

融合小波散射与 注意力时间卷积的非侵入式负荷监测方法

鄢仁武¹, 田堃霖¹, 吴慧敏¹, 梁涔², 李培强¹

(1. 福建理工大学 电子电气与物理学院, 福建 福州 350118;

2. 国网福建省电力有限公司 福州市长乐区供电公司, 福建 福州 350202)

摘要: 面对建筑能源消耗的持续增长和对精细用电管理的迫切需求, 现有的负荷监测方法在处理复杂信号时的计算复杂度和存储成本较高。为此, 提出一种经济高效且易于实现的非侵入式电力负荷监测方法。首先, 采用融合小波散射与时间卷积的神经网络架构精确捕捉总负荷数据的上下文信息; 通过残差连接实现多尺度特征的整合, 从而增强模型对负荷信号细节和整体模式的识别能力。其次, 引入高效通道注意力网络, 提升模型对负荷关键特征的识别精度, 并提高网络的训练效率。最后, 通过结合多任务学习框架和阈值分析法技术, 有效降低误判率, 增强对电器特征的识别准确性。在 UK-DALE 数据集上的比较实验结果验证了该方法相较于现有典型方法在负荷分解和负荷辨识任务上的优越性能。

关键词: 非侵入式负荷监测; 多任务学习; 时间卷积; 小波散射; 注意力网络

中图分类号: TP391

文献标志码: A

文章编号: 2097-3853(2026)01-0001-10

Non-intrusive load monitoring method integrating wavelet scattering and attention time convolution

YAN Renwu¹, TIAN Kunlin¹, WU Huimin¹, LIANG Cen², LI Peiqiang¹

(1. School of Electronic, Electrical Engineering and Physics, Fujian University of Technology, Fuzhou 350118, China;

2. Changle District Power Supply Company, State Grid Electric Power Co., Ltd., Fuzhou 350202, China)

Abstract: The continuous rise in building energy consumption and the growing demand for precise electricity management have exposed the limitations of existing load monitoring methods, which incur high computational and storage costs. To address these issues, a cost-effective and easily implementable non-intrusive load monitoring method is proposed. First, a neural network that integrates wavelet scattering and time convolution is used to capture the contextual information of total load data accurately. The integration of multi-scale features is realized by residual connection, so as to enhance the model's ability to identify both detailed and overall load patterns. Second, an efficient channel attention network is introduced to improve the recognition accuracy of the model for the key features of the load and improve the training efficiency of the network. Finally, the integration of a multi-task learning framework and threshold analysis method effectively reduces false detection rates and improves appliance recognition accuracy. Comparative experiments on the UK-DALE dataset demonstrate that the proposed model outperforms existing models in load disaggregation and identification tasks.

Keywords: non-intrusive load monitoring; multi-task learning; time convolution; wavelet scattering; attention network

收稿日期: 2025-01-06

基金项目: 国家自然科学基金资助项目(52377097)

第一作者简介: 鄢仁武(1981—), 男, 福建福州人, 副教授, 博士, 研究方向: 智能电网应用。

非侵入式负载监测 (NILM) 技术通过分析建筑物的综合电力消耗数据,可以精确地推断出单个设备的能耗情况,不仅有助于用户有效降低能源消耗,还能帮助电网运营商进行电网管理、识别故障电器,并分析电器的使用模式。因此 NILM 技术在多个领域中被推广,在节能监测、智能家居管理、电力系统优化等领域展现出巨大的应用潜力。

NILM 的主要研究问题之一是检测目标设备的开关状态。Yang 等^[1]利用多标签分类器识别设备的开关状态,通过状态与设备的平均能耗相乘采用均值估算的方法估算设备的实时能耗,但该方法精度不足。Gillis 等^[2]提出了一种基于半监督学习的多标签深度学习方法,能够同时从有标签和无标签的数据中学习特征,有效降低对大规模标记数据的依赖,但无标签数据容易导致模型准确度降低。He 等^[3]采用条件随机场 (CRF) 方法捕捉状态转换和预测状态分布的双卷积模型,证明了多状态设置在减少训练数据样本方差方面的优势,但其计算过于复杂,且需要手动调整大量参数。

NILM 的研究问题也被构建为估计目标电器实时能耗的回归问题。Kelly 等^[4]结合了 3 种前沿的神经网络架构和序列到序列学习技术,在 UK-DALE 数据集上的表现出色。Zhang 等^[5]探索了序列到点学习的方法,通过结合卷积神经网络 (CNN) 和序列到点学习技术,达到了领先的水平。两种方案都采用了大规模参数化的神经网络模型。受语音合成领域研究的启发,Jiang 等^[6]提出了基于因果卷积的神经网络,比较研究了两种不同深度学习框架在设备开关状态检测任务上的性能,但该方法的状态检测任务仅依赖于设备的开启功率阈值,在多状态设备上效果不明显。

综上,现有基于深度学习的方法仍面临以下问题:(1)数据规模庞大,为了获得更丰富的监测信息以实现更准确的负荷分析,NILM 方法呈现出了“时空密集化”采集的特点,监测地点日益广泛、监测间隔日益密集,进而导致了数据海量化和采样高频化等问题^[7];(2)算法模型复杂度高,精确的监测结果常依赖于冗余、庞大的计算模型,例

如神经网络模型通常需要大量的存储成本和计算资源需求。

针对以上问题,本研究综合考虑算法性能与计算开销,在适当保持神经网络模型性能的同时,降低模型的内存和计算开销,提出一种轻量化的基于多尺度特征融合与多任务学习框架的非侵入式负荷监测方法。该方法的结构主要由小波散射网络、时间卷积及通道注意力机制为基础模块构成,并采用多任务学习框架执行负荷监测任务。

