



# 基于细粒度迁移的跨场景认知能力评估方法

张迎伟<sup>1,2</sup>, 陈益强<sup>1,2,3\*</sup>, 于汉超<sup>4</sup>, 杨晓东<sup>1,2</sup>, 谷洋<sup>1,2,3</sup>

1. 中国科学院计算技术研究所移动终端与新型计算北京市重点实验室, 北京 100190

2. 中国科学院大学, 北京 100049

3. 鹏城实验室, 深圳 518066

4. 中国科学院前沿科学与教育局, 北京 100864

\* 通信作者. E-mail: yqchen@ict.ac.cn

收稿日期: 2021-08-10; 修回日期: 2021-11-10; 接受日期: 2022-02-22; 网络出版日期: 2022-04-12

广东省重点领域研发计划 (批准号: 2019B010109001)、国家自然科学基金 (批准号: 61972383, 61902377, 61902379)、山东省智研院创新课题 (批准号: SDAICT2191010) 和中国科学院青年创新促进会资助

**摘要** 认知健康是大脑健康的重要组成部分, 与个体全生命周期的正常发展密切相关. 目前, 非受限普适场景下的认知能力评估已成为实现认知障碍相关疾病早期预警的重要途径. 然而, 在医疗、家庭等不同计算场景下, 因感知设备、认知评估流程等的不同, 往往存在不同场景间评估模型相互适用难的问题. 针对以上跨场景认知能力评估挑战, 本文提出了一种参数自适应的细粒度迁移学习方法 PAFG-TL. PAFG-TL 基于随机森林模型实现, 通过参数自适应的个体分类器评估策略和领域自适应的决策树生长机制实现参数无关的个体分类器聚类评估和决策树细粒度进化生长. 通过在临床认知评估和公开基准数据集上进行实验验证, 证明了 PAFG-TL 方法在跨场景认知能力评估中的有效性.

**关键词** 参数自适应, 细粒度, 迁移学习, 跨场景, 认知能力评估

## 1 引言

认知健康是一种清晰地思考、学习和记忆的能力, 是大脑健康的重要组成部分. 认知健康关系到个体全生命周期的正常发展, 从认知能力不断发展的青少年阶段, 到认知能力达到最佳的成年阶段, 再到认知能力逐渐退化的老年阶段, 拥有健康认知能力是独立生活、学习技能和适应社会的基础<sup>[1]</sup>. 近年来, 许多与生活习惯、环境等相关的因素, 如睡眠不足、缺乏体育锻炼、吸烟、酗酒、摄入过量高脂肪或高钠食品、人口老龄化、环境变化, 使个体认知健康愈发成为一个严峻的家庭和社会危机. 德国海德堡大学 (Heidelberg University) 医学院的统计数据表明, 注意力缺陷多动症 (attention deficit hyperactivity disorder, ADHD) 在儿童群体的发病率约为 5%~7%, 全球约有 5110 万儿童受到注意力缺陷多动症的影响<sup>[2]</sup>. 此外, 2019 年世界卫生组织 (World Health Organization, WHO) 的统计数据

**引用格式:** 张迎伟, 陈益强, 于汉超, 等. 基于细粒度迁移的跨场景认知能力评估方法. 中国科学: 信息科学, 2023, 53: 647-662, doi: 10.1360/SSI-2021-0269  
Zhang Y W, Chen Y Q, Yu H C, et al. A fine-grained transfer learning method for cross-scenario cognitive-ability assessment (in Chinese). Sci Sin Inform, 2023, 53: 647-662, doi: 10.1360/SSI-2021-0269

表明, 全球约有 5000 万老年人存在认知能力衰退症状, 在 60 岁以上的老年群体中, 痴呆发病率高达 5%~8% [3]. 目前, 针对注意力缺陷多动症、老年痴呆等认知健康相关疾病尚无有效治疗方案. 关注发病早期阶段的训练与康复、缓解特异性症状、延缓病情发展, 是目前唯一可行的应对方法 [4].

普适计算 (ubiquitous computing) 是适应计算技术发展出现的一种新型计算模式, 强调融合信息、物理和社会空间, 通过无所不在的计算服务突破传统计算模式的时空限制 [5]. 区别于传统受限环境下的认知健康监护, 普适非受限环境下的认知健康监护能够通过可穿戴设备、自然人机交互和自适应计算实现无约束、长时间、跨场景认知健康监护 [6]. 然而, 在医疗问诊、社区生活、居家监护等不同普适计算场景下, 由于数据感知设备、感知流程的不同, 往往带来不同场景下数据分布不一致的问题, 使已有模型难以跨场景适配. 迁移学习放宽了传统机器学习中的数据分布一致性要求, 能够通过模拟人类类比推理过程, 将在已知领域上学习到的知识用于其他相关领域的推理分析上 [7], 为解决普适环境下认知能力评估的跨场景适用提供了新思路 [8]. 如, Long 等 [9] 融合特征迁移和深度神经网络, 实现了计算机视觉模型的跨域迁移. 尽管迁移学习在计算机视觉领域取得了优异的实验效果, 但受限于医疗问诊的可解释性需求强和数据的小样本等问题, 传统迁移学习方法在医疗健康领域的应用还存在一定挑战. 相较于计算机视觉领域常用数据集, 认知能力评估相关研究由于患者数量有限、标注困难、数据采集成本高昂等因素影响, 相关数据集往往规模更小, 通常只包含十余个受试者, 已有的研究往往受限于小样本问题, 难以实现高精度的认知能力评估模型构建. 传统迁移学习方法多基于深度神经网络设计, 而在如此小规模样本集合中, 难以使用梯度下降算法优化深度神经网络模型. 另外, 在医疗相关的认知评估模型构建过程中, 模型的可解释性也至关重要, 可解释能力是进一步发掘与健康相关诊疗指标的基础. 因此, 本文尝试将可解释性较强的随机森林模型与迁移学习思想相结合, 建立有效的自适应认知评估模型.

基于已有细粒度迁移随机森林相关研究工作 [10], 本文进一步研究了参数自适应的细粒度迁移学习方法 (parameter adaptive fine-grained transfer learning, PAFG-TL). PAFG-TL 同样基于随机森林实现, 由参数自适应的个体分类器评估策略 (parameter adaptation based model evaluation strategy, PAME) 和领域自适应的决策树生长机制 (domain adaptive tree growing mechanism, DATG) 两部分构成. 其中, PAME 基于决策树特征集合和阈值适应度, 以聚类方式对不同个体分类器进行聚类划分; DATG 通过不同策略实现决策树的细粒度进化生长. 本文在临床认知评估和公开基准数据集上分别进行了实验验证, 通过与已有 state of the art (SOTA) 方法的对比, 验证了 PAFG-TL 方法的有效性. 相较于文献 [10] 和已有其他研究者的工作, 本文的贡献主要包括以下 3 个方面: (1) 提出了一种基于随机森林的模型细粒度迁移策略, 能够适用于数据小样本、可解释性需求强的医学诊疗问题, 实现居家监护、医疗问诊等不同场景下认知评估模型的自适应, 解决了普适环境下认知评估模型跨场景适应难的问题; (2) 显示定义了特征集合适应度和分割阈值适应度概念, 能够通过二维坐标平面直观表示各个个体分类器的适应度; (3) 改进了文献 [10] 中基于信息增益的模型评估策略, 无需依赖全局最优阈值筛选和设置, 即可通过决策树特征集合和阈值适应度信息, 实现自适应的个体分类器分组划分.

本文的组织结构如下: 第 2 节介绍与本文相关的已有研究工作; 第 3 节对本文所用公式进行形式化定义; 第 4 节详细介绍本文提出的参数自适应细粒度迁移学习方法; 第 5 节对本文提出的算法进行了实验验证与评估; 第 6 节对本文进行了讨论、总结与展望.

## 2 相关工作

迁移学习的目标是构建一种具有良好泛化能力的识别模型, 使得能够将在一个场景中学习到的知

识用于具有不同程度相关度其他场景下的识别问题中. 在医疗健康领域, 可获取数据小样本、跨场景适应是智能医疗面临的重要挑战, 迁移学习通过利用来自相似领域的数据来解决给定领域中的数据不足, 近年来展现出了极好的科研潜力和实用前景. 广州医科大学的 Zhang 等<sup>[11]</sup> 将迁移学习应用于医学诊断, 利用光学相干断层成像术 (optical coherence tomography, OCT) 对患者眼底图像数据进行采集, 构建了来自 4686 例患者包含的 108312 张 OCT 图像的训练数据集, 其构建的识别模型在多分类任务上的分类准确率能够媲美人类专家, 达到了 96.6% 的分类识别精度; 另外, Zhang 等还将其所设计的分类模型迁移至儿童肺炎诊断任务, 通过胸部 X 光图像识别儿童患有肺炎的概率, 也取得了良好的识别效果. 为解决医疗数据的小样本和非均衡等问题, Khatami 等<sup>[12]</sup> 提出了一种基于分布检索的方式来分析医学图像、进而构建高精度的医疗诊断模型, 所提出的方法利用基于迁移学习的卷积神经网络和基于随机森林的选择池组成减少样本搜索空间, 构建描述样本的特征向量, 在包含 14400 张 X 光图像的公共数据集上的实验结果表明, 所提出的方法能够达到 90.30% 的分类准确率. 针对医疗数据的小样本问题, Yu 等<sup>[13]</sup> 探索了基于深度神经网络微调的模型适应方法, 他们使用 ImageNet 数据集预训练两种基于 VGGNet 和 ResNet 网络的深度卷积模型, 并利用包含 4532 张图像的 ImageCLEF2015 数据集和包含 6776 张图像的 ImageCLEF2016 数据集对预训练的网络模型进行微调, 取得了 76.87% 和 87.37% 的分类准确率. 针对同一领域不同任务间的模型适应性问题, Gao 等<sup>[14]</sup> 设计了一种跨任务迁移框架, 该学习框架能够通过特征迁移实现乳腺癌 X 光图像的分类、检测和分割, 利用低分辨率图像提升模型分类准确率. 针对乳腺癌诊断, Gao 等在分类、检测和分割 3 种任务上, 分别与已有的 8 种分类方法、4 种检测方法和 3 种分割进行实验对比, 验证了所提出的特征迁移学习框架的有效性. 同样地, 为了应对医疗健康数据的小样本挑战, 弥补人类专家感知到的高级信息与核磁共振成像设备捕捉到的低级视觉信息之间的信息差距, Swati 等<sup>[15]</sup> 将迁移学习和卷积神经网络结合起来, 提出了一种块级微调策略. 在磁共振造影增强基准数据集上的对比实验表明, Swati 等提出方法的平均精度达到 94.82%, 优于其他已有的方法. 为构建高精度的横纹肌肉瘤针对模型, Banerjee 等<sup>[16]</sup> 构建了包含 5 个卷积层的 AlexNet 深度学习网络, 并在通过对网络模型深层特征的调整, 使已有模型能够更好地适应横纹肌肉瘤分割任务. 实验结果表明, 通过交叉验证, 该方法的诊断准确率可达 85%, 能够在需要更少人工交互的情况下, 实现快速、高效、可重复的疾病诊断.

