

基于最优传输的多中心自闭症谱系障碍诊断

张俊艺, 万 鹏, 王明亮, 张道强

(南京航空航天大学计算机科学与技术学院, 南京, 211106)

摘 要: 融合来自多个中心的医学数据能够增加样本数量, 有助于研究自闭症谱系障碍 (Autism spectrum disorder, ASD) 的病理变化。因此, 如何有效地利用多中心数据, 提高对 ASD 诊断的准确性受到了越来越多的关注。然而, 以往的大部分研究忽略了多中心数据的异质性 (如受试者群体和扫描参数的不同), 这可能会降低模型在多中心数据上对疾病诊断的性能。为了解决这一问题, 提出一种基于联合分布最优传输 (Joint distribution optimal transport, JDOT) 的领域自适应模型鉴别 ASD。选择一个中心作为目标域, 其余的中心作为源域, 假设两个域的联合特征、标签空间分布之间存在非线性映射, 利用最优传输方法交替优化传输矩阵和分类器。结果表明, 在多中心静息态功能磁共振成像 (resting state functional magnetic resonance imaging, rs-fMRI) 数据上, 该模型能够有效提高对 ASD 鉴别的准确性。

关键词: 自闭症谱系障碍; 联合分布; 领域自适应; 多中心; 最优传输

中图分类号: TP391 **文献标志码:** A

Multi-center Autism Spectrum Disorder Diagnosis Based on Optimal Transport

ZHANG Junyi, WAN Peng, WANG Mingliang, ZHANG Daoqiang

(College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing, 211106, China)

Abstract: Effective fusion of medical data from multiple autism research centers contributes to the diagnosis of autism spectrum disorder (ASD), as large multi-site datasets increase the sample size, which facilitates the investigation of the pathological process of ASD. However, the existing methods generally ignore the heterogeneity (i.e., caused by subject populations and different scanning parameters) among diverse data sites, which degrades the effectiveness of model in ASD diagnosis based on multi-site datasets. To address this issue, we propose a novel domain adaptation method for ASD diagnosis based on joint distribution optimal transport (JDOT). Specifically, we alternately treat one site as target domain, and the rest are sources. Afterwards, we perform alignment in source-target domain by seeking a probabilistic coupling between joint feature and label distributions using optimal transport, which is optimized by an alternative minimization approach. Experimental results demonstrate the effectiveness of our method in ASD diagnosis based on multi-site resting-state functional magnetic resonance imaging (rs-fMRI) datasets.

Key words: autism spectrum disorder; joint distribution; domain adaptation; multi-site; optimal transport

引 言

自闭症谱系障碍(Autism spectrum disorder, ASD)是一系列复杂的神经发展障碍性疾病,主要症状表现为社交互动障碍、言语与非言语沟通困难以及局限性重复且刻板的行为。截至2015年,全球有6 220万例ASD病例^[1],但是ASD的病理生理学机制尚未明确^[2]。目前对ASD的诊断主要依靠临床特征,准确诊断ASD仍然是一项重要挑战。

静息态功能磁共振成像(Resting state functional magnetic resonance imaging, rs-fMRI)能够揭示大脑的固有自发活动规律、功能连接模式以及脑网络拓扑特征,已经被广泛应用于ASD的诊断。近年来,来自多中心的rs-fMRI数据量日益增长,如何有效地利用多中心数据,提高对ASD诊断的准确性受到了越来越多的关注。现有的方法^[3-6]一般是分别使用各个影像中心的数据,或者直接使用多个中心数据集的简单组合诊断ASD。但是,这些方法并没有考虑到自多个中心的数据存在异质性,这会直接影响对多中心疾病诊断的准确性。

