

控制与决策

Control and Decision

基于三维特征矩阵和冲压激励网络的多通道脑电情感识别

晁浩, 曹益鸣, 刘永利

引用本文:

晁浩,曹益鸣,刘永利. 基于三维特征矩阵和冲压激励网络的多通道脑电情感识别[J]. *控制与决策*, 2023, 38(12): 3427–3435.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2021.1386>

您可能感兴趣的其他文章

Articles you may be interested in

[基于批次图像化的卷积自编码故障监测方法](#)

Fault detection of batch image-based convolutional autoencoder

控制与决策. 2021, 36(6): 1361–1367 <https://doi.org/10.13195/j.kzyjc.2019.1342>

[区分交通模式的混合服务路口信号控制策略](#)

Signal control strategies of mixed service intersections to discriminate traffic flow patterns

控制与决策. 2021, 36(6): 1509–1515 <https://doi.org/10.13195/j.kzyjc.2019.1520>

[基于双分支特征融合的场景文本检测方法](#)

A scene text detection based on dual-path feature fusion

控制与决策. 2021, 36(9): 2179–2186 <https://doi.org/10.13195/j.kzyjc.2020.0002>

[基于多层次特征的机械臂单阶段抓取位姿检测](#)

Single-stage grasp pose detection of manipulator based on multi-level features

控制与决策. 2021, 36(8): 1815–1824 <https://doi.org/10.13195/j.kzyjc.2019.1840>

[结合注意力机制的循环神经网络复述识别模型](#)

Recurrent neural networks based paraphrase identification model combined with attention mechanism

控制与决策. 2021, 36(1): 152–158 <https://doi.org/10.13195/j.kzyjc.2019.0638>

基于三维特征矩阵和冲压激励网络的多通道脑电情感识别

晁浩[†], 曹益鸣, 刘永利

(河南理工大学 计算机科学与技术学院, 河南 焦作 454000)

摘要: 提出一种基于冲压激励网络的情感状态识别方法. 首先, 从不同通道的脑电信号中提取时域特征, 并根据电极通道的相对位置构造三维特征矩阵; 然后, 将冲压激励块与三维卷积神经网络相结合构建冲压激励网络进行高层抽象特征提取; 最后, 使用全连接层进行情感状态分类. 实验在 DEAP 数据集上开展, 实验结果表明, 冲压激励网络在利用脑电信号中的时域显著性信息和电极空间位置信息的基础上, 可自适应地纠正特征的注意力, 优化每个特征的权重并强化重要特征, 同时利用不同特征的互补信息来提高识别精度; 此外, 冲压激励网络的挤压操作可获取输入数据的全局信息, 具有较快的收敛速度.

关键词: 情感识别; 多通道脑电; 冲压激励网络; 三维特征矩阵; 注意力

中图分类号: TP18 文献标志码: A

DOI: 10.13195/j.kzyjc.2021.1386

引用格式: 晁浩, 曹益鸣, 刘永利. 基于三维特征矩阵和冲压激励网络的多通道脑电情感识别[J]. 控制与决策, 2023, 38(12): 3427-3435.

Emotion recognition from multi-channel EEG data through three-dimensional feature matrix and squeeze-and-excitation networks

CHAO Hao[†], CAO Yi-ming, LIU Yong-li

(School of Computer Science and Technology, Henan Polytechnic University, Jiaozuo 454000, China)

Abstract: An emotion state recognition method based on squeeze-and-excitation networks is proposed. Firstly, time-domain features are extracted from electroencephalogram (EEG) signals of different channels, and a three-dimensional feature matrix is constructed according to the relative position of electrode channels. Then, the squeeze-and-excitation networks are constructed by combining the squeeze-and-excitation block with the 3D convolutional neural network for high-level abstract feature extraction. Finally, fully connected layers are used for emotional state classification. The experiment is carried out on the DEAP data set. The experimental results show that squeeze-and-excitation networks can adaptively correct the attention of features and optimize the weight of each feature based on the time-domain saliency information and electrode spatial position information in the EEG signal. Meanwhile, the squeeze-and-excitation networks can also strengthen important features and improve the recognition accuracy by using the complementary information of different features. In addition, the squeeze operation of the squeeze-and-excitation networks can obtain the global information of the input data and have faster convergence speed.

Keywords: emotion recognition; multi-channel EEG; squeeze-and-excitation networks; three-dimensional feature matrix; attention

0 引言

情感识别技术在人机交互中具有广阔的应用前景. 目前的情感识别研究大多基于面部表情、语音信号和生理信号. 在这些方法中, 生理信号受神经和内分泌系统支配, 具有自发性和不受主观观念控制的特

点. 因此, 生理信号往往更能够表征人的真实情感状态. 与其他常见的生理信号如心电图、皮肤电导和脉搏相比, 脑电信号 (electroencephalogram, EEG) 由于可反映中枢神经系统的动态变化^[1], 常常用于情感状态的自动检测.

收稿日期: 2021-08-08; 录用日期: 2022-07-04.

基金项目: 国家自然科学基金项目 (61872126); 河南省高等学校重点科研计划项目 (19A520004); 河南省高校基本科研业务费专项资金项目 (NSFRF1616).

责任编辑: 胡清华.

[†]通讯作者. E-mail: chaohao1981@163.com.

EEG情感识别中常用的特征有3种:时域特征、频域特征和时频域特征. 常见的时域特征有平均值、功率、方差、标准差、第1差异值和第2差异值等统计值. 此外,还有非平稳指标^[2]、Hjorth特征^[3],高阶交叉特征^[4]和事件相关电位^[5]也用于情感识别. 频域特征是从EEG信号不同频带中提取的,不同空间位置和频段的脑电频域特征对识别结果的影响不同. Lin等^[6]发现在 δ 、 θ 和 α 分量上的功率谱不对称指数能够敏感地反映与情感反应相关的大脑活动. Davidson^[7]在实验中证明,相对较活跃的左前额叶的脑电活动与积极的情感相关,而相对较活跃的右前额叶的脑电活动与消极的情感相关. 时频域特征主要有离散小波变换特征^[8]和希尔伯特-黄谱(HHS)^[9].