然而, 区别于普适环境下的行为-认知关联分析问题, 现有面向医疗健康监护的领域自适应方法大多针对由 OCT 和 X 光图像等医学图像数据构成的数据集设计. 一般而言, 此类型数据集虽难以达到 ImageNet 等通用图片识别数据集的千万张数据规模, 但数据量也不会太小, 仍有数千张. 而由于数据获取的困难, 认知评估相关数据集一般规模十分小, 仅包含几十名受试者. 针对跨场景认知评估任务上的数据分布不一致挑战, 由于数据规模更小、一般由信号类数据构成等问题的影响, 目前相关的研究还十分缺乏. 本文拟尝试结合随机森林与迁移学习模型, 构建适用于小数据、可解释性较强的迁移学习模型. 基于上述思想, 目前已有部分相关研究, Sukhija 等<sup>[17]</sup> 针对源域与目标域数据的特征异构问题, 提出了一种基于随机森林的迁移学习方法, 但该方法需要使用跨域的共享标签分布以实现异构特征变换. Wen 等<sup>[18]</sup> 在目标域数据增量可及场景下, 利用源域数据提升目标域分类效果, 该方法中基础分类模型基于决策树构建, 但需使用多个源域数据以提升迁移效率. Ryu 等<sup>[19]</sup> 基于卷积神经网络训练随机森林模型以解决源域与目标域的数据分布不一致和神经网络模型的易过拟合问题, 但需较大数据以通过梯度下降方法调优模型结构. Segev 等<sup>[20]</sup> 将随机森林与迁移学习结合, 提出了结构优化和阈值更新两种随机森林增量更新方法以实现不同域数据之间的迁移, 但该方法并未考虑不同决策树迁移结果的个性化差异. 我们已有研究工作<sup>[10]</sup> 进一步提出了个体分类器评估策略, 为不同决策树制定个性化迁移策略, 但该方法具有较强的参数依赖性. 因此, 基于<sup>[10]</sup> 已有研究成果, 本文进一步提

出了一种参数无关的跨场景认知能力评估方法.

### 3 形式化

**迁移学习.** 迁移学习利用在源域学习的知识解决目标域的数据预测问题, 通常, 源域数据可表示为  $\mathcal{D}^S = \{(\mathbf{x}_1^S, \mathbf{y}_1^S), (\mathbf{x}_2^S, \mathbf{y}_2^S), \dots, (\mathbf{x}_{n_S}^S, \mathbf{y}_{n_S}^S)\}$ , 其中  $n_S = |\mathcal{D}^S|$  为源域的数据规模; 目标域数据可表示为  $\mathcal{D}^T = \{(\mathbf{x}_1^T, \mathbf{y}_1^T), (\mathbf{x}_2^T, \mathbf{y}_2^T), \dots, (\mathbf{x}_{n_T}^T, \mathbf{y}_{n_T}^T)\}$ , 其中  $n_T = |\mathcal{D}^T|$  为目标域的数据规模.  $\mathcal{X} \in \mathbb{R}^K$  为特征空间集合,  $\mathcal{Y} \in \mathbb{R}^L$  为类别空间集合,  $K$  和  $L$  分别为预测问题中特征数目和类别数目.

**随机森林.** 随机森林旨在学习包含  $M$  个决策树的集成分类器  $\mathcal{R} = \{h_1, h_2, \dots, h_M\}$ ,  $\mathcal{D}^S = \{(\mathbf{x}_1^S, \mathbf{y}_1^S), (\mathbf{x}_2^S, \mathbf{y}_2^S), \dots, (\mathbf{x}_{n_S}^S, \mathbf{y}_{n_S}^S)\}$  为训练数据集. 样本  $\mathbf{x}_i^S$  在个体分类器  $h_m$  上的预测结果可表示为  $(h_m^1(\mathbf{x}_i^S), h_m^2(\mathbf{x}_i^S), \dots, h_m^L(\mathbf{x}_i^S))$ , 其中  $h_m^l(\mathbf{x}_i^S)$  是个体分类器  $h_m$  在类别  $\mathcal{Y}_l$  上的预测结果. 在第  $m$  个个体分类器构建过程中, 随机森林使用 bootstrap 数据采样方法<sup>[21]</sup>, 从原始训练样本集合  $\mathcal{D}^S$  中筛选包含  $n_S$  个样本的训练子集  $\mathcal{D}_m^S$ ; 在选择分割属性时, 随机森林使用特征随机化方法从  $K$  个特征中选择  $k$  个特征作为候选特征, 之后根据信息增益等分割指标从候选特征集中选择最佳特征.  $\mathcal{R}$  的最终预测结果可通过平均法、投票法、学习法等不同模型融合策略计算, 以投票法为例, 最终计算结果为:  $\mathcal{H}(\mathbf{x}_i^S) = \mathcal{Y}_{\arg \max_i} \sum_{m=1}^M w_m h_m^l(\mathbf{x}_i^S)$ .

## 4 参数自适应的细粒度迁移学习方法

### 4.1 总体思路

参数自适应的细粒度迁移学习 (PAFG-TL) 方法是一种模型迁移方法, 使用小规模目标域数据通过有监督的方式实现源域到目标域的模型的适应. PAFG-TL 方法的总体框架图如图 1 所示, 由参数自适应的个体分类器评估策略 (PAME) 和领域自适应的决策树生长机制 (DATG) 两部分构成, 分别用于评估个体分类器和对个体分类器进行进化更新.

PAFG-TL 基于随机森林实现, 由多个决策树集成而成. 根据 Krogh 等<sup>[22]</sup> 的误差 - 分歧分解 (error-ambiguity decomposition) 理论, 集成学习模型的泛化误差可表示为:  $\text{Err} = \overline{\text{err}} - \bar{A}$ . 其中,  $\text{Err}$ ,  $\overline{\text{err}}$  和  $\bar{A}$  分别表示集成学习的模型误差、个体分类器的平均泛化误差和个体分类器的“分歧”程度, 集成学习模型的整体性能由个体分类器精准率和多样性两者决定, 提升目标域模型性能的关键是改善个体分类器在目标域上的适应度并增大个体分类器的多样性. 因此, PAFG-TL 方法的基本思路是分别评估各个个体分类器, 根据已有源域模型与目标域数据的匹配程度, 个性化制定不同的生长策略, 保证个体分类器的个性化进化更新并提升整体集成学习模型的多样性.

### 4.2 参数自适应的个体分类器评估策略

个体决策树呈二叉或多叉树形结构, 通过每个节点上的属性类别和分割阈值决定决策路径. 因而, 分割阈值的确定和特征集合的选择是衡量源域模型在目标域数据中适应度的关键, 本文使用源域 - 目标域构建特征的信息增益差和目标域数据的测试准确率分别衡量特征集合适应度和分割阈值适应度.

