

基于 Swin Transformer 的四维脑电情绪识别

陈宗楠, 金家瑞, 潘家辉*

(华南师范大学 软件学院, 广东 佛山 528225)

摘要:近年来,基于脑电图(Electroencephalogram, EEG)的情绪识别研究主要使用卷积神经网络、循环神经网络和深度信念网络模型。这些方法能利用全局差异来区分不同情绪状态,但忽视了局部脑电的变化对情绪状态的影响。针对上述问题,使用了一种基于 Swin Transformer 的 EEG 四维脑电情绪识别模型,能够更好地捕捉到细小的局部空间特征和复杂的时间序列特征。相较于其它情绪识别方法,该模型通过基于滑动窗口的自注意力机制提高了不同块之间的特征连接,使得模型的建模能力更强,也降低了计算的复杂度。此外,利用情绪脑电公开数据集 SEED 来评估该模型的可行性与有效性,在单被试情绪三分类的准确率为 $94.73\% \pm 1.72\%$, 跨被试情绪三分类的准确度为 $89.63\% \pm 3.42\%$, 并且测试速度能达到实时处理的水平。实验结果表明,基于 Swin Transformer 的 EEG 四维脑电情绪识别模型通过局部特征的学习在小样本训练上也能达到较高的情绪分类准确率和较快的测试速度。

关键词:深度学习;情绪识别;脑电图;特征融合;Swin Transformer

中图分类号:TP183

文献标识码:A

文章编号:1673-629X(2023)12-0178-07

doi:10.3969/j.issn.1673-629X.2023.12.025

Swin Transformer-based 4-D EEG Emotion Recognition

CHEN Zong-nan, JIN Jia-rui, PAN Jia-hui*

(School of Software, South China Normal University, Foshan 528225, China)

Abstract: In recent years, electroencephalogram (EEG)-based emotion recognition has focused on the use of convolutional neural networks, recurrent neural networks and deep belief network models. These methods can use global differences to distinguish between different emotional states, but ignore the effect of local EEG changes on emotional states. To address these issues, we use a 4-dimensional EEG emotion recognition model based on the Swin Transformer. The model can better capture both small local spatial features and complex time-series features. Compared with other emotion recognition methods, the model proposed improves the feature connectivity between different blocks through a self-attention mechanism based on shifted windows, which makes the model more modelable and also reduces the computational complexity. In addition, we use the public emotion EEG dataset SEED to evaluate the feasibility and effectiveness of this model, with an accuracy of $94.73\% \pm 1.72\%$ for single-subject emotion triple classification and $89.63\% \pm 3.42\%$ for cross-subject emotion triple classification, and the testing speed can reach the level of real-time processing. The experimental results show that 4-D EEG emotion recognition based on the Swin Transformer model can achieve high emotion classification accuracy and fast testing speed even with small sample training through local feature learning.

Key words: deep learning; emotion recognition; electroencephalogram (EEG); feature fusion; Swin Transformer

0 引言

情绪识别是计算机感知人类情感从而进行人机交互反馈的重要环节,在情感脑机接口领域得到了学术界和企业界的广泛关注。在众多情绪识别的方法中,脑电图(Electroencephalogram, EEG)是一种使用电生理指标记录大脑活动的方法,通过记录大脑正常活动的电信号变化,可以反映出脑神经细胞产生的各种生

理电信号在大脑皮层的变化情况。已有研究表明,脑电模式会随着生理与心理状态不同而变化,因此使用脑电图来进行情绪识别是可行的^[1],能够真实地反映人的情绪状态。

基于 EEG 的情绪识别是人工智能领域的一个重要分支。在各类基于 EEG 的情绪识别方法中,使用传统的机器学习如支持向量机(Support Vector Machine,

收稿日期:2023-03-04

修回日期:2023-07-05

基金项目:国家自然科学基金项目(62076103);科技创新2030项目-“脑科学与类脑研究”重点项目(2022ZD0208900)

作者简介:陈宗楠(1999-),男,硕士生,CCF会员(K6878G),研究方向为情感脑机接口、计算机视觉;通信作者:潘家辉(1982-),男,教授,博士,CCF会员(F7646M),研究方向为脑机接口、模式识别与智能系统。

