

# 基于卷积联合适应网络的脑电信号情感识别

陈景霞, 胡修文, 唐喆喆, 刘洋, 胡凯蕾

(陕西科技大学电子信息与人工智能学院, 西安 710021)

**摘要:** 提出一种基于深度卷积联合适应网络(Convolutional neural network-joint adaptation network, CNN-JAN)的脑电信号(Electroencephalogram, EEG)情感识别模型。该模型将迁移学习中联合适应的思想融合到深度卷积网络中,首先采用长方形卷积核提取数据的空间特征,捕捉脑电数据通道间的深层情感相关信息,再将提取的空间特征输入到含有联合分布的多核最大均值差异算法(Multi-kernel joint maximum mean discrepancy, MK-JMMD)的适配层进行迁移学习,使用MK-JMMD度量算法解决源域和目标域分布不同的问题。所提方法在SEED数据集上使用微分熵特征和微分尾端性特征分别进行情感分类实验,其中使用微分熵特征被试内跨试验准确率达到84.01%,与对比实验和目前流行的迁移学习方法相比,准确率进一步提高,跨被试实验精度也取得较好的性能,验证了该模型用于EEG信号情感识别任务的有效性。

**关键词:** 脑电信号;卷积神经网络;迁移学习;情感识别;联合适应网络

**中图分类号:** TP311      **文献标志码:** A

## EEG Emotion Recognition Based on Convolutional Joint Adaptation Network

CHEN Jingxia, HU Xiuwen, TANG Zhezhe, LIU Yang, HU Kailei

(College of Electronic Information and Artificial Intelligence, Shaanxi University of Science and Technology, Xi'an 710021, China)

**Abstract:** A new electroencephalogram (EEG) emotion recognition method based on deep convolutional neural network-joint adaptation network (CNN-JAN) is presented. It incorporates the idea of joint adaptation in transfer learning into deep convolutional networks. Firstly, the model uses a rectangular convolution kernel to extract the deep emotion-related spatial features between EEG data channels. Then, the extracted spatial features are input into the adaptation layer with multi-kernel joint maximum mean discrepancy (MK-JMMD) for transfer learning, aiming to reduce the distribution differences between the source and target domains. The experiments are carried out on differential entropy features and differential causality features of EEG data from the SEED dataset to verify the effectiveness and advantages of the proposed method. As a result, the within-subject emotion classification accuracy on differential entropy features reaches 84.01%, and the cross-subject emotion classification accuracy is also improved compared with other current popular transfer learning methods.

**Key words:** electroencephalogram (EEG); convolutional neural network; transfer learning; emotion recognition; joint adaptation network

## 引言

近年来,基于脑-机接口(Brain computer interface, BCI)提取脑电信号(Electroencephalogram, EEG)的情感识别成为情感计算领域中一个备受关注的研究热点。情感计算是一个是需要结合计算机科学、心理学以及社会学等多学科知识进行交叉研究的领域,其目的是通过让计算机学习人与人之间交流互动的特点,使得人机交互过程中计算机能够理解人产生的情感变化并做出适当的反应。情感计算研究可以借助很多生理信号,如心率血压、皮肤电活动、眼电信号和脑电信号等。研究表明,皮肤电反应是一种有效测量情感变化程度的生理信号指标,脑电信号就是皮肤电的一种,是被试者佩戴的脑电帽在头皮处采集的大脑发出的微弱电信号,其中包含大量人体生理信息。人类情感发生变化时,EEG信号也会随着发生变化<sup>[1]</sup>。当人类情绪变化大时,EEG信号也会产生较大波动;当人类情绪平稳时,EEG信号也趋近于平缓。由于EEG信号能够直接反应人的情感变化,具有客观的和无法伪装的特性,正在被越来越多的应用到情感计算领域中<sup>[2]</sup>。

由于脑电信号存在较大的个体差异,一个模型往往只适用于一个特定的任务,为了提高模型的泛化能力,迁移学习是解决这个问题的一种有效途径<sup>[3]</sup>。迁移学习可以把训练好的模型、参数等用于另一个相似任务上,大大减小了最优模型的训练时间,同时也减少了训练模型所需的数据量。随着迁移学习的快速发展,基于迁移学习的脑电情感识别成为情感计算领域的一项重要研究内容。传统的迁移学习方法不需要复杂的深度网络和大量的数据,就可以达到很好的迁移效果。Dai等<sup>[4]</sup>提出了TrAdaBoost迁移学习框架,利用少量带有标签的目标域数据和大量源域数据训练分类模型,增加有效参数的权重训练网络,完成从源域到目标域数据的迁移;Zhang等<sup>[5]</sup>进一步使用TrAdaBoost方法来训练基于个体相似度的模型,使用最大均值差异(Maximum mean discrepancy, MMD)<sup>[6]</sup>计算个体间分布的差异程度,该方法在DEAP数据集效价上情感二分类实验精度达到66.7%,在唤醒度上实验精度达到66.1%;Zheng等<sup>[7]</sup>对比了多种经典的迁移学习方法在SEED数据集上进行三分类情感识别实验,其中直推式参数迁移(Transductive parameter transfer, TPT)<sup>[8]</sup>在源域数据上训练支持向量机(Support vector machine, SVM)分类器,选取效果最好的分类器参数和数据分布的回归函数对目标域数据进行测试,达到了76.31%的分类精度。

