

面向类别不平衡 ECG 的快速患者间域适应 心律失常识别方法

冯建超^{1,2}, 吴龙文^{1,2}, 王 琪³, 李 涵¹, 何 欣¹, 赵雅琴^{1,2*}

(1. 哈尔滨工业大学电子与信息工程学院, 黑龙江哈尔滨 150001; 2. 哈尔滨工业大学郑州研究院, 河南郑州 450008;
3. 中国兵器工业集团航空弹药研究院有限公司, 黑龙江哈尔滨 150030)

摘 要: 心律失常作为心血管疾病中诱发心源性猝死的关键诱因,其早期精准识别与动态分类对改善患者预后具有重要临床意义。然而,受个体差异、采集设备差异及采集环境变化等因素影响,不同患者及不同数据库之间的心电图(ElectroCardioGram, ECG)信号普遍存在显著的域偏移问题;同时,临床 ECG 数据中普遍存在类别分布高度不平衡的现象,严重制约了现有模型在跨患者、跨数据库场景下的泛化性能。针对上述问题,本文提出了一种面向类别不平衡 ECG 信号的轻量级快速患者间域适应心律失常识别框架。首先,在特征层面,本文提出了一种基于宽度学习系统(Broad Learning System, BLS)的特征分布域适应方法 BLS-FDDA (BLS-Feature Distribution Domain Adaptation),基于协方差归一化和分布重构技术,深入分析了源域与目标域特征之间的偏移问题,成功实现了源目标域特征空间的对齐。该方法通过对 BLS 特征扩展矩阵的分布对齐,避免了传统深度学习框架中对复杂模型的依赖,并保证了特征信息的有效传递。其次,在数据层面,本文进一步提出了一种可逆数据域适应方法 BLS-DRDA (BLS-Data Reversible Domain Adaptation),结合误差扰动理论,深入推导了源域与目标域之间的数据变换关系。基于这一理论推导,该方法实现了在不重新训练 BLS 主模型的情况下,快速适应新数据域,大幅降低了迁移成本。此外, BLS-DRDA 方法在数据变换中保持了原始信号的判别能力,并有效避免了信息失真。在决策层,针对心律失常数据中类别分布高度不均衡的问题,设计了一种代价敏感判决算法,该算法通过引入类别中心距离和样本分布权重的概念,建立了一个加权决策机制,有效缓解了少数类样本在跨域迁移中的误判问题。最后,在 MIT-BIH 与 INCART 公开数据库构建的多患者、多数据库及连续域迁移实验中,所提方法在准确率、F1 值及 G_mean 等指标上均取得接近 100% 的识别性能,显著优于原始宽度学习模型及多种对比方法。理论分析与实验结果表明,所提出的 BLS-FDDA 和 BLS-DRDA 方法在多源数据、跨设备以及连续域适应场景中均表现出优越的性能,验证了该框架在复杂临床 ECG 应用中的有效性与实用性,特别是在多类别患者间心律失常的识别任务中,所提方法显著提升了少数类的识别能力,并在复杂域偏移和类别不平衡问题下展现出极强的鲁棒性。

关键词: 心律失常识别;域适应;类别不平衡;宽度学习系统

基金项目: 国家自然科学基金(No.62571163, No.62571167, No.62071153)

中图分类号: TP183; R541.7 **文献标识码:** A **文章编号:** 0372-2112(2026)02-0684-14

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.12263/DZXB.20251121

A Fast Inter-Patient Domain-Adaptive Arrhythmia Detection Method for Class-Imbalanced ECG Data

FENG Jianchao^{1,2}, WU Longwen^{1,2}, WANG Qi³, LI Han¹, HE Xin¹, ZHAO Yaqin^{1,2*}

(1. School of Electronics and Information Engineering, Harbin Institute of Technology, Harbin, Heilongjiang 150001, China;

2. Zhengzhou Research Institute, Harbin Institute of Technology, Zhengzhou, Henan 450008, China;

3. Norinco Group Air Ammunition Research Institute, Harbin, Heilongjiang 150030, China)

Abstract: Arrhythmia, as a key trigger for sudden cardiac death in cardiovascular diseases, holds significant clinical importance for improving patient outcomes through its early, precise identification and dynamic classification. However, due to factors such as individual variability, differences in recording devices, and variations in recording environments, significant domain shift issues are prevalent across ECG signals from different patients and databases. Additionally, clinical ECG data commonly exhibit highly imbalanced class distributions, severely limiting the generalization performance of existing models across patients and databases. To address these challenges, this paper proposes a lightweight, fast inter-patient domain adaptation framework for arrhythmia recognition tailored to class-imbalanced ECG signals. First, at the feature lev-

el, we introduce BLS-FDDA (broad learning system-feature distribution domain adaptation), a domain adaptation method based on the broad learning system (BLS). By leveraging covariance normalization and distribution reconstruction techniques, it thoroughly analyzes the offset between source and target domain features, successfully aligning their feature spaces. By aligning the distributions of the BLS feature expansion matrix, this method avoids reliance on complex models inherent in traditional deep learning frameworks while ensuring effective feature information transfer. Second, at the data level, this paper further proposes a reversible data domain adaptation method, BLS-DRDA (BLS-data reversible domain adaptation). Integrating error perturbation theory, it derives the data transformation relationship between source and target domains. Based on this theoretical derivation, the method achieves rapid adaptation to new data domains without retraining the BLS main model, significantly reducing transfer costs. Moreover, BLS-DRDA preserves the discriminative capability of the original signal during data transformation while effectively preventing information distortion. At the decision layer, addressing the severe class imbalance in arrhythmia data, a cost-sensitive decision algorithm is designed. By introducing concepts of class center distance and sample distribution weights, this algorithm establishes a weighted decision mechanism that effectively mitigates misclassification issues of minority samples during cross-domain transfer. Finally, multi-patient, multi-database, and continuous domain transfer experiments conducted on the MIT-BIH and INCART public datasets demonstrate that the proposed methods achieve recognition performance approaching 100% in metrics such as accuracy, F1_score, and G_mean, significantly outperforming the original width learning model and various comparative methods. Theoretical analysis and experimental results demonstrate that the proposed BLS-FDDA and BLS-DRDA methods exhibit superior performance across multi-source data, cross-device, and continuous domain adaptation scenarios. This validates the framework's effectiveness and practicality in complex clinical ECG applications, particularly in the task of identifying arrhythmias across multiple patient categories. The proposed methods substantially enhance recognition capabilities for minority classes and demonstrate exceptional robustness under complex domain shifts and class imbalance challenges.

Keywords: arrhythmia detection; domain adaptation; class imbalance; broad learning system

Foundation Item(s): National Natural Science Foundation of China (No.62571163, No.62571167, No.62071153)

0 引言

心律失常 (Arrhythmia) 是一类常见的心血管疾病,由心脏电活动异常引起,其早期检测与准确诊断对降低突发性心脏事件的风险、提高患者生存率具有重要意义^[1]。心电图 (ElectroCardioGram, ECG) 作为反映心脏电生理活动的非侵入性信号,是心律失常识别和临床诊断的主要依据。随着可穿戴设备和远程医疗技术的发展,来自不同采集设备、患者群体和环境条件的 ECG 数据被广泛收集。然而,这些数据存在显著的域差异,如采样率、噪声、导联配置和个体差异,导致模型在跨数据集或跨设备应用时性能下降^[2]。

为应对这一挑战,域适应 (Domain Adaptation, DA) 技术逐渐成为跨域心律失常识别的重要研究方向。域适应作为迁移学习的一个重要分支,旨在通过学习源域与目标域之间共享的特征表示,缩小源域与目标域之间的分布差异,从而提高模型在未见数据域中的泛化性能^[3]。对心律失常检测任务,基于 ECG 信号的域适应算法可有效解决跨数据库、跨设备及个体差异带来的性能退化问题,具有重要的理论意义和应用价值。

尽管深度学习在心律失常检测中取得了显著进展,但在跨域应用中仍面临许多挑战^[4]。首先,深度

网络具有庞大的参数规模和复杂结构,训练依赖于大数据和计算资源,难以在资源受限的医疗设备上部署。其次,深度模型在域偏移下容易过拟合源域特征,导致泛化能力不足。为缓解这一问题,研究者提出了多种域适应方法^[5],但这些方法往往依赖复杂的神经网络结构和对抗训练机制,训练过程不稳定、缺乏解释性,且在多源或异构 ECG 数据上的适应性有限。

宽度学习系统 (Broad Learning System, BLS) 作为一种新型浅层学习架构,为此类问题提供了新的解决思路^[6]。BLS 通过扩展网络的“宽度”而非“深度”增强特征,具备训练速度快、结构可扩展性强、模型可解释性好的优势,特别适合医疗信号处理中的高效性和可解释性需求。近年来,研究者尝试将域适应思想引入 BLS 框架,提出了基于宽度学习的域适应算法,如基于 BLS 的迁移学习算法^[7]。尽管这些方法在一定程度上提高了 BLS 的跨域迁移性能,但仍存在下列问题:首先,域适应模块与 BLS 结构的融合度不足,适应机制通常作为附加模块独立嵌入,缺乏端到端的联合优化,导致特征对齐与分类决策难以协同提升。其次,现有算法普遍缺乏持续适应能力,在源域与目标域差异较大时,模型容易过拟合源域特征,导致目标域性能下降。此外,ECG 数据存在显著的数据不平衡问题,进一步加剧了迁移中的偏倚风险,导致模型难

