

doi: 10.11835/j.issn.1000-582X.2022.207

# 运动想象脑机接口的判别迁移特征学习与分类

齐 垒, 陈民铀, 张 莉

(重庆大学 电气工程学院, 重庆 400030)

**摘要:**为了解决不同时间采集的运动想象脑电数据之间存在的分布差异,避免跨时段使用前长时间的重校准步骤,提出了一种基于判别迁移特征学习(discriminative transfer feature learning, DTFL)的运动想象分类方法。DTFL通过联合匹配源域和目标域之间的边缘分布和类条件分布来减少域间的差异,同时最大化类间距离和最小化类内距离来保留类判别信息,从而提升对运动想象的分类性能。基于DTFL的运动想象分类方法无需目标域脑电样本的类别信息,可以有效避免长时间的校准。在脑机接口竞赛数据集上的实验结果表明,DTFL显著优于其他迁移学习方法,有效缓解跨域分布的不一致性,提高了运动想象的分类正确率。

**关键词:**脑机接口;运动想象;迁移学习

中图分类号:TP391

文献标志码:A

文章编号:1000-582X(2024)03-086-10

## Discriminative transfer feature for motor imagery brain-computer interfaces

QI Lei, CHEN Minyou, ZHANG Li

(School of Electrical Engineering, Chongqing University, Chongqing 400030, P. R. China)

**Abstract:** To address the cross-sessions variability of motor imagery electroencephalogram (EEG) and eliminate the need for lengthy recalibration step, this study proposes a motor imagery classification method based on discriminative transfer feature learning (DTFL). DTFL aims to reduce domain differences by jointly matching the marginal distribution and class conditional distribution of both domains. Simultaneously, DTFL maximizes inter-class dispersion and minimizes intra-class scatter, preserving class discrimination information and improving classification performance. This method does not require class information for EEG samples in the target domain, effectively avoiding the need for long-term calibration. Experimental results on brain-computer interface competition datasets demonstrate that, compared with some transfer learning methods, the proposed DTFL mitigates cross-session variability and improves the classification accuracy of motor imagery EEG.

**Keywords:** brain-computer interfaces; motor imagery; transfer learning

脑机接口(brain-computer interface, BCI)旨在建立人脑与计算机之间的直接连接<sup>[1]</sup>。BCI能够通过特定心理任务期间记录的脑电(electroencephalogram, EEG)信号中解码大脑活动模式,将神经反应转化为计算

收稿日期:2022-01-15 网络出版日期:2022-04-02

基金项目:国家自然科学基金资助项目(51977020)。

Supported by National Natural Science Foundation of China(51977020).

作者简介:齐垒(1995—),男,硕士研究生,主要从事脑机接口研究,(E-mail)ql\_qilei@163.com。

通信作者:陈民铀,男,教授,博士生导师,(E-mail)minyouchen@cqu.edu.cn。

机命令。其中,基于运动想象的BCI系统广泛应用在医疗康复中,使重度残疾的人有恢复控制外部环境的能力的希望<sup>[2-3]</sup>。因此,该研究方向在医疗康复领域有重大的意义。

EEG具有非线性、非平稳性的特性,导致在不同时间记录的运动想象数据的分布会存在一定的差异<sup>[4]</sup>。为解决该问题,保证系统的鲁棒性,传统的BCI系统需要在每次运行前进行校准。校准阶段需要花费一定的时间来获取足够的训练数据,对分类模型进行重训练,这对于患者和普通人来说都是很不方便的。迁移学习<sup>[5]</sup>可以利用源域中带有类别信息的数据将有效的知识迁移到未标记的目标域中,这为解决BCI系统的重校准问题,提供了一种有效的思路<sup>[6-7]</sup>。

目前,迁移学习工作侧重于重新加权源数据以得到目标数据的加权分类器<sup>[8-9]</sup>或提取领域不变特征以减轻源域和目标域之间的差异,以便有用的源域知识可以有效地迁移到目标域<sup>[10-12]</sup>。基于特征的迁移学习方法通常侧重于利用不同类型的度量方法来减少领域间的分布差异,常用的度量方法有KL散度<sup>[8]</sup>和最大均值差异(maximum mean discrepancy, MMD)<sup>[10-11]</sup>。虽然基于特征的迁移学习可以有效减小领域之间的分布差异,但是通过最小化度量距离学习到的域不变特征,往往会出现特征失真的情况。特征失真可能会损坏源域和目标域固有的内在分类结构,从而会降低对目标样本的识别性能<sup>[12-13]</sup>。因此,学习到的特征不仅要保持域不变性,确保2个域的分布很好地对齐,而且还应具有类判别性。为此,Deng等<sup>[14]</sup>提出通过在源数据中引入类别聚类来最小化域内方差。此外,Li等<sup>[12]</sup>提出通过同时最大化类间散度和最小化类内散度来提高域不变特征的判别性。

为解决BCI系统的重校准问题,文中提出了一种基于判别迁移特征学习(discriminative transfer feature learning, DTFL)方法。采用MMD来度量不同领域之间的分布差异,通过联合域适应<sup>[10]</sup>来对齐领域间的边缘分布和条件分布。不同于文献[12]所提的方法,文中提出了一个简单但有效的约束项来提高特征的判别性,该方法在学习域不变特征的同时,对2个领域中同类样本对的最大距离和不同类样本对的最小距离进行了最小化和最大化约束,在学习到特征同时具备了域不变性和判别性。

## 1 基于DTFL的运动想象分类方法

### 1.1 算法模型

如图1(a)所示,传统的运动想象分类方法主要包括训练和测试2个阶段,每个阶段包含2个步骤。在训练阶段:第一步,对训练集中的EEG数据进行8~30 Hz的带通滤波,再通过特征提取方法提取与运动想象任务相关的脑电特征,本研究中采用的特征提取方法为广泛使用的公共空间模式(common spatial pattern, CSP)<sup>[15]</sup>;第二步,利用训练集特征训练分类器,本研究中采用的分类方法为支持向量机(support vector machine, SVM)。在测试阶段:第一步,对测试集中的EEG数据进行带通滤波,再通过CSP提取脑电特征,这里的CSP是在训练阶段学习得到的。第二步,利用训练阶段获得的SVM模型预测测试集特征的运动想象任务类别。传统的运动想象分类方法并未考虑训练集特征和测试集特征之间的分布差异,这将导致分类器不能准确地分类测试集的特征,进而会影响运动想象分类的正确率。