SVM)、决策树等模型^[2]对复杂函数表达能力有限,且不同人的脑电之间存在领域偏移(Domain Shift)问题^[3],这类浅层模型容易受到不同领域数据分布变化的影响,一定程度上限制了机器学习模型分类复杂模型的能力,使得跨被试情绪识别准确率不高。近年来的工作主要使用基于深度学习的卷积神经网络(Convolution Neural Network, CNN)、循环神经网络(Recurrent Neural Network, RNN)和深度信念网络(Deep Belief Network, DBN)等来分类情绪,通过多层非线性网络结构能够更好地拟合复杂函数^[4-5]。Wang 等^[6]利用 3DCNN 同时捕获时间和空间特征,并配合批标准化和密集预测解决数据分布偏移和参考标签不可靠问题,其消融实验证实了上述模块的有效性。陈景霞等^[7]基于 EEG 信号的时域、频域特征,提出一种基于 CNN 的 EEG 情感特征学习与分类算法,在效价维度上平均准确率最高达到 88.76%,在唤醒度上平均准确率最高达到 85.57%。Zhang 等^[8]提出时空 RNN,包含空间和时间两层 RNN 结构,实验策略上选择被试依赖,达到了 89.50% 的准确率。Chen 等^[9]提出了一种基于估计熵的深度信念网络情绪识别框架,对情绪四分类的识别准确率达到 83.34%。然而,由于 EEG 信号的非平稳性所造成的被试间的不同数据分布差异^[4],针对单被试训练的分类模型在跨被试情绪识别时存在困难。这种分布差异使得利用固定分类模型对脑电情绪的预测效果随时间改变而下降,极大地限制了脑电情绪识别技术的实际应用。

脑电图的非平稳特性和个体差异限制了情绪识别模型在不同时间、不同受试者之间的泛化,为了解决跨被试情绪识别问题,部分研究团队使用了领域适配(Domain Adaptation)^[10-11]和对比学习(Contrastive Learning)方法^[12]。李劲鹏^[13]为了解决 EEG 情绪识别模型的跨被试适配问题,提出了一种多源迁移学习框架。在每个合适的源上,减少目标和源域的差异,使源域上的分类器直接用于推断目标中样本的情绪标签。Jin 等^[14]利用域对抗神经网络(DANN)建立跨被试情感识别模型,在训练阶段保持特征区分性和领域不变性,在 SEED 数据集上的平均准确率为 $79.19\% \pm 13.14\%$ 。郭苗苗等^[15]提出全局域适应与相关子域自适应串联系统模型来实现跨被试的情感识别,以解决 EEG 的非平稳性及个体差异性造成的情绪识别模型跨被试泛化性能低的问题。Shen 等^[12]提出了基于对比学习的跨被试对齐方法,通过最大化被试间脑电的相似性,来降低跨被试脑电的差异。DANN 等模型适配了源域和目标域等边缘分布,但来自多名被试多源域场景的脑电数据则无法进行对抗学习。如果被试间的相关性较低,单源域迁移学习可能会造成负迁移现

象。同时,深度神经网络能够对脑电分类进行端到端的自动学习,却忽略了脑电特征和情绪之间的内在关系,导致深度学习模型自动提取情绪特征的可解释性不如人工设计的情绪特征。因此,如何将传统的脑电特征提取方式与深度学习算法相结合,提高模型的可解释性和跨被试情绪识别性能是该文研究的重点。

该文使用 Swin Transformer^[16]模型用于情绪识别, Swin Transformer 是在 Transformer^[17]和 ViT (Vision Transformer)^[18]模型上的改进。具体使用上,针对脑电信号特征,该文提出了改进的 Swin Transformer 框架,进行模型优化。为了适配脑电通道数以及特征维度,修改了输入尺寸和维度,并通过修改滑动窗口大小、特征块大小和优化结构层数来降低模型参数数量和浮点运算次数。针对脑电情绪识别任务,通过改进的 Swin Transformer 将人工特征生成二次特征,映射到一个公共的特征空间,在跨被试情绪识别中也有较高的准确率。为验证模型改进效果,改进后的模型在 SEED 数据集上测试。通过融合频域和空间域的人工情绪特征,模型速度效率和准确率均有提升,并进行了融合特征的有效性分析。研究结果表明,使用改进的 Swin Transformer 框架在跨被试情绪识别中有较高的准确率和较快的测试速度,这是将传统机器学习人工提取特征与深度学习特征分类在情感脑机接口的情绪识别任务上的一次有益尝试。

1 数据预处理

1.1 脑电情绪识别框架

基于 Swin Transformer 的 EEG 四维脑电情绪识别框架如图 1 所示,由三部分组成:脑电人工特征提取、特征组合与融合、改进的 Swin Transformer 情绪识别模型。在脑电特征提取中,将预处理后的脑电源数据从时域、频域和空间域特征提取,得到不同的人工脑电特征。在特征组合与融合中,将脑电特征进行多特征组合形成四维脑电特征数据,4 个维度包含单通道特征、脑电多通道、单特征维度、多特征组合。最后将组合特征输入改进的 Swin Transformer 情绪识别模型,输入的特征大小是单通道特征长为 56 的 56 个脑电通道集合,特征深度由组合的原始各个特征维度决定。依照情绪分类任务的类目数量,改变全连接层实现情绪多分类。