1 非侵入式负荷监测理论及分析

1.1 负荷分解原理

负荷分解任务旨在精确地区分和监测各个电器的能耗情况^[8]。在假定单个传感器在离散时间点 t 收到由源信号 $X(t)$ 形成的观测信号 $Y(t)$ 的情况下,负荷分解问题可以表述为:

$$Y(t) = \sum_{i=1}^I X_i(t) + e_t \quad (1)$$

其中,残差项 e_t 反映了在 t 时刻无法确定的电器能耗和背景噪声, i 的取值范围是 $\{1, 2, \dots, I\}$,即数量为 I 个不同类型的电器。

负荷分解的目标是构建一个能够从总的能耗读数 Y 中提取出各个电器 X_i 的能耗的模型。即求解以下分解映射关系:

$$f_i: Y \rightarrow X_i \quad (2)$$

其中,每个映射 f_i 的作用是将总的聚合读数序列 Y 转换为单个电器的能耗序列 X_i 。

1.2 负荷辨识原理

大多数设备都有一个启动/停止的能耗检测阈值表示设备启动所需的最小能量。如果设备的能耗低于该启动阈值,则认为设备处于关闭状态;反之则认为设备处于开启状态。负荷辨识的目标是基于总的能耗数据来估计电器的开启或关闭状态^[9]。

在形式上,考虑电能表的总读数序列 $Y = (y_1, y_2, \dots, y_T)$,其中 T 是序列的时间样本索引, y_i 表示某一时刻的能耗序列。开/关检测任务的目标是构建一个模型来识别单个电器的开关状态序列,表示为 $Z_i = (Z_{i1}, Z_{i2}, \dots, Z_{iT}), Z_{iu} \in \{0, 1\}, i \in \{1, \dots, I\}, t \in \{1, \dots, T\}$ 。其中 I 是已知电器的数

量, T 是时间域样本的索引, 0 和 1 分别代表关闭和开启状态。即求解以下映射关系:

$$f_i: Y \rightarrow Z_i \quad (3)$$

其中, 每个映射 f_i 从聚合读取序列 Y 映射到设备 Z_i 的开启/关闭状态序列。

2 融合小波散射与注意力时间卷积模型及方法

本研究提出的非侵入式电力负荷监测方法的模型详细结构如图 1 所示。

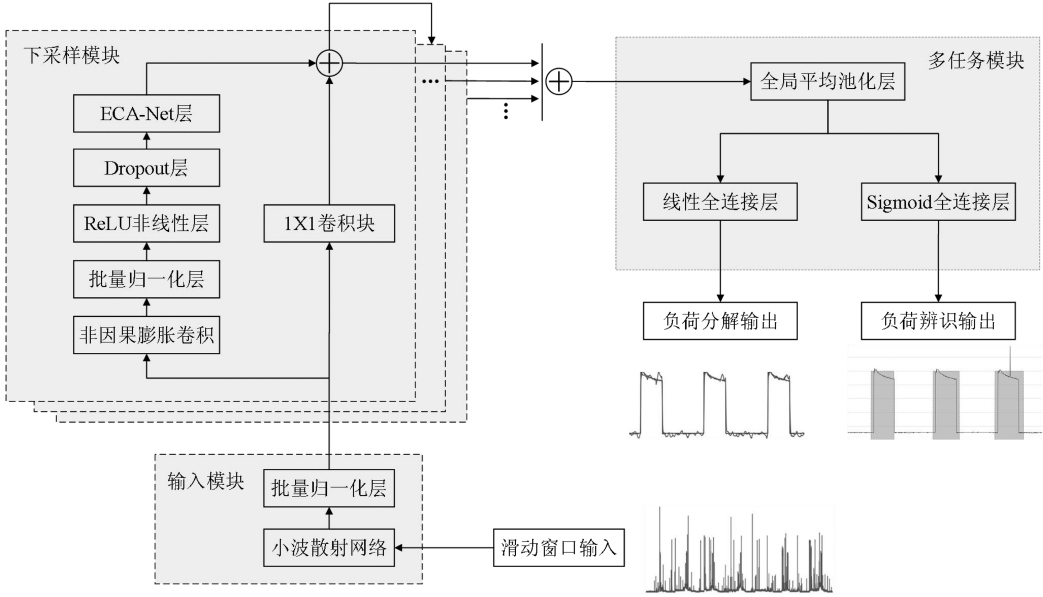


图 1 模型结构

Fig.1 Model structure

2.1 负荷序列输入模块

负荷序列输入模块采用固定滑动窗口下的功率序列片段为网络输入, 使用小波函数提取信号不同频度下的局部负荷特征。其中, 小波散射网络的计算结构在形式上与卷积神经网络相似, 区别在于小波散射网络不依赖于参数学习, 从而摒弃了繁琐的模型训练过程。小波散射网络采用分阶段处理方法, 每个阶段由 3 个操作组成, 如图 2 所示, 小波散射的级联过程分支的复杂度由滤波器数量和散射层数决定。

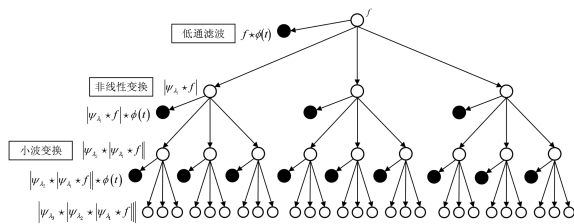


图 2 小波散射变换网络图

Fig.2 Wavelet scattering transform network diagram

2.1.1 卷积操作(小波系数)