上述研究表明,传统的迁移学习方法对提高情感识别分类任务是有效的。随着深度学习的快速发展,深度迁移学习方法与传统迁移学习方法相比表现出更好的性能<sup>[9-10]</sup>。Long等<sup>[11]</sup>提出了深度自适应网络(Deep adaptation network, DAN)方法,在AlexNet网络分类器的前一层加入多层多核MMD(Multi kernel MMD, MK-MMD)<sup>[12]</sup>的适配层,减小源域和目标域数据分布之间的差异,在图像领域取得了较好的效果;Li等<sup>[13]</sup>提出将DAN的思想应用在情感识别中,在多层感知机(Multilayer perceptron, MLP)网络中加入MK-MMD层,对源域和目标域数据进行适配,从而最大程度消除了个体数据差异,该方法准确率达到83.81%。Ganin等<sup>[14]</sup>提出了领域对抗神经网络(Domain-adversarial neural network, DANN),利用网络对抗训练的思想进行域间数据的适配,增加一个梯度反转层监督网络的对抗训练,当网络无法区分特征来自源域还是目标域时,便实现了有效的域间迁移,实验证明DANN在Office数据集上的图像分类任务表现良好;Ma等<sup>[15]</sup>将DANN改进后提出了一种领域残差网络(Domain residual network, DResNet)用于训练情感模型,在整个模型训练中目标域数据不需要带有标签,该方法在SEED-VIG数据集上对司机疲劳驾驶需警惕程度的预测准确率达到83.86%。Liu等<sup>[16]</sup>将联合适应网络(Joint adaptation network, JAN)<sup>[17]</sup>思想应用于情感识别中, JAN方法在深度网络中对数据进行联合分布自适应,采用AlexNet网络融入MMD度量算法进行情感识别实验,对SEED和SEED-V数据集上的DE特征进行三分类和五分类情感预测,被试内实验准确率分别达到了79.87%和76.46%。

虽然研究者们在使用迁移学习方法进行EEG情感识别方面取得了一定进展,但仍然存在两个重要问题需要进一步研究:一是如何减小源域和目标域EEG间分布差异,不同被试间EEG信号差异很大,而且随情绪变化波动明显,容易导致EEG情感分类模型的适应性和泛化能力较差;二是如何在网络中实现适配层之间联合分布的互相联系,进一步提高适配层的效率。通常在深度网络中进行深层特征提取后,用于在特征和标签之间减小数据分布的激活函数只会作用在适配层的最后一层,导致网络前面的训练没有把源域和目标域数据的信息关联起来。为了解决这两个问题,本文提出将JAN分别和CNN、门控循环单元(Gate recurrent unit,GRU)网络进行融合,同时使用含有联合分布的多核最大均值差异算法(Multi-kernel joint maximum mean discrepancy,MK-JMMD)对源域和目标域数据的概率分布进行度量,通过缩小数据间边缘概率分布和条件概率分布的差异来解决基于EEG的领域自适应情感识别问题。

## 1 本文方法

### 1.1 最大均值差异算法

最大均值差异(MMD)<sup>[6]</sup>是迁移学习训练中使用最广的损失函数之一,它通过核函数将两种不同但相关的数据映射到高维的再生核希尔伯特空间(Reproducing kernel Hilbert space,RKHS)<sup>[18]</sup>,在高维空间中来度量两种数据的分布距离。RKHS是由核函数构建的完备的内积函数空间。设 $H_s$ 是希尔伯特空间,存在函数 $k: X \times X \rightarrow \mathbf{R}, X \subseteq \mathbf{R}^n$ ,满足对任意 $x \in X, k(x, \cdot) \in H_s$ 且 $k(\cdot, \cdot)$ 具有再生性质,即

$$f(x) = \langle f(\cdot), k(x, \cdot) \rangle_H \quad (1)$$

那么称 $H$ 为再生核希尔伯特空间, $k(x, \cdot)$ 可以看作一个特征映射 $\Phi(x)$ ,其中 $k(x, x') = \langle \Phi(x), \Phi(x') \rangle_H$ 。

用MMD度量算法计算源域和目标域在RKHS空间映射的最小距离。从分布 $P(X^s)$ 和 $Q(X^t)$ 中分别取源域样本集 $D_{X^s} = \{x_1^s, \dots, x_n^s\}$ 和目标域样本集 $D_{X^t} = \{x_1^t, \dots, x_n^t\}$ ,则MMD的度量在RKHS中映射距离的计算公式为

$$D_H(P, Q) = \left\| \mu_{X^s}(P) - \mu_{X^t}(Q) \right\|_H^2 \quad (2)$$

当且仅当 $P = Q$ 时, $D_H(P, Q) = 0$ 。MMD的平方距离为

$$\hat{D}_H(P, Q) = \frac{1}{n_s^2} \sum_{i=1}^{n_s} \sum_{j=1}^{n_s} k(x_i^s, x_j^s) + \frac{1}{n_t^2} \sum_{i=1}^{n_t} \sum_{j=1}^{n_t} k(x_i^t, x_j^t) - \frac{2}{n_s n_t} \sum_{i=1}^{n_s} \sum_{j=1}^{n_t} k(x_i^s, x_j^t) \quad (3)$$

式中: $\hat{D}_H(P, Q)$ 为 $D_H(P, Q)$ 的无偏估计;核函数 $k(\cdot, \cdot)$ 在高维空间中计算源域 $s$ 和目标域 $t$ 中的数据间的距离。

### 1.2 联合适应网络

联合适应网络(JAN)是一种领域自适应的学习模型,通过学习达到减小数据间联合分布差异的目的。联合分布由边缘概率分布 $P(X^s, Y^s)$ 和条件概率分布 $Q(X^t, Y^t)$ 组成,上标 $s$ 和 $t$ 分别代表源域和目标域, $X$ 和 $Y$ 分别代表数据和标签。研究表明,将数据输入到深度网络中进行深度特征提取,联合分布的变化只会停留在激活层 $Z^1, Z^2, \dots, Z^L$ 的最高层网络 $L$ 层中<sup>[19]</sup>,因此,JAN利用层与层间的联合分布 $P(Z^{s1}, Z^{s2}, \dots, Z^{sL})$ 和 $Q(Z^{t1}, Z^{t2}, \dots, Z^{tL})$ 替换 $P(X^s, Y^s)$ 和 $Q(X^t, Y^t)$ 。

