

基于频空融合与3D-CNN-Attention的抑郁症识别

王建尚, 张冰涛, 王小敏, 严大川
兰州交通大学电子与信息工程学院, 甘肃 兰州 730070

【摘要】提出了一种基于频谱信息的三维特征构建方法,根据电极位置将每个通道的功率值排列成二维特征向量。将不同频段特征排列成三维积分特征张量,提取频域信息,同时,为了减少容积导体效应影响,利用功能连接将时序脑电(EEG)数据映射到空间脑功能网络,提取空间信息。通过对特征与目标类之间关系的分析,提出一种3D-CNN-Attention网络模型,在3D-CNN网络中加入Attention机制,以增强EEG特征学习能力。在公开数据集上的系列对比实验,结果表明基于3D-CNN-Attention网络框架的抑郁症检测性能优于其他方法,获得了最高为96.32%的准确率。本文方法能够为抑郁症检测提供一种有效的解决方案。

【关键词】抑郁症; EEG; 频谱; 脑功能网络; 3D-CNN-Attention

【中图分类号】R318; TP399

【文献标志码】A

【文章编号】1005-202X(2024)10-1307-08

Depression recognition based on frequency-space domain fusion and 3D-CNN-Attention

WANG Jianshang, ZHANG Bingtao, WANG Xiaomin, YAN Dachuan

School of Electronic and Information Engineering, Lanzhou Jiaotong University, Lanzhou 730070, China

Abstract: A three-dimensional feature construction method based on spectral information is presented, in which the power values of each channel are arranged into two-dimensional feature vectors based on electrode positions. The different frequency band features are arranged into a three-dimensional integral feature tensor to extract the information in frequency domain. Meanwhile, in order to reduce the influence of volume conductor effect, functional connectivity is utilized to map the temporal electroencephalogram data to the spatial brain functional network for extracting the spatial information. By analyzing the relationship between features and target classes, a 3D-CNN-Attention network model is proposed to incorporate an Attention mechanism in 3D-CNN network to enhance the electroencephalogram feature learning capability. A series of comparative experiments on publicly available datasets show that 3D-CNN-Attention network framework outperforms other methods in depression detection, obtaining an accuracy rate of up to 96.32%. The proposed method provides an effective solution for depression detection.

Keywords: depressive disorder; electroencephalogram; frequency spectrum; brain functional network; 3D-CNN-Attention

前言

抑郁症是一种典型的精神障碍,全世界约有2.8亿抑郁症患者^[1]。其精神方面的临床症状表现为内心悲伤、空虚、对生活失去兴趣、丧失自尊;而在身体方面则表现为睡眠中断、运动迟缓、疲劳、食欲明显下降、自残等。因此,寻找客观有效的抑郁症检测方法具有重要意义。现有疗法在一定程度上能够缓解

抑郁症,但由于医疗资源和经济发展水平差异,在经济欠发达国家,超过75%的抑郁症患者无法得到治疗^[2]。传统基于量表访谈诊断的方法耗时费力,且严重依赖于医生的主观经验^[3-4]。

随着生物医学工程、计算机科学、神经科学等跨学科前沿技术的发展,基于生理信号、行为、语音、面部表情等信息的计算机辅助抑郁症检测研究成果不断涌现。生理数据具有不可伪造性和不可掩饰性等特点,脑电(Electroencephalogram, EEG)是一种记录大脑神经活动电信号,常被用于精神疾病的诊断。与功能性核磁共振、弥散张量成像等生理数据相比,EEG具有非侵入性、时间分辨率高、以及成本低等特点,被广泛应用于精神疾病的诊断^[5]。因此,寻找一种能够有效筛查和诊断抑郁症的机器识别方法具有

【收稿日期】2024-05-10

【基金项目】国家自然科学基金(62362047, 61962034)

【作者简介】王建尚, 硕士研究生, 研究方向: 脑功能网络, E-mail: 1663465993@qq.com

【通信作者】张冰涛, 博士, 研究方向: 人工智能与医学信号处理, E-mail: zhangbingtao321@163.com

重要的应用价值和现实意义。近年来,EEG与人工智能结合在抑郁症诊断方面取得许多成果。Mumtaz等^[6]使用3种不同的时频分解方法,包括小波变换、短时傅里叶变换和经验模态分解,从EEG信号中提取时频特征,使用所有提取特征组合的逻辑回归分类器分类抑郁症和健康对照受试者,分类准确率为91.6%。Zhdanov等^[7]从加拿大第一个抑郁症生物标志物整合网络(CAN-BIND-1)研究的122名抑郁症患者的治疗前EEG信号中提取功率谱和时空复杂性特征,作为支持向量机(SVM)分类器的输入,并在预测抗抑郁药反应方面获得了79.2%的准确率。Jaworska等^[8]使用随机森林(RF)分类器将脑电图谱与多个人口统计学特征一起用于预测抑郁症治疗结果,并获得88%的准确率。Li等^[9]基于CNN相继完成了利用EEG数据对轻度抑郁症的检测。Sharma等^[10]创建了一个混合CNN-LSTM网络,用于从21名抑郁症患者和24名正常受试者获得的EEG信号中进行抑郁症检测。

