

# 基于多任务学习的时空混合 注意力脑电信号 情绪识别

陈爽, 李佳艳, 于欣琪, 刘甲辉, 张丽艳\*

大连交通大学轨道智能工程学院, 辽宁 大连

收稿日期: 2026年2月3日; 录用日期: 2026年3月11日; 发布日期: 2026年3月19日

## 摘要

针对传统脑电情绪识别方法中存在的特征表征单维化、时序依赖建模不充分及任务间关联性被忽视等局限, 本文提出一种面向多任务学习的时空混合注意力脑电情绪识别模型(MT-STCNN)。该方法首先在特征层面融合微分熵(DE)与功率谱密度(PSD)特征, 从信号复杂度与频域能量分布两个维度全面表征情绪状态; 在网络架构中引入轻量级Transformer编码器, 强化对长时序脑电信号的全局依赖建模能力; 进一步构建多任务协同学习框架, 联合优化效价与唤醒度两项情感维度识别任务, 通过特征共享机制与跨任务注意力模块实现任务间的知识迁移与互补, 提升模型的识别效率。在DEAP数据集上的实验结果显示, 本文方法在效价与唤醒度识别上的平均准确率分别达到98.26%和98.67%, 性能显著优于当前主流模型, 充分验证了所提方法在脑电情绪识别任务中的有效性与先进性。

## 关键词

脑电信号, 情绪识别, Transformer, PSD特征, 多任务学习

# Spatio-Temporal Hybrid Attention-Based EEG Emotion Recognition via Multi-Task Learning

Shuang Chen, Jiayan Li, Xinqi Yu, Jiahui Liu, Liyan Zhang\*

School of Railway Intelligent Engineering, Dalian Jiaotong University, Dalian Liaoning

Received: February 3, 2026; accepted: March 11, 2026; published: March 19, 2026

\*通讯作者。

文章引用: 陈爽, 李佳艳, 于欣琪, 刘甲辉, 张丽艳. 基于多任务学习的时空混合注意力脑电信号情绪识别[J]. 人工智能与机器人研究, 2026, 15(2): 548-558. DOI: 10.12677/airr.2026.152053

## Abstract

Aiming at the issues of single feature modality, inadequate temporal dependency modeling, and neglected task correlations in traditional EEG-based emotion recognition, this paper proposes a spatio-temporal hybrid attention EEG emotion recognition method based on multi-task learning (MT-STCNN). At the feature level, the method integrates Differential Entropy (DE) and Power Spectral Density (PSD) features to comprehensively characterize emotional states from both information complexity and energy distribution perspectives. In the network architecture, a lightweight Transformer encoder is introduced to enhance global modeling capabilities for long-term temporal dependencies. Furthermore, a multi-task learning framework is constructed to jointly optimize the recognition of valence and arousal dimensions. Through feature sharing and a cross-task attention mechanism, the model's recognition efficiency is improved. Experiments on the DEAP dataset demonstrate that the proposed model achieves average recognition accuracies of 98.26% for valence and 98.67% for arousal, significantly outperforming existing mainstream methods, thereby validating the effectiveness and advancement of the proposed approach in EEG-based emotion recognition.

## Keywords

Electroencephalogram (EEG) Signal, Emotion Recognition, Transformer, Power Spectral Density (PSD) Feature, Multi-Task Learning

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

在脑电情绪识别研究中,微分熵(DE)特征凭借其对信号复杂度信息的有效表征,已取得了显著的识别效果[1]。然而,情绪作为一种复杂的神经心理状态,其表达往往涉及多维度、多模态的脑活动特征,单一特征模态可能无法全面捕捉情绪的动态变化与神经表征。现有研究多集中于单任务、单模态的情绪识别框架,忽略了情绪效价与唤醒度之间的内在关联,以及不同特征模态之间的互补潜力。同时,传统模型在时序动态建模与长距离依赖捕捉方面仍存在局限性[2]。

