

基于 RFAM-TCN 模型和多模态特征优选的非侵入式负荷分解

邬峙涛¹, 李子晨², 夏杨红², 倪旖旎²

(1. 浙江大学工程师学院, 杭州 310015; 2. 浙江大学电气工程学院, 杭州 310027)

摘 要: 为实现准确的负荷感知和高效的设备管控, 该研究结合多模态特征优选方法与改进时间卷积网络提出一种新型的非侵入式负荷分解(NILD)方法以提高辨识精度与效率, 该方法利用负荷多尺度最大重叠离散小波变换结果作为输入特征并结合最大相关性与最小冗余度(mRMR)算法进行特征筛选以确保多模态特征的有效整合。同时在时域卷积网络(TCN)模型中加入感受野注意力机制(RFAM)以提高不同时间尺度特征的提取能力。最后, 通过PLAID等数据集等进行算法验证, 结果表明所提方法能在使用较少特征的前提下实现更为精确高效的负荷分解, 高于传统BiLSTM等模型5.31%的回归准确度, 对光伏等新能源户用设备的检出准确率达到96.6%以上。

关键词: 电力负荷管理; 小波变换; 特征选择; 非侵入式负荷分解; 感受野注意力机制; 时域卷积网络

中图分类号: TM732

文献标志码: A

0 引 言

随着“双碳”目标的提出和可再生能源替代行动的推进, 电力系统中具有波动性与随机性新能源的装机容量快速增长, 其大规模、高比例发展要求系统调节能力快速提升以实现系统的平衡与安全性要求^[1]。在此背景下, 挖掘可调节的灵活性资源, 实现供需侧资源协调优化将有助于推动清洁能源消纳, 与低碳节能目标的达成。非侵入式负荷分解(non-intrusive load decomposition, NILD)技术在1992年由Hart教授提出, 该技术依靠在家庭电表处测量的总负荷信号即可识别和监测家庭电器的使用情况, 达到监测电器用电情况的目的。与侵入式监测相比, NILD具有安装简单、成本低廉的优势。该技术为负荷侧资源的全面感知与分析提供了重要工具, 尤其可通过精细化分解用户侧柔性负荷, 支撑需求响应策略的实施, 将灵活用电需求与新能源发电高峰时段动态匹配, 显著提升风光等波动性电源的利用率。此外, 在户用场景中, NILD与家庭能源管理系统(home energy management system, HEMS)的深度融合, 可进一步优化灵活性资源调配, 促进用户侧清洁能源自发自用, 降低对电网的依赖和总体能源成本, 可为负荷侧参与电力系统调节提供有力支撑, 对于电力市场机制和用户互动策略的制定、电网优化运行、能源效率提升、网络安全监管等电力业务的高效开展具有重要意义。

近年来, 随着机器学习和深度学习技术的快速发展, NILD方法从传统的基于模式匹配、隐马尔可夫模型和聚类的方法, 逐渐转变为依赖于深度学习技术, 如卷积神经网络^[2-4]、残差网络^[5-6]、长短期记忆网络^[7-9]、深度神经网络^[10]等机器学习模式。此外, 多模态特征融合技术和图神经网络也在NILD中得到了较好的应用。文献[11]利用极限学习机建立了非侵入式负荷识别模型, 该模型具有较高的泛化性能, 应用来自美国匹兹堡的56个家庭多种不同设备的高频数据集(publicly available load identification dataset, PLAID)进行了验证。文献[12]采用基于卷积神经网络的学习模型进行了非侵入式负荷识别, 利用多任务聚合损失函数调整梯度反向传播过程, 初步解决了接入设备杂糅场景中的负荷分解问题, 但普通卷积未考虑来自不同位置特征的差异信息, 限制了特征提取的能力。文献[13]构建了基于时域卷积网络(temporal convolutional network, TCN)的NILD模型, 采用扩张因果卷积在总表功率序列上进行卷积运算, 扩大了感受野, 提取到更加丰富的特征, 但由于感受野的扩大, 容易出现对显著信息的忽视, 难以对时域特性施加不同关注度。

除此之外, 当前NILD研究中由于待识别电器的工作状态多且运行情况复杂, 多工作状态电器的类内和类间的特征量较为相似, 会产生待识别电器与其他电器发生混淆的情况, 使NILD在杂糅设备的接入场景中回归效果不尽如人意。由于NILD对于分解结果实时性的要求, 系统整体的轻量化

收稿日期: 2025-01-17

基金项目: 国家重点研发计划(2021YFB2401303)

通信作者: 夏杨红(1991—), 男, 博士、研究员、博士生导师, 主要从事可再生能源发电技术方面的研究。royxiayh@126.com

和高效化也应被纳入优化考虑的范畴。文献[14]利用小波变换提取了暂态小波特性作为神经网络的输入用于负荷识别,引入了时频域信息,获得了较好的回归效果;文献[15]利用互信息法进行特征选择,考虑了设备特征间的冗余度,精简了一部分重复输入特征;文献[16]运用最大相关最小冗余特征选择算法,从负荷的各工作状态中筛选出最优特征组合作为输入。

上述研究大多只关注高频数据的单个域特征(时域或频域),很少考虑双域特征的混合性能;而且在接入设备杂糅的情况下,对输入特征并未做出有效筛选,在设备暂稳态特性杂糅的监测终端容易产生漏检或误检的情况,且监测效率较低。综上所述,目前户用设备负荷分解研究的不足主要在于:1)对于繁杂特征图谱的提炼不到位,诸多暂稳态特性的输入导致模型处理效率低下;2)暂稳态特征时域跨度不定,发生时间不一,缺乏差异化的特性关注方法。为解决上述不足,本文提出一种基于包含感受野注意力机制 TCN 模型和多模态特征优选的新型 NILD 方法实现对户用设备的负荷分解。首先,针对设备能量流信息的预处理,利用最大重叠离散小波变换(maximal overlap discrete wavelet transform, MODWT)输入更多不同频段的时频数据,并通过 IRelief-F(improved Relief-F)计参与 MIC-Pearson 系数分别度量特征相关度和冗余度筛选最优特征子集,精炼输入的特征图谱。其次,为解决传统卷积运算中因参数共享带来的时域显著特征感知缺失,提出一种将感受野注意力机制(receptive field attention mechanism, RFAM)融入时域卷积网络中的方法,对不同时间域范围中的负荷特征进行更好地提取。最后,在公开数据集 PLAID 和自测户用新能源电气信息数据集^[17]上验证所提模型的有效性与优越性。

1 优化特征提取和特征筛选方法

在 NILD 任务中,特征的选取直接决定深度学习模型的回归效果,选择和提取最适合负荷分解的电气设备能量流信息有利于提升模型的性能、计算效率以及泛化能力。本文提出的多模态特征优选方法包含 MODWT 时频域特征提取方法,以及基于 IRelief-F 方法的特征相关度度量法和基于 MIC-Pearson 系数的特征冗余度度量方法。

1.1 基于 MODWT 的时频域特征提取方法

在传统 NILD 特征提取任务中,通常无法同时兼顾时域和频域上的信息,从而导致特征在回归效果上性能较差。为此,本文引入 MODWT 作为特征提取方法,在原先时域特征基础上增设频域特征,兼顾暂态和稳态分量,能有效捕捉信号中的局部特征和多尺度信息,并加以去噪,从而有效地提高 NILD 的性能。MODWT 是一种有效的非平稳突变信号的时频分析方法,是在离散小波变换(discrete wavelet transformation, DWT)基础上改良而来的。通过对电力负荷信号进行小波变