本文将源域和目标域EEG数据通过核函数映射到再生核希尔伯特空间,使用JMMD算法度量JAN网络所有适配层的联合分布差异,JMMD定义为

$$\hat{D}_H(P, Q) = \frac{1}{n_s^2} \sum_{i=1}^{n_s} \sum_{j=1}^{n_s} \prod_{\ell \in L} k^\ell(\mathbf{z}_i^{\ell}, \mathbf{z}_j^{\ell}) + \frac{1}{n_t^2} \sum_{i=1}^{n_t} \sum_{j=1}^{n_t} \prod_{\ell \in L} k^\ell(\mathbf{z}_i^{\ell}, \mathbf{z}_j^{\ell}) - \frac{2}{n_s n_t} \sum_{i=1}^{n_s} \sum_{j=1}^{n_t} \prod_{\ell \in L} k^\ell(\mathbf{z}_i^{\ell}, \mathbf{z}_j^{\ell}) \quad (4)$$

核函数  $k$  使用了数值不唯一的权重,体现了不同层  $\ell \in L$  中其他变量之间的相互作用。

### 1.3 深度卷积联合适应网络

卷积神经网络(Convolutional neural network, CNN)是标准全连接神经网络的一种变体,在图像、文本分类与识别领域均达到不错的性能,也越来越多地被使用在 EEG 相关任务中<sup>[20]</sup>。Yoon<sup>[21]</sup>用多个大小不同的长方形卷积核提取语义特征,验证了长方形卷积核可以更好地获取文本信息的局部相关性。Chakrabarty 等<sup>[22]</sup>证明使用大小为  $M \times 1$  的卷积核( $M > 1$ )进行训练比正方形卷积核可以在动态声学场景中达到更好的定位性能。

声学信号是波状数据,同样脑电信号也是波状数据。在上述工作启发下,本文采用具有正方形和长方形两种不同尺度卷积核构成的 CNN 模型,先从原始 EEG 数据中学习深层时空相关语义特征,然后再输入 JAN 进一步学习源域和目标域的联合分布,通过多核 JMMD 算法最小化两个联合分布的差异,使源域上训练的分类器能更好地适应目标域测试集上的数据。该组合模型被称为 CNN-JAN 模型,其结构如图 1 所示,先使用 32 个  $3 \times 3$  正方形卷积核学习通道间数据的相关性信息,再接入 64 个大小为  $3 \times 1$  的长方形卷积核,这一卷积层可以将数据信息合并到学习的空间特征中,两个卷积层均进行步长为 1 有填充的卷积操作,采用 ReLu 激活函数和 Adam 优化器。

CNN-JAN 模型的两个卷积层后分别连接两个池化层,选择  $2 \times 2$  大小且步长为 2 的 Maxpooling 过滤器对经过卷积核处理的数据进行下采样处理;将经过第二个池化层的数据输入 JAN 全连接适配层,适配层对经过 CNN 提取的空间特征进行迁移学习,减小源域和目标域联合分布的差异。多核 JMMD(Multi-kernel JMMD, MK-JMMD)是在 JMMD 的基础上由多个核线性组合得到的,具体表示为

$$K \triangleq \left\{ k = \sum_{u=1}^m \beta_u k_u; \beta_u \geq 0, \forall u \right\} \quad (5)$$

式中:  $K$  表示多个核函数经过线性组合的公式;  $k$  表示核函数;  $\beta$  表示核函数的数量。数据经过两层适配层后送入一个全连接层,并使用 Softmax 分类器进行三分类情感预测。CNN-JAN 模型的优化目标是 minimized MK-JMMD 的损失函数和网络自身损失函数之和, CNN 网络的交叉熵损失函数为

$$\min_f \frac{1}{n_s} \sum_{i=1}^{n_s} J(f(\mathbf{x}_i^s), \mathbf{y}_i^s) \quad (6)$$

式中  $J(\cdot, \cdot)$  为交叉熵损失函数。CNN-JAN 模型的最终目标函数定义为

$$\min_f \frac{1}{n_s} \sum_{i=1}^{n_s} J(f(\mathbf{x}_i^s), \mathbf{y}_i^s) + \lambda \hat{D}_L(P, Q) \quad (7)$$

式中  $\lambda > 0$  为 MK-JMMD 惩罚项权衡参数。所提 CNN-JAN 模型在训练中考虑到了数据不同通道间的相关性和深度网络层与层之间的相关性,能更充分地完成源域到目标域的迁移学习,从而提升模型在情感识别任务中的性能。

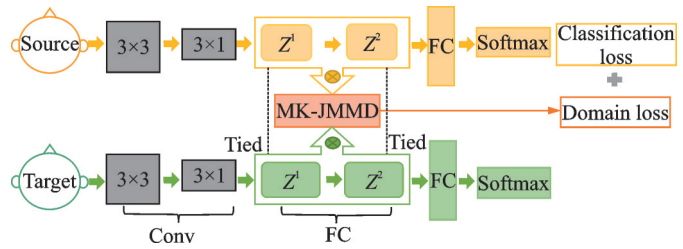


图1 CNN-JAN 模型

Fig.1 CNN-JAN model



本文同时提出用来提取 EEG 序列时间相关性特征的 GRU-JAN 组合网络作为对比模型。门控循环单元(Gate recurrent unit, GRU)<sup>[23]</sup>适合学习预测时间顺序数据之间的上下文相关信息, EEG 信号就是按照时间序列采集的信息。GRU-JAN 同 CNN-JAN 结构相似, 首先将源域和目标域特征分别输入到一层 GRU 中, 设 GRU 的当前节点的输入为  $x^t$ , 输入数据与保存之前所有参数的隐藏节点  $h^{t-1}$  共同在网络中学习, 经过一个节点的学习会得到输出  $y^t$  和将传递给下一个节点的隐藏节点  $h^t$ 。再将经过 GRU 学习到的时序相关特征输入适配层, 在适配层经过 MK-JMMD 度量最终接入全连接层和分类层进行情感分类。