以学习少数类特征。

针对上述问题,我们提出了一套面向跨域心律失常识别的宽度学习域适应与代价敏感决策框架,主要贡献如下:首先,提出了BLS-Feature Distribution Domain Adaptation(BLS-FDDA)算法,通过协方差归一化和分布对齐实现端到端的特征层域适应,避免了BLS结构的改动,有效迁移源域判别特征至目标域;其次,提出了BLS-Data Reversible Domain Adaptation(BLS-DRDA)算法,通过误差理论推导数据级变换关系,实现了在不重新训练BLS主模型的前提下,快速适应新设备、新数据库或新人群的ECG信号;最后,设计了基于类中心距离和样本量的Cost-Sensitive Decision Algorithm(CSDA),在判决阶段对少数类赋予更高权重,缓解了域偏移与类别不平衡带来的误判问题。实验结果表明,所提方法在多种患者群体、跨数据库及连续域迁移场景下,表现出优越的准确率和少数类识别能力,验证了框架的优越性与鲁棒性。

1 相关研究

1.1 域适应技术心律失常识别领域的研究

心律失常的自动检测是心血管疾病早期诊断的关键。早期的心律失常检测方法多依赖于手工特征提取技术,通常通过从ECG信号中提取时域、频域或时频域特征,并使用传统的机器学习算法进行分类^[8]。随着深度学习技术在心律失常检测中的应用,跨域迁移问题逐渐引起关注。域适应技术作为迁移学习的一种重要分支,旨在解决源域和目标域之间的数据分布差异问题,从而提升模型在新数据域上的适应能力^[9]。在心律失常检测中,域适应技术主要用于减少不同ECG数据集之间、不同设备采集的ECG信号之间的域偏移。

近年来,基于最大均值差异(Maximum Mean Discrepancy, MMD)、对比学习、特征对齐等方法的域适应技术被广泛应用于各个领域,取得了良好的性能^[10-11]。在心律失常识别方面,Kang等人^[12]提出了多查询频率提示(Multi-Query Frequency Prompting, MQFP)框架,利用频率域的可学习嵌入,提升了模型面对源域与目标域的域偏移时的性能,显著减少了误差并降低了计算复杂度。Msigwa等人^[13]提出了一种结合多流卷积递归神经网络和注意力机制的深度学习模型,采用了聚类技术和基于生成对抗网络的数据增强方法,实验结果表明,其F1得分从0.876提高到0.962,展示了在跨域迁移中有效应对噪声和数据不平衡的能力。韩昕哲等人^[14]提出的域适应分类模型MMD-Net,利用最大均值差异将源数据和目标数据映射至再生希尔伯特空间,提升了跨域学习能力。

当前ECG域适应研究主要集中于源域到目标域的单向迁移,通过对抗学习和特征对齐减少域间差异,提升目标域性能。然而,这些方法依赖目标域数据,且在目标域变化时难以快速迁移,需重新对齐和训练,限制了在临床场景中的应用。为避免此问题,需构建具有稳健泛化能力的模型框架,并实现与主任务的深度耦合。尽管现有方法缓解了跨域问题,但仍面临域适应模块与主模型融合度不足、对非线性分布偏移处理能力有限等挑战。

1.2 基于宽度学习的域适应技术研究

宽度学习系统是一种新型的浅层学习模型,通过扩展网络的宽度而非深度来增强特征表达能力。近年来,BLS已在多个领域表现出优异的性能,尤其在处理小样本、计算资源有限的场景下具有明显优势^[15]。针对ECG信号的分析,BLS能够快速学习并提取重要特征,减少对深度网络中庞大参数的依赖,且具有较好的模型可解释性。

近年来,已有研究尝试将域适应方法与BLS结合。Qin等人^[16]提出了部分域适应宽度学习系统(Partial Domain Adaptation Broad Learning System, PDA-BLS),通过结合源样本权重的最大均值差异和分类策略,有效减少源域和目标域之间的特征分布差异,提升了跨域故障诊断的准确性和泛化能力。Wang等人^[17]提出了类加权域适应网络(Class-Weighted Domain Adaptation Network, CWDAN),通过结合多核最大均值差异(Multi-Kernel Maximum Mean Discrepancy, MK-MMD)和Correlation Alignment(CORAL)损失进行域适应,并将CWMM引入BLS。该方法有效减少了源域和目标域之间的条件分布差异和类别偏差,并通过宽度扩展增强了对域不变特征的表示能力。此外,在跨域ECG识别任务中,同样有团队开展基于BLS的域适应算法研究。Fan等人^[18]提出了两种基于类别加权BLS的域适应算法:一步域适应(One-step CSWBLS-based Domain Adaptation, OCSWBLS-DA)和两步域适应(Two-step CSWBLS-based Domain Adaptation, TCSWBLS-DA)。OCSWBLS-DA从常见心跳中学习知识实现个体适应,而TCSWBLS-DA则先对常见心跳进行预训练,再针对患者特异性心跳进行微调。实验结果表明,两种方法在MIT-BIH心律失常数据库均表现出优秀的性能。

然而,现有的基于BLS的域适应研究仍然存在若干不足:首先,域适应模块通常作为附加机制嵌入BLS框架中,缺乏与BLS核心结构的深度融合,这使得跨域迁移的效果受到一定限制;其次,虽然有一些研究尝试引入对抗性损失来提高域适应能力,但在面对复杂的非线性域偏移时,BLS的特征映射层缺乏动

态自适应的能力,无法有效抑制负迁移,导致模型在多源或异构数据上的表现不稳定。与此同时,临床 ECG 数据普遍存在类别分布不均衡、少数类样本极度稀缺等问题,传统域适应方法往往侧重于对齐整体分布而忽略源、目标域中各类别样本的权重差异,容易在对齐过程中进一步“稀释”少数类信息,加剧少数类心律失常的误检与漏检风险。

2 本文方法

本文的核心工作聚焦于解决 ECG 信号心律失常识别中的域偏移问题,借助 BLS 的高效计算优势,设计了数据层面和特征层面两种域适应策略,并辅以特定类别决策修正机制,进一步提升识别性能。所提方法的整体框架如图 1 所示。

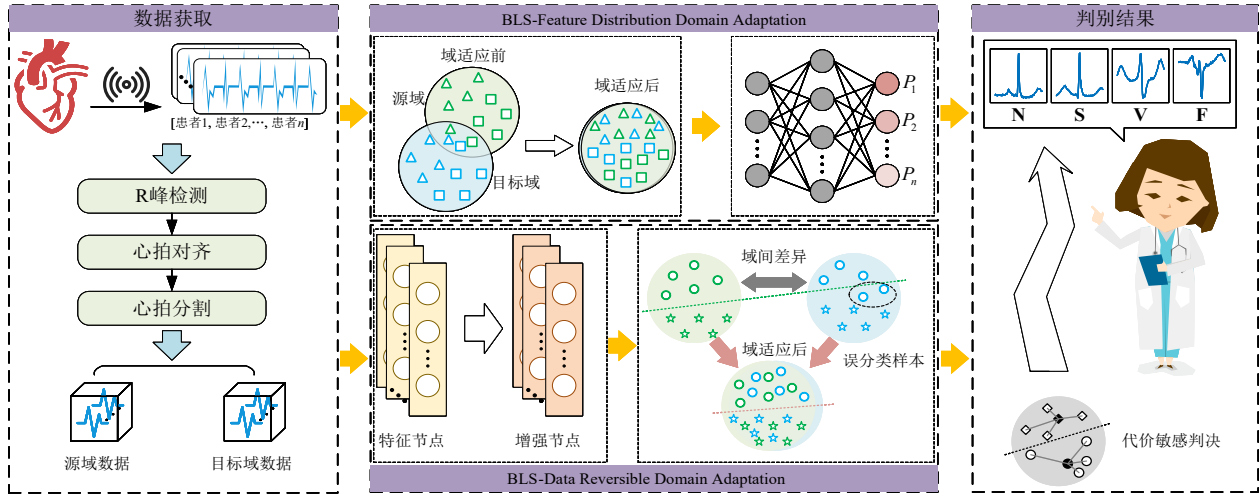


图 1 总体框架示意图

Figure 1 Overall framework

2.1 宽度学习算法

宽度学习系统的结构如图 2 所示,设输入数据矩阵为 $\mathbf{X} \in \mathbb{R}^{N \times M}$,其对应的标签矩阵表示为 $\mathbf{Y} \in \mathbb{R}^{N \times L}$ 。其中, N 为样本数量, M 表示 \mathbf{X} 中数据的维度, L 是类别数。

