

DOI: 10.19650/j.cnki.cjsi.J2210773

基于对比学习辅助训练的超短期风功率预测方法^{*}

王颖¹, 朱南阳¹, 谢浩川¹, 李健², 张凯锋¹

(1. 东南大学自动化学院复杂工程系统测量与控制教育部重点实验室 南京 210096; 2. 东南大学能源与环境学院
大型发电装备安全运行与智能测控国家工程研究中心 南京 210096)

摘要:利用深度学习方法提高风功率超短期预测精度能够给电力系统日内机组组合、超短期经济调度、和电力备用安排提供更精确的风功率预测结果,对进一步提高电力系统运行的安全性和经济性具有重要意义。本文针对当前深度学习特征提取模块对时序曲线中的隐式特征和趋势变化的相似性提取不充分的问题,提出一种基于对比学习辅助训练的超短期风功率预测模型,主要包括输入模块、特征提取模块、对比学习辅助模块和回归模块。该模型通过自监督的对比学习算法自主生成正负样本,并以拉开正负样本的映射空间距离为目标来辅助训练特征提取模块的网络参数,使得特征提取模块的映射结果中包含了输入信息相似性的隐式特征,进而减少数据冗余信息、增强样本关联性,最终提高风功率预测精度。实验结果表明,对比学习方法的平均绝对误差比长短期记忆网络和轻量梯度提升机方法分别下降了19.9%和6.5%,有效提高了风功率预测精度。

关键词:风功率预测;对比学习;深度学习;自监督学习;特征提取

中图分类号: TP181 TH702 **文献标识码:** A **国家标准学科分类代码:** 48060

Ultra-short-term wind power forecasting based on contrastive learning-assisted training

Wang Ying¹, Zhu Nanyang¹, Xie Haochuan¹, Li Jian², Zhang Kaifeng¹

(1. Key Laboratory of Measurement and Control of Complex Systems of Engineering of Ministry of Education, School of Automation, Southeast University, Nanjing 210096, China; 2. National Engineering Research Center of Power Generation Control and Safety, School of Energy and Environment, Southeast University, Nanjing 210096, China)

Abstract: The accuracy improvement of ultra-short-term wind power forecast with deep learning methods is of great significance for intraday unit commitment, ultra-short-term economic dispatch, and reserve scheduling of the power systems, which can further enhance the safety and efficiency. To address the problem in the existing feature extraction models that the similarity of implicit features and changing trends in time series curves have not been adequately extracted, this article proposes an ultra-short-term wind power forecast model based on contrastive learning-assisted training, which mainly consists of an input encoding module, a feature extraction module, a contrastive learning module, and a regression module. The self-supervised contrastive learning module autonomously generates positive and negative samples and enlarges the distance between the positive and negative samples in the projection space, which help to extract the implicit features of the similarity of the input information. In this way, the redundant information is reduced, the sample correlation is enhanced, and the accuracy of wind power forecast is ultimately improved. Compared with LSTM and Lightgbm methods, experimental results show that the mean absolute error of the proposed method is decreased by 19.9% and 6.5%, which effectively increase the wind power prediction accuracy.

Keywords: wind power forecast; contrastive learning; deep learning; self-supervised learning; feature extraction

0 引 言

近年来,风电迅速发展,已渐渐由补充能源转变为战略替代能源。然而,风电的不确定性对电力系统可靠运行和经济调度带来巨大的挑战。因此,提高风电功率预测精度,对于电力系统安全经济运行具有重要意义^[1]。其中,超短期预测通常指对风电场未来 1~4 h 以内的输出功率进行预测,对于电力系统实时调度和安全防控具有重要的实际意义。

作为一种典型的时间序列预测问题,业内已有大量学者研究了超短期风功率预测问题^[3-11]。文献[2]提出了基于灰色关联分析的风功率超短期预测方法,文献[3-5]分别研究了基于经验模态分解和小波分解的风功率预测和处理方法。随着大数据和人工智能技术的发展,机器学习方法也被大量应用于风功率预测领域。文献[6-7]提出了基于广义神经和支持向量机的短期风功率预测方法;文献[8]提出了基于卷积神经网络(convolutional neural network, CNN)和长短期记忆网络(long short-term memory, LSTM)的风功率预测模型;文献[9]提出了基于多元注意力(multi-head attention)框架与引导式监督学习的风功率超短期预测模型;文献[10]提出了结合 LSTM 和时域卷积网络(temporal convolutional network, TCN)的风功率预测方法;采用微软提出的基于决策树的轻量梯度提升机(light gradient boosting machine, Lightgbm),文献[11]提出了基于 Lightgbm 的风功率概率预测方法。

风功率的变化过程是一个自然物理过程,不同时段的变化趋势可能存在较强的相似性;并且这种变化趋势的相似性大多能从输入信息的相似性上反映出来,并用辅助预测任务。在风功率预测领域,目前已有许多利用相似性特征提高预测精度的工作。例如:文献[12-13]提出采用统计、指标计算、聚类等方法筛选出具有相似性的风功率历史数据作为训练样本,进而进行网络训练;文献[14]提出先用深度学习进行预测、再找相似历史数据进行修正的方法;文献[15]提出对气象信息的相似性进行聚类、进而采用神经网络的风功率预测模型。此外,相似性信息也被经常用于辅助电力负荷预测^[16-17]。然而,目前对于相似性信息的利用主要着眼于辅助“样本选取”和“预测后校正”这两个过程的工作,而对“特征提取”这一过程仍然大多基于传统的网络结构单元,如 CNN、LSTM 等。这些网络主要关注局部和周期性特征,对历史数据中的相似趋势信息和样本关联性特征关注不多;并且整个训练过程完全以风功率实际值作为正向引导来监督训练,其结果容易受到真实数据中的冗余细节和噪声的影响,产生局部过拟合现象,影响预测准确性。