上述研究方法从不同角度解决了抑郁症检测问

题,且各自具有优势,但很少有研究考虑频域信息和空间信息互补。为此,本文构建了基于多频谱图融合空间脑功能网络(Brain Functional Network, BFN)特征的3D-CNN-Attention网络框架,通过将不同频段的特征排列成三维积分特征张量,时序EEG数据映射到BFN,实现频域信息和空间信息互补。设计基于双流3D注意力模块的深度网络模型,该模型结合了CNN和注意力模块的优势,利用两个注意力模块分别专注于EEG通道和空间特征学习,以增强EEG特征学习能力。该框架考虑了频域和空间信息的互补性以及机器学习算法的优势,进而实现抑郁症的有效检测。

1 方法

1.1 多频谱图构建

本文中,EEG数据被划分为8 s的滑动窗口,重叠2 s,此滑动窗口长度在相关精神障碍研究中被广泛应用^[11]。接着将EEG电极映射至5×5的二维向量,其余位置用0填充,此映射过程如图1所示。

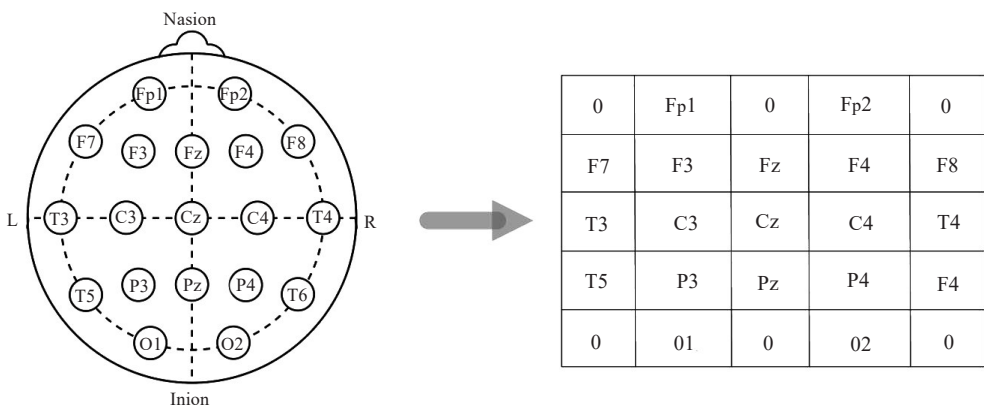


图1 EEG电极映射过程
Figure 1 EEG electrode mapping process

在二维向量上,基于Welch方法计算 δ (0.5~4 Hz)、 θ (4~7 Hz)、 α (7~12 Hz)、 β (12~30 Hz)等4个频段的功率谱密度(Power Spectral Density, PSD),进而使用PSD构建频谱图,此过程如图2所示,具体构建如下。

Welch通过将EEG信号分段、利用窗函数计算各频段PSD,各频段PSD计算过程如下:

$$P_i(f) = \frac{1}{N_w} \left| \sum_{n=0}^{N_w-1} x_i(n) e^{\frac{j2\pi fn}{N_w}} \right|^2 \tag{1}$$

其中, i 代表第*i*个频段, N_w 是EEG信号分段数量, n 是每段长度, $x_i(n)$ 是第*i*段中的第*n*个样本值, j 是虚数单位。为了提高频谱图的分辨率,径向基函数

(Radial Basis Function, RBF)插值法被用于生成各频段频谱图^[12],计算过程如下:

$$\hat{M}_i(p) = \sum_{j=1}^n \lambda_j \varphi((p - p^j)) \tag{2}$$

其中, n 是EEG电极数量, λ_j 是权重系数, φ 是RBF,计算过程如下:

$$\varphi(r) = \exp\left(-\frac{r^2}{2\sigma^2}\right) \tag{3}$$

1.2 BFN构建

本文中EEG电极被定义为网络节点,节点间同步性被定义为边,为了探索抑郁症识别的最佳同步性计算方法,使用锁相值(Phase Locked Value, PLV)、

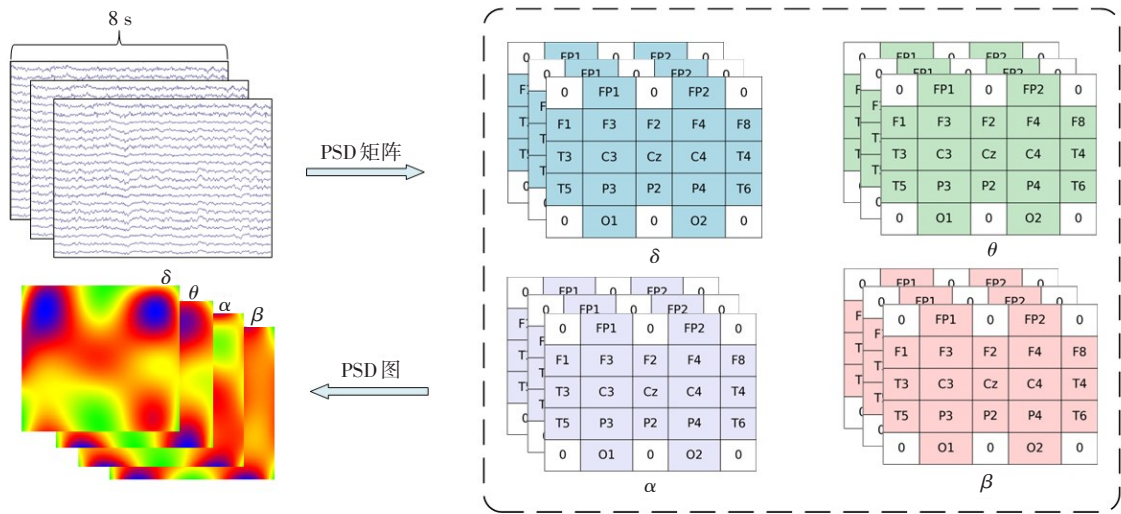


图2 频谱图构建过程

Figure 2 Frequency spectral construction process

皮尔逊相关系数 (Pearson Correlation Coefficient, PCC)、相位滞后系数 (Phase Lag Index, PLI) 3 种同步性计算方法。