1.2 脑电特征提取

为提高情绪识别的可解释性,并使传统的脑电特征提取方式与深度学习算法相结合,该文从以下人工特征中开展实验,包含时域、频域、空间域。时域特征包括一阶差分、二阶差分、Hjorth Mobility 特征、Hjorth Complexity 特征和不稳定指数(Non-Stationary Index,

NSI);频域特征包括微分熵(Differential Entropy, DE)和功率谱密度(Power Spectral Density, PSD);空间域特征包括不对称差(Differential Asymmetry, DASM)、不对称商(Rational Asymmetry, RASM)和不对称系数

(Asymmetry Index, AsI)。针对频域特征,脑电按频率分为 5 个频段: δ (1 ~ 4 Hz), θ (4 ~ 8 Hz), α (8 ~ 13 Hz), β (13 ~ 30 Hz), γ (30 ~ 47 Hz)。各情绪特征的特征参数根据具体特征维度有所不同,如表 1 所示。

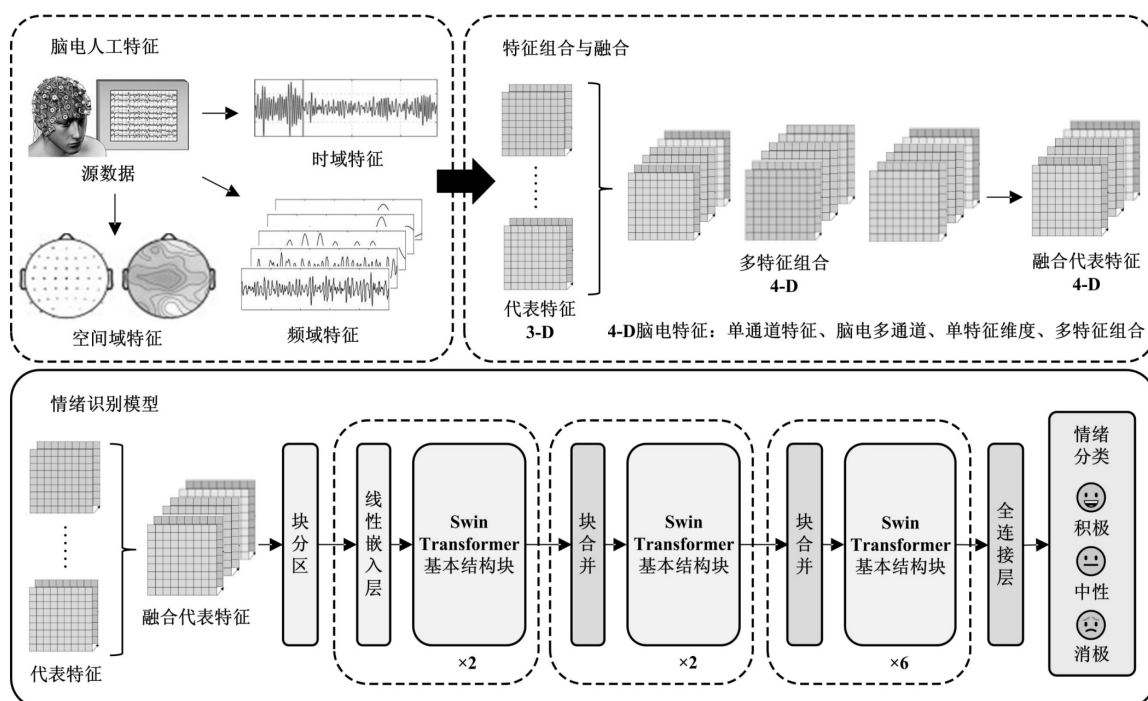


图 1 基于 Swin Transformer 的 EEG 四维脑电情绪识别框架

表 1 情绪特征参数

情绪特征	特征大小	情绪特征	特征大小
一阶差分 ^[19]	(56,56,25)	DE ^[4]	(56,56,5)
二阶差分 ^[19]	(56,56,25)	PSD ^[22]	(56,56,5)
Hjorth Mobility ^[20]	(56,56,1)	DASM ^[23]	(56,56,25)
Hjorth Complexity ^[20]	(56,56,1)	RASM ^[23]	(56,56,25)
NSI ^[21]	(56,56,1)	AsI ^[24]	(56,56,25)