对比学习(contrastive learning)作为一种典型的自监督学习方法,能够辅助特征提取模块捕捉相似性特征^[18]。对比学习的核心思想是自动构造相似实例(正样本)和不相似实例(负样本)来辅助训练特征提取模型。训练过程以拉开正负样本在高维映射空间中的距离为目标,使得样本映射到高维空间中的分布更好地反映出样本的隐式特征,从而更好地支持下游任务(如分类、预测等)。2018 年起,对比学习在自然语言处理和图像处理领域上已经取得了很大进展,包括 CPC、SimCLR^[19]、MoCo、SimCSE 等对比学习算法模型,InfoNCE^[19]、NCE loss^[20]等损失函数模型,以及基于对比学习的图像特征提取、聚类、自然语言数据增强等具体应用研究^[18]。在自然语言处理领域,对比学习的优势体现在自然语言本身经常出现重复和相似段落,通过对比学习能够让相似语义的特征分布更加聚集,不相似语义的特征分布更加远离,而整体的语义分布则更加均匀。对于时间序列而言,当通过特征提取将样本编码为向量矩阵后,其抽象形式和图像、自然语言编码类似,因而也可以采用对比学习辅助训练特征提取过程。目前,已有相关文献将对比学习方法用于时间序列预测,例如:文献[21]提出了一种通用的适用于时间序列的对比学习方法;文献[22]提出了一种考虑季节变化趋势的对比学习时间序列预测模型,并在 5 个时间序列预测问题上进行了测试并取得了较好的效果;文献[23]采用对比学习进行了股票趋势预测。然而,文献[21-22]均为面向通用的时间序列的对比学习预测模型,缺少针对风功率预测问题的适用性研究。因而,需要针对风功率数据特征,开展适合风功率预测的对比学习模型结构、正负样例生成方法、和损失函数模型等相关研究。

为此,本文将对对比学习方法引入超短期风功率预测模型,采用对比学习算法辅助特征提取模块的训练,使其能够更好地捕捉风功率时序曲线中的隐式特征和趋势变化的相似性,进而减少数据冗余信息、增强样本关联性,最终提高风功率预测精度。具体而言,该模型首先将不同高度的风速、风向、风功率等历史数据通过输入模块转化成矩阵向量;然后建立由 LSTM、CNN 和注意力机制构成的特征提取模块,并通过对比学习辅助训练优化特征提取模块的网络参数;最后利用回归模型将提取到的高维特征转化为风功率预测结果。

1 模型框架和训练思路

在超短期风功率预测中,常规的深度学习的方法大多采用监督学习算法,通过降低风功率实际值和预测值的偏差,整体训练深度学习模型;在此过程中,其特征提取模块和网络中的其他部分是同时训练的。相比于常规模

型,本模型的核心区别是增加了对比学习辅助模块,并采用该模块对特征提取模块的参数进行自监督学习、采用风功率实际值对回归模块进行监督学习。具体而言,本模型的整体框架如图1所示,包括4个部分:输入模块、特征提取模块、对比学习辅助模块和回归模块。其中,“输入模块”将输入数据进行编码,包括预测时间点前若干时段不同高度的风速、风向、风功率等数据;“特征提取模块”可以采用常见的深度学习模型,如LSTM、CNN等;“对比学习辅助模块”主要用于优化“特征提取模块”的网络参数;“回归模块”用于拟合和输出超短期风功率预测结果,一般可采用FFN、LSTM模型实现。

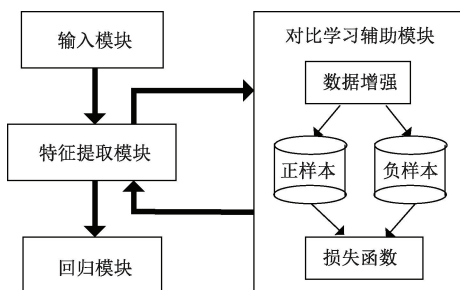


图1 模型整体框架

Fig. 1 Framework of the proposed model

“对比学习辅助模块”的建模思想是:首先对经过特征提取后的特征向量进行数据增强,得到正负样本集合;进而以对比学习损失函数为目标来训练“特征提取模块”的网络参数,从而拉开正负样本的映射空间距离,使得相似的样本在映射空间中更加接近,不相似的样本在映射空间中更加远离,其工作原理示意图如图2所示。

其中,风功率中具有某种相似性的段落如黑色和灰色曲线所示,其映射到高维空间(如球体)后变成高维空间的向量。如果采用常规的特征提取方法,这些向量在映射空间的距离可能较远,其样本之间的相似关系则被特征提取的过程给湮灭了。并且,由于高维输入时序数据中的隐式特征和相似关系是多样的,因而也很难预先通过简单地计算相似性指标来发现。对比学习的方法是以原样本为基础,采用特定方法生成正负样本,并以拉开正负样本之间的距离为目标来优化特征提取模型的参数,这样能够使得特征提取模型自主地发掘不同的隐式相似特征。

由于增加了“对比学习辅助模块”,本模型不能采用常规深度学习模型的端到端(end-to-end)训练方式,需要将训练过程分为两步。1)通过“输入模块”将时间信息、风功率、不同高度的风向和风速等数据转化为矩阵形式输入给“特征提取模块”,并采用自监督学习的“对比学习辅助模块”对“特征提取模块”的参数进行优化训练;2)在步骤1)“特征提取模块”训练完成后,采用“回归模