PLV 是一种衡量两个信号在相位上同步性的指标, 计算过程如下:

$$PLV = \left| \frac{1}{N} \sum_{t=1}^N e^{i(\phi_1(t) - \phi_2(t))} \right| \tag{4}$$

其中, N 是观测的时间点数, $\phi_1(t)$ 和 $\phi_2(t)$ 分别是两个电极在时间点 t 的瞬时点位。

PCC 反映了不同 EEG 信号之间线性相关性, 计算过程如下:

$$PCC = \frac{\text{cov}(w, z)}{\sigma_w \sigma_z} \tag{5}$$

其中, w 和 z 分别表示两个不同通道中的 EEG 信号, $\text{cov}(w, z)$ 为 w 和 z 的协方差, σ_w 、 σ_z 分别为 w 和 z 的标准差。

PLI 通过计算两个信号之间相位差以反映信号间同步性, 计算方法如下:

$$PLI = \frac{1}{M} \left| \sum_{n=1}^M \text{sign}(\Delta \phi_n) \right| \tag{6}$$

其中, M 是 EEG 信号采样点个数, sign 是符号函数, $\Delta \phi_n$ 是采样点之间的相位差。基于上述 3 种同步性度量方法的 BFN 构建过程如图 3 所示。

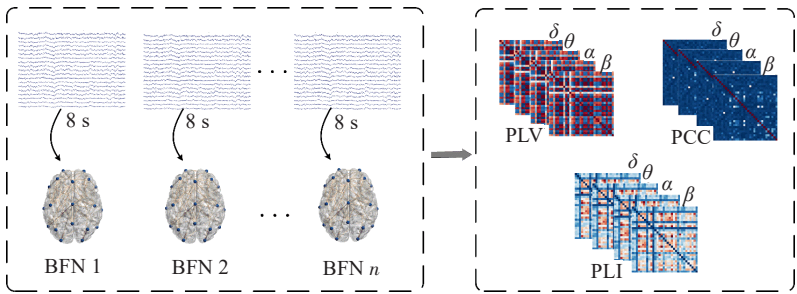


图3 BFN邻接矩阵的构建

Figure 3 Construction of BFN adjacency matrix

1.3 3D-CNN-Attention网络模型

频谱图和 BFN 邻接矩阵作为 3D-CNN 输入数据, 每个数据流通过 3D 注意力模块出来。第 1 个模块为 3D 卷积层, 其由 3D 批量归一化层、ReLU 激活函数和 3D 最大池化层组成。第 2~5 个模块由 4 个注意力模块组成, 每个模块包含两个 3D 卷积层和一个

3D 最大池化层。同时, 在第 2~5 个模块中加入通道注意力 (Channelwise Attention, CA) 模块和空间注意力 (Spatial Attention, SA) 模块。对于每个流, 采用全连接 (Fully Connected, FC) 层来学习样本属于各类别的概率, softmax 层融合两个流的概率获得最终识别结果。3D-CNN-Attention 网络模型如图 4 所示, 每个

模块的内核大小为 $3 \times 3 \times 3$,第一个模块的步幅为 $1 \times 2 \times 2$,其余两个模块的步幅为 $2 \times 2 \times 2$ 。5个模块的滤波器

数量分别为32、32、64、128和256。

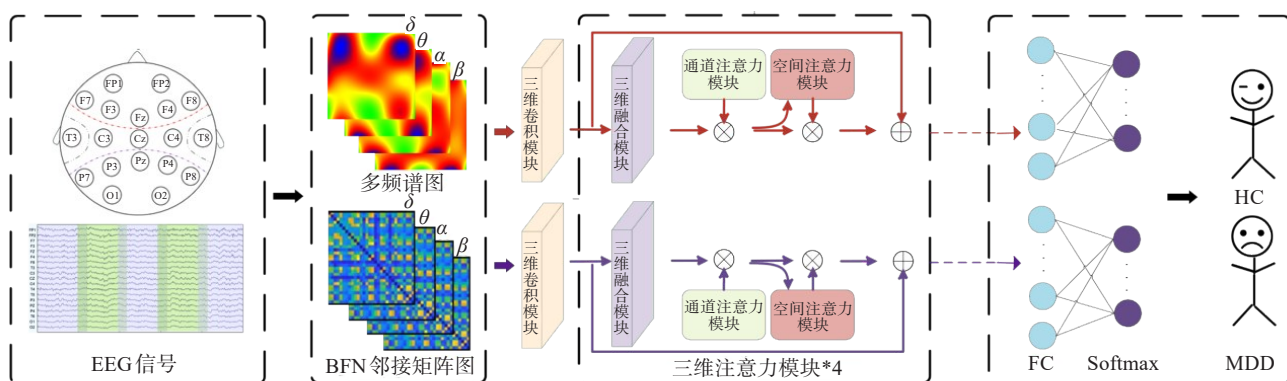


图4 基于3D-CNN-Attention的抑郁症识别模型

Figure 4 Depression recognition model based on 3D-CNN-Attention

1.3.1 3D-CNN网络结构 3D卷积能够有效地表征频域特征或空间特征的多通道信息,位置 (x, y, z) 处第 m 层第 n 特征图 S_{mn}^{xyz} 可以表示为:

$$S_{mn}^{xyz} = \tanh \left(b_{mn} + \sum_c \sum_{p=0}^{P_m-1} \sum_{q=0}^{Q_m-1} \sum_{r=0}^{R_m-1} w_{mnc}^{pqr} S_{(m-1)c}^{(x+p)(y+q)(z+r)} \right) \quad (7)$$

其中, b_{mn} 是第 n 个特征图的偏差, w_{mnc}^{pqr} 是连接第 $m-1$ 层第 c 个特征图和第 n 层第 m 个特征图 (p, q, r) 位置处的卷积核权重。 P_m, Q_m, R_m 分别是内核的高度、宽度和深度。3D批量标准化和ReLU激活函数计算过程如下所示:

$$y = \frac{s - \mu}{\sqrt{\sigma^2 + \varepsilon}} \cdot \gamma + \beta \quad (8)$$

$$\text{ReLU}(S) = \phi(S) = \max(0, S) \quad (9)$$

其中, μ, σ^2 为输入的均值和方差, ε 为避免分母为零的校正项, γ 和 β 为可学习变量,以增强模型灵活性。

池化层对于判别性特征提取和降低过度拟合风险非常有效^[13],本文使用尺寸为 $2 \times 2 \times 2$ 的3D最大池化。假设 $S_{m,n}$ 和 $S_{m-1,n}$ 分别是第 m 层和第 $m-1$ 层中的第 n 个特征,则3D最大池化为:

$$S_{m,n}^{xyz} = \max(C_{m,n}^{xyz}) \quad (10)$$

$$C_{m,n}^{xyz} = \left\{ S_{m-1,n}^{x'y'z'} \left| \begin{array}{l} x' \in \{2x-1, 2x\} \\ y' \in \{2y-1, 2y\} \\ z' \in \{2z-1, 2z\} \end{array} \right. \right\} \quad (11)$$

其中, $C_{m,n}^{xyz}$ 包括 $2 \times 2 \times 2$ 滑动窗口中的所有元素,第 m 层返回最大值,第 $m-1$ 层通过3D最大池化进行下采样以压缩特征。

FC层连接在模型的末尾,用于提取样本对每个类别的概率,FC层形式如下:

$$y = W^T x + b \quad (12)$$

其中, x 和 y 是输入和输出, W 和 b 是权重矩阵和偏差。对于3D-CNN中的每个流,应用softmax激活函数来导出FC层中的样本概率,过程如下所示:

$$P(y = i|x) = \frac{\exp(x^T w_i)}{\sum_{k=1}^K \exp(x^T w_k)} \quad (13)$$

其中, x 是FC层采样输出, k 是类别标签, $\sum_{k=1}^K \exp(x^T w_k)$ 是范围 $(0, 1)$ 的归一化项。本文中,选择交叉熵作为损失函数:

$$\text{Loss} = -\frac{1}{n} \sum_{i=1}^n \sum_{j=1}^m p(x_{ij}) \log(q(x_{ij})) \quad (14)$$

其中, n 是各批次的样本数, $p(x_{ij})$ 和 $q(x_{ij})$ 分别是真实结果和预测结果。

1.3.2 3D注意力模块 注意力机制具有深层次特征学习的优势,挤压和激励(SE)作为注意力模块,被广泛应用于二维CNN^[14]。为了提取深层次特征,本文将2D的SE模块扩展到3D-CNN。设计3D注意力模块分别包含CA模块和SA模块,如图5所示。

CA模块可以从不同的特征中捕获通道级EEG信息。在SE模块中保留全局平均池化(Global Average Pooling, GAP)的基础上,CA模块中还增加了全局最大池化(Global Max Pooling, GMP)。GAP和GMP的结合可以增强注意力模块的学习能力。假设特征图 $x^{l-1} \in \mathbb{R}^{C_{l-1} \times D_{l-1} \times H_{l-1} \times W_{l-1}}$ 为第 $l-1$ 层的输出, $C_{l-1}, D_{l-1}, H_{l-1}, W_{l-1}$ 分别表示通道尺寸、EEG通道/频段尺寸、特征高度和宽度。通过GAP和GMP的压缩操作,第 l 层第 c 个特征图的特征图可以表示为 $a_c^l \in \mathbb{R}^{C \times 1 \times 1 \times 1}$ 和 $m_c^l \in \mathbb{R}^{C \times 1 \times 1 \times 1}$,压缩过程如下所示:

