

# 基于ADASYN和图像分析的非侵入式负荷辨识方法研究

顾水福,周磊,李洁,李亚飞,李圆琪,朱超群

(国网江苏省电力有限公司 苏州供电分公司,江苏 苏州 215004)

**摘要:**为推广应用智能电能表负荷辨识技术及解决传统非侵入式负荷辨识算法在不平衡采样数据上辨识精度较低的问题,提出了一种基于自适应合成采样(adaptive synthetic, ADASYN)和图像分析的非侵入式负荷辨识方法。通过马尔可夫变迁场(markov transition field, MTF)编码将一维功率数据转换成二维MTF特征图像,作为图像识别网络的输入。基于密集连接网络(dense connectivity network, DenseNet)的深层信息挖掘能力,将二维图像输入至DenseNet121网络中提取特征信息,实现负荷类型的辨识。基于ADASYN算法对不平衡数据集进行过采样处理,消除数据不平衡分布带来的模型学习偏见。算例结果表明,ADASYN算法能够很好地解决非侵入式负荷监测数据不平衡问题,相对处理前的辨识准确率和 $F_1$ 得分,分别提升了0.247和0.267;同时,MTF图像具有明晰易辨的特征信息,结合DenseNet121网络强大的深层特征捕捉能力,其辨识准确率与 $F_1$ 得分均能达到0.952,有效提升了在不平衡采样数据上非侵入式负荷类型的辨识精度。

**关键词:**智能电能表;非侵入式负荷;自适应合成采样;马尔可夫变迁场;密集连接网络;负荷辨识

## Research on non-intrusive load identification method based on ADASYN and image analysis

GU Shuifu, ZHOU Lei, LI Jie, LI Yafei, LI Yuanqi, ZHU Chaoqun

(Suzhou Power Supply Branch, State Grid Jiangsu Electric Power Co., Ltd., Suzhou 215004, China)

**Abstract:** In order to popularize the load identification technology of smart meters and solve the problem of low identification accuracy of traditional non-intrusive load identification algorithm on unbalanced sampled data, a non-intrusive load identification method based on adaptive synthetic (ADASYN) and image analysis is proposed. 1D power data is converted into 2D MTF feature images by markov transition field (MTF) coding, which is used as the input of image recognition network. Based on the deep information mining capability of dense connectivity network (DenseNet), 2D images are input into DenseNet121 network to extract feature information and realize the identification of load types. Based on ADASYN algorithm, the unbalanced data set is oversampled to eliminate the model learning bias caused by the unbalanced data distribution. The results show that ADASYN algorithm can solve the non-intrusive load monitoring data imbalance problem well, and its identification accuracy and  $F_1$  score are increased by 0.247 and 0.267, respectively. At the same time, MTF images have clear and easily distinguishable feature information. Combined with the powerful deep feature capture capability of DenseNet121 network, the identification accuracy and  $F_1$  score can both reach 0.952, which effectively improves the identification accuracy of non-intrusive load types on unbalanced sampled data.

**Key words:** intelligent energy meter; non-intrusive load; adaptive synthetic; markov transition field; dense connectivity network; load identification

## 0 引言

随着“碳达峰,碳中和”的双碳战略目标的提出,在用户侧完善需求响应机制,发挥需求侧资源削峰填谷、节能降碳成为领域研究重点<sup>[1]</sup>。然而,电力行业造成的碳排放量具有较高占比<sup>[2]</sup>。随着中国居民生活水平的提高,居民用电量在国民经济总用

电量的占比不断增加,由此引发的碳排放问题引起广泛关注。据统计,近年来,我国居民家用电器的耗电量约占全国总用电量的12%<sup>[3]</sup>。以非侵入式负荷监测<sup>[4]</sup>(non-intrusive load monitoring, NILM)为代表的负荷监测技术通过监测各类用电设备的能耗数据,帮助用户培养绿色生活习惯,节约电能损耗,成为实现“碳达峰,碳中和”发展要求的重要举措。

NILM按照用途主要分为非侵入式负荷识别和非侵入式负荷分解。负荷分解通过回归模型从总功耗数据中获取各家用电器功耗情况进而实现其功耗监控的应用功能,而负荷识别侧重于识别出住

