

基于双卷积神经网络融合的注意力训练研究

徐欣, 张佳欣, 张如浩

(南京邮电大学通信与信息工程学院, 南京 210003)

摘要: 学生的学习情况与其课堂注意力状态密切相关。为了探寻注意力训练能否提高课堂注意力, 对10名在校学生进行了 α 音乐训练, 并收集了训练前后的非注意和注意状态的脑电(Electroencephalogram, EEG)信号进行对比研究。由于EEG信号本质上是动态的, 且具有低信噪比和高冗余度的特性, 为避免直接通过神经网络识别EEG信号效果差的问题, 提取了信号的样本熵(Sample entropy, SampEn)、各个波段的能量和能量比共11个特征, 并将这些特征进行融合转化为多特征图像, 作为神经网络模型的输入。此外, 将AlexNet和VGG11两个网络模型进行加权融合构成双卷积神经网络, 进一步提高了图像分类性能。结果表明, 与单个模型相比, 双卷积神经网络融合模型的性能更佳, 其识别准确率最高可达到97.53%。研究发现, 经过 α 音乐训练, 受试者的脑电特征与此前相比有显著性差异, 且网络模型的分​​类准确率比训练前提高了4%, 说明本文所提的 α 音乐训练能够提高健康学生的注意力水平。

关键词: 脑电信号; 注意力; α 音乐训练; 双卷积神经网络融合; 多特征图像

中图分类号: TN911.72 **文献标志码:** A

Attention Training Based on Double Convolutional Neural Network Fusion

XU Xin, ZHANG Jiaxin, ZHANG Ruhao

(College of Communication and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China)

Abstract: Students' learning situation is closely related to their classroom attention state. In order to explore whether attention training can improve classroom attention, the electroencephalogram (EEG) signals of non-attention and attention states of ten students before and after α music training are collected and compared. It is worth noting that EEG signal is dynamic in nature and has the characteristics of low signal-to-noise ratio and high redundancy. In order to avoid the problem of poor recognition of EEG signals directly through neural network, 11 features of signal sample entropy (SampEn), energy and energy ratio of each band are extracted, and these features are fused into multi-feature images as the input of neural network model. In addition, the weighted fusion of AlexNet and VGG11 network models is used to form a double convolution neural network (CNN), which can further improve the performance of image classification. The results show that the performance of the fusion model with double CNN can achieve a better performance compared with the model with single CNN. In particular, the recognition accuracy of the proposed model can reach 97.53%. It can be found that after α music training, the EEG features of the

subjects are significantly different from those before, and the classification accuracy of the network model can be 4% higher than that before training. This observations show that the considered α music training can improve the attention level of healthy students.

Key words: electroencephalogram (EEG) signal; attention; α music training; double convolutional neural network fusion; multi-feature image

引言

注意力是指人的心理活动集中指向于某种事物的能力,良好的注意力是人脑发生一系列认知活动的基础。课堂注意力是指在课堂教学期间学生的注意力水平,是学生学习的基础。注意力分散会对学生造成多方面的影响,不仅影响学生的学习效率和学习成绩,还会导致学生出现不遵守学校纪律、厌学自卑以及人际关系差等问题。通过注意力训练的方法,可以提升注意力水平,改善课堂学习的质量。同时,鉴于青少年中精神疾病的发病率不断上升,帮助学生训练注意力不仅有利于提高学业成绩,而且有利于促进心理和情感健康。目前注意力训练常被应用于帮助存在注意力缺陷多动障碍(Attention deficit hyperactivity disorder, ADHD)的儿童增强和维持注意力。2019年,Shereena等^[1]通过设计脑电神经反馈训练任务,改善ADHD儿童的注意力。同年,Lim等^[2]在ADHD患者的持续性注意力方面,采用3D游戏实施干预和治疗。2021年,吴佳武等^[3]针对ADHD儿童患者,提出了一种脑电生物反馈治疗的游戏系统,该系统通过设立奖励机制,提高治疗的效果。然而,很多家长和学生认为,这些训练方法不适用于健康学生在课堂上出现的注意力不集中的情况,其有效性和安全性无法保障。要弥合这一科学研究与服务应用之间的差距,需要发展可扩展的、以证据为基础的培训项目,这样才能使得培训项目切实地融入学校环境。 α 音乐是一种灵感音乐,在60~70节拍之间,频率在8~14 Hz之间,与人的 α 脑电波一致。将大脑频率调到 α 波状态,可以发挥学生潜意识的作用、提升悟性和创造力等。因而,基于 α 音乐的注意力训练引起了广泛的关注。

为了量化注意力训练的效果,可以通过分析脑电(Electroencephalogram, EEG)信号实现。EEG主要测量和人脑皮层突触活动相关的神经生理变化,是一种非侵入式采集方法。目前,许多特征提取和分类方法已经用于识别注意力脑电任务。Klimesch等的研究^[4]表明注意力水平与 α 节律密切相关;Lutsyuk等^[5]通过计算 θ 和 β 波的能量比值得出大脑注意力水平状态;Rebsamen等^[6]以脑电信号频谱的加权频率来表示注意力水平高低;Ko等^[7]发现脑电频谱的变化与持续注意任务中行为表现的受损有密切的关系。由于EEG信号是一种随机性很强的非平稳性、非线性的微弱信号,因此线性分析方法不可能包含全部特征,而EEG信号的非线性复杂性被认为包含了有关大脑神经网络的结构信息^[8]。清华大学的一项研究表明,当被试者注意力集中时,其EEG信号中的近似熵(Approximate entropy, ApEn)参数会随之上升^[9]。路荣等^[10]计算小波包分解后各频段信号的近似熵,以此作为支持向量机(Support vector machines, SVM)的输入,分类正确率达到了72.5%。燕楠等^[11]提取注意力相关脑电信号的样本熵(Sample entropy, SampEn)特征,并采用SVM算法进行分类,正确率最高可达到85.5%。Ming等^[12]提取多尺度熵(Multi-scale Entropy, MSE)特征,其分类准确率为63.158%。徐鲁强等^[13]提出改进模糊熵算法来表征脑电信号的注意力水平,平均分类准确率达到76.3%。由此可见,对注意力脑电的研究主要通过分析频带能量或非线性动力学特性来实现,其中样本熵区分注意力状态最为有效。在注意力脑电分类中,常用的分类器是K近邻(K-nearest neighbors, KNN)和SVM,然而该分类器的识别准确率不佳。近年来,深度网络模型被大量应用于识别EEG信号,卷积神经网络(Convolutional neural network, CNN)是其中的典型代表。因此,本文提出了采用 α 音乐对健康的学生群体进行注意力训练,所设计方

