

# 基于多注意力与多特征融合的脑电情感识别

吴家彬 阳小明 曹太强

(西华大学电气与电子信息学院 成都 610039)

**摘要:** 为充分融合脑电信号中多维度情感信息以提升情感识别性能,本文提出了一种多注意力与多特征融合的网络模型。该模型结合了脑半球非对称性以及脑电信号的空间、频谱和时域特性,通过并行双输入路径进行特征提取。利用并行注意力机制强化频率通道和空间信息的表达,同时通过动态核选择方式调整卷积核的尺寸,结合深度可分离卷积进一步提取和压缩特征。最后,通过Transformer编码层的融合,提取特征间的时序依赖关系和全局关联信息,从而实现情感分类。在SEED数据集上的三分类实验中,模型取得了98.53%的平均准确率,验证了该方法的优越性。此外,通过对注意力模块的可视化分析,进一步增强了模型的可解释性。

**关键词:** 脑电信号;情感识别;多特征融合;注意力机制;Transformer

中图分类号: TN911 文献标识码: A 国家标准学科分类代码: 510.4010

## Emotion recognition from EEG based on multi-attention and multi-feature fusion

Wu Jiabin Yang Xiaoming Cao Taiqiang

(School of Electrical and Electronic Information, Xihua University, Chengdu 610039, China)

**Abstract:** To fully integrate multi-dimensional emotional information from EEG signals and improve emotion recognition performance, this paper proposes a network model based on multi-attention and multi-feature fusion. The model combines the asymmetry of the brain hemispheres and the spatial, spectral, and temporal characteristics of EEG signals, performing feature extraction through parallel dual-input pathways. A parallel attention mechanism is used to enhance the expression of frequency channels and spatial information, while the size of convolution kernels is adjusted through dynamic kernel selection. Additionally, depthwise separable convolutions are employed to further extract and compress features. Finally, the temporal dependencies and global associations between features are captured through fusion in the Transformer encoding layer, enabling emotion classification. In three-class experiments on the SEED dataset, the model achieved an average accuracy of 98.53%, demonstrating the superiority of this approach. Furthermore, visual analysis of the attention module further enhances the interpretability of the model.

**Keywords:** EEG signals; emotion recognition; multi-feature fusion; attention mechanism; Transformer

## 0 引言

情感识别在智能人机交互、情感计算和心理健康监测等领域具有重要应用<sup>[1]</sup>。情感识别方法主要分为基于非生理信号和生理信号的方式。其中非生理信号如面部表情和语音信号等,易受主观因素干扰,被试者可以掩盖真实的情感状态<sup>[2]</sup>。而生理信号能够更为客观的反映情感状态,特别是脑电信号(electroencephalography, EEG)作为人脑活动的直接体现,能够直接反映个体情感变化,因而成为情感识别研究的热点<sup>[3-4]</sup>。

然而,EEG信号具有显著的非平稳性,并伴随大量噪

声和冗余信息,如何充分的利用有效的信息用于情感识别是该领域的一大挑战<sup>[5]</sup>。当前,脑电情感识别的研究框架主要包括特征提取和分类模型的构建。在特征提取方面,研究主要集中于时域、频域和时频域等维度<sup>[6]</sup>。其中微分熵(differential entropy, DE)因其能良好的表征EEG信号的非线性,为情感识别领域中最有效的特征之一<sup>[7]</sup>。此外,神经科学研究表明,情感状态下大脑皮层的电激活分布呈现不对称性。这一发现促使许多研究者通过分析大脑左右半球的对称性差异,提取非对称特征并构建相应的分类模型<sup>[8]</sup>。

随着深度学习技术的快速发展,端到端深度学习模型

因其能够有效处理复杂数据关系,逐渐成为脑电情绪识别的主流方法。An 等<sup>[9]</sup>提出了一种 3D 特征融合并通过卷积自编码器进行特征提取与分类的方法,在 DEAP 数据集上,效价和唤醒的准确率分别为 89.49% 和 90.76%。Li 等<sup>[10]</sup>采用深度可分离卷积网络,结合有序神经元长短期记忆网络 (long short-term memory, LSTM) 提取特征,在 SEED 数据集上实现了 94.61% 的三分类准确率。Su 等<sup>[11]</sup>提出了一种具有多尺度卷积内核的 3D 卷积神经网络模型实现情感分类,在 DEAP 数据集上实现了 95.67% 的四分类准确率。近年来,注意力机制被广泛的应用于深度学习模型中,使其能够聚焦输入中的关键信息。张冰雪等<sup>[12]</sup>利用一维卷积降低时序复杂度,并结合自注意力机制与 LSTM,在 DEAP 数据集上,效价和唤醒的准确率分别为 93.34% 和 91.76%。闫超等<sup>[13]</sup>提出了一种结合脑区注意力机制的一维卷积情感识别模型,在 DEAP 数据集上,效价和唤醒的准确率分别为 94.01% 和 93.55%。Qiao 等<sup>[14]</sup>通过生成对抗网络扩增数据样本,并利用空间和通道注意力网络进行分类,在 SEED 数据集上达到了 94.87% 的三分类准确率。Yao 等<sup>[15]</sup>利用 Transformer 的自注意力机制感知 EEG 信号的通道与时序信息,并结合卷积网络进行特征聚合,在 SEED 数据集上达到了 96.67% 的三分类准确率。Liu 等<sup>[16]</sup>提出了一种三维卷积注意力网络模型,将时空特征与注意力权重融合,在 SEED 数据集上实现了 97.35% 的三分类准确率。