宅内家用电器类型及其运行状态,一般用于物联网及智能家居领域,通过构建分类模型来解决。目前,国内外针对非侵入式负荷识别问题已进行了大量研究。文献[5]将高频的波形信息转换成结构化的图谱特征。文献[6]引入格拉姆角场(gramian angular field, GAF)分别对电流和电压信号进行图像编码并结合卷积神经网络完成负荷识别。文献[7]提出了一种基于特征融合与深度学习的负荷识别方法,在V-I轨迹中融入功率特征,提升了多状态负荷辨识能力。文献[8]提出一种香农熵加权投票算法,实现了非侵入式负荷的高精度辨识。文献[9]提出一种基于维特比算法改进的稳态混合非侵入式负荷识别方法,通过构建时序概率模型评价,实现识别结果的校核和修正。文献[10]基于希尔伯特变换对暂态特征进行提取,并结合Seq2Seq-LSTM网络对负荷进行识别。目前相关研究主要集中在辨识算法的选取与结构改进方面,虽引入大量的深度学习算法来捕捉特征信息,但其辨识精度仍普遍不高,主要是由于:①研究重点均放在算法改进上,未考虑数据源本身不平衡分布的现实问题;②已有二维可视化转换方法无法有效解决图像编码过程造成的信息丢失问题;③引入的神经网络结构较为简单,无法提取数据更深层次的语义信息,造成模型的泛化性能和辨识效果受限。

为解决上述问题,本文提出一种基于 ADASYN 和 MTF-DenseNet121 的负荷辨识框架。首先,通过能有效避免时序信息丢失的 MTF 可视化方法,将功率数据转换为具有丰富特征信息的二维纹理图像。接着,将图像数据输入至具有密集连接特征提取机制的 DenseNet121 模型中进行特征捕捉。然后,基于 ADASYN 算法对公共数据集 REDD 进行过采样扩充,以解决模型训练过程中由于数据不平衡带来的学习偏见。最后,结合其他先进图像识别算法,对本文提出的态势感知算法进行有效性验证。本研究旨在为智能电能表负荷辨识技术的推广应用及解决不平衡数据集上负荷辨识准确率偏低的问题提供一种新的解决方案。

## 1 基于 ADASYN 的不平衡数据处理方法

本文采用公开数据集 REDD 进行辨识框架性能验证。REDD 数据集为美国 6 个家庭的实测用电数据,数据包含 1 s 采样的总电源数据和 3 s 采样的单独设备数据,本文采用其中的洗碗机、照明灯、微波炉等 7 种非侵入式负荷设备。

当设备的状态发生改变时,可以通过分析功率

的变化来检测设备状态<sup>[11]</sup>。本文将 REDD 数据集中 6 个家庭的所有设备功率数据依次进行读取,提取设备事件发生前后的 64 个采样数据点。

不平衡数据集是指数据集中不同类别样本数量具有很大的差异。在 REDD 数据集中,各设备的采样数量差异显著,其中照明灯的采样数量最多,达到了 1 866 组,而烤箱的采样数量只有 73 组。深度学习网络在提取设备特征时,从采样数量多的设备中提取的特征要多于采样数量少的设备,造成特征学习偏见问题。已有应用的随机过采样、合成少数类过采样技术(synthetic minority oversampling technique, SMOTE)<sup>[12]</sup>将每一类样本都扩充到相同的数量,然而上述方法存在一定的样本合成盲目性及忽略了数据集的底层分布。ADASYN<sup>[13]</sup>方法能够根据每个样本周围的类别分布来决定生成合成样本的数量。与传统过采样方法相比,ADASYN 方法能够更好地处理那些在决策边界附近并且更难以分类的少数类样本,从而更适用于解决数据集不平衡带来的学习偏见问题。因此,本文采用 ADASYN 方法对不平衡数据集进行处理,其算法流程如下:

假设,  $x_i(i=1,2,\dots,m)$  为数据集中的  $m$  个样本。现用  $m_s$ 、 $m_d$  来分别表示少数类样本的数量和多数类样本的数量。

(1) 多数类与少数类的不平衡度为

$$D = m_s / m_d \quad (1)$$

式中:  $D \in (0, 1]$ 。

(2) 少数类样本需要合成补充的总个数为

$$G = (m_d - m_s)\beta \quad (2)$$

式中:  $\beta \in [0, 1]$ ,  $\beta$  为加入合成样本之后样本的不平衡度,  $\beta = 1$  时表示样本是完全平衡的。

(3) 找出少数类的每个样本  $x_i$  在  $n$  维空间的  $K$  个邻近,其比率为

$$r_i = \Delta_i / K \quad i = 1, 2, \dots, m \quad (3)$$

式中:  $\Delta_i$  为  $x_i$  的  $K$  个邻近中属于多数类的数量,  $r_i \in (0, 1]$ 。

(4) 按照  $\hat{r}_i = r_i / \sum_{i=1}^{m_s} r_i$  处理  $r_i$ , 则  $\sum \hat{r}_i = 1$ 。

(5) 少数类的每个样本  $x_i$  所需要的合成样本数量为

$$g_i = \hat{r}_i G \quad (4)$$

合成少数类的每个样本  $x_i$  所需的  $g_i$  个样本为

$$s_j = x_i + (x_{zi} - x_i) \times \lambda \quad j = 1, 2, \dots, g_i \quad (5)$$