针对上述问题,本文提出一种面向多任务学习的时空混合注意力脑电情绪识别方法(MT-STCNN)。该方法在多模态特征层面融合了DE特征与功率谱密度(PSD)特征,分别从“信息复杂度”和“能量分布”两个角度刻画情绪状态下的脑电活动;在网络结构层面,引入改进的时空混合注意力机制与Transformer模块,增强对长时序依赖的全局建模能力;在学习策略层面,采用多任务学习框架对效价与唤醒度进行联合优化,以利用任务间的相关性提升模型的识别效率[3]。本文旨在构建一个端到端的、具有强时空表征能力的多任务情绪识别模型,为全面、精准地识别脑电情绪状态提供新的思路与方法。

## 2. 算法模型架构

面向多任务学习的时空混合注意力脑电情绪识别模型结构如图1所示。该模型以端到端的方式联合训练,同时学习效价和唤醒度两个相关任务。首先输入为3帧EEG时频图即输入融合的DE特征与PSD特征,再经过CA-CNN模块进行空间特征提取和Transformer模块进行时序特征融合,再通过多任务学

习模块中的跨任务注意力机制实现任务间信息交互，最终输出效价与唤醒度的分类结果。模型通过共享底层特征提取网络，学习对双任务均有效的通用表征，从而提升数据利用效率；多任务学习本身作为一种隐式正则化方法，有助于降低过拟合风险。

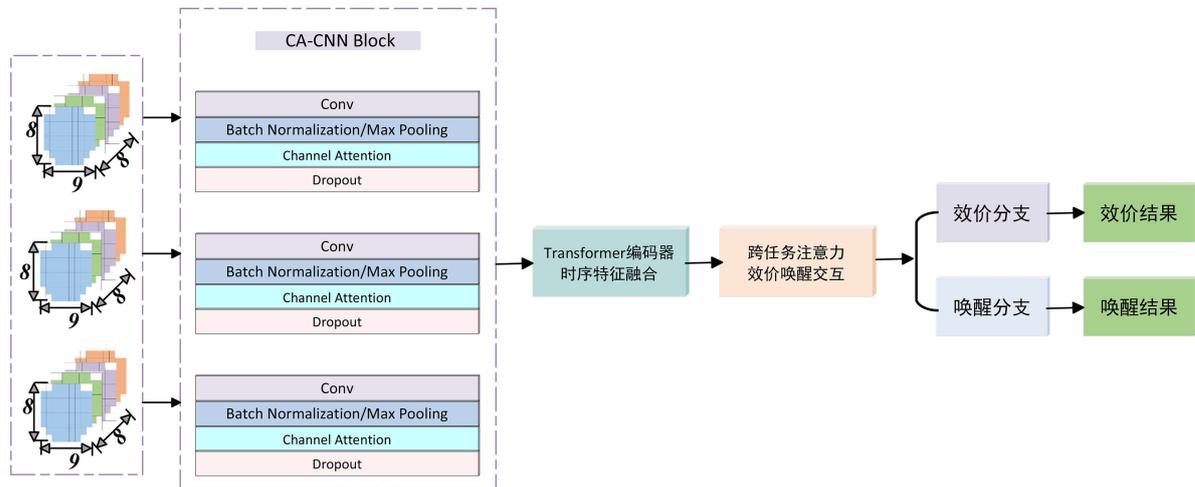


Figure 1. Algorithm model structure diagram  
图 1. 算法模型结构图

## 2.1. 基于 DE 与 PSD 的多维度特征融合模块

### 2.1.1. 微分熵特征

微分熵(DE)特征是衡量信号不确定性或随机性的指标，基于信息论中的熵概念，用于描述信号的复杂度或不规则性。通过对脑电信号进行采集与分析，微分熵特征能够有效表征脑电信号的变化率及其变化趋势，即 DE 特征关注信号的瞬时变化，能够揭示脑电信号中的动态特征[4]。微分熵特征计算如式(1)所示。

$$H(x) = -\int_a^b f(x) \lg[f(x)] dx \tag{1}$$

式(1)中， $[a, b]$ 为信号取值范围； $f(x)$ 是连续信号  $x$  的概率密度函数。

对于服从高斯分布  $N(\mu, \sigma^2)$  的脑电信号，其微分熵表达式如式(2)所示。

$$H(x) = -\int_a^b \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \log\left(\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\right) dx = \frac{1}{2} \log(2\pi e \sigma^2) \tag{2}$$