机器学习方法用于建立情感识别模型. 对于传统的分类方法, Samara等^[10]使用SVM融合统计测量值、来自 α 、 β 、 δ 和 θ 波的频带功率以及EEG信号的高阶交叉特征. Mehmood等^[11]采用Hjorth参数对所有EEG通道进行特征提取,并使用了支持向量机(SVM)和 K 近邻(KNN)两种不同的分类方法. 此外,决策树(DT)和多层感知器(MLP)也用于情感识别任务^[12]. 由于能够自动提取高层抽象特征,近年来深度学习也广泛应用于情感识别任务. Li等^[13]结合卷积神经网络(CNN)与循环神经网络(RNN)的混合深度模型提取特征,挖掘通道间的相关性. 文献[14]使用具有神经胶质链的深度信念网络挖掘通道间相关信息.

尽管基于EEG的情感识别已取得了很大进展,一些工作通过对脑电信号提取特征,构建经典分类算法和神经网络模型进行分类取得了不错的效果,但仍然存在一些挑战. 首先,EEG相邻电极通道蕴含与情感状态有关的空间位置信息,而以往的研究往往将不同的电极通道作为独立的个体;然后,不同特征对识别任务的贡献不同,使用CNN模型无法根据不同特征对于识别任务的贡献度进行建模,造成融合特征分类效果差于单一特征;最后,CNN中的卷积核只关注感受野中的信息,无法处理其他卷积核的上下文信息,因此无法学习输入数据的全局信息.

针对上述问题,本文根据电极在大脑中的排列位置,构建EEG信号时域特征组成的三维矩阵,然后利用3D-CNN网络挖掘电极通道的局部相关性信息. 使用冲压激励网络(squeeze-and-excitation networks, SENet)^[15]区分不同时域特征对情感识别结果的重要程度,即加强有用特征,抑制性能差的特征,同时利用不同特征间的互补信息提高识别精度和样本质量. SENet中的挤压操作可获取输入数据的全局信息,结合了不同卷积核的上下文信息,因此具有比

CNN更优异的学习能力.

1 相关工作

1.1 数据集和情感模型

DEAP是由伦敦玛丽皇后大学研究团队开发的用于分析人类情感状态的开放数据集^[16]. 该数据集记录了32名志愿者在选定视频刺激下的多模态生理信号,包括脑电和外周生理信号,32名志愿者中的22名还记录了正面视频. 每个志愿者需要观看40个1 min长的视频,使用32个有源AgCl电极记录脑电信号,这些电极根据国际10/20系统放置于头皮位置. 在每次实验结束时,参与者对他们的唤醒(arousal)、效价(valence)、喜好(liking)和支配(dominance)水平进行了0~9级的自我评估.

唤醒和效价是衡量情感状态的2个主要指标,根据这2个维度可建立二维情感模型. 图1为二维空间中的情感状态,其中2个维度分别为唤醒和效价. 因此,可在二维情感空间中定义4个标签,高唤醒-高效价(HAHV)、低唤醒-高效价(LAHV)、高唤醒-低效价(HALV)、低唤醒-低效价(LALV). 若唤醒和效价的自我评估值均大于5,则将实验样本分配至HAHV情感类别.

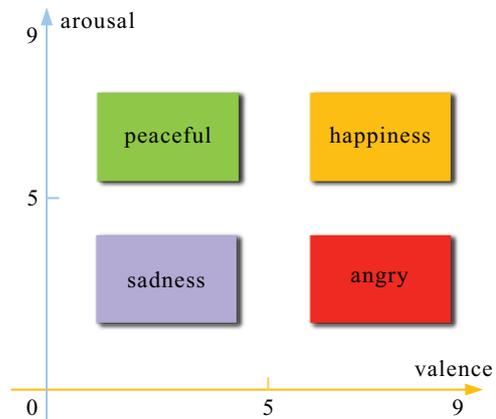


图1 arousal-valence情感空间

1.2 特征提取

在DEAP集中,每个人观看了40个情感诱导视频,每个视频中记录的脑电信号时长为60 s. 实验中将60 s脑电分为10个片段,因此可获得12 800个样本. 分别提取每个样本的32通道脑电信号的时域特征,包括均值、方差、标准差、第1差异值、第2差异值和模糊熵.

模糊熵是由Chen等^[17]于2009年提出的. 模糊函数的软连续边界保证了FuzzyEn在小参数下的连续性和有效性,因此,与近似熵和样本熵相比,模糊函数获得的更多细节也使得FuzzyEn成为更准确的熵定义.

2 基于冲压激励网络的情感识别模型

2.1 SENet模型框架

冲压激励网络模型包含了冲压激励块、卷积层和全连接层,通过SE块和卷积层的相互堆叠构成了一

个冲压激励网络,具体如图2所示. 卷积操作可从原始特征矩阵中挖掘不同电极通道的耦合关联性信息,冲压激励块可学习使用全局信息以选择性地强调信息特征,并抑制不太有用的特征.