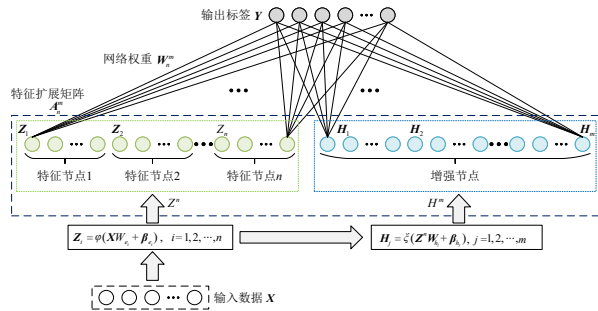


图 2 宽度学习系统结构图

Figure 2 Broad learning system architecture

在 BLS 中,设有 n 组特征映射节点和 m 组增强节点,则第 i 组特征映射节点 $\mathbf{Z}_i \in \mathbb{R}^{N \times k}$ 可由下式给出:

$$\mathbf{Z}_i = \varphi(\mathbf{X}\mathbf{W}_{e_i} + \boldsymbol{\beta}_{e_i}), i = 1, 2, \dots, n \quad (1)$$

其中: $\mathbf{W}_{e_i} \in \mathbb{R}^{M \times k}$ 、 $\boldsymbol{\beta}_{e_i} \in \mathbb{R}^{N \times k}$ 分别为初始权值和偏置; k 为 \mathbf{Z}_i 中的节点数; $\varphi(\cdot)$ 是非线性激活函数。将所有特

征节点 \mathbf{Z}_i 组合得到 $\mathbf{Z}^n \equiv [\mathbf{Z}_1, \mathbf{Z}_2, \dots, \mathbf{Z}_n] \in \mathbb{R}^{N \times (n \times k)}$ 。

对于每一组增强节点 $\mathbf{H}_j \in \mathbb{R}^{N \times p}$ 有

$$\mathbf{H}_j = \zeta(\mathbf{Z}^n \mathbf{W}_{h_j} + \boldsymbol{\beta}_{h_j}), j = 1, 2, \dots, m \quad (2)$$

其中: $\mathbf{W}_{h_j} \in \mathbb{R}^{(n \times k) \times p}$ 、 $\boldsymbol{\beta}_{h_j} \in \mathbb{R}^{N \times p}$ 分别为初始化随机权值和偏置; p 为 \mathbf{H}_j 中的节点数; $\zeta(\cdot)$ 为激活函数,通常使用 $\text{tansig}(\cdot)$ 函数。得到所有增强节点 $\mathbf{H}^m \equiv [\mathbf{H}_1, \mathbf{H}_2, \dots, \mathbf{H}_m] \in \mathbb{R}^{N \times (m \times p)}$ 后,将所有特征节点和增强节点连接成一个特征扩展矩阵,表示为

$$\mathbf{A}^m = [\mathbf{Z}^n | \mathbf{H}^m] \in \mathbb{R}^{N \times (n \times k + m \times p)} \quad (3)$$

即网络的输出模型表示为

$$\mathbf{Y} = \mathbf{A}^m \mathbf{W}_n^m \quad (4)$$

其中, \mathbf{W}_n^m 为待求的输出权重。

为求解 \mathbf{W}_n^m ,构造正则化最小二乘目标函数:

$$\arg \min_{\mathbf{W}_n^m} \Theta = \|\mathbf{A}_n^m \mathbf{W}_n^m - \mathbf{Y}\|_2^2 + \lambda \|\mathbf{W}_n^m\|_2^2 \quad (5)$$

其中,第一项是最小二乘误差项,第二项是正则化项, λ 是正则化系数,用于防止过拟合。设式(5)对 \mathbf{W}_n^m 的导数为 0,可以求得该优化问题的解析解如下:

$$\mathbf{W}_n^m = (\lambda \mathbf{I} + (\mathbf{A}_n^m)^T \mathbf{A}_n^m)^{-1} (\mathbf{A}_n^m)^T \mathbf{Y} \quad (6)$$

其中, \mathbf{I} 是单位矩阵。

在预测阶段,把测试样本 $x \in \mathbb{R}^{1 \times M}$ 输入训练良好

的模型,其对应的输出结果为

$$\mathbf{f}(x) = \mathbf{A}_x \mathbf{W}_n^m \quad (7)$$

其中, $\mathbf{A}_x \in \mathbb{R}^{1 \times l}$ 是测试样本 x 通过特征映射(1)和特征增强(2)得到的扩展特征(3)。 $\mathbf{f}(x) = [f_1(x), f_2(x), \dots, f_L(x)] \in \mathbb{R}^{1 \times (L+1)}$ 为输出向量。

最后, x 的预测标签为

$$\text{label}(x) = \arg \max_i f_i(x), i = 1, 2, \dots, L \quad (8)$$

2.2 基于BLS-FDDA的域适应算法

针对部分场景下原始ECG数据难以直接变换,如数据隐私限制、原始特征冗余的问题,本文提出BLS-FDDA算法。核心思路是在BLS的特征扩展矩阵 \mathbf{A} 层面进行分布对齐,利用特征更具判别性、冗余信息少的特点,实现更精准的域偏移校正,且无需修改BLS的基础网络结构。BLS的特征扩展矩阵 \mathbf{A} 融合了ECG信号的低维判别特征与高维扩展特征,其分布差异能更直接地反映域偏移对分类任务的影响。BLS-FDDA通过校准源域与目标域特征扩展矩阵的数据分布,对齐两者的特征空间,使源域训练的权重能直接适配目标域数据,避免原始数据变换带来的信息失真。

设源域数据为 $\mathbf{X}_S \in \mathbb{R}^{n \times m}$, 对应标签为 $\mathbf{Y}_S \in \mathbb{R}^{n \times L}$, 目标域数据为 $\mathbf{X}_T \in \mathbb{R}^{n' \times m}$, 对应标签为 $\mathbf{Y}_T \in \mathbb{R}^{n' \times L}$ 。根据式(1)~(3)可以求得源域与目标域的特征扩展矩阵分别为 \mathbf{A}_S 与 \mathbf{A}_T 。域适应的核心是对齐源域与目标域的特征分布,通常 \mathbf{C}_S 和 \mathbf{C}_T 为特征扩展矩阵的协方差矩阵,即标准化分布:

$$\mathbf{C}_i = \frac{1}{n-1} \mathbf{A}_i^T \mathbf{A}_i, i \in \{S, T\} \quad (9)$$

协方差矩阵的平方根逆 $\mathbf{C}_S^{-1/2}$ 可标准化源域分布, $\mathbf{C}_T^{1/2}$ 可将标准化后的分布映射到目标域分布,因此,域适应后的源域分布对齐变换可以表示为

$$\hat{\mathbf{A}}_S = \mathbf{A}_S \mathbf{C}_S^{-1/2} \mathbf{C}_T^{1/2} \quad (10)$$

由式(6)可知,初始源域的网络权重可以表示为 \mathbf{W}_S , 域适应后的理论网络权重表示为 $\hat{\mathbf{W}}_S$, 设 $\mathbf{M} = \mathbf{C}_S^{-1/2} \mathbf{C}_T^{1/2}$ 代入特征扩展矩阵的变换关系得

$$\hat{\mathbf{W}}_S = (\lambda \mathbf{I} + \mathbf{M} \mathbf{A}_S^T \mathbf{A}_S \mathbf{M})^{-1} \mathbf{M} \mathbf{A}_S^T \mathbf{Y}_S \quad (11)$$

进一步化简可得

$$\hat{\mathbf{W}}_S = \mathbf{M}^{-1} \mathbf{W}_S \quad (12)$$

即当需要对模型进行域适应时,只需对训练好的模型进行域适应变换即可,而无需重新训练。这大大提高了模型在域适应任务中的运算速度。

2.3 基于BLS-DRDA的域适应算法

为在ECG原始数据层面解决域偏移问题,本文提出一种基于宽度学习的可迁移数据对齐算法:BLS-DRDA。该方法在不改变既有BLS结构和参数的前提

下,从特征扩展矩阵出发,将域适应后的表示看作对原始表示的扰动,利用误差理论推导出源域与目标域之间的数据层变换关系,从而在信号层实现样本分布的快速对齐。得益于这一显式变换建模,面对新的数据域时,只需对源域数据和输出权值进行轻量更新即可完成域适应,无需对整个BLS模型重新训练,显著降低了跨域部署成本。

设源域数据为 $\mathbf{X}_S \in \mathbb{R}^{n \times m}$, 对应标签为 $\mathbf{Y}_S \in \mathbb{R}^{n \times L}$, 目标域数据为 $\mathbf{X}_T \in \mathbb{R}^{n' \times m}$, 对应标签为 $\mathbf{Y}_T \in \mathbb{R}^{n' \times L}$ 。基于样本分布进行数据对齐:

$$\mathbf{C}_i = \frac{1}{n-1} \mathbf{X}_i^T \mathbf{X}_i, i \in \{S, T\} \quad (13)$$

因此,域适应后的源域分布对齐变换可以表示为

$$\hat{\mathbf{X}}_S = \mathbf{X}_S \mathbf{C}_S^{-1/2} \mathbf{C}_T^{1/2} \quad (14)$$

通过式(1)~(3)求得 $\hat{\mathbf{A}}_S = [\hat{\mathbf{Z}}_S; \hat{\mathbf{H}}_S]$, 其中, $\hat{\mathbf{Z}}_S$ 为域适应后的特征节点, $\hat{\mathbf{H}}_S$ 为域适应后的增强节点。在计算特征节点和增强节点的同时,我们去除了偏置参数,由于数据已经分别进行中心化处理,这样做可以避免冗余参数的影响。