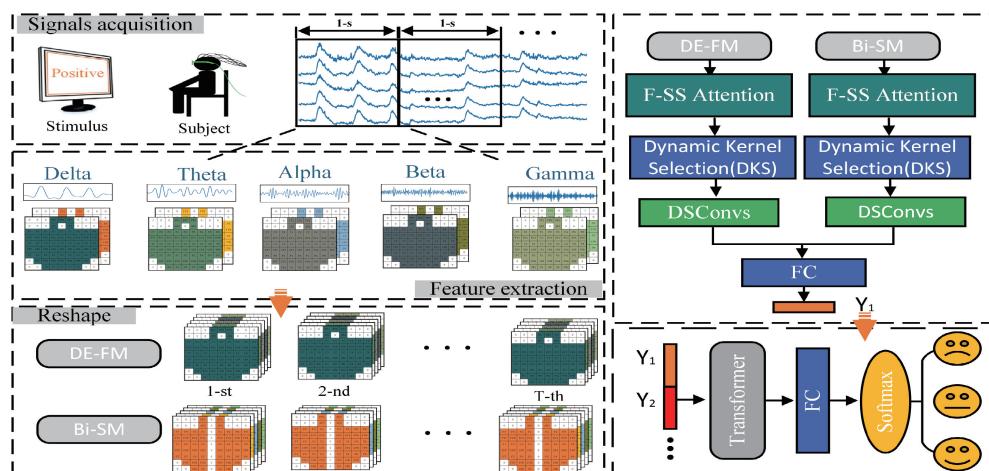


图 1 整体流程框架

Fig. 1 Overall process framework

首先,从 EEG 信号中提取不同频段,并计算每个电极在不同频段下的 DE 特征,构建 DE 特征矩阵 (DE feature matrix, DE-FM)。基于该矩阵,进一步构建双半球对称矩阵 (bi-hemisphere symmetric matrix, Bi-SM)。这两个矩阵作为双流输入,经过相同结构的特征提取网络进行处理。FSSA 模块通过自适应机制,动态调整频率与空间维度的权重,从而增强对关键频段和空间区域的关注。DKS 模块

尽管现有研究在脑电情感识别领域取得了一定进展,但大多数方法仍然存在以下不足:一是仅依赖单一特征或局部域信息,未能充分融合 EEG 信号中的多维度信息;二是对于 EEG 信号中的关键情感信息强调不足,导致分类准确率有待进一步提升。针对这些问题,本文提出了一种多注意力与多特征融合的网络模型,该模型采用并行双流输入结构,有效结合了脑电信号的时间、频谱、空间及脑半球非对称性等多维度信息。通过动态调整不同尺度卷积核的权重,扩展感受野的表示范围,并将注意力机制应用于双支路结构中的局部特征提取,以增强关键信息的捕捉能力。同时,结合 Transformer 中的自注意力机制进行全局信息交互,有效提升了各维度之间的关联性,显著增强了模型的表征能力,从而显著提高了情感识别的性能。

## 1 方 法

### 1.1 整体框架

为了有效提取脑 EEG 信号中的多维信息,并融合局部与全局特征,本文提出了一种基于双流输入的时空频特征融合网络模型。该模型包括频率-分离空间注意力模块 (frequency-separated spatial attention, FSSA)、动态核选择模块 (dynamic kernel selection, DKS)、深度可分离卷积模块 (depthwise separable convolutions, DSConvs)、Transformer 编码器及分类器等,整体流程框架如图 1 所示。

由多种尺寸的卷积核组成,旨在捕捉不同感受野下的局部信息,从而丰富整体特征。深度可分离卷积模块则用于进一步提取更为复杂的空间层级特征。

随后,融合来自两个模块的特征,构建该时间片段的特征表示。接着,将这些连续时间片段的特征输入 Transformer 编码器,进一步挖掘时序依赖关系及全局上下文信息。最终,使用 Softmax 分类器将提取的情感相关

信息映射到不同的标签空间,实现情感状态的推断。

## 1.2 特征提取

人类脑电信号通常分为 5 个主要频段 (Delta: 1~4 Hz, Theta: 4~8 Hz, Alpha: 8~13 Hz, Beta: 13~30 Hz, Gamma: 30~50 Hz), 这些频段与情感状态具有密切的关联。为提取这些频段的特征, 本文首先使用三阶带通巴特沃斯滤波器对原始脑电信号进行滤波, 后采用 1 秒的非重叠汉宁窗对 5 个频带信号进行分段处理, 并计算每个时间窗口的微分熵特征。微分熵用于量化随机变量的不确定性, 能够有效描述脑电信号的非线性特征, 在脑电情绪识别中表现出优异的性能。其计算公式如下:

$$DE = - \int_{-\infty}^{+\infty} f(x) \log(f(x)) dx \quad (1)$$

其中,  $f(x)$  表示连续信息的概率密度函数。对于近似服从高斯分布的 EEG 信号  $X \sim N(\mu, \sigma^2)$ , 其微分熵计算公式可简化为<sup>[17]</sup>:

$$DE(X) = \frac{1}{2} \log(2\pi e \sigma^2) \quad (2)$$

其中,  $e$  为欧拉常数,  $\sigma$  为 EEG 信号的方差。

### 1) DE 特征空间矩阵

为了保持脑电电极位置与其实际空间位置之间的对应关系, 本文对不同频段的 DE 特征进行了空间映射。具体映射过程如图 2 所示, 其中图中左侧的圆圈代表国际 10~20 系统中的一个电极通道, 依据脑电通道的相对空间位置, 将其投影到一个大小为  $H \times W$  的二维矩阵中, 其中  $H=W=9$ , 空白区域用零值填充。随后, 按频段顺序将各频段的特征矩阵拼接, 形成一个三维特征矩阵  $e_i$ 。接着, 将这些特征矩阵按时间顺序排列, 构建一个四维特征矩阵  $E = [e_1, e_2, \dots, e_T] \in \mathbb{R}^{T \times (C \times H \times W)}$ 。

