

DOI: 10.19783/j.cnki.pspc.220048

基于异构基 Stacking 机制的居民负荷特征图像识别方法

邱磊鑫, 余涛, 彭秉刚

(华南理工大学电力学院, 广东 广州 510640)

摘要: 负荷识别技术能将不同电器类型有效区分开, 对于用电策略制定、需求响应具有重要意义。针对当前负荷识别技术无法有效实现负荷特征融合以及不同识别器模型结合的问题, 提出一种基于异构基 Stacking 机制的居民用电负荷识别特征图像集成学习方法。该方法通过构建特征图像实现特征融合, 利用卷积神经网络充分挖掘特征图像中蕴含的深层次特征, 解决传统方法对特征挖掘不够深入的问题。同时引入集成学习 Stacking 方法将多种异质负荷识别模型结合, 综合各种模型的优势, 解决传统方法模型单一化的问题。最后使用公开数据集 PLAID 进行验证并在实验室电器设备上完成工程应用。结果表明, 所提方法具有较高的识别准确率和应用价值。

关键词: 负荷识别; 集成学习; 特征图像; 卷积神经网络; 特征筛选

Image recognition method of resident load characteristics based on heterogeneous basis Stacking mechanism

QIU Leixin, YU Tao, PENG Binggang

(School of Electric Power Engineering, South China University of Technology, Guangzhou 510640, China)

Abstract: Load identification technology can effectively distinguish different electrical types. This is of significance for power consumption strategy formulation and demand response. There is a problem in that the present load identification technology cannot effectively realize the load feature fusion and the combination of different recognizer models. Thus an integrated learning method of residential power load identification feature image based on heterogeneous basis Stacking mechanism is proposed. This method realizes feature fusion by constructing feature images and fully excavates the deep-seated features contained in feature images using a convolutional neural network, so as to solve the problem that the traditional methods are not deep enough in feature mining. At the same time, the ensemble learning Stacking method is introduced to combine a variety of heterogeneous load identification models to integrate the advantages of various models and overcome the problem of the simplification of traditional methods. Finally, the public data set PLAID is used to verify and complete the engineering application on laboratory electrical equipment. The results show that this method has high recognition accuracy and application value.

This work is supported by the Joint Project of the Commission of National Natural Science Foundation of China-Smart Grid of State Grid Corporation of China (No. U2066212).

Key words: load identification; integrated learning; feature image; convolutional neural network; feature selection

0 引言

在智能电网的建设中, 负荷识别^[1-2]技术能在电网与用户之间搭建起良好的信息互动机制。负荷识别技术能应用于即插即用设备, 有效帮助用户了解自己内部各种电器的实时使用情况和能耗对比; 同时负荷识别技术能够实现对用户用电的分类计量,

为需求侧响应技术提供必要的技术支持。

特征选择与识别算法的设计是负荷识别的两项关键技术。在特征选择上, 将多种负荷特征进行结合是常见的做法, 文献[3]便将负荷开关时的各种暂态功率信息组合作为指标进行识别; 文献[4]利用稳态电流谐波特性作为辨识特征; 文献[5]将电压和电流波形数据以 V-I 轨迹的形式组合, 并采用图像识别的方法解决, 该方法将特征组合转化为图像来挖掘特征, 但 V-I 轨迹没有将电流幅值大小考虑进来, 难以区分波形相似的电器; 文献[6]对 V-I 轨迹进行

基金项目: 国家自然科学基金委员会-国家电网公司智能电网联合基金项目资助(U2066212)

改进, 构建出电压-电流-功率三维轨迹特征用于识别。以上方法对于特征的选择与构建还停留在特征简单组合的层面, 没有进一步考虑特征的融合。

在识别算法上, 目前传统机器学习方法的应用已经较为成熟, 如聚类、决策树、支持向量机等。文献[7]采用 K 最近邻算法和人工神经网络两种方法对负荷的通断电暂态特性进行识别; 文献[8]对传统 K 均值聚类方法进行改进, 提高了聚类质量与识别精度; 文献[9]采用支持向量机方法对二值 V-I 轨迹进行识别。近年来, 随着深度学习和集成学习方法的发展, 卷积神经网络^[6,10]、递归神经网络^[11-12]、图卷积神经网络^[8]等方法也被用于负荷识别; 文献[13]运用集成学习的 bagging 思想, 相比于其他机器学习方法取得了较佳的效果; 文献[14]提出了一种平衡随机森林模型, 实现对类不平衡负荷数据集的识别, 但在空调、洗衣机等负荷上的识别精度还有待提高。以上方法虽然都取得了一定效果, 但单独某一种模型难以在所有电器上都取得较高的识别精度, 同时模型的泛化性能也不尽如人意。

基于对上述研究的分析可以发现, 目前研究过程中存在以下问题。

1) 单纯的特征堆叠会造成特征冗余, 难以挖掘出特征组合中蕴含的深层次信息, 阻碍负荷识别精度的提升;

2) 对单一识别算法进行优化改进从而提高分类精度的做法陷入了瓶颈, 引入 bagging 机制融合同质学习器的方法可以提升总体效果, 但对于部分电器存在识别精度不佳的问题, 可以考虑将多种异质分类器融合, 结合异质分类器在不同方面的优势来改进模型。

因此, 本文提出一种基于异构基 Stacking 机制的居民用电负荷识别特征图像集成学习方法, 该方法首先利用 mRMR 方法筛选出典型负荷特征, 根据负荷特征构建负荷特征向量和负荷图像, 其次引入集成学习 Stacking 机制将残差卷积神经网络与传统机器学习方法相融合构建负荷识别模型, 最后使用 PLAID 公开数据集和实验室自建数据集对模型进行训练和测试, 验证了本文方法的有效性以及工程价值。