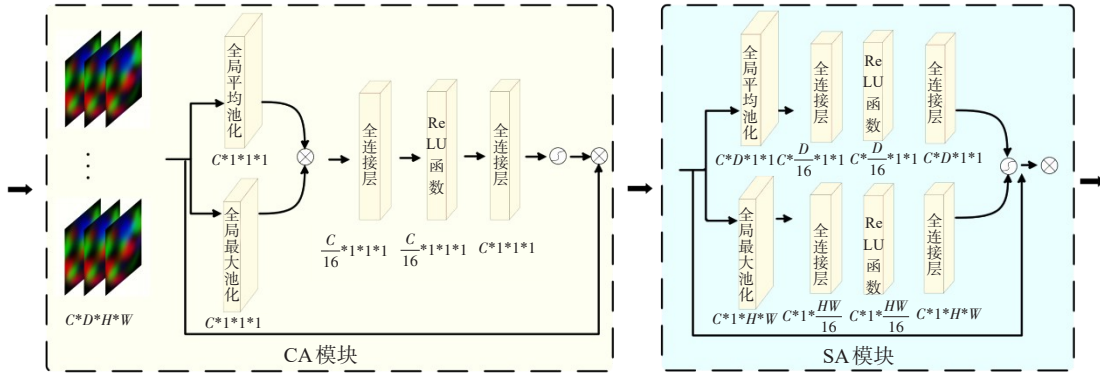


图5 3D注意力模块

Figure 5 3D attention module

$$\text{GAP: } a_c^l = F_{sq}(x_c^{l-1}) = \frac{1}{DHW} \sum_{d=1}^D \sum_{h=1}^H \sum_{w=1}^W x_c^{l-1}(d, h, w) \quad (15)$$

$$\text{GMP: } m_c^l = F_{sq}(x_c^{l-1}) = \max\{x_c^{l-1}(d, h, w)\} \quad (16)$$

然后按元素求和应用为:

$$s_c^l = a_c^l \oplus m_c^l \quad (17)$$

之后, 两个FC层用于激励操作:

$$\hat{s} = F_{ex}(s^l, W) = \sigma(W_2 \delta(W_1 s^l)) \quad (18)$$

其中 $s^l, \hat{s} \in R^{C \times 1 \times 1 \times 1}$ 是输入和输出, $W_1 \in RC \times (C/r)$ 和 $W_2 \in RC \times (C/r)$ 是两个FC层的权重, δ 和 σ 是ReLU和Sigmoid激活函数。减速比 γ 设定为16。

尺度运算将激励得到的权重向量 \hat{s} 通过乘法分配给每个通道的先前特征 x^{l-1} , 校准通过以下方式实现:

$$x' = F_{scale}(x^{l-1}, \hat{s}) = \hat{s} \otimes x^{l-1} \quad (19)$$

\hat{s} 中的每个元素反映了相应通道的重要性, 用于自适应地放大或抑制输入权重。

SA模块利用频域和空域进行抑郁症分析。研究两个空间维度, 一个为深度信息 S_{depth} (如频域多个EEG通道中的信息), 另一个为平面维度 S_{planar} (如空间特征的不同电极区域的激活程度), 在SA模块中, GAP用于聚合CA模块图的空间信息, 如下所示:

$$S_{depth} = d\text{GAP}(x') = \frac{1}{HW} \sum_{h=1}^H \sum_{w=1}^W x'(h, w) \quad (20)$$

$$S_{planar} = p\text{GAP}(x') = \frac{1}{D} \sum_{d=1}^D x'(d) \quad (21)$$

其中, $d\text{GAP}(x'), p\text{GAP}(x')$ 分别表示深度和平面维度上的3D平均池化, 分别为 $S_{depth} \in R^{C \times D \times 1 \times 1}$, $S_{planar} \in R^{C \times 1 \times H \times W}$, 执行与式(12)相同的激励, 输出分别为深度和平面维度的 $\hat{S}_{depth} \in R^{C \times D \times 1 \times 1}$, $\hat{S}_{planar} \in R^{C \times 1 \times H \times W}$, 最终的校准由以下方式完成:

$$S = \hat{S}_{depth} \otimes x', \hat{S} = \hat{S}_{planar} \otimes S \quad (22)$$

在初始块中, 内核专注于全局信息提取, 并且通常提取输入图像的特征。而在深度块中, 内核主要提取局部信息。每个注意力块可以自适应地学习通道级别和空间级别的权重值, 并且提取的特征在表示和抽象级别上更加复杂。

2 实验与结果

2.1 实验数据

实验数据来自Mumtaz等^[15]发布的公开数据集, 受试者是马来西亚理科大学医院(Hospital Universiti Sains Malaysia, HUSM)的门诊病人。该数据集来自34名抑郁症患者[女性17名, 男性17名, 年龄(40.3+12.9)岁]和30名年龄匹配的健康受试者[女性9名, 男性21名, 年龄(38.3+15.6)岁]。根据精神障碍诊断与统计手册第五版(The diagnostic and statistical manual of mental disorders IV, DSM-IV), 抑郁症患者符合诊断标准^[16]。所有参与者都签署了参与同意书, 并被告知实验数据采集所采用的实验程序, 并得到了HUSM伦理委员会的批准。数据包含闭眼和睁眼条件下抑郁症患者和健康受试者的EEG信号。实验采用19通道电极分布。传感器根据10-20电极放置标准和连耳参考放置。EEG信号分别从额叶(Fp1、Fp2、F3、F4、F7、F8、Fz), 颞叶(T3、T4、T5、T6)、顶叶(P3、P4、Pz), 枕叶(O1、O2)和中央(C3、C4、Cz)区域记录。