信息增益是衡量决策树中节点分割质量的常用指标:  $\text{IG}(\mathcal{D}, a) = \text{Ent}(\mathcal{D}) - \text{Ent}(\mathcal{D}|a)$ , 其中,  $a$  为分割属性,  $\text{Ent}(\mathcal{D}) = -\sum_{l=1}^L p_l \log_2 p_l$  为当前数据集  $\mathcal{D}$  上的信息熵,  $p_l$  为数据集  $\mathcal{D}$  上第  $l$  类样本的概率,  $\text{Ent}(\mathcal{D}|a) = \sum_{v=1}^V \frac{|\mathcal{D}^v|}{|\mathcal{D}|} \text{Ent}(\mathcal{D}^v)$  是当使用属性  $a$  划分数据集  $\mathcal{D}$  为  $V$  个子集合时的条件熵. 所有特征的信息增益可构成信息增益集合  $\text{IG}(\mathcal{D}, \mathcal{X})$ , 为评估各个特征的信息增益值, 本文对  $\text{IG}(\mathcal{D}, \mathcal{X})$  进行排

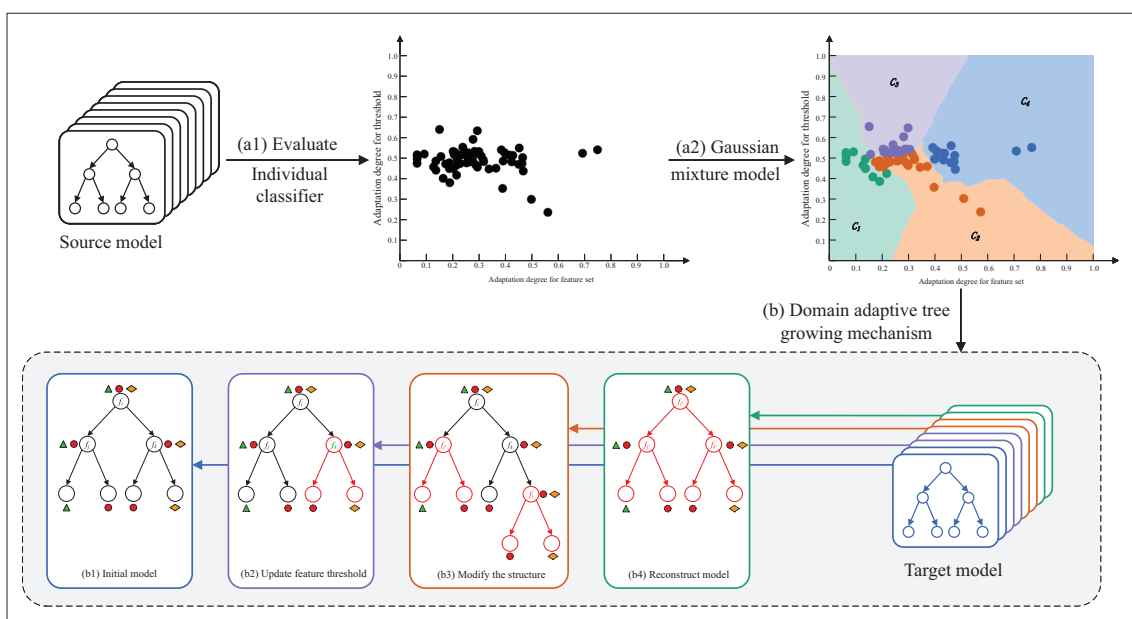


图 1 (网络版彩图) 参数自适应的细粒度迁移学习方法框架图

Figure 1 (Color online) The framework of PAFG-TL method

序, 并使用  $\text{IGR}(\mathcal{D}, \mathcal{X})$  表示各个特征信息增益的排序结果. 例如, 若  $\text{IG}(\mathcal{D}, \mathcal{X})$  为  $\{2.3, 6.7, 4.5, 1.2\}$ , 则  $\text{IGR}(\mathcal{D}, \mathcal{X})$  为  $\{3, 1, 2, 4\}$ . 假设  $A_m$  为构造第  $m$  棵个体决策树时的特征集合,  $\text{IG}(\mathcal{D}^S, A_m)$  为源域上的信息增益集合,  $\text{IG}(\mathcal{D}^T, A_m)$  为目标域集合上的信息增益集合,  $\text{IGR}(\mathcal{D}^S, A_m)$  和  $\text{IGR}(\mathcal{D}^T, A_m)$  分别是信息增益的排序结果. 为评估各个个体决策树的特征集合适度, 本文定义了如下指标:

$$F(h_m) = \sum_{a \in A_m} \log_K (|\text{IGR}(\mathcal{D}^S, a) - \text{IGR}(\mathcal{D}^T, a)| + 1), \quad (1)$$

其中,  $|\text{IGR}(\mathcal{D}^S, a) - \text{IGR}(\mathcal{D}^T, a)|$  用于衡量属性  $a$  在源域和目标域上的重要程度差异,  $+1$  用于避免  $\text{IGR}(\mathcal{D}^S, a) = \text{IGR}(\mathcal{D}^T, a)$  时出现的差值为零情况. 另外, 本文使用  $|A_m|$  归一化  $F(h_m)$  数值, 以将  $F(h_m)$  取值范围限制在  $[0, 1]$  之间, 最终特征集合适度指标定义为

$$F(h_m) = 1 - \frac{\sum_{a \in A_m} \log_K (|\text{IGR}(\mathcal{D}^S, a) - \text{IGR}(\mathcal{D}^T, a)| + 1)}{|A_m|}. \quad (2)$$

$F(h_m)$  数值越接近于 1, 表示源域模型  $h_m$  对目标域数据的特征集合适度越高; 否则, 数值越接近于 0, 表示源域模型  $h_m$  对目标域数据的特征集合适度越低. 本文使用目标域数据的测试准确率衡量分割阈值适应度, 定义为

$$T(h_m) = \frac{\sum_{(\mathbf{x}_i^T, \mathbf{y}_i^T) \in \mathcal{D}^T} \text{sgn}(h_m(\mathbf{x}_i^T), \mathbf{y}_i^T)}{|\mathcal{D}^T|}, \quad (3)$$

其中,  $\text{sgn}(h_m(\mathbf{x}_i^T), \mathbf{y}_i^T)$  是符号函数,  $|\mathcal{D}^T|$  用于将个体决策树  $h_m$  的识别准确率归一化在  $[0, 1]$  之间.  $T(h_m)$  数值越接近于 1, 表示源域模型  $h_m$  对目标域数据的分割阈值适应度越高; 否则, 数值越接近于 0, 表示源域模型  $h_m$  对目标域数据的分割阈值适应度越低.

完成特征集合适度和分割阈值适应度计算后, 可将  $M$  个个体分类器描述为二维集合  $\mathcal{S} = \{\mathbf{x}'_m = (F(h_m), T(h_m)) | m = 1, 2, \dots, M\}$ , 如图 1(a1) 所示. 为实现参数自适应的个体分类器分析, 本文通过

融合先验知识的高斯 (Gaussian) 混合聚类方法对各个个体分类器进行聚类划分. 这里, 所谓的先验知识表示对各个聚类簇的分布位置预估. 具体而言, 本文将所有个体分类器划分为 4 个不同类别, 并根据各个类别个体分类器的特征集合适应度和分割阈值适应度, 为其制定不同的领域自适应决策树生长机制. 则按照特征集合适应度和分割阈值适应度的数值, 可对 4 种类别分类器的大体位置进行如下预估, 如图 1(a2) 所示. (1) 高特征集合适应度、高分割阈值适应度类别  $C_4$ : 源域个体决策树对目标域数据适应度较好, 可通过不做更改或做较少更改, 实现源域模型对目标域数据适应; (2) 高特征集合适应度、低分割阈值适应度类别  $C_3$ : 源域个体决策树的特征选择基本合理, 但仍需调整阈值设置, 以实现源域模型对目标域数据适应; (3) 低特征集合适应度、高分割阈值适应度类别  $C_2$ : 源域个体决策树在目标域数据上能够取得较好分类准确率, 但特征选择不够合理, 需调整决策树结构, 以实现源域模型对目标域数据适应; (4) 低特征集合适应度、低分割阈值适应度类别  $C_1$ : 源域个体决策树对目标域数据适应度较差, 需通过较大幅度更改或重建, 实现源域模型对目标域数据适应.

为实现高斯混合聚类, 定义包含  $F$  个混合成分的高斯混合分布:  $p_G(\mathbf{x}) = \sum_{f=1}^F \alpha_f \cdot p(\mathbf{x}|\mu_f, \Sigma_f)$ , 其中  $\mu_f$  和  $\Sigma_f$  分别表示第  $f$  个高斯分布的参数,  $\alpha_f$  表示混合系数且满足  $\sum_{f=1}^F \alpha_f = 1$ . 若训练集  $\mathcal{D} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$  由高斯混合分布生成, 令随机变量  $z_i \in \{1, 2, \dots, F\}$  表示生成样本  $\mathbf{x}_i$  的高斯混合成分,  $f_i$  的先验概率  $P(z_i = f)$  为  $\alpha_f$ . 根据贝叶斯 (Bayes) 定理,  $z_i$  的后验分布为

$$\gamma_{if} = p_G(z_i = f|\mathbf{x}_i) = \frac{P(z_i = f) \times p_G(\mathbf{x}_i|z_i = f)}{p_G(\mathbf{x}_i)} = \frac{\alpha_f \times p(\mathbf{x}_i|\mu_f, \Sigma_f)}{\sum_{f'=1}^F \alpha_{f'} \times p(\mathbf{x}_i|\mu_{f'}, \Sigma_{f'})}. \quad (4)$$

另外, 当高斯分布已知时, 高斯混合聚类将样本集  $\mathcal{D}$  划分为  $F$  个聚类簇  $\mathcal{C} = \{C_1, C_2, \dots, C_F\}$ , 样本  $\mathbf{x}_i$  的簇标记可表示  $\lambda_i$  为

$$\lambda_i = \arg \max_{f \in \{1, 2, \dots, F\}} \gamma_{if}. \quad (5)$$

融合先验知识的高斯混合聚类 (prior knowledge based Gaussian mixture model, PGMM) 过程如算法 1 所示, 本文设置  $F = 4$ , 将  $M$  个个体分类器划分为 4 个聚类簇, 并定义各个聚类簇的高斯均值分布在  $\mu_f \in \mathbb{R}_f$  范围内, 其中  $\mathbb{R}_1 = \mathbb{R}_{[0,0.2] \times [0,0.2]}^2$ ,  $\mathbb{R}_2 = \mathbb{R}_{[0.3,0.8] \times [0.2,0.5]}^2$ ,  $\mathbb{R}_3 = \mathbb{R}_{[0.2,0.5] \times [0.3,0.8]}^2$ ,  $\mathbb{R}_4 = \mathbb{R}_{[0.5,0.8] \times [0.5,0.8]}^2$ . 高斯混合聚类过程中, 本文将对  $C_1, C_2, C_3, C_4$  4 个聚类簇的预估先验, 融入聚类分析过程. 在算法 1 的第 1 行将高斯分布均值初始化范围限制在  $\mathbb{R}_f$  范围内, 并在第 8 行根据  $\mathbb{R}_f$  进一步调整聚类中心位置.