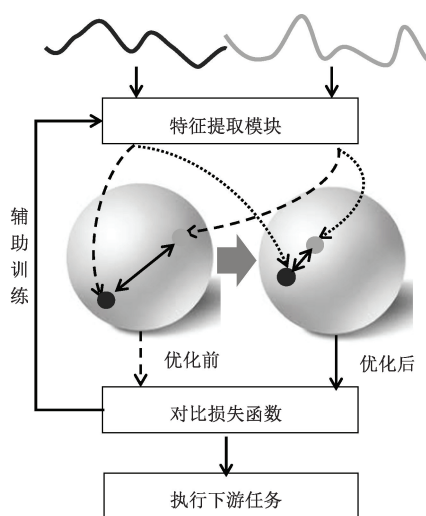


图2 对比学习建模原理示意图

Fig. 2 Schematic diagram of the contrastive learning assisted training process

块”对提取到的特征进行有监督学习的回归训练。在执行预测任务时,将需要已知信息提供给“输入模块”,然后经过“特征提取模块”和“回归模块”,最终输出风功率预测结果。

2 数据源及输入参数选取

数据是深度学习模型的基础。本文使用了中国北方某风电场从2019年1月1日~2020年12月31日共两年的实测数据,分辨率为1 h。本文采用了预测前36 h地面以上10 m处、30 m、风力涡轮机轮毂处的风向、风速和风功率测量值作为输入,输出为未来1~4 h的风功率预测结果。数据整体分为训练集、验证集和测试集。本文选择2019年1月1日~2020年10月31日的数据为训练集,2020年11月1日到2020年12月17日的数据为验证集,2020年12月18日~2020年12月31日的数据为测试集。

3 模型结构

3.1 模型整体结构

本文提出的超短期风电功率预测模型整体结构如图3所示,由输入模块、特征提取模块、对比学习辅助模块和回归模块组成。

输入模块:根据本文数据源的具体情况,本模型的训练输入包括:时间信息,预测时间点前1~36 h的4个高度(10、30、70 m、风力涡轮机轮毂处)的风速和风向,预测时间点前1~36 h的实测风功率。其中,时间变量拆分为

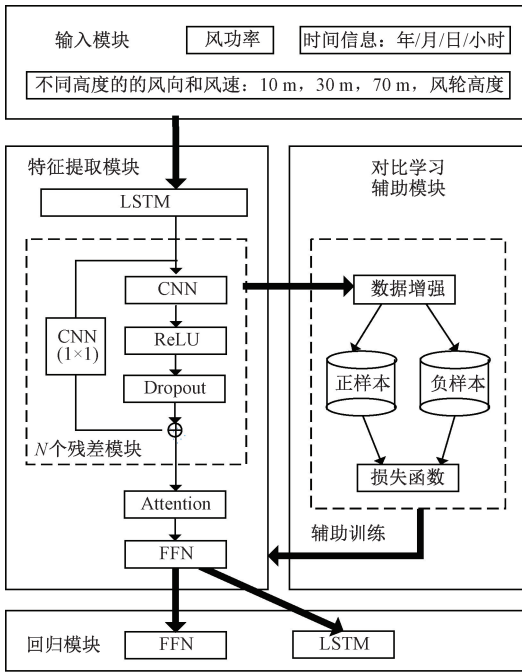


图3 超短期风电功率预测的对比学习框架

Fig. 3 The framework of representation learning for short term wind power forecasting

年、月、日和小时4个变量,从而更好地反映周期性和季节性,并采用 word2vec 量化,通过一键编码和映射层嵌入到一个矩阵中。

特征提取模块:特征提取模块本身的网络形式较为灵活,实际应用时可以根据计算能力和速度要求自主选择常见的网络模型,如 LSTM、CNN 等。本模型中借鉴了目前应用较多的 LSTM-CNN 模型^[8],并加入了注意力机制。其中, LSTM 网络是一种改进的循环神经网络 (recurrent neural network, RNN),常用于时间序列特征提取。一般情况下,不同高度的风速和风向对风力的影响具有时间上的滞后性,而 LSTM 具有反馈连接,并针对传统 RNN 可能存在的梯度消失问题进行了改进,增强了处理长时间序列的能力。经过 LSTM 后,数据进入由 CNN 构成残差网络和 Attention 等模块。本文采用重复多个相同的残差模块的方式,方便控制模型复杂度。其中, CNN 用于提取时间序列的局部信息, Attention 用于区分输入信息的重要性和关联性。具体来说,由于 LSTM 只能关注时间序列 $t \in T$ 之前的信息,且不能区分信息的重要性; CNN 模型只计算时间 $t \in T$ 上的局部窗口信息,且忽略了不同窗口之间的关联性;注意力机制 (attention) 则是模仿人类的注意力机制而提出的办法,对输入的每个元素采用不同的权重参数,从而自发地关注所需要的信息,抑制其它无用的信息,因此也被纳入本文的特征提取模型。

对比学习辅助模块:对比学习模块是一种自监督算法,其核心关键在于正负样本的生成方法和对比损失函数的构造方法,其详细模型将在 3.3 节中介绍。

回归模块:回归模块通常可以采用多种形式。为了评估对比学习辅助训练的特征提取结果对于不同回归模型的适应能力,本文建立了前馈神经网络 (feed forward networks, FFN) 和 LSTM 两个回归模型,并分别实现了整体预测功能。本文将所提出的基于对比学习辅助训练的方法命名为对比学习辅助训练 (contrastive learning assisted training, CLAT),分别与 FFN 和 LSTM 两种回归模型配合后命名为 CLATF 和 CLATL。

