

引用格式:蔡梓文,谈竹奎,赵云,等.基于 Multiformer-TSA 的光伏发电功率预测方法[J].电力科学与技术学报,2026,41(1):130-139.

Citation: CAI Ziwen, TAN Zhukui, ZHAO Yun, et al. Multiformer-TSA-based photovoltaic power forecasting method[J]. Journal of Electric Power Science and Technology, 2026, 41(1): 130-139.

基于 Multiformer-TSA 的光伏发电功率预测方法

蔡梓文^{1,2}, 谈竹奎³, 赵云^{1,2}, 张钥朗³, 刘熙鹏^{1,2}, 张后谊³

(1. 南方电网科学研究院有限责任公司, 广东 广州 510530; 2. 广东省电网智能量测与先进计量企业重点实验室, 广东 广州 510530;
3. 贵州电网有限责任公司电力科学研究院, 贵州 贵阳 550002)

摘要:针对当前光伏发电功率预测方法高度依赖气象监测与分类技术,无法实现大规模复杂数据下光伏发电全天候准确预测的问题,构建了基于两阶段注意力机制(two-stage attention, TSA)的 Multiformer-TSA 方法来预测光伏发电功率。首先,构建跨尺度嵌入层,生成分阶段采样标记,采集不同尺度光伏序列,获取跨尺度特征;然后,将多变量时间序列不同维度的点分段嵌入,组成新的特征向量,捕捉跨维度依赖性;最后,通过 TSA 将跨尺度与跨维度依赖信息融合,以实现准确的光伏发电全天候预测。利用澳大利亚爱丽斯泉的光伏发电公开数据集开展多尺度预测对比实验与消融实验,实验结果表明,所提方法能准确地捕捉多维时间序列的跨尺度与跨维度特征,提高光伏发电功率多尺度预测精度。与现有方法相比,预测结果的均方根误差和平均绝对误差指标皆表现最佳。

关键词: 光伏预测; 多变量时间序列; 跨维度特征; 跨尺度特征; 两阶段注意力机制

DOI: 10.19781/j.issn.1673-9140.2026.01.013 中图分类号: TM73; TK01 文章编号: 1673-9140(2026)01-0130-10

Multiformer-TSA-based photovoltaic power forecasting method

CAI Ziwen^{1,2}, TAN Zhukui³, ZHAO Yun^{1,2}, ZHANG Yuelang³, LIU Xipeng^{1,2}, ZHANG Houyi³

(1. CSG Electric Power Research Institute Co., Ltd., Guangzhou 510530, China; 2. Guangdong Provincial Key Laboratory of Intelligent Measurement and Advanced Metering of Power Grid, Guangzhou 510530, China; 3. Electric Power Research Institute, Guizhou Power Grid Co., Ltd., Guiyang 550002, China)

Abstract: To address the limitations of current photovoltaic power forecasting methods, which rely heavily on meteorological monitoring and classification techniques and are unable to achieve accurate all-weather forecasting under large-scale complex data, a Multiformer-TSA method based on a two-stage attention (TSA) mechanism is proposed for photovoltaic power forecasting. First, a cross-scale embedding layer is constructed to generate staged sampling tokens and collect photovoltaic sequences at different scales, thus extracting cross-scale features. Then, point-based segments of different dimensions in multivariate time series are embedded to form new feature vectors, capturing cross-dimensional dependencies. Finally, cross-scale and cross-dimensional dependency information is fused through the TSA mechanism to achieve accurate all-weather photovoltaic power forecasting. Multi-scale forecasting comparison experiments and ablation experiments are conducted on a publicly available photovoltaic power dataset from Alice Springs, Australia. The experimental results demonstrate that the proposed method accurately captures cross-scale and cross-dimensional features of multivariate time series and improves the multi-scale forecasting accuracy of photovoltaic power generation. Compared with existing methods, the proposed method achieves the best performance in terms of root mean square error (RMSE) and mean absolute error (MAE).

Key words: photovoltaic forecasting; multivariate time series; cross-dimensional feature; cross-scale feature; two-stage attention mechanism

随着“双碳”目标的持续推进,清洁能源占社会总能源比重日益增长^[1]。光伏发电作为清洁能源的

重要组成部分,受太阳辐射周期、地理环境及各种气象因素变化的影响,具有强波动性及随机性^[2-5]。

收稿日期: 2025-01-04; 修回日期: 2025-03-28

基金项目: 南方电网重点科技项目(GZKJXM20222127, GZKJXM20222133)

通信作者: 蔡梓文(1991—), 男, 硕士, 工程师, 主要从事新型电力系统绿色低碳用能等研究; E-mail: caizw@csg.cn

在大规模光伏并网背景下,提高光伏发电预测准确率对电网的平稳调度及电价制定至关重要^[6-10]。

天气变化是影响光伏发电最直接的因素,因此许多研究聚焦提升不同天气下光伏发电预测的准确率。文献[11]在提取关键气象要素的基础上,提出了改进模糊 C 均值聚类算法自主确定聚类数的方法,结合长短期记忆(long short-term memory, LSTM)神经网络,提高了光伏预测准确率。文献[12]使用平衡迭代还原和分层聚类(balanced iterative reducing and clustering using hierarchies, BIRCH)算法得到了 3 种典型天气,根据聚类结果划分测试集,对模型进行训练,提升了不同天气下的预测精度。文献[13]通过逐时计算历史日和待预测日的加权欧氏距离确定相似日,分别对功率和误差进行预测,最后将二者相加获得较高的光伏发电功率预测精度。文献[14]提出基于自适应混合核的相关向量机方法,捕捉光伏发电随机波动规律,通过历史相似日数据确定最优预测模型,提高非晴空天气的光伏功率短期预测准确率。然而,上述方法高度依赖历史数据的完整性和准确性来完成相似日的准确聚类,且多侧重局部特征的提取,难以处理好数据全局关系。

Transformer 模型基于注意力机制,在处理大规模复杂数据、捕捉长时依赖关系和提高计算效率方面具有明显优势^[15-16]。文献[17]构建了基于时序卷积网络(time-series convolutional network, TCN)的 TCN-Transformer 模型,实现了高精度的光伏功率预测。文献[18]提出了基于 Transformer 的增强型卷积变换器网络(convolutional-transformer-based network, CT-NET),同时增强对局部特征和上下文特征的提取能力,提高了光伏预测精度。文献[19]利用深度融合技术和多任务联合学习框架,构建了 TransPVP 模型捕捉光伏发电的长期依赖关系,提升了模型性能。文献[20]基于注意力机制的架构,提出时空融合 Transformer(temporal fusion Transformer, TFT)方法,对时序动态进行解析,在多个预测时段提供高性能预测。然而,上述研究侧重于对数据时序特征变换规律的捕捉,缺少对多维数据特征之间依赖关系的关注,在多变量时间序列(multivariate time series, MTS)下,预测模型的性能有待提高。