案的流程图如图1所示。具体过程如下:首先采集训练前后的EEG信号,再提取 δ 波、 θ 波、 α 波和 β 波4个波段的能量、能量比和样本熵等特征,然后将其融合为多特征图像。为了取得更好的实验效果和控制复杂度,分类器采用AlexNet和VGG11两个卷积神经网络的融合模型。与单个卷积神经网络方案相比,本文设计的方案可以取得更好的结果。

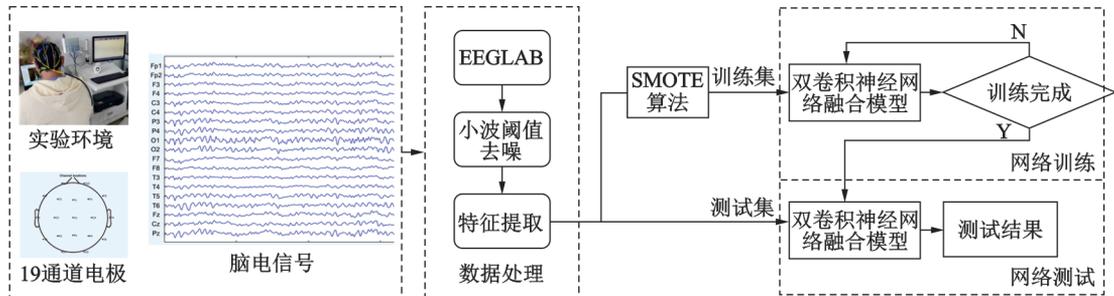


图1 本文所提方案

Fig.1 Illustration of the proposed scheme

1 注意力训练和脑电数据采集

本实验主要分为注意力训练和脑电信号采集两个部分。注意力训练方式为 α 音乐训练。脑电采集前首先进行注意网络测试(Attention network test, ANT)实验,之后采集静息态(非注意)和线上课程学习(注意)状态下的脑电数据。

1.1 实验前准备

(1) 实验对象

实验前,向志愿者简单介绍实验性质和流程,并获得其同意。实验共招募了10名志愿者,均为18~24岁的在校大学生,右利手,听力正常,视力正常或矫正正常,身体健康,无精神病和头部外伤史,且之前从未接受过注意力训练。

(2) 实验材料

①注意力训练过程中的 α 音乐,准备多个音乐歌单,志愿者可以根据喜好选择。

②注意力调查问卷,每周训练结束后志愿者填写注意力调查问卷,粗略评估训练效果。

③ANT实验设计,作为训练有效性的辅助支撑。实验程序由心理实验编程软件Eprime2.0完成编写,一次完整测验时长约为7 min。整个实验分为2个阶段,首先是练习阶段的1 min,帮助被试熟悉实验流程,接下来是6 min的正式实验试次,被试在测验过程中任务完成的反应时间及正确率均由软件自动记录。

④课程内容选取。为了更好地模拟课堂学习状态,注意脑电的采集在志愿者学习网课的状态下进行。本实验中选择了两节“中国大学MOOC”上的关于如何进行英文演讲的全英文课程,分别用于 α 音乐训练前和训练后。两节课的课时均为10 min左右,课程难度一致且难度不会太高或者太低而导致志愿者长期处于注意力不集中的状态。

1.2 实验流程

实验的总体流程如图2所示。实验中采用 α 音乐进行注意力训练,训练共持续3周,每周至少进行3次训练,每次至少保证15 min的听音乐时间。每周训练结束后,志愿者填写注意力调查问卷,粗略评

估训练效果,结合不同志愿者训练过程中的不同感受,及时调整听音乐时长和次数等训练细节,达到个性化训练。

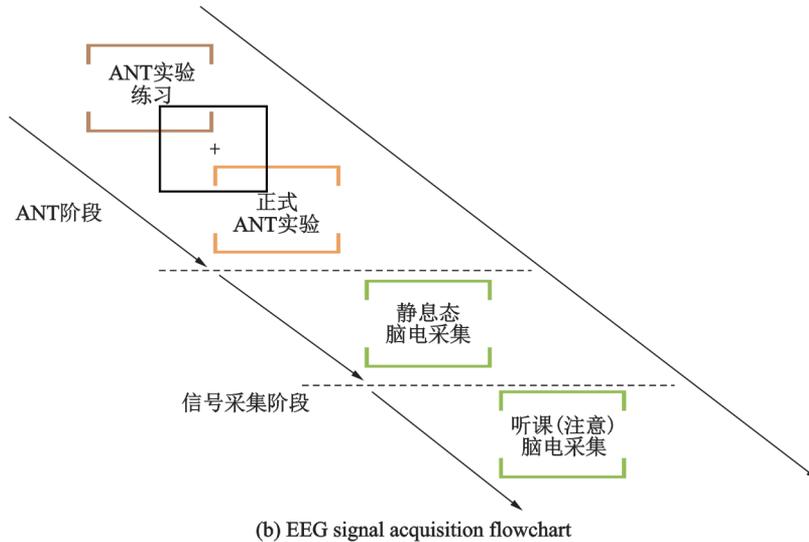
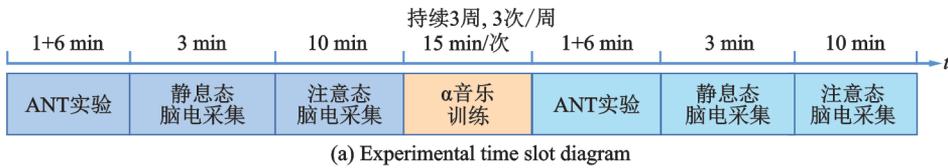