1.3 特征组合与融合

为了脑电数据能够符合文中 Swin Transformer 模型的输入大小,对源数据的通道和时长进行了选择和 处理。模型输入的尺寸大小是(56,56),第一维的 56 表示将选择 56 个脑电通道。因为大脑的额叶部分主要控制人们的情感,与情绪密切相关^[25],所以例如在 62 导脑电设备中剔除了非额叶部分 6 个对脑电特征影响较小的通道:P7, P8, PO7, PO8, CB1 和 CB2。第二维的 56 表示一个脑电通道在一个时间片下的特征长度为 56。该文对受试者每个实验单次诱导情绪片段截取了 175 秒脑电数据,共 35 000 个脑电采样点,按 7 秒为时间片切分为 25 组,每组有 1 400 个脑电采样点。以微分熵为例,将每 25 个采样点计算 1 次微分熵特征,使得该时间片微分熵特征是长度为 56 的向量,与 56 个通道和 5 个频段结合后,一个 7 秒时间片的特

征大小为(56,56,5)。最后将这 25 组时间片叠加成(25,56,56,5),代表了观看一段诱导视频中按 7 秒为一个时间片分为了 25 组,其中包含 56 通道脑电数据各自的长为 56 的特征,这些特征按频段或者特征组数在维度上进行叠加。

2 情绪识别模型

Swin Transformer^[16]是 2021 年提出的一种新型的 Transformer 架构,通过一个层次化的 Transformer 来解决二维数据的训练难点,其表示是用滑动窗口计算的。根据脑电特征图的大小和维度,对其模型结构进行了适配,图 2 是适配后的模型结构及其输入特征尺寸和维度的变化细节。针对脑电情绪识别任务, Swin Transformer 模型适配了人工脑电特征的输入尺寸和维度,将人工特征生成二次特征,映射到一个公共的特

征空间,使得跨被试情绪识别效果更好。模型通过优化与改进结构,修改特征块(Patch)的大小和通用结构

层数,降低了模型参数数量和浮点运算次数,使其有较快的运行速度。

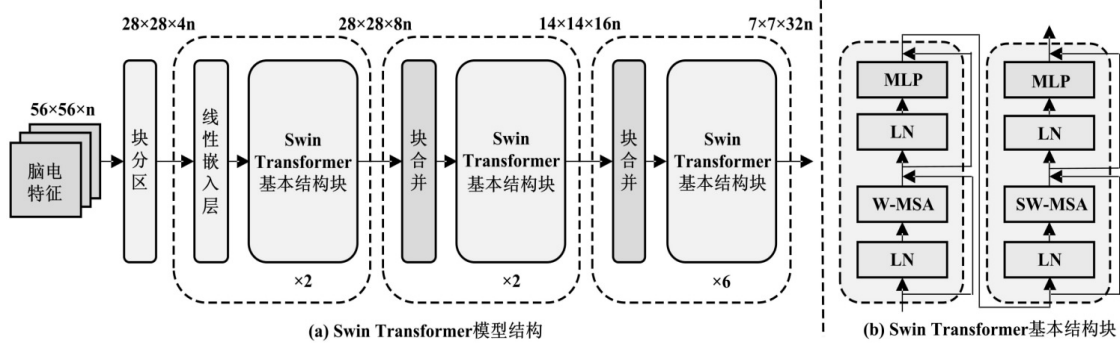


图 2 改进的 Swin Transformer 模型

为适配 Swin Transformer 模型的输入结构,输入模型的脑电特征大小定为 56×56 ,输入特征的维度由具体特征或特征组合而定,假定为 n 维,则通过 2×2 的特征块分区转化为 $28 \times 28 \times 4n$ 的大小。根据人工脑电特征的构建,每个 2×2 特征块包含相邻时间和相邻通道的特征数据。块合并(Patch Merging)类似于 CNN 的池化层,能够让模型产生多尺度的特征。随着模型的深入,经过块合并后每个块会拥有成倍增加的视野,使得表征能力更强。线性嵌入层(Linear Embedding)把特征维度升至 $8n$,提高了不同特征间的可区分性。在 Swin Transformer 基本模块中,使用基于窗口的掩码自注意力机制(W-MSA)对 2×2 特征块内计算自注意力,再使用基于滑动窗口的掩码自注意力机制(SW-MSA)使窗口之间进行互动,不同块的特征之间产生联系。滑动窗口方案将自注意力机制限制在非重叠的局部窗口上,通过允许跨窗口连接从而提高了效率。这种分层结构具有在不同尺度下建模的灵活性,并且相对于图像大小具有线性计算复杂性,能够提高模型的运行速度。通过线性嵌入层后,单层特征会被拉直为序列,使用层归一化(Layer Normalization, LN)在通道方向上进行归一化操作,最后通过一个全连接层(Multilayer Perceptron, MLP)输出到下一个模块。鉴于原始特征输入大小为 56,对最小的 Swin-Tiny 模型的模块进行修改,删去了最后一个块合并和基本结构块,整个模型由 3 个阶段组成,如图 2 所示,最后的输出再接上对应三分类的全连接层即可进行脑电情绪三分类。

