

# 基于主动迁移学习的负荷辨识泛化方法

蔡清淮<sup>1</sup>, 罗庆全<sup>1</sup>, 余涛<sup>1,2</sup>, 刘前进<sup>1</sup>, 刘熙鹏<sup>1</sup>, 潘振宁<sup>1</sup>

(1. 华南理工大学 电力学院, 广东 广州 510640; 2. 广东省电网智能量测与先进计量企业重点实验室, 广东 广州 510640)

**摘要:**为实现低成本地提升负荷辨识方法对新场景负荷样本的识别精度,提出一种基于主动迁移学习的负荷辨识泛化方法,利用极少量标签数据和无标签数据来高效提升方法的泛化性能。该方法利用异构模型的共识筛选高质量的伪标注样本,并对预训练模型进行更新;设计一种考虑模型分歧和样本多样性的主动学习策略来标注高价值样本,可在大幅降低样本标注成本的同时实现模型的高效迁移。在2个公开数据集集中的实验对比,验证了所提方法的优越性。

**关键词:**负荷辨识;泛化;主动学习;迁移学习;伪标注;异构模型

**中图分类号:**TM714;TP18

**文献标志码:**A

**DOI:**10.16081/j.epae.202411013

## 0 引言

能耗计量和节能减排已成为“双碳”目标背景下的发展趋势,2022年居民生活用电量高达 $1.3366 \times 10^{12}$  kW·h<sup>[1]</sup>,是节能减排工作的重点。负荷辨识技术通过电气数据快速识别负荷类型,可实现用户设备级的用电能效监测和感知,在此基础上设计家庭能量管理方案,优化用户用电模式,可达到高效节能的目标<sup>[2-3]</sup>。文献[4]表明,在负荷监测的支持下,不同用电系统可取得10%~40%的节能效果,因此负荷辨识技术成为当前研究的热点。

学者已提出许多负荷辨识方法并取得了初步成果。在结合电气领域特征的机器学习算法方面,文献[5-6]在提取有功功率、谐波幅值等典型的电气特征后,采用集成学习方法如随机森林(random forest, RF)等,设计多样化的基学习器来组合实现辨识。在深度学习算法方面:文献[7]利用双长短期记忆网络有效实现了对负荷功率序列的特征挖掘和辨识;文献[8]在将负荷 $V-I$ 轨迹图序列化后,利用一维卷积神经网络(1D convolutional neural network, 1D-CNN)以较小的计算量捕捉序列的细节变化。以

上方法虽能有效区分负荷,但高精度仅限于单一数据集,泛化能力不足。如文献[9]在将PLAID数据集训练的模型直接应用于WHITED数据集后,仅能取得23.8%的准确率。

目前很少有文献探讨如何提高负荷辨识模型的泛化能力。迁移学习可利用从旧任务学到的知识来改进新任务的泛化性能,常用的方法是微调,即在预训练模型参数的基础上,进一步利用新任务数据调整模型参数,以适应新任务,如文献[9-11]采用基于模型的迁移方法,利用负荷样本对AlexNet等预训练模型进行参数微调,将经典分类模型强大的表征能力和分类知识迁移至负荷辨识任务,虽然该方法可使模型在应用于新数据集时无须从头训练,但要达到令人满意的性能则需要依赖目标数据集中大量的有标签样本进行训练,存在成本过高的问题,尚不能满足泛化应用的要求。

在实际应用中,无标签样本相对容易获取且成本低廉,因此负荷辨识领域出现了利用无标签数据辅助模型训练的相关研究<sup>[12-13]</sup>,如文献[12]迭代执行基于马氏距离的最近邻算法,为无标签样本分配伪标签以扩充训练样本。然而,这类直接基于样本距离学习无标签样本的方法难以有效解决负荷特性相近但标签不同的问题,也不能有效提高算法的泛化性。

此外,负荷辨识领域也出现了以多学习器利用无标签数据的尝试,如文献[14]设置双模型对无标签样本进行筛选以及交换学习,但该工作在泛化应用上存在局限,无标签样本的筛选及伪标注仍存在改进空间。总体而言,利用无标签样本来提升负荷辨识泛化性能的研究仍较少。

除了利用无标签样本外,解决泛化问题的另一思路是研究如何在提升模型性能的同时,降低对有标签样本的需求量。主动学习技术通过智能选择最具价值的样本进行标注和学习,使识别模型可对样

收稿日期:2023-12-11;修回日期:2024-03-22

在线出版日期:2024-11-21

**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资助项目(U2066212);广州市基础研究计划基础与应用基础研究项目(SL2022A04J01135);广东省科技计划项目(广东省电网智能量测与先进计量企业重点实验室(2021年度))(2021B1212050014)

Project supported by the National Natural Science Foundation of China-Smart Grid Joint Fund of State Grid Corporation of China(U2066212), Guangzhou Basic and Applied Basic Research Program(SL2022A04J01135) and Guangdong Provincial Science and Technology Program(Guangdong Provincial Key Laboratory of Intelligent Measurement and Advanced Metering of Power Grid(2021))(2021B1212050014)

本集进行更加精准和针对性的训练,与传统监督学习相比,可显著降低训练成本<sup>[15]</sup>。文献[16-17]初步将主动学习应用于负荷辨识,依据样本信息量和不确定性来挑选样本进行标注,供模型学习,但这类学习策略仍较为简单,易导致所挑选的样本之间冗余度较高,如何以低成本高效提升负荷辨识的泛化性能仍有待进一步研究。

综上,现有负荷辨识方法普遍存在泛化能力不足的问题,严重阻碍了负荷辨识方法的推广应用<sup>[18]</sup>,而利用大量有标签样本来改善模型的方式与实际条件相悖,因此,研究一种无须增加过多标注成本即可有效提升负荷辨识泛化性的方法,对于促进负荷辨识的应用具有重要意义。为此,本文提出一种基于主动迁移学习的负荷辨识泛化框架,低成本地提高识别算法在不同数据中的可用性,解决常被忽视的泛化性难题。本文主要贡献如下:

1)所提出的方法框架针对有标签样本稀缺导致泛化能力不足的问题,利用异构模型间的共识对无标签样本进行伪标注,以辅助模型训练,使模型适应新应用场景的数据分布;

2)针对常用迁移学习方法过多依赖于有标签样本数量的问题,设计一种结合模型分歧和样本多样性的主动学习方法,使模型针对性地学习最能代表新场景特征空间的不确定样本,在高效提升模型性能的同时,节省了巨大的样本标注成本。