本文模型的数据处理方式如图 4 所示。在输入模块部分,对原始数据序列 (如各高度风速、风向、风功率等) 通过滑动窗口生成若干相同长度的序列,然后随机选择 N 条序列 $X_1, X_2, \dots, X_N$ 构成一个 batch 集合。对于一个 batch,对其中的各条序列进行特征提取,能够得到 N 个映射后的高维向量 $Y_1, Y_2, \dots, Y_N$ 。对映射后的向量进行数据增强,得到额外 N 个高维向量 $Y_1^+, Y_2^+, \dots, Y_N^+$,再选取其中的部分向量构成正负样本集合。

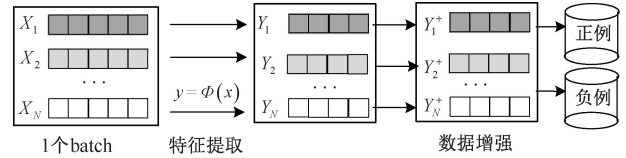


图4 数据处理方式

Fig. 4 The data processing method

3.2 特征提取模块

特征提取模块由 LSTM、若干个残差模块、Attention 和 FFN 构成。每个残差模块主要由 CNN 组成,并包含 1×1 卷积层的直连 (shortcut)。

1) LSTM^[24]: LSTM 是一种改进的循环神经网络,具有常规神经网络的重复模块链结构,被广泛用于风电、光伏等时间序列预测问题,其基本神经元包括遗忘门、输入门和输出门,分别决定了需要丢弃、更新、和输出的信息;各门控结构均含有一个 sigmoid 激活函数和一个点乘操作。

2) CNN^[25]: CNN 的基本结构包括卷积层和池化层,主要用于提取时间序列的局部信息。本模型中,首先通过一个滑动窗口将多维信息分成多个固定窗口的片段作为 CNN 的输入,然后提取每个片段的信息来获得局部信息。

3) Attention^[26]: 注意力机制可以获得一个序列和另一个序列之间的权重,并在全局信息之间建立关联。具体来说,本文中的序列是指经过 LSTM 和 CNN 获得的特征向量,权重是高维空间中两个序列之间的相似度矩阵,

其计算方法公式如式(1)和(2)所示。

$$\begin{cases} \mathbf{Q} = f(\mathbf{W}_q \mathbf{x} + \mathbf{b}_q) \\ \mathbf{K} = f(\mathbf{W}_k \mathbf{x} + \mathbf{b}_k) \\ \mathbf{V} = f(\mathbf{W}_v \mathbf{x} + \mathbf{b}_v) \end{cases} \quad (1)$$

$$\text{Attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{softmax}\left(\frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{d_k}}\right) \mathbf{V} \quad (2)$$

式中: $\mathbf{x}$ 是 CNN 模型的输出, $\mathbf{W}_q, \mathbf{W}_k, \mathbf{W}_v$ 是每个特征的权重, $\mathbf{b}_q, \mathbf{b}_k, \mathbf{b}_v$ 是偏置量, d_k 是 $\mathbf{K}$ 的维数, f 是激活函数。

3.3 对比学习模块

对比学习模块是一种自监督算法,其核心关键在于正负样本的生成方法和对比损失函数的构造方法。

1) 正负样本生成

图像和自然语言处理领域已经积累了大量生成正负样本的方法^[27-29]。例如,图像中经常采用旋转、翻转、裁剪、颜色抖动(jitter)等策略构造正样本,自然语言处理中通过删词、替换、回译、随机丢失(dropout)等方法生成正样本。这种通过对已有数据添加微小改动或从已有数据创建成新数据的方法,可以统称为数据增强(data augmentation, DA)方法^[30-31]。

针对时间序列问题,本文在经过特征映射后的矩阵上叠加采用了常用的数据增强技术来产生正样本。具体而言,对于形如 $Y_n(y_1, \dots, y_{t-1}, y_t, y_{t+1}, \dots, y_T)$ 的序列,随机采取以下一步或几步操作,其计算过程如算法1所示。(1) 滞后/超前操作:将每个 y_t 随机地将其替换为 y_{t-1}, y_t, y_{t+1} 中任意一个;(2) 幅值抖动操作:将每个 y_t 随机地加入少量高斯白噪声;(3) 时间抖动操作:采用 jitter 函数,可以看作添加时间上的噪声;(4) dropout 操作,相当于把 FFN 中部分隐节点以一定概率停止工作,通常在深度学习中用于增加输出多样性、减少过拟合。在随机地进行了上述操作后,得到的序列记为 Y_n^+ , 并与 Y_n 构成一个正样本对(positive pair)。对于一个包含 N 条序列的 batch,经过特征提取后包含 N 条 Y_n 序列;每一条 Y_n 经

过相同的数据增强后,可以生成一条 Y_n^+ , 因而一个 batch 最终包含的 Y_n 和 Y_n^+ 序列共有 $2N$ 个,如图4所示。

对应其中一个序列,例如 Y_n , 它与自己的 Y_n^+ 构成正样本对,并可以和其他所有 Y_k 及 $Y_k^+ (k \neq n)$ 构成负样本对(negative pairs)。为了防止负样本中碰巧存在与 Y_n 相似的序列、并减少负样本的个数,本文没有直接采用将所有其他的 Y_k 及 $Y_k^+ (k \neq n)$ 都放入负样本集合中,而是采用 DTW 算法^[32]进一步筛选负样本,其计算过程如算法2所示。

算法2

输入: Y_n and Y_n^+

输出: $\{Y_{n,k}^-\}$

- (1) For $k = 0$ to N do
- (2) 分值1 = DTW(Y_n, Y_k) $n \neq k$
- (3) 分值2 = DTW(Y_n, Y_k^+) $n \neq k$
- (4) 对所有 k 对应的分值1和2进行排序
- (5) 选择 K 个分值较小的序列构成负样本集合 $\{Y_{n,k}^-\}$
- (6) 返回 负样本集合 $\{Y_{n,k}^-\}$