换,可以提取出不同频率成分的细节系数和近似系数,进而揭示电力信号中的不同频率波段内的重要特征。

DWT 的实质是运用高通滤波器和低通滤波器将电力信号连续分解为近似系数和细节系数。对于每一层分解得到的近似系数,对其继续分解得到下一层的近似系数和细节系数,依此类推。然而,由于每次 DWT 操作将实现时间序列上的一次二分采样,导致分析的分辨率降低,并且信号的移位可能导致小波系数显著变化,即缺乏平移不变性。为解决这些问题,MODWT 在分解过程中,对小波函数和尺度函数在不同尺度上都以最大重叠的方式进行卷积,并在滤波器每一层上都经过适当的升采样扩展。基于此操作,不仅能在每个分解层次上保留原始信号的采样率,并且保证了信号的平移不变性,高度冗余且非正交,对序列长度无特殊限制^[18]。此外,这一步骤也确保了在后续深度学习模型中输入特征的长度与原始信号一致,在导入后续 TCN 网络中时满足全连接层中输入数据维度需一致的要求。针对长度为 N 的负荷功率序列 $x[n]$,MODWT 的高通滤波器 $h'[n]$ 和低通滤波器 $g'[n]$ 分别可表示为:

$$h'[n] = h[n] / \sqrt{2} \quad (1)$$

$$g'[n] = g[n] / \sqrt{2} \quad (2)$$

根据 Mallat 算法,在分解层次为任意正整数 H 时,MODWT 的近似系数 $A_H[k]$ 和细节系数 $D_H[k]$ 如式(3)、式(4)所示。经过 H 层的分解后,细节系数(高频)的对应频率范围如式(5)所示,近似系数(低频)的对应频率范围如式(6)所示。

$$A_H[n] = \sum_n g'[n] A_{H-1}[(n-l)\bmod N] \quad (3)$$

$$D_H[n] = \sum_n h'[n] A_{H-1}[(n-l)\bmod N] \quad (4)$$

$$D_H[n] \in \left[\frac{f_s}{2^{H+1}}, \frac{f_s}{2^H} \right] \quad (5)$$

$$A_H[n] \in \left[0, \frac{f_s}{2^{H+1}} \right] \quad (6)$$

式中:mod——取余操作; l ——滤波器长度; f_s ——采样频率,Hz。

理论上 MODWT 允许任意长度信号分解,分解的 H 达到 $\log_2 N - 1$ (N 为信号长度),然而若在高频采样的数据集中将所有层的分解结果均作为输入特征输入回归模型,容易造成计算资源消耗过大且过拟合的可能性增加。因此,后文将利用所提的负荷分解模型和典型数据集进行性能校验,从而确定最适合的 MODWT 分解层数加入到初始特征集合中。

1.2 基于最大相关最小冗余算法的特征筛选方法

由于通常的电气参数特征,比如总线有功、无功功率等均由波形数据进行逻辑运算得出,数据之间可能存在信息重

叠、关系扁平化等问题。因此,在初始特征集的基础上提出使用最大相关最小冗余(maximum redundancy maximum relevance, mRMR)特征选择算法筛选得到最优电气参数特征集。由于 MODWT 的分解层数量尚需通过校验来确定,因此先用 H 表示。结合电气参数特征和 MODWT 结果的初始特征集如表 1 所示。

表 1 初始特征集合 S Table 1 Initial feature set S

特征编号	名称
C_1	有功功率
C_2	无功功率
C_3	总线电流
C_4	总线电压
C_5	电流 3 次谐波
C_6	电流 5 次谐波
C_7	电流 7 次谐波
C_8	电流总谐波失真 THD
$C_9 \sim C_{(8+H)}$	MODWT 第 1~ H 层细节系数
$C_{(9+H)}$	MODWT 第 H 层近似系数

具体地, mRMR 特征选择算法采取基于 IRelief-F 方法的特征相关度量度和基于 MIC-Pearson 系数的特征冗余度量度,筛选出相互之间相关性最大,冗余度最小的特征,构成最优特征子集作为分解模型输入,从而有效提高识别精度、增加算法的鲁棒性、加快计算速度。

1.2.1 基于 IRelief-F 算法的特征相关度量度

Relief-F 算法根据特征对区分邻近样本的能力为其赋予权重,适用于多分类问题^[19]。该算法通过计算特征在邻近样本间的差异,评估每个特征的相关度,其特征排序策略是随机从训练样本集中选取一个样本 R ,随后在与 R 类别相同的样本中找到 k 个近邻样本 H ,并从 R 的不同类样本中找出 k 个近邻样本 M ,特征 γ 的 Relief-F 分值 $\omega(\gamma)$ 可根据样本 R 与样本 H 、 M 间的距离进行多次迭代更新求解得到,相关度越高的特征,对分类效果的贡献越大。由于 NILD 任务为时序回归任务,原先的 Relief-F 方法无法直接处理连续的数据,且其随机采样特性会导致包含时序逻辑的信息缺失。因此,本文在 Relief-F 方法的基础上提出 IRelief-F 方法,即使用滑动窗阵列(多个等长的积分滑动窗口)来提取时间序列数据的信息。对于某个特征对应的时间序列(如总线功率),使用滑动窗阵列进行预处理,将每个窗口内的子时间序列处理为一个对应的值,然后将同一滑动窗阵列中的多个值作归一化处理,得到窗口值为:

$$\hat{E}_{\varphi,\psi} = \frac{\sum_{n=(\alpha\varphi+\psi-1)W}^{(\alpha\varphi+\psi)W} x[n]^2}{\max_{\psi} \sum_{n=(\alpha\varphi+\psi-1)W}^{(\alpha\varphi+\psi)W} x[n]^2} \quad (7)$$

式中: $\hat{E}_{\varphi,\psi}$ ——输出中特定窗口的窗口值; φ ——滑动窗阵列的序号亦即其进行滑动的次数, $\varphi = 1, 2, \dots, \frac{N}{W} + \alpha - r$; ψ ——取到当前阵列中的第几个窗口值, $\psi = 1, 2, 3, \dots, r$; α ——每次滑动经过窗的宽度 W 的倍数; W ——滑动窗宽度; $x[n]$ ——输入/输出数据点。

经过这一处理,能将连续的时序数据输入转化为离散的窗口值。随后,按照 Relief-F 方法基本原理,取上述滑动窗阵列中的第一个窗口值作为目标样本,以与目标样本值的差值作为判别标准确定近邻样本和远邻样本。确认近、远邻样本,计算并归一化每组滑动阵列的窗口值后,计算输入与输出中的距离函数(式(8)),最终将各组的 IRelief-F 评分累计并归一化得到该组特征的相关度(式(9)),其中相关度 $\omega(\gamma)$ 越高的特征对分类效果的贡献越大。

$$d(\hat{E}) = \sum_{\varphi=1}^{\frac{N}{W} + \alpha - r} [(\max_{\psi} \hat{E}_{\varphi,\psi} - \hat{E}_{\varphi,1}) - (\min_{\psi} \hat{E}_{\varphi,\psi} - \hat{E}_{\varphi,1})] \quad (8)$$

$$\omega(\gamma) = \exp \{-[d(\hat{E}_{\varphi,\psi}) - d(E_{\varphi,\psi})]^2\} \quad (9)$$