如图1(b)所示,文中提出了基于特征迁移学习的运动想象分类方法,也由训练阶段和测试阶段组成。训练阶段不仅包含数据处理和特征提取的步骤,还包括了源域和目标域的特征迁移学习步骤,共有4个步骤。测试阶段在源域和目标域分布对齐的基础上进行运动想象分类。训练阶段:第一步,对源域和目标域的EEG数据分别进行8~30 Hz的带通滤波,再通过CSP分别提取源域和目标域的运动想象特征。第二步,源域和目标域的特征作为DTFL迁移学习方法的输入,共同学习特征变换矩阵。第三步,由特征变换矩阵对源域特征和目标域特征分别进行变换,得到分布对齐后的源域变换特征和目标域变换特征,再利用分布对齐后的源域变换特征作为训练集,对SVM进行训练。测试阶段:利用训练阶段获得的SVM模型预测目标域变换特征类别,同时更新目标域特征的伪标签。文中提出的基于DTFL的运动想象分类方法,利用迁移学习减小了运动想象特征之间的分布差异,不仅可以避免长时间的校准步骤,还能提高运动想象分类的正确率。在运动想象分类中,CSP和SVM得到了广泛的应用,文中不再对这2种方法进行赘述,将只对提出的DTFL迁移学习方法进行详细的介绍。

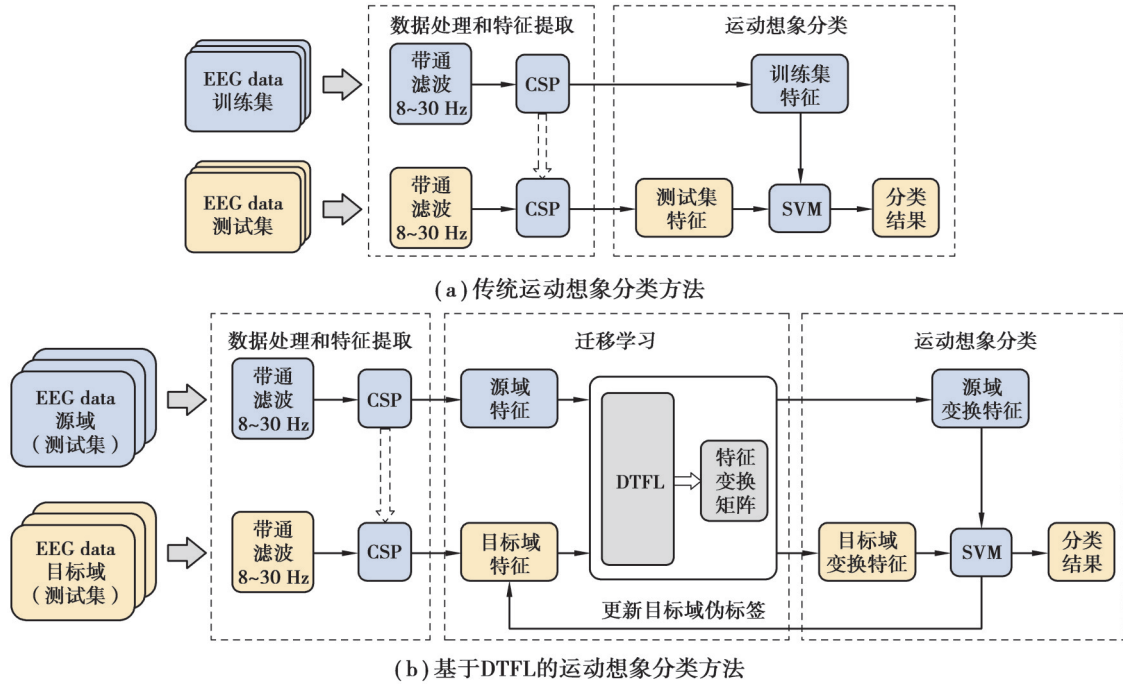


图1 运动想象分类方法流程图

Fig. 1 Flowchart of classification method of motor imagery

## 1.2 DTFL 方法

文中提出的DTFL方法由域不变特征学习和类判别特征学习2个部分组成。

### 1.2.1 迁移学习问题定义

带标签的源域数据  $\mathcal{D}_s = \{(\mathbf{x}_{si}, \mathbf{y}_{si})\}_{i=1}^{n_s} = \{\mathbf{X}_s, \mathbf{y}_s\}$  和无标签的目标域数据  $\mathcal{D}_T = \{\mathbf{x}_{Tj}\}_{j=1}^{n_t} = \{\mathbf{X}_T\}$  分别来自不同时段(比如不同天),  $n_s$  和  $n_t$  分别表示源域和目标域的样本数量。

### 1.2.2 域不变特征学习

域不变特征的学习由边缘分布对齐和类条件分布对齐2部分组成,首先,使用MMD来减小源域和目标域之间的边缘分布差异;其次,利用源域特征的真实标签和目标域特征的伪标签来改写MMD,进而减小两域之间的类条件分布差异。根据文献[10]中的方法,联合匹配MMD的计算公式为:

$$\mathcal{L}_{\text{MMD}} = \underbrace{\left\| \frac{1}{n_s} \sum_{i=1}^{n_s} \mathbf{P}^T \mathbf{x}_{si} - \frac{1}{n_t} \sum_{j=1}^{n_t} \mathbf{P}^T \mathbf{x}_{Tj} \right\|}_{\text{边缘分布差异}}^2 + \underbrace{\sum_{c=1}^C \left\| \frac{1}{n_s^{(c)}} \sum_{\mathbf{x}_{si} \in \mathcal{D}_s^{(c)}} \mathbf{P}^T \mathbf{x}_{si} - \frac{1}{n_t^{(c)}} \sum_{\mathbf{x}_{Tj} \in \hat{\mathcal{D}}_T^{(c)}} \mathbf{P}^T \mathbf{x}_{Tj} \right\|}_{\text{类条件分布差异}}^2 =$$