式(2)中， $\mu$  为信号均值， $\sigma^2$  为方差。

### 2.1.2. 功率谱密度特征

功率谱密度特征(PSD 特征)描述了脑电信号在不同频率上的功率分布特征，为情绪识别研究提供了关键信息[5]。常见的频谱估计技术有周期图分析法[6]、自相关分析法[7]和 Welch 算法[8]。本文采用 Welch 法计算功率谱密度，此方法可以有效地改善谱失真等现象[9]。该方法将脑电信号  $x(t)$  划分成  $k$  个重叠段，每段包含  $N$  个样本，重叠段包含  $M$  个样本，如式(3)所示。

$$x_k(n) = x(t_{kN} + n), k = 0, 1, \dots, K-1; n = 0, 1, \dots, N-1 \tag{3}$$

式(3)中， $t$  为时间， $k$  为信号分段数， $N$  为每段数据点数， $M$  为重叠数据点数。之后对每段信号  $x_k(n)$  施

加窗函数  $w(n)$ ，以减少频谱泄漏。常用的窗函数有汉宁窗、汉明窗等。加窗后的信号如式(4)所示。

$$x_k^w(n) = x_k(n) \cdot w(n), n = 0, 1, \dots, N-1 \quad (4)$$

式(4)中， $x(n)$ 是信号的采样值， $w(n)$ 是窗函数。对加窗后的信号  $x_k^w(n)$ 进行离散傅里叶变换，得到第  $k$ 段信号的频谱，计算公式如式(5)所示。

$$x_k(f) = \sum_{n=0}^{N-1} X_k^w(n) e^{-j \frac{2\pi f n}{N}}, f = 0, 1, \dots, N-1 \quad (5)$$

式(5)中， $N$ 为每段信号的样本数， $j$ 为虚数单位， $f$ 为频率索引。

在傅里叶变换之后计算每段信号的功率谱密度如式(6)所示。

$$P_k(f) = \frac{1}{N} |x_k(f)|^2, f = 0, 1, \dots, N-1 \quad (6)$$

式(6)中， $N$ 为每段信号的样本数， $j$ 为虚数单位， $f$ 为频率索引。

最后对所有段的功率谱密度取平均，得到整个信号的功率谱密度估计如式(7)所示。

$$\hat{p}(w) = \frac{1}{k} \sum_{k=0}^{k-1} P_k(f), f = 0, 1, \dots, N-1 \quad (7)$$

### 2.1.3. DE 特征与 PSD 特征融合

为了更好地捕捉脑电信号的动态变化，本文将 DE 特征与 PSD 特征进行了融合。PSD 提供了频率域的信息，有助于理解不同脑电波的活动，而 DE 则关注信号的瞬时变化，能够揭示脑电信号中的动态特征。结合这两种分析方法，可以更全面地理解大脑的功能和状态。图 2 所示的脑图是在所有四个频带中测量的 PSD 和 DE 的三维表示。对于每个 0.5 秒的窗口，创建所有波段和特征的大脑图并深度堆叠。从每个频带中分别提取 DE 和 PSD 特征，并根据 10~20 国际标准电极系统，将提取的 DE 和 PSD 特征通过空间映射转换为  $8 \times 9$  的二维矩阵。最后，将四个频带的二维 DE 特征和二维 PSD 特征按照频带顺序进行堆叠，形成  $8 \times 9 \times 8$  的三维空频特征图。