据此,本文在 Transformer 模型基础上,提出基于两阶段注意力机制(two-stage attention, TSA)的新型光伏发电功率预测方法 Multiformer-TSA,通

过在每一维度中对不同时间尺度的特征进行采样,获取跨尺度特征,再将不同维度序列重新分段嵌入,获得新的特征向量,捕捉跨维度依赖关系;设计 TSA,将多尺度特征信息进行融合以开展预测,并通过消融实验和对比实验对本文所提方法的有效性和优越性进行验证与分析。

1 数据准备

本文以澳大利亚沙漠知识太阳能中心(Desert Knowledge Australia Solar Center, DKASC)的光伏发电公开数据集为例开展相关研究。该数据集收集了多年度、多站点的光伏发电数据及对应站点所处位置的天气变量,采样间隔为 5 min。为优化数据输入,降低无效信息对模型预测性能的影响并减少计算量,本文对光伏发电原始数据进行了相关性分析。

1.1 光伏发电功率的影响因素

光伏发电功率受多种因素影响,本文所用数据集中包含风速、环境温度、相对湿度、风向、降水量、全球水平辐射和漫射水平辐射等数据。本文从跨尺度与跨维度思想出发,捕捉数据间的依赖关系。原数据中对光伏发电功率影响偏小的因素会引入弱相关特征,影响模型的预测性能,由此需要对原始多维数据进行相关性分析。本文使用皮尔逊相关性系数计算各输入数据间的相关性,皮尔逊相关性系数热力图如图 1 所示。

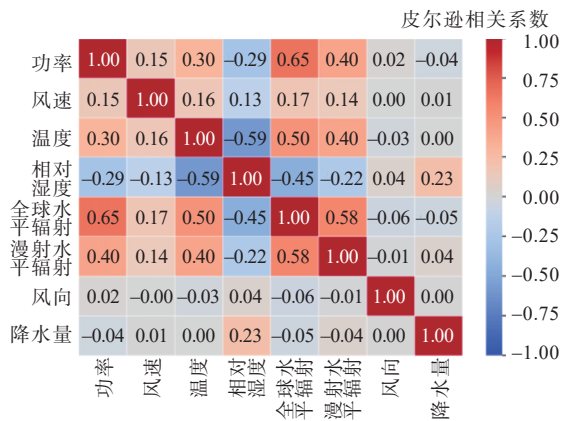


图 1 皮尔逊相关性系数热力图

Figure 1 Heat map of Pearson's correlation coefficient

由图 1 可知,在原始数据中,相对湿度与降水量对光伏发电功率的影响较小,故本文采用光伏发电有功功率、风速、环境温度、风向、全球水平辐射和漫射水平辐射作为输入数据。

1.2 数据预处理

光伏发电数据采集受气象条件及数据采集设

备性能的影响,在数据收集和传输过程中可能会出现缺失值与异常值,影响模型预测性能。为提升数据质量,使预测结果更加真实可信,本文对数据进行了预处理。

1) 缺失值填充。

传感器故障、信号传输中断或采集设备维护等因素会产生缺失值,使模型无法捕捉关键时间段内的趋势信息,影响时序依赖关系的学习。处理缺失值的方法有直接删除法、就近补齐法、线性插值法等。直接删除数据可能会导致有效信息丢失、遗漏关键信息,而就近补齐方法定义相似标准比较困难,故本文采用线性插值方法来进行缺失值的填补,计算如下:

$$x_i = x_j + \frac{x_k - x_j}{k - j} \cdot (i - j) \quad (1)$$

式中, x_i 为缺失点插值的值; x_j, x_k 为缺失点前、后最近的非缺失值; i, j, k 分别为缺失点插值和缺失点前、后最近的非缺失值的位置序号。

2) 异常值替换。

光伏发电数据中的异常值通常由极端天气条件、设备误差或数据录入问题引起,可能使模型过度关注不具代表性的极端数据,导致预测结果的波动性增大。异常值在整个功率时间序列中通常具有离散性,表现为与相邻数据有较大差异,故本文通过 3σ 定律来进行异常值的替换,计算如下:

$$x' = \begin{cases} \mu - 3\sigma, & x < \mu - 3\sigma \\ \mu + 3\sigma, & x > \mu + 3\sigma \end{cases} \quad (2)$$

式中, x 为原始值; x' 为替换后的值; μ 为样本均值; σ 为样本标准差。

3) 数据归一化。

光伏发电数据的特征值通常具有不同的单位和数值范围,会导致模型在训练时对某些特征的权重分配不均,预测性能下降。本文采用min-max归一化方法,将各特征值映射到相同的范围,统一数据尺度,如下:

$$x_n = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (3)$$

式中, x_n 为归一化后的值; x_{\min}, x_{\max} 分别为该特征的最小值和最大值。

2 Multiformer-TSA 光伏预测方法

2.1 Transformer 模型

光伏发电功率预测的本质为MTS预测,而Transformer模型具备强大的注意力机制,能有效处

理光伏发电功率等长时间序列数据,尤其适用于捕捉由日照、天气等外部因素引起的发电功率波动^[21-23]。Transformer的主要架构包括编码器和解码器两个部分,其模型结构如图2所示。

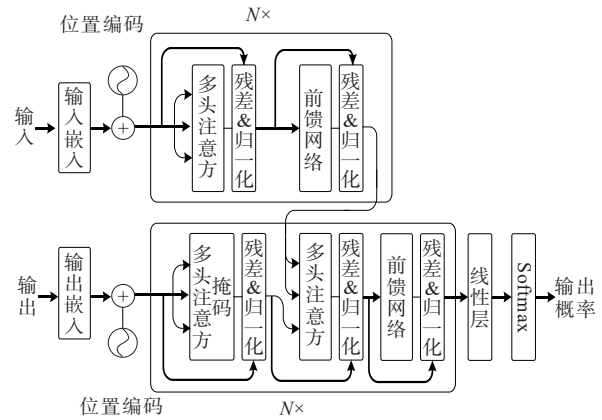


图2 Transformer模型结构

Figure 2 Transformer model structure

由图2可知,Transformer编码器由多个堆叠的层组成,每一层包含一个多头自注意力(multi-head self-attention, MHA)模块和一个前馈神经网络(feedforward neural network, FNN)模块。编码器采用正弦和余弦函数生成不同频率的编码,从而在光伏发电功率时间序列的长度增加时,仍能识别出数据中的相对和绝对位置信息。