$$\text{Tr}(\mathbf{P}^T \mathbf{X} \mathbf{W}_0 \mathbf{X}^T \mathbf{P}) + \sum_{c=1}^C \text{Tr}(\mathbf{P}^T \mathbf{X} \mathbf{W}_c \mathbf{X}^T \mathbf{P}) =$$

$$\sum_{c=0}^C \text{Tr}(\mathbf{P}^T \mathbf{X} \mathbf{W}_c \mathbf{X}^T \mathbf{P}) =$$

$$\text{Tr}(\mathbf{P}^T \mathbf{X} \mathbf{W} \mathbf{X}^T \mathbf{P}),$$

式中,  $n_s^{(c)}$  和  $n_t^{(c)}$  分别为源域和目标域的  $c$  类样本数量。其中,当  $c=0$  时,  $n_s^{(0)} = n_s$ ,  $n_t^{(0)} = n_t$  分别为源域和目标域的整体样本数量,  $\mathbf{W}_c = \mathbf{W}_0$  为边缘MMD矩阵。当  $c \in \{1, \dots, C\}$ ,  $\mathbf{x}_{si}^{(c)} \in \mathcal{D}_s^{(c)}$  为源域的  $c$  类样本。 $\hat{\mathcal{D}}_T^{(c)} = \{(\mathbf{x}_{Tj}^{(c)}, \hat{\mathbf{y}}_{Tj}^{(c)})\}_{j=1}^{n_t^{(c)}} = \{\mathbf{X}_T^{(c)}, \hat{\mathbf{y}}_T^{(c)}\}$ ,  $\hat{\mathbf{y}}_{Tj}^{(c)}$  为目标域  $c$  类样本的伪标签,由源域样本训练的分类器获得。 $\mathbf{P}$  为变换矩阵,  $\mathbf{X} = [\mathbf{X}_s, \mathbf{X}_T] \in R^{m \times (n_s + n_t)}$ ,  $\mathbf{W} = \sum_{c=0}^C \mathbf{W}_c$ ,  $\text{Tr}(\cdot)$  为矩阵的迹运算。 $\mathbf{W}_c$  为  $c$  类的条件MMD矩阵,计算如下:

$$(W_c)_{ij} = \begin{cases} \frac{1}{n_s^{(c)} n_s^{(c)}}, & \mathbf{x}_i, \mathbf{x}_j \in \mathcal{D}_s^{(c)}; \\ \frac{1}{n_t^{(c)} n_t^{(c)}}, & \mathbf{x}_i, \mathbf{x}_j \in \hat{\mathcal{D}}_t^{(c)}; \\ \frac{-1}{n_s^{(c)} n_t^{(c)}}, & \begin{cases} \mathbf{x}_i \in \mathcal{D}_s^{(c)}, \mathbf{x}_j \in \hat{\mathcal{D}}_t^{(c)}; \\ \mathbf{x}_j \in \mathcal{D}_s^{(c)}, \mathbf{x}_i \in \hat{\mathcal{D}}_t^{(c)}; \end{cases} \\ 0, & \text{otherwise.} \end{cases} \quad (2)$$

综上所述,最小化式(1)中的 $\mathcal{L}_{\text{MMD}}$ ,可以有效地减小源域和目标域之间的分布差异,源域和目标域的新表征 $\mathbf{Z} = \mathbf{P}^T \mathbf{X}$ 将是域不变的。

### 1.2.3 类判别特征学习

最小化式(1)可以在潜在特征子空间中对齐源域和目标域的分布,但不能保证学习到的特征在分类任务中具有足够的判别性。为了解决特征失真引起的判别性下降问题,文中根据文献[12-13]的工作,提出在学习域不变特征的同时,通过最小化同类样本对的最大距离和最大化不同类样本对的最小距离,来保持域不变特征的判别性,如图2所示。相比文献[12]的约束方法,文中只关注最特殊的情况,这样计算更加高效。

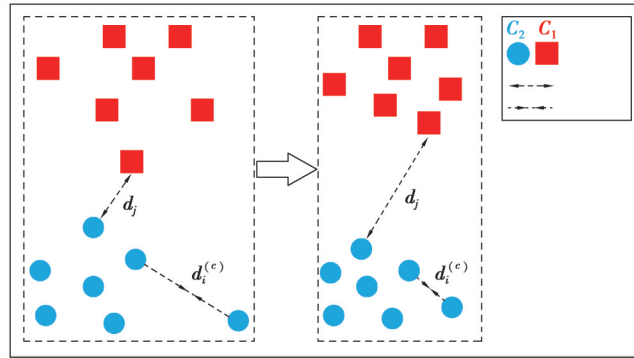


图2 源域样本对判别约束示意图

Fig. 2 Schematic diagram of source domain sample pair discrimination constraint

文中采用文献[13]中提到的方法来提高潜在空间中特征的判别能力。在源域中,距离计算公式为:

$$\begin{aligned} \mathcal{L}_{S, \text{distance}} &= \sum_{c=1}^C \arg \max_{\mathbf{x}_i, \mathbf{x}_j \in \mathcal{D}_s^{(c)}} \left\| \mathbf{P}^T \mathbf{x}_{si} - \mathbf{P}^T \mathbf{x}_{sj} \right\|^2 + \arg \min_{\mathbf{x}_i \in \mathcal{D}_s^{(c)}, \mathbf{x}_j \notin \mathcal{D}_s^{(c)}} \left\| \mathbf{P}^T \mathbf{x}_{si} - \mathbf{P}^T \mathbf{x}_{sj} \right\|^2 = \\ &= \text{Tr}(\mathbf{P}^T \mathbf{X}_s \mathbf{D}_{\text{same}}^{(S)} \mathbf{X}_s^T \mathbf{P}) + \text{Tr}(\mathbf{P}^T \mathbf{X}_s \mathbf{D}_{\text{diff}}^{(S)} \mathbf{X}_s^T \mathbf{P}) = \\ &= \text{Tr}[\mathbf{P}^T \mathbf{X}_s (\mathbf{D}_{\text{same}}^{(S)} - \mathbf{D}_{\text{diff}}^{(S)}) \mathbf{X}_s^T \mathbf{P}], \end{aligned} \quad (3)$$