在求得 $\hat{\mathbf{A}}_S$ 后,即可根据式(6)求得域适应后的特征扩展矩阵:

$$\hat{\mathbf{W}}_S = (\lambda \mathbf{I} + \hat{\mathbf{A}}_S^T \hat{\mathbf{A}}_S)^{-1} \hat{\mathbf{A}}_S^T \mathbf{Y}_S \quad (15)$$

基于上述域适应过程,我们视 $\hat{\mathbf{A}}_S$ 为 \mathbf{A}_S 的估计量,则可以记扰动误差为

$$\mathbf{E} = \hat{\mathbf{A}}_S - \mathbf{A}_S \quad (16)$$

根据式(15)和式(16),结合矩阵求逆的扰动公式,可得

$$\begin{aligned} \hat{\mathbf{A}}_S^T \hat{\mathbf{A}}_S &= (\mathbf{A}_S + \mathbf{E})^T (\mathbf{A}_S + \mathbf{E}) \\ &= \mathbf{A}_S^T \mathbf{A}_S + \mathbf{A}_S^T \mathbf{E} + \mathbf{E}^T \mathbf{A}_S + \mathbf{E}^T \mathbf{E} \end{aligned} \quad (17)$$

$$\hat{\mathbf{A}}_S^T \mathbf{Y} = (\mathbf{A}_S + \mathbf{E})^T \mathbf{Y} = \mathbf{A}_S^T \mathbf{Y} + \mathbf{E}^T \mathbf{Y} \quad (18)$$

令 $\mathbf{P} = \lambda \mathbf{I} + \mathbf{A}_S^T \mathbf{A}_S$, 则

$$\hat{\mathbf{W}}_S = (\mathbf{P} + \mathbf{A}_S^T \mathbf{E} + \mathbf{E}^T \mathbf{A}_S + \mathbf{E}^T \mathbf{E})^{-1} (\mathbf{A}_S^T \mathbf{Y} + \mathbf{E}^T \mathbf{Y}) \quad (19)$$

利用扰动公式近似展开:

$$(\mathbf{P} + \mathbf{\Delta})^{-1} \approx \mathbf{P}^{-1} - \mathbf{P}^{-1} \mathbf{\Delta} \mathbf{P}^{-1} \quad (20)$$

其中, $\mathbf{\Delta} = \mathbf{A}_S^T \mathbf{E} + \mathbf{E}^T \mathbf{A}_S + \mathbf{E}^T \mathbf{E}$ 。将其代入式(19)可得

$$\hat{\mathbf{W}}_S = \mathbf{W}_S + \mathbf{P}^{-1} (\mathbf{E}^T \mathbf{Y}_S - (\mathbf{A}_S^T \mathbf{E} + \mathbf{E}^T \mathbf{A}_S) \mathbf{W}_S) \quad (21)$$

只需通过上式对 \mathbf{W}_S 进行变换即可实现数据对齐,而无需对整个模型进行重新计算。同理, \mathbf{W}_S 也可由 $\hat{\mathbf{W}}_S$ 通过变换得到。

为明确BLS-DRDA“无需重新训练”的严格适用边界,量化域偏移对模型性能的影响,基于上述推导补充误差界分析如下。

对于基于BLS-DRDA的域适应算法,设有测试样本 x , 其对应的真实标签 $y_{\text{true}} = k$, 原得分满足 $s_k >$

$s_j(\forall j \neq k)$, 定义最小得分余量:

$$\Delta_{\min} = s_k - \max_{j \neq k} s_j > 0 \quad (22)$$

标签错误的充要条件为:若扰动后的得分条件满足 $\hat{s}_k < \max_{j \neq k} \hat{s}_j$, 则预测标签错误。结合 $\hat{s} = s + \Delta s$, 等价于:

$$(\Delta s_j - \Delta s_k) > \Delta_{\min} (\exists j \neq k) \quad (23)$$

为避免标签错误,需保证得分扰动不抵消原得分余量,即

$$|\Delta s_j - \Delta s_k| \leq \Delta_{\min} \quad (24)$$

结合得分扰动的范数上界 $|\Delta s_j - \Delta s_k| \leq \|\Delta s\| \leq \|\hat{W} - W\| \cdot \|x\|$, 代入参数扰动的范数,得:

$$\left(\frac{\|A_s\| + \|E\|}{\lambda} + \|A_s^{-1}\| \right) \|Y_s\| \cdot \|x\| \leq \Delta_{\min} \quad (25)$$

解此不等式,得到 E 的 2-范数边界:

$$\|E\| \leq \lambda \frac{\Delta_{\min}}{\|Y_s\| \cdot \|x\|} - \lambda \|A_s^{-1}\| - \|A_s\| \quad (26)$$

该边界是充分条件,实际中可根据允许的标签误差进行调整。

2.4 代价敏感判决算法

尽管前述 BLS-DRDA 与 BLS-FDDA 方法能够有效缓解整体域偏移问题,但在心律失常识别任务中,由于部分类别样本数量极少、分布具有高度不均衡性,域适应后的分类模型仍然容易对该类样本产生决策偏差。特别是在长尾类别的识别中,模型往往倾向于将测试样本判为样本量更大的主导类别,从而导致这些少样本类别的分类准确率较低。为此,本文在决策层进一步设计一种代价敏感判决算法 (Cost-Sensitive Decision Algorithm, CSDA), 在保证整体分类性能的同时,重点提升易混淆类别与小样本类别的识别能力。

首先,为每一类别在源域特征空间中计算其类中心,以表征该类别的特征分布位置:

$$\mu_{A_s}^k = \frac{1}{n_{S,k}} \sum_{i=1}^{n_{S,k}} A_{S,i}^k \quad (27)$$

其中: $k=1, 2, \dots, L$ 表示第 k 类数据; $A_{S,i}^k$ 表示源域第 k 类的第 i 条特征向量。

对域适应后的每一个测试样本 \hat{A}_{test} , 计算其预测度,由式(8)可得:

$$y = \arg \max_i f_i(\hat{A}_{\text{test}}) \quad (28)$$

同时,计算该样本特征与训练数据第 k 类中心的欧氏距离:

$$d_k = \|\hat{A}_{\text{test}} - \mu_{A_s}^k\|_2 \quad (29)$$

距离越小,说明样本越接近该类别的特征簇中心,从几何结构上更可能属于该类别。为了解决类别不平衡对判决的影响,引入类别分布权重:

$$w_k = \frac{\sum_j n_{s,j}}{n_{s,k}} \quad (30)$$

当目标类为少数类时,赋予其决策更大的类别权重。基于此构造修正后的“加权距离”,则修正后的判决结果为

$$\hat{y} = \arg \min_k (w_k d_k) \quad (31)$$

从理论上讲,CSDA 是对 BLS 判决机制的修正,而非重新学习特征或模型参数。在标准 BLS 中,由于类别不平衡,模型训练时偏向多数类,判决边界倾向主导类别。CSDA 通过引入基于源域类别分布的加权类中心距离,对判决边界进行修正,从而有效减少类别偏倚,尤其在目标域样本稀缺或分布差异较大的情况下,依然能够保持稳定的识别性能。

3 实验

3.1 实验设置

实验设备包括 Intel(R) Core(TM) Ultra 9 275HX (2.70 GHz)、16 GB RAM 和 Nvidia GeForce GTX 5080 图形处理器。我们直接对获取的 ECG 信号进行心拍分割,依据公开数据库的 R 峰标记,从 R 峰前后分别取 119 个和 120 个采样点,共 240 个采样点作为一个心拍。通过将不同数据库中的数据按患者划分为不同的数据域,分别将不同的数据域设置为训练集和测试集进行训练和评估。

3.2 评价指标

为了详细评估各类数据的识别效果,本文采用了领域内的常用指标^[2,18]:真阳性 (TP)、真阴性 (TN)、假阳性 (FP)、假阴性 (FN)、准确度 (Accuracy, Acc)、精确度 (Precision, Ppv)、敏感度 (Sensitivity, Sen)、总体准确率 (Overall Accuracy, OA)、F1 分数 (F1_score) 和 G_mean 值。其中,OA 和总体 F1 分数 (Overall F1_score, OF1) 来衡量模型的总体识别性能:

$$OA = \frac{\sum_{i=1}^L TP_i}{\text{all ECG beats}} \times 100\% \quad (32)$$

$$OF1 = \frac{\sum_{i=1}^L F1_score_i}{L} \quad (33)$$

3.3 数据库

为验证算法效果,我们从 Physionet 数据库选取了两个公开数据集进行性能评估^[19],主要包括 MIT-BIH 数据库和 INCART 数据库,这些数据集数据量大且普遍存在显著的不平衡问题。根据 AAMI 推荐标准,样本被划分为正常心拍 (N)、房性早搏 (S)、室性早搏 (V) 和融合心拍 (F) 四个类别^[20]。MIT-BIH 数据库被划分为训练集 (DS1) 和测试集 (DS2), INCART 数据库作为