然而,Transformer在处理多维信息时仍无法实现全局依赖关系的有效捕捉,为弥补现有模型对MTS维度内及维度间依赖关系捕捉能力不足的缺陷,本文通过提取跨尺度相关性与跨维度依赖性,构建Multiformer模型,提高预测性能。

2.2 光伏多维时间序列依赖信息捕捉方法

2.2.1 构建光伏数据跨维度分段层

光伏发电预测的目标是在给定历史数据的情况下,得到期望的时间序列。现有FNN、递归神经网络(recurrent neural network, RNN)、卷积神经网络(convolutional neural network, CNN)、Transformer等方法可捕捉数据在长时或短时维度上的特征依赖,如图3所示。

由图3可知,现有方法通常将光伏发电功率时间序列中同一时间步长的数据点嵌入一个向量 H 中,通过分析向量 H 内的跨时间依赖关系进行预测。这些方法虽有效捕捉了跨时间依赖性,但忽视了数据中外部因素与光伏发电功率之间的复杂关系,导致光伏发电功率的预测精度较低。为解决这一问题,本文构建了跨维度分段层(cross-dimensional

segmentation layer, CSL),以捕捉光伏发电数据中不同维度特征间的依赖关系,增强了预测模型的精度,跨维度依赖性捕捉原理如图 4 所示。

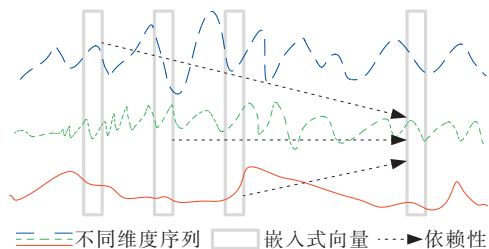


图 3 捕捉跨时间依赖性

Figure 3 Capturing cross-time dependency

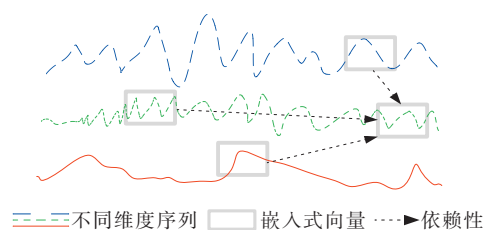


图 4 捕捉跨维度依赖性

Figure 4 Capturing cross-dimensional dependency

由图 4 可知,在光伏发电功率预测中,光伏发电功率不仅受历史时间序列数据影响,还与环境温度、风速、全球水平辐射等多维度的外部变量存在复杂的交互关系。因此,跨维度依赖性捕捉的核心目标是重新分段和嵌入不同时间步的多维特征,构建统一的特征向量,显示建模变量之间的交互特性。为有效捕捉光伏发电功率时间序列中的跨维度依赖信息,本文通过把输入序列按维度分段嵌入,将光伏发电相关的多维度特征嵌入一个统一的向量空间中,在每个采样片段进行线性变换,并添加时间位置编码,生成特征向量。最终,将所有维度的特征进行拼接,形成跨维度的二维特征矩阵,供后续注意力机制进行深层次建模。不同维度的特征会映射到相同的特征空间,不仅保留了各自维度内的信息,还包含了跨维度的依赖关系,从而增强了模型对不同外部条件下功率变化的理解,提高了光伏发电功率预测的精度和鲁棒性。具体操作流程如下。

1)对时间序列进行采样:

$$\chi_{1:T} = \left\{ \chi_{i,d} \mid 1 \leq i \leq \frac{T}{L}, 1 \leq d \leq D \right\} \quad (4)$$

式中, $\chi_{1:T}$ 为以 T 为总时长的光伏发电功率时间序列; L 为采样时间步; D 为光伏发电数据的维度数量; $\chi_{i,d}$ 表示第 d 个维度中被采样的第 i 个片段。

2)使用线性投影法将每个线段嵌入一个矢量

中,并加上位置嵌入,计算如下:

$$a_{i,d} = M\chi_{i,d} + M_{i,d}^{(\text{pos})} \quad (5)$$

式中, $a_{i,d}$ 为一个单变量光伏时间序列片段; M 为可学习的投影矩阵; $M_{i,d}^{(\text{pos})}$ 为在位置 (i,d) 的嵌入。

3)通过位置嵌入可得:

$$A = \left\{ a_{i,d} \mid 1 \leq i \leq \frac{T}{L}, 1 \leq d \leq D \right\} \quad (6)$$

式中, A 为 $a_{i,d}$ 组成的二维矢量数组。

由此,CSL 可将时间序列跨维度依赖关系嵌入同一个数组中,增强了模型对不同时间段光伏功率特征的定位能力。

2.2.2 光伏数据跨尺度扩散层构建

为有效捕捉光伏发电功率时间序列中同一维度数据的时序依赖特征,本文构建了跨尺度扩散层(cross-scale diffusion layer, CDL),对不同时间尺度的功率变化进行扩散采样,识别出光伏发电功率随时间变化的规律,并为每个阶段生成标签,增强了模型对发电功率动态变化的理解和预测能力,CDL 结构如图 5 所示。

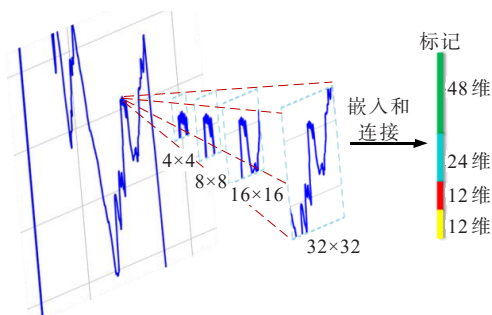


图 5 CDL 结构

Figure 5 CDL structure

由图 5 可知,CDL 采用长短距离注意力(long-short distance attention, LSDA)机制,捕捉光伏发电功率时间序列中的多尺度特征。其中,短距离注意力能够细化时序数据中的局部变化,防止模型在短时间波动下失真,长距离注意力则提供了数据全局视角,有助于捕捉周期性趋势变化,提高长期预测的准确性。在多尺度卷积核中,小卷积核有效提取了短时间波动特征,提高了模型在晴天条件下的预测精度;大卷积核捕捉了长期变化模式,使模型在复杂天气场景下依然具备较高稳定性。LSDA 机制自适应地调整了不同尺度特征的重要性,确保了在不同天气场景和时间段中,模型能够自动选择合适的特征进行预测。每 4 个对应的数据片段具有相同的采样中心和不同的采样尺度,能够从不同尺度捕