式中: $\mathbf{D}_{\text{same}}^{(S)}$ 为相同标签下的距离矩阵; $\mathbf{D}_{\text{diff}}^{(S)}$ 为不同标签下的距离矩阵,可计算为:

$$(\mathbf{D}_{\text{same}}^{(S)})_{ij} = \begin{cases} I(\mathbf{x}_{si} \in \mathcal{D}_s^{(c)}) + \sum_{\mathbf{x}_{sj} \in \mathcal{D}_s^{(c)}} I(\mathbf{x}_{si} = \arg \max_{\mathbf{x}_{sk} \in \mathcal{D}_s^{(c)}} \left\| \mathbf{P}^T \mathbf{x}_{sq} - \mathbf{P}^T \mathbf{x}_{sk} \right\|^2), & i=j; \\ -I(\mathbf{x}_{sj} \in \mathcal{D}_s^{(c)}, \mathbf{x}_{si} = \arg \max_{\mathbf{x}_{sk} \in \mathcal{D}_s^{(c)}} \left\| \mathbf{P}^T \mathbf{x}_{sj} - \mathbf{P}^T \mathbf{x}_{sk} \right\|^2); \\ -I(\mathbf{x}_{si} \in \mathcal{D}_s^{(c)}, \mathbf{x}_{sj} = \arg \max_{\mathbf{x}_{sk} \in \mathcal{D}_s^{(c)}} \left\| \mathbf{P}^T \mathbf{x}_{si} - \mathbf{P}^T \mathbf{x}_{sk} \right\|^2), & i \neq j. \end{cases} \quad (4)$$

$$(\mathbf{D}_{\text{diff}}^{(S)})_{ij} = \begin{cases} I(\mathbf{x}_{si} \in \mathcal{D}_S^{(c)}) + \sum_{\mathbf{x}_{sq} \in \mathcal{D}_S^{(c)}} I(\mathbf{x}_{si} = \arg \max_{\mathbf{x}_{sk} \in \mathcal{D}_S^{(c)}} \|\mathbf{P}^T \mathbf{x}_{sq} - \mathbf{P}^T \mathbf{x}_{sk}\|^2), & i=j; \\ -I(\mathbf{x}_{sj} \in \mathcal{D}_S^{(c)}, \mathbf{x}_{si} = \arg \max_{\mathbf{x}_{sk} \in \mathcal{D}_S^{(c)}} \|\mathbf{P}^T \mathbf{x}_{sj} - \mathbf{P}^T \mathbf{x}_{sk}\|^2); \\ -I(\mathbf{x}_{si} \in \mathcal{D}_S^{(c)}, \mathbf{x}_{sj} = \arg \max_{\mathbf{x}_{sk} \in \mathcal{D}_S^{(c)}} \|\mathbf{P}^T \mathbf{x}_{si} - \mathbf{P}^T \mathbf{x}_{sk}\|^2), & i \neq j. \end{cases} \quad (5)$$

式中,  $I(g)$  为指示函数。

类似式(3), 利用目标域数据的伪标签得到目标域的距离计算公式为:

$$\mathcal{L}_{T, \text{distance}} = \text{Tr}[\mathbf{P}^T \mathbf{X}_T (\mathbf{D}_{\text{same}}^{(T)} - \mathbf{D}_{\text{diff}}^{(T)}) \mathbf{X}_T^T \mathbf{P}]. \quad (6)$$

由式(3)和式(6), 得到2个域的整体距离公式为:

$$\mathcal{L}_{\text{distance}} = \mathcal{L}_{S, \text{distance}} + \mathcal{L}_{T, \text{distance}} = \text{Tr}[\mathbf{P}^T \mathbf{X} (\mathbf{D}_{\text{same}} - \mathbf{D}_{\text{diff}}) \mathbf{X}^T \mathbf{P}] = \text{Tr}(\mathbf{P}^T \mathbf{X} \mathbf{D} \mathbf{X}^T \mathbf{P}), \quad (7)$$

式中,  $\mathbf{D}_{\text{same}} = \text{diag}(\mathbf{D}_{\text{same}}^{(S)}, \mathbf{D}_{\text{same}}^{(T)})$ ,  $\mathbf{D}_{\text{diff}} = \text{diag}(\mathbf{D}_{\text{diff}}^{(S)}, \mathbf{D}_{\text{diff}}^{(T)})$ ,  $\mathbf{D} = \mathbf{D}_{\text{same}} - \mathbf{D}_{\text{diff}}$  为距离矩阵。

最小化式(7)中的  $\mathcal{L}_{\text{distance}}$  可以同时最小化类内距离和最大化类间距离, 从而增强了学习到的域不变特征的判别能力。

#### 1.2.4 优化求解

DTFL的优化目标是学习到一个变换矩阵  $\mathbf{P}$ , 使得新表征  $\mathbf{Z} = \mathbf{P}^T \mathbf{X}$  同时具有域不变和类判别特性, 不仅能对齐源域和目标域的分布, 同时又能提高潜在子空间的最终分类性能。整合式(1)和式(7)得到最终的优化目标函数为:

$$\begin{cases} \min_{\mathbf{P}} \mathcal{L}_{\text{MMD}} + \alpha \mathcal{L}_{\text{distance}} + \beta \|\mathbf{P}\|_{\text{F}}^2, \\ \min_{\mathbf{P}} \text{Tr}[\mathbf{P}^T \mathbf{X} (\mathbf{W} + \alpha \mathbf{D}) \mathbf{X}^T \mathbf{P}] + \beta \|\mathbf{P}\|_{\text{F}}^2, \quad \text{s.t. } \mathbf{P}^T \mathbf{X} \mathbf{H} \mathbf{X}^T \mathbf{P} = \mathbf{I}_d. \end{cases} \quad (8)$$

式中:  $\alpha$  和  $\beta$  是2个权衡参数;  $\mathbf{I}_d$  是  $d$  维单位矩阵。约束条件里的  $\mathbf{H}$  是中心矩阵, 为

$$\mathbf{H} = \mathbf{I}_{n_s + n_t} - \frac{1}{n_s + n_t} \mathbf{I}_{(n_s + n_t) \times (n_s + n_t)}, \quad (9)$$

式中:  $\mathbf{I}_{n_s + n_t}$  是  $n_s + n_t$  维的单位矩阵;  $\mathbf{I}_{(n_s + n_t) \times (n_s + n_t)}$  是元素全为1的  $(n_s + n_t) \times (n_s + n_t)$  矩阵。式(8)中的约束条件源自主成分分析方法, 旨在最大化源域和目标域潜在子空间中的数据方差<sup>[10, 12]</sup>。