3 实验结果与分析

3.1 数据集

实验数据集使用上海交通大学提供的情绪脑电数据库(Shanghai Jiao Tong University, Emotion Electroencephalogram Dataset, SEED)^[4]。SEED 包含 15 名受试者(7 名男性,平均年龄 23.27,标准差为 2.37)在观

看电影片段时收集的脑电图。电影片段诱发受试者产生不同类型的情绪:积极、中性和消极。15 个电影片段包含 5 个积极片段、5 个中性片段和 5 个消极片段。15 名受试者观看约 4 分钟的单个电影片段时,使用 62 通道的 ESI Neuro Scan 系统采集脑电信号,相应的 EEG 电极布置如图 3 所示。

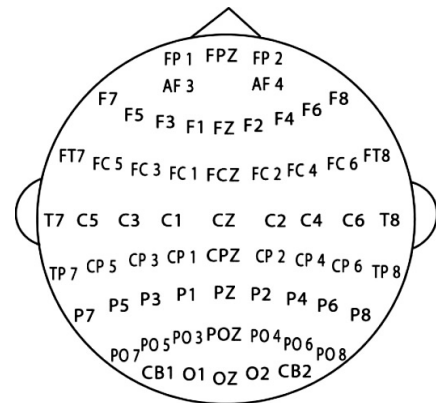


图 3 SEED 的 EEG 电极通道

SEED 提供了 15 名受试者各进行 3 次实验的脑电数据,共 45 次实验。每名受试者每次实验分为 15 段诱导脑电,其中表示积极、中性和消极的脑电数据各 5 段。单个脑电数据包含 62 通道上记录的脑电图,对其进行下采样处理,从原始 1 000 Hz 的 EEG 数据降采样到 200 Hz。为了滤除噪声和去伪影,使用 0.3 ~ 50 Hz 的带通滤波器对 EEG 数据进行预处理。该文将在 SEED 提供的预处理脑电上进行实验。

3.2 实验设计

为了研究 Swin Transformer 模型在不同特征组合下单被试、跨被试的情绪分类能力,设计了三种实验,分别为代表情绪特征的选取、单被试情绪识别和跨被试情绪识别,最后通过消融实验在单被试和跨被试中验证特征融合的有效性。

在代表情绪特征的选取中,使用 Swin Transformer 模型分别对各个特征单独做情绪识别测试,以选取有代表性的一种或多种特征来代表情绪,实验中训练集

和测试集的比例为 8 : 2。其中单被试准确率表示使用该特征分别测试 15 名受试者,取所有受试者准确率的平均值作为该特征的单被试准确率。在跨被试实验中,将 SEED 的 15 名受试者分别划分为 15 个独立的域。当 1 名受试者作为目标域时,其脑电数据将作为测试数据计算跨被试准确率,其余 14 名受试者将作为训练数据生成二次特征,映射到一个公共的特征空间,最终取各受试者作为目标域的测试准确率平均值作为该特征的跨被试准确率。通过对比各特征单被试和跨被试准确率,选取效果较好的人工特征融合为代表情绪特征,达到提高情绪识别效果的目的。

在代表情绪特征的单被试和跨被试情绪识别中,单被试和跨被试的实验设计与上述方法相同。其目的是从融合的不同人工脑电特征中获得最具差异性的信息,利用特征之间的互补性,融合特征之间的优点,进而提高模型的性能。为保证融合代表特征中每一个特征都起到了提升模型情绪识别能力的作用,消融实验则通过测试代表情绪特征不同组合的情绪识别准确率,以验证单个特征对模型情绪识别能力提升的贡献。

3.3 结果与分析

在代表情绪特征的选取中频域 DE 特征的单被试准确率最高(90.24%),频域 PSD 特征的跨被试准确率最高(82.16%)。空间域的特征表现不如频域特征优秀,但它能补充频域特征在空间分布上的位置信息。时域特征则表现较差,在情绪三分类任务里准确率较低,并不能达到很好的情绪表征能力。各个特征的情绪识别准确率如表 2 所示。结合不同特征的特点和测试准确度,选用单被试准确率高于 80% 且跨被试准确率表现较好的 3 个特征,即频域的微分熵和功率谱密度、空间域的不对称系数组成融合代表特征。

表 2 单特征情绪识别准确率 %

特征	单被试准确率	跨被试准确率
一阶差分 ^[19]	45.82	38.76
二阶差分 ^[19]	44.46	38.12
Hjorth Mobility ^[20]	43.63	36.78
Hjorth Complexity ^[20]	66.71	52.33
NSI ^[21]	41.87	33.46
DE ^[4]	90.24	81.62
PSD ^[22]	89.32	82.16
DASM ^[23]	73.31	61.24
RASM ^[23]	76.89	62.83
AsI ^[24]	81.43	72.38