捉光伏发电功率的变化模式。这4个数据片段会被嵌入并连接为一个统一的嵌入向量,用来捕捉功率数据在不同时间跨度上的依赖性。采样和嵌入过程通过4个卷积层来完成,在光伏发电功率预测中,直接使用不规则采样方式可能导致对特定时间尺度的过度关注。等分采样确保了每个尺度的特征被均衡采样,减少了模型对单一尺度特征的依赖。通过多尺度融合,模型能在多种复杂场景下保持较高的预测精度。本文以等分的方式分配每个尺度的嵌入维度,从而更高效地捕捉不同时间尺度上的发电功率特征,其卷积层计算成本为

$$C_{\text{cm}} = \lambda K^2 D^2 \quad (7)$$

式中, C_{cm} 为卷积层的计算成本; λ 为比例系数; K 、 D 分别为采样核尺寸和输入/输出维度。在光伏发电功率预测场景中,本文对大内核使用较低的维度,对小内核使用较高的维度,这样不仅能够提升处理光伏发电功率数据的效率,还可以在保持高精度的情况下减少资源消耗,适用于预测实时光伏功率。

CSL在每个时间步上整合不同维度的输入特征,使得CDL在进行多尺度采样时能够捕捉不同变量对光伏功率变化的综合影响,从而避免了单一尺度特征提取的片面性。CDL生成的多尺度嵌入向量为CSL提供了更丰富的上下文信息,使CSL能够更准确地捕捉多变量特征之间的依赖关系,从而提升模型在复杂场景下的鲁棒性。由此,CSL和CDL通过联合捕捉时间序列中的多维特征交互和多尺度时序模式,为Multiformer-TSA模型提供了更全面的输入特征表示。

2.3 基于Multiformer-TSA网络的光伏预测模型

为有效融合CSL与CDL获取的跨维度和跨尺度依赖信息,本文构建了TSA,融合了光伏发电功率时间序列中不同维度特征以及不同时间尺度上的依赖关系。

1) 跨时间阶段。

该阶段的目的是在光伏发电功率预测中,针对每个维度的数据序列,应用自注意力机制,捕捉同一维度下不同时间步之间的依赖关系,从而有效识别光伏发电功率在不同时间点的变化模式,帮助模型更准确地预测未来的功率输出,其计算如下:

$$\begin{cases} \hat{Z}_{\text{time},:,d} = \text{LN}(Z_{:,d} + \text{MHA}(Z_{:,d}, Z_{:,d}, Z_{:,d})) \\ Z_{\text{time}} = \text{LN}(\hat{Z}_{\text{time}} + \text{MLP}(\hat{Z}_{\text{time}})) \end{cases} \quad (8)$$

式中, $\hat{Z}_{\text{time},:,d}$ 为在时间步 L 和维度 d 上应用自注意力机制的结果,用于捕捉光伏发电功率在该维度下

不同时间步之间的依赖关系; $Z_{:,d}$ 为第 d 个维度上所有时间步的数据,代表光伏发电功率随时间变化的特征;MLP(\cdot)代表多层感知机(multi-layer perceptron, MLP)操作;MHA(\cdot)代表多头注意力机制(multi-head attention, MHA);LN(\cdot)代表层归一化(layer normalization, LN); Z_{time} 为时间阶段的输出; \hat{Z}_{time} 为应用自注意力机制的结果。

2) 跨维度阶段。

在跨维度阶段,模型通过跨不同维度的自注意力机制,捕捉光伏发电系统中各个维度特征之间的依赖关系,从而更好地理解不同外部环境因素对光伏发电功率的综合影响,提升模型功率预测的准确性和鲁棒性:

$$\begin{cases} B_{i,:} = \text{MHA}_{\text{dim}1}(R_{i,:}, Z_{\text{time},i,:}, Z_{\text{time},i,:}) \\ Z_{\text{dim},i,:} = \text{MHA}_{\text{dim}2}(Z_{\text{time},i,:}, B_{i,:}, B_{i,:}) \\ \hat{Z}_{\text{dim}} = \text{LN}(Z_{\text{time}} + Z_{\text{dim}}) \\ Z_{\text{dim}} = \text{LN}(\hat{Z}_{\text{dim}} + \text{MLP}(\hat{Z}_{\text{dim}})) \end{cases} \quad (9)$$

式中, \hat{Z}_{dim} 为结合跨时间注意力和跨维度注意力的输出; $B_{i,:}$ 为维度 i 上经过第一层跨维度注意力机制后的输出; $R_{i,:}$ 为维度 i 的查询项; $Z_{\text{time},i,:}$ 为通过跨时间阶段处理后维度 i 的表示; $Z_{\text{dim},i,:}$ 为维度 i 上经过第二层跨维度注意力机制的输出; Z_{dim} 为通过MLP进一步处理后的跨维度结果;MHA_{dim1}为第一层跨维度注意力机制。结合跨时间与跨维度的注意力机制,模型能够同时捕捉光伏发电功率时间序列中的时序依赖关系和各环境特征之间的相互影响,从而提升光伏发电功率预测的整体性能。跨时间阶段和跨维度阶段通过级联融合,逐步提升了特征的表达能力。最终输出的融合特征通过MLP生成预测结果,这种级联方式能够平衡时间和维度信息的权重,使模型在不同预测尺度下都具备良好的泛化能力。据此,本文提出Multiformer-TSA方法,其网络结构如图6所示。

由图6可知,Multiformer-TSA方法首先在CDL中利用多个不同尺度的卷积核对光伏历史功率数据进行采样,使每个特征包含来自不同时间尺度的信息,采集跨尺度特征,提升对光伏功率波动的捕捉能力。在CSL中,将光伏功率多维时间序列按维度划分,在多个时间段进行采样,并将每个时间段嵌入特征向量,保留时间和维度信息,构建出更全面的输入特征表示。然后,采用TSA和LSDA捕捉时间序列中的跨尺度和跨维度依赖关系。在后续功率预测中,TSA分为跨时间和跨维度两个阶段,