图2 实验流程图

Fig.2 Experimental flowchart

脑电信号的采集使用南京伟思医疗科技有限责任公司的大脑采集仪器实现,可设置脑电信号采集的速度和幅度,采样率为512 Hz。电极分布按照国际标准10~20系统放置,参考电极为A1、A2,选取额叶区电极F3、F4、F7、F8、Fz,颞叶区电极T3、T4、T5、T6,前额电极Fp1、Fp2,中央区电极C3、C4、Cz,顶叶区电极P3、P4、Pz,枕区电极O1、O2,共19个电极。

实验在安静环境内进行,实验室的温度控制在 $(23\pm 2)^{\circ}\text{C}$ 。数据采集之前,志愿者调整好坐姿和状态,进行ANT实验。首先呈现在屏幕中央的是“+”注视点,注视点消失后出现指示线索,被试准备好后按下“Enter”键,实验开始。在400 ms的时间间隔(“+”注视点)后,目标刺激(水平位置排列的5个方向随机的箭头)出现,受试者需要在700 ms内判断中间箭头的指向(箭头指向左点击鼠标左键;指向右点击鼠标右键),待被试做出反应后目标刺激消失,屏幕上会呈现反应正误、反应时间和平均准确率;700 ms以上无反应则目标刺激自动消失,此试次作为无效试次。每个试次结束后,再次呈现中心“+”注视点。整个实验被试全程注视屏幕,避免眼球的大幅度运动。

脑电信号采集阶段分为两个部分:3 min静息态脑电(非注意状态)和10 min线上课程学习脑电(注意力状态)。非注意状态下,被试闭目,保持头脑清醒,尽可能减少肢体动作,采集3 min的脑电。之后志愿者有1 min的时间稍作休息调整,在确保其状态良好的情况下,开始网课视频的学习,期间采集注意状态脑电信号。课程播放结束后,进行与课程内容相关的测试,询问被试对于学习过程中注意力变化的主观感受并记录。ANT实验和脑电信号采集在 α 音乐训练前后都要进行一次,用作前后对比。所有志愿者均完成了实验,没有中途离开者。

2 脑电信号处理

2.1 预处理

脑电信号微弱,且极易受外界干扰^[14]。为保证研究的有效性,根据“少坏段”和“少伪迹”的准则,从10名志愿者中选取了8名志愿者(S1~S8)的脑电数据用于后续处理分析。首先去除明显坏段,为了保留数据的有效频率,去除高频噪声,使用EEGLAB工具包去除50 Hz的工频干扰,并进行0.1~30 Hz的滤波。此后将信号降采样至256 Hz,进行独立成分分析(Independent component analysis, ICA),去除眼电、肌电、心电等伪迹成分。经过EEGLAB的系列处理后,对每组数据以4 s的滑动窗口进行截取,重叠率为50%,将信号分段。最终,训练前后都分别得到717个非注意状态样本,1 883个注意力状态样本。

小波滤波可以在最大限度保持真实信号的基础上,减少甚至完全滤除噪声。因此,对每个样本都进行小波阈值去噪,去除叠加在信号上的噪声。小波阈值去噪主要包含3方面:小波基选取、阈值选取和阈值函数选取。db小波分解能调整频域分辨率使之与脑电节律相一致,更适合用于处理脑电数据,本文采用db4小波去除脑电信号中的噪声干扰。阈值 λ 的选取会直接影响去噪效果,本文采用固定阈值,其计算公式为 $\lambda = \sqrt{2\lg N}$,其中 N 为信号的长度。最常用的阈值函数有硬阈值函数和软阈值函数两种,软阈值去噪后会丢失一些特征,误差相对较大,本文对信号进行硬阈值处理。其中一个样本经过小波阈值去噪的前后对比如图3所示,可以看出去噪后的重构信号具有良好的逼近性。

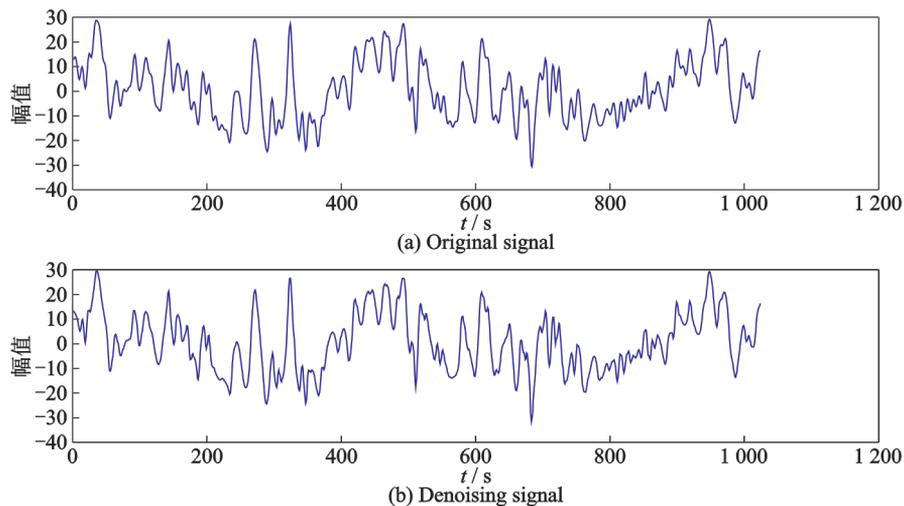


图3 小波阈值去噪

Fig.3 Wavelet threshold denoising

2.2 SMOTE算法

SMOTE(Synthetic minority oversampling technique)算法是对随机过采样算法的改进。由于随机过采样使用复制样本的方法来增加少数类,很可能造成模型过拟合的情况。而SMOTE算法通过分析少数类样本合成新样本,它以每个少数类样本点 x 的 k 个最近邻样本点为依据,根据样本不平衡比例确定采样率,从其 k 个最近邻样本点中随机选择 N 个邻近点。对于每个随机选出的邻近点 x_n 构建新样本 x_{new} ,公式为