## 1 问题描述及负荷辨识泛化框架

### 1.1 问题描述

负荷辨识的任务是根据已有的负荷数据集 $D = \{(\mathbf{x}_i, y_i)\}_{i=1}^n$ ,建立负荷电气数据到负荷类别的映射函数,其中: $\mathbf{x}_i$ 和 $y_i$ 分别为数据集中第 $i$ 个样本的数据和标签; $n$ 为数据集中的样本数量。目前,大多负荷辨识方法的高精度性能是建立在训练集和测试集数据分布相同的理想假设基础上的,然而,在实际应用中,由于数据采样方法、电网状态、负荷品牌、样本标注方法等方面的多元差异,预训练模型所使用的数据集和模型应用时所面对的样本集通常并不满足独立同分布的理想假设,因此,绝大多数负荷辨识方法在应用于新数据集时往往出现性能明显下降的问题。将用于预训练模型的数据集称为源域,将模型应用时所面对的样本集称为目标域,设 $\mathbf{x}$ 和 $y$ 分别为负荷样本数据和对应的类别标签,源域样本数据 $X_S$ 及其对应的标签 $Y_S$ 满足分布 $P_S(\mathbf{x}, y)$ ,目标域样本数据 $X_T$ 及其对应的标签 $Y_T$ 满足分布 $P_T(\mathbf{x}, y)$ 。由于域间差异的存在,通常有 $P_S(\mathbf{x}, y) \neq P_T(\mathbf{x}, y)$ ,这使预先在源域训练的负荷辨识模型的通用性难以得到保证<sup>[19]</sup>。

针对上述问题,本文研究负荷辨识的泛化问题,旨在提升模型在目标域应用时的识别精度。设源域数据集 $D_S = \{(\mathbf{x}_i^S, y_i^S)\}_{i=1}^{n_S}$ ,目标域数据集 $D_T = \{(\mathbf{x}_i^T, y_i^T)\}_{i=1}^{n_T}$ ,其中: $\mathbf{x}_i^S$ 和 $\mathbf{x}_i^T$ 分别为源域数据集和目标域数据集中第 $i$ 个样本的数据; $y_i^S$ 和 $y_i^T$ 分别为源域数据集和目标域数据集中第 $i$ 个样本的标签; $n_S$ 和 $n_T$ 分别为源域和目标域的样本数量。目标域样本未标注标签,即 $Y_T$ 未知。由于实际中目标域的无标签负荷数据容易在数据采集终端获得,假设目标域样本数据 $X_T = \{\mathbf{x}_i^T\}_{i=1}^{n_T}$ 可以用于模型的训练过程。本文旨在利用 $D_S$ 预训练得到的源域识别模型及 $X_T$ ,设计一种目标域标注成本低的负荷辨识泛化方法,提升模型在目标域的识别精度。

### 1.2 基于主动迁移学习的负荷辨识泛化框架

基于主动迁移学习的负荷辨识泛化框架如图1所示。

1)异构模型设计:使用源域数据集构建以电流周期为输入的深度学习模型和以特征量为输入的机器学习模型,目的是更好地发挥本文主动迁移学习方法的优势。

2)主动迁移学习方法:根据异构模型的预测结果,从无标签样本中筛选出可靠样本并给定伪标签,利用目标域伪标注样本对源域预训练模型的参数进行训练微调,以使模型迁移适应目标域。此外,结合主动学习方法,提出最大代表性抽样(maximum representative sampling, MRS)方法,从目标域不确定样本中挑选小样本集并标注真实标签,帮助模型对目标域进行针对性的学习。

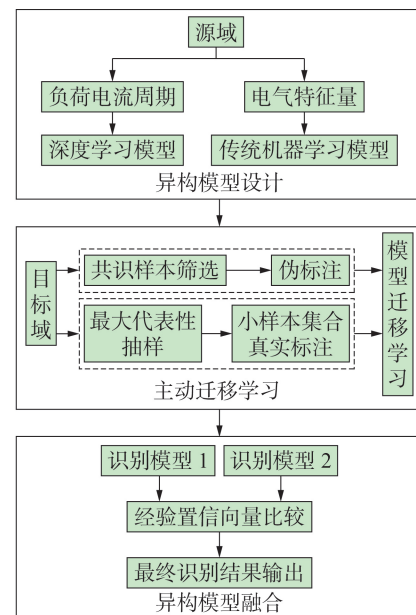


图1 本文方法总体框架

Fig.1 General framework of proposed method



3)异构模型融合:更新后的异构模型采用基于经验置信向量的融合策略,输出最终识别结果。

上述主动迁移学习的2个机制可优化模型参数的更新,从而使模型适应目标域,实现在极低的标注成本下大幅提高模型在目标域的泛化性能。

## 2 异构模型主动迁移学习方法

### 2.1 基于伪标注样本的迁移学习

根据第1章,源域预训练模型在新目标域的性能通常会出现明显下降,但不同模型由于在源域所拟合的分布存在差异,在目标域的识别能力也不同,而多模型的设置可从不同视角对负荷类型进行辨识,并对各自的识别结果进行相互验证或参考,从而减少误判<sup>[20]</sup>。文献[21]推导出的误差-分歧分解公式为:

$$E = \bar{E} - \bar{A} \quad (1)$$

式中: $E$ 为集成模型的泛化误差; $\bar{E}$ 为基模型泛化误差的均值; $\bar{A}$ 为基模型的分歧值,它反映基模型预测的不一致性和基模型的多样性。式(1)表明基模型的多样性越大,集成模型越优,因此,本文构建异构识别模型,使得识别一致的结果具有更高的置信度。将异构模型识别一致的结果作为无标签样本的伪标签,有利于辅助模型训练,以适应目标域。

基于伪标注样本的迁移学习流程如附录A图A1所示。首先,将目标域视为无标签样本库,将基模型预测结果一致的样本不放回取出,加入至可靠样本库,并以预测结果作为样本的伪标签。然后,异构模型基于可靠样本库进行参数更新,即模型在原有参数的基础上利用可靠样本进行训练。更新后的基模型对目标域样本再次进行识别,重复上述步骤,直至达到最大迭代次数 $\lambda$ 或无标签样本数量小于阈值 $\delta$ 。异构模型的参数在每次学习新“共识”样本的过程中进行调整,从而逐步适应目标域的数据分布。

### 2.2 主动学习

本文结合主动学习理论,最大限度降低标注成本,利用有限标注量指导目标域的差异化特性挖掘,以提升负荷辨识效果。

本文采用基于分歧的主动学习方法,为学习器预测不一致的样本标注真实标签并进行学习。与单学习器的主动学习相比,该方法能用更少的样本达到相当甚至更强的泛化性能<sup>[22]</sup>。具体地,若将学习器的期望误差记为 $\varepsilon$ ,则单学习器的主动学习样本复杂度为 $\tilde{O}(1/\varepsilon)$ ,而在不需要对数据分布进行强假设的情况下,基于分歧的主动学习可将复杂度降低至 $\tilde{O}(\log_q(1/\varepsilon))$ <sup>[23]</sup>,其中 $q$ 为任意大于0的数。因此,采用不同的特征输入和不同的模型结构来保证基模型间的分歧,可充分发挥基于分歧的主动学习方法

对模型性能的提高作用。

本文采用2.1节的双异构模型框架,在不增加模型开销的条件下实现基于分歧的主动学习。具体流程如图A1左边虚线框所示。首先,用2个基模型分别对目标域样本进行预测,将预测结果不一致的样本视为不确定样本;然后,按照特定的抽样策略从不确定样本集中选择少量样本,查询样本的真实标签并进行标注,将该样本作为可靠样本加入2.1节的可靠样本库中,用于异构模型的训练更新。在实际应用过程中,负荷样本的标注可通过移动端程序进行,将开启的且为待标注的负荷信息推送给用户完成标注。