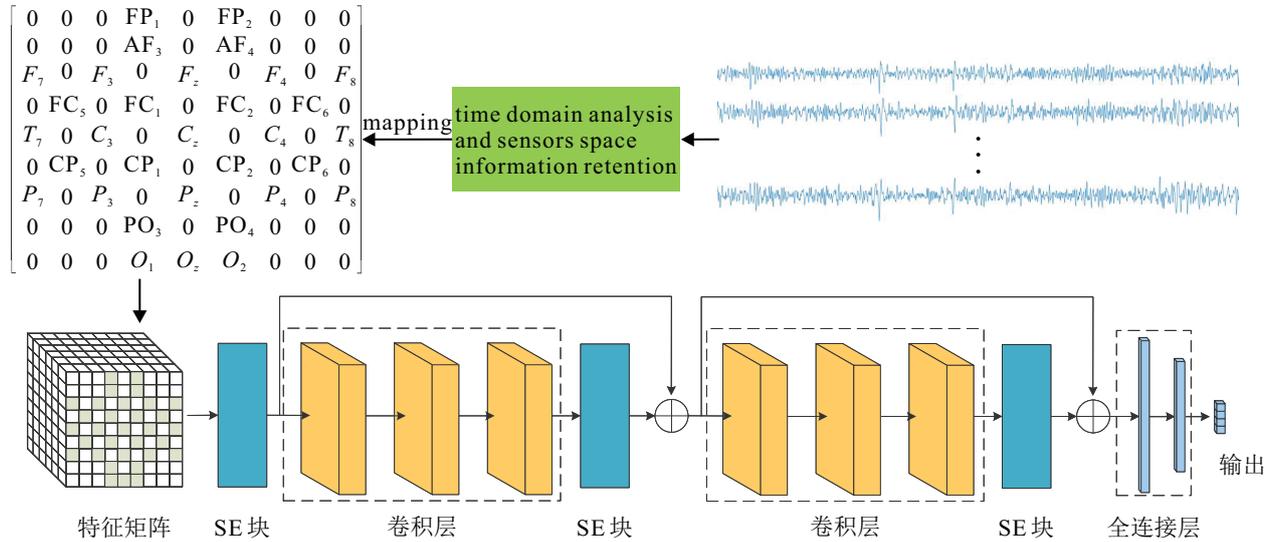


图2 SENet的结构

在该框架中,输入的是EEG信号时域特征构成的三维特征矩阵. 首先,对特征矩阵进行冲压激励运算,为每种特征分配一个权重;然后,将经过冲压激励操作后的特征矩阵输入至由卷积层和SE块构成的残差网络块,模型的卷积运算均使用RELU激活函数;最后,将残差网络块提取的高层抽象特征输入至2个全连接层进行分类. 为了防止网络过拟合,在全连接

层之后使用了2个dropout层. Softmax函数输出情感状态.

2.2 三维特征矩阵

图3为国际10/20系统的平面图及其二维矩阵. 图3左侧为国际10/20系统,其中实心圆圈标出的EEG电极为DEAP数据集中使用的测试点,空心圆圈为没有使用的测试点.

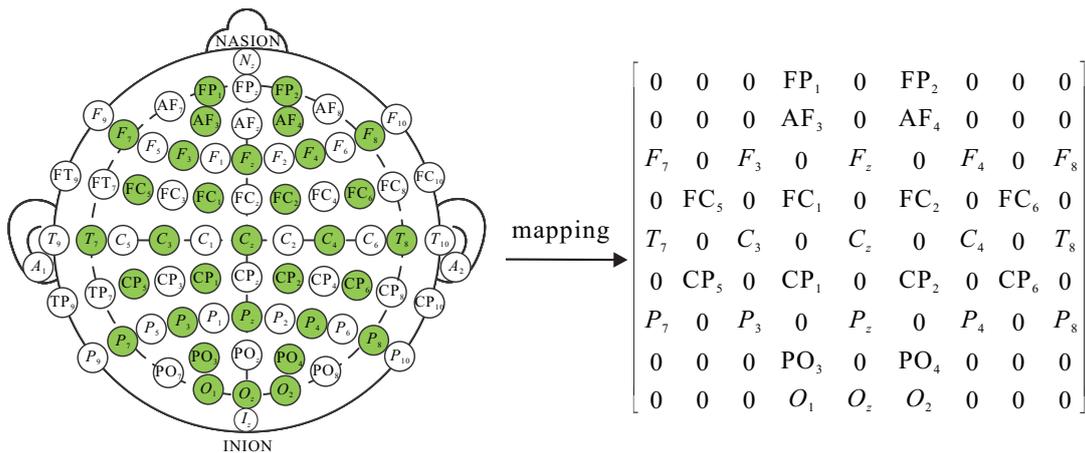


图3 国际10/20系统和9x9特征矩阵

利用脑电通道的全局信息和空间特征可提高情感识别的性能,为了表征所有脑电信号通道的空间信息和全局同步信息,根据电极在大脑上的位置构造特征矩阵,挖掘不同脑电信号通道的空间位置信息. 在特征矩阵中,将从不同脑电通道中提取的时域特征按

其相对位置坐标放入矩阵中对应的位置,矩阵中未使用电极的位置设为0.

对于每个样本,根据提取的6个时域特征按照图3所示的映射规则分别构建一个9x9x1矩阵,然后将6个时域特征映射的二维矩阵在第三维合并为一

个 $9 \times 9 \times 6$ 的三维特征矩阵,如图4所示.最后将构造的三维特征矩阵输入至神经网络模型.

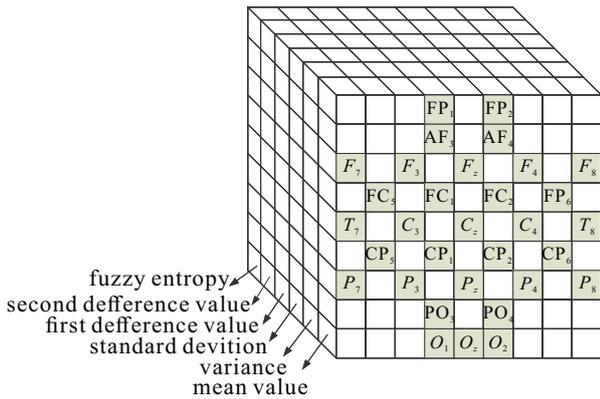


图4 3D特征矩阵

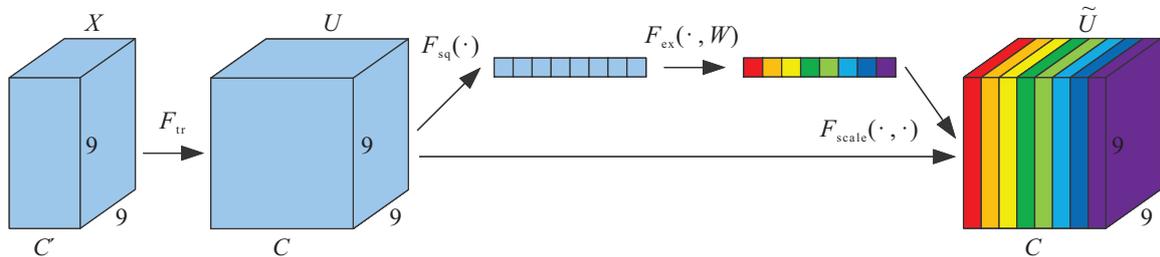


图5 冲压激励块

激励操作利用挤压操作收集到的信息来获得通道间的完全依赖,这是通过2个全连接层和2个激活层实现的.这样做的目的是:首先,它可学习不同特征间的非线性交互;然后,它关注所有特征,而不是强制执行单一热激活.激励运算如下式所示:

$$s = F_{\text{ex}}(z, W) = \sigma(g(z, W)) = \sigma(W_2(\delta(W_1 z))). \quad (2)$$