2.2 数据预处理

记录的19通道EEG信号被降采样到256 Hz, 并在0.5~70 Hz进行带通滤波以消除噪声。为了从混合信号中提取纯EEG信息, 利用独立成分分析(Independent Component Analysis, ICA)去除不感兴趣的眼球运动、眨眼、肌肉活动信号。在本文中, EEG数据划分为8 s的滑动窗口(段), 重叠2 s, 每次EEG

数据的记录时间约为 6 min,每个被试每个通道上的数据分为 50 段。

2.3 单特征模型与多特征模型比较

为获得具有统计学意义的实验结果,本方法在公开数据集上执行十折交叉验证。准确率、灵敏度和特异性用于检测性能评估。

设计对比实验,对比 PLV、PCC 和 PLI 3 种通道相关性以及多频谱图、多特征融合模型作为网络输入的检测性能,计算方法的实验结果显示,PLV 方法在抑郁症识别中表现最佳,准确率达到 91.24%,显著优于 PLI 的 88.28% 和 PCC 的 87.86%。单独多频谱图特征的准确率为 93.83%。多特征融合模型准确率最高,为 95.67%。结果如表 1 所示。

表 1 特征输入性能对比(%)
Table 1 Performance comparison of different feature inputs (%)

方法	准确率	灵敏度	特异性
PLV	91.24	92.46	90.13
PCC	87.86	89.04	88.15
PLI	88.28	87.92	86.72
多频谱图	93.83	88.30	91.41
多特征融合模型	95.67	94.48	96.11

为了进一步研究,使用 t 分布随机邻域嵌入(t -distributed Stochastic Neighbor Embedding, t -SNE)来了解视觉特征。特征分类如图 6 所示。结果显示,本文的特征融合模型比其他方法具有更好的检测性能。

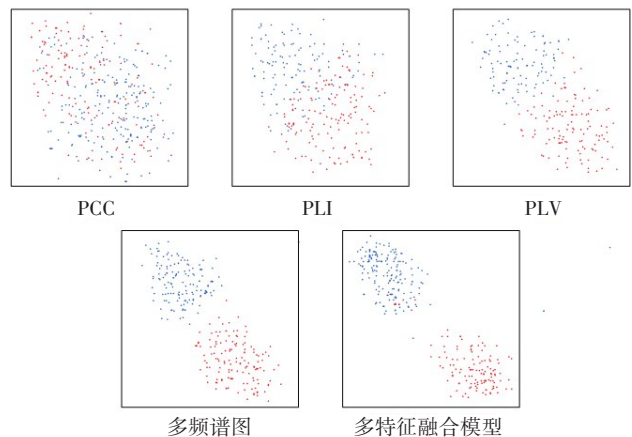


图 6 特征分类结果
Figure 6 Feature classification results

对比 PCC、PLV 和 PLI 3 种通道相关性计算方法的实验结果,揭示了它们在基于 EEG 数据的抑郁症识别中的表现差异。PLV 在捕捉 EEG 信号间复杂相

位关系方面具有独特优势,进一步突显了动态相位同步在抑郁症特征中的重要性。与之相对比,PLI 虽然也考虑了信号间的相位同步,但其仅关注相位差的存在而不考虑其大小,因此在准确性上略逊于 PLV。另一方面,PCC 在实验中的表现稍逊一筹,其 87.86% 的准确率可能是由于部分 EEG 信号特征的丢失而导致的。这种结果表明,传统的线性相关性分析方法在处理复杂的脑电信号时可能存在一定局限性。

对比多频谱图、通道间同步性载体和多特征融合模型实验结果表明了与通道间同步性载体相比,多频谱图特征包含了更多的抑郁症检测信息。由于特征融合可以提供信息互补,提高决策过程的准确性,因此多特征融合模型能够达到最高的检测精度。这进一步证明了在抑郁症诊断中,综合使用多种数据源和特征类型可以显著提升模型的性能和可靠性。通过将不同的信息维度相结合,可以捕获更全面的脑网络活动模式和潜在的病理特征,从而为精确诊断和个性化治疗提供更强有力的支持。

2.4 不同频段检测性能的比较

研究指出,不同频段的 EEG 信号在不同的精神状态或精神疾病下呈现出不同的变化趋势^[17-18]。为了确定哪个频段信息与抑郁症更相关,通过屏蔽除特定频段外的所有频段并评估特定频段的检测性能来实现这一目标。

这种操作被称为闭塞敏感性,已被成功用于理解神经网络如何进行睡眠分期和图像分类^[19]。本文利用不同频段的 EEG 信号特征,设计了基于 3D-CNN-Attention 的多特征融合框架下抑郁症检测性能评价的对照实验。准确性、灵敏度和特异性的比较结果如图 7 所示。从以上结果可以看出,低频段的检测性能要高于高频段。与其他频段相比, θ 频段的检测准确率最高为 96.32%。

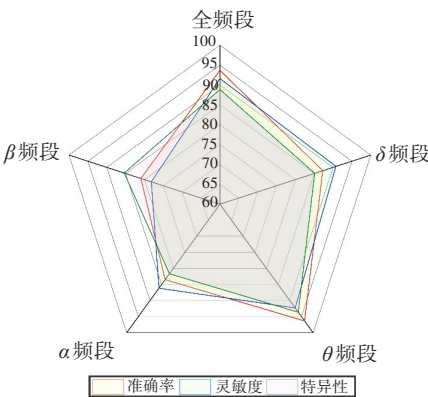


图 7 不同频段检测性能的比较
Figure 7 Comparison of detection performance in different frequency bands