主动学习降低标注成本的关键在于设计合适的抽样策略,即选择对模型提升最大的样本集进行标注。最常用的做法是采用基于不确定性的抽样策略进行标注<sup>[16-17]</sup>,使模型学习最不确定标签的样本,如边缘抽样(margin sampling, MS)法和信息熵抽样(information entropy sampling, IES)法,但这些抽样策略在选择样本的过程中未考虑样本间的相似性,易导致重复标注多个相似样本,不利于降低标注成本。针对以上不足,本文考虑样本的多样性,提出MRS策略,挑选能够广泛覆盖特征空间的样本,并对这些样本构成的样本集进行标注,以实现标注次数少、性能提升大的目标。

MRS策略的代表性是基于样本间的相似度进行计算的。首先,定义样本子集 $X_{\text{select}}$ 对目标域集合 $X_T$ 的代表性分值 $R(X_{\text{select}}, X_T)$ 为:

$$R(X_{\text{select}}, X_T) = \sum_{\mathbf{x}_j \in X_T} r(X_{\text{select}}, \mathbf{x}_j) \quad (2)$$

式中: $r(X_{\text{select}}, \mathbf{x}_j)$ 为集合 $X_{\text{select}}$ 对目标域样本 $\mathbf{x}_j$ 的代表性分值,其计算公式如式(3)所示。

$$r(X_{\text{select}}, \mathbf{x}_j) = \max_{\mathbf{x}_i \in X_{\text{select}}} s(\mathbf{x}_i, \mathbf{x}_j) \quad (3)$$

式中: $s(\mathbf{x}_i, \mathbf{x}_j)$ 为样本 $\mathbf{x}_i$ 和样本 $\mathbf{x}_j$ 的相似度,采用欧几里得范数进行度量。由上述代表性分值计算方式可知,为最大化 $R(X_{\text{select}}, X_T)$ ,集合 $X_{\text{select}}$ 所包含的样本多样性应尽可能丰富,以使更多的目标域样本有较高的 $r(X_{\text{select}}, \mathbf{x}_j)$ 。

基于上述样本代表性的度量方式,本文提出的结合模型分歧的MRS主动学习策略如附录A表A1所示。

## 3 异构模型设计

2.1节的伪标签设计方法和2.2节的基于分歧的主动学习方法均利用了基模型不同的识别能力,因此,构建存在足够差异的识别模型有利于充分发挥所提泛化方法的优势。1D-CNN模型和RF模型在具

备良好性能的同时,还兼顾了实用性<sup>[6,8]</sup>,而2种模型在结构和技术路线上不同:1D-CNN代表深度学习技术,通过拟合参数对输入进行变换,以此近似目标的概率分布;不同于深度学习路线,RF模型采用集成学习方法,通过采用Bagging算法结合多个决策树,并利用不同特征量逐步分割特征空间。因此,这2种模型在实用性和异构性上与本文方法适配。

对应于不同模型,本文在模型的输入量上进一步差异化,将负荷样本转换为电流周期序列和电压-电流特征量2种特征描述,其中电流周期序列可以更好地保留波形的变化特征,而电压-电流特征量则可在降低维度的同时考虑电压与电流的相关关系。因此,本文采用以电流周期序列为输入的1D-CNN模型和以电压-电流特征量为输入的RF模型作为异构模型。下面具体描述2个基模型的构建。

本文1D-CNN模型由卷积层Conv、批归一化层BatchNorm、最大池化层MaxPool、全连接层FC堆叠而成。卷积层能有效挖掘时间序列的局部特征,批归一化层可以减少模型训练过程中内部协变量偏移的影响,最大池化层在降低特征维度的同时能够扩大对序列的感知。所建立的模型结构如附录A图A2所示。通过最小化式(4)所示交叉熵损失函数 $L$ 训练模型参数。

$$L = -\frac{1}{m} \sum_{i=1}^m \sum_{c=1}^N l_{ic} \ln p_{ic} \quad (4)$$

式中: $m$ 为训练样本的数量; $N$ 为类别数; $l_{ic} \in \{0, 1\}$ ,若第*i*个样本真实标签为类别*c*,则其取值为1,否则其取值为0; $p_{ic}$ 为观测第*i*个样本真实标签为类别*c*的预测概率。

本文RF模型的输入量为13维的电压-电流特征向量,由电压、电流周期数据计算得到,包括电流峰峰值、电流有效值、电流波形因数、电流谐波畸变率、电流相对谐波畸变率、有功功率、无功功率、奇次(1、3、5次)谐波电流有效值及相位。在模型训练上,采用自举法将训练数据重采样为多个样本集后,从每个样本集中选择最优特征并对其进行逐步分割以构建对应的决策树,最终生成的决策树集合即为RF模型,识别结果由决策树以多数投票法的方式给出。

由于模型结构和技术路线的差异,在学习目标域样本时2个基模型的更新方式也不同。1D-CNN通过网络参数的梯度下降算法学习捕捉波形的细节特性,因此,其更新采用参数微调的策略,即在每次可靠样本库更新后,在原参数的基础上根据模型对可靠样本的损失值进行梯度下降。

而RF的输入特征结合了电气领域知识,对于不同负荷类型本身具有良好的可分性,并且维数更低,相比于深度学习方法,RF模型调节参数少,训练速度更快<sup>[24]</sup>,因此,本文采用重新拟合的策略对RF模

型进行更新,即在每次可靠样本库更新后,利用可靠样本库与源域数据集,训练得到一个新的RF模型,这种更新方式有利于保持RF模型在源域的性能,并避免模型对目标域数据集过拟合。

## 4 异构模型融合策略

为综合利用1D-CNN和RF的识别能力,并解决两者识别不一致时如何输出识别结果的问题,本章提出一种模型融合策略,用于融合两者各自的识别结果,以获得最终输出。最常用的模型融合策略是采用基于最大类概率的方式,设基模型 $h_a$ 在样本 $\mathbf{x}$ 上的预测输出为一个 $N$ 维向量 $h_a(\mathbf{x})$ ,设 $h_a^c(\mathbf{x})$ 为 $h_a$ 在类别*c*上的输出,即类概率,且 $h_a^c(\mathbf{x}) \in [0, 1]$ ,基于最大类概率的融合策略比较所有不同基模型的 $h_a(\mathbf{x})$ ,将最大类概率值所对应的类别作为最终结果的输出。但由于不同类型模型的概率估计方法和假设不同,输出的类概率具有不同的含义,不能直接比较类概率<sup>[21]</sup>,因此,该策略不适用于本文的异构模型。为此,本文提出以经验置信向量代替类概率向量的模型融合策略,以RF模型为例,设 $v_1 \sim v_N$ 分别表示识别模型对负荷类别1— $N$ 的经验置信度,则该模型经验置信向量 $\mathbf{v} = [v_1, v_2, \dots, v_N]$ 的计算步骤如下。