式中:  $x_{zi}$  为从  $x_i$  的  $K$  个邻近中随机选择的一个少数类样本,  $\lambda \in [0, 1]$ 。

## 2 基于MTF和DenseNet的负荷辨识框架

### 2.1 MTF算法原理

从一维NILM信号到二维图像的转换,现有转换方案主要是GAF。然而,GAF在处理含噪声干扰的实测信号数据时容易出现重要时序特征信息的丢失,例如幅值。此外,其计算流程较为繁琐,导致转换效率较低。为解决上述缺陷,本文提出了一种基于MTF的NILM数据可视化方法。

幅值与时间之间的相关性是时序数据中的主要关系,马尔可夫变迁场通过在马尔可夫链的基础上考虑时间位置信息,运用马尔可夫迁移概率进行编码,能够将一维时间序列转化为二维可视化图像,在生成的图像中保持时间顺序和统计动态性<sup>[14]</sup>。

给定一组时间序列  $X=\{x_1, x_2, \dots, x_n\}$ ,  $x_i$  为第  $i$  个采样点。首先定义  $Q$  个分位数单元  $q_j(j \in [1, Q])$ , 对时间序列进行离散化处理,量化其中的每一个值,任意数值点  $x_i$  可以映射到相应的分位数  $q_i$  上。然后,通过沿时间轴以一阶马尔可夫链的方式计算各分位数单元之间的迁移概率,经归一化处理得到  $Q \times Q$  的马尔可夫变迁矩阵  $\mathbf{W}$ , 其表达式如下

$$\mathbf{W} = \begin{bmatrix} w_{11}|P(x_i \in q_1|x_{i-1} \in q_1) & \cdots & w_{1Q}|P(x_i \in q_1|x_{i-1} \in q_Q) \\ w_{21}|P(x_i \in q_2|x_{i-1} \in q_1) & \cdots & w_{2Q}|P(x_i \in q_2|x_{i-1} \in q_Q) \\ \vdots & \ddots & \vdots \\ w_{Q1}|P(x_i \in q_Q|x_{i-1} \in q_1) & \cdots & w_{QQ}|P(x_i \in q_Q|x_{i-1} \in q_Q) \end{bmatrix} \quad (6)$$

式中:  $w_{ij}$  为在分位数  $q_j$  中的数值点被位于分位数  $q_i$  中的数值点逼近的概率,即  $w_{ij} = P(x_i \in q_i|x_{i-1} \in q_j)$ 。

然而,马尔可夫变迁矩阵对  $\mathbf{X}$  的空间分布不够敏感,同时忽略了其与时间步长之间的时间相关性,造成了大量的信息损失。为解决这一问题,利用MTF理论对其进行改进,通过考虑时间位置关系,将包含幅值轴上迁移概率的矩阵  $\mathbf{W}$  扩展至包含时间轴的MTF矩阵  $\mathbf{M}$ , 其表达式为

$$\mathbf{M} = \begin{bmatrix} M_{11} & M_{12} & \cdots & M_{1N} \\ M_{21} & M_{22} & \cdots & M_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ M_{N1} & M_{N2} & \cdots & M_{NN} \end{bmatrix} = \begin{bmatrix} P_{ij}(x_1 \in q_i|x_1 \in q_j) & \cdots & P_{ij}(x_1 \in q_i|x_N \in q_j) \\ P_{ij}(x_2 \in q_i|x_1 \in q_j) & \cdots & P_{ij}(x_2 \in q_i|x_N \in q_j) \\ \vdots & \ddots & \vdots \\ P_{ij}(x_N \in q_i|x_1 \in q_j) & \cdots & P_{ij}(x_N \in q_i|x_N \in q_j) \end{bmatrix} \quad (7)$$

式中:  $M_{mn}$  为从采样信号点  $x_m$  对应的分位数区域  $q_i$  转移至  $x_n$  对应的区域  $q_j$  的概率,即  $q_i$  与  $q_j$  对应分位数关系在矩阵  $\mathbf{W}$  上的转移概率,而对角线上的元

素则是对应的自转移概率。

这样经过MTF转换得到的二维图像具有以下特点与优势:

(1) 一维数据到二维数据为映射关系,图像的像素点位置信息即为功率数据的时序信息,避免了一维功率数据的时序信息丢失。

(2) 像素点幅值信息即为  $M_{ij}$  的值大小,其颜色深浅体现了  $q_i$  到  $q_j$  的转移概率,可视化后的二维图像非常利于深度密集连接网络的学习和识别。