在代表情绪特征的单被试情绪识别中,测试准确率如表 3 所示,该表包含了 15 名被试各自模型的情绪识别准确率。相较于单一特征,融合了 3 个特征的组

合特征在情绪识别准确率更高,平均准确率达到了 94.73%,标准差为 1.72%,其中最高的单被试准确率达到了 96.89%。

表 3 单被试情绪识别准确率 %

编号	准确率	编号	准确率	编号	准确率	编号	准确率
1	96.89	5	95.11	9	95.56	13	94.22
2	96.00	6	96.44	10	92.00	14	93.33
3	91.11	7	94.67	11	94.22	15	96.89
4	96.44	8	92.89	12	95.11	平均	94.73±1.72

模型在各个被试上训练测试的结果较为稳定,皆能达到 90% 以上的结果。各个被试的融合代表特征情绪识别准确率高于原先单一特征的 DE 最高平均准确率 90.24%,说明多个有效特征的融合有利于情绪识别能力的提升。

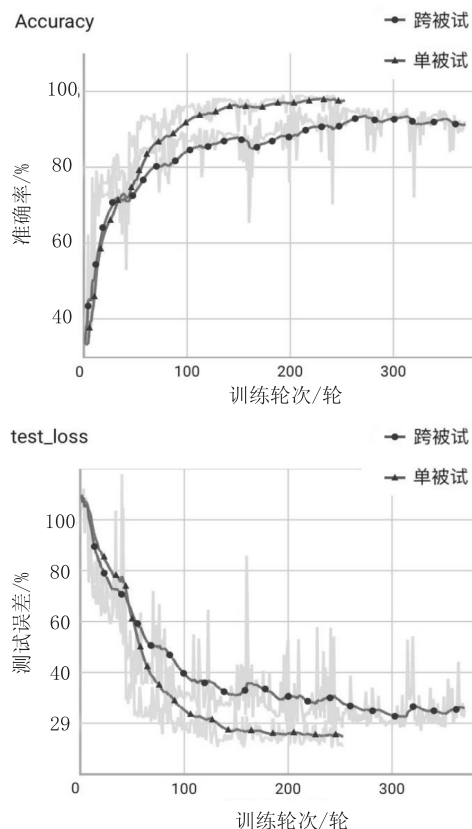


图 4 文中模型的准确率与测试误差

图 4 展示了以第一位受试者作为单被试训练以及跨被试的目标域时,其脑电数据作为测试数据计算单被试和跨被试的准确率和测试误差的结果。从结果来看,模型在第 150 轮左右开始收敛,在单被试测试上能达到约 95% 的准确率,在跨被试测试上能达到约 90% 的准确率。为保证模型测试的稳定性,对 15 名受试者进行单被试和跨被试测试的准确率取平均值记录,用该值代表模型的平均性能。将文中模型与近三年脑电情绪识别模型在 SEED 的识别准确率进行对比,如表

4 所示。

表 4 情绪识别的准确率对比 %

模型	单被试准确率	跨被试准确率
STNN ^[8]	89.50	—
BiDANN ^[3]	92.38±7.04	83.28±9.60
R2G-STNN ^[26]	93.38±5.96	84.16±7.63
BiHDM ^[27]	93.12±6.06	85.40±7.53
ATDD-LSTM ^[28]	91.08±5.34	90.92±1.05
DGCNN ^[29]	90.40±8.49	79.95±9.02
CLISA ^[12]	—	86.40±6.40
SVM-KNN ^[30]	—	88.84
DE-CNN-GRU ^[31]	—	86.50
SS_GDAN_RSAN ^[15]	91.66±7.32	84.05±5.91
文中模型	94.73±1.72	89.63±3.42

利用 Swin Transformer 使用层次化 Transformer 解决二维数据训练困难的特点,将脑电特征构造成相应的模型输入,在单被试和跨被试中都取得了较高的脑电情绪识别准确率。文中模型的输入融合了不同特征,相较于单个模态,有更好的情绪识别准确率。在单被试中,相较于 STNN^[8],R2G-STNN^[26],BiDANN^[3],BiHDM^[27],ATDD-LSTM^[28]的单模态模型输入,文中模型通过融合多个有效特征提升模型的情绪识别能力。与多模态输入的 DGCNN^[29]相比,DGCNN 通过动态图卷积神经网络将不同通道的 EEG 用邻接矩阵学习空间域特征,文中模型使用人工特征中的不对称系数来代表脑电的空间域特征,在单被试实验准确率和模型可解释性上都更优秀。在跨被试中,文中模型依然可以达到接近 90% 准确率,高于 PSD 单特征的 82.16%。由于没有在代表特征中融合进有效的时域人工特征,仅使用 7 秒的时间片来计算每种特征,文中模型在跨被试准确率上比 ATDD-LSTM^[28]的 90.92% 低 1.29 百分点,但仍然在跨被试情绪识别上有较高的分类准确率。