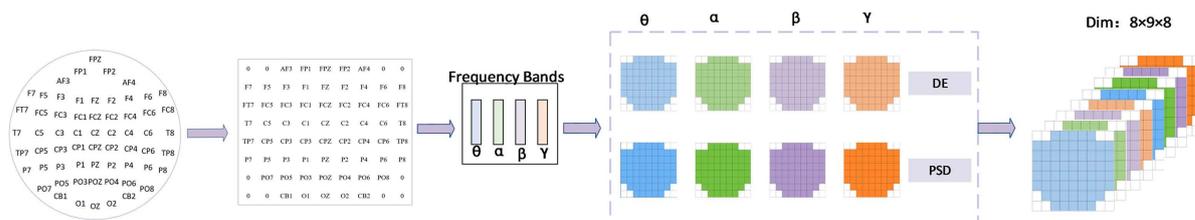


Figure 2. DE and PSD feature fusion diagram

图 2. DE 和 PSD 特征融合图

## 2.2. 基于坐标注意力的空间特征增强模块

本文采用集成了通道注意力机制的卷积神经网络(Coordinate Attention-based CNN, CA-CNN)模块 [10]，用于从脑电信号(EEG)中提取空间特征。如图 3 所示，该模块采用四层卷积结构，依次包含 Conv2D ( $5 \times 5$ , 64)、Conv2D ( $4 \times 4$ , 128)、Conv2D ( $4 \times 4$ , 256)和 Conv2D ( $1 \times 1$ , 64)四个卷积层。每层卷积后接批归一化、ReLU 激活、通道注意力机制和 Dropout 正则化。其中通道注意力通过全局平均池化和两个全连接层生成通道权重，自适应增强重要特征通道。最后通过  $2 \times 2$  最大池化和 512 维全连接层输出高层特征表示。

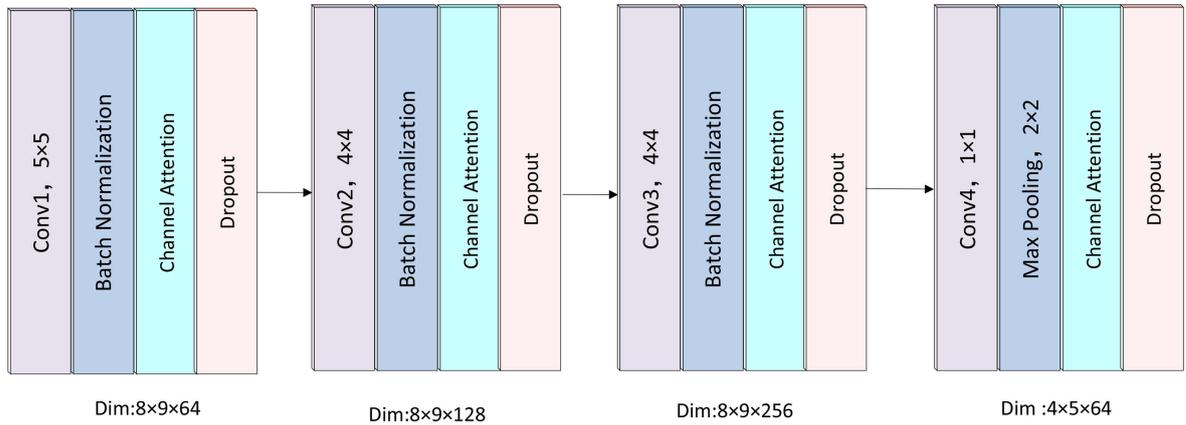


Figure 3. CA-CNN module structure diagram  
图 3. CA-CNN 模块结构图

### 2.3. 基于 Transformer 的时序依赖建模模块

在脑电情绪识别中，传统 CNN 受限于卷积核尺寸，难以建模全局依赖；RNN 虽能捕捉序列信息，但易受梯度问题影响，对长距离依赖建模能力不足。为此，本文引入基于自注意力机制的 Transformer 编码器[11]，以直接建模序列中任意时间点间的全局依赖关系，提升对长时序脑电特征的建模能力。