已有工作 [10] 将特征集合重要度差异和目标域测试准确度融合, 定义了  $S(h_m)$  指标, 其中  $S(h_m)$  可近似形式化表达为  $S(h_m) = F(h_m) + \lambda \times T(h_m)$ ,  $\lambda$  为权重系数. 然而, 该工作 [10] 并未显示定义特征集合适应度和分割阈值适应度, 仅给出了融合的个体分类器评价指标  $S(h_m)$ , 无法以如图 1(a1) 所示形象化表示各个个体分类器的适应度情况; 另外, 工作 [10] 中各个个体分类器的更新策略选择过程需依赖分割阈值  $\delta_1, \delta_2, \delta_3$ , 需事先进行全局最优阈值筛选和设置. 显式的特征集合适应度和分割阈值适应度定义, 以及参数自适应的个体分类器聚类划分, 是本文相对于工作 [10] 的主要改进.

### 4.3 领域自适应的决策树生长机制

#### 4.3.1 更新阈值

更新阈值 (update feature threshold, UFT) 过程使用散度指标微调已有随机森林模型的特征阈值, 该指标定义如下:

$$DI(\mathcal{D}_v, a(v), \tau(v), Q_L, Q_R)$$

**Algorithm 1** PGMM

---

**Require:** The adaption set  $\mathcal{S} = \{\mathbf{x}'_m = [F(h_m), T(h_m)] | m = 1, 2, \dots, M\}$ , the cluster number  $F$ , the prior space for every cluster  $\mathbb{R}_f, f = 1, 2, \dots, F$ , and the epoch of iteration epoch;

**Ensure:** Cluster result  $\mathcal{C} = \{\mathcal{C}_1, \mathcal{C}_2, \dots, \mathcal{C}_F\}$ ;

- 1: Initialize the parameter for gaussian mixture models  $\{(\alpha_f, \mu_f, \Sigma_f) | 0 \leq f \leq F, \mu_f \in \mathbb{R}_f\}$ ;
- 2: **while** epoch  $> 0$  **do**
- 3:   **for**  $m = 1, 2, \dots, M$  **do**
- 4:     Calculate the posterior probability of  $x_m$  generated by each mixture according to Eq. (4);
- 5:   **end for**
- 6:   **for**  $f = 1, \dots, F$  **do**
- 7:     Calculate the new mean vector  $\mu'_f = \frac{\sum_{m=1}^M \gamma_{mf} \mathbf{x}'_m}{\sum_{m=1}^M \gamma_{mf}}$ ;
- 8:     Adjust the position of  $\mu'_f$  according to  $\mathbb{R}_f$ ;
- 9:     Calculate the new covariance matrix  $\Sigma'_f = \frac{\sum_{m=1}^M \gamma_{mf} (\mathbf{x}'_m - \mu'_f)(\mathbf{x}'_m - \mu'_f)^T}{\sum_{m=1}^M \gamma_{mf}}$ ;
- 10:     Calculate the new mixture coefficient  $\alpha'_m = \frac{\sum_{m=1}^M \gamma_{mf}}{M}$ ;
- 11:   **end for**
- 12:   Update model parameter with  $\{(\alpha_f, \mu_f, \Sigma_i) | 1 \leq f \leq F\}$  with  $\{(\alpha'_f, \mu'_f, \Sigma'_f) | 1 \leq f \leq F\}$ ;
- 13:   epoch = epoch - 1;
- 14: **end while**
- 15:  $\mathcal{C}_f = \emptyset, 1 \leq f \leq F$ ;
- 16: **for**  $m = 1, 2, \dots, M$  **do**
- 17:   Determine the cluster  $\lambda_m$  of sample  $\mathbf{x}'_m$  according to Eq. (5);
- 18:   Divide  $\mathbf{x}'_m$  to corresponding cluster:  $\mathcal{C}_{\lambda_m} \cup \mathbf{x}'_m$ ;
- 19: **end for**

---

$$= 1 - \frac{|\mathcal{D}_L, \tau|}{|\mathcal{D}_v|} \text{JSD}(Q'_L, Q_L) - \frac{|\mathcal{D}_R, \tau|}{|\mathcal{D}_v|} \text{JSD}(Q'_R, Q_R), \quad (6)$$

$$\text{JSD}(P, Q) = (\text{KL}(P||M) + \text{KL}(Q||M))/2, \quad (7)$$

其中  $\mathcal{D}_v$  是可到达节点  $v$  的数据集,  $a(v)$  是节点  $v$  的分割属性,  $\tau(v)$  是节点  $v$  的现有分割阈值.  $Q_L$  和  $Q_R$  是当阈值为  $\tau(v)$  时数据集  $\mathcal{D}_v$  的标签分布,  $Q'_L$  和  $Q'_R$  是当使用新的阈值时数据集  $\mathcal{D}_v$  的标签分布.  $M = (P + Q)/2$  是分布  $P$  和  $Q$  的均值, Jensen-Shannon 散度 (Jensen-Shannon divergence, JSD) 是 Kullback-Leibler 散度 (Kullback-Leibler divergence, KL) 的对称扩展, 能够衡量两个分布的距离并计算出距离度量值. 特别地,  $\text{JSD}(Q'_L, Q_L)$  和  $\text{JSD}(Q'_R, Q_R)$  分别表示当节点  $v$  使用已有阈值  $\tau(v)$  和候选阈值时, 其左子树和右子树的 JSD 距离. 因此, 在考虑目标训练数据的情况下, 散度值可以优化节点  $v$  的特征阈值.

### 4.3.2 结构调整

结构调整 (modify the structure, MTS) 包括结构的扩展和缩减两部分, 能够通过减少无用的内部节点和分裂不纯的叶子节点更新已有模型 [20]. 结构调整为内部节点  $v$  定义了两种损失度量指标, 包括子树损失和叶子损失, 分别评估以  $v$  为根节点的子树的损失值和将节点  $v$  裁剪为叶子节点时的损失值, 子树损失大于叶子损失的节点或数据不可到达的节点将会被裁剪.

## 4.4 参数自适应的细粒度迁移随机森林

参数自适应的细粒度迁移随机森林实现过程如算法 2 所示, 其输入为源域的已有集成学习模型  $\mathcal{R}$ 、源域的信息增益集合  $\text{IG}(\mathcal{D}^S, \mathcal{X})$  和目标域数据  $\mathcal{D}^T$ ; 输出为更新后的集成学习模型. PAFG-TL 首先计

算目标域的信息增益集合  $IG(\mathcal{D}^T, \mathcal{X})$ , 并对  $IG(\mathcal{D}^S, \mathcal{X})$  和  $IG(\mathcal{D}^T, \mathcal{X})$  分别进行排序, 获取  $IGR(\mathcal{D}^S, \mathcal{X})$  和  $IGR(\mathcal{D}^T, \mathcal{X})$ ; 之后, 根据式 (2) 和 (3) 计算二维集合  $\mathcal{S}$ , 用于从特征集合适应度和分割阈值适应度两个方面描述各个个体分类器; 最后, 通过融合先验知识的高斯混合聚类方法对个体分类器进行聚类分析, 并使用不同更新策略对不同集合中个体分类器进行增量更新.

---

**Algorithm 2** PAFG-TL
 

---

**Require:** Ensemble classifier  $\mathcal{R}$ , information gain of source domain  $IG(\mathcal{D}^S, \mathcal{X})$ , and source domain  $\mathcal{D}^T$ ;

**Ensure:** Ensemble classifier  $\mathcal{R}$ ;

```

1: Calculate  $IG(\mathcal{D}^T, \mathcal{X})$  for target domain;
2: Calculate  $IGR(\mathcal{D}^S, \mathcal{X})$  and  $IGR(\mathcal{D}^T, \mathcal{X})$ ;
3: for each individual tree  $h_m \in \mathcal{R}$  do
4:   Calculate  $\mathcal{S}_m[1] = F(h_m)$  according to Eq. (2);
5:   Calculate  $\mathcal{S}_m[2] = T(h_m)$  according to Eq. (3);
6: end for
7:  $K = 4$ ;
8:  $\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3, \mathcal{C}_4 = \text{PGMM}(\mathcal{S}, K, \{\mathbb{R}_k | k = 1, 2, \dots, K\})$ ;
9: for each individual tree  $h_m$  in cluster  $\mathcal{C}_1$  do
10:   $h'_m = \text{buildTree}(\mathcal{D}^T)$ ;
11:  Update  $h_m$  with  $h'_m$ ;
12: end for
13: for each individual tree  $h_m$  in cluster  $\mathcal{C}_2$  do
14:   $h'_m = \text{UFT}(h_m, \mathcal{D}^T)$ ;
15:  Update  $h_m$  with  $h'_m$ ;
16: end for
17: for each individual tree  $h_m$  in cluster  $\mathcal{C}_3$  do
18:   $h'_m = \text{MTS}(h_m, \mathcal{D}^T)$ ;
19:  Update  $h_m$  with  $h'_m$ ;
20: end for
21:  $\mathcal{R} = \mathcal{C}_1 \cup \mathcal{C}_2 \cup \mathcal{C}_3 \cup \mathcal{C}_4$ .