领域自适应(Domain adaptation, DA)通过将源域知识迁移到同构的目标域,从而减少两个域之间的数据分布差异^[7],提高模型在目标数据集上的测试性能。现有的领域自适应方法可以分为半监督的领域自适应(Semi-supervised domain adaptation, SSDA)和无监督的领域自适应(Unsupervised domain adaptation, UDA)。半监督的领域自适应假设源域标签信息可得,并需要少量的目标域标签信息作为适配参照点。无监督的领域自适应不需要目标域的任何标签信息^[8],更符合实际应用的要求,因而获得了更为广泛的关注。Fernando等^[9]利用主成分分析(Principal component analysis, PCA)生成源域与目标域的子空间,通过最小化两个子空间的Bregman散度实现子空间对齐。Sun等^[10]通过源域协方差矩阵“白化”源域样本,然后用目标域协方差对其“重新着色”,利用源域与目标域的二阶统计量减小域间分布差异。近年来,无监督的领域自适应已经被成功应用到多中心医学数据的疾病分析。Wachinger等^[11]利用重加权的源域标记核磁共振成像(Magnetic resonance imaging, MRI)数据,训练适用于新数据集的分类器诊断阿尔兹海默症(Alzheimer's disease, AD)。Chambon等^[12]利用最优传输方法完成不同中心脑电图(Electroencephalography, EEG)数据的特征空间对齐,实现对睡眠障碍的诊断。但是,上述方法都隐含假设在源域数据的转换过程中不改变其标签信息。但是,没有明确的理由支持该假设成立。

基于联合分布最优传输(Joint distribution optimal transport, JDOT)的无监督领域自适应同时考虑了源域和目标域的边缘分布与条件分布^[13],并假设两个域的联合分布之间存在非线性映射,利用最优传输方法交替优化传输矩阵和分类器。通过样本对齐减少域间差异,提高分类器在目标域的性能。因此,可以利用JDOT减少多中心数据的异质性,提高分类器在目标中心对ASD鉴别的准确性。模型概述见图1。具体来说,首先对rs-fMRI图像进行预处理,然后选择一个中心作为目标域 X_t ,其余的中心作为源域 X_s ,假设目标域样本的标签信息未知,源域样本的标签信息已知;由源域样本训练获得分类器 f 的初始参数,用 f 生成目标域样本的初始标签。将分类器 f 嵌入到最优传输(Optimal transport, OT)求解器的代价函数中,求取传输矩阵 γ 。同时考虑两个域之间的样本距离、标签差异以及 f 的分类损失,由优化器交替优化传输矩阵 γ 和分类器 f 。Courty等^[14]证明了该优化问题等价于最小化分类器在目标域的误差边界。利用传输矩阵 γ 实现源域与目标域的非线性耦合,不仅将源域和目标域的样本对齐,也将源域的标签信息传递到目标域。最后,将已优化的分类器 f^* 用于目标域做最终的决策。

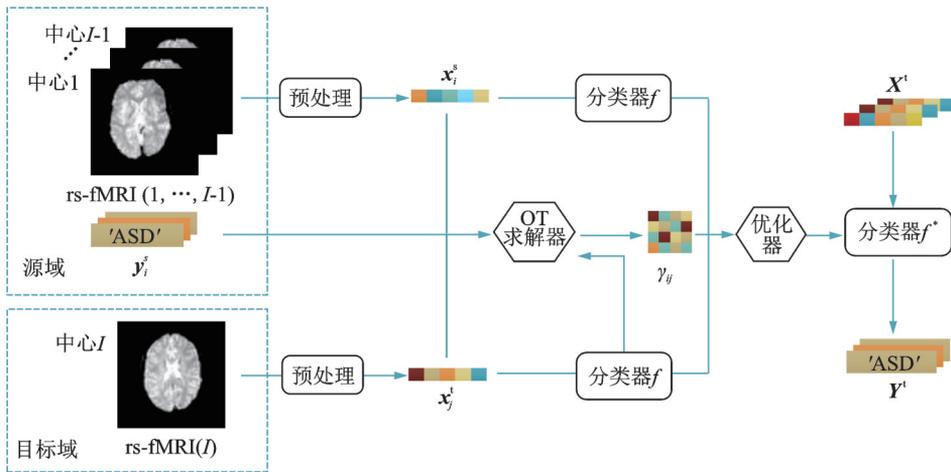


图1 JDOT模型

Fig.1 Overview of JDOT model

1 最优传输理论及方法

1.1 数据预处理

自闭症脑影像数据共享中心(Autism brain image data exchange, ABIDE)是一个大型的多中心自闭症数据库,共包含了17个来自不同中心的871例高质量rs-fMRI数据。由于部分中心数据量较少,选择了5个受试者数量超过50例的中心,包括Leuven, NYU, UCLA, UM和USM,共468例rs-fMRI数据。其中,ASD患者250例,正常对照者(Normal controls, NCs)218例,并且,各中心的患者和正常对照者比例相当。