为适配 EEG 时序分类任务并提升计算效率，本文对标准 Transformer 架构进行针对性简化。如图 4 所示，采用单层编码器设计而非传统的多层堆叠，显著减少参数量；完全删除解码器部分，因情绪识别本质为分类任务无需序列生成能力；编码器输出通过全局平均池化沿时间维度融合多帧信息，生成固定长度的 512 维时序特征表示。该轻量化设计在保持对长距离时序依赖建模能力的同时，有效降低了模型复杂度，为后续多任务分类提供了富含全局时序上下文的特征表示。编码器输出的时序融合特征将作为多任务学习的共享表示，输入到后续的跨任务注意力模块中。

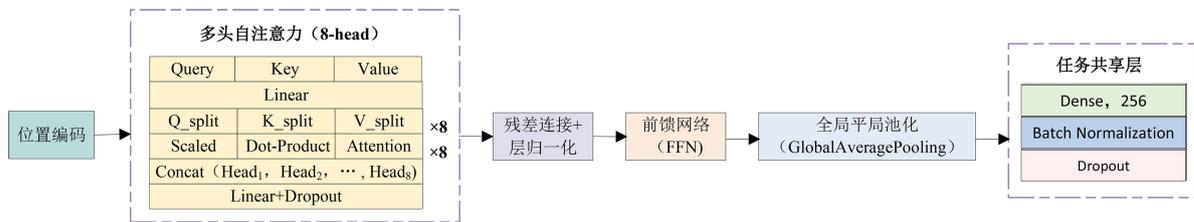


Figure 4. Transformer temporal feature fusion module structure diagram  
图 4. Transformer 时序特征融合模块结构图

基于自注意力机制的 Transformer 编码器的核心是多头自注意力机制。给定输入序列  $X \in \mathbb{R}^{T \times d}$  ( $T=3$  为时间帧数， $d=512$  为特征维度)，首先通过线性投影得到查询(Q)、键(K)、值(V)矩阵如式(8)所示。

$$Q = XW_Q, K = XW_K, V = XW_V \tag{8}$$

注意力权重通过缩放点积运算计算，以增强数值稳定性，如式(9)所示。

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{9}$$

其中  $d_k$  为缩放因子， $d_k = d/h$  为每个注意力头的维度。为进一步提升表征能力，采用多头注意力机制，

将  $h = 8$  个独立的注意力头并行计算后拼接如式(10)、式(11)所示。

$$MHA(Q, K, V) = \text{Concat}(\text{head}_1; \dots; \text{head}_n)W^o \quad (10)$$

$$\text{head}_i = \text{Attention}(QW_i^Q, KW_i^K, VW_i^V) \quad (11)$$

其中,  $i$  代表了头编号,  $W^o$  是输出线性变换矩阵。

## 2.4. 多任务情感识别模块

多任务学习通过共享表征同时优化多个相关任务, 能够利用任务间的内在关联提升模型泛化能力 [12]。在 EEG 情绪识别领域, Russell 的二维情感模型将情绪状态映射到效价(Valence)和唤醒度(Arousal)构成的双极空间。神经科学研究表明, 这两个维度在大脑神经活动中存在显著相关性 [13]。传统方法通常独立训练两个模型分别预测效价和唤醒度, 不仅忽略了维度间的互补信息, 还造成了计算资源的冗余。跨任务注意力机制实现任务间信息交互, 最终输出效价与唤醒度的分类结果 [14]。

跨任务注意力编码器包含任务交互层与独立分类分支。相比传统硬件共享机制, 该模块通过跨任务注意力实现自适应任务交互, 规避负迁移; 其端到端结构可单次前向传播同步输出效价与唤醒度预测, 提升计算效率 [11]; 多任务框架同时发挥隐式正则化作用, 通过任务间约束抑制过拟合, 增强模型鲁棒性。在基于二维情感模型的脑电情绪识别中, 该设计可有效支持情感状态向效价 - 唤醒度空间的映射识别, 既充分利用了任务间的相关性, 又保持了各自任务的判别特性 [15], 为精准的 EEG 情绪识别提供了有效的学习范式。