```

---

## 5 实验评估

### 5.1 数据集和预处理

**临床认知评估数据集.** 临床认知评估数据集使用基于触摸屏的认知评估测试衡量老年人的认知状态, 该测试由 4 种不同的单任务和 12 种不同的双任务构成<sup>[10]</sup>. 实验过程中, 在两种场景下分别进行数据采集. (1) 场景一 (标记为 TL): 61 名受试者参与数据采集, 包括 20 名轻度认知功能障碍患者 (年龄  $68.25 \pm 6.15$ , 8 名男性和 12 名女性) 和 41 名健康老年人 (年龄  $67.36 \pm 4.76$ , 21 名男性和 20 名女性). 测试系统运行在 Huawei M5 平板电脑上 (屏幕尺寸 10.1 英寸, 分辨率  $1920 \times 1200$ ); 数据采集过程中, 所有受试者均执行 4 种评估任务, 分别为单任务 I、单任务 II、单任务 III 和单任务 IV. (2) 场景二 (标记为 VS): 37 名受试者参与数据采集, 其中包括 25 名轻度认知功能障碍患者 (年龄  $65.08 \pm 9.68$ , 16 名男性和 9 名女性) 和 12 名健康人 (年龄  $39.44 \pm 2.31$ , 7 名男性和 5 名女性). 测试系统运行在 NanoPi M4 单片机上 (屏幕尺寸 21.5 英寸, 分辨率  $1920 \times 1200$ ); 数据采集过程中, 所有受试者均执行 4 种评估任务, 包括单任务 II、双任务 AII、双任务 BII 和双任务 CII. 针对临床认知评估数据集, 本文提取了 5 类特征, 包括基于数量的特征 ( $\in \mathbb{R}^{10}$ )、基于时间的特征 ( $\in \mathbb{R}^{50}$ )、基于速度的特征



表 1 4 种公开基准数据集

Table 1 Information for four public-available benchmark datasets

Dataset	#Subject	#Activity	#Sample	Sensor location	Sensor type
DSADS	8	19	1.14M	Torso, right arm, left arm, right leg, left leg	Acceleration, gyroscope, magnetometer
UCI-HAR	30	6	1.31M	Waist	Acceleration, gyroscope
USC-HAD	14	12	2.81M	Front right hip	Acceleration, gyroscope
PAMAP	9	18	2.84M	Wrist, chest, ankle	Temperature, acceleration, gyroscope, magnetometer, orientation

( $\in \mathbb{R}^{130}$ )、基于角度的特征 ( $\in \mathbb{R}^{60}$ ) 和基于落点分布的特征 ( $\in \mathbb{R}^{70}$ ) [23].

**公开基准数据集.** 目前, 针对认知能力评估, 尚缺乏基准的迁移学习数据集. 但认知障碍与行为能力密不可分, 瑞士伯尔尼大学 (University of Bern) 的 Roebbers 等 [24] 和荷兰鹿特丹伊拉斯姆斯大学 (Erasmus University Rotterdam) 医学中心的 Verlinden 等 [25] 的研究揭示了这种内在关联; 爱丁堡运动评估测试 (edinburgh motor assessment, EMAS)、共济失调评定量表 (scale for the assessment and rating of cerebellar ataxia, SARA) 和统一帕金森病评定量表 (the unified Parkinson's disease rating scale, UPDRS) 等多种临床常用认知能力评估量表也将行为能力作为衡量认知障碍的重要指标. 因此, 本文选择 4 种常用的行为能力评估数据集作为公开基准数据集, 很多迁移学习相关的研究也使用了相关数据集 [26]. 4 种公开基准数据集的简要介绍如表 1 所示, 包括: (1) DSADS 数据集 (标记为 D) 通过佩戴于 5 个位置的传感器采集 8 名受试者的 19 种行为 [27]; (2) UCI-HAR 数据集 (标记为 H) 通过佩戴于腰部的传感器采集 30 名受试者的 6 种行为 [28]; (3) USC-HAD 数据集 (标记为 U) 通过佩戴于前右臀部的传感器采集 14 名受试者的 12 种行为 [29]; (4) PAMAP 数据集 (标记为 P) 通过佩戴于 3 个位置的传感器采集 9 名受试者的 18 种行为 [30]. 由于不同数据集均采集自不同传感器、佩戴于不同位置并包含不同行为, 因此, 为实现不同数据集间迁移, 本文对初始数据集进行了统一化处理, 仅使用同种传感器、同种行为和同一位置的数据进行迁移. 具体而言, 本文使用三轴合成后的加速度和陀螺仪数据作为训练数据, 每种传感器分别提取 19 种、共 38 种时域和频域特征 [31, 32]; 使用右臂、腰部、右前腕部和右腕 4 种位置数据并筛选躺下、走路、上楼梯和下楼梯 4 种行为的数据作为训练数据集.

## 5.2 对比方法和参数设置

为验证 PAFG-TL 方法的实验效果, 本文选择 11 种已有方法作为对比方法, 包括:

- 分层迁移方法 (stratified transfer learning, STL) [26];
- 迁移成分分析方法 (transfer component analysis, TCA) [33];
- 测地线流式核方法 (geodesic flow kernel, GFK) [34];
- 仅使用目标域数据构建识别模型 (taronly, Tar);
- 仅使用源域数据构建识别模型 (srconly, Src);
- 同时使用源域和目标域数据构建识别模型 (comonly, Com);
- 决策树结构扩张/约减 (structure expansion/reduction, SER) [20];
- 结构迁移 (structure transfer, Struct) [20];
- SER 与 Struct 的混合方法 (mix of SER and Struct, Mix) [20];
- 参数全局优化设置的细粒度迁移学习方法 (fine-grained adaptation random forest, FAT) [10];

- 参数随机化设置的细粒度迁移学习方法 (FAT') [10].

其中, Src, Tar 和 Com 是批量学习方法; STL, TCA, GFK, SER, Struct, Mix 和 FAT 是迁移学习方法. SER, Struct, Mix 和 FAT 是基于随机森林的模型迁移方法; TCA, STL 和 GFK 是有代表性的特征迁移方法, 能够将源域和目标域数据映射至统一空间、减少数据分布差异. 实验环境为 Dell Precision 5530 (Intel Core i9-8950HK/32GB DDR3), 集成开发环境为 Matlab R2020a. 除 STL, Tar 和 GFK 外, 其余所有对比方法均使用随机森林作为基础分类器. 实验过程中, 设置随机森林中个体决策树个数为  $M = 30$ , 每个分割节点上候选特征为  $\sqrt{K}$  ( $K$  为特征总数), 节点停止分裂的最小样本数为 2, 单个个体决策树的最大深度为  $\max D = 10$ . 另外, STL, TCA 和 GFK 这 3 种基于属性迁移的学习方法需要对维度进行约减, 本文设置约减后维度为 30. 对 FAT 的  $\delta_1, \delta_2, \delta_3$  3 个参数进行全局优化设置, 分别设置为 0.6, 0.7, 0.8; 另外, 本文还对 FAT 的  $\delta_1, \delta_2, \delta_3$  3 个参数进行随机化设置并作为另一种对比方法, 后文记随意化参数设置的 FAT 方法为 FAT'.

### 5.3 对比实验结果

对比实验结果如表 2 所示, 每种迁移实验重复 10 次, 因此, 对于公开基准数据集, 表 2 中数值为 10 次实验平均识别准确率和标准差; 另外, 对于临床认知评估数据集, 由于不同场景 (即 TL 和 VS) 分别包含 4 种不同测试任务, 因此两两场景间的迁移过程 TL  $\rightarrow$  VS, VS  $\rightarrow$  TL, VS  $\rightarrow$  VS 和 TL  $\rightarrow$  TL 分别包含 16, 16, 12 和 12 种不同的迁移任务, 临床认知评估数据集准确率结果为同种迁移任务 10 次实验结果平均后不同迁移任务下识别准确率的平均, 标准差结果为同种迁移任务 10 次实验结果计算标准差后不同迁移任务下标准差的平均. 根据表 2 所示实验结果, 在临床认知评估数据集上, 本文所提出的 PAFG-TL 方法平均准确率为 79.4%, 优于 11 种对比方法, 对于不同场景间的迁移结果而言, PAFG-TL 在 4 种迁移过程中均优于对比方法. 相较于次优方法 FAT, 本文方法准确率提升 1.6%; 虽然 FAT 方法在进行全局参数优化设置情况下取得了次优的分类结果, 但在参数随机化设置情况下, 平均准确率仅为 73.8%, 本文提出的参数无关的 PAFG-TL 方法较 FAT' 准确率提升 5.6%. 另外, 通过识别标准差也可以看出, 本文方法具有一定的稳定性, 标准差小于 FAT 和 FAT' 方法. 对比 FAT 和 FAT' 方法的识别准确率和标准差, 可知已有工作 [10] 需进行参数优化设置才能具有较好的识别效果和稳定性.

在公开基准数据集上, 本文提出的 PAFG-TL 方法平均准确率为 69.8%, 优于前 9 种对比方法及随机化参数设置的 FAT' 方法; 在 12 种不同的迁移任务中, PAFG-TL 方法也均优于前 9 种对比方法及随机化参数设置的 FAT' 方法. 相较于进行全局参数调优的 FAT 方法, PAFG-TL 的平均准确率降低 0.7%, 但平均标准差略有改进; 在 12 种不同的迁移任务中的 4 种任务上识别准确率优于 FAT 方法, 12 种不同迁移任务中的 8 种任务上识别标准差优于 FAT 方法. 总体而言, 已有工作 [10] 中提出的 FAT 和 FAT' 方法, 需进行全局参数调优, 才能取得较好的识别效果 (FAT' 和 FAT 的平均准确率分别为 70.5% 和 66.3%) 和识别稳定性 (FAT' 和 FAT 的平均标准差分别为 7.9% 和 5.0%), 这也是该方法的主要不足, 而本文提出的 PAFG-TL 方法是一种参数无关的改进方法, 能够在无需全局参数调优情况下, 便取得较好的识别结果.