1 负荷特征构建

1.1 特征筛选

负荷特征的提取和选择是负荷识别的基础工作。通过对文献[15]进行分析, 选取了工程特征中的有功功率、无功功率、电流有效值、1~31 次谐波电流幅值和相角等运行特征进行计算, 以提取负荷

稳定运行过程中的特征信息。

虽然上述特征都具有一定的代表性, 但使用过多特征会增加问题的复杂程度。首先, 某些特征的加入对于负荷识别起到了阻碍作用; 此外, 具有相关关系的特征之间, 可能存在信息重复冗余的情况, 反而会使得识别精度下降。

常用的特征筛选方法包括相关系数法、互信息法等。这些方法的基本原理都是计算特征变量与目标值之间的相关程度, 从而筛选出最具代表性的特征, 但这些方法没有考虑特征变量之间存在的冗余现象, 因此本文采用文献[16]提出的 mRMR 方法筛选出与负荷类型结果相关度最高的一组特征。

mRMR 方法建立在互信息法的基础上, 基本原理是筛选出一组特征子集使得彼此的互信息最小而其与目标值的互信息最大, 这样就可以在确保识别精度的前提下减少信息冗余^[17-18]。

互信息被定义为两个随机变量之间的概率密度函数, 如式(1)所示。

$$I(x; y) = \iint p(x, y) \log \frac{p(x, y)}{p(x)p(y)} dx dy \quad (1)$$

式中: x 和 y 为两个随机变量; $p(x)$ 和 $p(y)$ 为概率密度函数。

本文应用 mRMR 方法进行特征筛选的具体步骤如下。

1) 确定所需的特征种类数 m , 并依次从总特征集中划分出所有可能的特征子集。

2) 计算特征子集 S 与标签 c 之间的相关度。

$$\max D(S, c), D = \frac{1}{|S|} \sum_{x_i \in S} I(x_i; c) \quad (2)$$

式中: c 为标签, 即负荷类型; $|S|$ 为特征子集 S 中的特征个数。

3) 计算每个特征子集 S 的冗余度。

$$\min R(S), R = \frac{1}{|S|^2} \sum_{x_i, x_j \in S} I(x_i, x_j) \quad (3)$$

式中, x_i 和 x_j 为属于特征子集 S 的任意两组数据。

4) 计算每个特征子集的 mRMR 值, 并筛选出 mRMR 值最大的特征子集作为目标特征子集。

$$\max \Phi(D, R), \Phi = D - R \quad (4)$$

式中: D 为步骤 2) 计算的相关度结果; R 为步骤 3) 计算的冗余度结果。

1.2 负荷特征向量及特征图像的构建

对于 1.1 节筛选出的 m 种负荷特征, 可以考虑采用以下两种方法对采集得到的负荷数据进行处理。

一种方法是计算各个电器一个周波内的 m 种稳态负荷特征, 并以此构建形状为 $(1, m)$ 的负荷特征向

量, 对此可以考虑采用传统机器学习方法如决策树进行识别。

另一种方法是将负荷特征向量理解为一个反映该电器用电行为特性的 RGB 图像, 同时考虑某些电器所具有的非周期性或长周期性的工作特性, 将多个周期内负荷运行特征的统计性数据如平均值、方差、峰峰值等加入原有的单周波数据中, 从而构建负荷特征图像。

负荷特征图像的具体构建方式如下所述。

对于每个电器, 采集其稳态运行时一个周波内的 m 个特征, 记作当前值, 同时计算各个特征在 N 个 (N 通常取 10) 周波内的最大值、平均值, 从而获得形状为 $(1, m, 3)$ 的三维矩阵。为了构建图像, 采用最大最小归一化方法对矩阵进行归一化, 然后对其进行缩放, 使其数值位于 $[0, 255]$ 内, 如式(5)所示, 最后将当前值、最大值、平均值各作为图像的一个通道, 以此构建长为 m 、宽为 1 的 RGB 负荷特征图像。

$$x^* = 255 \times \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (5)$$

式中: x 表示某一电器的某个特征未归一化和缩放前的数值; x_{\max} 表示所有采集的数据中该特征的最大值; x_{\min} 表示所有采集的数据中该特征的最小值; x^* 表示该特征进行归一化和缩放后的数值。

对于负荷特征图像可以引入 CNN、RNN 等神经网络方法提取电器独特的运行特性, 以实现负荷识别。

图 1 展示了几种典型居民用电负荷稳态运行时的负荷特征图像。可以发现, 运行特性差异较大的电器的特征图像有比较明显的区别, 例如微波炉和笔记本电脑亮度条的数量、位置和亮度都有所不同。相应地, 具有相近特性的电器的特征图像具有相似的特征, 例如吹风机和水壶从原理上都属于热阻类电器, 它们的亮度条具有比较一致的特点。上述分析直观地展示出特征图像的构建对于挖掘负荷特征的效果。

2 负荷识别模型

2.1 传统机器学习方法

传统机器学习从数据自身角度出发, 能够有效提取数据的特征, 发现同类数据中所具有的统计规律, 适用于解决负荷识别问题, 常用的机器学习方法包括支持向量机(support vector machine, SVM)、K 最近邻算法(K-nearest neighbour, KNN)、随机森林(random forest, RF)和极限树(extra trees, ET)等。