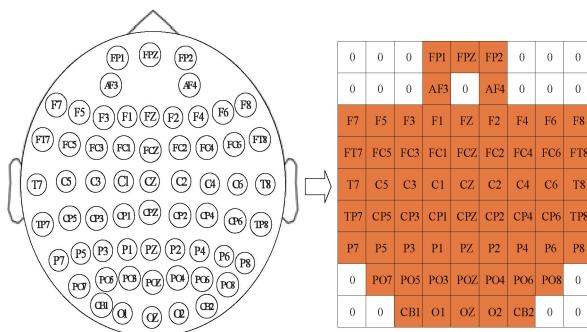


图 2 电极矩阵映射

Fig. 2 Electrode matrix mapping

### 2) 双半球对称矩阵

为了提取大脑半球在情感识别中的差异特征, 本文基于 DE 特征空间矩阵进一步构建了双半球对称矩阵。与先前的研究构建方法一致<sup>[18]</sup>, 首先, 将位于国际 10~20 系统中线上的电极 (Fpz, Fz, Fcz, Cz, Cpz, Pz, Poz, Cz) 值设为 0。剩余电极被划分为 27 对称电极对, 对称电极对是指相对中线的垂直距离相等的一对电极。(例如: FP1 与 FP2,

AF3 与 AF4)。对于每一对电极, 执行下式操作:

$$b_{f|t}^{p_j} = a \times (E_{f|t}^{l_j} - E_{f|t}^{r_j}), a = \begin{cases} 1, p \in left \\ -1, p \in right \end{cases} \quad (3)$$

其中,  $E_{f|t}^{l_j}, E_{f|t}^{r_j}$  分别代表第  $t$  秒时间, 第  $f$  频段上的第  $j$  对左右半球对称电极上的 DE 特征。得到的四维双半球对称矩阵数据  $B \in \mathbb{R}^{T \times C \times H \times W}$ , 其空间排列如式(4)所示。

$$B_{f|t} = \begin{bmatrix} 0 & 0 & 0 & b_{f|t}^{l_1} & 0 & b_{f|t}^{r_1} & 0 & 0 & 0 \\ 0 & 0 & 0 & b_{f|t}^{l_2} & 0 & b_{f|t}^{r_2} & 0 & 0 & 0 \\ b_{f|t}^{l_6} & b_{f|t}^{l_5} & b_{f|t}^{l_4} & b_{f|t}^{l_3} & 0 & b_{f|t}^{r_3} & b_{f|t}^{r_4} & b_{f|t}^{r_5} & b_{f|t}^{r_6} \\ b_{f|t}^{l_{10}} & b_{f|t}^{l_9} & b_{f|t}^{l_8} & b_{f|t}^{l_7} & 0 & b_{f|t}^{r_7} & b_{f|t}^{r_8} & b_{f|t}^{r_9} & b_{f|t}^{r_{10}} \\ b_{f|t}^{l_{14}} & b_{f|t}^{l_{13}} & b_{f|t}^{l_{12}} & b_{f|t}^{l_{11}} & 0 & b_{f|t}^{r_{11}} & b_{f|t}^{r_{12}} & b_{f|t}^{r_{13}} & b_{f|t}^{r_{14}} \\ b_{f|t}^{l_{18}} & b_{f|t}^{l_7} & b_{f|t}^{l_{16}} & b_{f|t}^{l_{15}} & 0 & b_{f|t}^{r_{15}} & b_{f|t}^{r_{16}} & b_{f|t}^{r_{17}} & b_{f|t}^{r_{18}} \\ b_{f|t}^{l_{22}} & b_{f|t}^{l_{21}} & b_{f|t}^{l_{20}} & b_{f|t}^{l_{19}} & 0 & b_{f|t}^{r_{19}} & b_{f|t}^{r_{20}} & b_{f|t}^{r_{21}} & b_{f|t}^{r_{22}} \\ 0 & b_{f|t}^{l_{25}} & b_{f|t}^{l_{24}} & b_{f|t}^{l_{23}} & 0 & b_{f|t}^{r_{23}} & b_{f|t}^{r_{24}} & b_{f|t}^{r_{25}} & 0 \\ 0 & 0 & b_{f|t}^{l_{27}} & b_{f|t}^{l_{26}} & 0 & b_{f|t}^{r_{26}} & b_{f|t}^{r_{27}} & 0 & 0 \end{bmatrix} \quad (4)$$

## 1.3 频率-分离空间注意力模块

输入的四维结构张量  $E$  和  $B$  包含了重要的空间和频率信息。FSSA 模块旨在增强有助于情绪识别的特征, 并抑制无效信息。为了减小不同注意力模块之间的相互干扰, 本文采用并行结构对注意力模块进行排列, 其结构如图 3 所示。

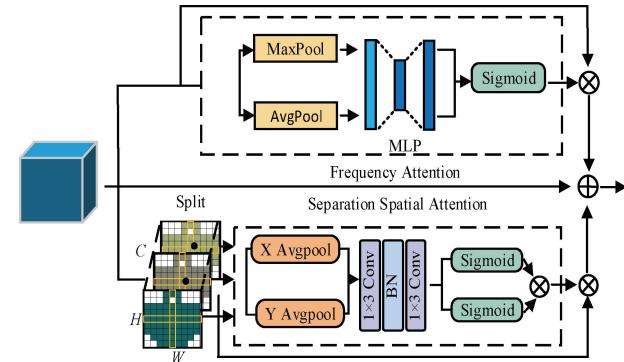


图 3 频率-分离空间注意力模块

Fig. 3 Frequency-separation spatial attention module

对于频率注意力模块, 将输入特征  $X$  的空间维度应用全局最大池化和平局池化, 然后采用多层感知机 (multi-layer perceptron, MLP) 进行处理, 并将输出结果相加, 并采用  $Sigmoid$  函数激活。公式可表示为:

$$A_F = \sigma(MLP(\text{MaxPool}(X)) + MLP(\text{AvgPool}(X))) \quad (5)$$

其中,  $\sigma()$  表示为  $Sigmoid$  激活函数。

空间注意力模块旨在聚焦于与情绪识别相关的关键电极位置。然而, 由于不同频段间关键电极位置的差异, 按传统方法通过将所有通道压缩为单一特征图, 并统一分

配每个通道的空间注意力权重,这种方法限制了注意力机制的自适应能力。为此,本文提出对不同频段的电极位置独立应用的分离空间注意力,以精细化调整各频段之间位置信息的重要性。