通过调整母小波的旋转角度和缩放比例, 形

成一簇小波函数 $\psi_{j,k}(t)$, 其中 j 表示尺度, k 表示平移。这些小波与输入信号进行卷积运算, 得到小波特征系数即尺度图系数, 该步骤可以捕获在信号中的高频细节。

给定一个输入信号 f , 定义其 J 阶散射系数对应的频率序列 $(\lambda_1, \dots, \lambda_J)$, 其中 ψ 表示母小波函数, $\lambda_j < J$, λ_j 表示小波的缩放比例, 决定分析的频率范围, C_j 表示尺度 j 的小波变换算子, 其输出 $C_j f(t)$ 表示信号在该尺度下的小波特征系数:

$$C_j f(t) = \psi_{\lambda_j}(t) * f \quad (4)$$

2.1.2 非线性变化(取模)

对小波变换的尺度图系数施加非线性算子(模数运算), 该步骤基于小波系数的零均值特性。非线性模运算能够产生具有信息量的非零系数, 为信号的特征提取提供了重要信息。 U_j 表示小波变换模系数, 小波变换模结果为:

$$U_j f(t) = |\psi_{\lambda_j}(t) * f| \quad (5)$$

2.1.3 平均池化(低通滤波)

平均池化阶段通过卷积和低通滤波器对非线性处理后的小波特征进行平均化处理, 从而有助

于提取具有平移不变性和对局部形变稳定性的小波特征。平均化操作可表示为:

$$S_j f(t) = |f * \psi_{\lambda_j}(t)| * \varphi(t) \quad (6)$$

其中, S_j 表示散射变换中尺度 j 的散射系数, $\varphi(t)$ 表示低通平滑滤波器即尺度函数。将所有 k 阶段的散射系数进行汇总, 形成最终的散射变换特征, 可以表示为:

$$Sf[\lambda_1, \lambda_2 \dots, \lambda_j, t] = \{S_{1f}, S_{2f}, \dots, S_{jf}\} \quad (7)$$

其中, $Sf[\lambda_j, t]$ 是散射变换的最终输出, 包括一系列散射系数 $S_j f(t)$ 的集合, 涵盖了不同的尺度 j 。

在理想情况下, 小波散射变换后各尺度下的小波系数是稀疏的, 但由于复杂的信道干扰和测量系统误差的存在, 采集信号中携带的噪声会降低小波系数整体的稀疏性, 导致小波系数表现为近似分量的形式, 并具有一定的结构特征。而深度神经网络能够识别特殊结构并去除信号噪声, 故利用下采样模块进一步恢复小波系数的稀疏性, 以此达到去噪的目的。

2.2 下采样网络模块

本模块采用 3 个不同的网络为基础模块搭建。

2.2.1 非因果膨胀卷积网络

因果卷积通过增加每层的扩张因子为 1, 逐步增加感受野, 从而在不增加计算复杂度的前提下增强网络对时间序列数据的分析能力。通过引入空洞卷积技术, 卷积核能够覆盖更广的输入范围, 从而提高模型对长距离依赖特征的学习能力。非侵入式负荷监测通常不需要考虑因果关系^[10], 因此本研究利用未来时间步的信息来优化当前时刻的预测, 从而实现更准确的负荷监测结果^[11]。

2.2.2 跨层连接网络

在本研究的模型中, 每一层的残差块输出都成为下一个残差块的输入, 不仅提高了网络的深度, 还缓解了梯度消失和梯度爆炸的问题。为了保留原始数据的细节和语义信息, 使用了跳跃连接将每一层下采样模块的输出直接添加到后续层的输出上进行求和操作, 使得网络能够整合不同层级的输出, 更充分地利用低级特征, 增强了模型对多尺度特征的捕捉能力, 并大幅减少计算的繁琐程度, 从而提升了模型训练速度。

2.2.3 高效通道注意力网络

通道注意力机制通过网络学习自动为每个特征通道分配权重, 以突显对当前任务更为重要的

特征, 使得网络能够集中关注对任务贡献较大的特征, 进而提升模型的整体性能。Wang 等^[12] 提出了高效通道注意力网络 (ECA-Net), 针对 CNN 进行了改进, 避免了使用全连接层和全局平均池化层, 直接通过一维卷积来学习每个通道的权重。通过这种方式, ECA-Net 显著减少了计算开销, 去除了繁琐的降维步骤, 在保障通道间信息交互的基础上降低了模型复杂度。

联合使用上述 3 个模块, 能够提取输入样本在多个感受野尺度下的详细特征。将卷积结果按对应通道进行加权组合, 能够在不增加特征维度的情况下, 提升整体特征的分辨率和语义丰富度, 相比直接拼接, 减少了后续模块的计算复杂度。

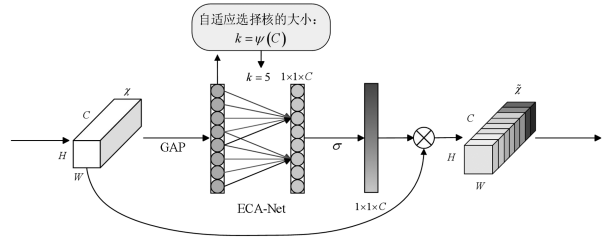


图 3 通道注意力网络结构图

Fig.3 Channel attention network structure

2.3 多任务学习模块

通过利用辅助任务提供的信息, 可以增强模型在目标任务上的表现^[13]。在 NILM 领域中, 负荷分解与负荷辨识两个任务之间是强相关的, 二者可以共享部分特征信息。在本研究提出的模型架构中, 卷积层被设计为在所有子任务中共享, 而每个子任务的特定需求则由各自独立的全连接层来满足。