通过 Preprocessed Connectomes Project (<http://preprocessed-connectomes-project.org>) 获取由脑连接组分析程序(Configurable pipeline for the analysis of connectome, C-PAC)预处理的rs-fMRI数据。如图2所示,对rs-fMRI图像的预处理过程包括时间层校正、头动校正和空间标准化,并通过干扰回归去除由扫描仪误差和受试者呼吸、心脏搏动引起的信号波动,然后将包含116个预定义感兴趣区域(Regions of interest, ROIs)的解剖学自动标记(Antomical automatic labeling, AAL)图谱对齐到每个图像上,提取每个ROI内的平均响应时间序列。最后通过计算ROI之间的皮尔逊相关系数获得功能连接矩阵,每条边的权值代表两个ROI之间的相关性。由于该矩阵具有对称性,即上三角部分与下三角部分的取值相同,而对角线上的取值仅表示每个ROI与自身的相关性。为了简单起见,选取功能连接矩阵去除对角线的下三角部分,并将其转换成向量作为样本特征。至此,对每一个受试者生成一个6 670维的特征向量。

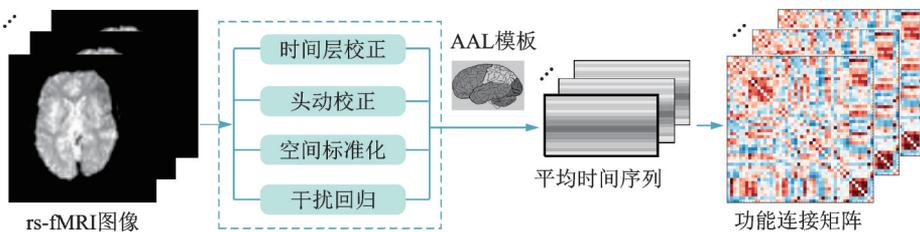


图2 rs-fMRI图像预处理

Fig.2 The preprocessing of rs-fMRI

1.2 基于联合分布最优传输的领域自适应方法

最优传输问题最早是由法国数学家蒙日(Monge)于17世纪80年代提出。定义 $\Omega \in \mathbf{R}^d$ 为一个 d 维的输入空间, C 为输出空间, $P(\Omega)$ 为输入空间 Ω 上的所有概率密度函数的集合, $u, v \in P(\Omega)$ 表示 Ω 上的边缘概率密度。Monge问题通过最小化概率密度 $u(x)$ 从 x 传输到 $T(x)$ 的代价,获得最优映射 T^* ,定义为式(1)。其中, $T\#u=v$ 表示映射 T 是保测度的,即总的概率密度相同,但是 T^* 一般不可解。

$$T^* = \operatorname{argmin}_T \int_{\Omega} D(x, T(x)) du(x) \quad \text{s.t. } T\#u=v \quad (1)$$

Kantorovitch问题可以看做是Monge问题的凸松弛^[15],求解空间 $\Omega \times \Omega$ 上所有边缘概率密度分别为 u, v 的最优联合概率耦合,定义为

$$\gamma^* = \operatorname{argmin}_{\gamma \in \Pi(u, v)} \int_{\Omega \times \Omega} D(x^u, x^v) d\gamma(x^u, x^v) \quad (2)$$

$$\text{s.t. } \Pi(u, v) = \{ \gamma \in P(\Omega \times \Omega) | \forall X_u, X_v \subset \Omega, P(X_u \times \Omega) = u, P(\Omega \times X_v) = v \}$$

假设存在一组标签信息已知的数据 $X_s = \{x_i^s\}_{i=1}^{N_s}, x_i^s \in \Omega$,相应的标签信息表示为 $Y_s = \{y_i^s\}_{i=1}^{N_s}, y_i^s \in C$;一组标签信息未知的数据 $X_t = \{x_i^t\}_{i=1}^{N_t}, x_i^t \in \Omega$,分别记作源域和目标域。联合分布最优传输同时考虑两个域的边缘分布和条件分布。 $P_s(X, Y), P_t(X, Y) \in P(\Omega \times C)$ 分别表示源域和目标域的联合特征、标签空间概率分布,通过优化式(3),在保持两个域的联合特征、标签空间概率密度不变的情况下,获得传输代价最小的联合分布传输矩阵,即

$$\gamma^* = \operatorname{argmin}_{\gamma \in \Pi(P_s, P_t)} D(x^s, y^s; x^t, y^t) d\gamma(x^s, y^s; x^t, y^t) \quad (3)$$

$$\text{s.t. } \Pi(P_s, P_t) = \{ \gamma \in P((\Omega \times C)^2) | \forall X \times Y \subset \Omega \times C, \gamma((X_s \times Y_s) \times (\Omega \times C)) = P_s, \gamma((\Omega \times C) \times (X_t \times Y_t)) = P_t \}$$