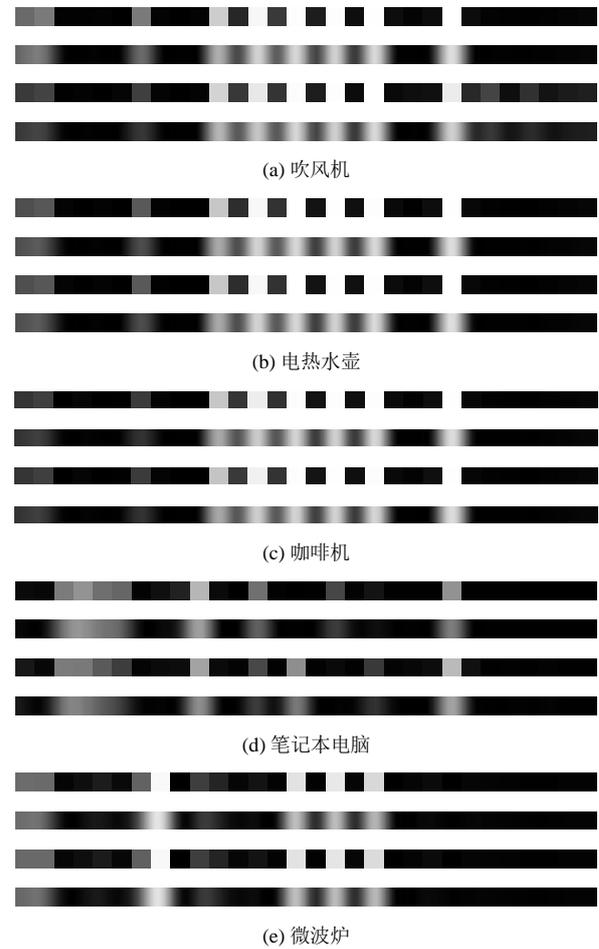


图 1 典型居民用电负荷特征图像

Fig. 1 Characteristic image of typical residential load

SVM 本质上是一种线性分类模型, 通过在特征空间构建超平面将不同类样本尽可能划分开, 进而实现分类。

KNN 算法的思想是找到待分类数据点在特征空间中最相邻的 k 个数据点, 然后根据多数原则, 将待分类点归入到 k 个点中占比最大的类别。

随机森林方法是对决策树(decision tree, DT)的拓展和优化, 通过随机选取特征、样本、参数和模型的方法构建多棵决策树, 将多棵树的分类结果整合在一起, 提升了树的分类性能。该方法广泛应用于电能质量研究^[19]、负荷预测^[20]和输电线路故障识别^[21]等领域。

极限树的方法与随机森林类似, 区别是极限树使用了所有样本进行训练, 同时在树的分支上的选择是随机的。

2.2 Att-ResNet 卷积神经网络

在负荷识别任务中, 从电器的原始电压电流波形中提取此类电器独有的运行特性信息是实现负荷

识别的关键。卷积神经网络(convolutional neural network, CNN)可以通过构造不同的卷积核实现对负荷特征图像中深层特性信息的提取。

卷积神经网络的性能与网络的深度息息相关,有关研究发现,随着网络层数的增加,梯度消失和梯度爆炸的现象将会阻碍模型的收敛。为了解决这一问题,本文引入文献[22]提出的残差机制和文献[23]提出的卷积块注意力机制。

残差机制引入了快捷连接的思想,原始输入通过快捷连接跳过多个卷积层,直接与残差部分叠加到输出中,这种操作使得残差卷积网络可以直接从输入中学习到原始特征,从而避免梯度消失或爆炸的问题。

卷积块注意力机制(convolutional block attention module, CBAM)通过对通道注意力子模块和空间注意力子模块的应用,关注数据在通道和空间这两个维度上对分类有意义的典型特征。

通道注意力子模块的原理:首先对于一个形状为(H, W, C)的特征图像,分别采用 Maxpool 和

AvgPool 压缩成两个形状为($1, 1, C$)的图像,然后两个图像经过神经元个数分别为 C/r 和 C 的两层全连接层,最后得到的结果直接相加并选取恰当的激活函数,完成通道注意力过程。

空间注意力子模块的原理:首先对于一个形状为(H, W, C)的特征图像,分别通过 Maxpool 和 AvgPool 压缩成两个形状为($H, W, 1$)的图像,然后这两个图像按照最后一个维度拼接起来,形成形状为($H, W, 2$)的图像,最后图像通过一个卷积层(卷积核大小一般为 7×7)和激活函数,完成空间注意力过程。

结合残差机制和卷积块注意力机制的卷积神经网络 Att-ResNet,可以充分挖掘负荷特征图像中的深层次特征信息,提高负荷识别的精度。本文用于负荷识别的 Att-ResNet 模型结构如图 2 所示,负荷特征图像先通过一个卷积层输入模型,然后通过两个残差块堆叠提取特征,CBAM 插入残差块中的两个卷积层之间,最后通过两个全连接层实现负荷图像到识别负荷类型的映射。

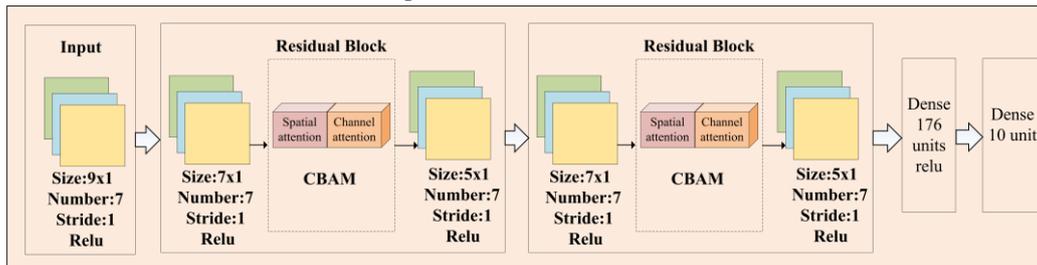


图 2 Att-ResNet 模型结构图

Fig. 2 Structure of Att-ResNet model

2.3 异构基 Stacking 机制

为了实现传统机器学习方法与卷积神经网络的模型融合,本文引入 Stacking 策略,将多种异质负荷识别模型作为基学习器,随机森林作为次级学习器对基学习器的分类结果再次进行训练,以获得更优的分类结果。