式(8)是有约束的非线性优化问题, 可以通过拉格朗日乘数法求解, 令对角矩阵  $\Phi = \text{diag}(\phi_1, \dots, \phi_d) \in \mathbb{R}^{d \times d}$  为拉格朗日乘子, 推导出式(8)的拉格朗日函数为:

$$L(\mathbf{P}, \Phi) = \text{Tr}\{\mathbf{P}^T [\mathbf{X} (\mathbf{W} + \alpha \mathbf{D}) \mathbf{X}^T + \beta \mathbf{I}_m] \mathbf{P}\} + \text{Tr}[(\mathbf{I}_d - \mathbf{P}^T \mathbf{X} \mathbf{H} \mathbf{X}^T \mathbf{P}) \Phi]. \quad (10)$$

令  $\frac{\partial L(\mathbf{P}, \Phi)}{\partial \mathbf{P}} = 0$ , 可得到一个经典的广义特征值分解问题:

$$[\mathbf{X} (\mathbf{W} + \alpha \mathbf{D}) \mathbf{X}^T + \beta \mathbf{I}_m] \mathbf{P} = \mathbf{X} \mathbf{H} \mathbf{X}^T \mathbf{P} \Phi. \quad (11)$$

最后, 将寻找最优矩阵  $\mathbf{P}$  的问题, 简化为求解式(11)的  $d$  个最小特征值问题。DTFL的具体算法步骤如表1所示。

类似文献[10-11], DTFL也可以通过核方法来解决非线性问题, 令  $\mathcal{H}$  表示再生核希尔伯特空间,  $\phi(\mathbf{X})$ :  $\mathbf{X} \rightarrow \mathcal{H}$  表示原始特征空间映射到再生核希尔伯特空间的映射函数, 源域和目标域的核矩阵利用核技巧可表示为  $\mathbf{K} = \phi(\mathbf{X})^T \phi(\mathbf{X}) \in \mathbb{R}^{(n_s + n_t) \times (n_s + n_t)}$ 。非线性版本的DTFL优化目标为:

$$\min_{\mathbf{P}_k} \text{Tr}[\mathbf{P}_k^T \mathbf{K} (\mathbf{W} + \alpha \mathbf{D}) \mathbf{K}^T \mathbf{P}_k] + \beta \|\mathbf{P}_k\|_{\text{F}}^2, \quad \text{s.t. } \mathbf{P}_k^T \mathbf{K} \mathbf{H} \mathbf{K}^T \mathbf{P}_k = \mathbf{I}_d. \quad (12)$$

与式(8)相似, 式(12)也可以转化为广义特征值分解问题来轻松求解。

## 2 实验研究

文中采用BCI Competition VI Dataset Ila<sup>[7]</sup>作为实验验证数据集, 验证文中所提出的迁移学习方法在解决

运动想象BCI重校准问题上的有效性,同时与联合分布适应(joint distribution adaptation, JDA)<sup>[10]</sup>和迁移成分分析(transfer component analysis, TCA)<sup>[11]</sup>2种先进的迁移学习方法进行比较。下面将详细介绍数据集和实验实现细节。

表1 DTFL的算法步骤

Table 1 Algorithm steps of DTFL model

**算法:**判别迁移特征学习(DTFL)

**输入:**带标签的源域样本  $\{X_S, Y_S\} = \{(x_{si}, y_{si})\}_{i=1}^{n_s}$ ;

不带标签的目标域样本  $\{X_T\} = \{x_{Tj}\}_{j=1}^{n_t}$ ;

权衡参数  $\alpha, \beta$ ;

迭代次数  $T$ ;

嵌入子空间的维数  $d$ 。

**输出:**映射矩阵  $P$ , 和最终目标域标签  $\hat{y}_T$ 。

- 1 开始
- 2 通过式(2)构造MMD矩阵  $W_0$ ;
- 3 初始化  $D_0 = 0_{(n_s + n_t) \times (n_s + n_t)}$ ;
- 4 循环
- 5 求解式(11)的广义特征值问题, 获得映射矩阵  $P$ ;
- 6 令  $[Z_S, Z_T] = P^T[X_S, X_T]$ , 在源域样本  $\{Z_S, Y_S\}$  上训练SVM, 预测目标域样本伪标签  $\hat{y}_T$ ;
- 7 更新矩阵  $W, D$ ;
- 8 循环结束 收敛或者达到最大迭代次数  $T$ 。
- 9 返回最终目标域标签  $\hat{y}_T$ 。
- 10 结束

## 2.1 实验数据和预处理

### 2.1.1 实验数据

BCI Competition VI Dataset Ila数据集由9名健康受试者的脑电数据组成。每名受试者被要求执行4类不同的运动想象任务,即想象左手、右手、双脚和舌头的运动。对于每名受试者,在不同的时间,记录了训练阶段和测试阶段2个时间段(session)的EEG数据。每个session由6组子实验组成,中间有短暂的休息,每组子实验包含48次运动想象任务(每类任务12次),即每阶段的实验可记录288个运动想象任务的实验数据。文中只采用了左右手2类运动想象任务的实验数据,每个session有144个EEG数据。脑电信号采用22个Ag/AgCl电极记录,采样频率为250 Hz,并进行了0.5~100 Hz带通滤波,50 Hz的陷波滤波。有关数据集的更多细节,请参阅网站 <http://www.bbc.de/competition/iv/>。

### 2.1.2 预处理

实验研究中对22导的EEG数据进行了预处理,截取任务提示后的0.5~2.5 s时间段的数据作为样本,再使用4阶Butterworth滤波器对EEG信号进行8~30 Hz的带通滤波。

## 2.2 实验对比设置

文中采用Ila数据集的9名受试者(A01, A02, ..., A09)的EEG数据做左/右手的二分类实验。在传统的运动想象分类中,每名受试者的session1的EEG数据集作为训练集,session2的EEG数据集作为测试集。为了验证文中提出的基于DTFL的运动想象分类方法(CSP+DTFL+SVM)的有效性,与以下3种方法进行了对比。

