可信的因果解释支撑, 有效提升了模型在实际工程应用中的透明性与信赖度。

2 实验验证

2.1 数据集说明

本研究所用数据集来源于某海上风电场的 SCADA (supervisory control and data acquisition) 系统, 具体选取了海装 5 MW 级风电机组的历史运行数据作为研究对象。原始数据由该风电场监控平台实时采集, 包含每个机组在不同运行状态下的多维参数。

数据集经 K 近邻插补 ($K=6$) 填补缺失值, 标准化异常时间格式, 通过正余弦编码提取时间周期性, 标签编码设备编号。通过滑动窗口技术构建 59 步历史时序样本, 将多源异构数据整合为三维张量输入 (样本量 $\times$ 时间步 $\times$ 特征维度), 最终形成包含时序特征、设备标识符与工况参数的多模态数据集。由于风电数据集的各个特征存在显著差异, 需要对数据进行归一化处理, 用公式表示为:

$$x_{\text{scaled}} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (5)$$

2.2 评价指标

本试验采用误差评价指标为:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (\bar{y} - \hat{y}_i)^2} \quad (6)$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (7)$$

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (8)$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (9)$$

$$\text{SMAPE} = \frac{100\%}{n} \sum_{i=1}^n \frac{|\hat{y}_i - y_i|}{\frac{|\hat{y}_i| + |y_i|}{2}} \quad (10)$$

式中: MAE 为平均绝对误差 (mean absolute error); MSE 为均方误差 (mean squared error); RMSE 为均方根误差 (root mean squared error); SMAPE 为对称平均绝对百分比误差 (symmetric mean absolute percentage error)。

2.3 实验模型参数设计

本实验使用 Adam 优化器, 初始学习率 (learning_rate) 设置为 0.001。Adam 结合动量与自适应学习率, 在时序预测任务中具有收敛快速且稳定的优点。采用均方误差 (MSE) 作为训练损失, 适用于回归问题, 能够对较大偏差给出较高惩罚。本试验中需要调节的参数和最优参数值如表 1 所示。

表 1 参数值

参数名称	参数取值
回溯窗口长度 (look_back)	59
CNN 层数	2
LSTM 层数	2
Transformer 多头注意力头数	8
风机编号嵌入维度	10
融合全连接正则化系数 (L2)	0.001
学习率	0.001

训练集与验证集的损失变化曲线如图 2 所示, 横轴为训练轮次 (Epoch), 纵坐标为 Loss 值。由图可见, 训练初期设置下, 训练集损失值红线线性下降, 验证集损失值蓝线同步降低, 显示模型抽取数据特征的能力较强。随后训练进程中, 两者曲线逐渐收敛并保持小距离, 无明显振荡或变异, 表明模型未发生过拟合。L2 正则化和 Dropout 等防过拟策略在培训过程中起到了有效抑制作用, 提升了模型的泛化能力。总体来看, 所构建的 CNN-LSTM-Transformer 模型训练收敛较为稳定, 具有良好的稳定性, 为后续功率预测操作奠定了基础。

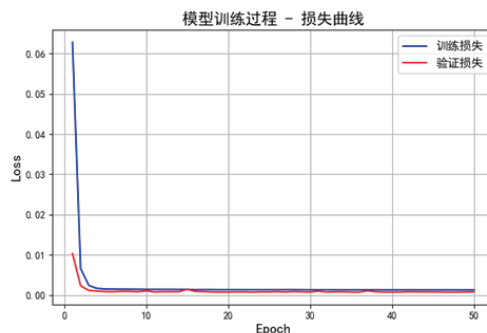


图 2 模型训练过程 - 损失曲线

2.4 实验结果分析

2.4.1 横向对比实验

为验证 CNN-LSTM-Transformer 模型的性能, 选取了多个具有代表性的时间序列预测模型作为对比。本文采用平均绝对误差 (MAE)、均方误差 (MSE)、均方根误差 (RMSE)、对称平均绝对百分比误差 (SMAPE) 与判定系数 (R^2) 作为评价指标, 对各模型在同一数据集上的预测表现进行全面比较, 如表 2 所示。

实验结果表明, 六种模型在风电机组变频器功率预测任务中表现各异, 其中 CNN-LSTM-Transformer 在所有评估指标上均优于其他模型。具体而言, TCN-Transformer 虽具备长程依赖建模能力, 但对局部特征的提取能力不足, 导致其 MAE 和 RMSE 相对较高; TCN-LSTM 通过引入 LSTM 增强了对时间序列的建模能力, 性能有所提升; CNN-LSTM-Attention 能够关注关键时刻信息, 但由于采用单头注意力机制, 对全局依赖的建模能力有限。