负荷分解属于典型的回归任务, 使用均方误差 (MSE) 作为其损失函数, 以衡量预测值与实际值之间的差异。回归分支的损失函数具体为:

$$l_{\text{reg}} = \frac{1}{N_{\text{train}}} \sum_{j=1}^{N_{\text{train}}} \frac{1}{T_{\text{out}}} \sum_{t=1}^{T_{\text{out}}} (y_{j,t} - \hat{y}_{j,t})^2 \quad (8)$$

其中, T_{out} 表示输出序列滑动窗口的长度, 在本研究使用的序列到点方法中 T_{out} 大小为 1。 N_{train} 表示训练数据的批量大小, $y_{j,t}$ 和 $\hat{y}_{j,t}$ 表示负荷分解的回归结果和设备的真实功率负荷。

负荷辨识是一个二分类任务。通过对数据集标签的统计分析表明, 某些设备例如洗碗机和洗衣机的启动频率较低, 属于典型的类别不平衡的问题, 因此采用权重化二元交叉熵损失 (BCE loss)^[14], 以增加对正样本的重视程度, 并按照每

个设备启动时间比例来计算权重比例,二元交叉熵损失函数具体表示为:

$$l_{\text{class}} = \frac{1}{N_{\text{train}}} \sum_{j=1}^{N_{\text{train}}} \frac{1}{T_{\text{out}}} \sum_{t=1}^{T_{\text{out}}} [y_{j,t} \cdot \log \hat{y}_{j,t} \cdot W_{\text{pos}} + (1 - y_{j,t}) \cdot \log(1 - \hat{y}_{j,t}) \cdot W_{\text{neg}}] \quad (9)$$

其中, W_{pos} 表示标签为正样本的权重值, W_{neg} 表示标签为负样本的权重值。

负荷分解和负荷辨识两个目标任务合成网络总损失以进行多任务学习模型联合训练。总损失 L_{total} 为:

$$L_{\text{total}} = \lambda L_{\text{reg}} + (1 - \lambda) L_{\text{class}} \quad (10)$$

其中, λ 为分解损失比例系数, $0 \leq \lambda \leq 1$ 。当 $\lambda = 1$ 或者 $\lambda = 0$ 的时候,多任务学习的模型将等价于单独的回归网络或者分类网络。

相对于单任务学习方式,通过共享提取特征的多任务学习能够极大减少计算的复杂度和网络的冗余程度,不同任务之间能够互相制约,防止模型发生过拟合;同时,针对不同任务的特殊性分别设计对应的输出层,能够体现出不同任务之间的异质性,提高模型处理对应任务的准确性。

2.4 负荷辨识结果的阈值分析法

非侵入式负荷监测获得的原始数据通常主要记录建筑总功耗,而非单个设备的能耗状态。为了将这些连续的功耗信号转换为可分类的状态,采用阈值化方法来界定设备的运行状态^[9]。

中点阈值法首先通过分析训练集中设备 i 的所有功率值集合,接着采用聚类方法将这些数据点分为两个群集,并识别出代表 m_0^i (对应关闭状态) 和 m_1^i (对应开启状态) 两个群集的质心。这两个状态的质心代表了各自状态的典型功率水平,最终将设备 i 的阈值 λ^i 设定在两个质心的中间值:

$$\lambda^i = \frac{1}{2} (m_0^i + m_1^i) \quad (11)$$

方差敏感阈值法在中点阈值法的基础上增加了对数据方差的考量,不仅依据聚类质心来确定状态,还考虑了群集内数据点的波动性,即考虑每个群集内数据点的平均值 m_k^i (表示设备 i 第 k 个群集的数据点集中趋势) 和标准差 σ_k^i (表示设备 i 数据点在第 k 个群集的离散程度)。如果 $\sigma_1^i > \sigma_0^i$, 那么阈值应该向 m_0^i 方向进行调整,防止距离质心 m_1^i 较远的点被错误地分类:

$$\lambda^i = \left(1 - \frac{\sigma_0^i}{\sigma_0^i + \sigma_1^i} \right) m_0^i + \frac{\sigma_0^i}{\sigma_0^i + \sigma_1^i} m_1^i \quad (12)$$

中点阈值法和方差阈值法对于处理单状态设备具有良好的效果,但在处理多状态设备时,没有考虑到多个状态之间的转换,以及每个状态的持续过程。在电力负荷数据集中,设备的用电行为通常呈现出明显的时间周期性和峰值特征。针对上述两种方法的不足,本研究采用激活时间阈值法,激活时间阈值法同时设定了功率阈值和时间阈值以获得更完整的负荷特征。规定设备必须保持某一状态的最短时间长度,如果一系列功率测量值在某一时间段内时低于功率阈值,则认为该序列仍处于前一状态。

3 算例分析

3.1 负荷数据预处理

为验证所提方法的有效性,考虑使用 UK-DALE 数据集中的水壶、微波炉、洗碗机、洗衣机、冰箱 5 台设备的电力数据进行模型训练与测试。为了避免模型过拟合到特定家庭的噪声或特定电器特征,采用房屋 1 的负荷数据作为训练集,验证集按 5 : 2 划分,而房屋 2 的负荷数据作为测试集,验证模型在未知环境中的泛化能力。