具体而言,首先将输入特征  $X$  按频率维度进行拆分。对于每个频段的特征,分别进行水平池化和垂直池化操作,以捕捉不同空间维度的信息。接着,将池化后的特征调整维度并拼接,通过卷积操作对其进行特征变换,后用参数分离得到水平和垂直方向的注意力向量,逐元素相乘生成对应的空间注意力图。最后,将各频段的空间注意力图拼接,得到最终的空间注意力权重。公式可表示为:

$$\begin{cases} M_h^i, M_w^i = f_2(\delta(f_1(\text{avgpool}(X_h^i); \text{avgpool}(X_w^i)))) \\ A_s^i = \sigma(M_h^i) \cdot \sigma(M_w^i) \\ A_s = \text{concat}(A_p^1, A_p^2, \dots, A_p^5) \end{cases} \quad (6)$$

其中,  $\text{avgpool}(X_h^i)$  和  $\text{avgpool}(X_w^i)$  分别表示在输入特征  $X$  的第  $i$  个通道上进行水平方向池化和垂直方向池化,  $\delta$  是非线性激活函数,  $f(\cdot)$  代表的是卷积操作, 卷积核大小为  $1 \times 3$ 。卷积 1 的输出通道数为 6, 卷积 2 的输出通道数为 2。随后, 将计算得到的频率和空间注意力权重应用于输入特征  $X$ , 并应用残差结构, 以生成最终的加权结果, 公式为:

$$A(X) = X + X \otimes A_F + X \otimes A_S \quad (7)$$

#### 1.4 动态核选择模块

不同频段的脑电特征在情绪表征能力上存在显著差异,而单一感受野的卷积操作难以全面捕获这些多样化的特征。为解决这一问题,本文提出了一种动态核选择模块,能够动态调整感受野大小,捕获脑电信号中不同空间尺度的特征,从而提升模型的情绪识别能力。如图 4 所示,该模块采用多分支结构,对输入脑电特征进行多尺度卷积操作,更全面地挖掘情绪相关的细节特征。

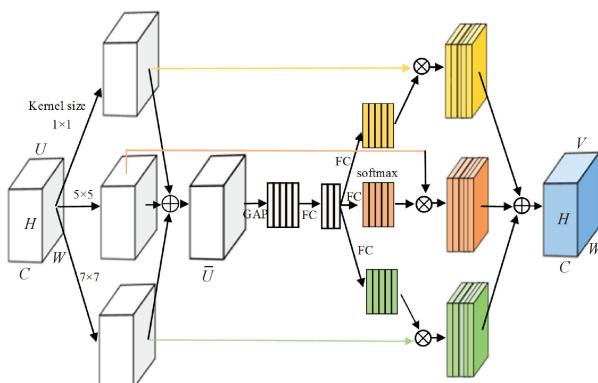


图 4 动态核选择模块

Fig. 4 Dynamic kernel selection module

具体而言,给定输入特征  $U \in R^{C \times H \times W}$ , 该模块首先通过尺寸分别为  $3 \times 3$ 、 $5 \times 5$  和  $7 \times 7$  的卷积核生成特征  $U_1$ 、 $U_2$  和  $U_3$ 。随后, 将这些特征逐元素求和得到  $\bar{U} \in$

$R^{C \times H \times W}$ 。接着, 对  $\bar{U}$  进行全局平均池化(global average pooling, GAP), 并经过全连接层(fully connected layer, FC)计算得到权重系数  $k_1, k_2, k_3$ 。最后, 使用 softmax 函数将对应位置的权重进行映射, 公式为:

$$w_c = \frac{e^{w_c}}{e^{k_{1,c}} + e^{k_{2,c}} + e^{k_{3,c}}}, w = k_1, k_2, k_3 \quad (8)$$

其中,  $k_{1,c}, k_{2,c}, k_{3,c}$  表示为  $k_1, k_2, k_3$  的第  $c$  个位置的元素。最终, 将得到的权重系数与卷积后的特征图逐元素相乘, 并将其求和以得到输出  $V$ :

$$V = k_1 \cdot U_1 + k_2 \cdot U_2 + k_3 \cdot U_3 \quad (9)$$

#### 1.5 深度可分离卷积模块

对于加权后的特征张量, 本文使用连续的深度可分离卷积模块进行更深层次的特征提取和降维处理。如图 5 所示, 深度可分离卷积对每个输入通道应用独立的卷积核进行特征提取, 并通过逐点卷积对各通道的输出进行线性组合, 从而生成最终的输出特征图。与传统卷积相比, 深度可分离卷积显著减少了模型的参数量和计算复杂度。该连续卷积结构由三层卷积组成, 每层卷积采用二维卷积并不进行填充, 同时通过 ReLU 激活函数引入非线性。卷积核的尺寸分别为  $3 \times 3$ 、 $3 \times 3$  和  $5 \times 5$ , 输出通道数依次为 16, 32 和 64。

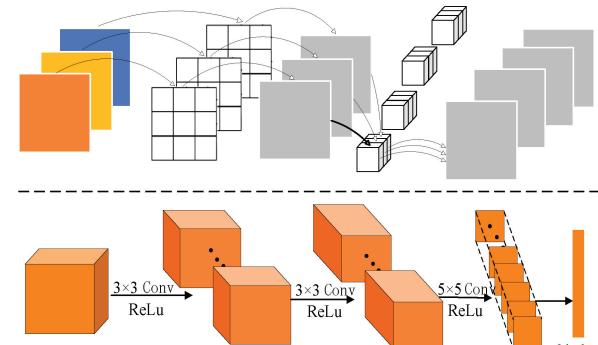


图 5 深度可分离卷积模块

Fig. 5 Depthwise separable convolution module

随后, 通过线性层将双流输入进行整合, 生成该时间片的综合表示  $Y_i$ 。接着, 将时间片特征按时间顺序连接, 公式表示为:

$$Y = [Y_1 \parallel Y_2 \parallel \dots \parallel Y_T] \quad (10)$$

其中,  $\parallel$  代表时间片的特征级联, 将时间段  $T$  内连续时间序列输入到 Transformer 特征融合模块, 以进一步提取时序信息和全局上下文。