此外在模型速度方面,文中模型的测试速度能达到实时处理的水平。相比起原本的 Swin-T 模型 29 M 模型参数量和 4.5 G 的浮点运算次数,修改后的模型使用的模型参数量为 12.1 M,浮点运算次数为 902.6 M。在实验训练条件的脑电识别测试时,处理一段 7 s 的脑电平均用时 0.37 s,说明在一些计算能力较弱的环境以及未来实际应用时能够较为容易达到实时处理的水平。

从消融实验结果中可知,使用多种人工特征组合后的效果要优于单一人工特征的脑电情绪识别效果。在单被试脑电情绪识别中,使用 PSD 和 AsI 能够给单一 DE 特征带来 3.57 百分点和 2.69 百分点的提升,使

用 DE 和 AsI 能够给单一 PSD 特征带来 4.49 百分点和 3.14 百分点的提升,使用 DE 和 PSD 能够给 AsI 特征带来 11.5 百分点和 11.03 百分点的提升。在跨被试情绪识别上也有同样的准确率提升,平均能够提高约 3.2 百分点。使用 3 种特征的组合特征效果最好,能够在单被试达到 94.73%、多被试达到 89.63% 的情绪识别准确率。这说明在融合代表特征种的 3 种人工特征对脑电情绪识别的结果都有各自的贡献,通过特征组合来提高情绪识别能力是有效的。

表 5 消融实验

DE	PSD	AsI	单被试 准确率/%	跨被试 准确率/%
✓			90.24	81.62
	✓		89.32	82.16
		✓	81.43	72.38
✓	✓		93.81	85.65
✓		✓	92.93	85.17
	✓	✓	92.46	84.56
✓	✓	✓	94.73	89.63

4 结束语

该文提出一种基于 Swin Transformer 的 EEG 四维脑电情绪识别模型来对脑电情绪进行分类研究。通过对 Swin Transformer 进行模型优化,修改了输入尺寸、维度和模型内滑动窗口、特征块的大小。研究结果表明,通过 Swin Transformer 进行跨时间或跨被试情绪识别,最终均得到了较高的准确率和较快的测试速度,在单被试达到了 94.73%±1.72% 的准确率,跨被试达到了 89.63%±3.42% 的情绪识别准确率。在与同类任务的其它模型的对比中,文中模型在效果上不仅有显著优势,并且在可解释性上也更好。测试速度上,模型优化结构后的参数量和浮点运算量能够满足实时脑电情绪识别的要求。这是将传统机器学习人工提取特征与深度学习特征分类在情感脑机接口的情绪识别任务上的一次有益尝试,为脑电情绪识别的实际应用提供了新的方法。

参考文献:

- [1] SAMMLER D, GRIGUTSCH M, FRITZ T, et al. Music and emotion: electrophysiological correlates of the processing of pleasant and unpleasant music[J]. Psychophysiology, 2007, 44(2):293-304.
- [2] 蒋小梅,张俊然,陈富琴,等. 基于 J48 决策树分类器的情绪识别与结果分析[J]. 计算机工程与设计, 2017, 38(3): 761-767.
- [3] LI Y, ZHENG W, CUI Z, et al. A novel neural network mod-