针对负荷辨识任务,使用了阈值化方法将连续的电力数据转化为离散的状态标签。在设备实时运行的状态下同时考虑开启持续时间和关闭持续时间,即:

$$S_t^i = \begin{cases} 1, & P_t^i \geq P_{\text{ON}}^i, T_{\text{ON}}^i \leq T_n^i \\ 0, & P_t^i < P_{\text{ON}}^i, T_{\text{OFF}}^i \leq T_f^i \end{cases} \quad (13)$$

其中, P_t^i 表示设备 i 在 t 时刻的负荷功率值, P_{ON}^i 表示判断该设备处于工作状态的负荷功率阈值, T_{ON}^i 表示判断该设备处于工作状态的时间阈值, T_{OFF}^i 表示判断该设备处于关闭状态的时间阈值, T_n^i 表示当设备功率达到开启阈值的持续时间, T_f^i 表示当设备功率未达到开启阈值的持续时间, S_t^i 表示在 t 时刻设备 i 的工作状态标签值。

3.2 评价指标

平均绝对值误差 (MAE) 能够客观地反映预测值与实际值之间的平均偏差,如式 (14) 所示。

$$\text{MAE} = \frac{1}{T} \sum_{t=1}^T |y_t - \hat{y}_t| \quad (14)$$

其中, y_t 表示预测值, \hat{y}_t 表示真实值或标签值, T 表示为滑动窗口的序列长度。

信号聚合误差 (SAE) 通过累加一段时间内预

测值与实际值之间的差异来评估算法性能,如式(15)所示。

$$\begin{cases} \text{SAE} = \frac{|\hat{a} - a|}{a}, \\ \hat{a} = \sum_t^T \hat{y}_t, \\ a = \sum_t^T y_t \end{cases} \quad (15)$$

其中, \hat{a} 表示一段时间内真实值的总和, a 表示一段时间内预测值的总和。

对于负荷辨识任务,本研究使用 F_1 分数来评估不同算法的性能。 F_1 分数是准确率和召回率的调和平均, F_1 越高代表算法性能越优,如式(16)所示。

$$\begin{cases} F_1 = 2 \cdot \frac{p \cdot r}{p+r}, \\ p = \frac{\text{TP}}{\text{TP}+\text{FP}}, \\ r = \frac{\text{TP}}{\text{TP}+\text{FN}} \end{cases} \quad (16)$$

其中,TP 表示真正例即被算法预测为正类的正样本,FP 表示假正例即被算法预测为正类的负样

本, FN 表示假反例即被算法预测为负类的正样本, p 表示查准率, r 表示召回率。

3.3 实验结果分析

本研究选取 Seq2seq^[4]、Seq2point^[5] 和 DAE^[4] 模型作为基准模型,并且将小波散射模块与 Seq2point 和 DAE 模型融合,即 WST-Seq2point 和 WST-DAE,观察小波散射网络在不同模型中的实际效果。如图 4 所示,所提的方法无论在设备开启或关闭状态下均能提供准确的分解结果。图 4(d) 中存在短暂的突变功率,并且冰箱的启停功率相对总功率较小,尽管预测曲线存在些许波动,但所提方法依旧能够较准确地识别。这是由于冰箱的功率曲线存在固定的运行模式特征,而本研究的下采样模块能够准确识别这种模式。图 4(e) 中 152 200 s 处也存在总功率突变,由于其平均功率过高且持续时间过短,并不符合洗碗机的功率曲线特征,因此并没有发生误判。洗衣机的负荷分解任务表现得相对复杂,一是由于洗衣机的使用频率相对较低,导致可用于训练的样本数量有限;二是由于洗衣机的工作模式较为复杂,其功耗在启动期间会出现频繁波动。