式中: r ——每个阵列中包含的积分窗口数量,为大于 3 的正整数。

综上所述,通过 IRelief-F 方法,可实现对连续特征的相关度量化。相比起原先方法,既实现了对连续数据的离散化和降采样和时间序列中的信息遍历,同时减少了进行该步处理时的资源消耗,效率更高。图 1 为 IRelief-F 方法原理示意图。

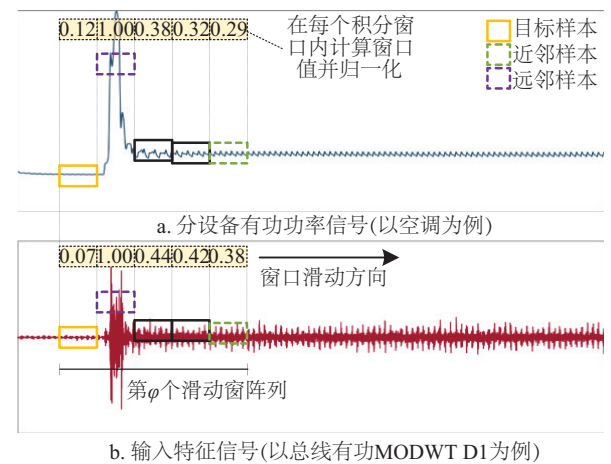


图 1 IRelief-F 方法原理示意图

Fig. 1 Overall process diagram of IRelief-F method

1.2.2 基于最大信息系数的特征冗余度量度

在特征选择中,不仅需要选择对目标变量有强相关性的特征,还需要避免特征之间的冗余(即特征之间的高度相

关)。最大信息系数(maximal information coefficient, MIC)不仅可以衡量两个随机变量之间的线性或非线形关系,而且可以揭示和量化变量间的非函数依赖关系,从而度量特征间的关联程度^[20]。MIC 主要利用互信息与网格划分方法进行计算,具体计算方法为给定两组变量 $X=\{x_i, i=1, 2, \dots, n\}$ 与 $Y=\{y_i, i=1, 2, \dots, n\}$, 定义互信息 $I(X;Y)$ 为:

$$I(X;Y)=\sum_{x \in X} \sum_{y \in Y} p(x,y) \log \frac{p(x,y)}{p(x)p(y)} \quad (10)$$

式中: $p(x,y)$ —— X 和 Y 的联合概率密度; $p(x)$ —— X 的边缘概率密度; $p(y)$ —— Y 的边缘概率密度。

假设 $D=\{(x_i, y_i), i=1, 2, \dots, n\}$ 为一个有限的有序对的集合, 定义 G 为由变量 X 和变量 Y 的值域构成 $x \times y$ 的网格, 将 X 与 Y 分别划分为 x 段与 y 段。在每个网格内部计算互信息, 选取不同划分方式下的最大互信息值作为 G 划分的互信息值, 然后通过归一化处理求得 MIC 值 $M_c(X,Y)$ 为:

$$M_c(X,Y)=\max_{|X||Y| \leq B(n)} \frac{\max(I(X;Y))}{\log_2(\min(I(X;Y)))} \quad (11)$$

式中: $B(n)$ ——网格划分 $x \times y$ 的上限值, 目的在于平衡网格划分的精确度与计算过程的效率, 文献[21]指出当 $B(n)=0.6$ 时能获得最佳效果, 故本文也采用该值。

通过上述方法计算特征集中任意两类特征的相关度, $M_c(C_i, C_j)$ 越小说明特征 i 与特征 j 间越独立, 反之亦然。

Pearson 相关性系数是一种衡量随机变量间线性的相关程度和方向的度量方法, 对于变量 $X=\{x_i, i=1, 2, \dots, n\}$ 与 $Y=\{y_i, i=1, 2, \dots, n\}$, Pearson 系数可由式(12)表示。 $P(X, Y)$ 取值范围为 $[-1, 1]$, 其绝对值越大表示两个变量间相关性越大。

$$P(X,Y)=\frac{\sum_{i=1}^n (x_i - \bar{X})(y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{Y})^2}} \quad (12)$$

式中: $\bar{X}$ 和 $\bar{Y}$ ——对应两个变量的平均值。

为了衡量数据集中的特征冗余度, 采用 MIC-Pearson 相关系数作为特征冗余度的评价指标。冗余度计算公式定义为:

$$P_M(X,Y)=\beta \cdot |P(X,Y)| + (1-\beta) \cdot M_c(X,Y) \quad (13)$$

对于一组特征集 $S=\{C_1, C_2, \dots, C_m\}$, 特征的冗余度大小 $r(C_i)$ 即为特征与子集中其他特征相关性之和, 将其定义为:

$$r(C_i)=\sum_{j=1}^m P_M(C_i, C_j), j \neq i \quad (14)$$

本文的特征排序与选择方法使用 IRelief-F 算法对特征进行相关度量, 采用 MIC-Pearson 相关系数对特征子集冗余度度量, 依次选择相关度最大且与其他所有特征冗余均值最小的特征入选特征子集, 入选的先后顺序即排序顺序, 排序越靠前的负荷特征重要性越高。

1.2.3 基于 SENet 的固定通道权重确定方法

在利用上述特征选择方法遴选出对应模型的最优特征子集后, 输入网络的特征数量大幅下降, 但各特征对于模型的权重是等同的, 缺乏特征对于回归结果的重要性刻画。如果使用通道注意力机制实时根据输入特性改变通道权重会导致模型测试资源消耗过大且可能发生拟合。因此, 本文在模型训练阶段为每个通道赋予特定的固定权重, 不仅能有效提高 NILD 回归效率, 并且在固定权重下可有效降低噪声敏感度, 避免因输入噪声产生无效波动且能够避免复杂动态调整带来的过拟合问题, 也能更为直接地反映出先验知识, 增加模型可解释性。本文利用通道注意力(squeeze and excitation network, SENet)方法在模型训练过程中为每个输入特征通道附加动态的权重, 然后对其结果取均值作为该通道最终的固定权重, 如式(15)所示。最终在模型测试过程中直接作为固定参数输入与输入特征相乘。

$$\omega_c = \frac{1}{N} \sum_{n=1}^N \sigma_c^{(n)}, \forall c \in [1, m] \quad (15)$$

式中: $\sigma_c^{(n)}$ ——第 n 个采样点时特征 C 的动态权重; c ——特征 C 的序号。

2 RFAM-TCN 神经网络原理

为捕捉不同尺度上的负荷特性, 选择将感受野注意力机制加入时域卷积网络中。本文构建的基于 RFAM-TCN 的非侵入式负荷分解模型主要由 TCN 模型和其中添加的 RFAM 模块组成。

2.1 基于时间卷积网络的时序特征提取模型

在处理特征时间序列数据时, 为避免时序信息丢失并增强模型训练效率, 采用 TCN 进行提取。TCN 融合了一维卷积和因果卷积的特点, 采用扩张因果卷积捕捉时序数据中的长期依赖关系。此外, 采用残差结构对网络输出进行修正, 在保证其性能的同时防止网络退化^[22]。