### 5.4 消融实验

消融实验过程中, 将 PAFG-TL 方法中的融合先验知识的高斯混合聚类方法逐一替换成 K 均值聚类 (K-means) 和传统高斯混合聚类 (Gaussian mixture model, GMM) 方法, 并分别计算识别准确率. 消融实验结果如图 2 所示, 同样, 每种不同迁移任务重复 10 次, 图 2(a) 中每个柱状图及误差线分别为

表 2 对比实验结果  
Table 2 The results of comparative experiments

	STL	TCA	GFK	Tar	Src	Com	SER	Struct	Mix	FAT'	FAT	PAFG
Clinical cognitive assessment dataset												
TL → VS	.474±.071	.533±.058	.561±.026	.702±.055	.605±.061	.713±.060	.713±.050	.746±.057	.735±.067	.769±.058	.805±.058	<b>.832±.042</b>
VS → TL	.509±.056	.510±.071	.556±.026	.633±.054	.537±.050	.663±.053	.659±.053	.671±.051	.668±.062	.694±.041	.741±.050	<b>.759±.036</b>
VS → VS	.554±.075	.537±.069	.553±.035	.696±.062	.671±.055	.750±.047	.670±.058	.815±.066	.784±.062	.788±.058	.855±.049	<b>.862±.042</b>
TL → TL	.481±.046	.552±.046	.584±.026	.654±.039	.572±.048	.685±.040	.645±.035	.734±.043	.709±.044	.700±.073	.712±.047	<b>.724±.036</b>
AVG	.505±.062	.533±.061	.563±.028	.671±.053	.596±.054	.703±.050	.672±.049	.741±.055	.724±.059	.738±.058	.778±.051	<b>.794±.039</b>
Public benchmark dataset												
D → H	.451±.000	.543±.055	.479±.003	.609±.055	.445±.025	.649±.028	.677±.050	.664±.075	.710±.043	.684±.099	<b>.732±.058</b>	.728±.043
D → U	.446±.004	.444±.026	.448±.003	.634±.027	.386±.031	.597±.020	.620±.023	.627±.043	.616±.043	.642±.054	<b>.708±.033</b>	.666±.058
D → P	.620±.005	.587±.013	.609±.005	.600±.047	.548±.023	.581±.019	.580±.039	.574±.033	.614±.035	.637±.065	<b>.666±.028</b>	.643±.039
H → D	.499±.004	.519±.037	.490±.003	.634±.111	.481±.067	.638±.037	.698±.041	.737±.034	.696±.067	.705±.088	.749±.064	<b>.782±.039</b>
H → U	.497±.002	.524±.019	.542±.003	.633±.024	.499±.034	.583±.023	.624±.026	.639±.018	.631±.015	.639±.058	.678±.040	<b>.683±.024</b>
H → P	.508±.003	.515±.009	.523±.003	.573±.052	.493±.024	.580±.029	.561±.044	.600±.037	.579±.040	.620±.060	<b>.650±.060</b>	.643±.031
U → D	.495±.006	.465±.2.00	.469±.005	.687±.093	.457±.006	.472±.019	.687±.079	.699±.052	.732±.033	.728±.104	<b>.770±.060</b>	.738±.102
U → H	.430±.003	.438±.026	.432±.003	.639±.088	.441±.007	.497±.049	.653±.041	.642±.052	.632±.052	.666±.077	.712±.077	<b>.716±.070</b>
U → P	.492±.003	.524±.017	.528±.004	.609±.051	.488±.007	.497±.011	.602±.019	.607±.018	.588±.022	.636±.066	<b>.661±.026</b>	.659±.026
P → D	.524±.006	.530±.056	.471±.008	.698±.060	.476±.029	.631±.068	.705±.055	.691±.064	.695±.115	.715±.108	<b>.780±.046</b>	.769±.062
P → H	.476±.002	.513±.020	.423±.001	.640±.059	.440±.006	.591±.058	.680±.044	.578±.050	.609±.057	.637±.107	<b>.685±.071</b>	.667±.047
P → U	.472±.004	.540±.027	.484±.003	.595±.043	.483±.020	.578±.023	.620±.025	.625±.029	.627±.025	.647±.061	.667±.035	<b>.683±.014</b>
AVG	.493±.004	.512±.192	.492±.004	.629±.059	.470±.023	.574±.032	.642±.040	.640±.042	.644±.046	.663±.079	<b>.705±.050</b>	.698±.046

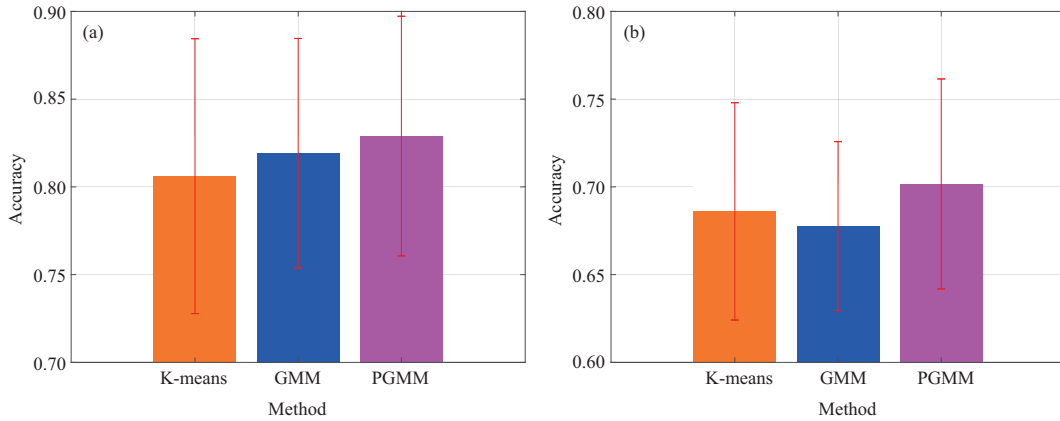


图 2 (网络版彩图) 消融实验结果

**Figure 2** (Color online) The results of ablation experiment for (a) clinical cognitive assessment dataset and (b) public benchmark dataset

56 种不同迁移任务下的实验结果平均值和标准差; 图 2(b) 中每个柱状图及误差线分别为 12 种不同迁移任务下的实验结果平均值和标准差. 临床认知评估数据集上的实验结果如图 2(a) 所示, K-means, GMM, PGMM 3 种聚类方法下的识别准确率分别为 80.6%, 81.9% 和 82.9%, 标准差分别为 7.8%, 6.5% 和 6.8%, 本文提出的 PGMM 方法优于其余两种对比方法, 准确率分别提升 2.3% 和 1.0%. 公开基准数据集上的实验结果如图 2(b) 所示, K-means, GMM, PGMM 3 聚类方法下的识别准确率分别为 68.6%, 67.8% 和 70.2%, 标准差分别为 6.2%, 4.8% 和 6.0%, 本文提出的 PGMM 方法优于其余两种对比方法, 准确率分别提升 1.6% 和 2.4%.

### 5.5 参数敏感性分析

PAFG-TL 是一种参数无关的迁移学习方法, 随机森林模型中的决策树数量  $M$  是该方法的显著影响变量; 同时, PAFG-TL 基于 FAT 方法提出,  $\delta_1$ ,  $\delta_2$ ,  $\delta_3$  是 FAT 方法中的显著影响变量. 因此, 本文将同时分析参数  $M$ ,  $\delta_1$ ,  $\delta_2$  和  $\delta_3$  对识别准确的影响. 参数  $M$  敏感性分析实验中, 设置  $M \in \{2, 4, 6, \dots, 28, 30\}$ , 依次计算  $M$  在不同取值下的识别准确率. 同样, 每种不同迁移任务重复 10 次, 实验结果如图 3 所示. 图 3(a) 中每条折线和误差带分别为 56 种不同迁移任务下的实验结果平均值和标准差; 图 3(b) 中每条折线和误差带分别为 12 种不同迁移任务下的实验结果平均值和标准差. 临床认知评估数据集上的实验结果如图 3(a) 所示, 从图 3(a) 可知, PAFG-TL 与 7 种对比方法的识别准确率均与参数  $M$  相关, 随着  $M$  的提升, 识别准确率逐步提升; PAFG-TL 较 7 种对比方法, 当参数  $M > 5$  时, 均能取得更好的识别准确率, 可能原因是随  $M > 5$  的提升, PAFG-TL 具有更好的模型多样性, 从而能够取得较好的识别准确率. 公开基准数据集上的实验结果如图 3(b) 所示, 同样, 随参数  $M$  的提升, 识别准确率逐步提升; 相较于 7 种对比方法, 当参数  $M > 9$  时, PAFG-TL 均能取得更好的识别准确率.