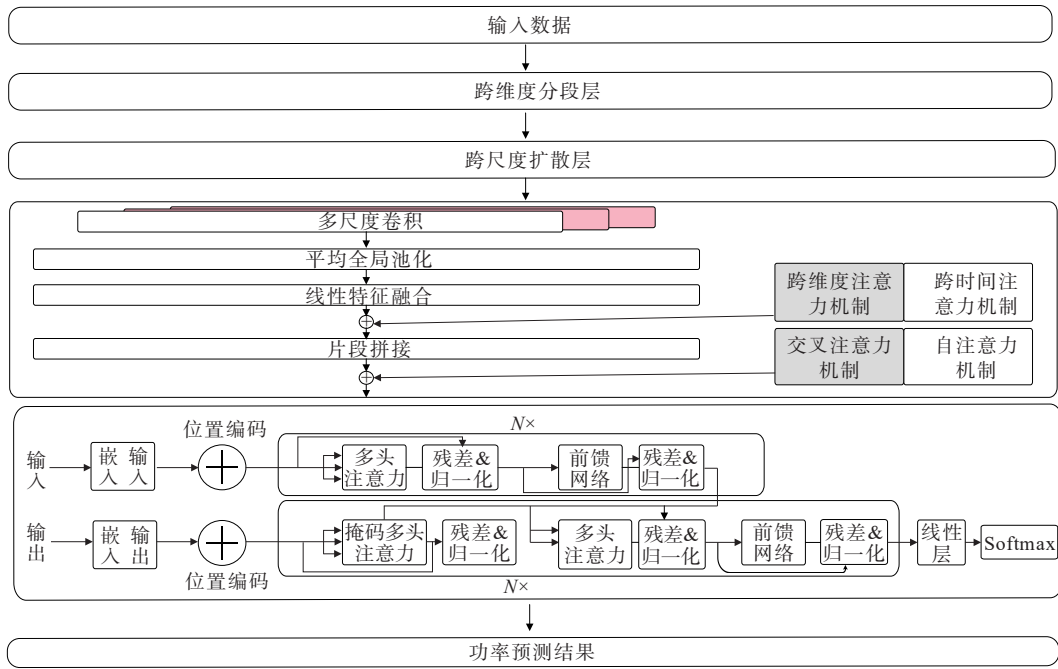


图 6 Multiformer-TSA 网络结构
Figure 6 Structure of Multiformer-TSA network

分别用于捕捉光伏发电功率的时间趋势依赖性和不同环境变量之间的相互关系,从而增强模型的预测能力。LSDA 通过将注意力机制按距离划分,分别捕捉相邻时刻和远距离时刻之间的依赖关系,使得模型在短期和长期功率预测中都具备良好的表现。最后采用分层编码器-解码器架构,逐步整合跨尺度和跨维度的信息,生成最终的光伏功率预测结果。综上所述,本文提出基于 Multiformer-TSA 的光伏发电功率预测方法的流程如图 7 所示。

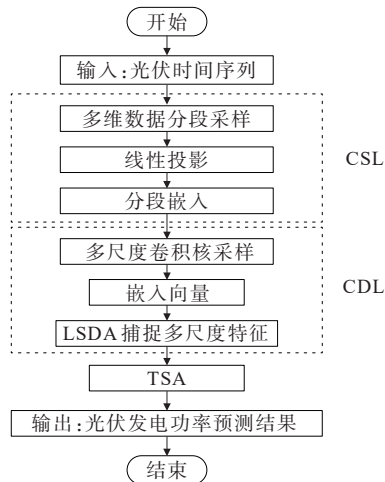


图 7 基于 Multiformer-TSA 的光伏发电功率预测方法的流程

Figure 7 Process of photovoltaic power forecasting method based on Multiformer-TSA

3 算例分析

本文采用澳大利亚爱丽斯泉的公开光伏发电数据集 DKASC-project-78-Site_DKA-M11_3-Phase 进行仿真,取其中 2022 年 6 月 1 日—9 月 7 日共 99 d 的相关数据,设置采样间隔为 5 min,并选择光伏发电有功功率、风速、环境温度、风向、全球水平辐射和漫射水平辐射这 6 个维度进行分析,取 2022 年 6 月 1 日—8 月 31 日的相关数据作为训练集,2022 年 9 月 1 日—9 月 7 日的相关数据作为测试集。

3.1 实验设置

本文实验均在 Python 3.8 环境下进行,模型采用 Pytorch 框架实现,使用的计算机硬件配置为 13th Gen Intel (R) Core (TM) i5-13400F CPU, NVIDIA GeForce RTX 4060 GPU, 32 GB RAM, 操作系统为 Windows 11。

本文所提方法的关键参数设置如下:输入特征维度为 6,输入序列长度为 12,输出序列长度为 12,采样长度为 12,缩放因子为 10,模型维度为 512,前馈网络维度为 1 024,注意力头数为 8,编码层数为 3,Dropout 概率为 0.1。模型初次训练采用均方误差损失函数,训练轮数为 180,优化器采用 Adam,学习率设置为 0.000 1。训练时每批次大小为 512。为验证 Multiformer-TSA 方法的优越性,本文选取 7 种现有方法(CNN、FNN、LSTM、Informer、TCN-

Transformer、TransPVP、CT-NET) 与该方法进行对比,对比模型超参数设置如表 1 所示。

表 1 对比模型超参数设置

模型	重要超参数	参数设置
CNN	卷积层个数	2
	卷积核大小	3
	激活函数	ReLU
	RNN层数	2
FNN	隐藏层层数	128
	LSTM层数	2
LSTM	Batch_size	512
	采样长度	12
	窗口大小	4
Informer	缩放因子	10
	Dropout 概率	0.1
	通道配置	[64, 128], [128, 256], [256, 512]
	卷积核大小	2
TCN-Transformer	Dropout 概率	0.2
	隐藏层层数	[64, 256]
	注意力头数	[1, 8]
	前馈网络维度	512
TransPVP	隐藏层层数	128
	注意力头数	4
	编码器层数	2
	前馈网络维度	512
CT-NET	Dropout 概率	0.1
	编码器个数	3
	注意力头数	3
	头大小	4
	前馈网络维度	32
	全连接层输出节点数	[832, 128, 64]

3.2 模型评估指标

本文评估所有模型的预测结果所采用的评估指标计算如下:

$$e_{\text{RMSE}} = \sqrt{\frac{1}{n} \sum_{t=1}^n (\tilde{y}_t - y_t)^2} \quad (10)$$

$$e_{\text{MAE}} = \frac{1}{n} \sum_{t=1}^n |\tilde{y}_t - y_t| \quad (11)$$

式中, e_{RMSE} 为均方根误差 (root mean square error, RMSE); e_{MAE} 为平均绝对误差 (mean absolute error, MAE); n 为预测样本总数; \tilde{y}_t 、 y_t 分别为光伏发电功率在时刻 t 的预测值和实际值。