其中: σ 为 Sigmoid 激活函数, δ 为 ReLU 激活函数, $W_1 \in \mathbf{R}^{r \times 6}$, $W_2 \in \mathbf{R}^{6 \times r}$. F_{ex} 中2个全连接层的输出维度与输入维度一致, r 的改变可增加全连接层的复杂性, r 的大小在性能与计算成本的平衡下选择. 最终输出 \tilde{U} 是通过将 F_{ex} 的输出 s 中的元素与 U 的通道(6种时域特征对应的二维特征矩阵)相乘得到

$$\tilde{x}_c = F_{\text{scale}}(u_c, s_c) = u_c s_c, \quad (3)$$

其中 $\tilde{X} = [\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_C]$, 到此冲压激励运算完成了对各特征分配权重的操作. 在反向传播中, 通过 RELU 和 Sigmoid 激活函数完成对权重参数的更新, 最终达到强化对分类作用比较大的特征, 弱化作用小的特征.

当 SE 块位于网络模型中时, C' 和 C 为提取到的高层抽象特征图的通道数, $X \in \mathbf{R}^{9 \times 9 \times C'}$. 以 $V = [v_1,$

2.3 冲压激励运算

SE block 侧重于通道关系, 是一个计算单元, 其结构如图5所示. 将6种时域特征构成的三维特征矩阵输入至 SE 块, 先进行冲压激励运算, 卷积运算被取消, 则 $U = [u_1, u_2, \dots, u_6]$, $C=6$, $U \in \mathbf{R}^{9 \times 9 \times 6}$. 其中 u_1 、 u_2 、 u_3 、 u_4 、 u_5 和 u_6 分别为由平均值、方差、标准差、第1差异值、第2差异值和模糊熵特征构成的二维特征矩阵. 然后挤压模块将输入数据的全局信息放入通道描述符中. 这是通过全局平均池化实现的. 通道依赖描述 $z \in \mathbf{R}^6$ 由 U 在 9×9 维度上的全局平均池化生成. 因此, z_i 的第 c 个元素计算如下:

$$z_c = F_{\text{sq}}(u_c) = \frac{1}{H \times W} \sum_{i=1}^9 \sum_{j=1}^9 u_c(i, j). \quad (1)$$

$v_2, \dots, v_C]$ 表示滤波器集合, 其中 v_i 为第 i 个滤波器的参数. 然后可将输出写为 $U = [u_1, u_2, \dots, u_C]$, 其中

$$u_c = v_c \times X = \sum_{s=1}^{C'} v_c^s \times x^s. \quad (4)$$

挤压操作将输入的特征图的全局信息放入通道描述符 $Z \in \mathbf{R}^C$ 中, 因此挤压操作中结合 F_{tr} 中不同卷积核的上下文信息, 避免了 CNN 中只关注卷积核感受野内的信息, 因此 SENet 与 CNN 相比能够更好地学习输入数据的分布, 在反向传播过程中有更快的收敛速度.

3 实验结果和分析

所提出方法在 DEAP 数据集上进行验证. 深度学习模型需要大量数据才能得到满意的结果, 因此采用时间分割的方法增加样本数量. 首先, 删除每个 EEG 信号中的前 3 s 基线; 然后, 通过滑动窗口技术划分每次实验的原始脑电信号. 为了保证提取特征的有效性和足够的样本数量, 滑动窗口的持续时间设置为 6 s. 对于提取的 12 800 个样本, 采用 10 折交叉验证技术对实验结果进行验证. 此外, 网络中的 SE 块中的 r 设置为 6, 交叉熵作为损失函数, 使用初始学习率为 0.001 的 Adam 优化器最小化损失函数.

3.1 CNN和SENet的结果分析

为了研究SENet的学习性能,实验中共设置了3组网络模型(G_1 、 G_2 、 G_3),每组网络均包含1个CNN和SENet。 G_1 、 G_2 和 G_3 间的区别是网络中的卷积核设置为不同大小,如表1所示, G_1 、 G_2 和 G_3 中的卷积核大小全部设置为 3×3 、 5×5 和 7×7 ,3组网络中SENet与CNN的卷积核大小保持一致。此外,同组中的SENet与CNN相比仅具有更多的SE块,即 G_1 、 G_2 和 G_3 中的CNN相比于同组中的SENet未进行冲压激励运算。其他超参数均设置为相同的值,这样可以非常客观地展示2个网络间的差异。6个模型执行相同的情感识别任务,识别性能的指标采用精度、 F_1 分数(F_1 score)、精确率(Precision)、召回率(Recall)和Kappa值。

表1 G_1 、 G_2 和 G_3 中CNN和SENet的超参数

	G_1		G_2		G_3	
	CNN1	SENet2	CNN2	SENet2	CNN3	SENet3
C_1	3×3	3×3	5×5	5×5	7×7	7×7
C_2	3×3	3×3	5×5	5×5	7×7	7×7
C_3	3×3	3×3	5×5	5×5	7×7	7×7
C_1	3×3	3×3	5×5	5×5	7×7	7×7
C_2	3×3	3×3	5×5	5×5	7×7	7×7
C_3	3×3	3×3	5×5	5×5	7×7	7×7
F_1	2.048	2.048	2.048	2.048	2.048	2.048
F_2	1.024	1.024	1.024	1.024	1.024	1.024