1)采用五折交叉验证的方式将源域数据集划分为训练集和验证集。

2)利用训练集建立RF模型。

3)利用验证集测试步骤2)所构建的RF模型性能,将RF模型在负荷类别*c*上的精确率作为该模型的经验置信度,即:

$$v_c = \frac{T_{p,c}}{T_{p,c} + F_{p,c}} \quad c = 1, 2, \dots, N \quad (5)$$

式中: $T_{p,c}$ 为负荷真实类别为*c*且预测类别也为*c*的样本数; $F_{p,c}$ 为负荷真实类别不为*c*而预测类别为*c*的样本数。

4)重复执行步骤1)–3),直至五折交叉验证结束,取这5次 $\mathbf{v}$ 的平均值作为经验置信向量。

在得到2个模型的经验置信向量后,当1D-CNN模型和RF模型的识别结果不一致时,比较两者识别结果所对应的经验置信度,将经验置信度较高的类别作为异构模型最终的识别结果。

## 5 算例分析

### 5.1 泛化实验设置

本文实验数据集来自PLAID<sup>[25]</sup>、WHITED<sup>[26]</sup>这2个公共数据集。PLAID为以30 kHz采样频率采集的美国宾夕法尼亚州56个不同用户的用电数据,包括329台电器的电压、电流波形数据。WHITED为以44.1 kHz采样频率采集的德国、澳大利亚、印度尼



西亚、加拿大4个区域的电器电压、电流波形数据。本文选取上述2个公共数据集的共有类别作为实验数据集,共有类别包括风扇、吸尘器、冰箱、热水器、荧光灯、电吹风、灯泡、微波炉、空调、洗衣机和笔记本电脑,实验数据集共有2174条样本,其中1824条样本来自PLAID,350条样本来自WHITED,所有样本的数据均降采样至7.5 kHz。

为构建存在域间分布差异的源域和目标域来验证本文泛化方法的有效性,本文从公共数据集、电器品牌、住宅3个角度将负荷样本划分为源域和目标域,将源域作为训练集,目标域作为测试集,3种角度的泛化实验设置方法分别如下。

1)数据集泛化:在PLAID中所选择的11种负荷样本数量较多,预训练的识别模型可以避免因训练集过度拟合而对实验造成干扰,因此,以PLAID为源域,WHITED为目标域。

2)品牌泛化:对于每类电器,随机抽取1/4的品牌作为目标域,其余品牌作为源域。

3)住宅泛化:随机抽取10个住宅的负荷数据作为目标域,其余住宅或地区的负荷数据作为源域。

## 5.2 实验平台

本文所提方法的实验验证环境如下:硬件环境为Intel(R) Core(TM) i7-10700F CPU @ 2.90 GHz, 16 GB DDR4内存的64位计算机。软件环境基于Windows 11操作系统,采用Python3.8及Pytorch深度学习框架。

## 5.3 评价指标

本文采用准确率 $I_{acc}$ 、宏平均F1分数 $F_1$ 、马修斯相关系数 $I_{mcc}$ 作为实验评价指标,各评价指标具体计算公式如附录A式(A1)~(A5)所示。

## 5.4 识别模型对比

为验证本文异构模型融合识别方法的有效性,以及所建立的1D-CNN和RF的实用性和优势,设置以下负荷辨识方法进行对比:本文方法,采用本文模型融合策略的1D-CNN和RF融合识别方法; $V-I$ 方法,文献[27]提出的以负荷 $V-I$ 轨迹为输入的卷积神经网络,该网络在对比实验中常被作为基线模型;自适应加权递归图(adaptive weighted recurrence graph, AWRG)方法,文献[28]中基于电流波形的AWRG和卷积神经网络结合的识别模型是目前负荷辨识领域的先进方法;AlexNet方法,文献[9]基于经典AlexNet深度学习模型的负荷辨识方法利用AlexNet表征能力和分类知识取得了优秀的性能;1D-CNN;RF。比较上述6种负荷辨识方法在非泛化实验中的识别精度。区别于5.1节的泛化实验设置方式,非泛化实验按照4:1的比例将实验数据集随机划分为训练集和测试集,分别用于识别模型的训练和测试,为避免实验的偶然性,对上述过程重复5次,取这5次测试集

的识别精度平均值作为各方法的最终识别精度。实验结果如表1所示。

表1 各识别方法的非泛化实验结果

Table 1 Results of non-generalization experiment for each identification method

方法	$I_{acc}$	$F_1$	$I_{mcc}$
RF	0.973	0.967	0.970
1D-CNN	0.931	0.929	0.923
$V-I$	0.843	0.829	0.826
AWRG	0.948	0.937	0.942
AlexNet	0.928	0.924	0.920
本文方法	0.977	0.976	0.974

由表1可知:本文方法比1D-CNN和RF这2个基模型的性能更优,且总体上取得了最高的识别精度,这验证了本文异构模型融合识别方法的有效性;1D-CNN的性能与AWRG方法的接近,优于AlexNet方法和 $V-I$ 方法;在模型复杂度上,1D-CNN是其中最轻量的神经网络,仅需134250个参数,该参数数量低至 $V-I$ 方法和AWRG方法参数数量的20%,可见1D-CNN在识别精度和模型复杂度上兼具优势,且其输入无需复杂的预处理,因此本文以1D-CNN作为异构模型之一,可使本文方法更具有轻量化的实用性优势。

进一步地,对以上6种负荷辨识方法进行泛化实验,泛化实验设置同5.1节,所有方法均在目标域数据集上进行测试,结果如图2所示。由图可知,传统负荷辨识方法的泛化能力不足,特别是在数据集泛化实验中,相较于上述非泛化实验的结果,6种方法的精度均下降明显,AlexNet方法仅得到0.35的准确率,这可能是由于PLAID和WHITED的负荷数据存在电网状态、品牌等多种差异,使得域间差异相对较高,识别难度高。

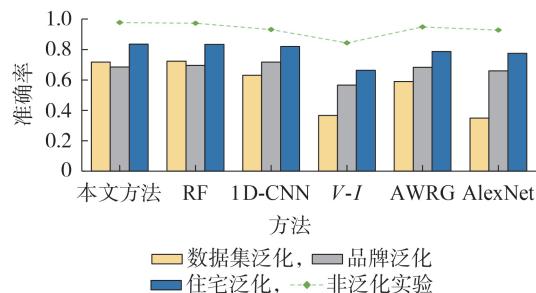


图2 6种识别方法的泛化实验结果

Fig.2 Results of generalization experiment for six identification methods

## 5.5 泛化性能对比

为验证本文主动迁移学习方法对于提升负荷辨识泛化能力的有效性和优势,将本文方法与当前负荷辨识中主流的迁移学习方法进行对比。选择文献[9]中基于AlexNet的迁移学习方法和文献[10]中基

于VGG16卷积神经网络的迁移学习方法进行实验对比。为对比分析各方法在不同目标域标注量下的泛化性能,设置5%、20%的标注比例分别作为目标域小标注量、大标注量的典型值进行讨论。具体设置如下:finetune-AlexNet方法,先根据文献[9]方法在源域数据集上预训练AlexNet模型,再依次从目标域随机抽样5%、20%的样本进行迁移学习,对模型参数进行微调;finetune-VGG16方法,先根据文献[10]方法在源域数据集上预训练VGG16模型,再依次从目标域随机抽样5%、20%的样本进行迁移学习,对模型参数进行微调;本文方法,在源域数据集上预训练1D-CNN和RF,随后结合模型分歧和MRS主动学习策略,从目标域中一次性抽样并标注一定比例的样本,并基于异构模型的共识从目标域迭代获取伪标注样本,将以上样本共同用于异构模型参数的训练更新,目标域标注比例同样设置为5%和20%分别进行实验。