表 2 横向对比实验

预测模型	MAE	MSE	RMSE	SMAPE/%	R^2
TCN-Transformer	293.36	211 364.45	459.70	36.75	0.922 1
TCN-LSTM	278.364	200 125.45	447.35	35.16	0.934 8
CNN-LSTM-Attention	286.627	201 312.78	448.68	36.30	0.947 7
BiLSTM	272.00	192 560.00	438.82	34.70	0.951 0
CNN-GRU	260.50	186 500.00	431.85	34.36	0.958 0
CNN-LSTM-Transformer	250.77	162 998.47	403.73	33.83	0.964 6

BiLSTM 通过双向结构强化了时序信息的捕捉，在 RMSE 和 R^2 表现上具有一定优势，但缺乏特征提取模块限制了其整体性能；CNN-GRU 在轻量化结构下取得了较低的 SMAPE，但在 MSE 和 R^2 方面略逊于最佳模型。

相比之下，CNN-LSTM-Transformer 融合了多尺度卷积提取、双层 LSTM 门控机制以及多头自注意力机制，兼具局部与全局时序特征学习能力。在所有模型中，该模型实现了最低的 MAE (250.77)、MSE (162 998.47)、RMSE (403.73)、SMAPE (33.83%) 及最高的 R^2 (0.964 6)，验证了其在风电功率预测任务中的优越性。

2.4.2 消融实验

本研究通过系统性消融实验验证了 CNN-LSTM-Transformer 混合模型各模块的有效性，实验结果表明多模态融合策略在预测精度与特征表达能力上的优势，如表 3 所示。

消融实验表明，CNN、LSTM 与 Transformer 在风电机组变频器功率预测中均为关键，三者协同最优，缺一模块均导致性能下降。去除 Transformer 的 CNN-LSTM (MAE=266.03, R^2 =0.954 2)，虽能提局部特征、做基础时序建模，但缺乏全局注意力，难捕长期趋势；移除 LSTM 的 CNN-Transformer (MAE=271.67, R^2 =0.952 1)，虽融卷积与自注意力，却中程依赖建模不足；去除 CNN 的 LSTM-Transformer (MAE=259.29, R^2 =0.955 1)，虽能捕全局时序信息，但对局部突变感知弱。

完整模型 (MAE=250.77, R^2 =0.964 6) 借 CNN 提升多尺度局部特征、LSTM 建长短期依赖、Transformer 强全局建模，实现优势互补，提升对复杂功率变化的感知力。综上所述，实验验证模块互补性，完整架构精度与稳健性最优。

如图 3 所示，预测值和真实值曲线基本重合，显示出紧密的拟合效果，直观地反映了模型对功率波动趋势的准确捕捉能力。该效果与 R^2 等综合指标相互印证，进一步验证了模型的预测性能优势，为上述比较分析提供了有效的可视化支撑。

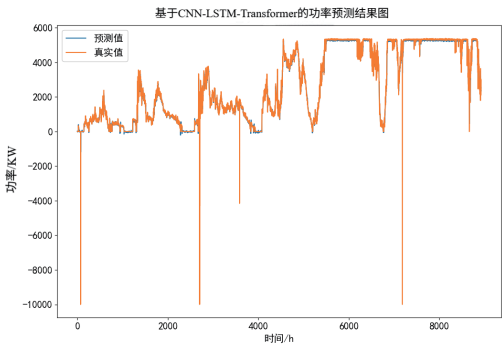


图 3 功率预测结果图

2.5 可解释性分析

在可解释性分析阶段，本文基于 SHAP 方法对 CNN-LSTM-Transformer 模型进行定量剖析，以揭示各输入特征对风电变频器功率预测的影响程度及方向。SHAP 值可有效衡量单一特征对预测结果的边际贡献，明确其正向或负向作用，直观反映模型的决策机制。

如图 4 所示，“变频器 1 网测有功功率”是对预测结果影响最大的特征，其平均 SHAP 值绝对值较高，进一步验证了其功率输出描述的核心关键。第二级特征包括“变频器机侧电流”“变频器 2 发电机转矩”和“变频器 2 机侧电流”等，这些特征和变频器运行过程中的功率和转动特性密切相关，有力支撑模型评估合理性。

表 3 消融实验

预测模型	MAE	MSE	RMSE	SMAPE/%	R^2
CNN-LSTM	266.03	204 964.42	452.72	34.63	0.954 2
CNN-Transformer	271.67	214 477.35	463.11	36.82	0.952 1
LSTM-Transformer	259.29	200 873.28	448.18	34.38	0.955 1
CNN-LSTM-Transformer	250.77	162 998.47	403.73	33.83	0.964 6