DS3,用于验证所提方法在跨数据集任务中的识别效果。数据集结构见表1。

表1 数据分布

Table 1 Data distribution

数据集	记录	N	S	V	F
DS1(MIT-BIH)	101, 106, 108, 109, 112, 114, 115, 116, 118, 119, 122, 124, 201, 203, 205, 207, 208, 209, 215, 220, 223, 230	45 850	944	3 788	414
DS2(MIT-BIH)	100, 103, 105, 111, 113, 117, 121, 123, 200, 202, 210, 212, 213, 214, 219, 221, 222, 228, 231, 232, 233, 234	44 244	1 837	3 220	388
DS3(INCART)	所有记录	150 330	3 172	20 005	1 942

4 实验结果与讨论

4.1 患者间域适应实验结果

为了测试所提算法在患者间域适应场景下的有效性,我们以DS1和DS2分别作为源域和目标域进行轮换测试。如图3所示,BLS-FDDA和BLS-DRDA在DS1为源域、DS2为目标域的实验,展现了针对患者间的优秀识别效果,并有效应对数据分布不平衡问题,识别了少数类。图4展示了DS2为源域、DS1为目标域的实验结果。综合图3和图4,原始BLS在DS1→DS2任务中存在类别混淆和少数类识别不足,而BLS-FDDA和BLS-DRDA通过数据对齐显著减少误判,提升准确率。在更具挑战性的DS2→DS1任务中,BLS-FDDA和BLS-DRDA仍保持优异表现,尤其是BLS-DRDA在两组跨域任务中实现了接近100%的识

别率。

结合表2和表3的实验指标可知,原始BLS在患者间域适应任务中表现不佳,其对少数类S、V、F的各项指标均表现不佳,在DS1→DS2中,S类的Ppv仅7.09%、F类的Sen仅0.52%;表3中三大总体指标也远低于改进方法,凸显其在域适应和类别不平衡场景下的局限性。

反观BLS-FDDA与BLS-DRDA,两者在表2的详细类别指标中,对N、S、V、F四类的Acc、Ppv、Sen、F1_score几乎均接近或达到100%,无论DS1→DS2还是DS2→DS1的域迁移方向,均展现出对各类别的优异识别能力。虽然所提方法总体表现优异,但通过观察表2可知,模型在V类上的表现略逊于其他三类。这个问题主要是由V类心跳形态的复杂性以及其与正常心跳的相似性所致。

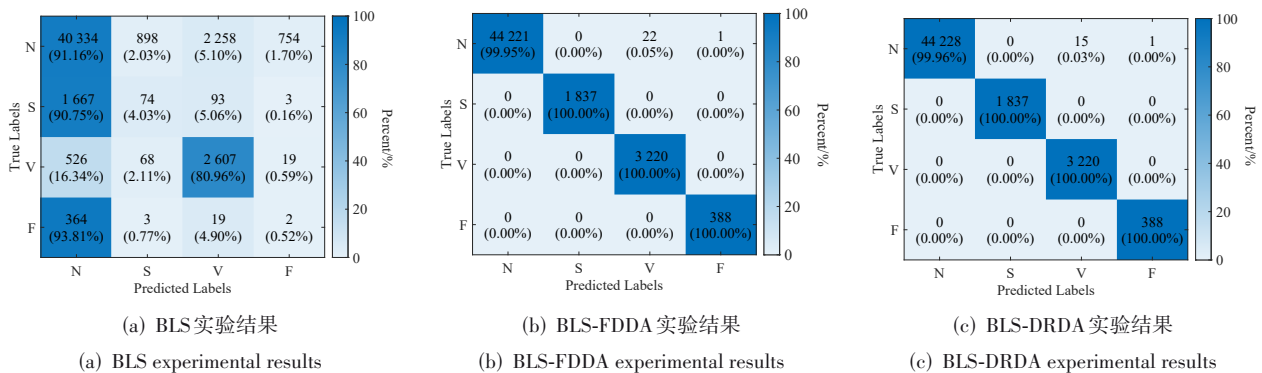


图3 DS1→DS2(以DS1为源域,DS2为目标域)实验结果

Figure 3 DS1→DS2 (With DS1 as the source domain and DS2 as the target domain) experimental results

表3所示的总体指标中,二者的OA、OF1、G_mean也都接近100%。实验结果表明,BLS-FDDA通过特征层对齐、BLS-DRDA通过数据重加权与代价敏感机制,均能有效解决域间特征差异与类别不平衡问题,在患者间域适应任务中实现了性能的大幅跃升,其中BLS-DRDA在总体指标的鲁棒性上表现更为突出。

表4展示了本文所提方法与三种模型的运行效率对比。可以看出,相较于传统域适应方法,基于

BLS的方法具备优秀的运算速度。基于BLS的训练架构避免了数据迭代的过程,大大减少了训练时长。同时,相较于BLS,BLS-FDDA提高了准确率,但训练时间增加;BLS-DRDA在保持较高准确率的同时,训练时间显著低于BLS-FDDA,表现出更好的性能与效率平衡。总体而言,BLS-FDDA和BLS-DRDA既保留了BLS的快速运算特性,又在准确率上实现了突破,其中BLS-DRDA的综合表现最为突出。

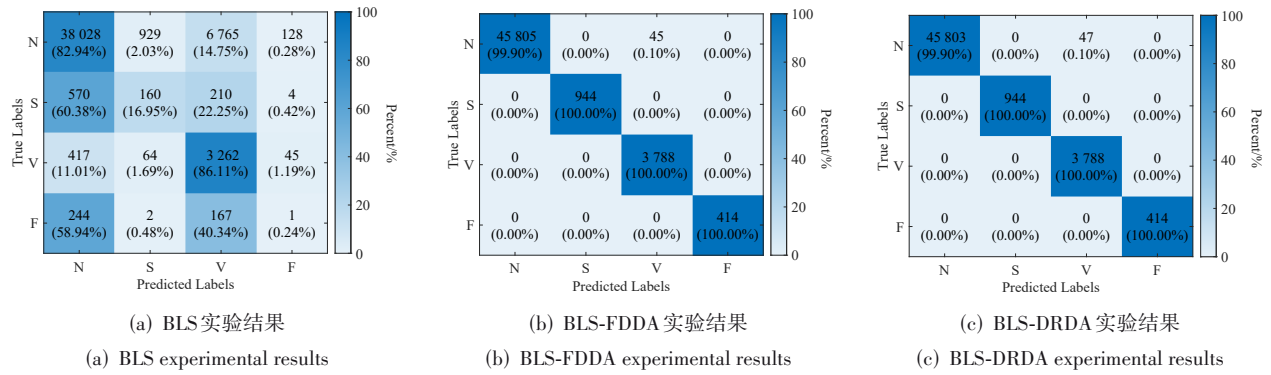


图 4 DS2→DS1(以 DS2 为源域,DS1 为目标域)实验结果

Figure 4 DS2→D1 (With DS2 as the source domain and DS1 as the target domain) experimental results

表 2 BLS 与所提方法在患者间域适应实验上的详细指标

单位:%

Table 2 Detailed metrics for BLS and the proposed method in inter-patient domain adaptation experiments

unit:%

方法	指标	DS1→DS2				DS2→DS1			
		N	S	V	F	N	S	V	F
BLS	Acc	86.99	94.50	94.00	97.66	82.25	96.51	84.96	98.84
	Ppv	94.04	7.09	52.38	0.26	96.86	13.85	31.35	0.56
	Sen	91.16	4.03	80.96	0.52	82.94	16.95	86.11	0.24
	F1_score	92.58	5.14	63.61	0.34	89.36	15.25	45.97	0.34
BLS-FDDA	Acc	99.95	100	99.96	100	99.91	100	99.91	100
	Ppv	100	100	99.32	99.74	100	100	98.83	100
	Sen	99.95	100	100	100	99.90	100	100	100
	F1_score	99.97	100	99.66	99.87	99.95	100	99.41	100
BLS-DRDA	Acc	99.97	100	99.97	100	99.91	100	99.91	100
	Ppv	100	100	99.54	99.74	100	100	98.77	100
	Sen	99.96	100	100	100	99.90	100	100	100
	F1_score	99.98	100	99.77	99.87	99.95	100	99.38	100

表 3 BLS 与所提方法在患者间域适应实验上的总体指标

单位:%

Table 3 Overall metrics of BLS and the proposed method in inter-patient domain adaptation experiments

unit:%

数据集	BLS			BLS-FDDA			BLS-DRDA		
	OA	OF1	G_mean	OA	OF1	G_mean	OA	OF1	G_mean
DS1→DS2	86.57	40.42	11.13	99.95	99.88	99.99	99.97	99.91	99.99
DS2→DS1	81.28	37.73	13.08	99.91	99.84	99.98	99.91	99.83	99.97

表 4 运行效率对比

Table 4 Comparison of computational efficiency

方法	数据集	OA/%	训练时长/s	测试时长/s
1DCNN-MMD	DS1→DS2	92.35	1 002	0.89
PCA-MMD-SVM	DS1→DS2	89.61	2 547	0.55
BLS	DS1→DS2	86.57	0.76	0.47
BLS-FDDA	DS1→DS2	99.95	3.83	0.64
BLS-DRDA	DS1→DS2	99.97	1.35	1.04