需要说明的是:在异构模型的MRS主动学习策略计算相似度的过程中,负荷样本数据 $x$ 由1D-CNN和RF的输入量共同表征,但为减少电流周期序列中高频噪声对相似度计算的影响, $x$ 由电流周期在1D-CNN第2个全连接层的输出向量(64维)和RF模型的输入向量(13维)组成;为有效利用无标签样本,本文方法设置 $\lambda=10$ 、 $\delta=5$ ;为了避免上述神经网络模型在迁移学习过程中对少量样本过拟合,当模型对训练集的准确率保持在0.95以上时停止训练。

为避免实验结果的偶然性,从数据集、品牌、住宅3个角度的泛化实验对3种方法进行对比,所有方法均在目标域数据集上进行测试,结果如附录B表B1所示。5%的目标域标注量下3种方法模型更新前后的识别性能增益如附录B图B1所示,性能增益等于各模型参数更新前后的各项评价指标值之差。根据实验结果可知:本文方法在仅5%的目标域标注量下取得了最优的泛化性能,且综合性能增益最大;在5%的目标域标注量下,finetune-AlexNet和finetune-VGG16效果不理想,主要原因是两者的迁移学习方式未考虑有标签样本的来源,同时识别模型参数复杂,因而迁移学习须依赖大量的有标签样本进行参数调优,且在迁移学习前,两者对目标域的初始性能过低;当目标域标注量达到20%时,finetune-AlexNet和finetune-VGG16的泛化性能明显提升,但仍低于5%目标域标注量下本文方法的泛化性能,本文方法在大幅节省样本标注工作的同时,高效提升了模型的泛化性能。

为更进一步细化分析在目标域小标注量条件下本文方法泛化性能的增益,选择从3%开始逐渐提高目标域标注比例进行实验,直至目标域的综合识别精度超过0.9。以准确率 $I_{acc}$ 、宏平均F1分数 $F_1$ 、马

修斯相关系数 $I_{mcc}$ 三者的平均值作为综合识别精度,数据集泛化实验结果如附录B表B2所示。由表可知,要实现从PLAID迁移至WHITED并达到良好的精度(综合识别精度超过0.9),本文方法对WHITED的标注量需求可以低至6%,可见,本文主动迁移学习方法仅需增加极少量的标注成本即可高效提升泛化性。

## 5.6 主动学习策略对比

为单独验证本文MRS主动学习策略的有效性和优越性,选择文献[16-17]的主动学习策略进行对比。为避免实验的偶然性,并对比分析在不同主动学习策略下识别模型性能随着目标域标注量的变化趋势,分别设置0%、3%、6%、9%、12%的目标域抽样比例进行多组实验。按照数据集泛化实验设置源域和目标域,以源域数据预训练RF和1D-CNN,仅利用主动学习策略采样并标注的目标域样本进行模型更新,并在目标域数据集上测试模型的识别精度。图3给出了在数据集泛化实验中,RF、1D-CNN在不同策略、不同标注量条件下的准确率对比。

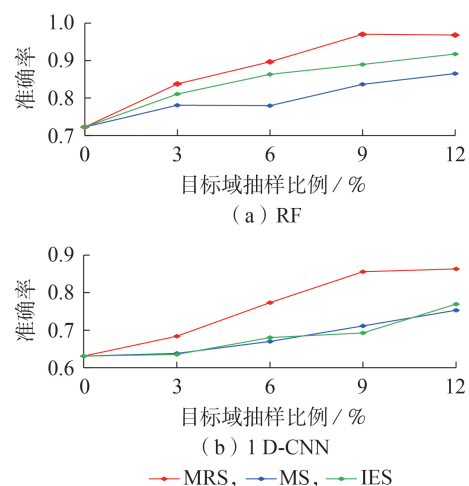


图3 不同主动学习策略对比

Fig.3 Comparison of different active learning strategies

由图3可知:由于IES和MS的低价值、重复采样,其性能难以快速提升,而须抽样并标注更多的样本才能达到与本文MRS策略相当的性能;本文MRS策略在主动采样的过程中考虑样本多样性,减少了重复抽样,有利于模型学习目标域的样本分布,从而高效提高泛化性能。

## 5.7 消融实验

为进一步验证本文异构模型、伪标注样本学习、主动学习等多个环节对提升泛化性的作用,本节设计对照实验,在本文方法中替换或去除部分环节作为对照方法,验证各环节对于泛化识别的改进作用,具体实验设置及分析过程如附录D所示。由实验结



果可知,本文结合模型分歧和MRS策略进行主动学习的方式,有利于更高效地提升模型的泛化能力,所提出的伪标注样本学习环节实现了无标签数据的有效利用,且本文方式与文献[14]的方式相比更具优势。

## 6 结论

本文针对传统负荷辨识方法泛化能力不足的问题展开研究,提出一种基于主动迁移学习的负荷辨识泛化方法,以无标签样本和极少量的目标域标注样本大幅提高了源域预训练模型在目标域的泛化能力。本文通过利用PLAID和WHITED公共数据集的数据,设置3种泛化实验进行验证,由实验结果可得如下结论。

1)所设计的异构模型能够有效地将无标签样本应用于泛化性能的提升,根据异构模型间的共识可为目标域样本分配高质量的伪标签,从而有效辅助模型训练,使模型根据目标域负荷数据自动迭代更新。

2)所提出的MRS方法与异构模型分歧结合的主动学习策略,可挑选广泛覆盖目标域样本空间且不确定性高的负荷样本进行学习,相比于其他主动学习策略,能更充分发挥主动学习降低标注成本的优势,更高效地提升识别模型的性能。

3)与负荷辨识中主流的迁移学习方法相比,本文以异构模型为框架的主动迁移方法是一种低成本的负荷辨识泛化方法,仅需利用极少量的标注样本和无标签样本即可高效提升模型的泛化性能。在本文的泛化实验中,本文方法仅需5%的目标域标注量和无标签样本即可超越使用20%目标域标注量的其他迁移学习方法。

总体而言,本文提出的泛化方法符合负荷辨识方法低成本、高性能的实际需求,为负荷辨识方法泛化性差的问题提供了解决思路。尽管如此,本文方法在模型设计和模型融合策略方面仍需进一步完善。未来研究工作主要包括:在保持泛化性的同时,进一步降低异构模型的计算复杂度;研究目标域样本存在标签噪声时的泛化方法;研究如何在主动迁移学习过程中动态调整模型的经验置信向量,以更好地表征异构模型预测结果的可信度。

附录见本刊网络版(<http://www.epae.cn>)。

## 参考文献:

- [1] 国家能源局. 2022年全社会用电量同比增长3.6%[EB/OL]. [2023-03-21]. [https://www.nea.gov.cn/2023-01/18/c\\_13106-91508.htm](https://www.nea.gov.cn/2023-01/18/c_13106-91508.htm).
- [2] 徐伟枫,华锦修,余涛,等. 计及电器状态关联规则的非侵入式负荷分解[J]. 电力自动化设备,2020,40(4):197-203.  
XU Weifeng, HUA Jinxiu, YU Tao, et al. Non-intrusive load

decomposition considering association rules of appliances' state[J]. Electric Power Automation Equipment, 2020, 40(4): 197-203.