使用唤醒、效价和四分类的情感标记方案的识别结果如表2~表4所示。在唤醒维度上,3个SENet模型的识别精度相比于CNN分别提升了3.04%、4.75%和3.79%, F_1 分数分别提升了1.57%、6.00%

表2 G_1 、 G_2 和 G_3 在唤醒维度上的情感识别结果

	Acc	F_1	Pre	Recall	Kappa
CNN1	0.8752	0.8589	0.8403	0.8784	0.7481
SENet1	0.9056	0.8746	0.9063	0.8452	0.7861
CNN2	0.8518	0.8341	0.8395	0.8288	0.7109
SENet2	0.8993	0.8941	0.9127	0.8761	0.8173
CNN3	0.8670	0.8642	0.8651	0.8634	0.7623
SENet3	0.9049	0.8860	0.8942	0.8780	0.8018

表3 G_1 、 G_2 和 G_3 在效价维度上的情感识别结果

	Acc	F_1	Pre	Recall	Kappa
CNN1	0.8763	0.8647	0.8431	0.8875	0.7578
SENet1	0.9039	0.8932	0.8761	0.9112	0.8097
CNN2	0.8601	0.8154	0.8263	0.8067	0.6801
SENet2	0.8867	0.8782	0.8492	0.9093	0.7804
CNN3	0.8384	0.8337	0.8145	0.8539	0.7029
SENet3	0.8904	0.8911	0.8770	0.9056	0.8064

表4 G_1 、 G_2 和 G_3 在四分类维度上的情感识别结果

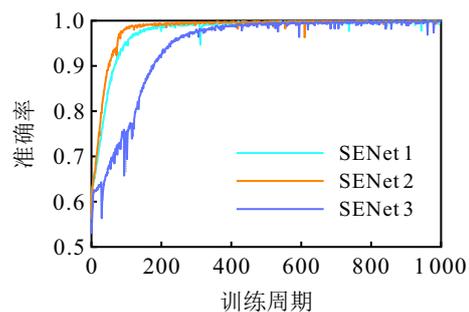
	Acc	F_1	Pre	Recall	Kappa
CNN1	0.8438	0.8403	0.8729	0.8677	0.8270
SENet1	0.8862	0.8799	0.8804	0.8798	0.8373
CNN2	0.8093	0.8048	0.8087	0.8014	0.7397
SENet2	0.8404	0.8519	0.8514	0.8530	0.8024
CNN3	0.8105	0.8016	0.8037	0.8000	0.7316
SENet3	0.8510	0.8397	0.8386	0.8416	0.7871

和2.18%;在效价维度上,3个SENet模型的识别精度相比于CNN分别提升了2.76%、2.66%和5.20%, F_1 分数分别提升了2.85%、6.28%和5.74%;在四分类中,3个SENet模型的识别精度相比于CNN分别提升了4.59%、3.11%和4.05%, F_1 分数分别提升了3.96%、4.71%和3.81%。所提出神经网络在其他4种评价方法上的表现也优于对比神经网络。

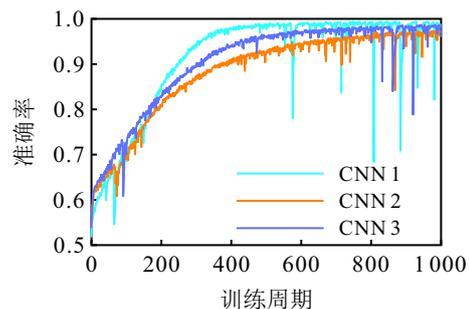
由表2~表4可见,无论采用哪种特征或超参数,在唤醒、效价和四分类维度上,SENet的识别性能均明显优于CNN。实验结果验证了SE块可提高神经网络模型的学习性能,通过SE块自适应地根据每个特征的贡献度加权,达到强化对识别作用大的特征,弱化对任务作用小或有负作用的特征,从而提高模型识别精度。

3.2 SENet的收敛速度分析

神经网络模型的训练包括前向传播、反向传播和权值更新过程。SENet可对不同的特征加权,因此具有更好的学习能力。为了验证SENet的高效性,记录了3组网络模型在3个维度上训练集的前1000个训练周期准确率的变化过程,如图6~图8所示。