4.2 跨数据集域适应实验结果

为了进一步测试所提算法的域适应能力,我们采

用 DS1→DS3 和 DS2→DS3 两种跨数据库模式开展进一步实验,即以 MIT-BIH 数据库中的数据作为源域,

INCART 数据库中的数据作为目标域。

图 5 与图 6 展示了 DS1→DS3 与 DS2→DS3 两种跨数据库实验的混淆矩阵。原始 BLS 在跨数据库迁移中受到明显域偏移影响,尤其对 S、V、F 等少数类别的识别能力较弱。相比之下, BLS-FDDA 和 BLS-DRDA 的识别性能全面提升。在 DS1→DS3 实验中,

BLS-DRDA 的表现接近理想;在更具挑战性的 DS2→DS3 场景中, BLS-DRDA 对 V、F 等小样本类别仍保持 100% 的正确率,而 BLS-FDDA 在 DS2 数据质量较差的情况下性能有所下降。整体结果表明, BLS-FDDA 与 BLS-DRDA 能有效提升跨数据库的域适应能力, BLS-DRDA 在特征差异大、分布不平衡的情况下表现最优。

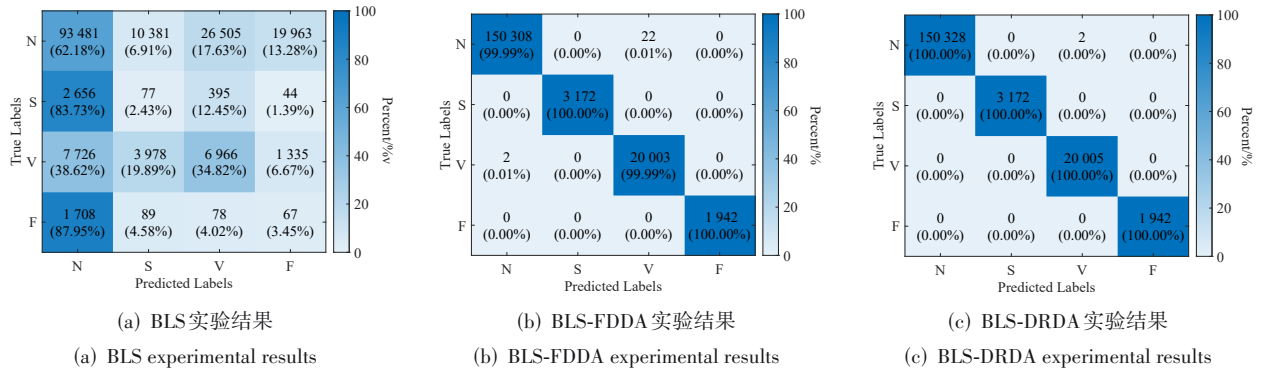


图 5 DS1→DS3(以 DS1 为源域,DS3 为目标域)实验结果

Figure 5 DS1→DS3 (With DS1 as the source domain and DS3 as the target domain) experimental results

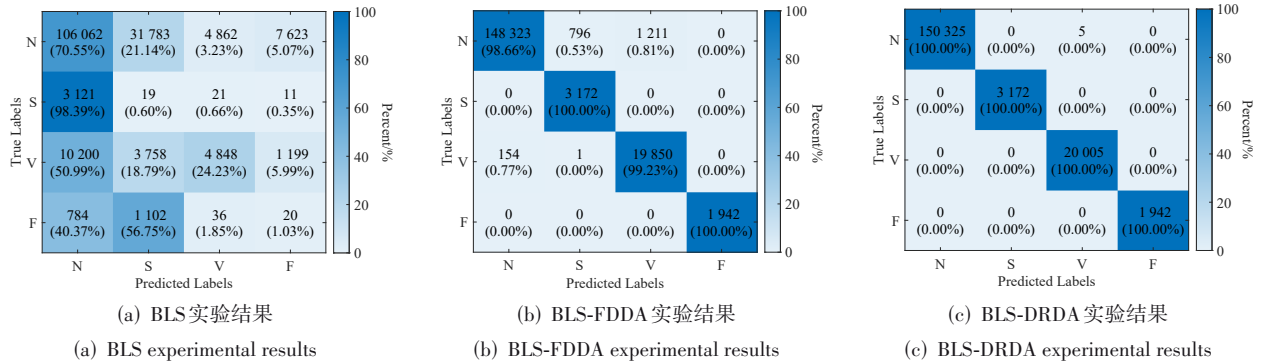


图 6 DS2→DS3(以 DS2 为源域,DS3 为目标域)实验结果

Figure 6 DS2→DS3 (With DS2 as the source domain and DS3 as the target domain) experimental results

综合图 5、图 6 及表 5、表 6 的实验结果可以明确:跨数据库迁移任务的挑战性显著高于同数据库迁移。在跨数据库患者间域适应场景中,原始 BLS 同时面临域偏移与类别不平衡的双重挑战,导致其性能存在显著短板。从表 5 的详细类别指标来看,该模型对 S、V、F 等少数类别的 Ppv、Sen 及 F1_score 均处于个位数甚至零点的极低水平;而表 6 的总体性能指标进一步验证了其局限性,即在两项跨数据库任务中,原始 BLS 的 OA 仅为 57.73% 和 63.24%,OF1 仅为 25.08% 和 27.86%,G_mean 更是低至 1.16% 和 0.57%,这充分暴露了其在复杂域适应场景下的适配能力不足。

与之形成鲜明对比的是,本文提出的 BLS-FDDA 通过引入特征层对齐机制,有效弥合了跨数据库间的特征分布差异,显著提升了域适应性能。表 5 的详细

类别指标显示, BLS-FDDA 对 N、S、V、F 四类目标的 Acc、Ppv、Sen 及 F1_score 几乎均接近或达到 100%;表 6 的总体指标同样印证了其优势,两项任务的 OA 分别达到 99.99% 和 98.77%,OF1 分别为 99.98% 和 96.20%,在缓解域偏移与类别混淆问题上成效显著。

表现更为突出的是 BLS-DRDA 模型:由表 5 可知,无论在 DS1→DS3 还是 DS2→DS3 迁移任务中,该模型对所有类别的 Acc、Ppv 及 F1_score 均达到 100%,仅在 DS2→DS3 任务中 V 类的 Ppv 为 99.98%。表 6 的总体性能指标进一步彰显了其卓越性能,充分展现了其对各类目标的精准识别能力与极强的环境鲁棒性。BLS-FDDA 在 DS2→DS3 任务中的效果较 DS1→DS3 任务有所下降,这是由于 DS1 的数据分布与 DS3 更接近,而 BLS-FDDA 在特征层面进行数据对齐时,能够

表 5 BLS 与所提方法在跨数据库域适应实验上的详细指标

单位:%

Table 5 Detailed metrics for BLS and the proposed method in cross-database domain adaptation experiments

unit:%

方法	指标	DS1→DS3				DS2→DS3			
		N	S	V	F	N	S	V	F
BLS	Acc	60.71	90.00	77.19	86.77	66.73	77.32	88.56	93.87
	Ppv	88.55	0.53	20.52	0.31	88.26	0.05	49.64	0.23
	Sen	62.18	2.43	34.82	3.45	70.55	0.60	24.23	1.03
	F1_score	73.06	0.87	25.82	0.57	78.42	0.10	32.57	0.37
BLS-FDDA	Acc	99.99	100	99.99	100	98.77	99.55	99.22	100
	Ppv	100	100	99.89	100	99.90	79.92	94.25	100
	Sen	99.99	100	99.99	100	98.66	100	99.23	100
	F1_score	99.99	100	99.94	100	99.28	88.84	96.67	100
BLS-DRDA	Acc	100	100	100	100	100	100	100	100
	Ppv	100	100	99.99	100	100	100	99.98	100
	Sen	100	100	100	100	100	100	100	100
	F1_score	100	100	100	100	100	100	99.99	100

更好地适应源域与目标域之间的特征差异。当源域与目标域的分布差异较大时,模型的对齐效果相对较

弱,导致在 DS2→DS3 任务中的性能下降,这在 S 类的表现上异常明显。

表 6 BLS 与所提方法在跨数据库域适应实验上的总体指标

单位:%

Table 6 Overall metrics for BLS and the proposed method in cross-database domain adaptation experiments

unit:%

数据集	BLS			BLS-FDDA			BLS-DRDA		
	OA	OF1	G_mean	OA	OF1	G_mean	OA	OF1	G_mean
DS1→DS3	57.73	25.08	1.16	99.99	99.98	99.99	100	100	100
DS2→DS3	63.24	27.86	0.57	98.77	96.20	99.47	100	100	100

4.3 连续域适应实验结果

通常而言,在域适应之后,数据的分布发生了改变,原有模型在源域上学习到的判别边界可能被削弱,甚至出现性能退化的风险。为了测试所提方法在连续域适应任务中的表现,本文构造了多阶段串联的域迁移场景,在每一阶段将前一阶段适配后的特征作为新的源域输入,并与单次域适应结果进行对比分析。