- el based on cerebral hemispheric asymmetry for EEG emotion recognition [C]//Proceedings of the 27th international joint conference on artificial intelligence. Stockholm: International Joint Conferences on Artificial Intelligence Organization, 2018: 1561–1567.
- [4] ZHENG W, LU B. Investigating critical frequency bands and channels for EEG-based emotion recognition with deep neural networks [J]. IEEE Transactions on Autonomous Mental Development, 2015, 7(3): 162–175.
- [5] LI J, ZHANG Z, HE H. Hierarchical convolutional neural networks for EEG-based emotion recognition [J]. Cognitive Computation, 2018, 10(2): 368–380.
- [6] WANG Y, HUANG Z, MCCANE B, et al. EmotioNet: a 3-D convolutional neural network for EEG-based emotion recognition [C]//2018 international joint conference on neural networks (IJCNN). Rio: International Neural Network Society, 2018: 1–7.
- [7] 陈景霞, 王丽艳, 贾小云, 等. 基于深度卷积神经网络的脑电信号情感识别 [J]. 计算机工程与应用, 2019, 55(18): 103–110.
- [8] ZHANG T, ZHENG W, CUI Zhen, et al. Spatial-temporal recurrent neural network for emotion recognition [J]. IEEE Transactions on Cybernetics, 2019, 49(3): 839–847.
- [9] CHEN T, JU S, YUAN X, et al. Emotion recognition using empirical mode decomposition and approximation entropy [J]. Computers and Electrical Engineering, 2018, 72: 383–392.
- [10] 王 翎. 域适应算法研究及应用 [D]. 南京: 南京航空航天大学, 2019.
- [11] 柴 鑫. 基于领域适配的跨个体脑电情绪识别方法研究 [D]. 哈尔滨: 哈尔滨工业大学, 2018.
- [12] SHEN X, LIU X, HU X, et al. Contrastive learning of subject-invariant EEG representations for cross-subject emotion recognition [J]. IEEE Transactions on Affective Computing, 2022, 4(4): 1–12.
- [13] 李劲鹏. 脑电情绪识别中跨被试迁移学习方法研究 [D]. 北京: 中国科学院大学, 2019.
- [14] JIN Y, LUO Y, ZHENG W, et al. EEG-based emotion recognition using domain adaptation network [C]//2017 international conference on orange technologies (ICOT). Singapore: Singapore Section SMC Chapter, 2017: 222–225.
- [15] 郭苗苗, 陈昕彤, 王 磊, 等. 子域自适应网络跨被试情绪识别算法 [J]. 信号处理, 2022, 38(10): 2211–2220.
- [16] LIU Z, LIN Y, CAO Y, et al. Swin transformer: hierarchical vision transformer using shifted windows [C]//IEEE/CVF international conference on computer vision. Montreal: IEEE, 2021: 10012–10022.
- [17] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [C]//Advances in neural information processing systems. California: Curran Associates, 2017: 30.
- [18] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16x16 words: transformers for image recognition at scale [J]. arXiv:2010.11929, 2020.
- [19] BASTOS-FILHO T F, FERREIRA A, ATENCIO A C, et al. Evaluation of feature extraction techniques in emotional state recognition [C]//2012 4th international conference on intelligent human computer interaction (IHCI). Kharagpur: [s. n.], 2012: 1–6.
- [20] HJORTH B. EEG analysis based on time domain properties [J]. Electroencephalography and Clinical Neurophysiology, 1970, 29(3): 306–310.
- [21] KROUPI E, YAZDANI A, EBRAHIMI T. EEG correlates of different emotional states elicited during watching music videos [C]//Affective computing and intelligent interaction. Berlin: Springer, 2011: 457–466.
- [22] JENKE R, PEER A, BUSS M. Feature extraction and selection for emotion recognition from EEG [J]. IEEE Transactions on Affective Computing, 2014, 5(3): 327–339.
- [23] DUAN R, WANG X, LU B. EEG-based emotion recognition in listening music by using support vector machine and linear dynamic system [C]//Neural information processing: 19th international conference. Doha: Neural Information Processing, 2012: 468–475.
- [24] KIM M K, KIM M, OH E, et al. A review on the computational methods for emotional state estimation from the human EEG [J]. Computational and Mathematical Methods in Medicine, 2013(4): 13.
- [25] ALARCÃO S M, FONSECA M J. Emotions recognition using EEG signals: a survey [J]. IEEE Transactions on Affective Computing, 2019, 10(3): 374–393.
- [26] LI Y, ZHENG W, WANG Lei, et al. From regional to global brain: a novel hierarchical spatial-temporal neural network model for EEG emotion recognition [J]. IEEE Transactions on Affective Computing, 2022, 13(2): 568–578.
- [27] LI Y, WANG L, ZHENG W, et al. A novel bi-hemispheric discrepancy model for EEG emotion recognition [J]. IEEE Transactions on Cognitive and Developmental Systems, 2021, 13(2): 354–367.
- [28] DU X, MA C, ZHANG G, et al. An efficient LSTM network for emotion recognition from multichannel EEG signals [J]. IEEE Transactions on Affective Computing, 2022, 13(3): 1528–1540.
- [29] SONG T, ZHENG W, SONG P, et al. EEG emotion recognition using dynamical graph convolutional neural networks [J]. IEEE Transactions on Affective Computing, 2020, 11(3): 532–541.
- [30] 滕凯迪, 赵 倩, 谭浩然, 等. 基于 SVM-KNN 算法的情绪脑电识别 [J]. 计算机系统应用, 2022, 31(2): 298–304.
- [31] 赵丹丹, 赵 倩, 董宜先, 等. 基于 EEG 和 DE-CNN-GRU 的情绪识别 [J]. 计算机系统应用, 2022, 32(4): 206–213.