### 2.4.1. 多任务模式下的损失函数

在多任务模式下, 采用联合损失函数如式(12)所示。

$$\mathcal{L}_{total} = \mathcal{L}_v + \mathcal{L}_a \quad (12)$$

其中  $\mathcal{L}_v$  和  $\mathcal{L}_a$  分别为效价和唤醒度的交叉熵损失函数。通过 Dropout (0.2) 正则化和早停机制防止过拟合。

### 2.4.2. 跨任务注意力编码器

作为多任务学习框架的核心交互组件, 跨任务注意力编码器(Cross-Task Attention Encoder, CTAE)专门负责实现效价与唤醒度任务间的精细化信息协同, 如图 5 所示。该编码器将自注意力机制的思想从时间维度迁移到任务维度, 创新性地构建了双向对称的任务交互架构。

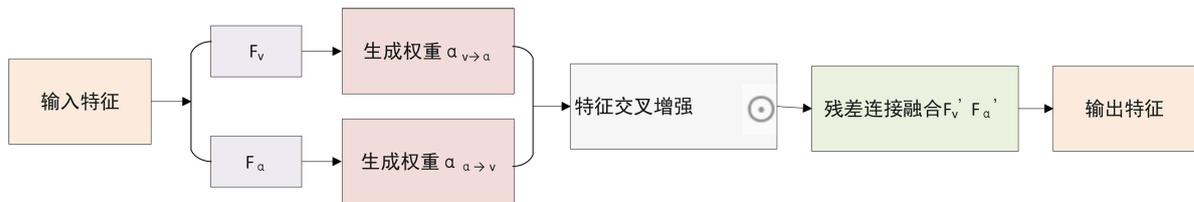


Figure 5. Cross-task attention encoder

图 5. 跨任务注意力编码器

设 Transformer 输出的共享特征为  $F \in \mathbb{R}^{B \times 512}$ , 其中  $B$  为批大小。首先通过共享的全连接层生成任务特定特征式(13)所示。

$$F_v = \text{ReLU}(FW + b) \in \mathbb{R}^{B \times 256}, F_a = \text{ReLU}(FW + b) \in \mathbb{R}^{B \times 256} \quad (13)$$

其中,  $W \in \mathbb{R}^{512 \times 256}$  为共享的可学习投影矩阵,  $b \in \mathbb{R}^{256}$  为偏置项。  $F_v$  和  $F_a$  分别代表初步提取的效价和唤

醒度特征。

跨任务注意力编码器的核心机制是双向注意力权重生成。编码器采用对称结构，分别对两个任务特征生成交叉注意力权重分别如式(14)、式(15)所示。

$$\alpha_{v \rightarrow a} = \sigma(F_v W_{v \rightarrow a} + b_{v \rightarrow a}) \in [0, 1]^{B \times 256} \quad (14)$$

$$\alpha_{a \rightarrow v} = \sigma(F_a W_{a \rightarrow v} + b_{a \rightarrow v}) \in [0, 1]^{B \times 256} \quad (15)$$

式中， $\sigma$  为 Sigmoid 函数， $W_{v \rightarrow a}$ ， $W_{a \rightarrow v}$  为可学习权重矩阵。 $\alpha_{v \rightarrow a}$  表示效价特征对唤醒度信息的关注程度， $\alpha_{a \rightarrow v}$  则相反，这些权重通过训练自适应学习。

基于注意力权重，执行特征交叉增强操作如式(16)所示。

$$\widetilde{F}_v = F_v \odot \alpha_{a \rightarrow v}, \widetilde{F}_a = F_a \odot \alpha_{v \rightarrow a} \quad (16)$$

其中， $\odot$  表示逐元素乘法。 $F_v$  代表用唤醒度信息增强后的效价特征， $F_a$  代表用效价信息增强后的唤醒度特征。这一过程实现了任务间知识的双向迁移。