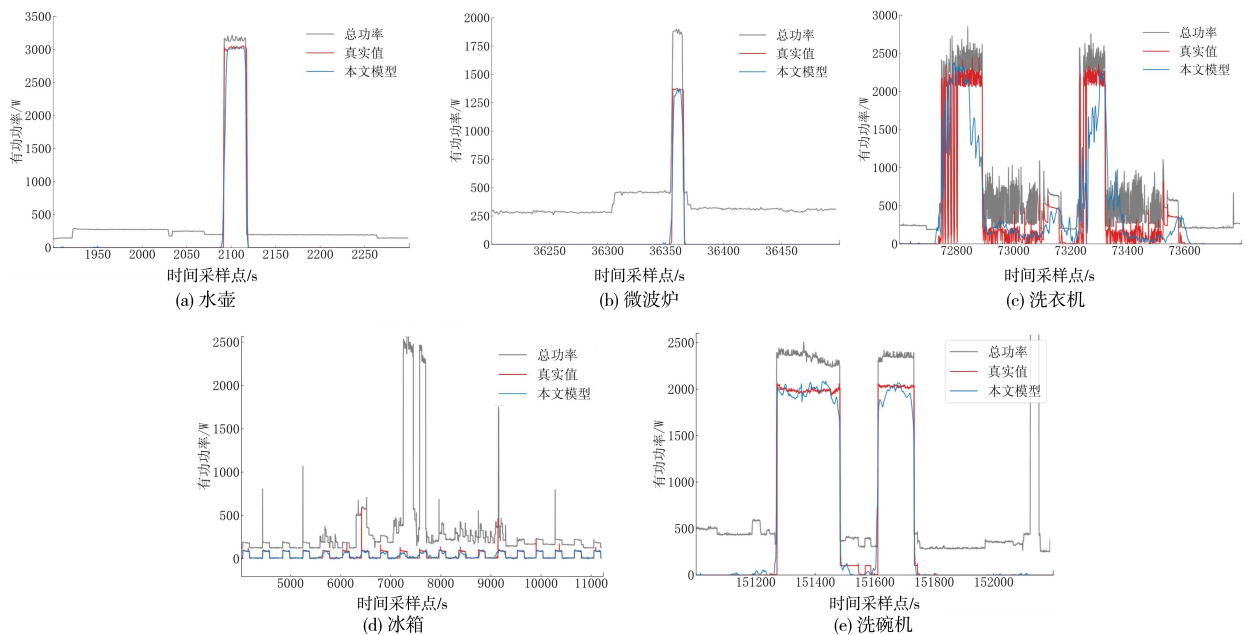


图 4 负荷分解的结果

Fig.4 Results of load decomposition

不同的 NILM 方法在负荷分解评价指标上的表现如表 1 所示。与 Seq2seq、Seq2point、DAE 相比,所提方法提高了在所有设备上的指标性能,MAE 总体分别降低了 37%、26%、39%。在 SAE

指标上,所提方法将性能总体分别提高了 58%、44%、51%,进一步验证了其在处理不同设备负载特征时的有效性和泛化能力。

表 1 不同 NILM 方法在 UK-DALE 数据集上的表现

Tab.1 Performance of different NILM models on UK-DALE dataset

方法 指标	本研究模型		Seq2point		Seq2seq		DAE		WST-Seq2point		WST-DAE	
	MAE	SAE	MAE	SAE	MAE	SAE	MAE	SAE	MAE	SAE	MAE	SAE
水壶	5.825	0.041	6.026	0.053	5.276	0.003	12.531	0.231	14.431	0.199	10.709	0.108
冰箱	18.446	0.076	20.894	0.121	24.489	0.373	19.573	0.259	21.344	0.112	22.323	0.121
洗碗机	21.153	0.336	27.704	0.645	32.515	0.779	26.652	0.254	19.977	0.186	25.421	0.231
洗衣机	7.834	0.182	12.663	0.284	10.153	0.453	18.834	0.495	7.631	0.285	5.847	0.188
微波炉	3.698	0.172	8.661	0.486	13.619	0.205	16.753	0.445	6.013	0.182	6.319	0.263
平均值	11.391±	0.162±	15.472±	0.321±	17.999±	0.423±	18.869±	0.336±	13.879±	0.193±	14.124±	0.182±
	5.232	0.204	7.718	0.217	9.063	0.194	5.284	0.218	6.537	0.334	6.243	0.287

不同设备下负荷辨识的性能如表 2 所示。现有的方法大都采用文献[4]中经过实验统计的经

验阈值,因此表 2 也采用了经验阈值法作为基准方法。

表 2 不同方法的分类效果(经验值)

Tab.2 Classification effects of different models (empirical value)

对象	F_1 分数					
	本研究模型	Seq2point	Seq2seq	DAE	WST-Seq2point	WST-DAE
水壶	0.947	0.299	0.468	0.926	0.870	0.864
冰箱	0.816	0.869	0.880	0.801	0.754	0.705
洗碗机	0.770	0.561	0.510	0.462	0.507	0.704
洗衣机	0.734	0.077	0.224	0.159	0.335	0.845
微波炉	0.701	0.799	0.826	0.624	0.602	0.434
平均值	0.794±0.223	0.521±0.424	0.582±0.312	0.598±0.244	0.613±0.246	0.710±0.259

在 F_1 分数指标中,Seq2seq、Seq2point 和 DAE 在所有设备上的性能基本相同,但是所提方法在 F_1 分数上实现了显著提升,证明了其在负荷辨识方面的优越性能。

几种不同的阈值判断法应用在不同设备上的效果如表 3 所示。其中,在洗衣机上采用不同阈值法判断设备状态的具体效果如图 5 所示。

从表 3 和图 5 可知,对于工作频率、用电功率较为明显和周期较为固定的设备,一般使用中点阈值法已能达到良好的效果;方差敏感阈值法在某些具有间歇性周期特点的设备上,相比于中点阈值法会有更良好的效果。而激活时间阈值法相比其他的阈值判断方法,在性能指标上均有所提升,验证了其优越性。

表 3 不同阈值方法的分类 F_1 -score 对比Tab.3 Comparison of classification F_1 -scores of different threshold methods

对象	F_1 分数			
	经验 值法	中点 阈值法	方差敏感 阈值法	激活时间 阈值法
水壶	0.712	0.958	0.939	0.947
冰箱	0.816	0.808	0.794	0.816
洗碗机	0.729	0.859	0.878	0.982
洗衣机	0.493	0.695	0.761	0.861
微波炉	0.627	0.638	0.677	0.701
平均值	0.675±	0.791±	0.809±	0.861±
	0.485	0.187	0.205	0.193