2.1.1 TCN 建模

对于一个时间图数据输入 $h=\{h^1, h^2, \dots, h^V\}$, 其对应的期望输出为 $\hat{h}=\{\hat{h}^1, \hat{h}^2, \dots, \hat{h}^V\}$ 。其中, V 表示每个节点的时间维数即滑动序列的长度, 则输入与期望输出序列的关系为:

$$\hat{h}^1, \hat{h}^2, \dots, \hat{h}^V = f(h^1, h^2, \dots, h^V) \quad (16)$$

式(16)的约束为 $\hat{h}$, 且仅取决于 $\{h^1, h^2, \dots, h^V\}$, 这意味着 TCN 在预测时只能使用当前时间及其更早的信息。TCN 建立在一维全卷积网络的基础上, 融合了因果卷积未循环连接的结构优点, 使时间序列数据能并行输入以获得更快的训练速度, 并且使用一系列扩张卷积来解决网络模型深度大、易过拟合的问题。TCN 通过以指数形式增加扩张因子 d , 使 TCN 以较少的网络层即可达到较大的感受野。但单纯的扩大感受野不能够对不同时域阶段上的能量流特性施加关注, 因此在下文中引入 RFAM 改进。

2.1.2 残差结构

残差结构是 TCN 模型的另一个重要部分,它通过跳跃连接操作,避免深度学习训练过程中随着网络深度的增加导致的梯度消失和退化问题^[23]。此外,在 TCN 残差块中融合了感受野注意力机制,其结构如图 2 所示。

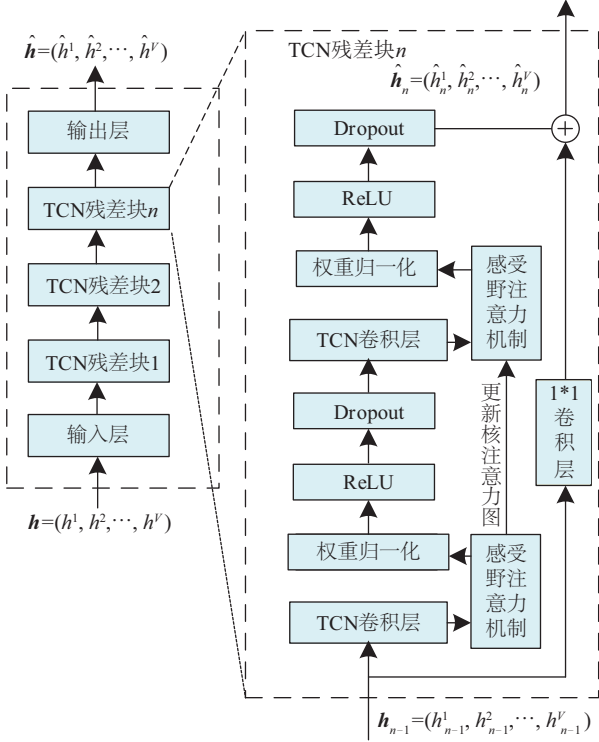


图2 RFAM-TCN 残差块结构

Fig. 2 TCN network and TCN residual block structure

残差块输出依赖于输入和一系列的变换 ρ , 定义为:

$$O_b = \text{Activation}(h_b + \rho(h_b)) \quad (17)$$

式中: O_b ——残差块输出; h_b ——残差块输入; $\text{Activation}()$ ——激活函数。

依据上文中特征集提取暂态事件后各类特征的时间序列窗, 利用滑动时间窗口建立时序特征图 $h = \{\vec{h}_1, \vec{h}_2, \dots, \vec{h}_m\}$, $h \in \mathbb{R}^{m \times v}$, 即将滑动窗口内的多维特征时间序列 $a_i = \{a_i^1, a_i^2, \dots, a_i^v\}$ 视为节点, $h_i = \{h_i^1, h_i^2, \dots, h_i^v\}$ 作为 RFAM-TCN 模型的输入, 其中 $i = 1, 2, \dots, m$, (其中 m 为节点数量, 即 1.2 节中特征子集中的各类特征数量)。模型输入滑动时间窗口的过程如图 3 所示。

时序特征图的节点代表各类不同特征, 多维特征间的边代表模型的通道相关性, 相邻节点间的边则代表两个变量间的空间相关性。通过 2 节中 IRelief-F 系数和 MIC-Pearson 相关系数的综合评价指标以及 SENet 附加的固定通道权重进行特征的筛选和排列实现通道间的加权, 但在空间层面没有注意力机制加持易导致训练效率低下, 导致回归结果不理想, 需通过加权感受野内的空间特征提高网络性能。

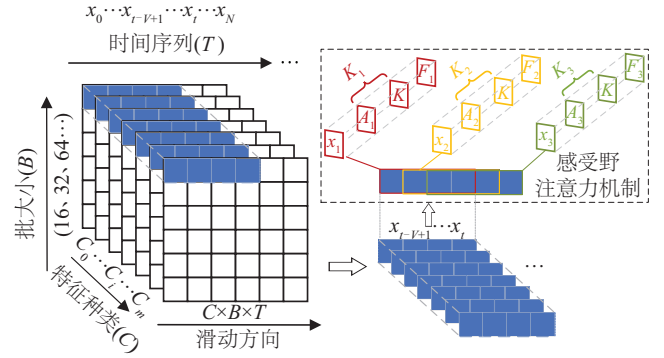


图3 滑动窗口法处理 TCN 输入数据和感受野注意力机制原理示意图

Fig. 3 Schematic diagram of sliding window method processing TCN input data and RFAM

2.2 感受野注意力机制

在谐波含量高和启停能量流跃变大的设备杂糅接入时(如微波炉和笔记本电脑), 由于有较大的瞬态电流, 深度学习模型易将其误识别为噪声纹波, 从而导致负荷分解出现问题。但是如果能动调整深度学习模型提取时间序列特征时的侧重, 提高对输入特征跳变、骤降等表征设备投切操作部分的感知权重, 可以有效减小误检、漏检概率。因此, 本文引入基于 RFAM 的空间注意力机制, 通过动态改变卷积感受野参数, 赋予对不同时间上的暂稳态特性的不同关注度。该机制解决了传统空间注意力机制中卷积核的参数共享问题, 并考虑到每个暂稳态特性在感受野中的重要性。RFAM 设计的卷积操作可根据输入数据的复杂性和重要性动态调整对应感受野区域的卷积核参数, 为其生成特定的权重。仅需额外的少量参数和计算开销即可使网络更加有效地捕捉和利用输入数据中的信息, 提高网络性能。

具体地说, 这个过程将注意力机制与卷积核相结合, 为每个感受野位置产生一个定制化的卷积核而非共享卷积核。如图 3 所示, 其中的 K_1 、 K_2 和 K_3 分别是通过将通用卷积核参数 K 与对应的注意力权重相乘得到的。图 4 为 RFAM 作用的基本流程。以 3×3 的卷积核为例, 首先对维度为 $C \times B \times T$ 的特征图使用平均池化聚合每个感受野特征的全局信息, 使用 1×1 的分组卷积交互信息, 生成 $9C \times B \times T$ 的注意力图用于后续的卷积核分配权重。与此同时, 原特征图经过 3×3 分组卷积得到与注意力图尺寸和维度相同的感受野空间特征图。最后, 感受野空间特征图根据注意力图的权重提取特征, 即可得到感受野注意力卷积的输出结果, 其维数与输入特征图一致, 但解决了不同卷积核之间部分参数共享的问题。感受野注意力机制的计算过程为:

$$\begin{aligned} F' &= \text{Softmax}(g^{1 \times 1}(\text{AvgPool}(h))) \times \text{ReLU}(\text{Norm}(g^{L \times L}(h))) \\ &= A_i \times F_i \end{aligned} \quad (18)$$