## 2 数据集及数据预处理

### 2.1 数据集

本文采用上海交通大学 BCMI 实验室提供的 SEED 数据集(The SJTU emotion EEG dataset)进行实验。该数据集记录了 7 名男性被试和 8 名女性被试观看 15 个不同的中国电影剪辑片段时, 根据电影情节变化刺激大脑所产生的脑电信号, 这些视频引发 3 种类型的情绪, 包括中立、积极和消极。所有被试都进行了 3 次脑电图记录实验, 每次实验之间间隔两周, 因此, 每个被试都有 3 次实验数据, 每次实验包含 15 次 62 个通道的脑电信号数据。数据采集时, 每个电影片段播放前均有 5 s 的提示时间, 电影片段 4 min 的播放时间, 自我完成问卷评估 45 s, 休息时间 15 s。采集装置采用“10/20”国际导联标准的 62 导联电极帽, 采样频率为 1 000 Hz, 下采样到 200 Hz。

### 2.2 数据预处理

已有的研究结果<sup>[9,13]</sup>表明微分熵(Differential entropy, DE)特征在脑电情感识别实验中取得了较优的性能。本研究在上述成果的启发下, 使用 SEED 数据集提供的微分熵和微分尾端性(Differential causality, DCAU)两种频域相关特征进行情感分类实验。该数据集共有 45 个 .mat 文件, 每一个 .mat 文件记录了每个被试每次实验观看 15 个电影片段所采集脑电数据的 5 个特征。其中, DE 特征的数据格式为 62(Channels) × 样本数 × 5(Frequency bands), 5 个频带分别为  $\delta$ 、 $\theta$ 、 $\alpha$ 、 $\beta$ 、 $\gamma$ <sup>[24]</sup>。使用线性动态系统(Linear dynamic system, LDS)算法对这些特征进行了平滑处理。由于 15 名被试分别进行了 3 组不同时间的实验, 为了提高计算效率, 将每组实验观看 15 视频(15 trials)采集的 EEG 数据分别降采样至 185 个采样点, 再将每个被试每次实验 EEG 数据格式整理为 62(Channels) × 5 × 2 775(Samples = 185 Samples/Trial × 15 Trials)。

DCAU 是大脑额叶和后叶电极位置的微分熵进行差分计算得到的特征数据, 反应了不同脑区产生的电信号间的相互联系, 它的数据结构为 32(Channels) × 样本数 × 5(Frequency bands), 同样提取 LDS 算法处理过的数据, 对每个 Trial 分别下采样至 185 个采样点, 每组数据大小整理为 32(Channels) × 5 × 2 775(Samples)。SEED 数据集将实验中刺激产生的情绪分为积极、中立和消极, 使用 {1, 0, -1} 组成的数组作为标签表示这 3 种不同的情绪, 将标签大小整理为 2 775 × 1。

## 3 情感分类实验及结果分析

本实验基于上述 SEED 数据集里 15 名被试的脑电特征来验证所提出方法的可行性和有效性, 采用交叉验证来测试算法准确性。所提模型在 pytorch 1.8.1 框架和 GeForce GTX 3090 GPU 上实现。CNN-JAN 模型的适配层选择径向基函数(Radial basis function, RBF), 经过参数调优后学习率设置为 0.001, 迭代次数设置为 10 000 次。本文选择 3 种传统迁移学习方法和 1 种深度自适应网络作为基准方法展开对比实验, 4 种对比方法的基本原理和配置如下:

(1) 迁移成分分析法(Transfer component analysis, TCA)<sup>[25]</sup>将特征映射到RKHS空间中,可以最大化不同类别数据之间的方差和最小化源域和目标域数据分布之间的距离,最大程度的保留使学习任务有效的数据属性。

(2) 联合分布适配法(Joint distribution adaptation, JDA)<sup>[26]</sup>同时适配源域和目标域的边缘分布概率和条件分布概率。

(3) 平衡分布适配(Balanced distribution adaptation, BDA)<sup>[27]</sup>对 JDA 方法进行改进,根据不同的数据领域对不同分布的权重进行动态自适应调整。

TCA、JDA 和 BDA 三种方法均选择线性核作为核函数,都使用 KNN 分类器进行三分类实验。TCA、JDA 和 BDA 的正则化参数分别设为 5.0、0.8 和 0.7。

(4) DAN:由两层卷积层后分别连接两层池化层,再加上两层含有 MK-MMD 的全连接适配层和一层全连接层构成。两个卷积层分别由 32 个  $3 \times 3$  和 64 个  $2 \times 2$  大小的正方形卷积核构成,进行步长为 1 的有填充的卷积操作。两个池化层都采用  $2 \times 2$  大小的 Maxpooling 过滤器对经过卷积核处理的数据进行下采样处理,步长设置为 2。将经过第二个池化层池化过的数据输入采用 RBF 作为核函数的全连接适配层,经过两层适配层后接入全连接层,使用 Softmax 分类器进行三类情感分类。卷积层均采用 ReLu 激活函数,选择 Adam 优化器,经过参数调优后学习率设置为 0.001,迭代次数设置为 10 000 次。