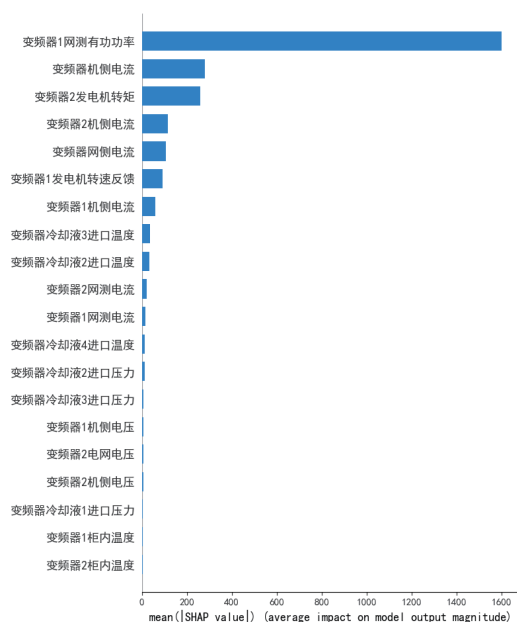


图4 全局 SHAP 特征重要性条形图

3 结论

在“双碳”目标下,风电稳定运行与功率预测对能源结构优化、电网调度意义重大。本文针对风电机组变频器功率预测,构建融合 CNN、LSTM 与 Transformer 的混合神经网络,引入风机编号嵌入表征设备差异,结合正则化与 Dropout 提升鲁棒性。横向对比实验中,该模型 MAE 达 250.77、 R^2 为 0.964 6,显著优于 TCN、BiLSTM 等模型;消融实验证实三模块互补且关键。通过 SHAP 分析量化特征影响,明确“变频器 1 网测有功功率”等关键特征,增强模型透明度。模型表现优异,未来可拓展至多风场、多气象区域等场景。

参考文献:

- [1] 龚佳辉,王磊.“双碳”目标下风电设备及技术发展分析[J].中国设备工程,2025(9):221-223.
- [2] 强涵悦,任惠,于光发,等.基于区域迁移学习与动态图网络建模的超短期海上风电功率预测方法[J].电力科学与工程,2025,41(6):19-29.
- [3] Global wind energy council: China leads global growth in offshore wind power for third consecutive years[R].Brussels: Global Wind Energy Council, 2021.
- [4] 罗敏,程将南,王毅,等.基于小波去噪和决策树的个性化大用户负荷预测[J].南方电网技术,2016,10(10):37-42.
- [5] 孙海蓉,曹瑶佳,张雨晴.基于 WKFCM-SMOTE 和随机森林的风电机组故障诊断[J].山东电力技术,2022,49(3):65-70.
- [6] 杜尊峰,邵玄玄,王晓梅.基于 CP 结合 DE-GWO-SVR 的

海上风电基础结构损伤识别[J].振动与冲击,2020,39(22):110-118.

- [7] 苏国梁,汪健冬,付恩强,等.XGBoost 算法在风电机组发电机故障监测预警中的应用研究[J].太阳能,2021(9):78-84.
- [8] 赵娟娟,刘广臣,王瑞桃,等.基于 LightGBM 的风电机组齿轮箱油温故障预警研究[J].电力大数据,2021,24(11):76-84.
- [9] 刘冰冰,王安琪,张传江,等.基于随机森林的风电变频器故障预警模型研究与应用[J].船舶工程,2020,42(S2):181-186.
- [10] 李森娟,张萍,岳大为,等.基于支持向量机的风电机组故障预测[J].计算机仿真,2022,39(5):84-88.
- [11] REN H, XIA J, TANG H X, et al. Active learning-based wind turbine failure prediction with consideration of data from different wind farms[J]. Journal of physics:conference series, 2023, 2427: 012044.
- [12] 张西晨,徐璐.基于边缘智能的风电机组设备监测系统[J].电子测量技术,2023,46(2):52-58.
- [13] 徐武,范鑫豪,沈智方,等.多尺度特征提取的 Transformer 短期风电功率预测[J].太阳能学报,2025,46(2):640-648.
- [14] CAI L, JANOWICZ K, MAI G C, et al. Traffic transformer: capturing the continuity and periodicity of time series for traffic forecasting[J]. Transactions in GIS, 2020, 24(3):736-755.
- [15] WANG B Y, CHEN Z Y, ZHANG P H, et al. The lithium-ion battery temperature field prediction model based on CNN-Bi-LSTM-AM[J]. Sustainability, 2025, 17(5):2125.
- [16] 向玲,高鑫,姚青陶,等.基于视觉 Transformer 多模型融合的风电机组异常状态监测[J].太阳能学报,2025,46(4):522-529.
- [17] 苏向敬,程子凡,聂良钊,等.基于 AGCN-LSTM 模型的海上风电场功率概率预测[J].电力系统自动化,2024,48(22):140-149.

【作者简介】

罗威(2001—),男,湖北黄冈人,硕士研究生,研究方向:智能预测, email:15571393558@163.com。

(收稿日期:2025-09-16 修回日期:2026-01-08)