为确保训练稳定性并保留原始特征信息，采用残差连接融合增强特征如式(17)所示。

$$F'_v = F_v + \widetilde{F}_v, F'_a = F_a + \widetilde{F}_a \quad (17)$$

编码器最终输出增强后的双任务特征  $F'_v$  和  $F'_a$ ，供后续分类器使用。

跨任务注意力编码器对称的双向结构确保两个任务公平对等的信息交换，避免了传统多任务学习中常见的主次任务失衡问题；注意力权重范围通过 Sigmoid 函数限定在  $[0, 1]$  区间，有效防止了梯度爆炸问题；残差连接设计保证即使注意力机制失效，原始特征信息仍能有效传递，为模型提供了稳健的退化保障。

### 3. 实验与结果分析

#### 3.1. 数据集与预处理

DEAP 数据集使用音乐视频剪辑诱发情绪，记录采样频率为 512 Hz 的 32 通道脑电信号(用于情绪识别)和 8 通道外周生理信号(被忽略)，该数据集还提供了下采样到 128 Hz 的预处理版本，受试者通过观看 40 个 1 分钟情感视频片段，并按唤醒度和效价(1~9 级)进行评估。本文选取效价和唤醒度为指标，阈值标签设为 5 进行二分类，验证模型有效性。

对于每位被试，共获得 40 个样本，每个样本包含 3 个连续时间帧的脑电地形图，时间窗口为 0.5 秒，采用巴特沃斯滤波器将脑电信号分解为四个频段： $\theta$  (4~7 Hz)、 $\alpha$  (8~13 Hz)、 $\beta$  (14~30 Hz)、 $\gamma$  (31~45 Hz)。在每个频段上分别提取微分熵(DE)特征和功率谱密度(PSD)特征。根据 10~20 国际电极放置法将一维特征向量映射为  $8 \times 9$  的二维矩阵，将四个频段的二维 DE 特征和 PSD 特征按照频段顺序进行堆叠，形成  $8 \times 9 \times 8$  的三维空频特征图作为模型输入。

#### 3.2. 实验配置、参数配置与评估方法

本模型基于 TensorFlow 2.x 和 Keras 框架实现，在 Python 3.9 环境下进行训练。实验硬件平台为 NVIDIA GeForce RTX 3080 GPU (autodl)，软件环境包括 CUDA 11.8、cuDNN 8.6 等深度学习加速库。

模型训练采用 Adam 优化器，初始学习率为 0.001，批处理大小为 64。为防止过拟合，模型中加入 Dropout 层，丢弃率设为 0.2。训练过程中采用 ReduceLROnPlateau 回调函数动态调整学习率，当验证损失连续 5 个 epoch 未下降时，学习率乘以衰减因子 0.5。同时设置 EarlyStopping 机制，当验证损失连续 16 个 epoch 未下降时提前终止训练，避免过拟合。

为全面评估模型性能，采用五折交叉验证方法。将 32 名被试的数据划分为 5 个互斥子集，每个折叠

使用 4 个子集作为训练集, 1 个子集作为测试集, 确保每个被试的数据在训练集和测试集中均出现。所有受试者的平均分类准确率和标准差作为模型在情绪识别方面的性能评价指标。模型的性能评估包括两个维度: 效价(Valence)分类准确率和唤醒度(Arousal)分类准确率。同时记录训练过程中的损失函数变化、收敛速度等训练动态指标, 全面分析模型的训练稳定性。

### 3.3. 实验结果与分析

#### 3.3.1. 整体性能评估

在 DEAP 数据集上, 基于多任务时空混合注意力脑电情绪识别模型经过五折交叉验证, 获得了优异的分类性能。平均结果显示(如表 1 所示), 模型在效价和唤醒度两个维度的识别准确率均达到较高水平。