### 3.1 同一被试跨时间的实验结果与分析

同一被试跨时间的实验用来预测被试在不同时间的情绪变化,检测所提模型随时间变化的稳定性。在实验中,分别选择其中一组数据作为源域,其他两组数据分别作为目标域构建交叉验证集,将三折交叉验证的平均分类准确率作为每个被试最终的跨时间实验结果。所提的 CNN-JAN 模型以及各种基准模型的整体实验性能如表 1 所示。

观察发现,在 DCAU 特征上,传统迁移学习方法 TCA 和 JDA 的平均准确率分别为 60.52% 和 59.54%,BDA 模型的性能较优,其平均准确率为 60.73%,说明 BDA 模型加入对数据分布的权重进行动态调整的方法是有效的。进一步观察发现 DAN 模型的平均准确率为 76.65%,表明深度迁移学习方法明显优于传统的迁移学习方法,而所提 CNN-JAN 模型的平均准确率达到 78.38%,在对比方法中表现最优。

在 DE 特征上,BDA 也取得传统迁移学习方法里最好的准确率 73.14%,较 TCA 和 JDA 分别高出 9.92% 和 0.40%。深度迁移学习方法

DAN 模型准确率为 81.08%,说明使用 CNN 提取空间相关性后再进行迁移学习可以提升情感分类任务的准确率。同时比较深度模型的参数量和运行时间,DAN 模型参数总量为 15 174 个,单次一组数据的实验运行时间 738 s,CNN-JAN 模型参数总量为 13 126 个,单次实验运行时间 5 771 s,虽然所提 CNN-JAN 模型耗时较多,但其参数总量少于 DAN 模型,且 CNN-JAN 达到最优准确率 84.01%,比 DAN 的准确率高出 2.93%。CNN-JAN 模型性能提升显著,表明 CNN-JAN 在 DAN 考虑不同分布差异的基础上同时考虑适配层间的分布可以进一步提高 EEG 情感分类准确率。

同时观察发现,使用 DE 特征进行实验的准确率在各个模型上都高于 DCAU 特征,这种现象可能是因为 DCAU 特征是在 DE 特征基础上进行通道间的差分计算得到的特征数据,EEG 信号的各个通道采

表 1 同一被试跨时间实验的平均分类准确率

Table 1 Average classification accuracy of the within-subject across time experiments

模型	准确率	
	DCAU 特征	DE 特征
TCA	0.605 2	0.632 2
JDA	0.595 4	0.727 4
BDA	0.607 3	0.731 4
DAN	0.766 5	0.810 8
CNN-JAN	0.783 8	0.840 1

集的是不同大脑区域的生理电信号,数据本身具有较大差异性,通道间的差异经过差分计算之后反而更加凸显,使得 DCAU 特征的分类准确率低于 DE 特征。15 名被试以 DE 特征为输入进行跨时间情感三分类实验的平均准确率对比情况如图 2 所示。观察图 2 发现,传统迁移学习方法在每个被试上的准确率大多数都低于深度迁移学习方法,但在一个两个被试上性能反超深度迁移学习模型,如 sub3 和 sub13 上 JDA 的准确率均超过 80%,说明源域和目标域的边缘分布和条件分布权重大体相同,适合采用 JDA 对数据进行处理,但对于源域和目标域上这两种分布差异较大的数据,JDA 方法的准确率并没有达到最优。总体来看,与其他方法对比,虽然 CNN-JAN 模型没有在每个被试上都取得最好的准确率,但该模型对各种不同分布的数据均表现出不错的性能,其平均精度达到了最高,证明了该模型的有效性和优越性。

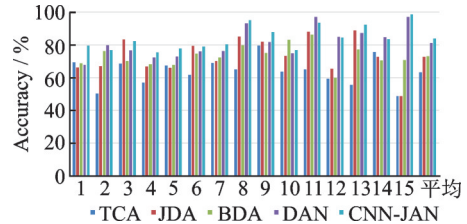


图 2 15 名被试 DE 特征被试内跨时间实验结果对比

Fig.2 Statistics chart of classification accuracy results of experimental test set on the DE feature for 15 subjects

同时在图 2 中可以看出第 15 名被试使用 CNN-JAN 模型在 DE 特征上的分类准确率达到 98.85%,为所有被试最高。图 3 为 sub15 训练集上效果最优的一次学习曲线,橙线表示训练准确率 accuracy,蓝线表示平均误差 loss。从图 3 可以看出,随着迭代次数的增加,训练准确率从 0.35 增长至 0.991 7 处收敛;平均误差 loss 整体呈现下降趋势,从 0.13 逐步降低向 0 逼近。在这一过程中,Iteration 在 6 000 到 7 500 轮之间 loss 出现小幅度波动,但并未对训练准确率造成影响。

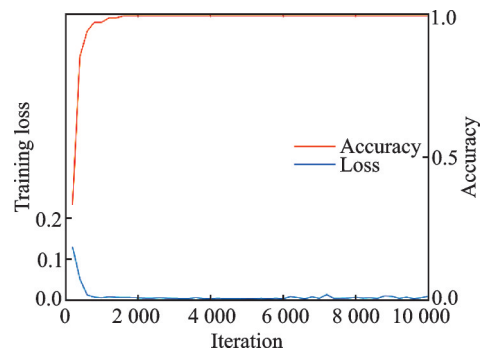


图 3 sub15 上 CNN-JAN 模型训练过程示意图

Fig.3 sub15 training process of CNN-JAN

### 3.2 跨被试的实验结果与分析

跨被试的实验用来检测不同个体间的情感识别模型的稳定性。在实验中,依次将每一名被试的 EEG 数据作为测试集即目标域,其他 14 名被试的 EEG 数据作为训练集即源域,为了避免时间因素影响,均使用 SEED 数据集中每个被试第一组时间采集的 EEG 数据进行实验。所提模型以及各种基准模型的整体实验结果如表 2 所示。