#### 1.6 Transformer 特征融合模块

Transformer 编码器通过自注意力机制实现了在多个子空间中的并行信息交互, 有效地融合了局部信息与全局上下文。本文采用的 Transformer 编码器包含三层, 每层由多头自注意力模块和前馈全连接网络(feed-forward network, FFN)构成, 并结合层归一化(layer normalization,

LN) 和残差连接。该编码器的结构如图 6 所示。

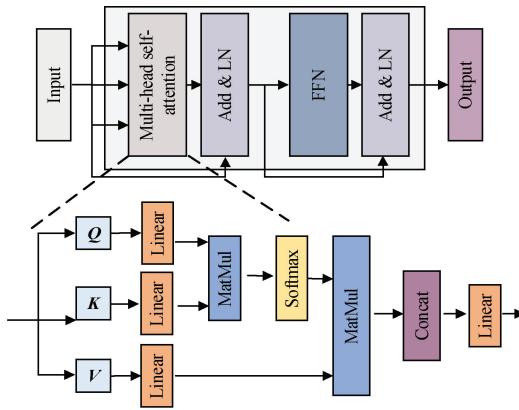


图 6 Transformer 特征融合模块

Fig. 6 Transformer feature fusion module

多头自注意力计算公式如下：

$$A_h(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (11)$$

其中,  $A_h$  代表每个注意力头的计算结果,  $h=1, 2, \dots, n$ 。在本文中  $n$  为 4, 其中  $Q, K, V$  分别代表查询向量, 键向量和值向量, 每一层的  $Q, K, V$  由上一层的编码器的输出线性映射得到。 $QK^T$  为点积操作, 用于计算输入序列中不同位置之间的相似性。 $\sqrt{d_k}$  为  $K$  的最后一个维度平方根, 可作为一种调节参数, 用于防止注意力权重集中在极少数位置上, 失去模型捕捉全局依赖关系的能力, 所有注意力头的输出经过连接后, 输入到全连接层进行进一步处理, 即:

$$M(Y) = \text{Linear}(\text{Concat}(A_1, A_2, \dots, A_n)) \quad (12)$$

然后通过残差结构和 LN:

$$F = LN(M(Y) + Y) \quad (13)$$

前馈全连接网络由两个线性映射层组成, 用于进一步的非线性变换和特征重组, 通过 FFN 层和 LN 可以得到  $Y'$ :

$$Y' = LN(\text{Linear}(\text{RELU}(\text{Linear}(F))) + F) \quad (14)$$

将得到的  $Y'$  输入下一个编码层, 将通过所有编码层提取的特征输入分类器进行情绪分类, 该分类器由全连接层和 SoftMax 操作组成。

## 2 实验与分析

### 2.1 数据集

本文实验使用了国际公开的脑电情感数据集 SEED<sup>[19]</sup>。SEED 数据集的实验流程如图 7 所示, 包含 15 名年龄约为 23 岁的受试者(7 名男性, 8 名女性)的 EEG 信号数据。每位受试者在不同时间段进行了 3 次实验, 每次实验间隔约一周, 共收集了 45 组实验数据。在每次实验中, 受试者观看了 15 段约 4 min 的电影片段, 每段影片前有 5 s 的提示, 影片结束后设有 45 s 的主观评价时间和

15 s 的休息时间。所选电影片段分为 3 种情绪类型: 消极、中性和积极, 每种类型包含 5 个片段。采集的脑电信号经过下采样处理, 采样率为 200 Hz, 并通过带通滤波处理至 0~75 Hz 范围, 数据通过国际 10~20 系统的 62 个电极通道进行记录。

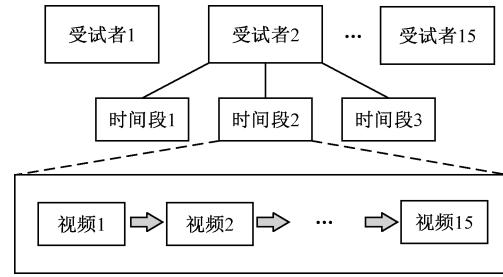


图 7 SEED 数据构建流程

Fig. 7 SEED data construction process

### 2.2 实验环境与参数

本文模型基于 PyTorch 框架, 并在 RTX 4060 GPU 上完成训练。在优化过程中, 采用 Adam 优化器, 并使用交叉熵作为损失函数。随机丢弃率(drop out)为 0.3, 时间片  $T$  被设置为 2 s, 则每位受试在不同时期分别可截取 1 692 个数据样本, 整体数据可获取 76 140(1 692 × 15 × 3)个数据样本。模型性能通过十折交叉验证进行评估, 其中数据集被随机划分为 10 个等大小的子集。每次从中选取 1 个子集作为测试集, 剩余 9 个子集作为训练集, 并重复该过程, 直至每个子集都作为测试集使用。最终, 通过计算所有测试集上的准确率的平均值, 并统计所有受试者数据样本总和的宏观精确率、召回率和 F1 分数, 作为评估模型的性能指标。

为进一步优化模型配置, 本文对前 3 位受试者(受试者 1、受试者 2 和受试者 3)的 EEG 数据进行了调优实验, 实验过程中, 批次大小(batch size)分别设置为 64、128 和 256, 训练周期(epoch)分别设置为 60、80 和 100。调优结果如表 1 所示, 结果表明, 当批次大小为 128 且训练周期为 100 时, 模型的各项指标表现最佳。