2) 对比损失函数(contrastive loss)

对比损失函数表征了正负样本之间的距离,通过优化对比损失函数能够优化特征提取模块的网络参数。常用的对比损失函数大多基于互信息概念而建立,包括 NCE loss^[20], InfoNCE^[19], ProtoNCE^[33] 及一系列变种函数^[18]。本文采用的是目前最常用的 InfoNCE 损失函数,并由于在生成负样本后本模型采用了 DTW 算法对负样本进行了筛选,因此对分母进行了部分调整,其具体形式如式(3)所示,旨在最大化所有正样本对之间的相似性并最小化所有负样本对之间的相似性。其中,相似性的常见度量方法包括欧几里得距离和余弦相似度等,这里采用余弦相似度,其具体计算方法如式(4)所示。

$$\min l = - \sum_{n=1}^N \log \left(\frac{\exp[\text{sim}(Y_n, Y_n^+)/\tau]}{\sum_{k=1}^K \exp[\text{sim}(Y_n, Y_{n,k}^-)/\tau]} \right) \quad (3)$$

$$\text{sim}(\mathbf{u}, \mathbf{v}) = \mathbf{u}^T \mathbf{v} / \|\mathbf{u}\| \|\mathbf{v}\| \quad (4)$$

式中: (Y_n, Y_n^+) 为正样本对,而 $(Y_n, Y_{n,k}^-)$ 为负样本对, sim 为余弦相似度; τ 是温度超参数。

3.4 超参数设置

超参数设置主要包括网络超参数和训练超参数。在网络超参数方面,本文的输入数据是 36 h 不同高度的风速、风向等,整体数据量不大,为了防止过拟合,网络规模不宜过大;同时为了充分提取数据深层特征,本文在模型训练过程中,对部分超参数在其常见范围内进行了若干

算法1

输入: $Y_n(y_1, \dots, y_{t-1}, y_t, y_{t+1}, \dots, y_T)$

输出: $Y_n^+(y_1^+, \dots, y_{t-1}^+, y_t^+, y_{t+1}^+, \dots, y_T^+)$

- (1) For $t=0$ to T do
- (2) 随机采取以下一步或几步操作:
- (3) $y_t^+ \leftarrow \text{random}(y_{t-1}, y_t, y_{t+1})$
- (4) $y_t^+ \leftarrow y_t + \text{noise}$
- (5) $y_t^+ \leftarrow \text{jitter}(y_t)$
- (6) $y_t^+ \leftarrow \text{dropout}(y_t)$
- (7) 返回 正样本 $Y_n^+(y_1^+, \dots, y_{t-1}^+, y_t^+, y_{t+1}^+, \dots, y_T^+)$

尝试,最终设置的超参数如表 1 所示。在训练超参数方面,综合考虑训练时间和预测精度需要,本文采用均匀分布初始化函数(uniform)对网络参数进行初始化,选择批样本数量(batch size)为 128,采用学习率衰减的 Adam 方法作为训练优化器,学习率为 2×10^{-5} 。此外,为了评价本文所提模型的预测性能,本文将所提出的方法与 Lightgbm^[11]和 LSTM^[24]的预测结果进行对比,对比模型所采用的超参数如表 2 所示。

表 1 本文模型超参数

Table 1 Model hyperparameters of the proposed model		
模块	网络单元	参数
特征提取模块	LSTM	隐藏层节点个数=128 隐藏层层数=2
	残差网络	残差网络个数=3 dropout 比例=0.1 CNN 卷积核 1×2,步长=1
	Attention	单元数=128
对比学习 辅助模块	对比损失函数	温度参数温度参数 $\tau=0.1$
回归模块	FFN	隐藏层节点个数为 64 激活函数 ReLU 损失函数 MSE
		隐藏层节点个数为 128 隐藏层层数为 2 损失函数 MSE

表 2 性能比较模型的超参数

Table 2 Hyperparameters of performance comparison models	
模型名称	超参数
LSTM	隐藏层节点个数=128;隐藏层层数=2;学习率= 2×10^{-5} ;损失函数 MSE
Lightgbm	弱分类器数目=200;最大深度=30;学习率= 2×10^{-5} ;损失函数 MSE

3.5 训练过程分析

本文给出的两种基于对比学习辅助训练的风功率预测模型(CLATF 和 CLATL)具有相同的特征提取模块和不同的回归模块。图 5 给出了 CLATF 和 CLATL 两种方法的训练损失函数。可以看出,随着训练轮数增加,损失函数均快速下降。具体而言,CLATF 大约在 200 轮左右趋于收敛,CLATL 方法大约在 400 轮最优趋于收敛。本文选取训练 400 轮的模型作为最终的预测模型。

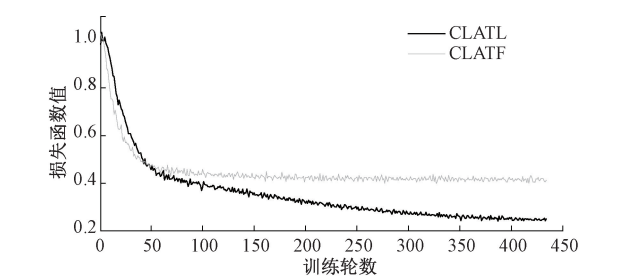


图 5 模型训练损失函数

Fig. 5 Training loss function

4 模型性能评价

4.1 评价指标

为了评估风功率预测模型的性能,本文采用以下评价指标:均方根误差(root mean square error, RMSE)、平均绝对误差(mean absolute error, MAE)和决定系数(R^2)^[34],其定义如式(5)~(7)所示。

$$RMSE = \sqrt{\frac{1}{N_s} \sum_{n=1}^{N_s} (y_n - \hat{y}_n)^2}$$
 (5)

$$MAE = \frac{1}{N_s} \sum_{n=1}^{N_s} |y_n - \hat{y}_n|$$
 (6)

$$R^2 = \left(\frac{\sum_{n=1}^{N_s} (y_n - \bar{y})(y - \hat{y}_n)}{\sqrt{\sum_{n=1}^{N_s} (y_n - \bar{y})^2} \sqrt{\sum_{n=1}^{N_s} (y - \hat{y}_n)^2}} \right)^2$$
 (7)