表 2 跨被试实验的平均分类准确率

Table 2 Average classification accuracy across subjects

模型	准确率	
	DCAU 特征	DE 特征
TCA	0.562 6	0.569 1
JDA	0.582 9	0.600 4
BDA	0.600 7	0.622 2
DAN	0.727 2	0.722 6
CNN-JAN	0.740 0	0.742 2

观察发现 JDA 和 BDA 模型的性能在 DCAU 和 DE 特征上均优于 TCA 模型,BDA 在 DE 特征准确率达到 62.22%,在 DCAU 特征上实验准确率为 60.07%,再一次验证动态适配边缘分布和条件分布比默认数据分布权重相同的方法性能优异。DAN 模型较 BDA 性能有所提升,在 DE 和 DCAU 特征上的分类性能分别提升了 10.04% 和 12.65%;CNN-JAN 模型在 DCAU 特征上准确率达到 73.02%,在 DE 特征进行实验取得最好准确率 73.60%,表明所提模型在跨被试实验中也取得了较好的准确率。CNN-JAN 模型在 DE 特征上进行跨被试实验中,15 名被试分别作为目标域测试集的平均分类精度如图 4 所示。由图 4 观察到,跨被试实验中第 3 名被试作为目标域的交叉验证实验准确率达到 84.29%,为所有被试最高,其在最优训练集上的学习曲线如图 5 所示。

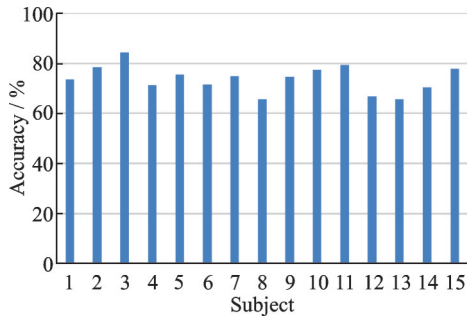


图4 15名被试基于CNN-JAN模型和DE特征的跨被试情感分类结果

Fig.4 Cross-subject sentiment classification results based on CNN-JAN model and DE features for 15 subjects

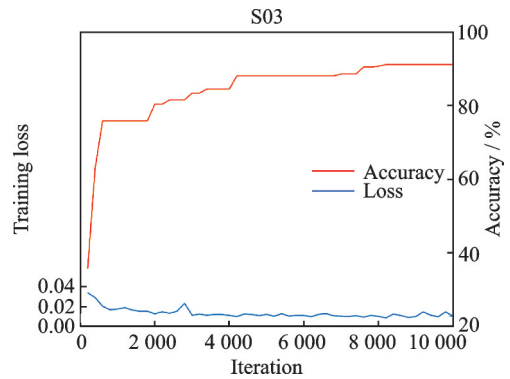


图5 sub03上CNN-JAN模型跨被试训练过程示意图  
Fig.5 sub03 training process of CNN-JAN across subjects

从图5看出,随着平均误差 Loss从0.034逐步降低至0.01,训练准确率逐步提升并收敛至0.910 4。在这一过程中,Iteration从2 000到4 000轮之间 loss出现一定幅度的波动,同时训练准确率也呈现震荡上升趋势;Iteration从5 000到10 000轮 Loss波动较小趋于平稳,训练准确率也趋于稳定。

### 3.3 不同JAN变种模型的实验结果与分析

为了验证所提具有长方形卷积核CNN-JAN模型的优势,本文与只含有正方形卷积核的CNN-JAN模型的变体(简称JAN-1)和1.3节中提出的GRU-JAN模型进行对比实验,参数设置如下:

(1) 所提 CNN-JAN和GRU-JAN模型的适配层均选择RBF核函数,经过参数调优后学习率均设置为0.001,迭代次数设置为10 000次,优化器Momentum动量设为8,L2正则化参数设置为0.000 5。

(2) JAN-1模型由两层3×3大小的卷积核构成,为了公平起见,池化层和适配层的结构与CNN-JAN一致,适配层同样选择RBF核函数,迭代次数和学习率等参数设置一致。

被试内跨时间的实验结果如表3所示,跨被试的实验结果如表4所示。

表3 不同JAN模型被试内实验的分类性能对比  
Table 3 Comparison of classification performance in within-subject experiments of different JAN models

模型	准确率	
	DCAU特征	DE特征
JAN-1	0.720 4	0.788 1
GRU-JAN	0.721 3	0.804 5
CNN-JAN	0.783 8	0.840 1

表4 不同JAN模型跨被试实验的分类性能对比  
Table 4 Comparison of classification performance of different JAN models across subjects

模型	准确率	
	DCAU特征	DE特征
JAN-1	0.656 2	0.675 7
GRU-JAN	0.601 3	0.626 9
CNN-JAN	0.740 0	0.742 2

由表3可知,JAN-1模型在DCAU特征上准确率为72.04%,在DE特征上准确率为78.81%,CNN-JAN分别在这两个特征上比JAN-1性能提升了6.34%和5.20%。由此可以看出,针对脑电数据,长方形卷积核可以更好的提取数据的时空特征。同时观察到GRU-JAN模型在DCAU和DE特征上准确率分别为72.13%和80.45%,由于GRU-JAN模型侧重于提取EEG数据的时序特征,该模型也取得了较好的性能,但是其性能仍然低于CNN-JAN模型,表明CNN模型所提取的深层时空特征比GRU模型所提取的深层时序特征具有更好的可分性。