式中: $D(x^s, y^s; x^t, y^t) = \alpha d(x^s, x^t) + \lambda L(y^s, y^t)$ 是样本距离与标签差异的联合代价度量, d 表示平方欧氏距离, α 和 λ 分别是平衡特征空间和标签差异的正值参数。也就是说,当源域样本与目标域样本相近且标签相同时,耦合代价较低。

对ASD的鉴定是一个多源域的无监督领域自适应的分类问题,共有 I 个中心,每个中心都可以表示为有 N_i 个样本的集合 $X_i = \{x_j\}_{j=1}^{N_i}$,相应的标签表示为 $Y_i = \{y_j\}_{j=1}^{N_i}$ 。选择 X_t 作为目标域 $X_t = \{x_i^t\}_{i=1}^{N_t} = \{x_j\}_{j=1}^{N_t}$,标签为 $Y_t = \{y_i^t\}_{i=1}^{N_t} = \{y_j\}_{j=1}^{N_t}$,其余的中心作为源域 $X_s = \{x_i^s\}_{i=1}^{N_s} = \sum_{i=1}^{I-1} \{x_j\}_{j=1}^{N_i}$,标签为 $Y_s = \{y_i^s\}_{i=1}^{N_s} = \sum_{i=1}^{I-1} \{y_j\}_{j=1}^{N_i}$,假设目标域样本的标签信息 $Y_t = \{y_i^t\}_{i=1}^{N_t}$ 是未知的。考虑到最终的目的是学习一个适用于目标域的分类器 f ,可以用 $f(x^t)$ 替代 y^t ,将目标域的联合概率分布表示为 $P_t^f(X, f(X))_{X \sim v, o}$ 。一般取 P_s 和 P_t^f 的经验值 P_s^* 和 P_t^{f*} ,认为 $P_s^* = \frac{1}{N_s} \delta x_i^s, y_i^s, P_t^{f*} = \frac{1}{N_t} \sum_{i=1}^{N_t} \delta x_i^t, f(x_i^t)$ ^[14]。那么, γ 可看作是在两个均匀分布之间的非负耦合。在对 γ 优化的过程中,找到预测值与目标样本适配的源域样本标签最相近的分类器,Courty等^[14]证明了该优化问题等价于最小化分类器在目标域的误差边界。目标函数定义为

$$\min_{f, \gamma \in \Pi(P_s, P_t)} \sum_{i, j} D(x_i^s, y_i^s; x_j^t, f(x_j^t)) \gamma_{ij} \quad (4)$$

式(4)直接依赖于在目标域中学习的分类器,但是仅考虑分类器在目标域的预测性能,可能会导致分类器在源域表现不佳,这被称为“灾难性遗忘”^[16]。Damodaran等^[17]认为,可直接在式(4)中增加一项 f

在源域的分类损失,以避免这种退化。最终的目标函数定义为

$$\min_{f, \gamma \in \Pi(P_s, P_t)} \frac{1}{N_s} \sum_i L_s(y_i^s, f(x_i^s)) + \sum_{i,j} (\alpha \|x_i^s - x_j^t\|^2 + \lambda L_t(y_i^s, f(x_j^t))) \gamma_{ij} \quad (5)$$

式(5)中包含两组待优化的变量:传输矩阵 γ 和分类器 f , 可以采用交替最小化的方法优化该问题。当 f^* 确定时,对式(5)的求解就是代价函数为 $D_{ij} = \alpha \|x_i^s - x_j^t\|^2 + \lambda L_t(y_i^s, f^*(x_j^t))$ 的标准最优传输问题。当 γ^* 确定时,对 f 的优化就是传统的机器学习问题。为了使 JDOT 更好地嵌入到深度学习框架,分别从源域和目标域随机抽取大小为 m 的 mini-batch,将优化问题定义为式(6), E 为从源域和目标域随机抽取的 mini-batches 的期望值,即有

$$\min_f E \left[\frac{1}{m} \sum_i^m L_s(y_i^s, f(x_i^s)) + \min_{\gamma \in \Pi(P_s, P_t)} \sum_{i,j}^m (\alpha \|x_i^s - x_j^t\|^2 + \lambda L_t(y_i^s, f(x_j^t))) \gamma_{ij} \right] \quad (6)$$