(3) 较GAF而言,MTF整个转换过程十分简洁,计算量较小。分位数划分的引入使MTF对干扰和噪声的抵抗力变强。

### 2.2 密集连接网络

近年来,基于卷积神经网络的图像识别方法<sup>[15]</sup>愈加成熟,其能够直接对二维图像进行处理,很好地进行图像分析与辨识,然而浅层的卷积神经网络无法有效提取功率数据深层的语义信息,进而造成模型泛化性能较差,因此本文采用深层密集连接网络对MTF编码后的图像数据进行特征提取,该网络基于其特殊的密集连接机制,能够保留前层所有的特征信息,因此能够全面挖掘和提取图像数据语义信息,从而提升泛化性能与辨识效果。密集连接网络架构如图1所示。模型方面,应用基于迁移学习的DenseNet121模型对特征图像进行分析。DenseNet121是DenseNet的一个具体实现,特征重用的密集连接方式,使得其在ImageNet数据集上取得了超越深度残差网络等先进图像分类模型的性能表现。通过将预训练权重加载至DenseNet121网络中能够加快模型训练过程中的收敛速度,提升模型的训练质量和效率。

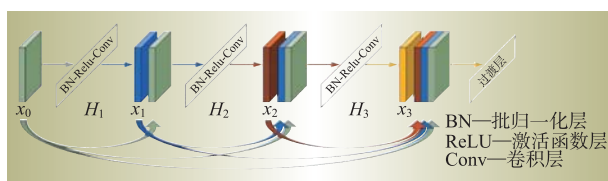


图1 密集连接网络架构

Fig. 1 Structure of DenseNet

在进行分类之前,需对模型进行训练并在验证集中保留最优模型,进而使模型在测试集中的分类准确率达到最高。模型训练步骤如下:

(1) 将采集到的设备功率数据集通过ADASYN方法进行自适应合成扩充处理。

(2) 将处理好的一维功率数据集通过MTF视觉转换方法转化成二维特征图像。

(3) 将设备的类别编号作为负荷标签。

(4) 将图像数据和对应的负荷标签分别输入到构建好的DenseNet121中进行训练,在每一次迭代训练中在验证集中对比性能评价指标,并保留最优训练模型。

(5) 基于最优训练模型,对测试集数据进行负荷辨识性能测试与验证。

### 3 算例分析

#### 3.1 评价指标

为全面评价本文所提方法的负荷识别性能,采用准确率、精确率、召回率和F1得分(分类任务中常用指标,文中用以衡量分类模型的精确率与召回率的综合表现)作为模型识别结果的性能评价指标。计算公式如下

$$A_{accuracy} = \frac{T_p + T_n}{T_p + T_n + E_p + E_n} \quad (8)$$

$$P_{recision} = \frac{T_p}{T_p + F_p} \quad (9)$$

$$R_{ecall} = \frac{T_p}{T_p + F_n} \quad (10)$$

$$F1 = 2 \times \frac{P_{recision} R_{ecall}}{P_{recision} + R_{ecall}} \quad (11)$$

式中:  $T_p$  和  $T_n$  分别为准确识别的属于正样本和负样本的样本数量;  $E_p$  和  $E_n$  分别为正样本和负样本的错误识别数量。

#### 3.2 数据处理结果

经过ADASYN方法对数据集处理后,各类样本的数量如表1所示。烤箱的样本数量从73组扩充至1 866组,火炉的样本数量也增至1 890组,每类设备的样本数量基本保持均衡,使得DenseNet网络对每一类样本的学习度基本保持一致,实现模型辨识精度的进一步提升。

表1 经ADASYN处理前后的数据集中各设备采样数量

Table 1 Number of samples for each device in the dataset before and after ADASYN processing

设备名称	处理前样本个数	处理后样本个数
洗碗机	608	1 864
照明灯	1 866	1 866
烤箱	73	1 866
微波炉	1 107	1 824
冰箱	1 307	1 859
火炉	152	1 890
洗衣机	1 787	1 863

#### 3.3 模型超参数设置

本文模型的各项超参数如表2所示。网络全连

接层的隐含层部分的激活函数都统一设置为ReLU,输出的激活函数设置为Softmax,分类网络中的损失函数采用交叉熵损失函数,优化方法采用SGD优化方法,学习率设置为0.001,训练次数为100次。