表 1 不同批次和时期的实验结果

Table 1 Experimental results of different batches and periods

Batchsize epoch	准确率	精确率	召回率	F1 值	%
(64,60)	96.82	96.80	96.81	96.80	
(64,80)	97.34	97.33	97.33	97.33	
(64,100)	97.58	97.58	97.57	97.58	
(128,60)	97.03	97.03	97.02	97.01	
(128,80)	97.54	97.53	97.53	97.53	
<b>(128,100)</b>	<b>97.64</b>	<b>97.63</b>	<b>97.64</b>	<b>97.63</b>	
(256,60)	95.57	95.55	95.56	95.56	
(256,80)	96.10	96.10	96.08	98.08	
(256,100)	96.46	96.48	96.46	96.45	

### 2.3 实验结果分析

为进一步评估模型在不同测试时期的性能,图8展示了各受试者在不同测试阶段的分类准确率差异。结果显示,所有受试者的平均准确率均超过95%,其中,受试者4的最低准确率为95.83%,而受试者15的最高准确率达到99.92%。同一受试者在不同会话时期的表现在一定波动,这可能是由于受试者在部分测试期间未能充分集中注意力,从而导致标注的脑电信号情绪标签与实际诱发的情感之间存在偏差。

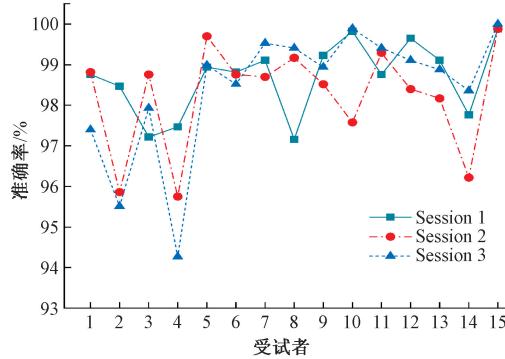


图8 受试者平均分类准确率

Fig. 8 Average classification accuracy of subjects

为更全面地评估本文所提出模型在各情感类别上的识别能力,并揭示其可能的偏差,图9展示了所有受试者情感样本测试结果的汇总的混淆矩阵。结果表明,模型在积极情感的识别上表现最佳,其次为中立情感和消极情感。

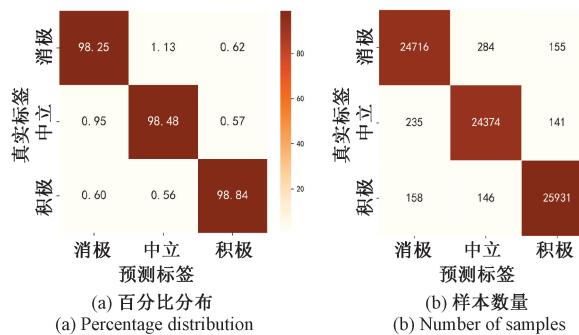


图9 汇总混淆矩阵

Fig. 9 Summary confusion matrix

### 2.4 注意力机制验证

#### 1) 注意力机制对比

本文模型引入了FSSA模块对特征进行加权,以增强有效信息。为了验证该模块的有效性,分别将其替换为传统的压缩激励模块(squeeze-and-excitation, SE)、空间注意力模块(spatial attention module, SAM)和卷积注意力模块(convolutional block attention module, CBAM)模块进行对比实验,实验结果如表2所示。

表2 不同注意力的准确率对比

Table 2 Accuracy comparison of different attention mechanisms

注意力	准确率	精确率	召回率	F1值	%
无注意力	97.64	97.62	97.64	97.63	
SE	98.18	98.18	98.19	98.18	
SAM	98.12	98.11	98.12	98.11	
CBAM	98.25	98.24	98.25	98.24	
FSSA	98.53	98.52	98.53	98.53	

从表2可以看出,SE和SAM分别在频率通道层次和空间层次上进行了增强,而CBAM则同时对频率通道和空间层次进行了增强,均有效地提升了模型性能。与此相比,本文提出的FSSA注意力模块减少了通道层次之间的干扰,并在空间层次上进行了更为精细的独立增强,其性能指标均优于其他注意力模块,验证了其有效性。

#### 2) 空间和频谱注意力分析

为了进一步揭示本文方法的潜在推理过程,本文在模型中可视化了空间和频谱注意力权重。将DE-FM和Bi-SM的权重掩码合并后绘制脑热图,如图10所示。可以看出,模型对各频段所关注的电极位置存在差异,这一现象验证了采用分离空间注意力机制的合理性。整体而言,模型更为关注前额叶、外侧颞叶等与情感更为相关的大脑区域。

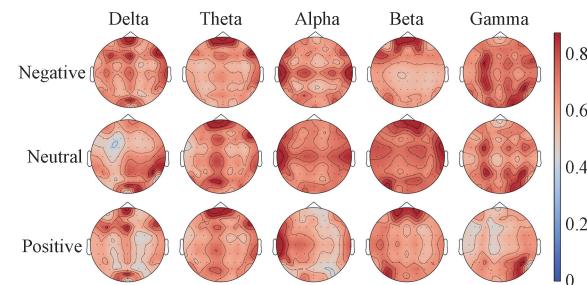


图10 空间权重可视化

Fig. 10 Spatial weight visualization

此外,针对频谱注意力权重,本文可视化了每位受试者在不同脑电频段的权重占比,如图11所示。结果表明,高频段的权重占比更高,特别是Gamma波段,其分配的权重最高,表明该波段可能与情感状态密切相关,这与现有研究结果一致<sup>[20]</sup>。

#### 2.5 消融实验

为了评估各模块对整体模型性能的影响,本文分别移除各模块,并与完整模型进行对比分析,其测试结果如表3所示。

可以看出去除DE-FM和Bi-SM模块后,模型的分类准确率分别下降了4.52%和0.43%,表明DE特征对情感识别具有更显著的贡献,而Bi-SM信息的引入能够进一步提升情感分类性能。当移除其中的DKS、Transformer和