参数  $\delta_1$ ,  $\delta_2$ ,  $\delta_3$  敏感性分析实验中, 设置  $\delta_1 \in \{0.12, 0.18, \dots, 0.66\}$ ,  $\delta_2 \in \{0.61, 0.63, \dots, 0.79\}$ ,  $\delta_3 \in \{0.71, 0.73, \dots, 0.89\}$ , 遍历不同  $\delta_1$ ,  $\delta_2$ ,  $\delta_3$  取值情况下 FAT 方法的迁移准确率并与 PAFG-TL 方法的结果进行对比. 类似于参数  $M$  敏感性分析过程, 本小节对各种迁移场景下的不同迁移任务依次进行实验, 每种实验重复 10 次并将同一类型迁移任务结果汇总, 实验结果如图 4 所示. 另外, 由于  $\delta_1$ ,  $\delta_2$ ,  $\delta_3$  非 PAFG-TL 方法的影响变量, 因此, 为便于对比不同参数下 FAT 方法与 PAFG-TL 方法的实

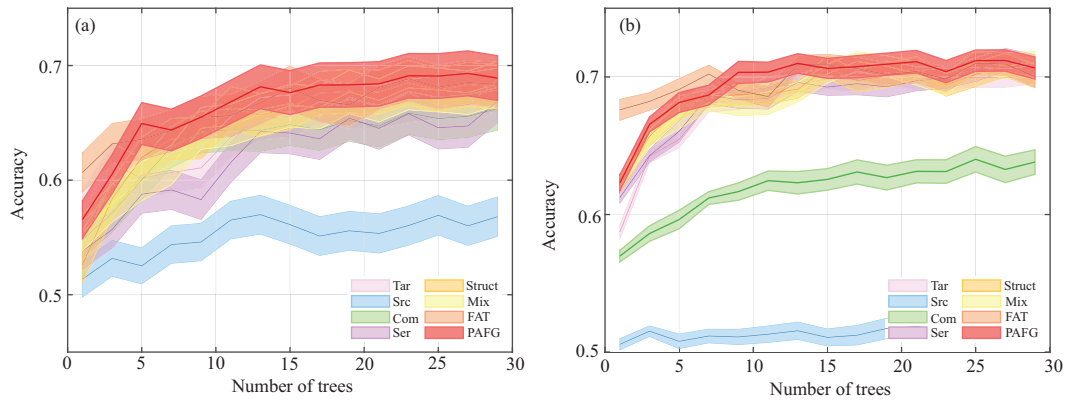


图 3 (网络版彩图)  $M$  参数敏感性分析实验结果

Figure 3 (Color online) The results of parameter  $M$  analysis experiment for (a) clinical cognitive assessment dataset and (b) public benchmark dataset

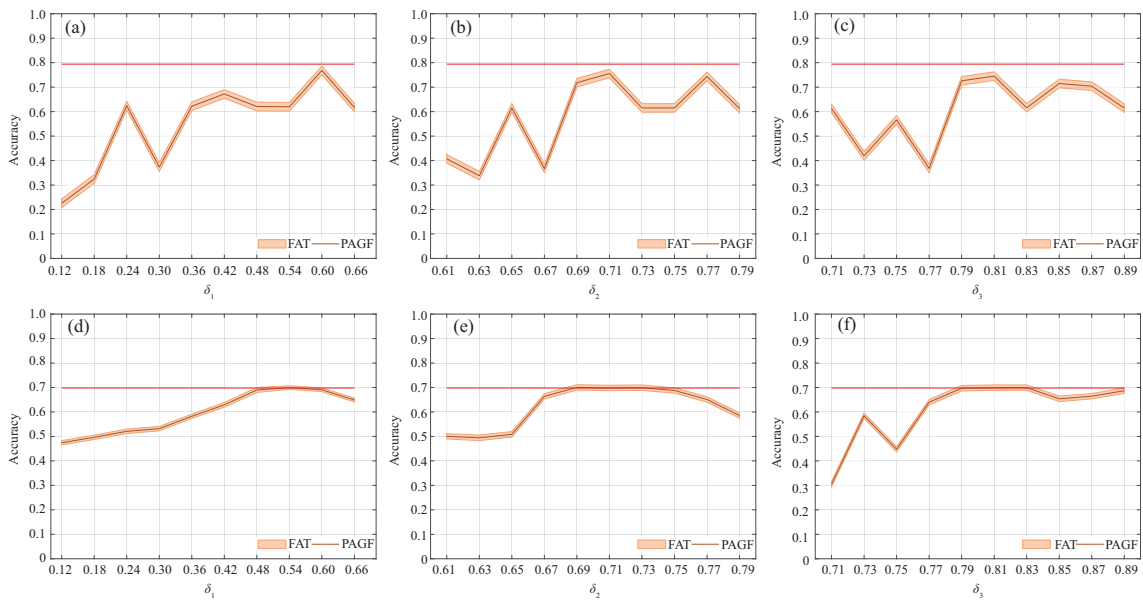


图 4 (网络版彩图) 参数敏感性分析实验结果

Figure 4 (Color online) The results of parameters analysis experiment for (a) clinical cognitive assessment dataset  $\delta_1$ , (b) clinical cognitive assessment dataset  $\delta_2$ , (c) clinical cognitive assessment dataset  $\delta_3$ , (d) public benchmark dataset  $\delta_1$ , (e) public benchmark dataset  $\delta_2$ , and (f) public benchmark dataset  $\delta_3$

验结果, 取 PAFG-TL 方法多次实验下的均值绘制于图 4. 从图 4 可知, 在较优的参数设置下, FAT 方法甚至可取得优于 PAFG-TL 方法的迁移效果, 但 FAT 方法对  $\delta_1, \delta_2, \delta_3$  参数敏感, 需进行参数调优以提升迁移效果, 参数无关、适应性强是 PAFG-TL 方法相较于 FAT 方法的主要优势.

### 5.6 时间复杂度分析

本文依次分析了当参数  $M$  取不同数值时, Tar, Src, Com, Ser, Struct, Mix, FAT 和 PAFG-TL 方法在临床认知评估数据集和公开基准数据集上的模型训练、更新和测试耗时, 实验结果如图 5 所示. 从图 5 可知, 参数  $M$  与模型训练、更新和测试耗时显著相关, 随着参数  $M$  的提升, 模型训练、更新和

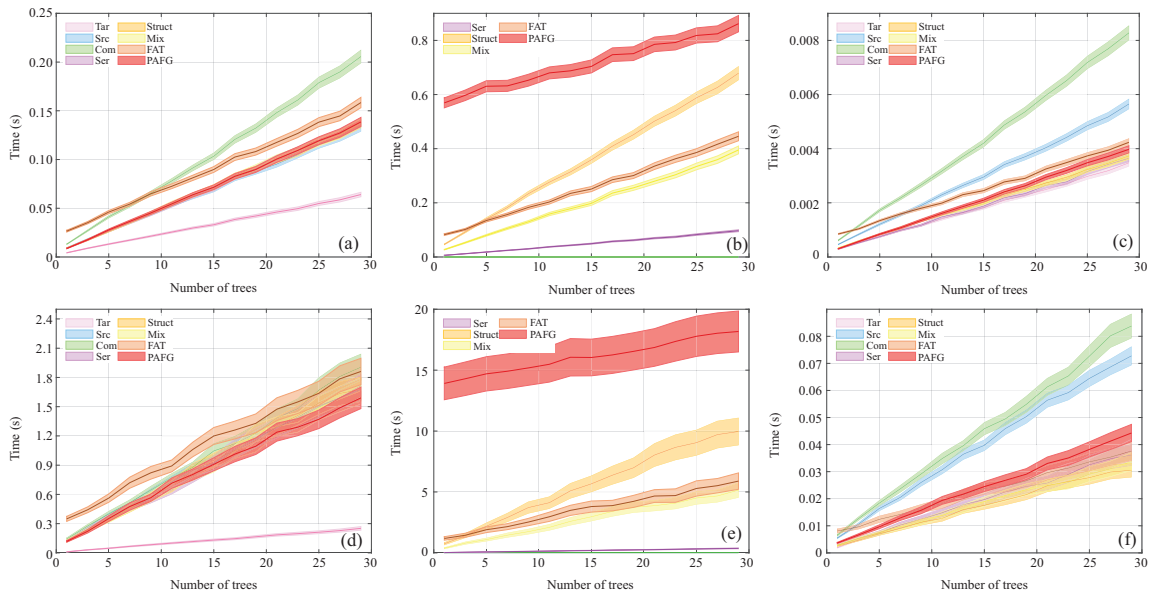


图 5 (网络版彩图) 时间复杂度分析实验结果

Figure 5 (Color online) The time consuming analysis experiment results of (a) training, (b) updating, (c) testing time consuming in clinical cognitive assessment dataset, as well as (d) training, (e) updating, and (f) testing time consuming in public benchmark dataset

测试耗时显著提升. 在模型训练耗时方面, PAFG-TL 与 FAT 方法接近, 训练耗时长于 Src 和 Com 方法, 短于 Struct, Mix, Tar 和 Ser 方法; 在模型更新耗时方面, PAFG-TL 显著高于其他 7 种对比方法, 一种可能原因是 PAFG-TL 方法中个体分类器模型聚类过程较为耗时; 模型测试耗时的结果与模型训练耗时结果接近, 测试耗时长于 Src 和 Com 方法, 短于 Struct, Mix, Tar 和 Ser 方法.

## 6 结语

认知能力由语言能力、思维能力、记忆力、执行能力、判断能力、注意力和视觉搜索能力等多种能力构成, 认知健康关系个体一生的正常发展. 近年来, 随着可穿戴感知、智能计算等技术的发展, 智慧医疗的概念逐渐受到越来越多相关从业人员的关注. 然而, 局限于传统受限环境下数据感知、穿戴方式、适用时间和适用场景的不足, 智慧医疗难以突破传统医学诊疗的设备和时空限制, 实现疾病的早发现、早诊断. 普适计算强调通过无所不在的计算服务为用户提供便利, 能够拓展传统计算模式: 通过手表、手环、步态鞋等可穿戴设备扩展数据感知方式, 满足穿戴方式不受限、长时间监护、多场景应用需求. 如何构建通用的认知健康评估模型, 实现医疗问诊、社区生活和居家监护等不同场景下的跨场景适配, 是普适环境下认知健康评估面临的一个重要挑战.