3.3 不同模型对比分析

为验证 Multiformer-TSA 方法的优越性, 本文将其与现有方法进行对比并分析, 不同模型预测全天候光伏发电功率的结果如图 8 所示。

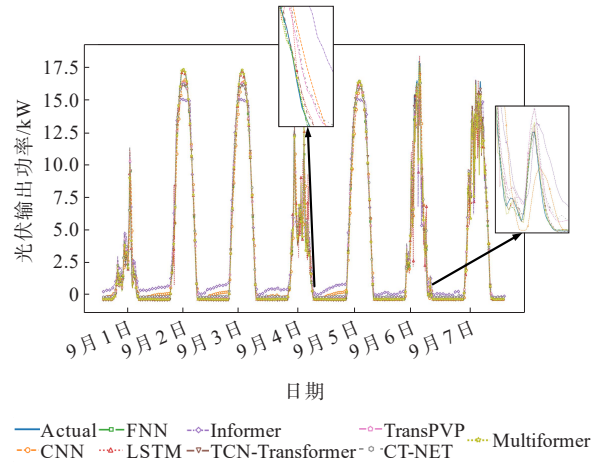


图 8 不同模型预测全天候光伏发电功率结果

Figure 8 All-weather photovoltaic power forecasting results of different models

由图 8 可知, CNN、FNN、LSTM 在面对复杂的时间序列数据时, 易受训练数据的局部模式影响, 在处理快速变化的光伏发电功率时, 存在响应延迟的问题, 预测误差较大。而基于 Transformer 改进的模型, 在处理光伏发电数据多维度特征和多尺度问题时, 未能充分挖掘特征的深度依赖关系, 降低了复杂环境下的预测精度。

基于 Multiformer-TSA 的光伏发电功率预测方法通过跨维度和跨尺度的特征融合, 避免了对突发变化响应滞后的问题, 其预测曲线能较好地逼近真实值, 尤其在发电功率波动较大的时段, 预测值能迅速响应实际值的变化, 验证了该方法能够准确捕捉短期内的突发波动和长期的变化趋势, 体现了较好的动态响应能力, 实现了对真实值的最高拟合。而其他对比模型在功率变化剧烈时段, 则未能及时跟踪实际发电功率波动。

综上所述, 本文所提方法凭借其多维度特征融合和跨尺度处理能力, 在 7 d 的全天候光伏发电功率预测中, 能够准确反映实际波动, 解决了其他模型在动态变化、长期依赖和多尺度特征处理上的不足, 表现出了更优的预测准确性和稳定性。

为进一步证明所提方法的预测性能, 本文进行了不同尺度的预测对比实验, 即输出序列长度分别为 12、24、72、96、144、288, 不同预测长度下的 RMSE、MAE 结果如表 2 所示。

表 2 不同预测长度下的 RMSE、MAE 结果
Table 2 RMSE and MAE with different forecasting lengths

模型	不同预测长度下的 RMSE						不同预测长度下的 MAE					
	12	24	72	96	144	288	12	24	72	96	144	288
CNN	0.123 4	0.157 6	0.209 7	0.234 5	0.265 4	0.309 8	0.098 7	0.123 4	0.165 3	0.189 2	0.215 6	0.254 3
FNN	0.129 8	0.165 4	0.223 1	0.250 7	0.284 2	0.331 5	0.103 4	0.130 9	0.176 8	0.201 3	0.229 8	0.270 4
LSTM	0.114 6	0.150 3	0.198 7	0.223 5	0.256 4	0.301 7	0.091 7	0.118 4	0.156 7	0.180 7	0.207 6	0.247 8
Informer	0.118 7	0.154 3	0.206 8	0.231 2	0.263 7	0.308 4	0.094 3	0.121 5	0.162 7	0.1873	0.214 3	0.253 7
TCN-Transformer	0.112 3	0.145 7	0.196 4	0.220 8	0.253 7	0.298 3	0.089 7	0.116 3	0.153 4	0.176 8	0.203 4	0.243 7
TransPVP	0.109 8	0.143 2	0.193 4	0.217 6	0.250 7	0.295 4	0.087 6	0.113 4	0.149 8	0.172 7	0.198 3	0.238 7
CT-NET	0.111 4	0.146 8	0.197 3	0.221 7	0.254 7	0.300 2	0.088 4	0.115 3	0.154 3	0.178 6	0.205 3	0.245 7
Mutiformer-TSA	0.104 3	0.136 4	0.184 7	0.209 4	0.240 3	0.284 3	0.083 7	0.107 6	0.143 4	0.165 7	0.190 3	0.229 4

由表 2 可知,基于 Multiformer-TSA 的光伏发电功率预测方法在所有预测尺度上均表现最优,尤其在序列长度为 144 和 288 时, RMSE 和 MAE 显著低于其他模型,验证了该方法能有效捕捉光伏发电功率时间序列中不同维度和不同时间尺度的依赖关系,增强了对动态变化的理解和建模能力。

TransPVP 和 CT-NET 的表现次优,特别是在短尺度预测中,相关数据接近本文所提 Multiformer-TSA 方法,说明它们的时间依赖建模能力较强,能在较短时间跨度内准确捕捉光伏发电功率的波动。然而,在长尺度预测中它们的性能略逊于 Multiformer-TSA 方法,表明这两种方法在整合多尺度特征时仍存在一定局限性。

Informer 和 TCN-Transformer 在中长尺度预测上表现稳定,表明它们在时间序列建模中具有优势,能捕捉较长时间依赖关系,但在短时间波动的精细建模上稍有不足,导致短尺度预测误差较高。

综上所述,以上对比实验结果充分验证了本文所提出的 Multiformer-TSA 方法的预测性能优越性。该方法通过 CSL、CDL 和 TSA,可有效捕捉光伏发电功率数据中的复杂动态波动和长期趋势变化,规避了传统模型在处理多维度、多尺度特征时的局限性,具有较高的预测精度和稳定性。