对目标优化函数式(6)的求解过程如下:

(1) 固定分类器 f^* 的参数和每一个随机抽取的 mini-batch, 计算传输矩阵 γ 。传输矩阵 γ 不仅将源域和目标域的样本对齐,也将源域的标签信息传递到目标域。分类器 f 使用神经网络模型,初始参数由源域样本训练获得,即有

$$\min_{\gamma \in \Pi(P_s, P_t)} \sum_{i,j}^m (\alpha \|x_i^s - x_j^t\|^2 + \lambda L_t(y_i^s, f^*(x_j^t))) \gamma_{ij} \quad (7)$$

(2) 固定上一个步骤中计算获得的传输矩阵 γ^* ,对式(8)定义的 mini-batches 上的损失优化分类器 f ,更新 f 的参数,同时在源域和目标域训练分类器 f ,即有

$$\frac{1}{m} \sum_i^m L_s(y_i^s, f(x_i^s)) + (\alpha \|x_i^s - x_j^t\|^2 + \lambda L_t(y_i^s, f(x_j^t))) \gamma_{ij}^* \quad (8)$$

2 实验分析与结果

2.1 评估参数

对二分类模型进行性能评估时,通常会根据标签的实际值与预测值给出混淆矩阵的概念,由混淆矩阵就可以很容易地用数学公式描述各个评估参数^[18]。混淆矩阵见表1。基于 ABIDE 数据库中多个中心的 rs-fMRI 数据评估 JDOT 模型对 ASD 和正常人分类的性能,以分类准确性(Accuracy, ACC)、敏感性(Sensitivity, SEN)、特异性(Specificity, SPE)3个标准作为评估参数。

表1 混淆矩阵

Table 1 Confusion matrix

类别	预测阳性	预测阴性	合计
实际阳性	真阳性(TP)	假阴性(FN)	TP + FN
实际阴性	假阳性(FP)	真阴性(TN)	FP + TN
合计	TP + FP	FN + TN	TP + FN + FP + TN

ACC 是使用最普遍,也是最直观的性能评估参数,定义为预测正确的样本占全部样本的比例,见式(9)。准确性评估了模型的区分能力,但是在实际应用中,特别是在数据严重不平衡的情况下,只关注预测正确的结果可能会产生一定程度的误导性。因此,需要更有效的评估参数作为补充,以评估预测错误的结果需要承担的代价。

$$ACC = \frac{TP + TN}{TP + FN + FP + TN} \quad (9)$$

SEN定义为真阳性的样本占实际阳性样本的比例,见式(10)。在医学领域,敏感性又被称为真阳性率,是实际患病且被诊断为患病的概率,评估了模型对疾病的鉴别能力。敏感性越高,漏诊的概率越低。

$$\text{SEN} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (10)$$

SPE定义为真阳性的样本占预测阳性样本的比例,见式(11)。在医学领域,特异性又被称为真阴性率,是实际未患病(称之为正常)且被诊断为正常的概率,评估了模型对非患者的鉴别能力。特异性越高,误诊的概率越低。

$$\text{SPE} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (11)$$

2.2 实验设置与分析

在实验中设置分类器 f 的网络架构包括1个输入层、3个隐含层和1个输出层。输入层的维度为6 670;3个隐含层的维度分别为10,5,5,激活函数为ReLU;输出层的维度为2,激活函数为softmax。损失函数选择交叉熵损失,使用随机梯度下降优化算法更新模型参数。轮流选择1个中心作为目标域,其余的中心作为源域。JDOT模型的 L_s, L_t 均选择交叉熵损失函数^[17]。为方便起见,从源域和目标域随机抽取的mini-batch的大小设置为 $m = N_t$ 。设模型超参数 α, λ 的取值范围分别为 $\alpha = \{0.001, 0.01, 0.1, 0.5, 1\}, \lambda = \{0.001, 0.01, 0.1, 0.5, 1\}$ 。超参数 α, λ 的取值对JDOT模型性能的影响见图3。在取值范围内,当 α 取值为1时,JDOT模型在5个目标中心对ASD和NC分类的平均准确性、敏感性和特异性都达到了较优的水平,且当 λ 的取值为1时,模型的性能最优。