Stacking 是集成学习思想的其中一种体现,它让多个异质学习器并行进行学习,最后再通过投票法或是再训练一个元模型将多个基学习器的分类结果组合起来,以期获得一个比单个基学习器更强的模型^[24]。Stacking 方法已经在配电网网损预测^[25]和停电空间预测^[26]等领域有所应用,但在负荷识别上的研究还不常见。Stacking 能够成功的关键在于多个基学习器能针对原始训练数据得出有差异性且预测能力好的输出值,这样通过次级学习器进一步学习后,能够在多个基学习器中取长补短,提升预测的准确度和稳定性,因此,基学习器一般会选择拟合

度高的模型,而次级学习器则一般选择简单模型。

3 基于异构基 Stacking 机制的负荷识别模型设计

3.1 基于异构基 Stacking 机制的负荷识别模型构建

本文提出的基于异构基 Stacking 机制的负荷识别模型,如图 3 所示。

将 Att-ResNet、RF、SVM、KNN 和 ET 作为基学习器,DT 作为次级学习器,以此构建负荷识别模型。其中,特征数据以负荷特征图像输入 Att-ResNet,以负荷特征向量输入其他基学习器。

3.2 Stacking 框架效果比较

以 Att-ResNet 作为基础,依次将 SVM、KNN、ET 和 RF 方法添加进 Stacking 框架中,模型的负荷识别精度变化情况如图 4 所示。

可以看出,融合后的模型识别精度相较于融合

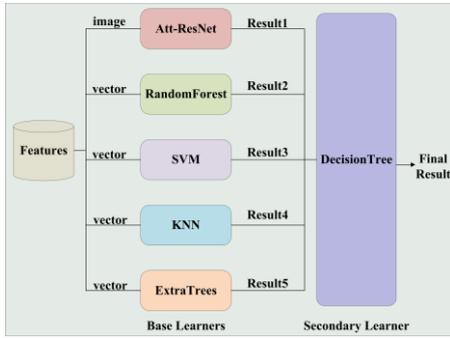


图 3 基于异构基 Stacking 机制的负荷识别模型结构
Fig. 3 Structure of load identification model based on heterogeneous basis Stacking mechanism

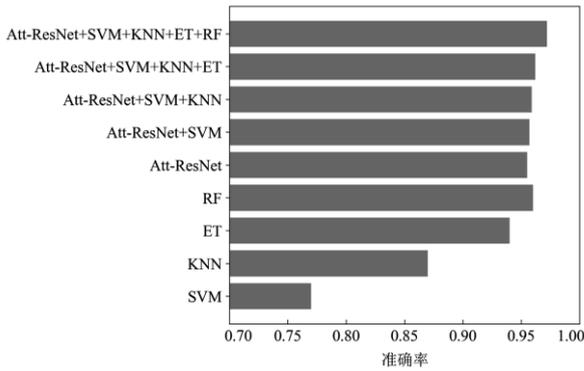


图 4 基学习器对 Stacking 模型识别精度的影响
Fig. 4 Influence of base learners on the recognition accuracy of Stacking model

前的单一模型均有所提升；同时，随着 Stacking 框架中融合方法的增加，识别精度也在提高。

3.3 训练方法

由于负荷识别是典型的多分类问题，采用交叉熵作为模型训练过程中的损失函数，损失函数如式(6)所示。同时对于 Att-ResNet 模型采用 Early-Stopping 机制防止数据过拟合。

$$L_{\text{loss}} = -\sum_{i=1}^N \sum_{j=1}^K y_{i,j} \log(p_{i,j}) \quad (6)$$

式中： $p_{i,j}$ 表示样本 i 属于第 j 类的概率； $y_{i,j}$ 表示第 i 个样本是否属于第 j 类，属于则 $y_{i,j} = 1$ ，否则 $y_{i,j} = 0$ ； N 为样本数； K 为样本标签种类数。