$$x_{new} = x + \text{rand}(0, 1) \times (x - x_n) \quad (1)$$

在原样本的基础上,将邻近点与原样本进行差值并乘上一个 $[0, 1]$ 范围的阈值,从而达到合成数据

的目的。该算法的核心是:特征空间上的邻近点其特征都是相似的。它并不是在数据空间上进行采样,而是在特征空间中进行采样,所以准确率会高于传统的采样方式。

由于非注意状态和注意状态2种类别的EEG样本数量严重不平衡,故本文使用SMOTE算法来解决类别不均衡的问题。首先,将数据集根据9:1的比例分为训练集和测试集,即71个非注意样本和188个注意样本划分为测试集,其余为训练集。然后将训练集中的非注意状态样本经过SMOTE处理,合成新的样本,生成的新训练集包含1695个非注意状态样本和1695个注意状态样本。

2.3 特征提取

由于EEG信号的随机性、不稳定性以及信号采集过程中产生的各种噪声干扰无法全部去除,使得EEG具有非常高的空间特性且冗余度高,直接将预处理后的脑电信号样本送入神经网络的效果比较差,导致网络模型理解其含义比较困难。脑电特征提取的目的是为了选取出具有代表性的特征数据,以更全面有效地表征脑电信号。本文在相关工作中手工提取信号特征作为深度神经网络模型的输入,进行注意力脑电识别。

2.3.1 频带能量

小波变换只分解信号的低频部分,忽略了高频部分。对于EEG信号来说,小波包变换^[15]不仅可以分解信号的低频部分,也分解高频部分,没有冗余和疏漏,可以很好地进行时频分析,保证了信号的不丢失。

将小波包树节点记为 (j, n) , j 表示小波包树的层数, n 表示对应层数的节点个数, h_{l-2k} 和 g_{l-2k} 为一对正交镜像滤波器,可以将信号划分到任意频段。第 k 点的小波包分解系数为

$$d_j^{2n}(k) = \sum_{l \in Z} h_{l-2k} d_{j-1}^n(l) \quad (2)$$

$$d_j^{2n+1}(k) = \sum_{l \in Z} g_{l-2k} d_{j-1}^n(l) \quad (3)$$

经过小波包分解,信号采样率为 f_s ,则第 j 层各个节点对应的频段分别为:

$$\left[0, \frac{f_s}{2^{j+1}}\right]; \left[\frac{f_s}{2^{j+1}}, \frac{2f_s}{2^{j+1}}\right]; \left[\frac{2f_s}{2^{j+1}}, \frac{3f_s}{2^{j+1}}\right]; \dots; \left[\frac{(2j-1)f_s}{2^{j+1}}, \frac{f_s}{2}\right]。选取合适的节点进行脑电信号的重构,$$

第 $j-1$ 层、第 k 点的重构公式为

$$d_{j-1}^n(k) = \sum_{l \in Z} h_{k-2l} d_j^{2n}(l) + \sum_{l \in Z} g_{k-2l} d_j^{2n+1}(l) \quad (4)$$

根据提取信号的频域特性,选择相应的频带对应的小波分解系数,利用式(4)进行重构,即可得到特定的频段信号。

本文使用db4小波对预处理之后的信号样本进行5层分解,分别提取了每个样本的 α 、 θ 、 β 和 δ 节律并重构 α 波、 θ 波、 β 波和 δ 波。分解后,各个节点及其对应的频带范围如图4所示。

由于小波包将能量不重叠地正交分解到相邻频带上,因此遵循能量守恒定律,有

$$E_j = \sum_{i=1}^N \|n_i\|^2 \quad (5)$$

式中: j 表示相应频带; N 表示原数据点数; n_i 表示小波包分解后的系数。根据式(5),容易求得 α 、 θ 、 β 和 δ 波段对应的频带能量,并由此计算将各个频段两两之间的能量比值,每个样本的19个通道的数据分别得到10个特征值。

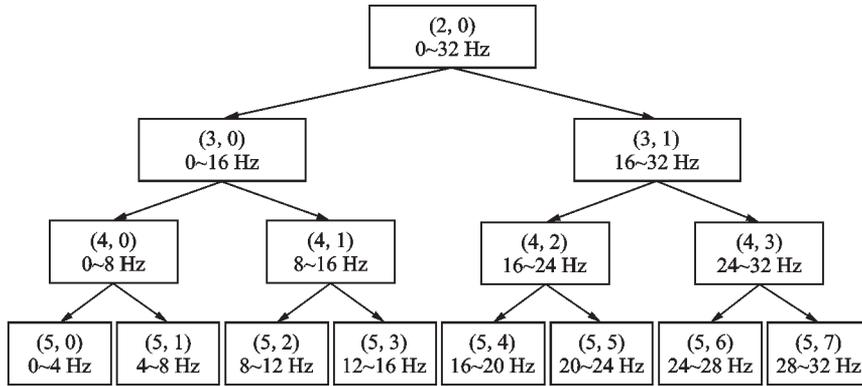


图4 小波包分解对应的频率空间

Fig.4 Frequency space decomposed by wavelet packets

2.3.2 样本熵(SampEn)

研究表明,脑电信号有很强的非线性结构^[16],SampEn^[17]是脑电信号常用的非线性特征之一。SampEn是由ApEn改进得到的,时间序列的复杂性和维数改变时序列产生新模式的概率越大,序列的复杂度越高,熵值越大。

假定由 N 个数据组成的时间序列 $x(n) = x(1), x(2), \dots, x(N)$, 其 SampEn 的计算方式如下:

(1)按序号组成 1 组维数为 m 的向量序列, $X_m(1), \dots, X_m(N - m + 1)$, 其中

$$X_m(i) = [x(i), x(i + 1), \dots, x(i + m - 1)] \quad 1 \leq i \leq N - m + 1 \tag{6}$$

(2)定义 $X_m(i)$ 与 $X_m(j)$ 之间的距离为 $d[X_m(i), X_m(j)]$, 则

$$d[X_m(i), X_m(j)] = \max_{k=0, \dots, m-1} (|x(i+k) - x(j+k)|) \quad i \neq j \tag{7}$$

(3)给定 1 个阈值 $r (r > 0)$, 统计 $X_m(i)$ 与 $X_m(j)$ 之间距离小于等于 r 的数量, 记作 B_i 。对于 $1 \leq i \leq$

$N - m$, 定义 $B_i^m(r) = \frac{1}{N - m - 1} B_i$, 求其对所有 i 的平均值为

$$B^m(r) = \frac{1}{N - m} \sum_{i=1}^{N-m} B_i^m(r) \tag{8}$$

(4)增加维数到 $m + 1$, 统计 $X_{m+1}(i)$ 与 $X_{m+1}(j)$ 之间距离小于等于 r 的数量, 记作 A_i 。对于 $1 \leq i \leq$

$N - m$, 定义 $A_i^m(r) = \frac{1}{N - m - 1} A_i$, 求其对所有 i 的平均值为

$$A^m(r) = \frac{1}{N - m} \sum_{i=1}^{N-m} A_i^m(r) \tag{9}$$

(5) $B_m(r)$ 和 $A_m(r)$ 是两个序列在相似容限下分别匹配 m 个点和 $m + 1$ 个点的概率, 将 SampEn 定义为

$$\text{SampEn}(m, r) = \lim_{N \rightarrow \infty} \left\{ - \ln \left[\frac{A^m(r)}{B^m(r)} \right] \right\} \tag{10}$$

(6)当 N 为有限值时, 可将长度为 N 的序列的 SampEn 估计为

$$\text{SampEn}(m, r, N) = -\ln \left[\frac{A^m(r)}{B^m(r)} \right] \quad (11)$$

本文选取参数 $m=2, r=0.1\sigma$, 计算各个样本中各个通道的 SampEn, 其中 σ 为相应样本的标准差。本研究中有用电极共 19 个, 将每个样本中的各个电极按顺序进行排列。将每个电极的数据经过小波包分解重构得到的 10 个频带能量与能量比特特征 (4 个频带能量特征: α 频带能量、 θ 频带能量、 β 频带能量和 δ 频带能量, 6 个能量比: $E_\alpha/E_\theta, E_\alpha/E_\beta, E_\alpha/E_\delta, E_\theta/E_\beta, E_\theta/E_\delta, E_\beta/E_\delta$), 和计算所得的 SampEn 共 11 个特征结合在一起, 组成 19×11 的融合特征矩阵, 并对矩阵进行归一化处理, 然后将归一化后的矩阵转化为对应的多特征彩色图像。样本图像分辨率为 543 像素 \times 429 像素, 作为双卷积神经网络融合模型的输入, 进行二次特征提取。

3 双卷积神经网络融合模型

与传统的分类方法相比, CNN 作为深度学习的主要算法之一, 能够有效地保留图片的特征。典型的 CNN 模型主要由 3 个部分组成: 卷积层、池化层和全连接层。卷积层一般设置多个不同的卷积核对输入进行操作从而获得不一样的特性, 达到提取局部特征的效果; 池化层用来大幅降低参数量级; 全连接层负责预测分类, 输出结果。由于本文中数据集相对较小, 且多模型融合训练相对于单一网络模型训练耗时比较长, 因此选用 CNN 模型中结构较为简单的 AlexNet 网络和 VGG 网络进行融合, 以进一步提高识别效果。

3.1 AlexNet 网络

AlexNet 的整个网络结构由 5 个卷积层和 3 个全连接层组成, 深度总共 8 层, 层与层之间使用最大池化 (Max-pooling) 分开。该模型主要有以下特点: (1) 使用 ReLU 函数进行激活操作, 确保输出特征的值在合理范围之内, 防止梯度消失, 也能解决一定的梯度爆炸问题, 从而加快训练速度; (2) AlexNet 中采用的是最大池化, 从而保留最显著的特征, 同时 AlexNet 中步长小于池化核窗口, 使池化有重叠部分; (3) 全连接层采用 Dropout 方法, 避免过拟合。

本文在 AlexNet 网络的基础上, 介于卷积层 (Conv) 和激活层 (ReLU) 之间加入了 BN (BatchNorm) 层, 批正则化, 让数据分布均匀, 且卷积层不使用偏置 bias。全连接层之间的 Dropout 层, 失活比例设置为 0.5, 随机失活一些神经元, 减少神经网络过拟合。为了加快训练速度, 在保证识别率不受较大影响的情况下, 将每一层的卷积核个数减半。AlexNet 网络结构简图如图 5 所示。

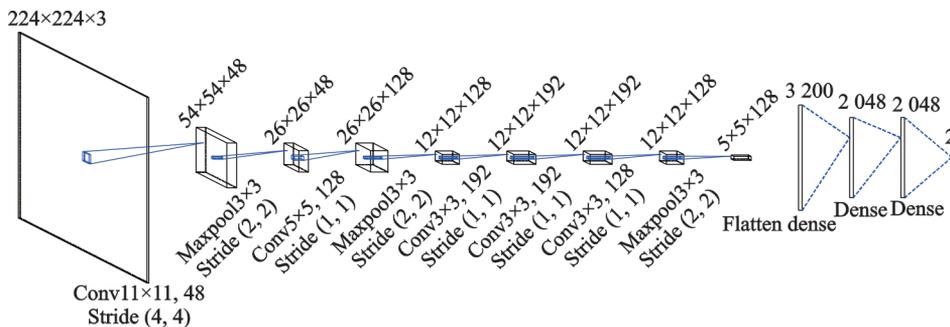


图 5 AlexNet 网络结构

Fig.5 AlexNet network structure

4 结果分析

4.1 实验过程数据

4.1.1 ANT 实验数据

将E-prime软件后台记录的ANT实验过程中的反应正误和反应时间进行汇总,计算每个受试者训练前后的平均反应时间和反应正确率,将结果用柱状图表示,由图8可以清晰地看出:大多数受试者经过 α 音乐训练后,ANT实验的平均反应时间变短,反应正确率提高。只有受试者S5训练后反应时间反而变长,正确率稍有下降,这可能与受试者进行ANT实验时的情绪状态有关。对训练前后的反应时间和正确率分别进行配对样本t检验,得到 $p < 0.05$,说明训练前后的平均反应时间和正确率都有显著差异,证实了 α 音乐训练能够有效提高学习者的注意力。