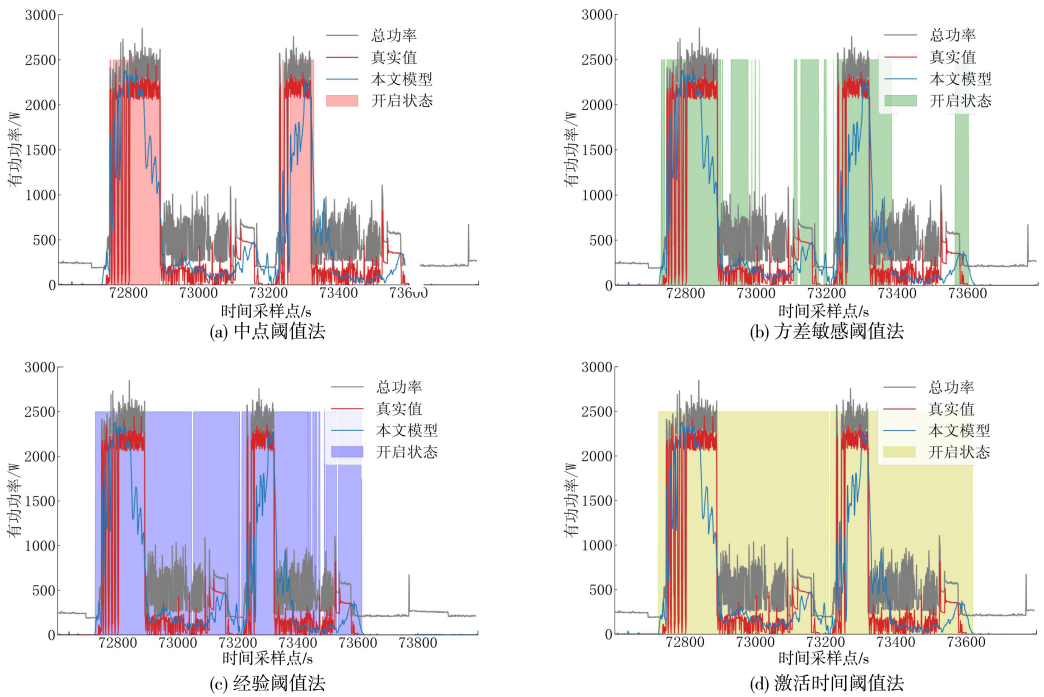


图 5 负荷辨识结果(洗衣机)

Fig.5 Load identification results (washing machines)

3.4 消融实验分析

为了系统评估模型中各个模块的贡献,本研

究在 UK-DALE 数据集上进行了消融实验。其中,水壶设备的消融实验可视化结果如图 6 所示。

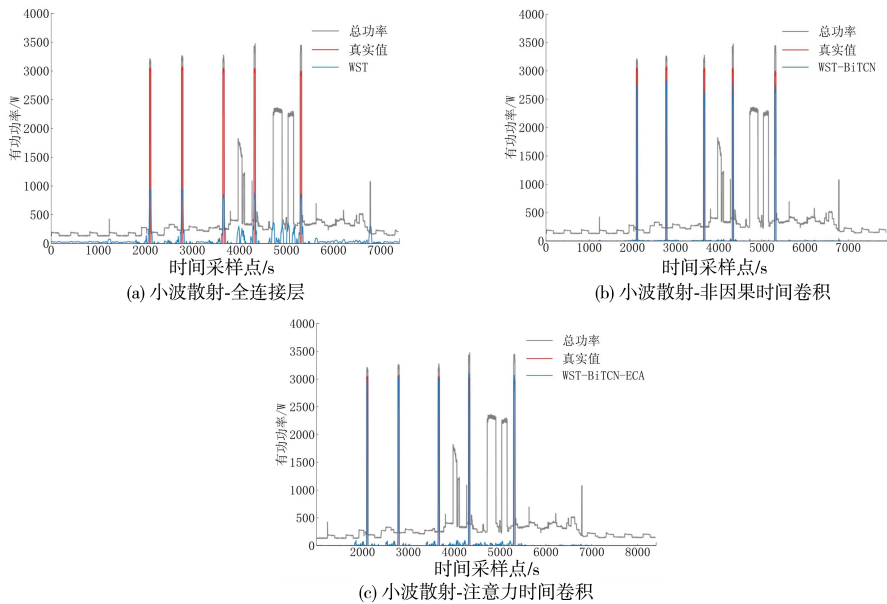


图 6 消融实验的结果(水壶)

Fig.6 Results of ablation experiments (kettle)

首先,信号经过小波散射网络后使用全连接网络进行非线性映射,即 WST-FC。从图 6 可知,即便在只使用小波散射和全连接网络的简化模型

中,也能有效识别负荷特征,验证了小波散射网络在特征提取上的实用性。但 WST-FC 对于突变的总功率信号没有表现出稳健的响应特性,导致输

出假阳性片段的占比较高。

通过引入时间卷积模块即 WST-TCN,模型性能在各指标上相比 WST-FC 都得到提升,显著提高了设备功率曲线的检测精度,有效抑制了瞬态总功率对模型性能的干扰,从而大幅降低了信号突变过程中的误判概率,如表 4 所示。这表明 TCN 在处理具有时间周期性的负荷监测任务上具有相当大的潜力,但在功率的精度大小上仍有所不足。

最后,在 WST-TCN 模型的基础上增加 ECA-Net 模块即 WST-TCN-ECA 之后,模型整体在 MAE 指标上提升了 19.24%,在 SAE 上提升了 39.16%,其性能展示在图 6 中则表现为拟合了更精准的预测曲线,同时该模型还展现出对瞬态功率扰动及宽频噪声干扰的鲁棒性特征。这得益于 ECA-Net 模块能够为不同特征通道分配适应性权重,增强了对关键特征的关注,同时抑制了相对不重要的特征的权重占比,从而提升了模型的整体性能。

表 4 消融实验的结果

Tab.4 Results of ablation experiments

模型	MAE	SAE	F_1 分数	训练时间/min
WST-FC	69.653	0.856	0.271	2.59
WST-TCN	9.833	0.143	0.923	43.36
WST-TCN-ECA	7.941	0.087	0.947	28.61