本文针对普适环境下认知健康的跨场景评估问题, 基于在相关领域的已有研究基础<sup>[10]</sup>, 提出了一种参数自适应的细粒度迁移学习方法 PAFG-TL. PAFG-TL 基于随机森林模型实现, 由参数自适应的个体分类器评估策略和领域自适应的决策树生长机制两部分构成, 能够分别分类评估不同的个体分类器并使用不同策略对不同分类器进行个性化更新. 总体而言, PAFG-TL 仅需少量样本便可实现源域模型的个性化更新, 能够适用于数据小样本和可解释性要求高的应用场景; PAFG-TL 方法显式定义了特征集合适应度和分割阈值适应度, 能够为个体分类器评估提供一种直观的解释; PAFG-TL 方法是一

种参数无关的方法, 无需通过全局参数优化, 便可以取得较好的识别效果和稳定性. 在临床认知评估和公开基准数据集上的实验验证, 均证明了 PAFG-TL 方法在跨场景认知评估中的有效性、鲁棒性和高效性.

## 参考文献

- 1 Hendrie H C, Albert M S, Butters M A, et al. The NIH cognitive and emotional health project. *Alzheimers Dement*, 2006, 2: 12–32
- 2 Banaschewski T, Becker D M, Holtmann M, et al. Attention-deficit/hyperactivity disorder. *Dtsch Arztebl Int*, 2017, 114: 149–159
- 3 World Health Organization. Dementia. 2019. <https://www.who.int/en/news-room/fact-sheets/detail/dementia>
- 4 Fang C, Li C, Cabrerizo M, et al. A Gaussian discriminant analysis-based generative learning algorithm for the early diagnosis of mild cognitive impairment in Alzheimer’s disease. In: *Proceedings of IEEE International Conference on Bioinformatics and Biomedicine*, 2017. 538–542
- 5 Lyytinen K, Yoo Y. Ubiquitous computing. *Commun ACM*, 2002, 45: 63–96
- 6 Kjeldskov J, Skov M B. Exploring context-awareness for ubiquitous computing in the healthcare domain. *Pers Ubiquit Comput*, 2007, 11: 549–562
- 7 Pan S J, Yang Q. A survey on transfer learning. *IEEE Trans Knowl Data Eng*, 2010, 22: 1345–1359
- 8 Zhang Y, Chen Y, Yu H, et al. Dual layer transfer learning for sEMG-based user-independent gesture recognition. *Pers Ubiquit Comput*, 2022, 26: 575–586
- 9 Long M, Cao Y, Cao Z, et al. Transferable representation learning with deep adaptation networks. *IEEE Trans Pattern Anal Mach Intell*, 2019, 41: 3071–3085
- 10 Zhang Y, Chen Y, Yu H, et al. Bridging cross-tasks gap for cognitive assessment via fine-grained domain adaptation. In: *Proceedings of the 29th International Conference on International Joint Conferences on Artificial Intelligence*, 2021. 4330–4337
- 11 Kermany D S, Goldbaum M, Cai W, et al. Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell*, 2018, 172: 1122–1131
- 12 Khatami A, Babaie M, Tizhoosh H R, et al. A sequential search-space shrinking using CNN transfer learning and a Radon projection pool for medical image retrieval. *Expert Syst Appl*, 2018, 100: 224–233
- 13 Yu Y, Lin H, Meng J, et al. Deep transfer learning for modality classification of medical images. *Information*, 2017, 8: 91
- 14 Gao F, Yoon H, Wu T, et al. A feature transfer enabled multi-task deep learning model on medical imaging. *Expert Syst Appl*, 2020, 143: 112957
- 15 Swati Z N K, Zhao Q, Kabir M, et al. Brain tumor classification for MR images using transfer learning and fine-tuning. *Comput Med Imag Graphics*, 2019, 75: 34–46
- 16 Banerjee I, Crawley A, Bhethanabotla M, et al. Transfer learning on fused multiparametric MR images for classifying histopathological subtypes of rhabdomyosarcoma. *Comput Med Imag Graph*, 2018, 65: 167–175
- 17 Sukhija S, Krishnan N C, Singh G. Supervised heterogeneous domain adaptation via random forests. In: *Proceedings of International Joint Conference on Artificial Intelligence*, 2016. 2039–2045
- 18 Wen Y, Qin Y, Qin K, et al. Online transfer learning with multiple decision trees. *Int J Mach Learn Cyber*, 2019, 10: 2941–2962
- 19 Ryu J, Bae J, Lim J. Collaborative training of balanced random forests for open set domain adaptation. 2020. [ArXiv:2002.03642](https://arxiv.org/abs/2002.03642)
- 20 Segev N, Harel M, Mannor S, et al. Learn on source, refine on target: a model transfer learning framework with random forests. *IEEE Trans Pattern Anal Mach Intell*, 2016, 39: 1811–1824
- 21 Wehrens R, Putter H, Buydens L M C. The bootstrap: a tutorial. *Chemometr Intell Lab Syst*, 2000, 54: 35–52
- 22 Krogh A, Vedelsby J. Neural network ensembles, cross validation, and active learning. In: *Proceedings of the 7th International Conference on Neural Information Processing Systems*, 1994. 231–238
- 23 Zhang Y, Chen Y, Yu H, et al. What can “drag & drop” tell? Detecting mild cognitive impairment by hand motor function assessment under dual-task paradigm. *Int J Hum-Comput Stud*, 2021, 145: 102547

- 24 Roebers C M, Röthlisberger M, Neuenschwander R, et al. The relation between cognitive and motor performance and their relevance for children's transition to school: a latent variable approach. *Hum Movement Sci*, 2014, 33: 284–297
- 25 Verlinden V J A, van der Geest J N, Hofman A, et al. Cognition and gait show a distinct pattern of association in the general population. *Alzheimers Dement*, 2014, 10: 328–335
- 26 Wang J, Chen Y, Hu L, et al. Stratified transfer learning for cross-domain activity recognition. In: *Proceedings of IEEE International Conference on Pervasive Computing and Communications*, 2018. 1–10
- 27 Barshan B, Yüsek M C. Recognizing daily and sports activities in two open source machine learning environments using body-worn sensor units. *Comput J*, 2014, 57: 1649–1667
- 28 Anguita D, Ghio A, Oneto L, et al. Human activity recognition on smartphones using a multiclass hardware-friendly support vector machine. In: *Proceedings of International Workshop on Ambient Assisted Living*, 2012. 216–223
- 29 Zhang M, Sawchuk A A. USC-HAD: a daily activity dataset for ubiquitous activity recognition using wearable sensors. In: *Proceedings of ACM Conference on Ubiquitous Computing*, 2012. 1036–1043
- 30 Reiss A, Stricker D. Introducing a new benchmarked dataset for activity monitoring. In: *Proceedings of the 16th International Symposium on Wearable Computers*, 2012. 108–109
- 31 Lu W, Chen Y, Wang J, et al. Cross-domain activity recognition via substructural optimal transport. *Neurocomputing*, 2021, 454: 65–75
- 32 Qin X, Chen Y, Wang J, et al. Cross-dataset activity recognition via adaptive spatial-temporal transfer learning. *Proc ACM Interact Mob Wearable Ubiquit Technol*, 2019, 3: 1–25
- 33 Pan S J, Tsang I W, Kwok J T, et al. Domain adaptation via transfer component analysis. *IEEE Trans Neural Netw*, 2010, 22: 199–210
- 34 Gong B, Shi Y, Sha F, et al. Geodesic flow kernel for unsupervised domain adaptation. In: *Proceeding of Conference on Computer Vision and Pattern Recognition*, 2012. 2066–2073

## A fine-grained transfer learning method for cross-scenario cognitive-ability assessment

Yingwei ZHANG<sup>1,2</sup>, Yiqiang CHEN<sup>1,2,3\*</sup>, Hanchao YU<sup>4</sup>, Xiaodong YANG<sup>1,2</sup> & Yang GU<sup>1,2,3</sup>

1. *Beijing Key Laboratory of Mobile Computing and Pervasive Devices, Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190, China;*

2. *University of Chinese Academy of Sciences, Beijing 100049, China;*

3. *Peng Cheng Laboratory, Shenzhen 518066, China;*

4. *Bureau of Frontier Sciences and Education, Chinese Academy of Sciences, Beijing 100864, China*

\* Corresponding author. E-mail: yqchen@ict.ac.cn

**Abstract** Cognitive health is an important aspect of brain health and is closely related to an individual's development throughout their life. Currently, cognitive-ability assessment in unrestricted ubiquitous scenarios has gained attention for realizing early warnings of cognitive impairment-related diseases. However, the perception devices and differences in the cognitive assessment processes in different application scenarios (e.g., hospital and home) influence cross-scenario usage. To address these challenges, we propose a parameter-adaptive, fine-grained transfer-learning method (PAFG-TL), which is designed based on the random forest algorithm. It is composed of a parameter adaptation-based model-evaluation strategy and a domain-adaptive tree-growing mechanism to realize a parameter-independent clustering evaluation of individual classifiers and the fine-grained evolutionary growth of decision trees. In this study, the experimental verification of clinical cognitive assessment and public benchmark datasets demonstrated the effectiveness of PAFG-TL in cross-scenario cognitive-ability assessment.

**Keywords** parameter adaptive, fine-grained, transfer learning, cross-scenario, cognitive ability assessment