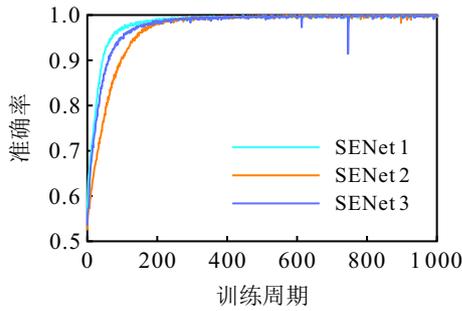


(a) SENet在唤醒维度上的训练过程

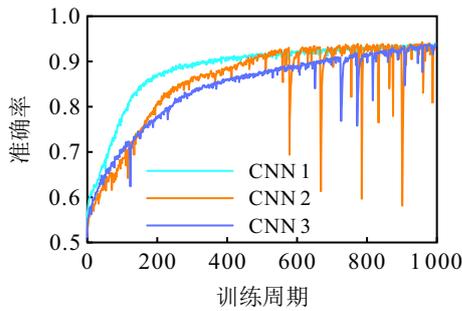


(b) CNN在唤醒维度上的训练过程

图6 唤醒维度上的训练过程

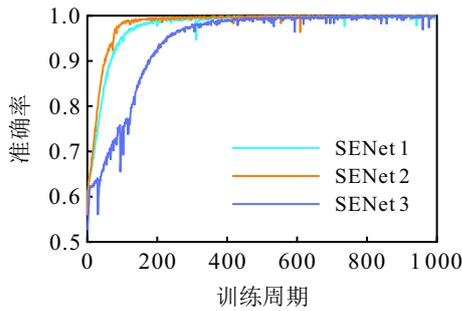


(a) SENet在效价维度上的训练过程

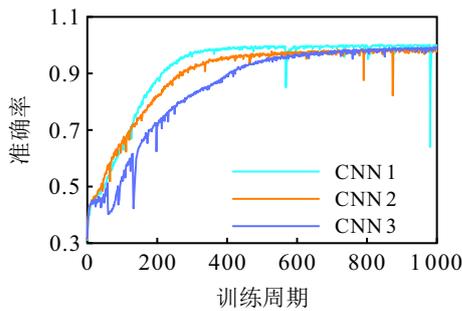


(b) CNN在效价维度上的训练过程

图7 效价维度上的训练过程



(a) SENet在四分类上的训练过程



(b) CNN在四分类上的训练过程

图8 四分类上的训练过程

由图6~图8可见,SENet 1、SENet 2和SENet 3在唤醒维度和效价维度上,当训练周期小于500时,训练集上的精度随着训练周期的增加而增加,当训练周期大于500时,精度不再增加.3个SENet模型在四分类维度上,当训练周期大于500时精度便不再增加.而3个CNN模型当训练周期小于1000时训练集上的准确率一直在增加,因此CNN完全拟合训练数据需要更多的训练周期,SENet仅仅略微增加了计算量而取得了优异的学习能力.这主要是由于不同特

征间会存在冗余信息,特征的融合造成了CNN模型出现过拟合,SENet的冲压和激励操作通过对不同的特征重新校准以提升网络的表示能力,从而获得了相对于CNN更为强大的学习能力.

3.3 SENet网络结构通用性分析

为了验证SENet框架的通用性,探索冲压激励块在不同深度和不同结构卷积网络中的适用性,即在其他网络结构中加入冲压激励块构成冲压激励网络相对于原始的卷积神经网络是否依然能够起到特征重校准的作用,另设计了3组不同结构的网络,用以探索SENet在不同网络结构中是否仍然具有很好的情感状态分类性能,最后与对应的CNN在识别精度上进行对比. G_4 和 G_5 的主干网络的结构参考VGGNet. G_5 中主干网络结构相比于 G_1 、 G_2 和 G_3 去除了跳跃连接结构,网络深度保持不变. G_4 网络深度为 G_5 的一半,包含1个block. G_6 的主干网络结构参考GoogleNet,为多分支结构.3组网络中的SENet仅比对应的CNN多了SE块,卷积操作中卷积核的大小全部设置为 3×3 ,其他超参数与之前保持一致.

使用唤醒、效价和四分类维度上的情感标记方案的情感识别结果如表5所示.在唤醒分类任务中,SENet 4的准确率最高为0.9052.在效价分类任务中,SENet 4的准确率最高为0.8972.在四分类任务中,SENet 6的识别准确率最高为0.8806. SENet 5在3个情感识别任务上的表现差于SENet 4与SENet 6,因此单纯增加网络深度并不能提高网络性能,反而会造成网络过拟合.与CNN相比,在唤醒维度上,3个SENet的识别准确率平均提高了约4.39%,在效价维度上平均提高了约5.55%.在四分类任务中,两种网络的差距更为明显,SENet的识别准确率相较于CNN平均提高了12.64%.表5中对比结果表明,由冲压激励块和卷积层构成的不同网络结构的SENet相较于对应结构的CNN均具有良好的分类性能,表明了SENet模型结构在脑电情感识别任务中具有较好的通用性.

表5 SENet与CNN在3个维度上的识别精度对比

分类维度	SENet			CNN		
	SENet 4	SENet 5	SENet 6	CNN 4	CNN 5	CNN 6
唤醒	0.9052	0.8787	0.8947	0.8537	0.8366	0.8566
效价	0.8972	0.8715	0.8956	0.8278	0.8201	0.8498
四分类	0.8795	0.8371	0.8806	0.7339	0.7060	0.7781

3.4 SENet的全局信息提取能力分析

为了验证SE block在学习全局信息的有效性,设计了专门的实验进行验证,对SENet的学习效果进

行可视化分析. 首先, 构造表1中SENet1模型, 并分别加载之前在唤醒和效价维度上训练的模型参数; 然后, 将一个值全为1的 $9 \times 9 \times 6$ 的三维张量输入进SENet1, 提取第1个SE块中 F_{ex} 的输出, 如表6所示.

表6 三维张量冲压激励运算结果

情感维度	特征					
	平均值	方差	标准差	第1差异值	第2差异值	模糊熵
唤醒	0.99	0.29	0.50	0.42	0.50	0.66
效价	1.00	0.12	0.12	0.88	0.74	0.56

将三维张量输入至SE块中. 首先, 进行全局平均池化运算, 此时变为维度为 $1 \times 1 \times 6$ 的张量; 然后, 在激励操作中输入至RELU和Sigmoid激活函数中得到表6中的结果. 在第3.1节和第3.3节的实验中平均值特征放置于三维特征矩阵的第1层, 索引为 $[:, :, 0]$, 因此输入值全为1的 $9 \times 9 \times 6$ 的三维张量, 输出的 $1 \times 1 \times 6$ 张量的第1个值与平均值特征对应. 显而易见, 在唤醒维度上平均值特征得到了强化, 方差特征相比于其他特征得到了弱化. 在效价维度上也是如此.

图9和图10为在唤醒和效价维度上单个特征的分类准确率. 由图9和图10可见, 对于表6中SE块强化的特征, 其识别准确率较高, 对于弱化的特征, 识别准确率较低, 这是由于其能够捕捉一些对于给定的任务很重要的特征, 通过这种自注意力机制, 网络可学会使用全局信息以选择性地强调信息性特征并抑制

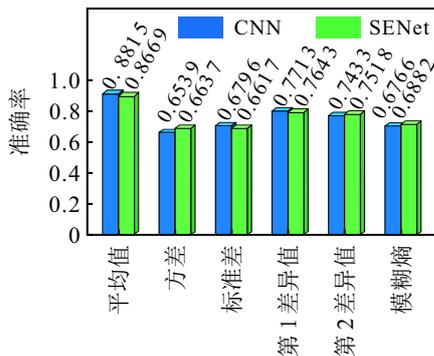


图9 唤醒维度上 SENet 和 CNN 在各特征上的准确率

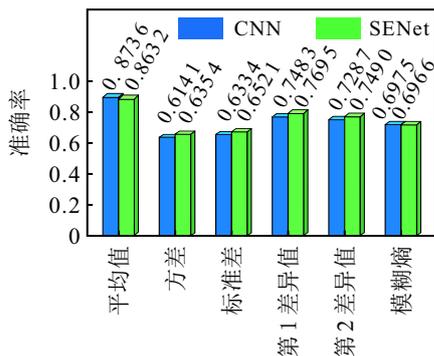


图10 效价维度上 SENet 和 CNN 在各特征上的准确率

不太有用的特征, 且SENet能够提取到不同特征间的互补性信息以提高性能.