同时，为了减小过拟合的风险，Stacking 一般使用 K -折交叉训练方法进行训练，如图 5 所示。

Stacking 的具体训练过程如下。

1) 将训练集划分为 K 折；

2) 针对每个基学习器，将每份数据作为验证集，其余 $K - 1$ 份作为训练集，从而对模型进行 K 次训练，将 K 个验证集的验证结果拼接在一起作为

该基学习器对训练集的分类结果；

3) 上一步得到 N 个基学习器对于训练集的 N 组分类结果，将其作为训练样本输入次级学习器，进行训练再预测，得到最终分类结果。

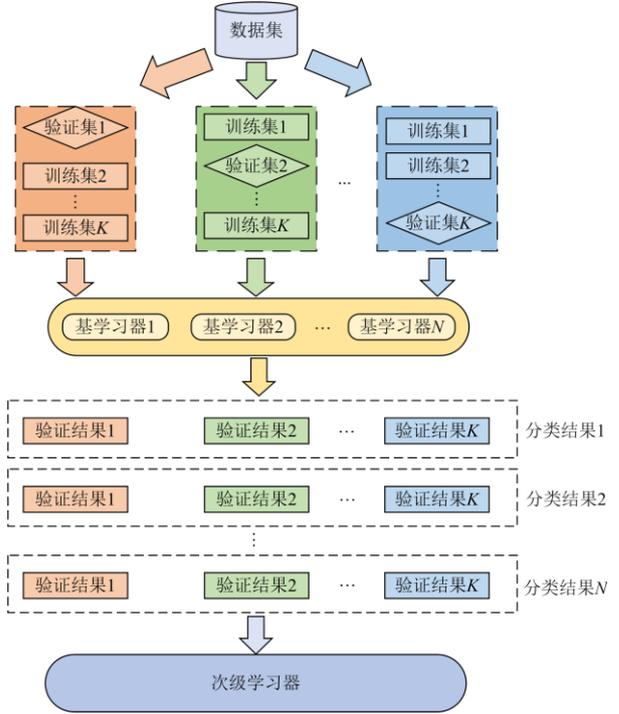


图 5 Stacking 机制训练过程
Fig. 5 Training process of Stacking mechanism

4 实验及分析

4.1 实验步骤

本文提出的负荷识别方法具体包括以下步骤。

步骤 1：对于待识别负荷的原始电压和电流波形数据进行处理。计算负荷运行过程中每个周期的有功功率、无功功率、视在功率、功率因数和电流的平均值、峰峰值、有效值、峰均比、波形畸变率等特征，同时采用 FFT 算法计算负荷 1~31 次电流谐波幅值和相位、1~31 次谐波阻抗幅值和相位。

步骤 2：使用 mRMR 方法筛选出典型特征。使用 mRMR 算法对步骤 1 计算得到的特征进行筛选，得到其中得分最高的 m 个特征，以此作为负荷识别的典型特征。本文提出的模型在不同特征数目下的性能如图 6 所示。

从图 6 可以看出，使用 30 个特征的模型宏观 F1 值可达 0.9 以上，这表示通过 mRMR 方法筛选出的排名靠前的 30 个特征已经包含数据集的大部分信息。最终筛选得到的 30 个特征如表 1 所示。

步骤 3：取步骤 3 获得的典型特征构建负荷特征向量和特征图像。

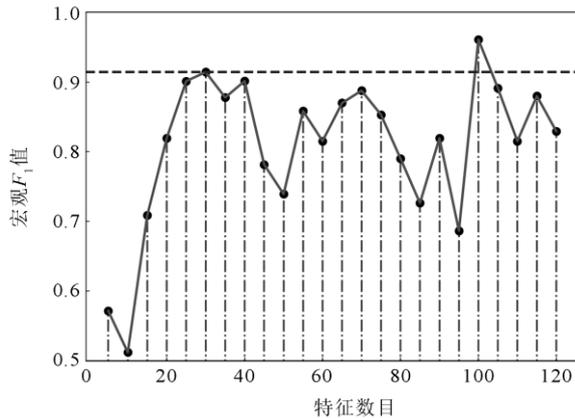


图 6 不同特征数目下模型的性能

Fig. 6 Performance of the model with different number of features

表 1 mRMR 方法筛选出的特征

Table 1 Characteristics screened by mRMR method

电流有效值	电流平均值	电流峰均比	电流峰峰值	功率因数
有功功率	无功功率	波形因数	电流谐波畸变率	电流整流平均值
电流基波幅值	电流基波相位	基波阻抗幅值	电流 3 次谐波幅值	电流 3 次谐波相位
3 次谐波阻抗幅值	3 次谐波阻抗相位	5 次谐波阻抗幅值	5 次谐波阻抗相位	7 次谐波阻抗幅值
7 次谐波阻抗相位	7 次谐波阻抗幅值	9 次谐波阻抗幅值	9 次谐波阻抗相位	11 次谐波阻抗幅值
13 次谐波阻抗幅值	15 次谐波阻抗幅值	15 次谐波阻抗相位	17 次谐波阻抗幅值	19 次谐波阻抗幅值

步骤 4: 将负荷特征向量和特征图像分别送入传统机器学习模型和 Att-ResNet 模型进行训练。

步骤 5: 基于 Stacking 思想, 将步骤 4 的基分类器模型分类结果代入次级学习器, 进行训练再预测, 完成最终模型的训练。

步骤 6: 使用测试数据对训练后模型的准确度

和泛化能力进行测试。常用的评价指标包括 Recall、Precision、F₁-score 等, 具体公式为

$$R_{\text{recall}} = \frac{T_p}{T_p + F_N} \quad (7)$$

$$P_{\text{recision}} = \frac{T_p}{T_p + F_p} \quad (8)$$

$$F_1 = 2 \times \frac{P_{\text{recision}} \times R_{\text{recall}}}{P_{\text{recision}} + R_{\text{recall}}} \quad (9)$$

式中: T_p 为真正值; T_N 为真负值; F_p 为假正值; F_N 为假负值。

4.2 公开数据集实验

采用 PLAID(plug load appliance identification dataset)^[27] 公开数据集的电气数据对本文提出的负荷方法进行验证, PLAID 是由 Jingkun Gao 等人建立的电器运行数据集, 以 30 kHz 的采样频率采集了美国 55 个家庭中 10 余种家用电器的电压和电流数据, 其中每种电器都对应多个不同的品牌和型号。

本文选取其中 10 种电器, 包括空调、荧光灯、风扇、冰箱、吹风机、加热器、白炽灯、笔记本电脑、微波炉和吸尘器作为研究对象。采用 3:1:1 的比例划分训练集、验证集和测试集。为了更好地测试模型的泛化能力, 在数据集的划分时不会将每一电器的所有品类都加入训练集。同时由于 PLAID 数据集中不同种类的电器数目不平衡, 本文采用 SMOTE 方法^[28]对训练集中的少数类样本进行过采样, 从而使各类样本达到平衡。

为了验证本文所提模型的优越性能, 以文献[7]提出的 KNN 方法、文献[14]提出的随机森林方法和本文提出的方法进行比较, 使用同样的训练集、验证集和测试集进行实验, 不同方法的各项性能指标如表 2 所示。由实验结果可知:

表 2 PLAID 数据集各类指标比较

Table 2 Comparison of various indicators of different methods in the PLAID dataset

电器	Recall				Precision				F1			
	KNN ^[7]	RF ^[14]	Att-ResNet	Stacking	KNN ^[7]	RF ^[14]	Att-ResNet	Stacking	KNN ^[7]	RF ^[14]	Att-ResNet	Stacking
空调	0.81	0.90	0.91	0.94	0.72	0.96	0.95	0.96	0.76	0.93	0.93	0.95
荧光灯	0.99	0.99	1.00	1.00	0.90	1.00	0.97	0.98	0.94	1.00	0.99	0.99
风扇	0.79	0.98	0.94	0.98	0.76	0.90	0.90	0.94	0.77	0.94	0.92	0.96
冰箱	0.54	0.89	0.89	0.97	0.67	0.95	0.87	0.95	0.60	0.92	0.88	0.96
吹风机	0.85	0.97	0.94	0.95	0.89	0.93	0.96	0.96	0.87	0.95	0.95	0.95
加热器	0.78	0.81	0.88	0.88	0.80	0.97	0.96	0.96	0.79	0.88	0.92	0.92
白炽灯	0.95	0.99	1.00	1.00	0.93	0.99	1.00	1.00	0.94	0.99	1.00	1.00
笔记本电脑	0.82	1.00	0.93	0.98	0.95	0.97	0.95	0.97	0.88	0.98	0.94	0.98
微波炉	0.99	0.99	0.99	0.99	0.98	0.99	1.00	0.99	0.98	0.99	1.00	0.99
吸尘器	0.98	1.00	1.00	1.00	0.82	1.00	0.92	1.00	0.89	1.00	0.96	1.00
洗衣机	0.75	0.91	0.92	0.91	0.92	0.94	0.88	0.96	0.82	0.93	0.90	0.93
总体	0.87	0.96	0.95	0.97	0.87	0.96	0.95	0.97	0.87	0.96	0.95	0.97

1) 在 KNN、随机森林、Att-ResNet 这 3 种单一模型中, KNN 效果最差, 冰箱的识别准确率仅有 54%; 而随机森林和 Att-ResNet 的效果比较接近, 多种电器的识别准确率达到 90% 以上, 总体准确率达到 95% 以上, 其中随机森林方法对于风扇、吹风机和笔记本电脑等电器的识别优于 Att-ResNet, 而 Att-ResNet 则对于加热器、洗衣机的识别略胜一筹。

2) 在应用了基于异构基 Stacking 机制的负荷识别方法后, 大部分电器的识别准确率都有所提高, 比如冰箱的识别准确率相较于随机森林和 Att-ResNet 提高了 8% 左右, 总体准确率也高于单一模型, 达到了 97%, 充分说明了本文方法的有效性。

4.3 在实际设备上的工程应用

此外, 由于 PLAID 数据集测量的是国外的电器数据, 电气环境与国内有所区别, 同时数据集中选取的某些电器在国内并不常用, 为了更好地验证模型的泛化性与实用性, 在实验室中搭建负荷识别平台, 采用高精度采样仪采集了一批常用家用电器的电压电流波形数据, 并建立数据集对本文所提模型进行测试, 实验室平台如图 7 所示。

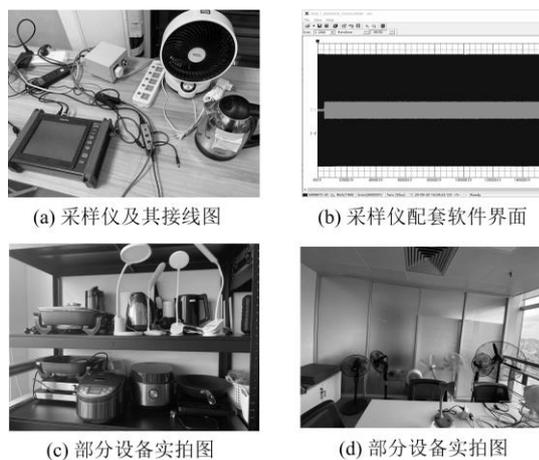


图 7 实验室平台展示图

Fig. 7 Display diagram of laboratory platform

实验波形数据采集仪器是 HIOKI 公司型号为 MR8875-30 的采样仪, 采样频率为 20 kHz, 采集的

电器类型包括了多种常用家用电器, 部分电器的电流波形如图 8 所示。为了增加采样数据的普遍程度, 使结果更加具有推广意义, 每类电器都有两种以上不同品牌样本, 每个样本都重复采样数据 10 次以上。

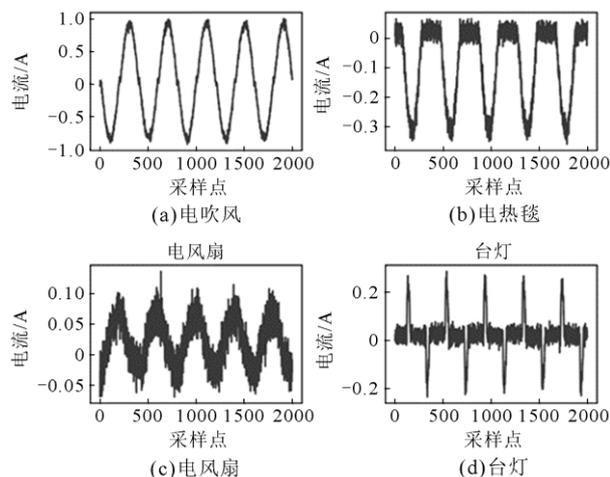


图 8 实验室部分电器电流波形图

Fig. 8 Current waveform of some electrical appliances in the laboratory

选取了实验室数据集中的部分电器, 包括电热毯、电吹风、电风扇、台灯和电热水壶, 对本文提出的负荷识别方法进行泛化验证。本实验中的 Att-ResNet 模型, 除了改变最后一层全连接层的神经元个数以外(与电器类别数一致), 结构与公开数据集实验完全一致, 该实验也能验证 Att-ResNet 模型的迁移学习能力与通用性。