- [3] 冯昌森,刘攀,王佳颖,等. 利用有限低频信息的居民用户非侵入负荷监测算法[J]. 电力自动化设备,2023,43(11):181-187.  
FENG Changsen, LIU Pan, WANG Jiaying, et al. Non-intrusive load monitoring algorithm of residential users using limited low-frequency information[J]. Electric Power Automation Equipment, 2023, 43(11): 181-187.
- [4] LEE D S, CHENG C C. Energy savings by energy management systems: a review[J]. Renewable and Sustainable Energy Reviews, 2016, 56: 760-777.
- [5] 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 ensemble non-intrusive load identification method based on Shannon entropy weighted voting algorithm[J]. Proceedings of the CSEE, 2022, 42(24): 8876-8888.
- [6] TABANELLI E, BRUNELLI D, ACQUAVIVA A, et al. Trimming feature extraction and inference for MCU-based edge NILM: a systematic approach[J]. IEEE Transactions on Industrial Informatics, 2022, 18(2): 943-952.
- [7] 周步祥,赵雯雯,臧天磊,等. 基于低频功率差量特征与双长短期记忆网络的非侵入式负荷监测方法[J]. 电力自动化设备, 2023, 43(8): 167-173, 209.  
ZHOU Buxiang, ZHAO Wenwen, ZANG Tianlei, et al. Non-intrusive load monitoring method based on low-frequency power difference characteristic and dual long short-term memory network[J]. Electric Power Automation Equipment, 2023, 43(8): 167-173, 209.
- [8] 杨桂兴,王维庆,姚红雨,等. 基于1DCNN-BP的非侵入式负荷识别算法[J]. 高压技术,2023,49(7):3031-3039.  
YANG Guixing, WANG Weiqing, YAO Hongyu, et al. Research on non-intrusive load identification method based on 1DCNN-BP[J]. High Voltage Engineering, 2023, 49(7): 3031-3039.
- [9] LIU Y C, 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.
- [10] 崔昊杨,蔡杰,陈磊,等. 基于颜色编码的非侵入式负荷细粒度识别方法[J]. 电网技术,2022,46(4):1557-1567.  
CUI Haoyang, CAI Jie, CHEN Lei, et al. Non-intrusive load fine-grained identification based on color encoding[J]. Power System Technology, 2022, 46(4): 1557-1567.
- [11] 鲍光海,黄逸欣. 基于ResNeXt网络和迁移学习的非侵入式负荷监测[J]. 电力系统自动化,2023,47(13):110-120.  
BAO Guanghai, HUANG Yixin. Non-intrusive load monitoring based on ResNeXt network and transfer learning[J]. Automation of Electric Power Systems, 2023, 47(13): 110-120.
- [12] IWAYEMI A, ZHOU C. SARAA: semi-supervised learning for automated residential appliance annotation[J]. IEEE Transactions on Smart Grid, 2017, 8(2): 779-786.
- [13] 周勇军,吴元香,董智华,等. 基于模体挖掘与调和函数半监督学习的非侵入式负荷监测[J]. 电力自动化设备,2022,42(7): 3-10.  
ZHOU Yongjun, WU Yuanxiang, DONG Zhihua, et al. Non-intrusive load monitoring based on motif mining and harmonic function based semi-supervised learning[J]. Electric Power Automation Equipment, 2022, 42(7): 3-10.
- [14] LIU H, LIU C Y, ZHAO H R, et al. Non-intrusive load monitoring method for multi-energy coupling appliances considering spatio-temporal coupling[J]. IEEE Transactions on Smart Grid, 2023, 14(6): 4519-4529.

- [15] SETTLES B. Active learning literature survey[J]. *Machine Learning*, 2010, 15(2): 201-221.
- [16] GUO L Y, WANG S X, CHEN H W, et al. A load identification method based on active deep learning and discrete wavelet transform[J]. *IEEE Access*, 2020, 8: 113932-113942.
- [17] LIEBGOTT F, YANG B. Active learning with cross-dataset validation in event-based non-intrusive load monitoring[C]//2017 25th European Signal Processing Conference (EUSIPCO). Kos, Greece: IEEE, 2017: 296-300.
- [18] KASELIMI M, PROTOPAPADAKIS E, VOULODIMOS A, et al. Towards trustworthy energy disaggregation: a review of challenges, methods, and perspectives for non-intrusive load monitoring[J]. *Sensors*, 2022, 22(15): 5872.
- [19] LI X, ZHANG W. Deep learning-based partial domain adaptation method on intelligent machinery fault diagnostics[J]. *IEEE Transactions on Industrial Electronics*, 2021, 68(5): 4351-4361.
- [20] 李典阳, 张育杰, 冯健, 等. 变压器故障样本多维诊断及结果可信度分析[J]. *电工技术学报*, 2022, 37(3): 667-675.  
LI Dianyang, ZHANG Yujie, FENG Jian, et al. Multi-dimensional diagnosis of transformer fault sample and credibility analysis [J]. *Transactions of China Electrotechnical Society*, 2022, 37(3): 667-675.
- [21] 周志华. *机器学习*[M]. 北京: 清华大学出版社, 2016: 183-186.
- [22] 周志华. 基于分歧的半监督学习[J]. *自动化学报*, 2013, 39(11): 1871-1878.  
ZHOU Zhihua. Disagreement-based semi-supervised learning [J]. *Acta Automatica Sinica*, 2013, 39(11): 1871-1878.
- [23] WANG W, ZHOU Z H. On multi-view active learning and the combination with semi-supervised learning[C]//Proceedings of the 25th international conference on Machine learning. Helsinki, Finland: ACM, 2008: 1152-1159.
- [24] 李兵, 韩睿, 何怡刚, 等. 改进随机森林算法在电机轴承故障诊断中的应用[J]. *中国电机工程学报*, 2020, 40(4): 1310-1319, 1422.  
LI Bing, HAN Rui, HE Yigang, et al. Applications of the improved random forest algorithm in fault diagnosis of motor bearings[J]. *Proceedings of the CSEE*, 2020, 40(4): 1310-1319, 1422.
- [25] MEDICO R, DE BAETS L, GAO J K, et al. A voltage and current measurement dataset for plug load appliance identification in households[J]. *Scientific Data*, 2020, 7: 49.
- [26] KAHL M, HAQ A U, KRIECHBAUMER T, et al. Whited—a worldwide household and industry transient energy data set [C]//3rd International Workshop on Non-Intrusive Load Monitoring. Vancouver, BC, Canada: Simon Fraser University, 2016: 1-4.
- [27] DE BAETS L, RUYSSINCK J, DEVELDER C, et al. Appliance classification using VI trajectories and convolutional neural networks[J]. *Energy and Buildings*, 2018, 158: 32-36.
- [28] FAUSTINE A, PEREIRA L, KLEMENJAK C. Adaptive weighted recurrence graphs for appliance recognition in non-intrusive load monitoring[J]. *IEEE Transactions on Smart Grid*, 2021, 12(1): 398-406.