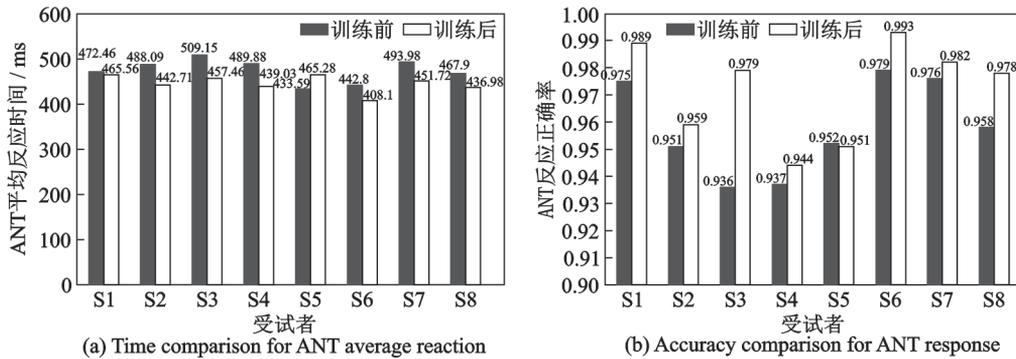


图8 ANT结果分析

Fig.8 ANT result analysis

4.1.2 课程内容测试结果

本研究中每位志愿者结束网课学习后,都要进行与课程内容相关的小测试,根据测试的得分判断其学习效果和注意力水平。测试共设置了5个问题,其中包括2个判断题(2×10 分),1个单选题(10分),1个多选题(答案为4选3,答对1个10分,共30分)和2个简答题(每个简答题包括4个小点,每个点记5分,共 2×20 分),题目难度逐渐增加。训练前后的两次测试中问题的难易程度相当,因此可以根据2次测试的得分情况检测训练效果。结果如表1所示,可以看出经过训练,8名志愿者的测试得分都有一定程度的提高,从一定程度上反应出受试者听课的注意力水平有所提升。

4.2 特征分析

研究采用非参数检验——相关样本威尔科克森符号秩检验来判断注意力训练前后各特征的非差异性,从而判断 α 音乐训练对学生注意力的影响。将训练前注意态-训练后注意态各特征的渐进显

表1 课程内容测试结果

Table 1 Test results of course content

受试者	判断	单选	多选	简答	总分
S1	20	10	20	25	75
	20	10	20	30	80
S2	20	0	20	25	65
	20	10	30	30	90
S3	10	10	30	25	75
	20	10	30	40	100
S4	20	10	10	30	70
	20	10	20	35	85
S5	20	10	20	20	70
	20	10	30	30	90
S6	10	10	30	30	80
	20	10	30	35	95
S7	20	10	20	25	75
	20	10	30	30	90
S8	20	10	30	15	75
	20	10	20	30	80

著性(双侧检验)记为 a-diff,结果如表 2 所示:除 E_δ/E_α 、 E_δ/E_θ 这两个能量比特征的渐进显著性大于 0.05,其余特征的渐进显著性均小于 0.05,表明经过注意力训练后,大部分特征产生了显著变化。

表 2 各特征差异性分析

Table 2 Analysis of differences between each feature

特征	E_α	E_β	E_δ	E_θ	E_α/E_β	E_δ/E_α	E_δ/E_β	E_δ/E_θ	E_θ/E_α	E_θ/E_β	样本熵
a-diff	<0.001	<0.001	<0.001	0.014	0.000	0.291	0.041	0.157	<0.001	<0.001	0.000

此外,研究选取样本熵和 E_θ/E_β 能量比这两个常用于注意力识别的特征,分别进行各个通道的平均数值比较。由图 9 可知,经过注意力训练,大部分通道数据的样本熵有了提高, E_θ/E_β 能量比下降,代表注意力有了一定水平的提高。