式中:Softmax()——归一化指数函数;Avgpool()——平均池化层运算;ReLU()——线性整流函数; $i \times i$ ——尺寸为 $i \times i$ 的分组卷积; L ——卷积核的尺寸;Norm——归一化处理; A_r ——注意力图; F_r ——感受野空间特征图。

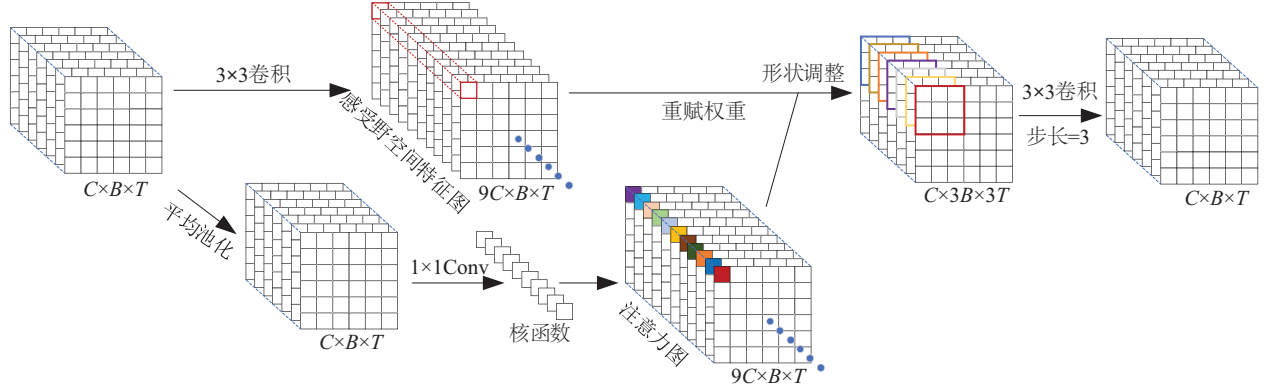


图4 感受野注意力机制作用基本流程

Fig. 4 Basic process of RFAM in network

3 算例分析

3.1 实验平台

为验证算法的有效性,采用 Matlab R2023b 及其深度学习包实现深度神经网络的建模。硬件环境为 13th Gen Intel®i9-13980HX@2.2 GHz 的 CPU, NVIDIA GeForce RTX 4090 的 GPU, 和 16 G RAM 的 64 位 Windows 系统计算机。在训练网络时使用 GPU 进行硬件加速。

3.2 实验数据

为获取高采样率的电气信息数据集以提取暂态特性,采用 NILDT 领域中经常使用的 PLAID 数据集以及自测户用新能源电气信息数据集。PLAID 数据集记录了美国宾夕法尼亚州匹兹堡 56 个家庭 11 种电器的高频电压、电流数据,其中电器工频为 60 Hz,采样频率为 30 kHz。由于该数据集采样频率较高,数据量较大,因此需在不影响结果的情况下适当进行降采样处理。根据奈奎斯特定理,在原数据的基础上每 20 个点提取 1 个数据点,取 1500 Hz 的采样频率不影响暂态特性提取。自测户用新能源电气信息数据集在常规户用设备的基础上增加了户用分布式光伏等新能源设备,由基于 TMS320F28335 的实验平台获取,采样频率为 1500 Hz。数据预处理中,随机选择多种不同设备接入,将能量流信息合并得到总线采样信息,将其与分设备信息分别作为负荷分解框架训练集中的输入和输出。

3.3 实验结果和分析

通过实验,将得到的经过多模态特征提取优选以及 RFAM-TCN 模型的负荷分解效果,与传统卷积神经网络(convolutional neural network, CNN)、隐马尔科夫模型(hidden Markov model, HMM)、双向长短期记忆网络(bi-directional

输出特征 F' 是通过将注意力图 A_r 与变换的感受野空间特征 F_r 相乘获得。RFAM 的加入在参数和计算复杂度方面略有增加,但能够显著提高网络在接入设备繁多等不良条件下的负荷分解和感知能力^[24]。

long short-term memory, BiLSTM)等模型进行对比。实验中采用均方误差(mean squared error, MSE, 记作 E_M);均方根误差(root mean squared error, RMSE, 记作 E_R);平均绝对误差(mean absolute error, MAE, 记作 E_A)和决定系数 R^2 等指标评估模型性能。

$$E_M = \frac{1}{N} \sum_{i=1}^N (w_i - \widehat{w}_i)^2 \quad (19)$$

$$E_R = \sqrt{\frac{1}{N} \sum_{i=1}^N (w_i - \widehat{w}_i)^2} \quad (20)$$

$$E_A = \frac{1}{N} \sum_{i=1}^N |w_i - \widehat{w}_i| \quad (21)$$

$$R^2 = 1 - \frac{\sum_i (\widehat{w}_i - w_i)^2}{\sum_i (\overline{w_i} - w_i)^2} \quad (22)$$

式中: w_i ——户用设备用电数据的实际值; $\widehat{w}_i$ ——户用设备用电数据的预测值。

3.3.1 多模态特征优选效果验证

首先,确定最适合模型的 MODWT 分解层数量以完成初始特征集的建立。采用 db6 小波函数对 PLAID 数据集总线有功功率信号进行多尺度 MODWT 并输入到所提 NILDT 模型中,得到回归结果如表 2 所示。表 2 表明,从第 4 层的 MODWT 分解开始,随着更多分解层的特性被输入模型,由于得到的细节系数所覆盖的频率范围会逐步向低频区过渡,决定系数反而下降,说明 MODWT 分解层数的增加对于一些富含高次谐波的电器设备的特性提取鲜有帮助。并且,特征子集中的特征数量会随着分解层数的增加模型处理的信息量和过拟合的风险显著增大。因此,由于第 4 层分解相比 3 层带来的决定系数收益不显著但效率明显降低,本文应用 3 次 MODWT 得到的 1~3 层细节系数和第 3 层近似系数加入初始特征集中分别作为 $C_9 \sim C_{12}$ 用于后续辨识。

表2 MODWT分解层数与回归结果关系

Table 2 Relationship between number of decomposition layers of MODWT and regression results

分解层数	D_h 频率范围/Hz	$R^2/\%$	测试所需时间/s
1	375~750	81.09	5.13
2	187.5~375	88.72	5.68
3	93.75~187.5	95.64	6.27
4	46.88~93.75	95.82	8.09
5	23.44~46.88	93.77	10.78
6	11.72~23.44	90.09	14.01

通过 MODWT 得到的高频分量能够获取与设备投切强相关的暂态特征分量,低频分量能够获取与设备长期运行状态和功率消耗模式相关的稳态特征分量。结果如图 5 所示。接下来,对包含 12 个特征的初始特征集合 S 进行 mRMR 特征筛选。为了筛选最优特征子集作为后续辨识的依据并验证本文所提特征相关度量方法 IRelief-F 的合理性和有效性,将特征根据两类特征排序方法(IRelief-F 和 Lasso)的相关度排序结果,按照相关度由高到低依次添加到特征子集中,并输入 RFAM-TCN 模型对比其辨识结果。