#### 作者简介:

蔡清淮(1998—),男,硕士研究生,主要研究方向为基于机器学习的负荷辨识(**E-mail**: caiqh1998@foxmail.com);

余涛(1974—),男,教授,博士,主要研究方向为复杂电力系统的非线性控制理论、优化与机器学习等(**E-mail**: taoyu1@scut.edu.cn)。

(编辑 王锦秀)

## Generalization method for load identification based on active transfer learning

CAI Qinghui<sup>1</sup>, LUO Qingquan<sup>1</sup>, YU Tao<sup>1,2</sup>, LIU Qianjin<sup>1</sup>, LIU Xipeng<sup>1</sup>, PAN Zhenning<sup>1</sup>

(1. School of Electric Power, South China University of Technology, Guangzhou 510640, China;

2. Guangdong Provincial Key Laboratory of Intelligent Measurement and Advanced Metering of Power Grid, Guangzhou 510640, China)

**Abstract:** In order to realize identification accuracy improvement of load identification methods to load samples in new scenarios with low cost, a load identification generalization method based on active transfer learning is proposed, which utilizes very few labelled data and unlabelled data to efficiently improve the generalization performance of the method. The method utilizes the consensus of heterogeneous models to select high-quality pseudo-labelled samples and updates the pre-trained model. An active learning strategy considering model disagreement and sample diversity is designed to label high-value samples, which can realize efficient model transfer while significantly reducing sample labeling cost. Experimental comparison of two public datasets verifies the superiority of the proposed method.

**Key words:** load identification; generalization; active learning; transfer learning; pseudo-labeling; heterogeneous model



附录 A:

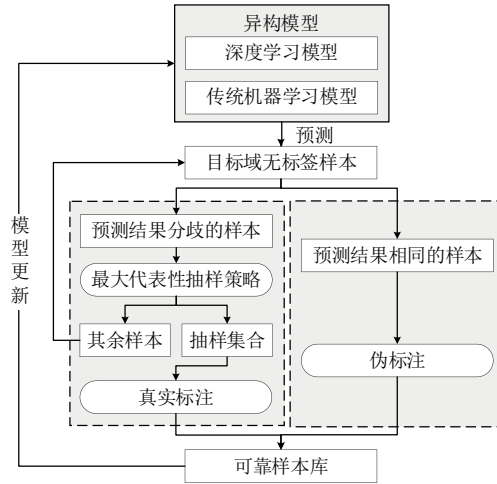


图 A1 主动迁移学习方法流程图  
Fig.A1 Active transfer learning methodology flowchart

表 A1 所提出的主动学习策略  
Table A1 Proposed active learning strategy

**算法:** 结合模型分歧的 MRS 主动学习策略

**输入:** 目标域样本  $X_T$ ; 1D-CNN 模型和 RF 模型对目标域的识别结果  $\hat{y}_1$  和  $\hat{y}_2$ ; 选择样本的数量  $n_q$ ;

**输出:** 目标域抽样集合  $X_{select}$ ;

- 1: 把目标域样本按  $1, 2, \dots, n_T$  进行编号, 编号集合记为  $G$ ;
- 2: 初始化:  $X_{select} = \emptyset, R(X_{select}, X_T) = 0$ ;
- 3: 存储双模型识别结果不一致的编号集合:  $G_d = \{i | \hat{y}_{1i} \neq \hat{y}_{2i} \text{ 且 } i \in G\}$
- 4: MRS 策略:
- 5: for  $k = 1, 2, \dots, n_q$  do
- 6:  $w = \arg \max_{i \in G_d} R(X_{select} \cup \{x_i^T\}, X_T)$
- 7: 将编号为  $w$  的样本加入至选择集合:  $X_{select} = X_{select} \cup \{x_w^T\}$ ;
- 8: 将编号  $w$  从集合  $G_d$  中移除;
- 9: end for
- 10: 输出抽样集合  $X_{select}$  并标注其真实标签。

注:  $\arg \max_{i \in G_d}$  表示函数取得最大值时变量  $i$  的取值。

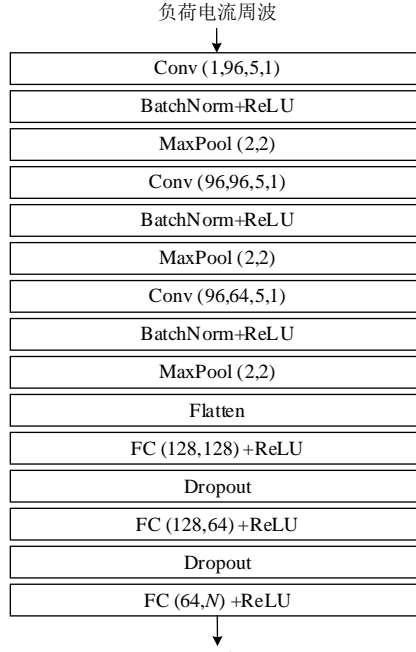


图 A2 1D-CNN 模型结构

Fig.A2 The model structure of 1D-CNN

其中 Conv 的四个参数分别表示输入通道数、输出通道数、卷积核大小及其步长，MaxPool 两个参数分别表示窗口的大小及其步长，FC 两个参数分别表示输入和输出的维度数。

$$I_{acc} = \frac{T_p + T_N}{T_p + F_p + T_N + F_N} \quad (A1)$$

$$F_1 = \frac{2 \times I_{precision} \times I_{recall}}{I_{precision} + I_{recall}} \quad (A2)$$

$$I_{recall} = \frac{T_p}{T_p + F_N} \quad (A3)$$

$$I_{precision} = \frac{T_p}{T_p + F_p} \quad (A4)$$

$$I_{mcc} = \frac{T_p \times T_N - F_p \times F_N}{\sqrt{(T_p + F_p)(T_p + F_N)(T_N + F_p)(T_N + F_N)}} \quad (A5)$$

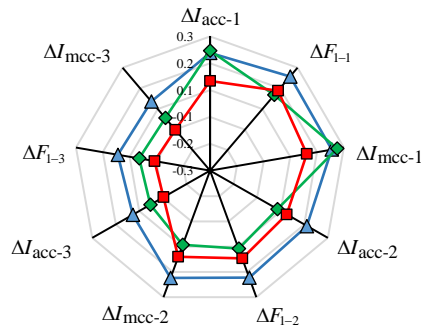
其中  $T_p$  正确识别的正类数量， $T_N$  为正确识别的负类数量， $F_p$  为错误识别的负类数量， $F_N$  为错误识别的正类数量， $I_{recall}$  和  $I_{precision}$  分别表示查全率和查准率。 $I_{acc}$ 、 $F_1$ 、 $I_{mcc}$  取值范围均为  $0 \leq 1$ ，且值越大表明识别效果越好。