3.5 计算复杂度分析

多种 NILM 方法在 UK-DALE 数据集上实验时的资源占用情况如表 5 所示。

表 5 不同 NILM 方法在 UK-DALE 数据集上计算资源

Tab.5 Computing resources of different NILM methods on UK-DALE datasets

方法	参数量/ 10^6	训练时间/min	预测时间/s	模型大小/Mb
Seq2seq	31.32	84.16	28.71	119.00
Seq2point	30.71	10.53	14.25	117.00
DAE	23.96	4.89	11.83	91.40
WST-Seq2point	15.39	7.81	5.01	58.80
WST-DAE	6.25	2.06	3.49	23.90
本研究模型	0.38	24.51	22.29	1.65

所提方法的训练时间和预测时间相对 Seq2point 和 DAE 方法有所增加,主要是由于使用了多个残差连接和跳跃连接,增加了一些中间层的操作,但训练参数相比其他方法显著减少了至少 98%。虽然 Seq2point 和 DAE 方法的训练时间与预测时间比本研究所提方法少,但其可训练参数量也明显更大,因此更容易过拟合。由于 Seq2point 方法的网络架构是使用多个卷积模块进行特征提取,利用多个全连接层进行功率分解,且 Seq2point 使用的是 2D 卷积块,相对于 1D 卷积增加了过多参数,而 DAE 中使用了多个全连接层来筛选特征,导致模型十分臃肿。

本研究还将小波散射网络融合到了 Seq2point 和 DAE 方法中,WST-Seq2point 和 WST-DAE 不仅显著降低了训练参数,还加速了模型的训练和预测时间。因此小波散射网络还具有能够将复杂模型轻量化的作用。

4 结束语

本研究提出的非侵入式负荷监测方法通过引入小波散射网络,有效地降低了模型的输入维度,加速了网络的收敛速度,同时显著减少了训练过程中的时间复杂度和空间复杂度开销。此外,使用非因果时间卷积网络学习小波散射后的特征,并通过加入通道注意力机制,增强了神经网络的表征能力,从而提供更细粒度和更精确的结果;通过对几种不同的阈值判断方法进行比较发现,大部分设备在使用激活时间阈值法时均能够达到最佳的分类效果。在 UK-DALE 数据集上的实验结果表明,本研究方法在负荷分解与负荷辨识任务上实现了与训练有素的深度神经网络相媲美甚至更优的效果。

在低分辨率下的高频功率设备,例如洗衣机等设备,由于工作周期的波动频率较大,导致功率监测的难度进而增加。在后续 NILM 研究中,将考虑结合数据增强等方法进一步增加信号的分辨率,或者采用一些后处理方法以增强方法的准确性。

参考文献:

- [1] YANG Y D, ZHONG J, LI W, et al. Semisupervised multilabel deep learning based nonintrusive load monitoring in smart grids[J]. IEEE Transactions on Industrial Informatics, 2020, 16(11): 6892-6902.
- [2] GILLIS J M, MORSE W G. Non-intrusive load monitoring using semi-supervised machine learning and wavelet design[J]. IEEE Transactions on Smart Grid, 2017, 8(6): 2648-2655.
- [3] HE J L, LIU J M, ZHANG Z J, et al. MSDC: exploiting multi-state power consumption in non-intrusive load monitoring based on a dual-CNN model [C] // AAAI Conference on Artificial Intelligence. Menlo Park: AAAI Press, 2023: 5078-5086.
- [4] KELLY J, KNOTTENBELT W. Neural NILM: deep neural networks applied to energy disaggregation [C] // In Proceedings of the 2nd ACM International Conference on Embedded Systems for Energy-Efficient Built Environments, NY: ACM, 2015: 55-64.
- [5] ZHANG C Y, ZHONG M J, WANG Z Z, et al. Sequence-to-point learning with neural networks for non-intrusive load monitoring[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2018, 32(1): 2604-2611.
- [6] JIANG J, KONG Q Q, PLUMBLEY M D, et al. Deep learning-based energy disaggregation and on/off detection of household appliances[J]. ACM Transactions on Knowledge Discovery from Data, 2021, 15(3): 1-21.
- [7] 袁博, 葛少云, 刘洪, 等. 基于压缩感知的非侵入式负荷监测[J]. 中国电机工程学报, 2025, 45(3): 1205-1219.
- [8] 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.
- [9] PRECIOSO D, GÓMEZ-ULLATE D. Thresholding methods in non-intrusive load monitoring[J]. The Journal of Supercomputing, 2023, 79(13): 14039-14062.
- [10] HUBER P, CALATRONI A, RUMSCH A, et al. Review on deep neural networks applied to low-frequency NILM[J]. Energies, 2021, 14(9): 2390.
- [11] JIA Z Y, YANG L F, ZHANG Z R, et al. Sequence to point learning based on bidirectional dilated residual network for non-intrusive load monitoring[J]. International Journal of Electrical Power & Energy Systems, 2021, 129: 106837.
- [12] WANG Q L, WU B G, ZHU P F, et al. ECA-net: efficient channel attention for deep convolutional neural networks [C] // 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2020: 11531-11539.
- [13] KIM H, PARK J, BENNIS M, et al. Blockchain-based on-device federated learning[J]. IEEE Communications Letters, 2019, 24(6): 1279-1283.
- [14] BRUNDAGE M, AVIN S, CLARK J, et al. The malicious use of artificial intelligence: forecasting, prevention, and mitigation [J]. ArXiv e-Prints, 2018; arXiv: 1802.07228.

(责任编辑: 方素华)