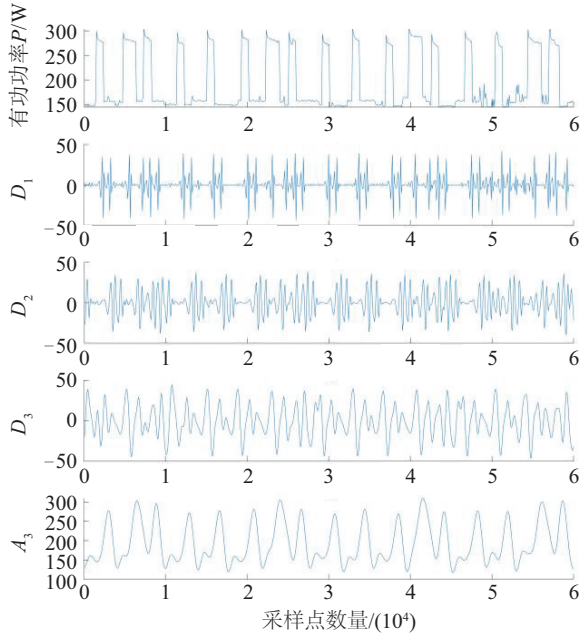


图5 对总线有功做多尺度最大重叠离散小波变换的结果

Fig. 5 Results of multi-scale maximum overlapping discrete wavelet transform for bus active power

通过图 6 中 IRelief-F 和 Lasso 的对比可得,利用 IRelief-F 方法时选择有功功率、MODWT 第 1 层细节系数等 9 个特征作为特征子集,回归决定系数最高达到 93.78%,优于 Lasso 所选的 10 个特征达到的 92.22%。此外,还需对特征冗余度

度量方法的有效性做出检验。在以 IRelief-F 方法作为特征相关度量方法的前提下,利用 MIC-Pearson 系数对 12 类特征之间冗余度进行综合度量。如图 6 所示,选取前 7 项特征即可使所提模型达到最佳负荷分解准确度,决定系数为 96.45%,且极值点两侧单调,相较于只用 IRelief-F 方法提高了 2.67%。因此,在 mRMR 筛选后的特征子集表示识别的最佳子集,即选择有功功率、MODWT 第 1 层细节系数、电流 5 次谐波、电流 3 次谐波等 7 个特征作为特征子集时,能够达到最大识别正确率,并由此得到最优特征子集 $S'=\{C_1, C_9, C_6, C_5, C_{12}, C_2, C_8\}$ 。可以看出,所提特征排序方法在大幅降低特征输入维度的同时也提升了负荷分解的准确性。

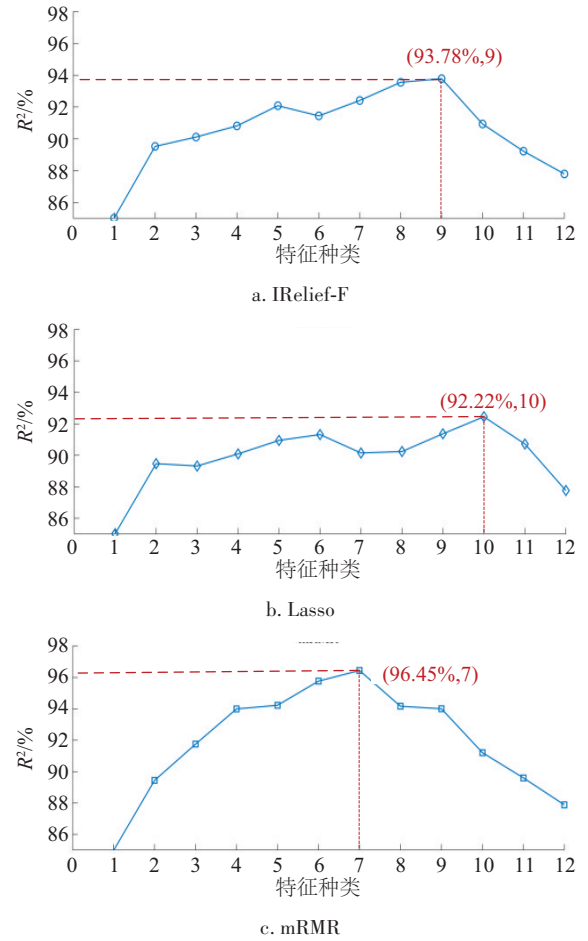


图6 mRMR方法辨识回归计算结果

Fig. 6 Regression results of identification by mRMR method

另外,为了进一步提升回归效率,还需确定每个输入通道的权重。基于 SENet 的固定通道权重确定方法,最终确定该最优特征子集 S' 中 7 项特征权重 $\omega_c=\{0.91, 0.76, 0.64, 0.66, 0.59, 0.38, 0.21\}$,回归效果的 R^2 指标由调节权重前的 96.45% 提高到 97.21%,提升较大;测试所需时间则由 6.24 s 变为 6.27 s,未见显著变化。可见,固定权重在几乎不影响回归速度的前提下对于回归准确度的提高有较大帮助。

3.3.2 负荷分解模型效果验证

为验证本文所提负荷分解模型效果,首先对输入参数进行预处理,其中,将每组序列的数据点样本数量 N 定为 15000,拆分为 4 个批次处理($B=4$)。由于时间序列仍较长,因此考虑设置多组周期性重置的扩张因子防止因指数增长导致训练的不稳定。其中,每一组内的扩张因子 d_i 设置为 1、2、4,一组包含 3 个残差块,分成 4 组加工,因此残差块数量共为 12 个;由于在每个残差块中包含两层加载 RFAM 的扩张卷积,因此扩张卷积层数为 24;感受野覆盖时间步长根据公式 $R_F = 1 + 2 \times \sum_{i=1}^n d_i$ 得到为 4093,累积到 4 个批次中能够保证感受野覆盖整个时间序列,无信息流失。为验证 RFAM 机制的加入带来的回归收敛,利用可视化注意力权重分布的方法监测在时域卷积的过程中,卷积核参数 K 的变化,如图 7 所示,当能量流数据发生陡增或突降时,卷积核权重参数响应幅值显著增大,表明了 RFAM 对于关键的暂态能量流变化有着较强的敏感度。

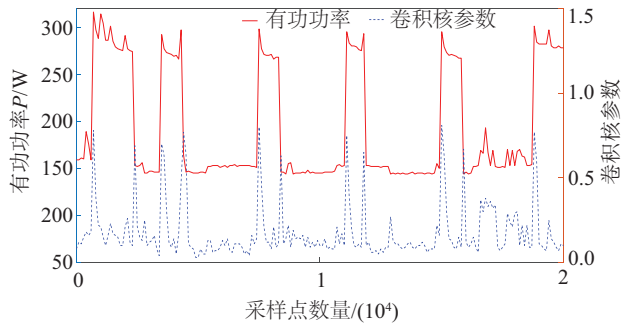


图 7 TCN 和 RFAM-TCN 对暂稳态数据回归对比

Fig. 7 Regression comparison between TCN and RFAM-TCN for transient data

针对 RFAM 机制对于不同时间尺度下的回归特性,本文提取了典型的稳态特性和暂态特性的能量流数据并分别利用 TCN 和 RFAM-TCN 加以验证。回归结果如表 3 所示。可以看出,加入 RFAM 模块的 TCN 模型在稳态特性的回归上优于原 TCN 模型 0.19%,但对于暂态特性的回归能力优于原 TCN 模型 5.45%。以上结果证明本文所提模型充分挖掘了各类特征的时空特性信息,使回归模型的性能得到了进一步提升。