由表4可知,CNN-JAN比JAN-1在DACU特征上性能提高了8.38%,在DE特征上提高了6.65%,



再一次验证了本实验采用长方形卷积核进行情感分类性能较优。但在跨被试实验中,GRU-JAN模型的性能略低于JAN-1,这种现象可能是因为EEG数据量相对较小,GRU更适合针对大规模长时序列进行特征提取;也可能是因为跨被试EEG数据在时序上下文之间的相关性相对通道间的空间相关性较弱,后续将进一步对该模型进行优化。

进一步比较各个模型参数总量,JAN-1模型参数总量为25 222个,GRU-JAN模型为48 073个,CNN-JAN模型为13 126个,所提CNN-JAN模型参数量最少,表现出较高的计算性能。以上实验结果表明,所提CNN-JAN模型无论在被试内跨时间实验还是跨被试实验的性能均优于对比模型,验证了长方形与正方形组合的卷积核在提取EEG数据深层时空相关性特征方面的优势,同时也证明通过融合联合适应层确实能够减小源域和目标域深层特征联合分布之间的差异,有效提高模型的分类准确性和鲁棒性。

### 3.4 采用不同核函数的实验结果与分析

CNN-JAN是通过适配层的核函数对EEG深层特征进行空间映射,通过MK-JMMD算法缩小源域和目标域特征分布间的差异,从而完成特征的迁移学习。为了进一步证明所提方法的优越性,本文对采用不同核函数的CNN-JAN模型进行了对比实验,对比模型分别采用线性(Linear)核函数和RBF。其中RBF-5和RBF-10分别表示每个适配层采用数量为5个和10个RBF核函数。为了确保实验公平,除了选取不同核函数外,其余实验参数与上文CNN-JAN相关实验保持一致,均使用DE特征,实验结果如表5所示。观察表5发现,采用RBF核的实验准确率比Linear线性核的高,证明RBF能更好地将EEG特征映射到高维空间并进行分布计算,在EEG特征适配方面具有较强的鲁棒性。RBF-5被试内和被试间分类准确率分别达到80.73%和69.70%,也取得了不错的效果,表明在适配层加入适当数量核函数的模型更能充分利用CNN学习到的深层时空相关语义特征,显著提高特征迁移学习的效果和情感识别的性能。

表5 使用不同核函数的实验结果对比

**Table 5 Comparison of experimental performance of CNN-JAN model by using different kernel functions**

核函数	准确率	
	被试内	被试间
Linear	0.780 7	0.605 8
RBF-5	0.807 3	0.697 0
RBF-10	0.840 1	0.742 2

## 4 结束语

本文采用了卷积神经网络和联合适应网络相结合的CNN-JAN模型来处理基于EEG信号的情感识别任务。利用含有正方形卷积核与长方形卷积核的CNN对EEG数据进行深层时空特征提取,再使用多核JMMD算法度量其源域和目标域特征的联合分布差异,从而完成源域到目标域深层特征的迁移学习。通过与经典的传统迁移学习方法和流行的深度迁移学习方法在SEED数据集EEG信号的DE特征和DCAU特征上进行对比实验和结果分析,发现所提CNN-JAN模型在被试内和被试间的实验中都体现出较好的情感分类性能,EEG信号情感分类的准确率和稳定性得到了进一步的提高,证明该模型确实能够通过深层特征提取与联合适应学习,减少源域与目标域特征分布之间的特征,有效提高了模型的精度和自适应性,为后续开发鲁棒的情感脑-机接口应用打下了良好的基础。针对现存迁移学习解释性弱的问题,下一步将对该方法进行优化,尝试提取相关性更高、解释性更强的脑电特征进行更深层次的研究,同时开展基于迁移学习的EEG信号跨数据集情感识别研究,在确保模型简单的基础上获得较高的性能。