同时,笔者还采用配对样本 t 检验来检验 θ 频段与其他频段的检测准确率是否存在显著差异,结果如表2所示。

表2 θ 频段与其他频段检测性能差异分析(P 值)
Table 2 Analysis of differences in detection performance between θ wavebands and other wavebands (P value)

方法	vs. δ 频段	vs. α 频段	vs. β 频段	vs.全频段
十折交叉验证	0.020 6	0.051 7	0.005 9	0.059 1
留一交叉验证	0.018 4	0.052 8	0.004 6	0.061 3

θ 频段与 δ 、 β 频段之间存在显著差异($P<0.05$),而与 α 频段和全频段之间差异不显著,说明与全频段相比, θ 频段和 α 频段可以提高检测性能,有效减少计算量。

2.5 不同检测算法的比较

通过比较不同的检测方法,并将本文算法与传统算法进行比较,验证了所提模型的性能。

为了验证特征融合的检测性能,在十折交叉验证数据策略下,将多频谱图特征和相应的BFN连接特征作为典型检测算法输入,然后将典型检测算法与本文所提模型的抑郁症检测结果进行比较。本文中典型的检测算法包括K-近邻算法(KNN)、RF和SVM。表3中给出了典型检测算法的主要参数设置。

表3 3种检测算法主要参数
Table 3 Main parameters of 3 detection algorithms

算法	主要参数
KNN	k_neighbor=3, algorithm='kd_tree', metric='euclideanDis'
RF	n_estimators=20, criterion='gini', max_depth=None
SVM	C=1.0, kernel='rbf'

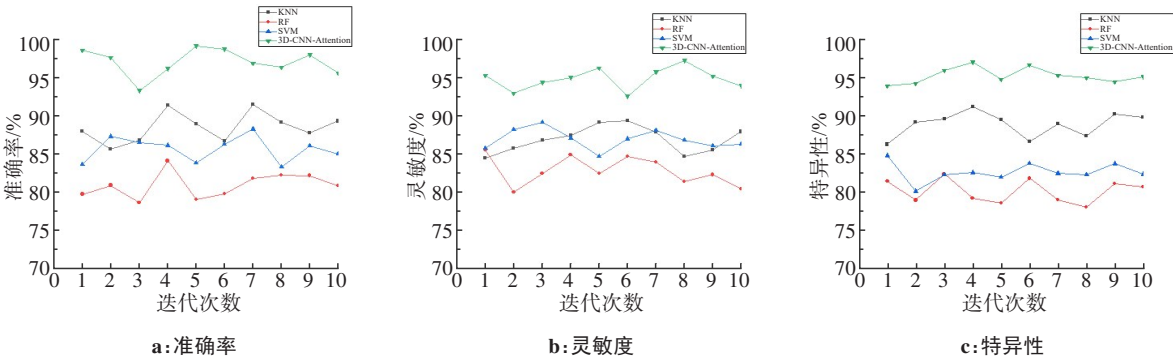


图8 不同检测算法的十折交叉验证结果

Figure 8 Ten-fold cross-validation results of different detection algorithms

为了清晰展示不同方法的检测性能,图8列出了十折交叉验证每次迭代的性能指标,3D-CNN-Attention模型最高准确率为98.74%,对应灵敏度、特异性分别为96.47%和94.11%(见迭代次数5);KNN的最高准确率为93.91%,对应灵敏度、特异性分别为84.27%和77.93%(见迭代次数4);RF的最高准确率为84.30%,对应灵敏度、特异性分别为84.51%和86.61%(见迭代次数4);SVM的最高准确率为87.38%,对应灵敏度、特异性分别为88.05%和84.53%(见迭代次数7)。

表4列出了不同检测算法的平均准确率对比结果,可以看出3D-CNN-Attention模型获得了最高的检测准确率。特别是与RF相比,提高了15.15%。

这种显著的性能提升可能归因于3D-CNN-Attention模型在处理具有复杂空间关系和特征层次的数据时,能够更有效地学习和提取重要信息。相较于传统的机器学习方法如RF,3D-CNN-Attention通过深度学习框架深层次地理解数据结构,同时注意力机制的加入使得模型能够聚焦于最为关键的数据特征,从而在各种性能指标上都实现了更好的结果。此外,深度学习方法通常在拥有大量数据进行训练时性能更佳,能够捕捉到更加微妙和复杂的数据模式,这可能是其相对于基于规则的方法和传统统计学习方法具有更高准确率的另一个原因。

3 结论

针对抑郁症的检测问题,提出了一种基于多频谱图与BFN邻接矩阵融合的3D-CNN-Attention网络检测框架,所得主要结论如下:(1)为了实现抑郁症的准确检测,本文将不同频段的特征排列成三维积分特征张量以及将时序EEG数据映射到空间BFN,提取多频谱图特征和BFN邻接矩阵,输入至3D-CNN-Attention网络,从而完成了高准确率的抑郁症检测,提高了治疗抑郁症的便捷性。(2)通过设计基

表4 不同检测算法的检测性能对比($\bar{x}\pm s$, %)