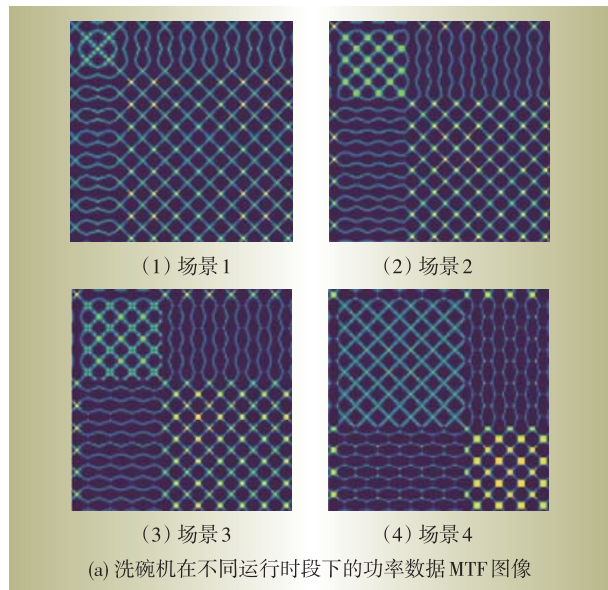
表2 DenseNet121模型超参数设置

Table 2 Hyperparameter settings of DenseNet121

网络层	DenseNet-121(增长率 $k=32$ )
卷积层	7×7 Conv, stride=2
池化层	3×3 最大池化层, stride=2
Dense Block1	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$
过渡层	1×1 卷积层 2×2 平均池化层, stride=2
Dense Block2	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$
过渡层	1×1 卷积层 2×2 平均池化层, stride=2
Dense Block3	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 24$
过渡层	1×1 卷积层 2×2 平均池化层, stride=2
Dense Block4	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 16$
分类层	自适应平均池化层 1 000D 全连接, Softmax

#### 3.4 图像编码结果

图2为洗碗机和照明灯两种非侵入式负荷设备分别在4个不同时段及运行工况下采集到的一维功率数据对应的可视化转换图像。由图3可以看出,在经过MTF对功率序列进行图像编码之后,增加了颜色和纹理特征,使得各类别区分度增加,也为后续深度学习图像识别算法的应用奠定了良好基础。



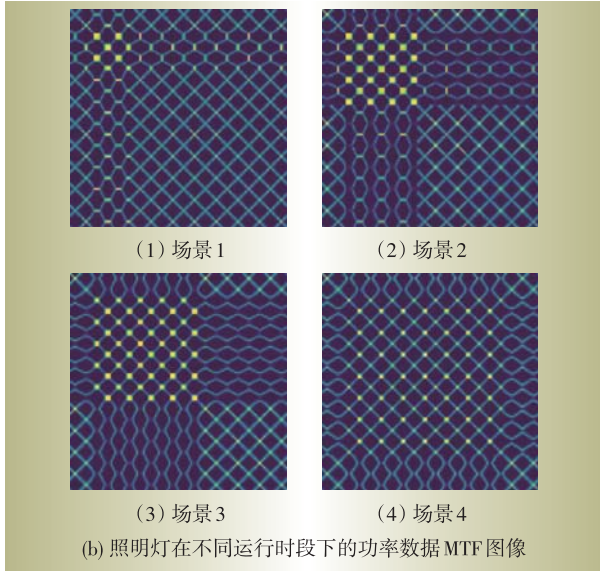


图2 洗碗机和照明灯数据经MTF编码后的图像

Fig. 2 MTF-encoded images of dishwasher and lamp data

3.5 数据集不平衡对结果的影响

使用ADASYN技术对数据集进行处理之后,包含合成数据,数据集中一共有13 032组样本数据。从不包含合成数据的原始数组中,每个类别各随机抽取30组样本数据作为测试集,其余包含合成数据的12 822组样本以9:1划分为训练集和验证集。

对于未进行ADASYN处理的原始数据集,同样的从每个类别的电器样本中随机抽取30组数据作为测试集,其余原始数据仍以9:1划分为训练集和验证集。其中,数据集经过ADASYN技术进行不平衡处理之后,模型的负荷识别准确率达到0.952,而在未经过处理的数据集上的识别准确率仅为0.705。为了对识别结果进行可视化分析,本文使用混淆矩阵来更加直观地展示识别结果。上述两种处理方式的识别结果对应的混淆矩阵如图3所示,模型在各类负荷上的识别性能表现见表3。其中,0代表洗碗机,1代表照明灯,2代表烤箱,3代表微波炉,4代表冰箱,5代表