1) CSP+SVM: 该方法是传统的运动想象分类方法,使用session1的EEG数据训练模型,并在不进行任何迁移学习的情况下对session2的EEG数据进行分类。其中,采用CSP对EEG数据进行空间滤波和特征提取,使用线性SVM对EEG特征进行分类。

2)CSP+TCA+SVM:该方法包含特征迁移学习的步骤。首先,利用CSP提取EEG特征。然后,利用TCA对齐源域和目标域的边缘分布。最后,使用源域数据集(每名受试者的session1的EEG数据集)训练线性SVM,对目标域样本(session2的EEG数据)进行预测。

3)CSP+JDA+SVM:该方法也包含特征迁移学习的步骤。使用JDA同时对源域和目标域的边缘分布和类条件分布进行对齐。

在迁移学习过程中,TCA方法实现了对源域和目标域的边缘分布对齐,但缺乏对类条件分布的对齐,而JDA方法则在TCA方法的基础上进行了改进,增加了对源域和目标域的类条件分布的对齐,实现了联合分布对齐。但是EEG信号具有非线性、非平稳的特点,严格的联合分布对齐可能会导致特征的失真,降低特征的可区分性。文中提出的DTFL方法则很好地解决了此问题。DTFL方法是在JDA方法的基础上加入了对类内距和类间距的约束,不仅实现了对源域和目标域的联合分布对齐,同时保留了特征的可区分性,防止了特征的失真。

为了实验的公平性和有效性,3种基于迁移学习的运动想象分类方法的正确率均是在最优参数下得到的,文中采用网格搜索法进行寻优。所有实验均在Win10操作系统上完成,采用的软件是Matlab 2018b,电脑配置为Intel(R) Core(TM) i5-7200U,2.50 GHz,12 GB RAM。

### 2.3 实验结果分析

表2列出了4种算法在数据集II a上的分类正确率。表格中,列出了每一种方法对9名受试者的分类正确率和平均值。每行表示单个受试者在4种方法下的正确率,为方便比较,最高的分类正确率用粗体突出显示。结果显示,相较其他3种方法,文中所提出的方法在每一名受试者上都取得了最佳的分类正确率,在平均正确率上分别提高了8.2%、2.8%和4.8%。为进一步研究所提出的基于DTFL的方法与其他方法在正确率差异上的显著性( $p < 0.01$ ),进行了配对 $t$ 检验( $\alpha = 0.05$ )。在表2的最后一行给出了配对 $t$ 检验的结果,显著性结果用粗体突出显示,结果表明,所提出的基于DTFL的运动想象分类方法显著优于其他3种方法。

从表2中可以看出,增加了特征迁移学习步骤的方法(CSP+TCA+SVM、CSP+JDA+SVM、CSP+DTFL+SVM)取得的正确率均显著( $p < 0.03$ )高于不采用迁移学习的方法(CSP+SVM),在平均值上3种方法分别提高了5.4%、3.4%和8.2%。结果表明迁移学习可以有效减小2个session之间的分布差异,使得在源域上训练得到的分类器(SVM),也能很好地预测目标域的标签。

表2 4种算法的分类正确率

Table 2 Classification accuracy of the 4 algorithms

%

受试者	算法			
	CSP+SVM	CSP+TCA+SVM	CSP+JDA+SVM	CSP+DTFL+SVM
A01	84.0	83.3	86.8	<b>91.0</b>
A02	58.3	59.7	60.4	<b>66.7</b>
A03	92.4	94.4	93.8	<b>97.2</b>
A04	60.4	79.2	65.3	<b>80.6</b>
A05	56.3	65.3	65.3	<b>66.7</b>
A06	70.1	72.2	71.5	<b>73.6</b>
A07	78.5	78.5	81.3	<b>82.0</b>
A08	97.2	<b>97.9</b>	97.2	<b>97.9</b>
A09	80.6	<b>95.8</b>	86.8	<b>95.8</b>
平均值	75.3	80.7(↑5.4)	78.7(↑3.4)	<b>83.5(↑8.2)</b>
P值	<b>0.004</b>	<b>0.018</b>	<b>0.017</b>	

文中采用了  $t$ -SNE-可视化方法<sup>[16]</sup>,对受试者 1(subject 1, A01)和受试者 4(subject 4, A04)在 4 种方法下的特征分布进行了可视化,如图 3 和图 4 所示。图 3 和图 4 中,蓝色表示 session1(源域)的数据,其中,浅蓝色实心圆和三角形分别表示源域中的 2 类样本,深蓝色实心圆和三角形分别表示源域中的 2 类样本的类中心,深蓝色四边形表示源域的所有样本的中心,红色表示 session2(目标域)的数据。

图 3(a)是 CSP 提取的 subject 1 的脑电特征的分布图,图 3(b)、图 3(c)、图 3(d)是经过 CSP 特征提取之后,再利用 3 种不同方法对脑电特征进行迁移学习之后的特征分布图。图 3(b)中,TCA 有效的减小了 session1 和 session2 的样本中心之间的距离,源域和目标域的边缘分布得到了对齐,但是条件分布差异并未被减小,所以未能很好地提高分类正确率。图 3(c)、图 3(d)中,session1 与 session2 同类样本中心之间的距离和整体样本中心之间的距离都得到了减小。但是,相较于图 3(c),图 3(d)中的不同类的样本中心之间的距离更大,而且类内散度更小,说明了 DTFL 不仅具有 JDA 联合分布匹配的能力,还具有提高特征判别性的能力。因此,CSP+JDA+SVM 和 CSP+DTFL+SVM 虽然都能提高分类正确率,但 CSP+DTFL+SVM 的分类性能更佳。