3.4 消融实验

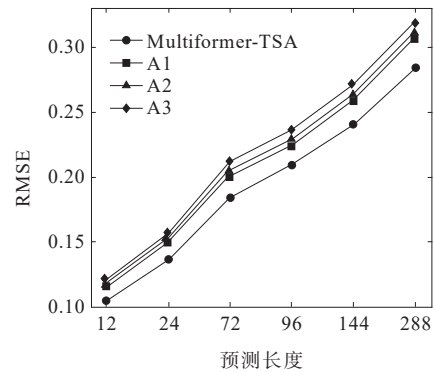
为评估 CSL、CDL 和 TSA 的有效性,本文进行了消融实验,实验设置如表 3 所示,实验结果如图 9 所示。由图 9 可知,在依次去除 CSL、CDL 和 TSA 后,各模型的光伏发电预测性能皆有不同程度的下降。

表 3 消融实验设置

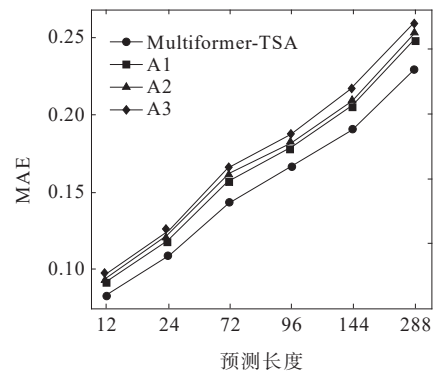
Table 3 Ablation experiment settings

实验编号	消融部分
A1	CSL
A2	CDL
A3	TSA

Multiformer-TSA



(a)RMSE



(b)MAE

图 9 消融实验结果对比

Figure 9 Comparison of ablation experiment results

CSL 被用于建模光伏发电功率与外部环境变量的交互关系,缺失 CSL 后,模型难以捕捉不同变

量间的复杂依赖性,导致短期波动预测不准确。实验 A1 在 12 和 24 时间步长的短期预测中的 RMSE 和 MAE 显著增加,预测精度较完整模型下降了约 9.6% 和 11.2%。此外,光伏功率的变化不仅由时间序列内在趋势决定,还受到环境因素的非线性耦合影响。CSL 通过跨维度特征建模,有效补充了这一信息,验证了 CSL 在快速变化的功率波动场景下可增强模型对多维特征交互的理解。

CDL 通过多尺度卷积采样捕捉光伏发电功率的长期趋势和季节性变化特征,去除 CDL 后,模型在较长时间跨度内难以准确识别周期性变化,预测结果波动较大,实验 A2 在时间步长为 72 和 96 的中长期预测中,预测误差显著增加, RMSE 和 MAE 分别上升约 7.5% 和 8.3%,特别是在复杂天气条件下,模型对功率变化模式的理解能力明显不足,表现出较高的预测误差,验证了 CDL 可通过建模多尺度动态变化,帮助模型更好地应对光伏发电的长期趋势和复杂波动。

TSA 被用于融合 CSL 和 CDL 提取的多尺度与多维度特征,缺失 TSA 后,模型难以有效整合跨时间和跨维度依赖信息,导致预测结果对时间趋势变化和多维特征交互缺乏足够的敏感性。实验 A3 在时间步长为 144 和 288 的长尺度预测中,性能下降最为明显, RMSE 和 MAE 分别增加约 13.2% 和 14.8%。此外, A3 在日间突发功率波动和长期趋势变化的预测中表现出明显的滞后现象,特别是在功率快速变化时,预测曲线无法及时响应实际变化,验证了 TSA 能够通过多阶段注意力融合特征,增强模型对短期波动和长期趋势变化的综合理解能力。

CSL、CDL 和 TSA 这 3 个模块在模型中具有协同效应。CSL 主要负责短期波动建模,增强模型对多维变量依赖关系的捕捉能力;CDL 专注于提取长期趋势,使模型能够更好地理解光伏功率的周期性变化;TSA 通过融合多尺度和多维度特征,实现了对功率变化模式的全局建模,显著提升了模型在复杂场景下的预测精度。综上所述, Multiformer-TSA 方法通过跨维度与跨尺度建模,有效解决了现有方法在多变量时间序列预测中的局限性。

4 结 语

针对现有光伏发电功率全天候预测精度较低的问题,本文提出了基于 Multiformer-TSA 的光伏发电功率预测方法,构建了跨维度分段层与跨尺度扩散层,增强了捕捉多维时间数据跨维度与跨尺度

信息的能力。通过分析对比实验与消融实验的结果,本文得出以下结论。

1) 本文提出的 Multiformer-TSA 方法通过 CDL 和 CSL,有效捕捉了光伏时间序列中跨维度与跨尺度依赖信息,并通过消融实验证明了所提方法各模块的功能有效性。

2) 本文建立的 TSA 机制将 CDL 与 CSL 捕捉的多尺度信息进行融合,显著提升了 Multiformer-TSA 方法对关键特征的选择和学习能力。实验结果表明,该方法相较于现有方法和其他基于 Transformer 预测的方法在多个预测尺度上,均具有最佳 RMSE 和 MAE 指标表现,验证了本文提出的 Multiformer-TSA 方法能够有效应对光伏发电数据中的复杂动态波动和长期变化,解决了现有方法在复杂环境下全天候预测准确性不足的问题。

参 考 文 献:

- [1] 丁明,王伟胜,王秀丽,等. 大规模光伏发电对电力系统影响综述[J]. 中国电机工程学报,2014,34(1):1-14.
DING Ming, WANG Weisheng, WANG Xiuli, et al. A review on the effect of large-scale PV generation on power systems[J]. Proceedings of the CSEE,2014,34(1): 1-14.
- [2] 赖昌伟,黎静华,陈博,等. 光伏发电出力预测技术研究综述[J]. 电工技术学报,2019,34(6): 1201-1217.
LAI Changwei, LI Jinghua, CHEN Bo, et al. Review of photovoltaic power output prediction technology[J]. Transactions of China Electrotechnical Society,2019, 34 (6): 1201-1217.
- [3] 黄立滨,蔡海青,顾浩瀚,等. 计及分布式光伏和储能主动支撑的配电网日前日内协调优化运行策略[J]. 南方电网技术,2024,18(8): 51-60.
HUANG Libin, CAI Haiqing, GU Haohan, et al. Coordinated optimal strategy for day-ahead and intra-day operation of distribution network considering the active support of distributed photovoltaic and energy storage system[J]. Southern Power System Technology,2024,18(8): 51-60.
- [4] 张静,熊国江. 考虑季节特性与数据窗口的短期光伏功率预测组合模型[J]. 电力工程技术,2025,44(1): 183-192.
ZHANG Jing, XIONG Guojiang. Short-term photovoltaic power prediction combination model considering seasonal characteristic and data window[J]. Electric Power Engineering Technology,2025,44(1): 183-192.
- [5] 韩晓,王涛,韦晓广,等. 考虑阵列间时空相关性的超短期光伏出力预测[J]. 电力系统保护与控制,2024,52(14): 82-94.
HAN Xiao, WANG Tao, WEI Xiaoguang, et al. Ultrashort-term photovoltaic output forecasting considering spatiotemporal correlation between arrays[J]. Power System Protection and Control,2024,52(14): 82-94.