我们以 DS1 为源域,DS2 为第一目标域,DS3 为第二目标域进行连续域适应实验。图 7 展示了本文所提的两种方法在连续域适应任务中的表现。从混淆矩阵的可视化结果来看,与传统域适应方法在多阶段迁移中易出现的类别混淆、判别边界模糊问题不同,所提方法均保持了极高的分类准确率。对比图 7 与图 5 可知,所提方法有效抵御了连续域迁移带来的性能退化风险。

表 7 与表 8 分别从类别细分指标和总体指标层面提供了数据支撑。表 7 的类别指标显示,BLS-DRDA 在四类信号的 Acc、Ppv、Sen、F1_score 四项核心指标中均达到 100%,无论经过多少次域迁移,模型对各类别信号的识别精度、阳性预测值、敏感性及综合评价分数均未出现任何衰减;BLS-FDDA 虽存在微量错

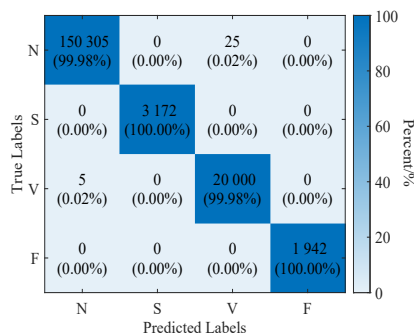
误,但 N 类、S 类、F 类的各项指标均保持在 99.95% 以上,与单次域适应(DS1→DS3)的结果几乎一致,未出现明显性能下滑。

表 8 的总体性能指标进一步验证了方法的优越性:BLS-FDDA 的 OA、OF1、G_mean 分别达到 99.98%、99.98%、99.99%,BLS-DRDA 则实现三项指标全 100%。对比单次域适应任务中两种方法的表现,连续域适应后的总体性能未出现任何衰减,反而因多阶段适配进一步贴合目标域分布,部分指标略有提升。

这一结果充分说明,本文所提的 BLS-FDDA 和 BLS-DRDA 方法不仅能够解决单阶段域分布差异问题,更能通过对特征分布的深度校准与判别信息的保留机制,有效规避连续域迁移中常见的“判别边界削弱”与“性能退化”问题。模型在多阶段域适配过程中,能够持续吸收新目标域的分布信息,同时保留源域与前序目标域的有效判别知识,具备优异的持续域泛化能力,为实际临床中多中心、多批次数据的自适应建模提供可靠支撑。

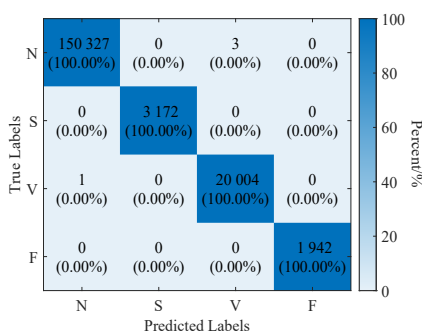
4.4 消融实验

为了验证所提方法中各个主要部分的贡献,我们针对 BLS-FDDA、BLS-DRDA 和 CSDA 在 DS1→DS2 任



(a) BLS-FDDA 实验结果

(a) BLS-FDDA experimental results



(b) BLS-DRDA 实验结果

(b) BLS-DRDA experimental results

图7 DS1→DS2→DS3(以DS1为源域,DS2为第一目标域,DS3为第二目标域)实验结果

Figure 7 DS1→DS2→DS3 (DS1 as the source domain, DS2 as the first target domain, DS3 as the second target domain) experimental results

表7 BLS与所提方法在连续域适应实验上的详细指标

单位:%

Table 7 Detailed metrics for the BLS and proposed method in continuous domain adaptation experiments unit:%

数据集	类别	BLS-FDDA				BLS-DRDA			
		Acc	Ppv	Sen	F1_score	Acc	Ppv	Sen	F1_score
DS1→	N	99.98	100	99.98	99.99	100	100	100	100
	S	100	100	100	100	100	100	100	100
DS2→	V	99.98	99.88	99.98	99.93	100	99.99	100	99.99
	F	100	100	100	100	100	100	100	100

表8 BLS与所提方法在连续域适应实验上的总体指标

单位:%

Table 8 Overall metrics for the BLS and proposed method in continuous domain adaptation experiments unit:%

数据集	BLS-FDDA			BLS-DRDA		
	OA	OF1	G_mean	OA	OF1	G_mean
DS1→DS2→DS3	99.98	99.98	99.99	100	100	100

务中开展了消融实验,实验结果如表9所示。

实验结果表明,CSDA对数据不平衡问题具有显著的矫正效果,且在基础BLS模型上,两种域适应方法(BLS-FDDA和BLS-DRDA)均借助CSDA实现了显著的性能提升。具体来看表9中的数据:BLS-DRDA without CSDA组有两个指标无法计算,其核心原因是该方法仅通过数据层面进行对齐,改变了原始数据的分布结构,进而导致模型对少数类别完全无法识别;但从总体准确率指标可以明确,数据对齐方法本身显著提升了模型在目标域的整体识别能力。

表9 消融实验结果的详细指标

单位:%

Table 9 Detailed metrics of the melting experiment results unit:%

方法	OA	OF1	G_mean
BLS without CSDA	86.57	40.42	11.13
BLS with CSDA	97.26	92.18	99.09
BLS-FDDA without CSDA	92.62	67.31	42.44
BLS-FDDA with CSDA	99.95	99.88	99.99
BLS-DRDA without CSDA	91.53	NaN	NaN
BLS-DRDA with CSDA	99.97	99.91	99.99

4.5 对比分析

为验证所提出模型的有效性,我们将其与近年来具有代表性的多种心搏分类方法进行了系统性对比,实验结果如表10所示。整体来看,以往研究在不同心搏类别上的表现差异明显,尤其是在应对类别不平衡和跨患者差异方面,仍存在不足。

以往的研究方法在心搏分类中虽有突破,但普遍面临类别不平衡、跨域适应不足及特征表达有限等挑战。Mondéjar-Guerra等人^[21]的多特征集成SVM方法

提高了正常心搏的识别能力,但对S类和F类的识别效果较差;Li等人^[22]构建的三维CNN模型改善了S类和V类的敏感度,但S类的阳性预测率仍低;Shi等人^[23]采用加权XGBoost方法提升了正常心搏和V类的预测率,但F类的识别仍有限;Zhai等人^[24]的半监督CNN具有较好的灵活性,但在F类的表现不理想;He等人^[25]的多级域自适应框架提高了整体准确率,但对于融合类仍不稳定;Liu等人^[26]结合SE-ResNet和专家特征取得较均衡表现,但S类与F类的识别仍

不足; Fan 等人^[18]提出的域自适应策略在检测室上性和室性早搏中表现出色,但整体类别均衡性仍欠缺。相比之下,本研究提出的 BLS-FDDA 与 BLS-DRDA 模型在各项指标上展现出显著优势,特别是在四类心搏

的阳性预测率和敏感度上接近完美,且具有较高的稳定性与泛化能力,综合表现优于现有主流方法,验证了其在心搏分类任务中的优越性及潜在临床应用价值。

表 10 在 MIT-BIH 数据集上的患者间域适应文献对比
Table 10 Comparison of inter-patient domain adaptation methods on the MIT-BIH dataset

单位:%
unit:%

方法	N		S		V		F		OA	OF1	G_mean
	Ppv	Sen	Ppv	Sen	Ppv	Sen	Ppv	Sen			
文献[21]	98.20	95.94	49.75	78.10	93.79	94.75	23.65	12.37	94.47	67.09	54.55
文献[22]	98.90	91.80	35.40	89.00	90.10	95.10	20.30	32.20	91.45	65.87	70.75
文献[23]	99.45	91.90	44.86	91.67	88.09	95.12	12.16	61.60	91.87	67.89	83.82
文献[24]	99.02	96.38	59.05	93.79	91.02	87.52	19.84	31.19	95.20	70.91	70.48
文献[25]	99.35	98.00	56.98	88.65	91.38	87.18	37.28	47.56	96.76	74.77	77.47
文献[26]	98.92	97.70	64.58	85.64	91.46	91.46	75.60	65.46	96.60	83.39	84.12
文献[18]	97.05	98.38	67.35	75.72	95.51	81.99	30.00	0.77	95.72	64.68	26.21
BLS-FDDA	100	99.95	100	100	99.32	100	99.74	100	99.95	99.88	99.99
BLS-DRDA	100	99.96	100	100	99.54	100	99.74	100	99.97	99.91	99.99

5 结论

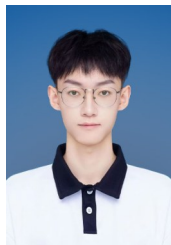
本文针对心律失常跨域识别中的域偏移与类别不平衡问题,提出了一套基于宽度学习系统的轻量级域适应与代价敏感决策框架。通过 BLS-FDDA 在特征层实现精确的源、目标域特征对齐,结合 BLS-DRDA 在数据层快速实现域适应,无需重新训练主模型。针对临床 ECG 数据中的类别不平衡问题,设计了 CSDA 算法,通过调整判决阶段的权重,显著提高了长尾类别和易混淆类别的识别能力。在 MIT-BIH 与 INCART 的多域、多数据库实验中,所提方法在患者间与跨数据库任务中均表现优异,取得接近 100% 的 OA、OF1 和 G_mean,相较于原始 BLS 及其他方法,在总体性能和少数类识别方面均有显著提升,尤其在准确率与运行效率平衡上,BLS-DRDA 展现了突出表现,验证了方法在复杂域偏移和严重类别不平衡情况下的鲁棒性与实用性。