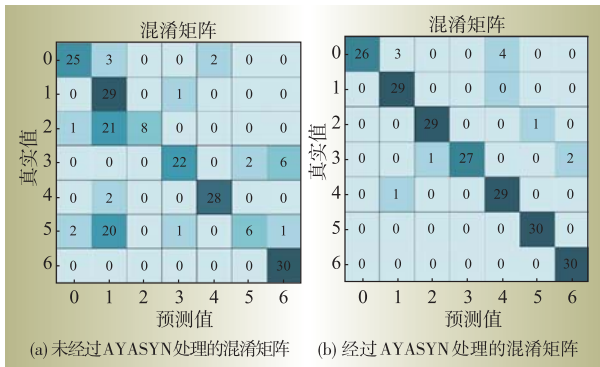


图3 有无数据集不平衡处理的混淆矩阵

Fig. 3 Confusion matrix with or without unbalanced handling of dataset

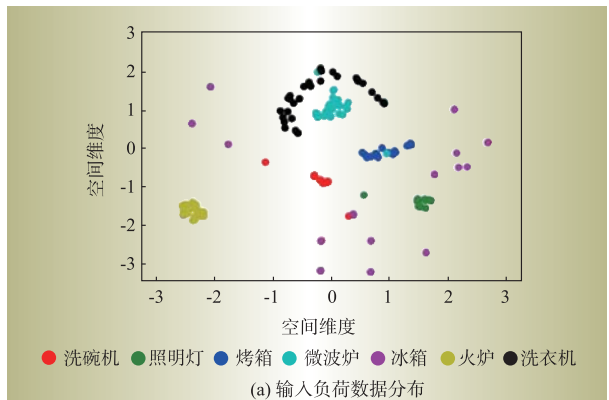
表3 经ADASYN处理前后的各类负荷识别结果  
Table 3 Various load identification results before and after ADASYN processing

负荷类别	未经ADASYN处理			经ADASYN处理		
	精确率	召回率	F1得分	精确率	召回率	F1得分
洗碗机	0.893	0.833	0.862	1.000	0.867	0.929
照明灯	0.387	0.967	0.553	0.967	0.967	0.967
烤箱	1.000	0.267	0.422	0.967	0.967	0.967
微波炉	0.917	0.733	0.815	1.000	0.900	0.947
冰箱	0.933	0.933	0.933	0.853	0.967	0.906
火炉	0.750	0.200	0.316	0.968	1.000	0.983
洗衣机	0.811	1.000	0.896	0.938	1.000	0.968
均值	0.813	0.705	0.685	0.956	0.953	0.952

表火炉,6代表洗衣机。对角线中的数字代表着准确识别的各个相应类别的设备数量。

由图3和表3可以看出,烤箱和火炉的正确识别数量与数据集未经ADASYN处理对比,得到了显著提升,其召回率和F1得分分别提升了0.7/0.545和0.8/0.667,说明基于ADASYN的不平衡数据集处理能够有效提升模型对于少数类样本的识别效果。此外,由表3可知,数据集在经过ADASYN处理后的模型在各类负荷上的精确率、召回率和F1得分基本上均得到了明显的提升,其中3种指标的整体平均值分别提高了0.143/0.248/0.267,且识别准确率提升了0.247,由此验证了基于ADASYN的不平衡数据集处理能够有效提升模型对于非侵入式负荷类型的识别性能。

为进一步验证本文所提算法的有效性,将原始输入负荷数据和模型全连接层的输出特征可视化,以检验所提算法的负荷识别能力。对于空间维度的可视化和降维处理,使用t分布随机近邻嵌入(t-distributed stochastic neighbor embedding, t-SNE)方法,其迭代次数和分辨率分别调整为5 000次和100次。图4显示了7种非侵入式负荷的输入数据和模型全连接层提取特征的可视化效果。可以看出,相对输入负荷数据的不规则空间分布,经DenseNet121



(a) 输入负荷数据分布

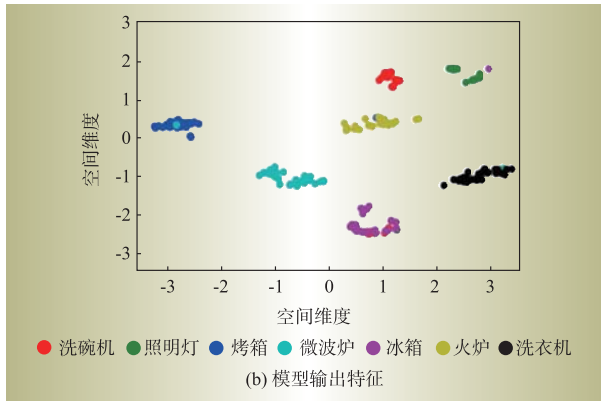


图4 基于t-SNE的输入负荷数据分布和模型输出特征可视化  
Fig. 4 Input load data distribution and model output features visualization based on t-SNE

模型进行特征提取后的相同类型数据基本均被清晰地分为一组,从而进一步验证了所提方法在非侵入式负荷识别方面的有效性与合理性。

### 3.6 与不同先进辨识算法的对比

为验证所提算法的优越性,本文针对经 ADASYN 方法处理后的相同数据集,运用不同的深度卷积神经网络进行对比分析。

选用两种先进图像识别网络 AlexNet 和 GoogleNet 进行对比验证,两者均采用由 MTF 图像编码技术转换成的二维图像数据作为输入,其超参数的设置和深度学习试验环境与本文所用的 DenseNet121 模型一致,最终负荷识别结果的混淆矩阵如图 5 和图 6 所示。不同负荷辨识算法的识别性能表现如表 4 所示。

由图 5 和图 6 可以看出, AlexNet 和 GoogleNet 两种图像辨识算法对于 7 种负荷设备的正确识别数量均低于所提出的 DenseNet121 模型,同时结合表 4 可知,本文算法较两种对比算法的负荷识别准确率分别提升了 0.090 和 0.066,且在各类负荷上的精确率、召回率和 *F1* 得分基本上均得到了全面提升,这是因