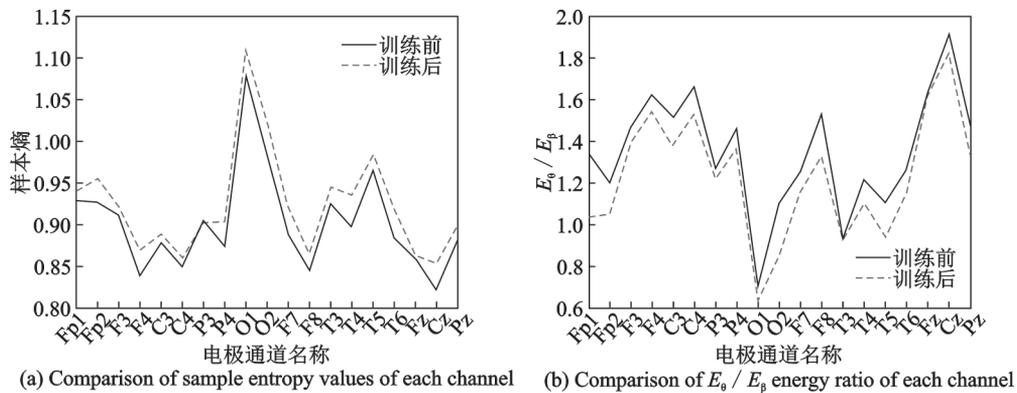


图 9 样本熵和 E_θ/E_β 数值的对比

Fig.9 Comparison of sample entropy and E_θ/E_β values

4.3 模型效果

将此前划分好的经过扩充的训练集图像(以注意力训练前的数据集为例)按 224 分辨率进行随机裁剪和翻转等处理后,输入到设定好的网络中用于模型训练,学习率为 0.000 02。不同的模型融合权重对模型训练效果有很大的影响,设 AlexNet 网络的权重为 x ,VGG11 网络的权重为 y ,模型训练收敛后,不同 x 和 y 取值下的训练平均准确率见表 3。因此,将 AlexNet 和 VGG11 的融合权重分别设为 0.4 和 0.6,迭代 100 次,网络准确率得以收敛。为了验证双卷积神经网络融合模型的有效性,将本文所提出的模型与单一模型进行对比,训练准确率如图 10 所示,可以看出本文所提出的多模型融合网络的识别率明显高于单一网络识别率。

在测试集上评估训练好的模型效果,得出: α 音乐训练前,单一的 AlexNet 网络模型测试集上的识别准确率为 83.36%,VGG11 的识别准确率为 91.06%,本文提出的双卷积神经网络融合模型识别率为 93.49%。 α 音乐训练后,AlexNet 网络模型的识别准确率为 86.18%,VGG11 的识别准确率为 92.62%,本文提出的多模型融合网络识别

表 3 不同融合权重对比

Table 3 Comparison of different fusion weights

x	y	准确率/%
0.2	0.8	88.76
0.3	0.7	82.14
0.4	0.6	94.17
0.5	0.5	91.89
0.6	0.4	91.52
0.7	0.3	90.18
0.8	0.2	81.70

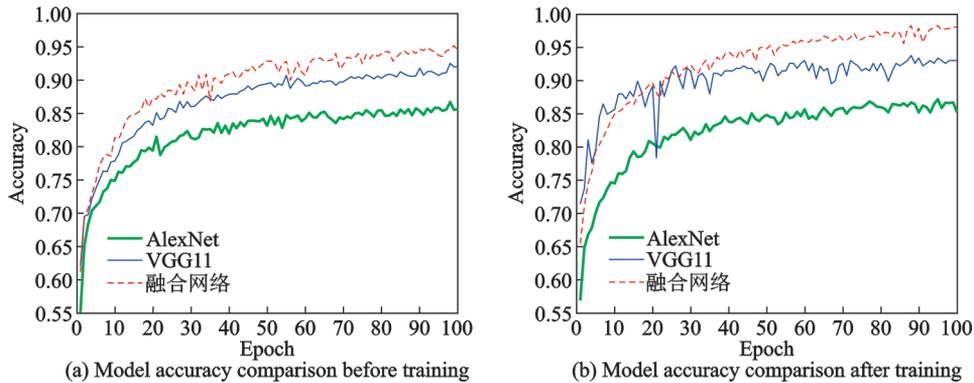


图10 模型训练结果

Fig.10 Model training results

率为97.53%。经过 α 音乐训练,双卷积神经网络融合模型的分类准确率提高了4%,结合4.2中各特征差异性对比结果,说明经过训练后,注意态和非注意态的脑电差别变大,含有更丰富的识别信息,证明这种注意力训练方法的有效性。

4.4 模型性能对比

二分类中,可以将样本根据真实类别和模型预测类别分为4种:真正类(True positive, TP),真实样本类别和模型预测类别均为正;假正类(False positive, FP),真实样本类别为负,模型预测类别为正;真负类(True negative, TN),真实样本类别和模型预测类别均为负;假负类(False negative, FN),真实样本类别为正,模型预测类别为负。为了能够清晰地表述模型性能,本文采用模型的精确率(Precision),召回率(Recall), F_1 分数(F_1 -score)和准确率(Accuracy)作为模型性能评价指标。

在模型预测是正类数的所有结果中,模型预测对的比重为精确率,即

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (12)$$

真实值是正类数的所有结果中,模型预测对的比重为召回率,即

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (13)$$

F_1 -score是精确率和召回率的调和平均数,与模型输出性能优劣成正比,即

$$F_1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (14)$$

分类模型所有预测正确的结果占总数的比重为准确率,即

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (15)$$

将注意力状态记为正类,非注意状态记为负类。为了证明模型性能的提升,本文的分类模型与常见的传统分类方法进行对比,各项指标如表4所示。相对于朴素贝叶斯方法和SVM,双卷积神经网络融合模型除了在精确率方面低于SVM,在召回率和 F_1 -score方面均有不同程度的提高,整个模型系统的准确率比性能较好的SVM提高了4%~7%。可以看出,本文所提出的双卷积神经网络融合模型具有显著的优势,可以更好地用于注意力脑电的识别。

表4 模型性能对比

Table 4 Comparison of classification methods

%

模型	Precision		Recall		F ₁ -score		Accuracy	
	训练前	训练后	训练前	训练后	训练前	训练后	训练前	训练后
朴素贝叶斯	68.60	77.80	88.70	87.50	77.40	82.37	79.90	83.40
SVM	1.00	1.00	81.56	83.42	89.80	90.96	88.70	90.10
本文方法	89.80	95.40	96.50	99.20	93.03	97.30	93.49	97.53

5 结束语

本研究针对正常学生上课过程中出现的注意力分散的现象,提出了 α 音乐的注意力训练方法,并采集了注意力训练前后的EEG信号。针对EEG信号直接作为神经网络的输入进行端到端的识别效果不理想,本研究提取了 δ 、 θ 、 α 和 β 波4个波段的能量、能量比和样本熵等特征,将其融合为多特征图像,作为神经网络模型的输入。同时,本文将AlexNet和VGG11两个卷积神经网络进行融合,从而达到对分类效果优化的目的。结果表明,双卷积神经网络融合模型能够提高识别准确率,优于单个模型的性能,也优于传统分类器的性能,可以用于注意力脑电分类。同时,本文通过对比注意力训练前后提取的各特征差异性、注意力脑电的分类效果,以及每位志愿者ANT实验的反应平均时间和反应正确率,发现 α 音乐训练方法确实能够提高健康学生的课堂注意力水平。同时,听音乐的形式具有较高的安全性和可行性,为注意力提升方案的制定提供了依据和思路。