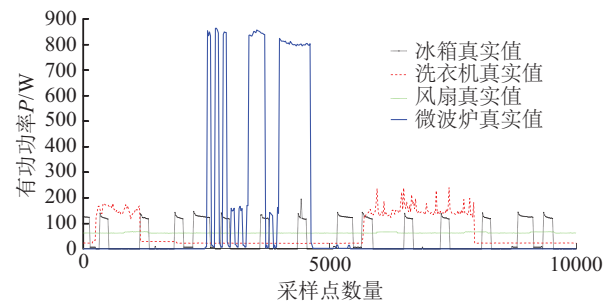
表 3 TCN 和 RFAM-TCN 对暂稳态数据回归对比

Table 3 Regression comparison between TCN and RFAM-TCN for transient data

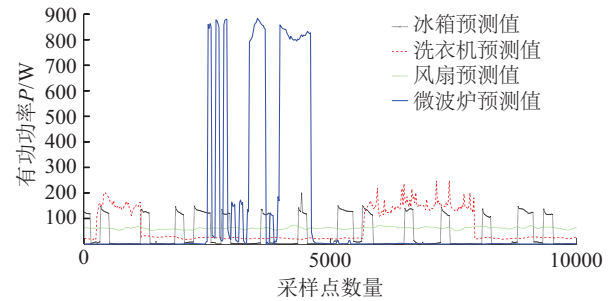
时域尺度	$R^2/\%$	E_M/W	E_R/W	E_A/W
TCN(暂态)	91.07	21.091	4.592	4.112
RFAM-TCN(暂态)	96.52	7.286	2.699	2.378
TCN(稳态)	97.09	4.579	2.139	1.992
RFAM-TCN(稳态)	97.28	3.781	1.944	1.771

图 8 为基于本文所提模型处理后部分设备的负荷分解效果。可以看出,所应用的基于 RFAM-TCN 模型和多模态特征优选的 NILD 方法对数据集里包含较大启停能量流瞬变过程或包含丰富谐波的微波炉和笔记本电脑等设备,都能实现较好的负荷分解回归效果。将所提方法与常见负荷分解的模型进行对比。在对各模型都输入最优特征子集 S' 中的 7 项特征的情况下,表 4 说明本文所提模型相较其他方法而言能够达到更好的 NILD 回归效果,回归准确率高于 BiLSTM 等传统模型 12.98%。如果对每个分解模型各自输入其通过 mRMR 筛选得到的最优特征子集。由表 5 可知,本文所提负荷分解模型利用最少的输入特征,各项回归性能指标均优于所列常见方法,并且在测试时间上也较有优势。

对比表 4 与表 5,既可以说明 mRMR 特征筛选方法对于不同模型的负荷分解效果均有提升,也证明了本文所提的



a. 分设备有功功率真实值



b. 分设备有功功率分解结果

图 8 负荷分解效果(分设备有功,降采样后)

Fig. 8 Load decomposition effect (active power of devices)

表 4 相同输入特征下的不同分解模型效果对比

Table 4 Comparison of different decomposition model identification methods under same input features

分解模型	决定系数 $R^2/\%$
HMM	86.09
CNN	91.91
TCN	92.11
BiLSTM	84.23
所提模型	97.21

表5 不同分解模型辨识效果和回归速率对比

Table 5 Comparison of identification effects and regression rates of different decomposition models

分解模型	所需特征数	$R^2/\%$	E_M/W	E_R/W	E_A/W	测试所需时间/s
HMM	8	90.08	13.982	3.739	3.441	5.82
CNN	7	91.90	8.702	2.949	2.673	5.38
TCN	8	94.57	6.451	2.540	2.182	6.35
BiLSTM	10	93.74	9.710	3.116	2.763	13.57
所提模型	7	97.21	4.511	2.228	1.873	6.27

RFAM-TCN 模型相较其他模型具有综合最优的回归效果,说明该模型充分降低了数据输入的复杂性和计算过程的资源消耗,使得回归性能得到进一步提升。考虑到未来大规模高比例新能源设备接入的场景下,本文进一步对包含光伏等新能源设备的自测户用新能源电气信息数据集进行模型的回归效果校验,实验结果如表 6 所示。在对常规户用设备保持较好回归效力的情况下,对光伏等新型户用设备的检出率达到 96.6% 以上,能有效避免对负荷特性的误判,支撑配电网源荷的协同调控,从而有效提升网侧新能源的利用率,并挖掘负荷侧参与电网调控的潜力。

表6 户用新能源电气信息数据集负荷分解结果

Table 6 Load decomposition effect of residential integrated new energy dataset

负荷种类	$R^2/\%$	E_M/W	E_R/W	E_A/W
冰箱	97.23	6.011	2.452	2.191
照明灯具	98.11	0.989	0.994	0.603
空调	94.82	5.189	2.278	1.972
微波炉	97.43	8.302	2.881	2.571
热水壶	98.32	5.826	2.414	2.198
光伏	96.60	101.445	10.072	8.671
综合回归评估 (均值)	97.32	20.153	4.489	4.007

4 结 论

本文提出一种基于 RFAM-TCN 模型和多模态特征优选的 NILD 模型,能将总表功率数据分解为单个电器的功率信息并通过数据集上的测试验证了这种方法的有效性。

1)在特征提取中,采用 MODWT 方法获取时频域特性,将最适合模型的三层分解结果加入到特征集中,并利用 mRMR 算法优化筛选暂稳态负荷能量流特性形成最优子集,较 Lasso 等特征选择方法提高了模型 4.23% 的回归准确率,也减少了模型用于实际场景的资源消耗和辨识时间。

2)在深度学习框架的建立上,选用在时域序列回归领域

效果出色的时域卷积网络并添加了 RFAM 机制。与传统神经网络模型相比,RFAM-TCN 在捕捉复杂特征关系和时序依赖性方面表现出显著优势,对于暂态特性的回归能力优于原 TCN 模型 5.45%;在设备能量流信息上也有了更好的回归效果,高于传统 CNN 等模型 5.31% 的回归准确度。

3)在保证较高分解准确率的前提下,该模型实现了轻量化和高效化上的进步,相比 BiLSTM 传统模型减少了 7.3 s 的检出用时。在高渗透新能源电力系统的背景下,对光伏等新能源户用设备的检出率也能够达到 96.6% 以上。

4)今后将考虑对模型的结构参数进一步优化,以适应更多不同的应用场景。

[参考文献]

- [1] 程祥,李林芝,吴浩,等.非侵入式负荷监测与分解研究综述[J].电网技术,2016,40(10):3108-3117.
CHENG X, LI L Z, WU H, et al. A survey of the research on non-intrusive load monitoring and disaggregation [J]. Power system technology, 2016, 40(10): 3108-3117.
- [2] 李利娟,刘海,刘红良,等.融合外部注意力机制的序列到点非侵入式负荷分解[J].上海交通大学学报,2024,58(6):846-854.
LI L J, LIU H, LIU H L, et al. Non-intrusive load disaggregation using sequence-to-point integrating external attention mechanism [J]. Journal of Shanghai Jiao Tong University, 2024, 58(6): 846-854.
- [3] ZHANG Y M, YANG G H, MA S D. Non-intrusive load monitoring based on convolutional neural network with differential input[J]. Procedia cirp, 2019, 83: 670-674.
- [4] 林顺富,詹银枫,李毅,等.基于 CNN-BiLSTM 与 DTW 的非侵入式住宅负荷监测方法[J].电网技术,2022,46(5):1973-1981.
LIN S F, ZHAN Y F, LI Y, et al. Non-intrusive residential load monitoring method based on CNN-BiLSTM and DTW [J]. Power system technology, 2022, 46(5): 1973-1981.
- [5] ZHANG Y S, QIAN W Y, YE Y J, et al. A novel non-