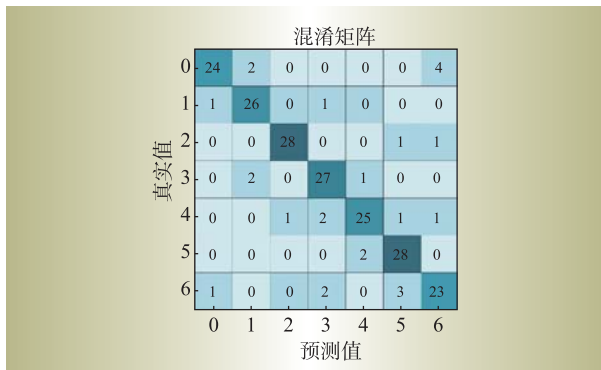


图5 采用 AlexNet 图像识别网络的混淆矩阵  
Fig. 5 Confusion matrix of AlexNet image recognition network

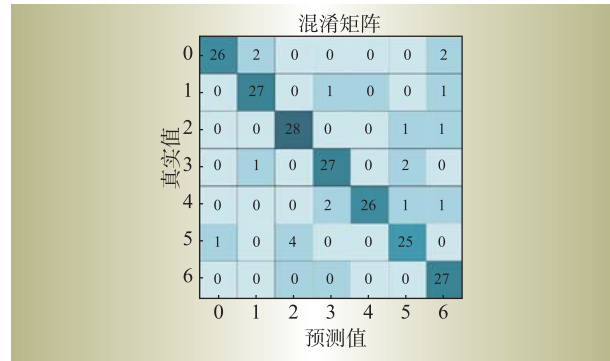


图6 采用 GoogleNet 图像识别网络的混淆矩阵  
Fig. 6 Confusion matrix of GoogleNet image recognition network

为 DenseNet121 网络的密集连接机制能够捕捉更深层次的特征信息,从而有效提高模型的负荷辨识精度性能。综上,本文所提出的辨识算法通过 DenseNet 强大的密集连接机制和信息流改善能力,能够有效减少梯度消失问题,且有利于提取更深层次的图像语义信息,因此对于非侵入式负荷的识别效果优于其他两种算法。

表4 与不同负荷辨识算法的辨识性能对比

Table 4 Comparison with identification performance of different load identification algorithms

负荷类别	AlexNet			GoogleNet		
	精确率	召回率	<i>F1</i> 得分	精确率	召回率	<i>F1</i> 得分
洗碗机	0.923	0.800	0.857	0.963	0.867	0.912
照明灯	0.867	0.867	0.867	0.900	0.900	0.900
烤箱	0.903	0.933	0.918	0.848	0.933	0.888
微波炉	0.844	0.900	0.871	0.844	0.900	0.871
冰箱	0.862	0.833	0.847	0.963	0.867	0.912
火炉	0.848	0.933	0.888	0.862	0.833	0.847
洗衣机	0.793	0.767	0.780	0.844	0.900	0.871
均值	0.863	0.862	0.861	0.889	0.886	0.886

## 4 结束语

为解决传统非侵入式负荷辨识算法在不平衡数据集上辨识精度偏低的问题,本文提出了一种基于 ADASYN 和图像分析的非侵入式负荷辨识算法。该算法能有效解决数据不平衡带来的模型学习偏见问题,同时结合能够保留一维时序信息的 MTF 和具有强大特征捕捉能力的 DenseNet121 网络,实现对非侵入式负荷类型的高精度辨识。首先,通过 MTF 图像编码将一维功率数据转化为特征明晰的二维图像。然后,将图像数据及对应标签输入到 DenseNet121 中进行训练学习,提取图像深层特征信息进行负荷辨识,解决了一般 CNN 捕捉特征信息能力不足的问题。之后,基于 ADASYN 算法对

公开数据集 REDD 进行过采样处理,解决了由于数据不平衡带来的学习偏见问题。最后,基于处理前后的 REDD 数据集,验证了 ADASYN 算法对于解决数据不平衡问题的有效性,同时结合两种先进图像分类网络 AlexNet 和 GoogleNet 进行性能对比,验证了所提 MTF-DenseNet121 算法在非侵入式负荷辨识方面的有效性与优越性。[D]