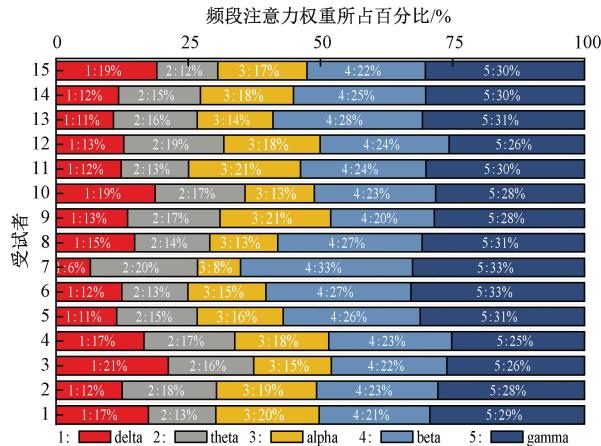


图 11 频段权重可视化

Fig. 11 Frequency band weight visualization

表 3 模块消融实验

Table 3 Module ablation experiment %

模型	准确率	精确率	召回率	F1 值
无 DE-FM	94.01	93.98	93.97	93.98
无 BI-SM	98.10	98.09	98.09	98.10
无 DKS	97.65	97.64	97.64	97.64
无 Transformer	97.83	97.82	97.81	97.82
无 DSconv	96.45	96.43	96.44	96.44
All	98.53	98.52	98.53	98.53

DSconv 等模块分类准确率分别下降了 0.88%、0.7% 和 2.08%，证明了各模块的有效性。实验结果表明，完整模型结构在所有评估指标上均表现最佳，去除任意模块均会导致整体性能下降。

## 2.6 同类型研究对比

为了验证本文所提出方法的有效性，本文将与近几年在 SEED 数据上进行研究的方法进行了对比，不同的方法得到的实验结果如表 4 所示。

表 4 同类型研究对比

Table 4 Comparison with similar studies

方法	特征	准确率/%
GECNN <sup>[21]</sup>	能量分布图像	92.93
ResNet <sup>[22]</sup>	2D-DE	94.88
MDGCRN-SRCNN <sup>[23]</sup>	DE+PCC	95.08
Conformer <sup>[24]</sup>	EEG	95.30
TDMNN <sup>[25]</sup>	4D-DE	97.20
Bi-ViTNet <sup>[26]</sup>	EEG+PSD	97.55
本文	DE-FM+Bi-SM	98.53

文献[21]将 EEG 信号转化为能量分布图像，并嵌入卷积神经网络中进行特征提取与分类，在 SEED 数据集上

达到了 92.93% 的准确率。文献[22]提取脑电信号的 DE 特征，并用差值法提高特征图分辨率，采用深度残差网络进行情感分类，在 SEED 数据上实现了 94.88% 的准确率。文献[23]通过构建 DE 和皮尔逊相关系数 (pearson correlation coefficient, PCC) 基础上的图域特征，利用动态图卷积网络与卷积神经网络提取深层次特征进行分类，在 SEED 数据集上取得了 95.08% 的准确率。文献[24]采用卷积和 Transformer 结构相结合的方法，利用卷积网络提取局部特征，采用 Transformer 捕获全局相关性，在 SEED 数据集上实现了 95.3% 的准确率。文献[25]通过最小化不同时段 EEG 情绪特征的时间差异，结合卷积递归神经网络与多分支 LSTM，有效提取情绪特征进行分类，在 SEED 数据集上取得了 97.2% 的准确率。文献[26]提出了一种并行双分支结构，分别提取时空特征与空频特征用于情绪判别，在 SEED 数据集上的准确率为 97.55%。本文提出的模型结合脑半球对称性，综合提取 EEG 信号的时空频多维度特征，并在 SEED 数据集上达到了 98.53% 的准确率。与现有研究相比，本文模型的准确率更高，验证了该方法的有效性。

## 3 结论

为充分提取 EEG 信号中的关键信息用于情感分类任务，本文提出了一种基于双分支结构的网络模型，该模型将包含时空频维度的四维特征 DE-FM 和考虑双半球非对称性的 Bi-SM 作为输入。为了提升注意力机制的自适应能力，鉴于不同频率通道中重要电极位置有所差异，设计了一种并行结构的 FSSA 注意力模块，用于加权重要频率和空间位置。特征提取过程中，结合动态核选择与离散卷积模块，提取有效特征，并通过 Transformer 融合双流输入特征，以增强脑电信号上下文关系的建模能力，提升特征之间的全局交互。本文模型在 SEED 数据集上进行验证，获得了 98.53% 的平均准确率，相较于现有方法，表现出更优的情感识别性能。此外，为进一步提升模型的可解释性，本文对 FSSA 注意力模块进行了可视化分析，结果表明，模型能够有效揭示不同频率通道中关键电极位置的差异性，且高频波段在情感预测中展现出更强的影响力。

## 参考文献

- [1] 韩凌, 李栋, 王增霖, 等. 基于人脸表情与脑电信号的情绪识别系统[J]. 国外电子测量技术, 2023, 42(10): 190-195.
- [2] HAN L, LI D, WANG Z L, et al. Emotion recognition system based on facial expressions and EEG signals [J]. Foreign Electronic Measurement Technology, 2023, 42(10): 190-195.
- [3] 李奇, 常立娜, 武岩, 等. 基于深层图卷积的 EEG 情绪识别方法研究[J]. 电子测量技术, 2024, 47(4): 18-22.
- [4] LI Q, CHANG L N, WU Y, et al. Research on EEG