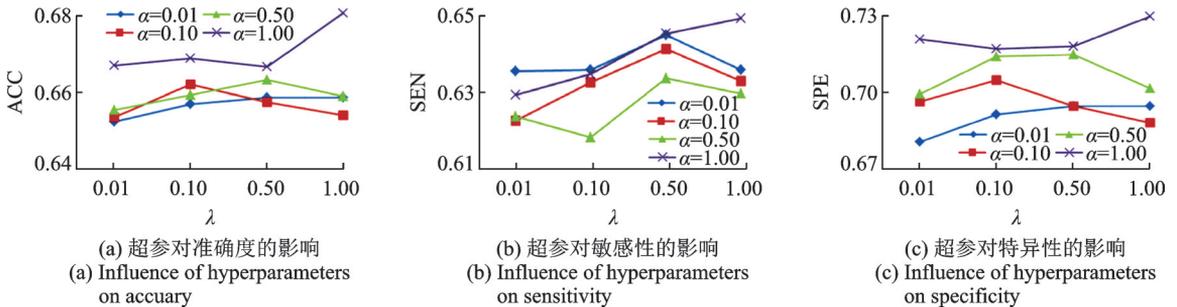


图3 超参数对模型性能的影响

Fig.3 Influence of hyperparameters on model performance

轮流在 Leuven, NYU, UCLA, UM 和 USM 中选择1个中心作为目标域,其余的中心作为源域。选择3种用于无监督领域自适应的模型和JDOT模型作横向比较,分别是基于子空间对齐(Subspace alignment, SA)^[9]的领域自适应、基于二阶统计量的领域自适应(Correlation alignment, CORAL)^[10]、基于最优传输的领域自适应(Optimal transport domain adaptation, OTDA)^[12]。其中,SA的子空间维度通过最小化子空间之间的Bregman散度来选择。模型在5个目标中心的分类准确性、敏感性和特异性分别见图4—6,性能评估参数值详见表2。相比之下,JDOT模型取得了最高的平均准确性(68.11%)。在对Leuven, NYU, UM, USM 4个中心鉴别ASD时,JDOT模型均取得了最高的准确性(分别是70.18%, 67.07%, 68.14%, 71.67%);在Leuven, NYU和USM 3个中心,模型均取得了最高的敏感性(分别是68.00%, 67.61%, 60.53%);在UM, USM两个中心,模型均取得了最高的特异性(分别是69.23%, 90.91%)。特别地,在对USM中心鉴别ASD时,模型取得了71.67%的准确性、60.53%的敏感性和90.91%的特异性,均高于其他领域自适应方法。综合上述实验结果,可以说明同时考虑联合特征、标签