3.5 时域特征的优势

为了表明时域特征的优势, 从原始脑电信号中提取频域特征进行分类. 这些频域特征是从32通道脑电信号5个频带中提取的功率谱密度特征, 原始脑电信号已经过 $4 \sim 45$ Hz的带通滤波处理, 因此5个频带分别设定为 $4 \sim 8$ Hz (theta)、 $8 \sim 12$ Hz (alpha)、 $12 \sim 30$ Hz (beta)、 $30 \sim 45$ Hz (gamma)和 $0 \sim 45$ Hz. 使用频域特征在唤醒、效价和四分类情感维度上的识别结果如表7~表9所示.

表7 SENet 1、CNN 1使用时域和频域特征的识别结果

情感维度	时域		频域
	SENet 1	SENet 1	CNN 1
唤醒	0.9056	0.7077	0.6264
效价	0.9039	0.6933	0.6293
四分类	0.8862	0.4889	0.4130

表8 SENet 2、CNN 2使用时域和频域特征的识别结果

情感维度	时域		频域
	SENet 2	SENet 2	CNN 2
唤醒	0.8993	0.6756	0.6195
效价	0.8867	0.6546	0.6385
四分类	0.8404	0.4656	0.4415

表9 SENet 3、CNN 3使用时域和频域特征的识别结果

情感维度	时域		频域
	SENet 3	SENet 3	CNN 3
唤醒	0.9049	0.6417	0.6204
效价	0.8904	0.6517	0.6374
四分类	0.8510	0.5023	0.4494

表7~表9中的识别结果清楚地表现出使用时域特征的优势, G_1 、 G_2 和 G_3 三组网络中SENet在时域特征上的识别精度高于频域特征, 且在频域特征上SENet的表现依然优于对应结构的CNN.

SENet在时域和频域特征上的识别结果验证了时域特征在刻画脑电信号中情感相关特征方面的优势. 同时在频域特征上, SENet依然能够保持自身优势, 在识别精度上高于相对应的CNN.

3.6 SENet与其他网络模型进行对比

将所提出方法在两种情感标注方案下的识别性能与现有的一些研究进行了比较. 对于所有对比的研究, 数据集(DEAP)与情感状态标记方案是相同的. 表10详细说明了对比研究中使用的特征和分类器. 以识别精度进行比较, 因其是最常用的.

表10 所提出方法与其他方法对比

方法	特征和分类器	准确率		
		唤醒	效价	四分类
Shahnaz等 ^[18]	离散小波系数,支持向量机	0.665 1	0.647 1	—
Zubair等 ^[19]	统计特征和基于小波变换提取的特征,支持向量机	—	—	0.497 0
Wichakam等 ^[20]	功率谱密度,支持向量机	0.650 0	0.649 0	—
Kolestra等 ^[16]	功率谱特征,朴素贝叶斯分类器	0.620 0	0.576 0	—
Bahari等 ^[21]	非线性动力学特性, K 近邻	0.645 6	0.580 5	—
Mohammadi等 ^[22]	小波变换, K 近邻	0.840 5	0.867 5	—
Hatamikia等 ^[23]	熵和分形维数,自组织映射网络	0.712 5	0.689 2	0.551 5
Xu等 ^[24]	功率谱特征, DBN	0.698 8	0.668 8	—
Chen等 ^[25]	功率谱密度,双向GRU+注意力机制	0.665 0	0.679 0	—
Kwon等 ^[26]	脑电的小波变换谱图和皮肤电的短时过0率, CNN	0.765 6	0.804 6	0.734 3
Mei等 ^[27]	皮尔逊相关系数矩阵, CNN	0.867 0	0.858 0	0.731 0
Xing等 ^[28]	功率谱密度,堆叠自编码器+LSTM	0.811 0	0.743 8	—

通过表10中的对比研究,在DEAP数据集上验证了经典分类算法和神经网络模型. SENet在使用时域特征进行分类时,与目前的一些方法相比在识别精度上提升许多. 由于时域特征和频域特征蕴含了对分类有效的信息不一致,在使用频域特征时表现较差. 对比结果表明,所提出方法在多通道脑电情感识别中表现优秀.

4 结论

本文提出了一种基于三维特征表示和冲压激励神经网络的多通道脑电情感识别方法. 所提出网络用于从三维特征矩阵中提取每个脑电通道的唯一信息,且网络能够对不同特征进行重校准,通过这种机制,网络可学习使用全局信息以选择性地强调信息性特征,并抑制不太有用的特征.

实验结果表明,根据提取到的时域特征构造三维特征矩阵表示多通道脑电信号的情感相关特征,所提出神经网络能够充分利用这些特征进行情感识别. 此外,将所提出方法的情感识别性能与现有的一些方法进行了比较,显示了其优越性,验证了该情感识别方法的可行性和有效性,所提出网络能够对不同特征间的相互依赖性进行建模以提高准确率.

参考文献(References)

- [1] Heller W. Neuropsychological mechanisms of individual differences in emotion, personality, and arousal[J]. *Neuropsychology*, 1993, 7(4): 476-489.
- [2] Kroupi E, Yazdani A, Ebrahimi T. EEG correlates of different emotional states elicited during watching music videos[C]. *International Conference on Affective Computing and Intelligent Interaction*. Springer, 2011: 457-466.