本实验同样以 KNN 方法和随机森林方法作比较, 得到不同负荷识别方法的性能指标, 如表 3 所示。由实验结果可知:

1) 在 KNN、随机森林、Att-ResNet 这 3 种单一模型中, KNN 方法在电热毯上的识别效果较差, 仅有 77% 的准确率, 随机森林在电风扇上的准确率为 85%, 除此之外其他方法的准确率都达到 90% 以上, 3 种方法的总体准确率也都在 90% 以上。

2) 在应用了基于 Stacking 原理的负荷识别方法

表 3 实验室数据集各类指标比较

Table 3 Comparison of various indicators of different methods in the laboratory dataset

电器	Recall				Precision				F1			
	KNN ^[7]	RF ^[14]	ResNet	Stacking	KNN ^[7]	RF ^[14]	ResNet	Stacking	KNN ^[7]	RF ^[14]	ResNet	Stacking
电热毯	0.77	1.00	0.95	1.00	1.00	0.95	1.00	1.00	0.87	0.98	0.97	1.00
电吹风	1.00	1.00	1.00	1.00	0.82	0.98	0.98	0.99	0.90	0.99	0.99	0.99
电风扇	1.00	0.85	1.00	1.00	1.00	1.00	0.91	1.00	1.00	0.92	0.95	1.00
台灯	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
电热水壶	0.90	0.95	0.975	0.975	1.00	1.00	1.00	1.00	0.95	0.97	0.99	0.99
总体	0.925	0.979	0.983	0.996	0.94	0.98	0.98	1.00	0.92	0.98	0.98	1.00

后,所有电器的识别准确率都有所提高,电热毯、电吹风、电风扇和台灯的识别准确率达到100%,总体准确率也高于单一模型的效果,达到了99%以上,说明了本文方法的有效性以及实用价值。

5 结论

本文提出了一种基于异构基 Stacking 机制的居民用电负荷识别特征图像集成学习方法,在特征提取上,该方法选取了常规的电气参数如有功功率、无功功率等,并利用 mRMR 算法对于负荷特征进行初步筛选,并将负荷组合特征用 RGB 图像的形式表征出来,更好地提取负荷的深层特性;在分类器的选择上,引入了集成学习 Stacking 思想,融合多种异构模型,将 Att-ResNet 模型与多种传统机器学习算法作为基学习器对居民电器特征数据进行训练,再构建次级学习器,对于基学习器的结果进行学习,实现对电器类型的有效辨识。该模型在 PLAID 数据集和实验室数据集上都进行了仿真实验,结果表明该模型有较好的识别精度和泛化能力,未来期望能进一步缩小模型的规模,提高模型的计算速度,从而能在实际居民用电负荷识别场景中发挥作用。

参考文献

- [1] 时光远,王云飞,姜世公,等.基于融合模糊识别和人工神经网络的交直流配电系统发展阶段评估研究[J].供用电,2020,37(5):11-17.
SHI Guangyuan, WANG Yunfei, JIANG Shigong, et al. Research on developmental stage evaluation of AC/DC distribution system based on fuzzy recognition and artificial neural network[J]. Distribution & Utilization, 2020, 37(5): 11-17.
- [2] 江帆,杨洪耕.基于广义回归神经网络的非侵入式负荷识别方法[J].电测与仪表,2020,57(3):1-6.
JIANG Fan, YANG Honggeng. RNon-intrusive load identification method based on general regression neural network[J]. Electrical Measurement & Instrumentation, 2020, 57(3): 1-6.
- [3] 高云,杨洪耕.基于暂态特征贴度匹配的家用负荷识别[J].电力系统自动化,2013,37(9):54-59.
GAO Yun, YANG Honggeng. Household load identification based on closeness matching of transient characteristics[J]. Automation of Electric Power Systems, 2013, 37(9): 54-59.
- [4] 李如意,王晓换,胡美璇,等.RPROP神经网络在非侵入式负荷分解中的应用[J].电力系统保护与控制,2016,44(7):55-61.
LI Ruyi, WANG Xiaohuan, HU Meixuan, et al. Application of RPROP neural network in nonintrusive load decomposition[J]. Power System Protection and Control, 2016, 44(7): 55-61.
- [5] WANG Longjun, CHEN Xiaoming, WANG Gang, et al. Non-intrusive load monitoring algorithm based on features of V-I trajectory[J]. Electric Power Systems Research, 2018, 157: 134-144.
- [6] LIU Jianguo, LIU Ning, SONG Huina, et al. Non-intrusive load identification model based on 3D spatial feature and convolutional neural network[J]. Energy and Power Engineering, 2021, 13(4): 30-40.
- [7] TSAI M S, LIN Y H. Modern development of an adaptive non-intrusive appliance load monitoring system in electricity energy conservation[J]. Applied Energy, 2012, 96: 55-73.
- [8] GAO Mingze, PAN Shibo, CHEN Sirui, et al. Identification method of electrical load for electrical appliances based on K-Means++ and GCN[J]. IEEE Access, 2021, 9: 27026-27037.
- [9] REN Zhiren, TANG Bo, WANG Longfeng, et al. Household appliance identification based on a novel load signature processing framework[C] // 2019 IEEE 3rd Conference on Energy Internet and Energy System Integration (EI2), November 8-10, 2019, Changsha, China: 2076-2080.
- [10] ZHANG Ying, YIN Bo, CONG Yanping, et al. Multi-state household appliance identification based on convolutional neural networks and clustering[J]. Energies, 2020, 13(4): 792-803.
- [11] KIM J, LE T T H, KIM H. Nonintrusive load monitoring based on advanced deep learning and novel signature[J]. Computational Intelligence and Neuroscience, 2017, 2017: 31-42.
- [12] 刘恒勇,史帅彬,徐旭辉,等.一种关联RNN模型的非侵入式负荷辨识方法[J].电力系统保护与控制,2019,47(13):162-170.
LIU Hengyong, SHI Shuaibin, XU Xuhui, et al. A non-intrusive load identification method based on RNN model[J]. Power System Protection and Control, 2019, 47(13): 162-170.
- [13] HIMEUR Y, ALSALEMI A, BENSAAALI F, et al. Robust event-based non-intrusive appliance recognition using multi-scale wavelet packet tree and ensemble bagging tree[J]. Applied Energy, 2020, 267: 1-10.
- [14] LIU Hui, WU Haiping, YU Chenming. A hybrid model for appliance classification based on time series features[J]. Energy and Buildings, 2019, 196: 112-123.
- [15] GAO Jingkun, KARA E C, GIRI S, et al. A feasibility study of automated plug-load identification from high-