Table 4 Performance comparison among different detection algorithms (*Mean* \pm *SD*, %)

指标	KNN	RF	SVM	3D-CNN-Attention
准确率	88.42 \pm 2.26	81.17 \pm 1.86	85.89 \pm 3.02	96.32 \pm 2.31
灵敏度	86.92 \pm 2.72	82.81 \pm 2.57	87.53 \pm 1.76	94.73 \pm 2.99
特异性	89.03 \pm 3.21	79.91 \pm 3.14	82.88 \pm 3.44	95.36 \pm 2.87

于CA和SA的3D注意力模块,增强EEG特征学习能力,充分发挥CNN在图像处理方面的优势。显著增强了抑郁症检测的精度。

【参考文献】

[1] World Health Organization. Depressive disorder (depression)[EB/OL]. (2023-03-31). <https://www.who.int/news-room/fact-sheets/detail/depression>.

[2] Evans-Lacko S, Aguilar-Gaxiola S, Al-Hamzawi A, et al. Socio-economic variations in the mental health treatment gap for people with anxiety, mood, and substance use disorders: results from the WHO World Mental Health (WMH) surveys[J]. Psychol Med, 2018, 48(9): 1560-1571.

[3] 张冰涛, 周文颖, 李延林, 等. 基于脑功能网络的抑郁症识别研究[J]. 生物医学工程学杂志, 2022, 39(1): 47-55.
Zhang BT, Zhou WY, Li YL, et al. Research on depression recognition based on brain function network [J]. Journal of Biomedical Engineering, 2022, 39(1): 47-55.

[4] 陈海霞, 马欣旭, 王艺霖, 等. 线粒体功能障碍与抑郁症发病机制的研究进展[J]. 神经解剖学杂志, 2023, 39(4): 476-480.
Chen HX, Ma XX, Wang YL, et al. The research progress of mitochondrial dysfunction and the pathogenesis of depression [J]. Chinese Journal of Neuroanatomy, 2023, 39(4): 476-480.

[5] Khan DM, Masroor K, Jailani MFM, et al. Development of wavelet coherence EEG as a biomarker for diagnosis of major depressive disorder[J]. IEEE Sens J, 2022, 22(5): 4315-4325.

[6] Mumtaz W, Xia LK, Mohd Yasin MA, et al. A wavelet-based technique to predict treatment outcome for major depressive disorder[J]. PLoS

One, 2017, 12(2): e0171409.

[7] Zhdanov A, Atluri S, Wong W, et al. Use of machine learning for predicting escitalopram treatment outcome from electroencephalography recordings in adult patients with depression [J]. JAMA Netw Open, 2020, 3(1): e1918377.

[8] Jaworska N, de la Salle S, Ibrahim MH, et al. Leveraging machine learning approaches for predicting antidepressant treatment response using electroencephalography (EEG) and clinical data [J]. Front Psychiatry, 2019, 9: 768.

[9] Li XW, La R, Wang Y, et al. EEG-based mild depression recognition using convolutional neural network[J]. Med Biol Eng Comput, 2019, 57(6): 1341-1352.

[10] Sharma G, Parashar A, Joshi AM. DepHNN: a novel hybrid neural network for electroencephalogram (EEG) -based screening of depression[J]. Biomed Signal Process Control, 2021, 66: 102393.

[11] Sun JF, Hong XF, Tong SB. Phase synchronization analysis of EEG signals: an evaluation based on surrogate tests[J]. IEEE Trans Biomed Eng, 2012, 59(8): 2254-2263.

[12] Lee YJ, Yoon J. Nonlinear image upsampling method based on radial basis function interpolation[J]. IEEE Trans Image Process, 2010, 19(10): 2682-2692.

[13] Liu SQ, Wang X, Zhao L, et al. 3DCANN: A spatio-temporal convolution attention neural network for EEG emotion recognition[J]. IEEE J Biomed Health Inform, 2022, 26(11): 5321-5331.

[14] Su EZ, Cai SQ, Xie LH, et al. STAnet: a spatiotemporal attention network for decoding auditory spatial attention from EEG[J]. IEEE Trans Biomed Eng, 2022, 69(7): 2233-2242.

[15] Mumtaz W, Xia LK, Ali SSA, et al. Electroencephalogram (EEG)-based computer-aided technique to diagnose major depressive disorder (MDD)[J]. Biomed Signal Process Control, 2017, 31: 108-115.

[16] Sharma M, Achuth PV, Deb D, et al. An automated diagnosis of depression using three-channel bandwidth-duration localized wavelet filter bank with EEG signals[J]. Cogn Syst Res, 2018, 52: 508-520.

[17] Zhang XW, Lu DW, Pan J, et al. Fatigue detection with covariance manifolds of electroencephalography in transportation industry[J]. IEEE Trans Industr Inform, 2021, 17(5): 3497-3507.

[18] Zhang XW, Pan J, Shen J, et al. Fusing of electroencephalogram and eye movement with group sparse canonical correlation analysis for anxiety detection [J]. IEEE Trans Affect Comput, 2022, 13(2): 958-971.

[19] Chambon S, Galtier MN, Arnal PJ, et al. A deep learning architecture for temporal sleep stage classification using multivariate and multimodal time series[J]. IEEE Trans Neural Syst Rehabil Eng, 2018, 26(4): 758-769.

(编辑:薛泽玲)