其中, N_s 为风电序列的时间长度, y_n 为第 n 个样本的真实值, $\hat{y}_n$ 为第 n 个样本的模型预测值, $\bar{y}$ 真实值的平均值, y 是预测值的平均值。

4.2 预测结果分析

不同模型在测试数据集上的评价指标结果如表 3 所示,其中 48 h 的预测结果如图 6 所示。为了比较不同模型的预测指标。可以看出,对于不同预测方法,CLATF 和 CLATL 整体预测误差较小、回归拟合性能较好。具体来看,在预测误差 RMSE 方面,对比学习方法相比于 LSTM 平均下降 14.0%,相比 Lightgbm 平均下降 3.6%;在预测误差 MAE 方面,对比学习方法相比于 LSTM 平均下降 19.9%,相比 Lightgbm 平均下降 6.5%,总体预测精度提升显著;在回归指标方面,对比学习方法相比于 LSTM 平均提升 5.8%,相比 Lightgbm 平均提升 18.8%,拟合效果提升明显。

对比不同提前时间长度的预测结果可以看出,各个模型都在提前 1 h 预测精度和回归效果相对最好;随着预测时间尺度增加,预测精度和回归精度会有所降低。

表 3 不同模型的表现

Table 3 Performance of different models

指标	提前预测小时	LSTM	Lightgbm	CLATF	CLATL
RMSE	1 h	10.08	10.28	9.15	9.05
	2 h	14.67	13.07	13.22	13.09
	3 h	17.41	14.38	15.00	14.03
	4 h	19.39	16.41	15.17	16.09
MAE	1 h	6.45	7.33	5.81	6.41
	2 h	10.23	8.88	8.49	8.55
	3 h	13.14	9.90	10.05	9.69
	4 h	15.12	10.68	9.70	10.70
R ²	1 h	0.84	0.86	0.87	0.88
	2 h	0.76	0.68	0.79	0.80
	3 h	0.72	0.60	0.75	0.76
	4 h	0.67	0.49	0.73	0.74

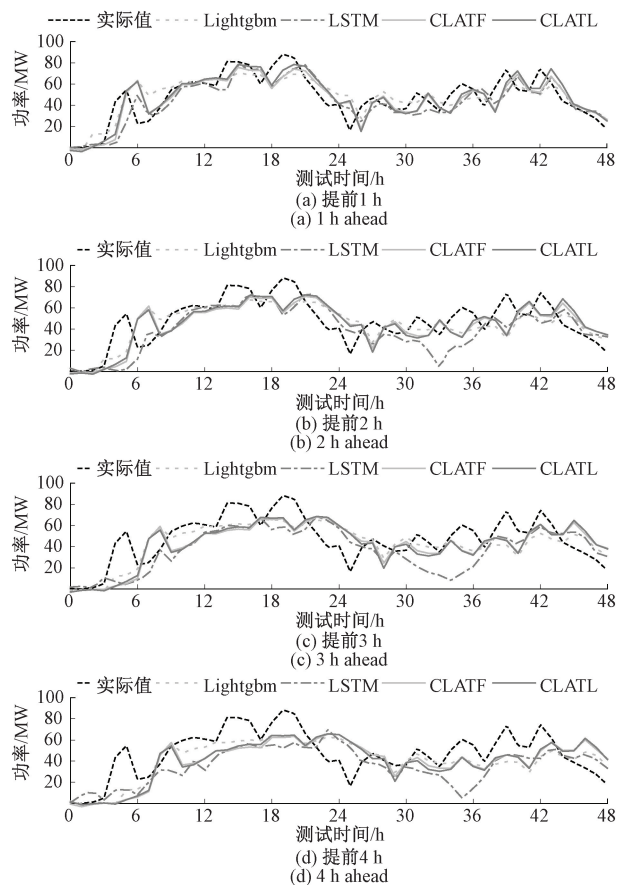


图 6 提前 1~4 h 不同模型的风电功率预测曲线
Fig. 6 Wind power prediction curves of different models
in 1~4 h ahead

并且,在不同时间尺度上,对比学习方法的回归精度相比于 LSTM 和 Lightgbm 均有显著提升,且提升程度随着预

测时间尺度增加而进一步扩大,整体效果呈现出良好的稳定性;在不同时间尺度的预测精度方面,对比学习在各个时间尺度的预测精度相比 LSTM 都有较为显著的优势,相比 Lightgbm 在 1 和 4 h 尺度上具有一定优势。综上所述,对比学习在超短期风电功率预测方面,在预测精度和回归效果上相比于对照模型均有显著的改善,且部分指标在较长时间尺度预测上依然具有稳定且扩大的优势。

5 结 论

针对当前深度学习风功率预测方法中的特征提取模块对历史数据中的相似趋势信息和样本关联性特征关注较少的问题,本文提出了一种基于对比学习辅助训练的超短期风功率预测方法,主要结论包括:

- 1) 本文提出了一种基于对比学习辅助训练的超短期风功率预测方法,利用对比学习算法优化特征提取模块参数,能够更好地捕捉风功率时序曲线中的隐式特征和趋势变化的相似性,进而减少数据冗余信息、增强样本关联性,最终提高风功率预测精度。
- 2) 本方法相比于在“样本选取”和“预测后校正”中考虑相似性特征的风功率预测方法,能够在特征提取模块中更好地捕捉风功率时序曲线的相似性信息;相比于应用于图像处理、自然语言处理等领域的对比学习方法,本文建立了适合风功率预测的对比学习模型结构、正负样本生成方法和损失函数模型。
- 3) 基于实际数据的预测结果表明,基于对比学习辅助训练的特征提取模块与回归模块配合,能够有效提高风功率预测精度和回归拟合性能。相比于 LSTM 和 Lightgbm,对比学习方法的预测误差 RMSE 指标分别下降 14.0% 和 3.6%,MAE 指标分别下降 19.9% 和 6.5%,预测精度提升显著。

参考文献

[1] XIE L, CARVALHO P M S, FERREIRA L A F M, et al. Wind integration in power systems: Operational challenges and possible solutions[J]. Proceedings of the IEEE, 2010, 99(1): 214-232.

[2] SHI J, DING Z, LEE W J, et al. Hybrid forecasting model for very-short term wind power forecasting based on grey relational analysis and wind speed distribution features[J]. IEEE Transactions on Smart Grid, 2013, 5(1): 521-526.

[3] 杨茂,陈郁林. 基于 EMD 分解和集对分析的风电功率实时预测[J]. 电工技术学报, 2016, 31(21): 86-93.

YANG M, CHEN Y L. Real-time prediction for wind

- power based on EMD and set pair analysis [J]. Transactions of China Electrotechnical Society, 2016, 31(21): 86-93.
- [4] 曲正伟, 张坤, 王云静, 等. 基于小波-原子分解的超短期风电出力预测模型[J]. 仪器仪表学报, 2016, 37(10): 2251-2258.
- QU ZH W, ZHANG K, WANG Y J, et al. Ultra short-term wind power output forecast model based on wavelet decomposition and atomic decomposition [J]. Chinese Journal of Scientific Instrument, 2016, 37 (10): 2251-2258.
- [5] 陈科彬, 邱晓燕, 刘波. 风电场混合储能的小波包-SOC 分区功率控制 [J]. 仪器仪表学报, 2017, 38(8): 2078-2085.
- CHEN K B, QIU X Y, LIU B. Wavelet packet-SOC partitioned power control of hybrid energy storage for wind farm[J]. Chinese Journal of Scientific Instrument, 2017, 38(8): 2078-2085.
- [6] 王慧莹, 吴亮红, 梅盼盼, 等. 果蝇优化广义神经网络的风电功率短期预测 [J]. 电子测量与仪器学报, 2019, 33(6): 177-183.
- WANG H Y, WU L H, MEI P P, et al. Fruit fly optimization based generalized regression neural network for wind power short-term prediction [J]. Journal of Electronic Measurement and Instrumentation, 2019, 33(6): 177-183.
- [7] 孙驷洲, 付敬奇, 朱峰. CGAQPSO 优化 LSSVM 短期风电预测[J]. 电子测量与仪器学报, 2016, 30(11): 1718-1725.
- SUN S ZH, FU J Q, ZHU F. Short-term wind power prediction based on LSSVM optimized by chaos Gauss attractor quantum-behaved particle swarm optimization[J]. Journal of Electronic Measurement and Instrumentation, 2016, 30(11): 1718-1725.
- [8] WU Q, GUAN F, LYU C, et al. Ultra-short-term multi-step wind power forecasting based on CNN-LSTM [J]. IET Renewable Power Generation, 2021, 15 (5): 1019-1029.
- [9] 崔杨, 王议坚, 黄彦浩, 等. 基于多元注意力框架与引导式监督学习的闭环风电功率超短期预测策略[J]. 中国电机工程学报, 2023, 43(3): 1334-1347.
- CUI Y, WANG Y J, HUANG Y H, et al. Closed-loop wind power ultra-short-term forecasting strategy based on multi-attention framework and guided supervised learning[J]. Proceedings of the CSEE, 2023, 43(3): 1334-1347.
- [10] XIANG L, LIU J, YANG X, et al. Ultra-short term wind power prediction applying a novel model named SATCN-LSTM[J]. Energy Conversion and Management, 2022, 252, 115036.
- [11] HONG T, PINSON P, FAN S, et al. Probabilistic energy forecasting: Global energy forecasting competition 2014 and beyond [J]. International Journal of Forecasting, 2016, 32(3): 896-913.
- [12] 师洪涛, 张智峰, 丁茂生, 等. 基于相似性指数聚类的风电功率预测方法研究[J]. 计算机仿真, 2022, 39(8): 121-125.
- SHI H T, ZHANG ZH F, DING M SH, et al. Research on wind power prediction method based on similarity index clustering [J]. Computer Simulation, 2022, 39(8): 121-125.
- [13] 阎洁, 许成志, 刘永前, 等. 基于风速云模型相似日的短期风电功率预测方法[J]. 电力系统自动化, 2018, 42(6): 53-59.
- YAN J, XV CH ZH, LIU Y Q, et al. Short-term wind power prediction method based on wind speed cloud model in similar day[J]. Automation of Electric Power Systems, 2018, 42(6): 53-59.
- [14] 仲悟之, 李崇钢, 崔杨, 等. 考虑历史相似性加权的超短期风电功率组合预测[J]. 太阳能学报, 2022, 43(6): 160-168.
- ZHONG W ZH, LI CH G, CUI Y, et al. Combined prediction of ultra-short term wind power considering weighted historical similarity[J]. ACTA Energiæ Solaris Sinica, 2022, 43(6): 160-168.
- [15] 符杨, 郑紫宸, 时帅, 等. 考虑气象相似性与数值天气预报修正的海上风功率预测[J]. 电网技术, 2019, 43(4): 1253-1260.
- FU Y, ZHEN Z CH, SHI SH, et al. Offshore wind power forecasting considering meteorological similarity and NWP correction [J]. Power System Technology, 2019, 43(4): 1253-1260.
- [16] 黎灿兵, 李晓辉, 赵瑞, 等. 电力短期负荷预测相似日选取算法[J]. 电力系统自动化, 2008(9): 69-73.
- LI C B, LI X H, ZHAO R, et al. A novel algorithm of selecting similar days for short-term power load forecasting[J]. Automation of Electric Power Systems, 2008(9): 69-73.
- [17] 李滨, 高枫. 基于虚拟相似日与 DA-LSTPNet 的地区电网短期负荷预测[J]. 电力系统自动化, 2021, 45(22): 55-64.
- LI B, GAO F. Short-term load forecasting for regional power grid based on virtual similar days and dual-stage attention-based long and short time pattern network[J]. Automation of Electric Power Systems, 2021, 45(22): 55-64.