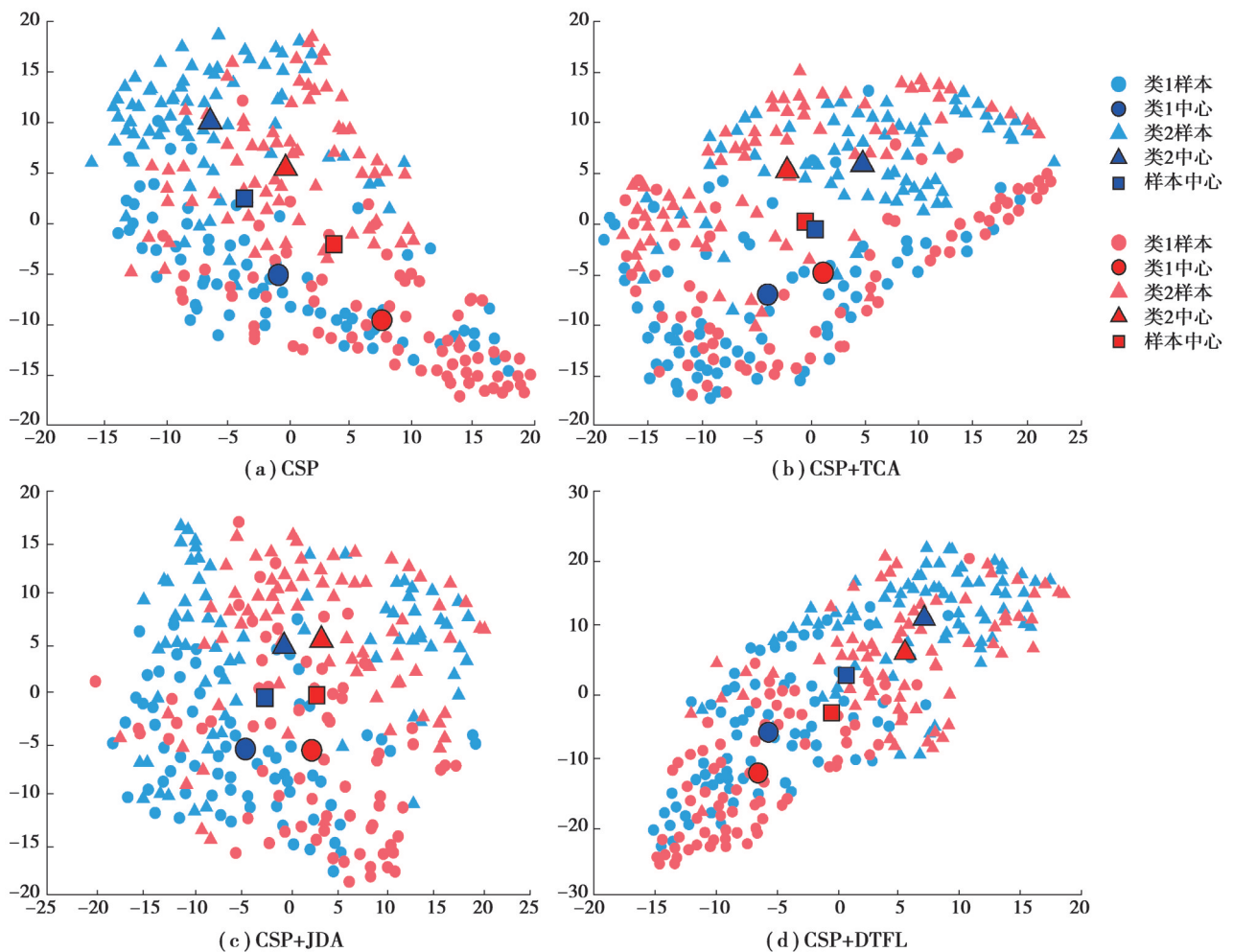


图 3 Subject 1 数据分布可视化图

Fig. 3 Visualization of data distribution on subject 1

如图 4(c)所示,JDA 虽然有效减小了 session1 和 session2 之间的条件分布差异和边缘分布差异,但是也减小了不同类样本之间的距离,削弱了特征的判别能力。相比之下,图 4(d)中不同类样本之间的距离并未被减小,进一步说明了 DTFL 不仅能减小域间的分布差异,更能保持特征的判别能力。从表 2 中的分类结果也能看出,CSP+DTFL+SVM 在 subject 4 上的分类正确率要比 CSP+JDA+SVM 高出 15.3%,更加证明了特征的判