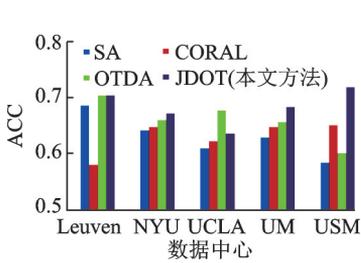


图4 模型的准确度

Fig.4 The accuracy of each model

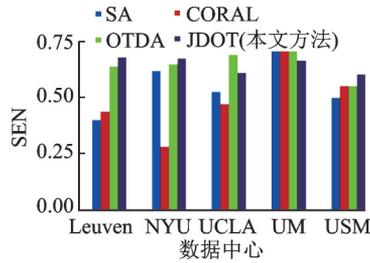


图5 模型的敏感性

Fig.5 The sensibility of each model

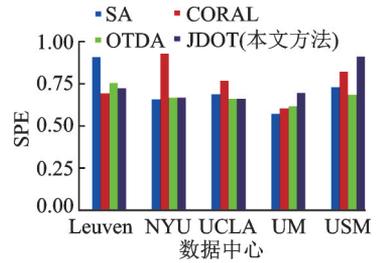


图6 模型的特异性

Fig.6 The specificity of each model

表2 模型在目标域的性能评估值

Table 2 Performance evaluation values for each model on target domains

目标域		Leuven	NYU	UCLA	UM	USM	平均值
准确性	SA	68.42	64.02	60.81	62.83	58.33	62.88
	CORAL	57.89	64.63	62.16	64.60	65.00	62.86
	OTDA	70.18	65.85	67.57	65.49	60.00	65.82
	JDOT(本文方法)	70.18	67.07	63.51	68.14	71.67	68.11
敏感性	SA	40.00	61.97	52.78	70.83	50.00	55.11
	CORAL	44.00	28.17	47.22	70.83	55.26	49.10
	OTDA	64.00	64.79	69.44	70.83	55.26	64.86
	JDOT(本文方法)	68.00	67.61	61.11	66.67	60.53	64.78
特异性	SA	90.63	65.59	68.42	56.92	72.73	70.86
	CORAL	68.75	92.47	76.32	60.00	81.82	75.87
	OTDA	75.00	66.67	65.79	61.53	68.18	67.43
	JDOT(本文方法)	71.88	66.67	65.79	69.23	90.91	72.90

空间的概率分布,利用最优传输方法,实现源域和目标域之间的非线性映射,能够在一定程度上减少域间差异,提高模型对多中心 ASD 鉴别的有效性。此外,OTDA 模型取得了仅次于 JDOT 模型的平均准确性(65.82%)和最高的平均敏感性(64.86%),在对 UCLA 中心鉴别 ASD 时,模型取得了最高的准确性(67.57%);在 UCLA 和 UM 两个中心,模型均取得了最高的敏感性(分别是 69.44%,70.83%)。这在一定程度上说明了最优传输方法对减少源域和目标域的数据差异、提高模型性能的有效性。

选择 3 种用于 ASD 鉴别的最先进的模型和 JDOT 模型作横向比较,分别是基于铰链损失和全局损失的图卷积网络(sGCN-1, sGCN-2)^[3]、基于 LOO 分类器的功能连通性关联分析(Functional connectivity association analysis with leave-one-out classifier, FCA)^[4]、以及基于两个自动编码器的去噪编码器(Denoising autoencoder, DAE)^[5]。以 Leuven, UCLA, UM 和 USM 4 个中心作为源域,使用上述 3 种方法对目标中心 NYU 进行 ASD 鉴别,得到的评估值见表 3。相比之下,JDOT 模型取得了更高的准确性(67.07%)和特异性(66.67%)。可

表3 ASD 鉴别模型的性能比较

Table 3 Performance comparison of models for ASD identification

方法	准确性	敏感性	特异性
sGCN-1	60.50	—	—
sGCN-2	63.50	—	—
FCA	63.00	72.00	58.00
DAE	66.00	66.00	65.00
JODT(本文方法)	67.07	67.61	66.67

见,分别使用各个影像中心的数据,或者直接使用多个中心数据集的简单组合诊断 ASD,会直接影响对多中心疾病诊断的准确性。而 JDOT 模型考虑到自多个中心的数据存在的异质性,利用最优传输方法,实现各个中心之间的非线性映射,能够在一定程度上减少域间差异,提高模型对多中心 ASD 鉴别的有效性。

3 结束语

本文提出一种基于联合分布最优传输的无监督领域自适应模型,利用多中心数据鉴别 ASD。具体来说,为了减少多中心数据的异质性,提高分类器在目标中心对 ASD 鉴定的准确性,首先在源域训练初始分类器,生成目标域样本的伪标签;然后利用最优传输方法,以两个域之间的样本距离和标签差异为代价,求解二者特征、标签空间的最优耦合,同时交替优化分类器,不仅将源域和目标域的样本对齐,也将源域的标签信息传递到目标域。为了避免灾难性遗忘导致分类器在源域的性能退化,在优化分类器时,同时考虑其在源域和目标域上的分类损失。最终获得的分类器能够有效提高在多中心 rs-fMRI 数据上鉴别 ASD 的准确性。

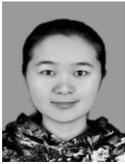
参考文献:

- [1] FERRÁN C L, HUTTON B. Risk of mortality among children, adolescents, and adults with autism spectrum disorder or attention deficit hyperactivity disorder and their first-degree relatives: A protocol for a systematic review and meta-analysis of observational studies[J]. *Systematic Reviews*, 2017, 6(1): 189.
- [2] WANG J, WANG Q, PENG J, et al. Multi-task diagnosis for autism spectrum disorders using multi-modality features: A multi-center study[J]. *Human Brain Mapping*. 2017, 38(6): 3081-3097.
- [3] K TENA S I, PARISOT S, FERRANTE E, et al. Metric learning with spectral graph convolutions on brain connectivity networks[J]. *NeuroImage*, 2017, 169: 431-442.
- [4] NIELSEN J A, ZIELINSKI B A, THOMAS F P, et al. Multisite functional connectivity MRI classification of autism: ABIDE results[J]. *Frontiers in Human Neuroscience*, 2013, 7(599): 1-12.
- [5] HEINSFELD A S, FRANCO A R, CRADDOCK R C, et al. Identification of autism spectrum disorder using deep learning and the ABIDE dataset[J]. *Neuroimage Clinical*, 2017, 17: 16-23.
- [6] WANG M L, ZHANG D Q, HUANG J S, et al. Low-rank representation for multi-center autism spectrum disorder identification[C]//*Proceedings of 21st International Conference on Medical Image Computing and Computer-Assisted Intervention*. Granada, Spain:[s.n.], 2018: 647-654.
- [7] PAN S J, YANG Q. A survey on transfer learning[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2009, 22(10): 1345-1359.
- [8] BENDAVID S, BLITZER J, CRAMMER K, et al. Analysis of representations for domain adaptation[C]//*Proceedings of 14th International Conference on Neural Information Processing Systems*. Kitakyushu, Japan:[s.n.], 2007: 137-144.
- [9] FERNANDO B, HABRARD A, SEBBAN M, et al. Unsupervised Visual domain adaptation using subspace alignment[C]//*Proceedings of 21st International Conference on Computer Vision*. Sydney, Australia: [s.n.], 2013: 2960-2967.
- [10] SUN B C, FENG J S, SAENKO K. Return of frustratingly easy domain adaptation[C]//*Proceedings of 13th AAAI Conference on Artificial Intelligence*. Arizona, USA:[s.n.], 2016.
- [11] WACHINGER C, REUTER M. Domain adaptation for Alzheimer's disease diagnostics[J]. *NeuroImage*, 2016, 139: 470-479.
- [12] CHAMBON S, GALTIER M N, Gramfort A. Domain adaptation with optimal transport improves EEG sleep stage classifiers[C]//*Proceedings of 8th International Workshop on Pattern Recognition in Neuroimaging*. Singapore:[s.n.], 2018: 1-4.
- [13] LONG M, WANG J, DING G, et al. Transfer feature learning with joint distribution adaptation[C]//*Proceedings of 21th International Conference on Computer Vision*. Sydney, Australia:[s.n.], 2013: 2200-2207.
- [14] COURTY N, FLAMARY R, HABRARD A, et al. Joint distribution optimal transport for domain adaptation[C]//*Proceedings of 24th International Conference on Neural Information Processing Systems*. Guangzhou, China:[s.n.], 2017: 3730-

3739.

- [15] KANTOROVICH L V. On the translocation of masses[J]. Journal of Mathematical Sciences, 2006, 133(4): 1381-1382.
- [16] SHMELKOV K, SCHMID C, ALAHARI K. Incremental learning of object detectors without catastrophic forgetting[C]// 23th International Conference on Computer Vision. Venice, Italy:[s.n.], 2017: 3400-3409.
- [17] DAMODARAN B B, KELLENBERGER B, FLAMARY R, et al. DeepJDOT: Deep joint distribution optimal transport for unsupervised domain adaptation[C]//Proceedings of 15th European Conference on Computer Vision. Munich, Germany:[s.n.], 2018: 467-483.
- [18] ZHU W, ZENG N, WANG N. Sensitivity, specificity, accuracy, associated confidence interval and ROC analysis with practical SAS implementations[J]. NESUG Proceedings: Health Care and Life Sciences, 2010, 19: 67.

作者简介:



张俊艺(1996-),女,硕士研究生,研究方向:机器学习和医学影像分析,E-mail: zhangjunyyyy@163.com。



万鹏(1993-),男,博士研究生,研究方向:机器学习、模式识别和医学影像分析。



王明亮(1989-),男,博士研究生,研究方向:机器学习、模式识别和数据挖掘。



张道强(1978-),通信作者,男,教授,博士生导师,研究方向:机器学习、模式识别以及数据挖掘和医学影像分析,E-mail: dqzhang@nuaa.edu.cn。

(编辑:刘彦东)