- [18] 张重生, 陈杰, 李岐龙, 等. 深度对比学习综述[J]. 自动化学报, 2023, 49(1): 15-39.
ZHANG CH SH, CHEN J, LI Q L, et al. Deep contrastive learning: A survey [J]. Acta Automatica Sinica, 2023, 49(1): 15-39.
- [19] CHEN T, KORNBLITH S, NOROUZI M, et al. A simple framework for contrastive learning of visual representations[C]. Proceedings of the 37th International Conference on Machine Learning (ICML), 2020: 1597-1607.
- [20] GUTMANN M, HYVÄRINEN A. Noise-contrastive estimation: A new estimation principle for unnormalized statistical models [C]. Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics, JMLR Workshop and Conference Proceedings, 2010: 297-304.
- [21] YUE Z, WANG Y, DUAN J, et al. TS2VEC: Towards universal representation of time series [C]. Proceedings of the AAAI Conference on Artificial Intelligence, 2022, 36(8): 8980-8987.
- [22] WOO G, CHENGHAO L, DOYEN S, et al. CoST: Contrastive learning of disentangled seasonal-trend representations for time series forecasting [C]. The International Conference on Learning Representations, ICLR, 2022.
- [23] HOU M, XU C, LIU Y, et al. Stock trend prediction with multi-granularity data: A contrastive learning approach with adaptive fusion [C]. Proceedings of the 30th ACM International Conference on Information & Knowledge Management, 2021: 700-709.
- [24] SHAHID F, ZAMEER A, MUNEEB M. A novel genetic LSTM model for wind power forecast [J]. Energy, 2021, 223: 120069.
- [25] HE K, ZHANG X, REN S Q, et al. Deep residual learning for image recognition [C]. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015: 770-778.
- [26] SHIH S Y, SUN F K, LEE H. Temporal pattern attention for multivariate time series forecasting [J]. Machine Learning, 2019, 108(8): 1421-1441.
- [27] WI Z, WANG S, GU J, et al. Contrastive learning for sentence representation [J]. ArXiv Preprint, 2020, ArXiv:2012.15466.
- [28] MENG Y, XIONG C, BAJAJ P, et al. COCO-LM: Correcting and contrasting text sequences for language model pretraining [J]. Advances in Neural Information Processing Systems, 2021, 34: 23102-23114.
- [29] ELDELE E, RAGAB M, CHEN Z, et al. Time-series

representation learning via temporal and contextual contrasting [J]. ArXiv Preprint, 2021, ArXiv: 2106.14112.

- [30] TIAN Y, SUN C, POOLE B, et al. What makes for good views for contrastive learning? [J]. Advances in Neural Information Processing Systems, 2020, 33: 6827-6839.
- [31] KHARITONOV E, RIVIÈRE M, SYNNAEVE G, et al. Data augmenting contrastive learning of speech representations in the time domain [C]. 2021 IEEE Spoken Language Technology Workshop (SLT), IEEE, 2021: 215-222.
- [32] MÜLLER M. Dynamic time warping [J]. Information Retrieval for Music and Motion, 2007: 69-84.
- [33] LI J N, ZHOU P, XIONG C, et al. Prototypical contrastive learning of unsupervised representations [C]. Proceedings of 9th International Conference on Learning Representations, ICLR, 2021.
- [34] KRAUSE P, BOYLE D P, BÄSE F. Comparison of different efficiency criteria for hydrological model assessment [J]. Advances in Geosciences, 2005, 5: 89-97.

作者简介



王颖, 2011 年于南京航空航天大学获得学士学位, 分别于 2014 和 2018 年于东南大学获得硕士和博士学位, 现为东南大学自动化学院讲师, 主要研究方向为电力调度优化和电力市场。

E-mail: wyseu@seu.edu.cn.

Wang Ying received her B.Sc. degree from Nanjing University of Aeronautics and Astronautics in 2011, and received her M.Sc. and Ph.D. degrees both from Southeast University in 2014 and 2018, respectively. She is currently a lecturer with the School of Automation at Southeast University. Her main research interests include the power system dispatch optimization and electricity market.



张凯锋 (通信作者), 2004 年于东南大学获得博士学位, 现为东南大学自动化学院教授, 博士生导师, 主要研究方向为电力系统频率控制、新能源发电以及人工智能在电力系统中的应用。

E-mail: kaifengzhang@seu.edu.cn.

Zhang Kaifeng (Corresponding author), received his Ph.D. degree from Southeast University in 2004. He is currently a professor and Ph.D. supervisor with the School of Automation at Southeast University. His research interests include frequency control of power systems, renewable energy generation, and the application of artificial intelligence in power systems.