- intrusive load monitoring method based on ResNet-seq2seq networks for energy disaggregation of distributed energy resources integrated with residential houses [J]. *Applied energy*, 2023, 349: 121703.
- [6] YIN L F, GE W. Mobileception-ResNet for transient stability prediction of novel power systems [J]. *Energy*, 2024, 309: 133163.
- [7] ZHOU X X, FENG J R, LI Y. Non-intrusive load decomposition based on CNN - LSTM hybrid deep learning model[J]. *Energy reports*, 2021, 7: 5762-5771.
- [8] HWANG H, KANG S. Nonintrusive load monitoring using an LSTM with feedback structure[J]. *IEEE transactions on instrumentation and measurement*, 2022, 71: 2509111.
- [9] 周亚军. 面向单体非侵入式负荷监测的神经网络数据特征分析[D]. 北京: 华北电力大学, 2023.
- ZHOU Y J. Analysis of neural network data characteristics for non-invasive load monitoring of single unit [D]. Beijing: North China Electric Power University, 2023.
- [10] SIROJAN T, PHUNG B T, AMBIKAI RAJAH E. Deep neural network based energy disaggregation [C]//2018 IEEE International Conference on Smart Energy Grid Engineering (SEGE). Oshawa, ON, Canada, 2018: 73-77.
- [11] 莫浩杰, 彭勇刚, 蔡田田, 等. 一种基于时频特征融合和极限学习机的非侵入式负荷识别方法[J]. *电工电能新技术*, 2023, 42(3): 85-96.
- MO H J, PENG Y G, CAI T T, et al. Nonintrusive load identification method based on time-frequency feature fusion and extreme learning machine [J]. *Advanced technology of electrical engineering and energy*, 2023, 42(3): 85-96.
- [12] 邓旭晖, 陈中, 杨凯, 等. 基于多任务学习卷积网络的非侵入式负荷监测方法[J]. *电力系统自动化*, 2023, 47(8): 189-197.
- DENG X H, CHEN Z, YANG K, et al. Non-intrusive load monitoring method based on multi-task learning convolutional network [J]. *Automation of electric power systems*, 2023, 47(8): 189-197.
- [13] 刘仲民, 侯坤福, 高敬更, 等. 基于时间卷积神经网络的非侵入式居民用电负荷分解方法[J]. *电力建设*, 2021, 42(3): 97-106.
- LIU Z M, HOU K F, GAO J G, et al. Non-intrusive residential electricity load disaggregation based on temporal convolutional neural network [J]. *Electric power construction*, 2021, 42(3): 97-106.
- [14] 李想, 刘宇航, 张琪, 等. 小波检测和特征图谱决策的非侵入电动自行车充电实时监测系统[J]. *电力系统自动化*, 2023, 47(19): 177-186.
- LI X, LIU Y H, ZHANG Q, et al. Non-intrusive real-time monitoring system for electric bicycle charging based on wavelet detection and feature graph decision [J]. *Automation of electric power systems*, 2023, 47(19): 177-186.
- [15] REN Y H, ZHANG S, WANG J, et al. TSC-MI: a temporal spatial convolution neural network fused with mutual information for motor imagery based EEG classification [C]//Brain Informatics. Cham: Springer, 2021: 474-485.
- [16] 王磊, 马佳琪, 韩肖清, 等. 考虑多状态特征的非侵入式负荷识别方法[J]. *电网技术*, 2024, 48(11): 4720-4728.
- WANG L, MA J Q, HAN X Q, et al. Non-intrusive load monitoring method with multi-state characterization of loads [J]. *Power system technology*, 2024, 48(11): 4720-4728.
- [17] NI Y N, XIA Y H, LI Z C, et al. A non-intrusive identification approach for residential photovoltaic systems using transient features and TCN with attention mechanisms[J]. *Sustainability*, 2023, 15(20): 14865.
- [18] 马乐乐, 孔小兵, 郭磊, 等. 基于最大重叠离散小波变换和深度学习的光伏功率预测[J]. *太阳能学报*, 2024, 45(5): 576-583.
- MA L L, KONG X B, GUO L, et al. Photovoltaic power forecasting based on maximum overlap discrete wavelet transform and deep learning [J]. *Acta energiae solaris sinica*, 2024, 45(5): 576-583.
- [19] KONONENKO I. Estimating attributes: analysis and extensions of RELIEF [C]//Machine Learning: ECML-94. Berlin, Heidelberg: Springer, 1994: 171-182.
- [20] 温廷新, 郭晓赛. 基于特征组合的ECA-TCN光伏发电功率预测模型[J]. *太阳能学报*, 2024, 45(12): 94-100.
- WEN T X, GUO X S. Eca-tcn photovoltaic power prediction model based on feature combination [J]. *Acta energiae solaris sinica*, 2024, 45(12): 94-100.
- [21] LIANG J, NG S K K, KENDALL G, et al. Load signature study: part I: basic concept, structure, and methodology [J]. *IEEE transactions on power delivery*, 2010, 25(2): 551-560.
- [22] MONTEMURRO A, ISAKSEN J L, HANSEN T, et al. Temporal convolutional networks for automatic ECG features extraction[J]. *Journal of electrocardiology*, 2019, 57: S106-S107.
- [23] 时培明, 郭轩宇, 杜清灿, 等. 基于TCN-BiLSTM-Attention-ESN的光伏功率预测[J]. *太阳能学报*, 2024, 45(9): 304-316.
- SHI P M, GUO X Y, DU Q C, et al. Photovoltaic power prediction based on TCN-BiLSTM-Attention-ESN [J]. *Acta*

energiae solaris sinica, 2024, 45(9): 304-316.

- [24] 刘云翔, 马海力, 朱建林, 等. 基于感受野注意力卷积的自动驾驶多任务感知算法[J]. 计算机工程与应用, 2024, 60(20): 133-141.

LIU Y X, MA H L, ZHU J L, et al. Autonomous driving multi-task perception algorithm based on receptive-field attention convolution [J]. Computer engineering and applications, 2024, 60(20): 133-141.

NON-INTRUSIVE LOAD DECOMPOSITION APPROACH BASED ON RFAM-TCN MODEL AND MULTI-MODAL FEATURE OPTIMIZATION

Wu Zhengtao¹, Li Zichen², Xia Yanghong², Ni Yini²

(1. Polytechnic Institute of Zhejiang University, Zhejiang University, Hangzhou 310015, China;

2. College of Electrical Engineering, Zhejiang University, Hangzhou 310027, China)

Abstract: To achieve accurate load awareness and efficient device control, this study proposes a novel non-intrusive load decomposition (NILD) method by combining multi-modal feature selection techniques with an improved time convolutional network to enhance identification accuracy and efficiency. In this method, the load multi-scale maximum overlapping discrete wavelet transform results are introduced as input features, and the feature is combined with the mRMR algorithm to ensure the effective integration of multi-modal features. At the same time, the adaptive receptive field attention mechanism (RFAM) is added to the TCN method, improving the ability to extract features at different time scales. Finally, the algorithm is validated on datasets such as PLAID. The results show that the proposed method can achieve more accurate and efficient load decomposition with fewer features, which is higher than the regression accuracy of the traditional BiLSTM model of 5.31%, and the detection rate of new energy household equipment such as photovoltaic power has also reached over 96.6%.

Keywords: electric load management; wavelet transforms; feature selection; non-intrusive load decomposition; receptive field attention mechanism; temporal convolutional network