## 附录 B:

表 B1 迁移学习方法的泛化性能对比

Table B1 Comparison of generalization performance of transfer learning methods

方法	目标域标注量	数据集泛化			品牌泛化			住宅泛化		
		$I_{acc}$	$F_1$	$I_{mcc}$	$I_{acc}$	$F_1$	$I_{mcc}$	$I_{acc}$	$F_1$	$I_{mcc}$
finetune-AlexNet	5%	0.546	0.419	0.498	0.696	0.660	0.670	0.778	0.709	0.756
finetune-AlexNet	20%	0.811	0.734	0.782	0.857	0.860	0.802	0.844	0.811	0.826
finetune-VGG16	5%	0.461	0.433	0.420	0.667	0.656	0.642	0.728	0.671	0.705
finetune-VGG16	20%	0.802	0.773	0.729	0.812	0.791	0.795	0.868	0.842	0.853
本文方法	5%	0.906	0.865	0.895	0.858	0.860	0.846	0.913	0.899	0.904
本文方法	20%	0.931	0.916	0.937	0.916	0.919	0.909	0.921	0.909	0.914



—▲— 本章方法    —◆— finetune-AlexNet    —■— finetune-VGG16

注：下标 1、2、3 分别对应数据集泛化实验、品牌泛化实验、住宅泛化实验。

图 B1 泛化性能增益雷达图

Fig.B1 Radar chart of generalization performance gains

表 B2 本文方法在不同标注比例下的性能

Table B2 Performance at different labelling ratios

标注比例%	$I_{acc}$	$F_1$	$I_{mcc}$	平均值
0	0.7663	0.6972	0.7398	0.7345
3	0.8479	0.7390	0.8297	0.8055
4	0.8743	0.8026	0.8594	0.8454
5	0.9057	0.8650	0.8947	0.8885
6	0.9130	0.8849	0.9035	0.9005
7	0.9181	0.8904	0.9090	0.9058

## 附录 C:

为验证本文的经验置信向量相比于模型直接输出的类概率向量，可更好地利用 1D-CNN 和 RF 的识别能力进行融合识别，将本文所提出的模型融合策略与基于模型类概率的融合策略进行对比实验，对比方法设置如下：

融合策略 1：第 4 节所述的基于最大类概率的融合策略。具体地，从两个模型输出的类概率向量中，取最大类概率所对应的类别作为融合识别的结果，见式 (11)。

融合策略 2：根据模型类概率之和输出模型融合识别结果。具体地，将两个模型输出的类概率向量相加后，取和最大所对应的类别作为融合识别的结果，见式 (12)。



$$c = \arg \max_b \{h_1^b(\mathbf{x}), h_2^b(\mathbf{x})\} \quad (11)$$

$$c = \arg \max_b (h_1^b(\mathbf{x}) + h_2^b(\mathbf{x})) \quad (12)$$

其中  $h_a^b(\mathbf{x})$  是识别模型  $h_a$  在类别  $b$  上的输出，即类概率，且  $h_a^b(\mathbf{x}) \in [0,1]$ ； $\arg \max_b$  表示函数取得最大值时变量  $b$  的取值。

以上融合策略均采用与本文方法相同的异构模型，并直接进行数据集角度的泛化实验，所有方法识别模型的训练均仅在源域进行，并在目标域数据集上进行测试，实验结果如表 3。

由实验结果可知，基于经验置信向量的融合策略得到的识别结果更加准确，表明大部分情形下，本文经验置信向量相比于模型的类概率，更能表征模型在新目标域识别结果的可信度。进一步深入分析基于模型类概率的融合策略效果不佳的原因，当模型的类型和结构不同时，基于模型类概率进行融合识别易出现如下情形：虽然  $h_1^c(\mathbf{x}) > h_2^c(\mathbf{x})$ ，但是模型 2 预测结果正确的可能性更大，这是因为异构模型的类概率不可直接比较。如在该实验中，1D-CNN 预测结果为冰箱时，其平均类概率为 0.828，但其对冰箱的宏平均 F1 分数仅为 0.18；RF 预测为冰箱时，其平均类概率为 0.600，而其对冰箱的宏平均 F1 分数为 0.56。因此不能直接地将异构模型的类概率对比，近似为预测结果可信度的对比。而本文融合策略中异构模型经验置信度的对比，是基于在多次预设识别任务中所得的经验，相比之下更为可靠。

表 C1 融合策略对比

Table 3 Comparison of fusion methods

方法	$I_{acc}$	$F_1$	$I_{mcc}$
本文方法	0.718	0.657	0.690
融合策略 1	0.660	0.562	0.627
融合策略 2	0.663	0.571	0.632

## 附录 D:

设置以下对照方法进行消融实验:

方法 1: 不结合模型分歧进行主动学习，其余环节与本文方法一致;

方法 2: 在本文方法的基础上将主动学习策略替换为 MS;

方法 3: 在本文方法的基础上将主动学习策略替换为 IES;

方法 4: 去除主动学习环节，其余与本文方法一致;

方法 5: 去除完整的主动迁移学习环节，其余与本文方法一致;

方法 6: 本文方法基础上将无标签数据学习替换为同构模型互学机制<sup>[14]</sup>，并去除主动学习环节。

以上方法采用主动学习时目标域的抽样比例均为 5%，各方法的数据集泛化准确率如图 D1 所示。

根据实验结果，比较本文完整方法和对照方法 1 可以看出，结合模型分歧进行主动学习的方式，有利于帮助模型针对性学习目标域的不确定样本，由于异构模型预测不一致的样本通常在分类边界附近出现，将有限的标注资源用于该部分样本可帮助模型拟合分类边界，从而提高泛化能力；将对照方法 1 与 2、3 比较，可以进一步验证本文 MRS 主动学习策略的优越性；比较对照方法 4 和 5，两者的区别在于对照方法 4 保留了本文基于伪标注样本的迁移学习方法，实验结果验证了本文伪标注样本学习环节实现了无标签数据的有效利用，提升了模型的泛化性能；比较对照方法 4 和 6 可看出，相比于与文献[14]所提出的同构模型互学机制(mutual learning)，本文利用无标签样本辅助训练的方式更具优势，这得益于本文异构模型的设计使得伪标签质量更高，从而更适用于泛化应用场景的识别精度提升。

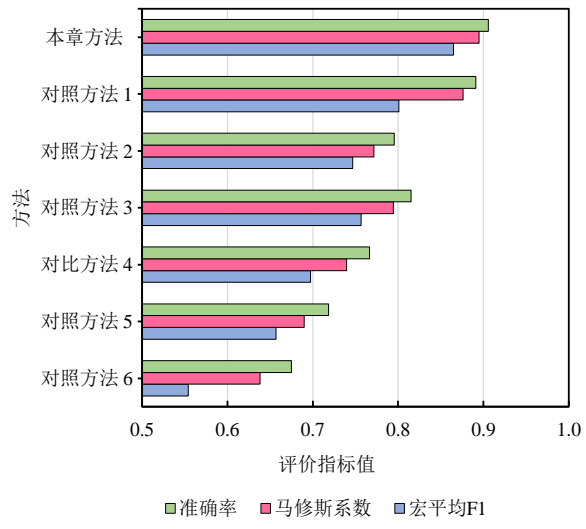


图 D1 消融实验结果

Fig.D1 Ablation experiment results