- [6] 郑珂,王丽婕,郝颖,等.基于数据集蒸馏的光伏发电功率超短期预测[J].中国电机工程学报,2024,44(13): 5196-5208.
ZHENG Ke, WANG Lijie, HAO Ying, et al. Ultra-short-term prediction of photovoltaic power based on dataset distillation[J]. Proceedings of the CSEE, 2024, 44(13): 5196-5208.
- [7] 杨建,常学军,姚帅,等.基于WT-CNN-BiLSTM模型的日前光伏功率预测[J].南方电网技术,2024,18(8): 61-69+79.
YANG Jian, CHANG Xuejun, YAO Shuai, et al. Day-ahead photovoltaic power forecasting based on WT-CNN-BiLSTM model[J]. Southern Power System Technology,2024,18(8): 61-69+79.
- [8] 王舒雨,李豪,马刚,等.基于TCN和DLinear的光伏发电功率多步预测模型[J].电力建设,2025,46(4): 173-184.
WANG Shuyun, LI Hao, MA Gang, et al. Multistep prediction model for photovoltaic power generation based on time convolution and DLinear[J]. Electric Power Construction,2025,46(4): 173-184.
- [9] 贾伟,付鹏,王正风,等.基于动态时间规整聚类的VMD-KELM光伏发电功率预测[J].电网与清洁能源,2025,41(11): 104-112.
JIA Wei, FU Peng, WANG Zhengfeng, et al. Photovoltaic power generation forecasting based on VMD-KELM with dynamic time warping clustering[J]. Power System and Clean Energy,2025,41(11): 104-112.
- [10] 周伊,肖先勇,赵清华,等.基于组合数据清洗与改进注意力机制的光伏发电功率预测[J].供用电,2024,41(10): 31-37+49.
ZHOU Yi, XIAO Xianyong, ZHAO Qinghua, et al. Photovoltaic power forecasting based on combined data cleaning and improved attention mechanism[J]. Distribution & Utilization,2024,41(10): 31-37+49.
- [11] 秦宇,许野,王鑫鹏,等.基于改进FCM-LSTM的光伏出力短期预测研究[J].太阳能学报,2024,45(8): 304-313.
QIN Yu, XU Ye, WANG Xinpeng, et al. Study on short-term photovoltaic output prediction based on improved FCM-lstm[J]. Acta Energetica Sinica, 2024, 45(8): 304-313.
- [12] 董俊,刘瑞,束洪春,等.基于BIRCH聚类的L-Transformer分布式光伏短期发电功率预测[J].高电压技术,2024,50(9): 3883-3893.
DONG Jun, LIU Rui, SHU Hongchun, et al. Short-term distributed photovoltaic power generation prediction based on BIRCH clustering and L-transformer[J]. High Voltage Engineering,2024,50(9): 3883-3893.
- [13] 李超然,潘鹏程,杨伟荣,等.基于改进相似日优化HBA-BiLSTM-KELM的光伏发电功率预测[J].太阳能学报,2024,45(5): 508-516.
LI Chaoran, PAN Pengcheng, YANG Weirong, et al. Research on pv system power prediction based on improved similar day and HBA-BiLSTM-KELM neural network[J]. Acta Energetica Sinica,2024,45(5): 508-516.
- [14] 王清亮,杨博,应欣峰,等.非晴空条件下光伏发电短期功率预测方法[J].太阳能学报,2022,43(3): 188-196.
WANG Qingliang, YANG Bo, YING Xinfeng, et al. Short-term photovoltaic power forecasting method under non-clear sky condition[J]. Acta Energetica Sinica,2022, 43(3): 188-196.
- [15] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]// Proceedings 31st Annual Conference on Neural Information Processing Systems' 17, Long Beach, CA, USA, 2017.
- [16] 杨鹏伟,赵丽萍,陈军法,等.基于相似时段匹配与Transformer网络建模的分布式光伏超短期功率预测方法[J].中国电力,2024,57(12): 60-70.
YANG Pengwei, ZHAO Liping, CHEN Junfa, et al. Distributed photovoltaic ultra-short-term power forecasting method based on temporal analog matching approach and transformer network modeling[J]. Electric Power,2024,57(12): 60-70.
- [17] 盛端祥,张啸宇.基于概率TCN-Transformer的短期光伏功率预测模型[J].综合智慧能源,2024,46(11): 10-18.
SHENG Ruixiang, ZHANG Xiaoyu. Photovoltaic power forecasting model based on probabilistic TCN-Transformer[J]. Integrated Intelligent Energy, 2024, 46(11): 10-18.
- [18] WEI K K, LUO W Q, TAN S Q, et al. CTNet: a convolutional transformer network for color image steganalysis[J]. Journal of Computer Science and Technology,2025,40(2): 413-427.
- [19] WANG J F, HU W S, XUAN L F, et al. TransPVP: a transformer-based method for ultra-short-term photovoltaic power forecasting[J]. Energies,2024,17(17): 4426.
- [20] LÓPEZ SANTOS M, GARCÍA-SANTIAGO X, ECHEVARRÍA CAMARERO F, et al. Application of temporal fusion transformer for day-ahead PV power forecasting[J]. Energies,2022,15(14): 5232.
- [21] LIM B, ARIK S O, LOEFF N, et al. Temporal Fusion Transformers for interpretable multi-horizon time series forecasting[J]. International Journal of Forecasting,2021, 37(4): 1748-1764.
- [22] 彭嘉杰,王滨,王佶江,等.基于Transformer的门控多任务暂态稳定评估[J].电力大数据,2023,26(5): 25-33.
PENG Jiajie, WANG Bin, WANG Jijiang, et al. Transformer based gated multi-task transient stability assessment[J]. Power Systems and Big Data,2023,26(5): 25-33.
- [23] 张禄晔,石研,曹志远,等.基于深层感知交互Transformer的光伏功率预测[J].智慧电力,2025,53(8): 70-78.
ZHANG Luxi, SHI Yan, CAO Zhiyuan, et al. Photovoltaic power forecasting based on deep perception-interaction transformer[J]. Smart Power,2025,53(8): 70-78.