- [3] Petrantonakis P C, Hadjileontiadis L J. Emotion recognition from EEG using higher order crossings[J]. *IEEE Transactions on Information Technology in Biomedicine: A Publication of the IEEE Engineering in Medicine and Biology Society*, 2010, 14(2): 186-197.
- [4] Hjorth B. EEG analysis based on time domain properties[J]. *Electroencephalography and Clinical Neurophysiology*, 1970, 29(3): 306-310.
- [5] Frantzidis C A, Bratsas C, Papadelis C L, et al. Toward emotion aware computing: An integrated approach using multichannel neurophysiological recordings and affective visual stimuli[J]. *IEEE Transactions on Information Technology in Biomedicine: A Publication of the IEEE Engineering in Medicine and Biology Society*, 2010, 14(3): 589-597.
- [6] Lin Y P, Wang C H, Wu T L, et al. Support vector machine for EEG signal classification during listening to emotional music[C]. *IEEE the 10th Workshop on Multimedia Signal Processing*. Cairns, 2008: 127-130.
- [7] Davidson R J. Anterior electrophysiological asymmetries, emotion, and depression: Conceptual and methodological conundrums[J]. *Psychophysiology*, 1998, 35(5): 607-614.
- [8] Murugappan M, Rizon M, Nagarajan R, et al. EEG feature extraction for classifying emotions using FCM and FKM[J]. *International Journal of Computers and Communications*, 2007, 1(2): 21-25.
- [9] Hadjidimitriou S K, Hadjileontiadis L J. Toward an EEG-based recognition of music liking using time-frequency analysis[J]. *IEEE Transactions on Bio-Medical Engineering*, 2012, 59(12): 3498-3510.
- [10] Samara A, Menezes M L R, Galway L. Feature extraction for emotion recognition and modelling using neurophysiological data[C]. *The 15th*

- International Conference on Ubiquitous Computing and Communications and International Symposium on Cyberspace and Security. Granada, 2016: 138-144.
- [11] Mehmood R M, Lee H J. Emotion classification of EEG brain signal using SVM and KNN[C]. IEEE International Conference on Multimedia & Expo Workshops. Turin, 2015: 1-5.
- [12] Thammasan N, Moriyama K, Fukui K I, et al. Familiarity effects in EEG-based emotion recognition[J]. Brain Informatics, 2017, 4(1): 39-50.
- [13] Li X, Song D W, Zhang P, et al. Emotion recognition from multi-channel EEG data through convolutional recurrent neural network[C]. IEEE International Conference on Bioinformatics and Biomedicine. Shenzhen, 2016: 352-359.
- [14] 晁浩, 刘永利, 连卫芳. EEG情感识别中基于集成深度学习模型的多分析域特征融合[J]. 控制与决策, 2020, 35(7): 1674-1680.
(Chao H, Liu Y L, Lian W F. Multi-analysis domain feature fusion of EEG emotion recognition based on integrated deep learning model[J]. Control and Decision, 2020, 35(7): 1674-1680.)
- [15] Hu J, Shen L, Sun G. Squeeze-and-excitation networks[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, 2018: 7132-7141.
- [16] Koelstra S, Muhl C, Soleymani M, et al. DEAP: A database for emotion analysis; Using physiological signals[J]. IEEE Transactions on Affective Computing, 2012, 3(1): 18-31.
- [17] Chen W T, Zhuang J, Yu W X, et al. Measuring complexity using FuzzyEn, ApEn, and SampEn[J]. Medical Engineering & Physics, 2009, 31(1): 61-68.
- [18] Shahnaz C, Shoaib-Bin-Masud, Hasan S M S. Emotion recognition based on wavelet analysis of empirical mode decomposed EEG signals responsive to music videos[C]. IEEE Region 10 Conference. Singapore, 2016: 424-427.
- [19] Zubair M, Yoon C. EEG based classification of human emotions using discrete wavelet transform[M]. IT Convergence and Security. Singapore: Springer Singapore, 2017: 21-28.
- [20] Wichakam I, Vateekul P. An evaluation of feature extraction in EEG-based emotion prediction with support vector machines[C]. The 11th International Joint Conference on Computer Science and Software Engineering. Chon Buri, 2014: 106-110.
- [21] Bahari F, Janghorbani A. EEG-based emotion recognition using recurrence plot analysis and K nearest neighbor classifier[C]. The 20th Iranian Conference on Biomedical Engineering. Tehran, 2013: 228-233.
- [22] Mohammadi Z, Frounchi J, Amiri M. Wavelet-based emotion recognition system using EEG signal[J]. Neural Computing and Applications, 2017, 28(8): 1985-1990.
- [23] Hatamikia S, Nasrabadi A M. Recognition of emotional states induced by music videos based on nonlinear feature extraction and SOM classification[C]. The 21th Iranian Conference on Biomedical Engineering. Tehran, 2014: 333-337.
- [24] Xu H Y, Plataniotis K N. EEG-based affect states classification using deep belief networks[C]. Digital Media Industry & Academic Forum. Santorini, 2016: 148-153.
- [25] Chen J X, Jiang D M, Zhang Y N. A hierarchical bidirectional GRU model with attention for EEG-based emotion classification[J]. IEEE Access, 2019, 7: 118530-118540.
- [26] Kwon Y H, Shin S B, Kim S D. Electroencephalography based fusion two-dimensional (2D)-convolution neural networks model for emotion recognition system[J]. Sensors, 2018, 18(5): 1383.
- [27] Mei H, Xu X M. EEG-based emotion classification using convolutional neural network[C]. International Conference on Security, Pattern Analysis, and Cybernetics. Shenzhen, 2017: 130-135.
- [28] Xing X F, Li Z Q, Xu T Y, et al. SAE+LSTM: A new framework for emotion recognition from multi-channel EEG[J]. Frontiers in Neurobotics, 2019, 13: 37.

作者简介

晁浩(1981—), 男, 副教授, 博士, 从事模式识别、情感计算、语音识别等研究, E-mail: chaohao1981@163.com;

曹益鸣(1997—), 男, 硕士生, 从事情感计算、深度学习等研究, E-mail: caoyiming1997@163.com;

刘永利(1980—), 男, 教授, 博士, 从事信息检索、数据挖掘、大数据等研究, E-mail: yongli.buaa@gmail.com.