参考文献

- [1] Jaiswal V, Kumar D, Mashkoor Y, et al. Demographics and trends of sudden cardiac death-related mortality in the United States, 1999 to 2022[J]. Journal of the American Heart Association, 2025, 14(18): e040340.
- [2] 吕杭, 蒋明峰, 李杨, 等. 基于混合时频域特征的卷积神经网络心律失常分类方法的研究[J]. 电子学报, 2023, 51(3): 701-711.
Lü Hang, Jiang Mingfeng, Li Yang, et al. Research on arrhythmia classification by using convolutional neural network with mixed time-frequency domain features[J]. Acta Electronica Sinica, 2023, 51(3): 701-711. (in Chinese)
- [3] Xue Linlin, Huang Yukun, Wang Chenglong, et al. Cross-domain fault diagnosis of RV reducer rolling bearings based on relative shaft synchronous resampling and adaptive gradient freezing transfer learning[J]. Advanced Engineering Informatics, 2026, 69: 103987.
- [4] 王红宇, 曹春歌. 人工智能心电图拓展临床诊疗的新视野[J]. 中国心血管病研究, 2025, 23(6): 485-489.
Wang Hongyu, Cao Chungu. Artificial intelligence electrocardiograms have expanded new horizons in clinical diagnosis and treatment[J]. Chinese Journal of Cardiovascular Research, 2025, 23(6): 485-489. (in Chinese)
- [5] Malik J, Devocioglu O C, Kiranyaz S, et al. Real-time patient-specific ECG classification by 1D self-operational neural networks[J]. IEEE Transactions on Biomedical Engineering, 2022, 69(5): 1788-1801.
- [6] Philip Chen C L, Liu Zhulin. Broad learning system: An effective and efficient incremental learning system without the need for deep architecture[J]. IEEE Transactions on Neural Networks and Learning Systems, 2018, 29(1): 10-24.
- [7] Yang Le, Song Shiji, Philip Chen C L. Transductive transfer learning based on broad learning system[C]//2018 IEEE International Conference on Systems, Man, and Cybernetics. Piscataway: IEEE, 2018: 912-917.
- [8] 吴新根, 吕维雪, 罗立民. 一种改进的 ECG 分类神经网络方法[J]. 电子学报, 1997, 25(10): 44-47.
Wu Xingeng, Lü Weixue, Luo Limin. An improved method for ECG classification based on neural network[J]. Acta

- Electronica Sinica, 1997, 25(10): 44-47. (in Chinese)
- [9] 潘杰, 刘波, 邹筱瑜. 基于特征异常检测与伪标签回归的无监督对抗域适应[J]. 电子学报, 2025, 53(1): 128-140.
Pan Jie, Liu Bo, Zou Xiaoyu. Feature anomaly detection and pseudo-label regression for adversarial domain adaptation[J]. Acta Electronica Sinica, 2025, 53(1): 128-140. (in Chinese)
- [10] 赵宇, 舒巧媛. 基于渐进式混合对比学习的无监督领域自适应行人再识别[J]. 电子学报, 2025, 53(6): 1829-1846.
Zhao Yu, Shu Qiaoyuan. Unsupervised domain adaptive person re-identification based on progressive hybrid contrastive learning[J]. Acta Electronica Sinica, 2025, 53(6): 1829-1846. (in Chinese)
- [11] 韩少勇, 周国华, 殷新春. 基于重构迁移子空间多视角领域适应的脑电情感分类方法[J]. 电子学报, 2025, 53(8): 2830-2842.
Han Shaoyong, Zhou Guohua, Yin Xinchun. Reconstructed transfer subspace based multi-view domain adaptation method for electroencephalogram emotion classification[J]. Acta Electronica Sinica, 2025, 53(8): 2830-2842. (in Chinese)
- [12] Kang J, Byun H, Kim T, et al. Multi-query frequency prompting for physiological signal domain adaptation[J]. Knowledge-Based Systems, 2025, 327: 114082.
- [13] Msigwa G, Ntambala E, Yun J. Enhancing atrial fibrillation classification from single-lead electrocardiogram signals using attention-based networks and generative adversarial networks with density-based clustering[J]. Engineering Applications of Artificial Intelligence, 2024, 133: 108607.
- [14] 韩昕哲, 尚莉伽, 张宏坡, 等. 基于MMD心电域适应学习的分类模型[J]. 计算机应用与软件, 2023, 40(1): 117-121.
Han Xinzhe, Shang Lijia, Zhang Hongpo, et al. Domain adaptive ECG classification model based on mmd[J]. Computer Applications and Software, 2023, 40(1): 117-121. (in Chinese)
- [15] Xia Heng, Tang Jian, Yu Wen, et al. Tree broad learning system for small data modeling[J]. IEEE Transactions on Neural Networks and Learning Systems, 2024, 35(7): 8909-8923.
- [16] Qin Aisong, Hu Qin, Zhang Qinghua, et al. A partial domain adaptation broad learning system for machinery fault diagnosis[J]. Measurement, 2025, 243: 116437.
- [17] Wang Haoyu, Cheng Yuhu, Wang Xuesong. A novel hyperspectral image classification method using class-weighted domain adaptation network[J]. Remote Sensing, 2023, 15(4): 999.
- [18] Fan Wei, Si Yujuan, Sun Meiqi, et al. Class-specific weighted broad learning system-based domain adaptation for patient-specific ECG classification[J]. Expert Systems with Applications, 2025, 273: 126824.
- [19] Goldberger A L, Amaral L A N, Glass L, et al. PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals[J]. Circulation, 2000, 101(23): E215-20.
- [20] Hu Y H, Palreddy S, Tompkins W J. A patient-adaptable ECG beat classifier using a mixture of experts approach[J]. IEEE Transactions on Biomedical Engineering, 1997, 44(9): 891-900.
- [21] Mondéjar-Guerra V, Novo J, Rouco J, et al. Heartbeat classification fusing temporal and morphological information of ECGs *via* ensemble of classifiers[J]. Biomedical Signal Processing and Control, 2019, 47: 41-48.
- [22] Li Feiteng, Xu Yin, Chen Zhijian, et al. Automated heartbeat classification using 3-D inputs based on convolutional neural network with multi-fields of view[J]. IEEE Access, 2019, 7: 76295-76304.
- [23] Shi Haotian, Wang Haoren, Huang Yixiang, et al. A hierarchical method based on weighted extreme gradient boosting in ECG heartbeat classification[J]. Computer Methods and Programs in Biomedicine, 2019, 171: 1-10.
- [24] Zhai Xiaolong, Zhou Zhanhong, Tin C. Semi-supervised learning for ECG classification without patient-specific labeled data[J]. Expert Systems with Applications, 2020, 158: 113411.
- [25] He Ziyang, Chen Yufei, Yuan Shuaiying, et al. A novel unsupervised domain adaptation framework based on graph convolutional network and multi-level feature alignment for inter-subject ECG classification[J]. Expert Systems with Applications, 2023, 221: 119711.
- [26] Liu Jinlei, Liu Yunqing, Jin Yanrui, et al. A novel diagnosis method combined dual-channel SE-ResNet with expert features for inter-patient heartbeat classification[J]. Medical Engineering & Physics, 2024, 130(1): 104209.

作者简介



冯建超 男,1999年7月生,云南祥云人。哈尔滨工业大学电子与信息工程学院博士研究生。主要研究方向为人工智能与机器学习、生物医学信号处理。

E-mail: 25B905065@stu.hit.edu.cn



李涵 男,2006年8月生,河南周口人。哈尔滨工业大学电子与信息工程学院本科生。主要研究方向为生物医学信号处理、毫米波雷达。

E-mail: lihan.hnzk@stu.hit.edu.cn



吴龙文 男,1988年9月生,江西樟树人。博士,哈尔滨工业大学高级工程师。主要研究方向为辐射源识别、辐射源个体识别、无源定位、医学信号处理。中国电子学会会员编号: E190036434M。

E-mail: wulongwen@hit.edu.cn



何欣 女,1994年1月生,黑龙江齐齐哈尔人。博士,哈尔滨工业大学副研究员。主要研究方向为遥感图像智能处理。在国内外发表学术论文120余篇。中国电子学会会员编号: E190035066M。

E-mail: hexin1@hit.edu.cn



王琪 男,1988年3月生,山东济宁人。中国兵器工业集团航空弹药研究院有限公司高级工程师,哈尔滨工业大学电子与信息工程学院博士研究生。主要研究方向为高超声速飞行器多物理场耦合力学。中国电子学会会员编号: E190000000N。

E-mail: genius@uestc.edu.cn



赵雅琴 女,1976年2月生,黑龙江兰西人。博士,哈尔滨工业大学教授。主要研究方向为辐射源识别、辐射源个体识别、无源定位、光通信、医学信号处理。中国电子学会会员编号: E190012844S。

E-mail: yaqinzhao@hit.edu.cn