别能力对分类的重要性。

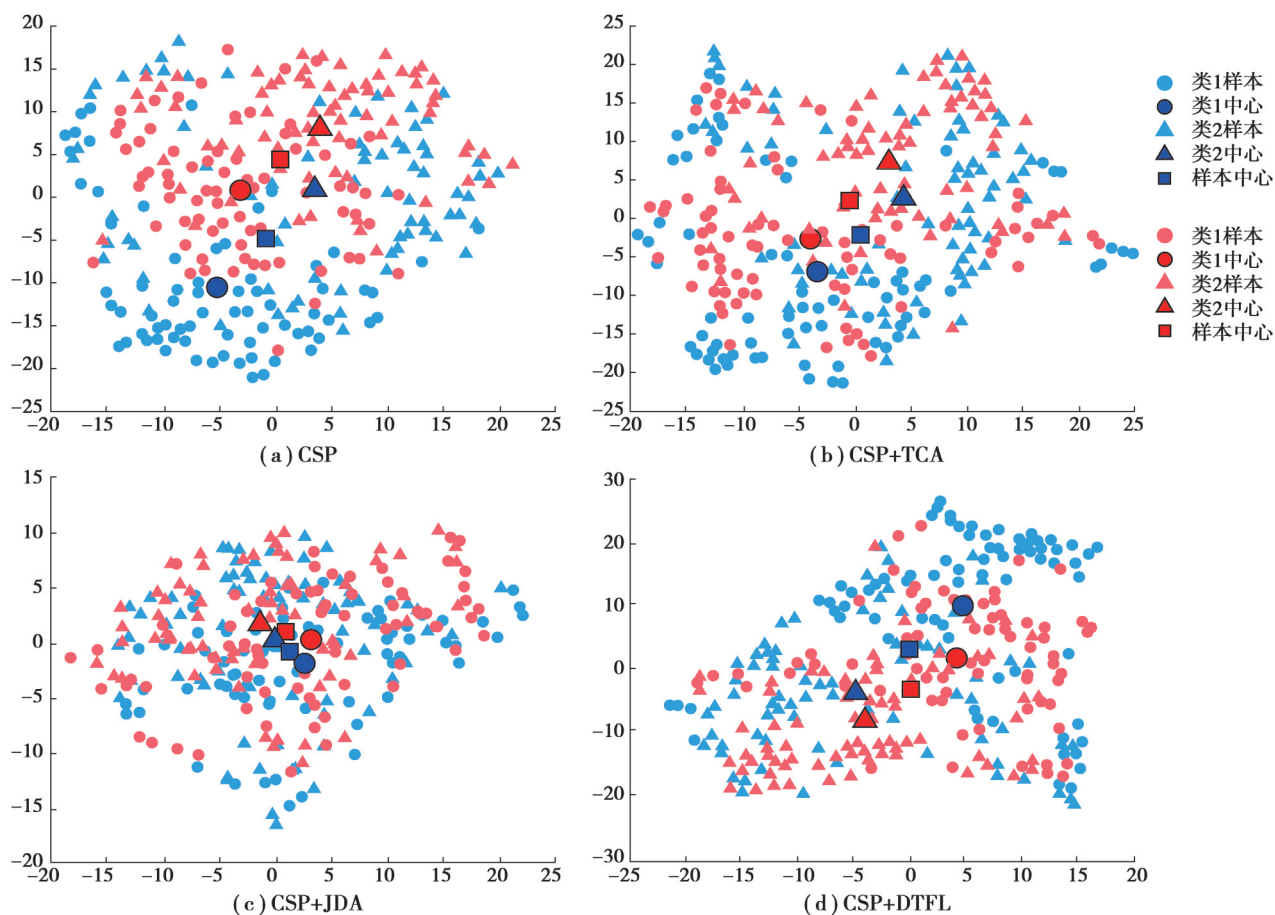


图4 Subject 4数据分布可视化图

Fig. 4 Visualization of data distribution on subject 4

### 3 结论

文中提出了一种基于判别迁移特征学习(DTFL)的运动想象分类方法,用于解决运动想象BCI的重校准问题。DTFL不仅减小了源域和目标域之间的分布差异,同时对源域和目标域的内散度和类间散度进行了最小化和最大化约束,有效的避免了域不变特征的失真,提高了分类性能。通过实验证明,特征的域不变性和类判别性是相互补充的,对于解决运动想象BCI分布差异问题具有重要意义。运动想象数据集的实验结果表明,在不进行校准的情况下,文中提出的方法可以有效地解决运动想象BCI的重校准问题,提高了运动想象分类的正确率。并且在运动想象分类中,文中提出的DTFL方法显著的优于其他几种先进的迁移学习方法。

### 参考文献

- [ 1 ] Brusini L, Stival F, Setti F, et al. A systematic review on motor-imagery brain-connectivity-based computer interfaces[J]. IEEE Transactions on Human-Machine Systems, 2021, 51(6): 725-733.
- [ 2 ] 刘霞,张萍,李云杰,等.基于运动想象的脑机接口技术运用于脑卒中瘫痪患者脑功能激活和神经网络重塑的研究进展[J]. 中华神经科杂志, 2021, 54(10): 1089-1093.  
Liu X, Zhang P, Li Y J, et al. Advances in the application of motor imagery based brain computer interface systems for brain function activation and neural network remodeling in patients with paralysis after stroke[J]. Chinese Journal of Neurology, 2021, 54(10): 1089-1093.(in Chinese)

- [ 3 ] 蒋勤,张毅,谢志荣. 脑机接口在康复医疗领域的应用研究综述[J]. 重庆邮电大学学报(自然科学版), 2021, 33(4): 562-570.  
Jiang Q, Zhang Y, Xie Z R. A review on brain-computer interfaces for rehabilitation application[J]. Journal of Chongqing University of Posts and Telecommunications (Natural Science Edition), 2021, 33(4): 562-570.(in Chinese)
- [ 4 ] Peterson V, Nieto N, Wyser D, et al. Transfer learning based on optimal transport for motor imagery brain-computer interfaces[J]. IEEE Transactions on Bio-Medical Engineering, 2022, 69(2): 807-817.
- [ 5 ] Zhuang F Z, Qi Z Y, Duan K Y, et al. A comprehensive survey on transfer learning[J/OL]. (2021-10-11). <https://doi.org/10.48550/arXiv.1911.02685>.
- [ 6 ] Azab A M, Ahmadi H, Mihaylova L, et al. Dynamic time warping-based transfer learning for improving common spatial patterns in brain-computer interface[J]. Journal of Neural Engineering, 2020, 17(1): 016061.
- [ 7 ] He H, Wu D. Transfer learning for brain-computer interfaces: a euclidean space data alignment approach[J]. IEEE Transactions on Biomedical Engineering, 2020, 67(2): 399-410.
- [ 8 ] Azab A M, Mihaylova L, Ang K K, et al. Weighted transfer learning for improving motor imagery-based brain-computer interface[J]. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2019, 27(7): 1352-1359.
- [ 9 ] Peng J, Sun W, Ma L, et al. Discriminative transfer joint matching for domain adaptation in hyperspectral image classification[J]. IEEE Geoscience and Remote Sensing Letters, 2019, 16(6): 972-976.
- [ 10 ] Long M, Wang J, Ding G, et al. Transfer feature learning with joint distribution adaptation[C]//IEEE International Conference on Computer Vision. IEEE, 2013: 2200-2207.
- [ 11 ] Pan S J, Tsang I W, Kwok J T, et al. Domain adaptation via transfer component analysis[J]. IEEE Transactions on Neural Networks, 2011, 22(2): 199-210.
- [ 12 ] Li S, Song S, Huang G, et al. Domain invariant and class discriminative feature learning for visual domain adaptation[J]. IEEE Transactions on Image Processing, 2018, 27(9): 4260-4273.
- [ 13 ] Li S, Liu C H, Su L M, et al. Discriminative transfer feature and label consistency for cross-domain image classification[J]. IEEE Transactions on Neural Networks and Learning Systems, 2020, 31(11): 4842-4856.
- [ 14 ] Deng W, Liao Q, Zhao L, et al. Joint clustering and discriminative feature alignment for unsupervised domain adaptation[J]. IEEE Transactions on Image Processing, 2021, 30: 7842-7855.
- [ 15 ] Mishuhina V, Jiang X D. Feature weighting and regularization of common spatial patterns in EEG-based motor imagery BCI[J]. IEEE Signal Processing Letters, 2018, 25(6): 783-787.
- [ 16 ] Van Der Maaten L, Hinton G. Visualizing data using t-SNE[J]. Journal of Machine Learning Research, 2008, 9: 2579-2605.

(编辑 詹燕平)