参考文献:

- [1] SHEREENA E A, GUPTA R K, BENNETT C N, et al. EEG neurofeedback training in children with attention deficit/hyperactivity disorder: A cognitive and behavioral outcome study[J]. *Clinical EEG and Neuroscience*, 2019, 50(4): 242-255.
- [2] LIM C G, POH X W W, FUNG S S D, et al. A randomized controlled trial of a brain-computer interface based attention training program for ADHD[J]. *PLoS One*, 2019, 14(5): e0216225.
- [3] 吴佳武, 刘金明, 许泽举, 等. 针对ADHD患者脑电生物反馈治疗的游戏系统设计[J]. *新型工业化*, 2021, 11(4): 56-60, 63. WU Jiawu, LIU Jinming, XU Zeju, et al. Design of game system for EEG biofeedback therapy of ADHD patients[J]. *The Journal of New Industrialization*, 2021, 11(4): 56-60, 63.
- [4] KLIMESCH W, DOPPELMAYR M, RUSSEGER H, et al. Induced alpha band power changes in the human EEG and attention[J]. *Neuroscience Letters*, 1998, 244(2): 73-76.
- [5] LUTSYUK N V, ISMONT E V, PAVLENKO V B. Modulation of attention in healthy children using a course of EEG-feedback sessions[J]. *Neurophysiology*, 2006, 38(5/6): 389-395.
- [6] REBSAMEN B, BURDET E, GUAN C, et al. A brain-controlled wheelchair based on P300 and path guidance[C]// *Proceedings of the First IEEE/RAS-EMBS International Conference on Biomedical Robotics and Biomechanics*. Washington, DC: IEEE Computer Society, 2006: 1101-1106.
- [7] KO L W, KOMAROV O, HAIRSTON W D, et al. Sustained attention in real classroom settings: An EEG study[J]. *Front Hum Neurosci*, 2017, 11: 388.
- [8] BOSL W, TIERNEY A, TAGER-FLUSBERG H A. EEG complexity as a biomarker for autism spectrum disorder risk[J]. *BMC Med*, 2011, 9(1): 18.
- [9] HONG Bo, TANG Qingyu, YANG Fusheng, et al. ApEn and Cross-ApEn: Property, fast algorithm and preliminary application to the study of EEG and cognition[J]. *Signal Processing*, 1999, 15(2): 100-108.
- [10] 路荣, 黄力宇, 晋琅. 小波包分解脑电复杂性特征提取的注意状态实时识别[J]. *医疗卫生装备*, 2013, 34(2): 1-5. LU Rong, HUANG Liyu, JIN Lang. Real-time attention state recognition based on EEG complexity feature extraction using wavelet packet decomposition[J]. *Chinese Medical Equipment Journal*, 2013, 34(2): 1-5.

- [11] 燕楠, 王珏, 魏娜, 等. 基于样本熵的注意力相关脑电特征信息提取与分类[J]. 西安交通大学学报, 2007, 10(41): 1237-1241.
YAN Nan, WANG Jue, WEI Na, et al. Feature exaction and classification of attention related electroencephalographic signals based on sample entropy[J]. Journal of Xi'an Jiaotong University, 2007, 10(41): 1237-1241.
- [12] MING D, ZHANG M, XI Youyuan, et al. Multiscale entropy analysis of attention ralated EEG based on motor imaginary potential[C]//Proceedings of International Conference on Computational Intelligence for Measurement Systems and Applications. Hong Kong, China: [s.n.], 2009: 24-27.
- [13] 徐鲁强, 刘静霞, 肖光灿, 等. 脑电注意水平的特征识别[J]. 计算机应用, 2012, 32(11): 3268-3270.
XU Luqiang, LIU Jingxia, XIAO Guangcan, et al. Characterization and classification of EEG attention level[J]. Journal of Computer Applications, 2012, 32(11): 3268-3270.
- [14] 吴金玉, 张文强, 潘玲佼, 等. 一种结合自适应噪声完备经验模态分解和盲反卷积去除脑电中眼电伪迹的新方法[J]. 数据采集与处理, 2020, 35(4): 720-729.
WU Quanyu, ZHANG Wenqiang, PAN Lingjiao, et al. A new method for electroencephalogram artifact automatic removal based on CEEMDAN and BD in EEG signals[J]. Journal of Data Acquisition and Processing, 2020, 35(4): 720-729.
- [15] 谷学静, 位占锋, 刘海望, 等. 基于小波包和串并行 CNN 的脑电信号分类[J]. 微电子学与计算机, 2021, 38(6): 60-65.
GU Xuejing, WEI Zhanfeng, LIU Haiwang, et al. EEG signal recognition based on wavelet packet and serial parallel CNN[J]. Microelectronics & Computer, 2021, 38(6): 60-65.
- [16] WANG Fei, WU Shichao, ZHANG Weiwei, et al. Multiple nonlinear features fusion based driving fatigue detection[J]. Biomedical Signal Processing and Control, 2020, 62: 102075.
- [17] JIE Xiang, CAO Rui, LI Li. Emotion recognition based on the sample entropy of EEG[J]. Bio-medical Materials and Engineering, 2014, 24(1): 1185-1192.

作者简介:



徐欣(1978-),男,副教授,研究方向:信号与信息处理、人工智能、生物电信号分析与处理及其应用, E-mail:xuxin@njupt.edu.cn。



张佳欣(1998-),通信作者,女,硕士研究生,研究方向:生物电信号分析与处理, E-mail: zhangjiaxinjy@163.com。



张如浩(1998-),男,硕士研究生,研究方向:生物电信号分析与处理。

(编辑:张黄群)