- emotion recognition method based on deep graph convolution[J]. *Electronic Measurement Technology*, 2024, 47(4): 18-22.
- [3] 冯笑,代少升,黄炼.基于可解释深度学习的单通道脑电跨被试疲劳驾驶检测[J].*仪器仪表学报*,2023,44(5):140-149.
- FENG X, DAI SH SH, HUANG L. Cross-subject fatigue driving detection based on interpretable deep learning with single-channel EEG[J]. *Chinese Journal of Scientific Instrument*, 2023, 44(5): 140-149.
- [4] 唐意恒,王永雄,王哲,等.基于子领域适应和时空学习的脑电情感识别[J].*应用科学学报*,2024,42(6):1016-1026.
- TANG Y H, WANG Y X, WANG ZH, et al. EEG emotion recognition based on subdomain adaptation and spatio-temporal learning[J]. *Journal of Applied Science*, 2024, 42(6): 1016-1026.
- [5] MAHMOUD A, AMIN K, AIRAHHAL M M, et al. A CNN approach for emotion recognition via EEG[J]. *Symmetry*, 2023, 15(10): 1822.
- [6] 秦天鹏,生慧,岳路,等.脑电信号情绪识别研究综述[J].*计算机工程与应用*,2023,59(15):38-54.
- QIN T P, SHENG H, YUE L, et al. A review of EEG signal emotion recognition research[J]. *Computer Engineering and Applications*, 2023, 59(15): 38-54.
- [7] 王泽田,张学军.基于多域信息融合的脑电情感识别研究[J].*电子测量技术*,2024,47(2):168-175.
- WANG Z T, ZHANG X J. Research on EEG emotion recognition based on multi-domain information fusion[J]. *Electronic Measurement Technology*, 2024, 47 (2): 168-175.
- [8] GONG L L, CHEN W Z, ZHANG D G. An attention-based multi-domain bi-hemisphere discrepancy feature fusion model for EEG emotion recognition [J]. *IEEE Journal of Biomedical and Health Informatics*, 2024, 28(10): 5890-5903.
- [9] AN Y L, HU S H, DUAN X Y, et al. Electroencephalogram emotion recognition based on 3D feature fusion and convolutional autoencoder[J]. *Frontiers in Computational Neuroscience*, 2021, 15: 743426.
- [10] LI Q, LIU Y Q, LIU Q Y, et al. Multidimensional feature in emotion recognition based on multi-channel EEG signals[J]. *Entropy*, 2022, 24(12): 1830.
- [11] SU Y, ZHANG Z X, LI X, et al. The multiscale 3D convolutional network for emotion recognition based on electroencephalogram [J]. *Frontiers in Neuroscience*, 2022, 16: 872311.
- [12] 张冰雪,李文楷.少量通道脑电信号的实时情绪分类模型[J].*小型微型计算机系统*,2024,45(2):271-277.
- ZHANG B X, LI W K. Real-time emotion classification model for a small number of EEG channels [J]. *Mini-micro Systems*, 2024, 45 ( 2 ): 271-277.
- [13] 闫超,张雪英,张静,等.结合注意力机制和特征融合1DCNN的脑电情感识别[J].*计算机工程与应用*,2023,59(13):171-177.
- YAN CH, ZHANG X Y, ZHANG J, et al. EEG emotion recognition combined with attention mechanism and feature fusion 1D CNN[J]. *Computer Engineering and Applications*, 2023, 59 ( 13 ): 171-177.
- [14] QIAO W X, SUN L, WU J H, et al. EEG emotion recognition model based on attention and GAN[J]. *IEEE Access*, 2024, 12: 32308-32319.
- [15] YAO X Z, LI T W, DING P, et al. Emotion classification based on transformer and CNN for EEG spatial-temporal feature learning[J]. *Brain Sciences*, 2024, 14(3): 268.
- [16] LIU S Q, WANG X, ZHAO L, et al. 3DCANN: A spatio-temporal convolutional attention neural network for EEG emotion recognition [J]. *IEEE Journal of Biomedical and Health Informatics*, 2021, 26 ( 11 ): 5321-5331.
- [17] LI D D, XIE L, CHAI B, et al. Spatial-frequency convolutional self-attention network for EEG emotion recognition [J]. *Applied Soft Computing*, 2022, 122: 108740.
- [18] HUANG D M, CHEN S T, LIU C, et al. Differences first in asymmetric brain: A bi-hemisphere discrepancy convolutional neural network for EEG emotion recognition[J]. *Neurocomputing*, 2021, 448: 140-151.
- [19] ZHENG W L, LU B L. 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.
- [20] YANG K, TONG L, SHU J, et al. High gamma band EEG closely related to emotion: evidence from functional network [J]. *Frontiers in Human Neuroscience*, 2020, 14: 89.
- [21] SONG T F, ZHENG W M, LIU S Y, et al. Graph-embedded convolutional neural network for image-based EEG emotion recognition [J]. *IEEE Transactions on Emerging Topics in Computing*,

- 2021, 10(3): 1399-1413.
- [22] 王竟茜, 苗敏敏, 徐宝国, 等. 基于脑电空间域表征可视化的情感识别研究[J]. 传感技术学报, 2023, 36(9): 1385-1394.
- WANG J Q, MIAO M M, XU B G, et al. Emotion recognition based on EEG spatial domain representation visualization[J]. Journal of Sensors, 2023, 36(9): 1385-1394.
- [23] BAO G C, YANG K, TONG L, et al. Linking multi-layer dynamical GCN with style-based recalibration CNN for EEG-based emotion recognition [J]. Frontiers in Neurorobotics, 2022, 16: 834952.
- [24] SONG Y H, ZHENG Q Q, LIU B C, et al. EEG conformer: Convolutional transformer for EEG decoding and visualization[J]. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2022, 31: 710-719.
- [25] JU X Y, LI M, TIAN W L, et al. EEG-based emotion recognition using a temporal-difference minimizing neural network [J]. Cognitive Neurodynamics, 2024, 18(2): 405-416.
- [26] LU W, TAN T P, MA H. Bi-branch vision transformer network for EEG emotion recognition[J]. IEEE Access, 2023, 11: 36233-36243.

### 作者简介

吴家彬,硕士研究生,主要研究方向为脑机接口技术、智能信息处理。

E-mail: xhueduwjb@163.com

阳小明(通信作者),教授,硕士生导师,主要研究方向为无线射频技术、智能信息处理。

E-mail: 0120030022@xhu.mail.edu.cn

曹太强,教授,硕士生导师,主要研究方向为电力电子技术。

E-mail: ctq815@163.com