- frequency measurements[C] // IEEE Global Conference on Signal & Information Processing (GlobalSIP), December 14-16, 2015, Orlando, USA: 220-224.
- [16] PENG Hanchuan, LONG Fuhui, DING C. Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2005, 27(8): 1226-1238.
- [17] RAMIREZ-GALLEGO S, LASTRA I, MARTINEZ-REGO D, et al. Fast-mRMR: fast minimum redundancy maximum relevance algorithm for high-dimensional big data[J]. International Journal of Intelligent Systems, 2017, 32(2): 1-13.
- [18] 胡威, 张新燕, 李振恩, 等. 基于优化的VMD-mRMR-LSTM模型的短期负荷预测[J]. 电力系统保护与控制, 2022, 50(1): 88-97.
HU Wei, ZHANG Xinyan, LI Zhen'en, et al. Short-term load forecasting based on an optimized VMD-mRMR-LSTM model[J]. Power System Protection and Control, 2022, 50(1): 88-97.
- [19] KIRANMAI S A, LAXMI A J. Data mining for classification of power quality problems using WEKA and the effect of attributes on classification accuracy[J]. Protection and Control of Modern Power Systems, 2018, 3(3): 303-314.
- [20] 李焱, 贾雅君, 李磊, 等. 基于随机森林算法的短期电力负荷预测[J]. 电力系统保护与控制, 2020, 48(21): 117-124.
LI Yan, JIA Yajun, LI Lei, et al. Short term power load forecasting based on a stochastic forest algorithm[J]. Power System Protection and Control, 2020, 48(21): 117-124.
- [21] 杨杰, 吴浩, 董星星, 等. 基于电流故障分量特征和随机森林的输电线路故障类型识别[J]. 电力系统保护与控制, 2021, 49(13): 53-63.
YANG Jie, WU Hao, DONG Xingxing, et al. Transmission line fault type identification based on the characteristics of current fault components and random forest[J]. Power System Protection and Control, 2021, 49(13): 53-63.
- [22] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, et al. Deep residual learning for image recognition[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, June 27-30, 2016, Las Vegas, USA: 770-778.
- [23] WOO S H, PARK J, LEE J Y, et al. CBAM: convolutional block attention module[C] // 15th IEEE European Conference on Computer Vision, 2018, Munich, Germany: 3-19.
- [24] ZHOU Zhihua. Ensemble methods: foundations and algorithms[M]. Taylor & Francis, 2012.
- [25] 邓威, 郭钊秀, 李勇, 等. 基于特征选择和 Stacking 集成学习的配电网网损预测[J]. 电力系统保护与控制, 2020, 48(15): 108-115.
DENG Wei, GUO Yixiu, LI Yong, et al. Power losses prediction based on feature selection and Stacking integrated learning[J]. Power System Protection and Control, 2020, 48(15): 108-115.
- [26] 侯慧, 陈希, 李敏, 等. 一种 Stacking 集成结构的台风灾害下停电空间预测方法[J]. 电力系统保护与控制, 2022, 50(3): 76-84.
HOU Hui, CHEN Xi, LI Min, et al. A space prediction method for power outage in a typhoon disaster based on a Stacking integrated structure[J]. Power System Protection and Control, 2022, 50(3): 76-84.
- [27] GAO Jingkun, GIRI S, KARA E C, et al. PLAID: a public dataset of high-resolution electrical appliance measurements for load identification research: demo abstract[J]. ACM, 2014.
- [28] HAN Hui, WANG Wenyuan, MAO Binghuan. Borderline-SMOTE: a new over-sampling method in imbalanced data sets learning[C] // Proceedings of the 2005 International Conference on Advances in Intelligent Computing, Volume Part I, 2005: 878-887.

收稿日期: 2022-01-11; 修回日期: 2022-03-14

作者简介:

邱磊鑫(1998—), 男, 硕士研究生, 研究方向为非侵入式负荷分解、负荷识别; E-mail: 810236717@qq.com

余涛(1974—), 男, 通信作者, 博士, 教授, 博士生导师, 研究方向为复杂电力系统的非线性控制理论和仿真、优化及机器学习; E-mail: taoyu1@scut.edu.cn

彭秉刚(1997—), 男, 硕士研究生, 研究方向为非侵入式负荷分解、机器学习。E-mail: 344532641@qq.com

(编辑 许威)