### 参考文献:

- [1] 王恺,赵文会,张伟时,等. 双碳背景下微电网的优化配置研究[J]. 电力需求侧管理,2023,25(4):86-92.  
WANG Kai, ZHAO Wenhui, ZHANG Weishi, et al. Optimal al location of microgrid under the background of carbon peaking and neutralization [J]. Power Demand Side Management, 2023, 25(4):86-92.
- [2] 姚志华,余波,袁法培,等. 考虑碳排放目标的配电网源网荷储优化调度探究[J]. 电力需求侧管理,2023,25(3):80-86.  
YAO Zhihua, YU Bo, YUAN Fapei, et al. Optimal scheduling of distribution network side source network load storage considering dual-carbon targets in distribution market [J]. Power Demand Side Management, 2023, 25(3):80-86.
- [3] 汤奕,邓克愚,孙华东,等. 智能家电参与低频减载协调配合方案研究[J]. 电网技术,2013,37(10):2861-2867.  
TANG Yi, DENG Keyu, SUN Huadong, et al. Research on coordination scheme for smart household appliances participating under frequency load shedding [J]. Power System Technology, 2013, 37(10):2861-2867.
- [4] 王昇成,罗华峰,马骏超,等. 一种小微负荷智能监测与聚合分析系统研究[J]. 浙江电力,2021,40(12):37-45.  
WANG Yicheng, LUO Huafeng, MA Junchao, et al. Research on smart monitoring and aggregate analysis system for small and micro loads [J]. Zhejiang Electric Power, 2021, 40(12):37-45.
- [5] 武昕,严萌,郭一凡,等. 基于结构化特征图谱的组合支持向量机非侵入式负荷辨识[J]. 电力系统自动化,2022,46(12):210-219.  
WU Xin, YAN Meng, GUO Yifan, et al. Non-intrusive load identification by combined support vector machine based on structured characteristic spectrum [J]. Automation of Electric Power Systems, 2022, 46(12):210-219.
- [6] 崔昊杨,吴轶凡,江友华,等. 非侵入式负荷识别的电流序列可视化方法[J]. 电力自动化设备,2022,42(7):40-45.  
CUI Haoyang, WU Yifan, JIANG Youhua, et al. Current sequence visualization method of non-intrusive load recognition [J]. Electric Power Automation Equipment, 2022, 42(7):40-45.
- [7] 王守相,郭陆阳,陈海文,等. 基于特征融合与深度学习
- 习的非侵入式负荷辨识算法[J]. 电力系统自动化,2020,44(9):103-110.  
WANG Shouxiang, GUO Luyang, CHEN Haiwen, et al. Non-intrusive load identification algorithm based on feature fusion and deep learning [J]. Automation of Electric Power Systems, 2020, 44(9):103-110.
- [8] 魏广芬,赵航,胡春华,等. 基于香农熵加权投票算法的集成式非侵入式负荷识别方法[J]. 中国电机工程学报,2022,42(24):8876-8888.  
WEI Guangfen, ZHAO Hang, HU Chunhua, et al. An integrated non-intrusive load identification method based on shannon entropy weighted voting algorithm [J]. Proceedings of the CSEE, 2022, 42(24):8876-8888.
- [9] 刘宇,刘丛笑,赵欣,等. 基于维特比算法改进的稳态混合非侵入式负荷识别方法[J]. 电工技术学报,2023,38(19):5241-5255.  
LIU Yu, LIU Conglin, ZHAO Xin, et al. An improved steady-and transient-state mixed non-intrusive load monitoring using viterbi algorithm [J]. Transactions of China Electrotechnical Society, 2023, 38(19):5241-5255.
- [10] LE T, HEO S, KIM H. Toward load identification based on the Hilbert transform and sequence to sequence long short-term memory [J]. IEEE Transactions on Smart Grid, 2021, 12(4):3252-3264.
- [11] LIU Y, WANG X, YOU W. Non-intrusive load monitoring by voltage-current trajectory enabled transfer learning [J]. IEEE Transactions on Smart Grid, 2019, 10(5):5609-5619.
- [12] CHAWLA N, BOWYER K, HALL L, et al. SMOTE: synthetic minority over-sampling technique [J]. Journal of artificial intelligence research, 2002, 16:321-357.
- [13] HE H, BAI Y, GARCIA E A, et al. ADASYN: adaptive synthetic sampling approach for imbalanced learning [C]//Proceedings of International Joint Conference on Neural Networks, New York:IEEE, 2008:1322-1328.
- [14] WANG Z, OATES T. Spatially encoding temporal correlations to classify temporal data using convolutional neural networks [J]. Journal of Computer and Systems Sciences, 2015, arXiv:1509.07481.
- [15] XIE G, ZHANG X, YANG W, et al. LG-CNN: from local parts to global discrimination for fine-grained recognition [J]. Pattern Recognition, 2017, 71:118-131.

### 作者简介:

顾水福(1966),男,江苏苏州人,研究员级高级工程师,研究方向为电力系统自动化;

周磊(1989),男,山西临汾人,博士,高级工程师,研究方向为电力需求侧响应及其在电力系统中的应用;

李洁(1972),女,江苏苏州人,高级工程师,研究方向为电力负荷管理及需求响应。

(责任编辑 水 鹤)