## 参考文献:

- [1] LIN Chongde, YANG Zhiliang, HUANG Xiting. The comprehensive dictionary of psychology[M]. Shanghai: Shanghai Educational Publishing House, 2004.
- [2] 张家瑞,王刚.基于脑电信号的情感识别研究[J].计算机应用研究,2019,36(11): 3306-3309.  
ZHANG Jiarui, WANG Gang. Research on emotion recognition based on EEG signals[J]. Application Research of Computers, 2019, 36(11): 3306-3309.
- [3] 林佳伟,王士同.用于迁移学习的多尺度领域对抗网络[J].数据采集与处理,2022,37(3): 555-565.  
LIN Jiawei, WANG Shitong. Multi-scale domain adversarial network for transfer learning[J]. Journal of Data Acquisition and Processing, 2022, 37(3): 555-565.
- [4] DAI Wenyuan, YANG Qiang, XUE Guirong, et al. Boosting for transfer learning[C]//Proceedings of the 24th International Conference on Machine Learning. New York, USA: Association for Computing Machinery, 2007: 193-200.
- [5] ZHANG Xiaowei, LIANG Wenbin, DING Tingzhen, et al. Individual similarity guided transfer modeling for EEG-based emotion recognition[C]//Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine. San Diego, CA: IEEE, 2019: 1156-1161.
- [6] BORGWARDT K M, GRETTON A, RASCH M J, et al. Integrating structured biological data by kernel maximum mean discrepancy[C]//Proceedings of the 14th International Conference on Intelligent Systems for Molecular Biology. Fortaleza, Brazil: [s.n.], 2006: 49-57.
- [7] ZHENG Weilong, LU Baoliang. Personalizing EEG-based affective models with transfer learning[C]//Proceedings of the 25th International Joint Conference on Artificial Intelligence. New York, USA: [s.n.], 2016: 2732-2739.
- [8] SANGINETO E, ZEN G, RICCI E, et al. We are not all equal: Personalizing models for facial expression analysis with transductive parameter transfer[C]//Proceedings of ACM International Conference on Multimedia. Orlando, USA: ACM, 2014: 357-366.
- [9] 权学良,曾志刚,蒋建华,等.基于生理信号的情感计算研究综述[J].自动化学报,2021,47(8): 1769-1784.  
QUAN Xueliang, ZENG Zhigang, JIANG Jianhua, et al. Physiological signals based affective computing: A systematic review [J]. Acta Automatica Sinica, 2021, 47(8): 1769-1784.
- [10] 高凌飞,王海龙,王海涛,等.基于轻量级卷积神经网络的人证比对[J].南京航空航天大学学报,2021,53(5): 751-758.  
GAO Lingfei, WANG Hailong, WANG Haitao, et al. Scene-identity face matching based on efficient convolutional neural network[J]. Journal of Nanjing University of Aeronautics & Astronautics, 2021, 53(5): 751-758.
- [11] LONG Mingsheng, CAO Yue, WANG Jianmin, et al. Learning transferable features with deep adaptation networks[C]//Proceedings of International Conference on Machine Learning. Lille, France: [s.n.], 2015: 97-105.
- [12] GRETTON A, SRIPERUMBUDUR B K, SEJDINOVIC D, et al. Optimal kernel choice for large-scale two-sample tests [C]//Proceedings of Advances in Neural Information Processing Systems. Lake Tahoe, USA: [s.n.], 2012: 1214-1222.
- [13] LI He, JIN Yiming, ZHENG Weilong, et al. Cross-subject emotion recognition using deep adaptation networks[C]//Proceedings of International Conference on Neural Information Processing. Siem Reap, Cambodia: [s.n.], 2018: 403-413.
- [14] GANIN Y, USTINOVA E, AJAKAN H, et al. Domain-adversarial training of neural networks[J]. The Journal of Machine Learning Research, 2016, 17: 2096-2030.
- [15] MA Boqun, LI He, ZHENG Weilong, et al. Reducing the subject variability of EEG signals with adversarial domain generalization[C]//Proceedings of International Conference on Neural Information Processing. Sydney, Australia: [s.n.], 2019: 30-42.
- [16] LIU Hong, GUO Hong, HU Wei. EEG-based emotion classification using joint adaptation networks[C]//Proceedings of IEEE International Symposium on Circuits and Systems. [S.l.]: IEEE, 2021.
- [17] 王掬辰,杨士俊.再生核希尔伯特空间连续线性泛函的范数及其应用[J].高校应用数学学报(A辑),2022,37(1): 123-126.  
WANG Yuchen, YANG Shijun. The norm of continuous linear functional on the reproducing kernel Hilbert space and its applications[J]. Applied Mathematics A Journal of Chinese Universities(Ser.A), 2022, 37(1): 123-126.
- [18] LONG Mingsheng, ZHU Han, WANG Jianmin, et al. Deep transfer learning with joint adaptation networks[C]//Proceedings

- of the 34th International Conference on Machine Learning.[S.l.]: [s.n.], 2017.
- [19] YOSINSKI J, CLUNE J, BENGIO Y, et al. How transferable are features in deep neural networks[C]//Proceedings of the 27th International Conference on Neural Information Processing Systems (NIPS).[S.l.]: [s.n.], 2014: 3320-3328.
- [20] 陈景霞,郝为,张鹏伟,等.基于混合神经网络的脑电时空特征情感分类[J].软件学报,2021,32(12): 3869-3883.  
CHEN Jingxia, HAO Wei, ZHANG Pengwei, et al. Emotion classification of spatiotemporal EEG features using hybrid neural networks[J]. Journal of Software, 2021, 32(12): 3869-3883.
- [21] YOON K. Convolutional neural networks for sentence classification[C]//Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP). Doha, Qatar: Association for Computational Linguistics, 2014: 1746-1751.
- [22] SOUMITRO C, EMANUEL A P H. Multi-speaker DOA estimation using deep convolutional networks trained with noise signals[J]. IEEE Journal of Selected Topics in Signal Processing, 2019. DOI: 10.1109/JSTSP.2019.2901664.
- [23] CHUNG J, GULCEHRE C, CHO K, et al. Empirical evaluation of gated recurrent neural networks on sequence modeling[J]. Neural and Evolutionary Computing, 2014. DOI:10.48550/arXiv.1412.3555.
- [24] SHI Lichen, JIAO Yingying, LU Baoliang. Differential entropy feature for EEG-based vigilance estimation[C]//Proceedings of the 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). Osaka, Japan: [s.n.], 2013: 6627-6630.
- [25] 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.
- [26] LONG Mingsheng, WANG Jianmin, DING Guiguang, et al. Transfer feature learning with joint distribution adaptation[C]//Proceedings of IEEE International Conference on Computer Vision. Sydney, Australia: IEEE, 2013: 2200-2207.
- [27] WANG Jindong, CHEN Yiqiang, HAO Shuji, et al. Balanced distribution adaptation for transfer learning[C]//Proceedings of 2017 IEEE International Conference on Data Mining(ICDM).New Orleans,LA: IEEE, 2017: 1129-1134.

#### 作者简介:



陈景霞(1979-),通信作者,女,博士,副教授,硕士生导师,研究方向:机器学习、脑电信号处理、脑机接口和情感计算, E-mail: chenjingxia@sust.edu.cn。



胡修文(1997-),女,硕士研究生,研究方向:脑电信号处理、迁移学习和情感计算。



唐喆喆(1998-),女,硕士研究生,研究方向:数据增强、情感计算。



刘洋(1996-),男,硕士研究生,研究方向:多模态、脑机接口。



胡凯蕾(1999-),女,硕士研究生,研究方向:睡眠检测、情感计算。

(编辑:王静)