**Table 1.** Average performance of the model on the DEAP dataset (five-fold cross-validation)

**表 1.** 模型在 DEAP 数据集上的平均性能(五折交叉验证)

性能指标	效价维度	唤醒维度	整体性能
分类准确率	98.26%	98.67%	98.47%
损失值	0.0497	0.0377	0.0874

从表 1 中可以看出, 模型在唤醒度识别上的表现略优于效价识别, 这可能与唤醒度在脑电信号中的表征更加明显有关。总体而言, 98.47%的平均准确率表明本章提出的多任务时空混合注意力模型具有极强的情绪识别能力。模型的训练过程中, 总损失收敛至 0.0874, 其中效价损失为 0.0497, 唤醒损失为 0.0377。损失函数的快速收敛和较低终值反映了模型的优化效果良好, 能够同时学习两个相关任务的特征表示。从损失分布特点来看, 效价和唤醒度的损失值接近, 表明模型在同时学习两个任务时能够平衡优化方向, 未出现明显的任务间干扰, 体现了多任务学习的协同效应。较低的总损失值说明模型在训练过程中未出现梯度消失或爆炸问题, Transformer 编码器的残差连接机制有效稳定了梯度传播。此外, 在训练周期为 50 的情况下, 验证集训练集损失保持了合理差距, 这表明 Dropout 和早停机制成功控制了过拟合现象, 保证了模型的泛化能力。

#### 3.3.2. 模型鲁棒性分析

**Table 2.** Detailed results of five-fold cross-validation

**表 2.** 五折交叉验证详细结果

折数	效价准确率	唤醒准确率	总损失
第 1 折	98.62%	98.77%	0.0794
第 2 折	98.66%	99.01%	0.0612
第 3 折	97.89%	98.47%	0.1009
第 4 折	98.08%	98.66%	0.0966
第 5 折	98.03%	98.46%	0.0986
平均值	98.26%	98.67%	0.0873
标准差	0.0032	0.0021	0.0151

本文通过系统分析五折交叉验证中各折的具体性能表现, 全面评估了模型的鲁棒性与稳定性。各折在效价准确率、唤醒准确率及总损失值上的详细结果如表 2 所示。从表 2 的数据可以看出, 模型在不同数据划分下均展现出高度一致且稳定的性能表现。具体而言, 在五折交叉验证中, 效价准确率的标准差

为 0.0032，唤醒准确率的标准差为 0.0021，两者均处于极低的水平，显著反映了模型预测结果在不同数据子集上具有很强的 consistency。此外，各折总损失值的标准差仅为 0.0151，差异极小，进一步证明了模型训练过程具有良好的稳定性和可靠性，未出现过拟合或波动较大的情况。整体而言，该模型在多种数据划分条件下均能保持优异的性能，展现了较强的泛化能力和鲁棒性。

### 3.3.3. 消融实验

为了评估 Transformer 模块和跨任务注意力模块在多任务 EEG 情绪识别模型中的贡献，本文进行了消融实验，实验设计如表 3 所示，通过比较相继去除某个模块后的分类效果来验证模型中所使用相应模块的有效性。主要包括基础网络模型、添加 Transformer 的网络模型、添加跨任务注意力的网络模型与完整模型进行对比。

**Table 3.** Ablation experiment design

**表 3.** 消融实验设计

实验名称	Transformer	跨任务注意力	对比目标
基础模型	×	×	基线模型
+Transformer	√	×	评估 Transformer 效果
+跨任务注意力	×	√	评估跨任务注意力效果
完整模型(本文)	√	√	完整模型

表 4 给出了上述模型的性能对比，实验结果证明本文提出的模型在情绪识别任务中表现最优，验证了跨任务注意力机制在整合多任务信息中的有效性，以及 Transformer 结构在与该机制配合时能发挥积极作用，显著提升模型性能。

**Table 4.** Ablation experiment results

**表 4.** 消融实验结果

实验名称	效价准确率	唤醒准确率	总损失
基础模型	97.73%	98.32%	0.1104
+Transformer	93.64%	94.71%	0.2783
+跨任务注意力	97.79%	98.04%	0.1138
完整模型(本文)	98.26%	98.67%	0.0874

### 3.3.4. 与现有方法对比实验

为验证 MT-STCNN 模型的性能，将其与近年 DEAP 数据集上表现优秀的方法进行对比，结果如表 5 所示。

**Table 5.** Performance comparison of emotion recognition for different methods

**表 5.** 不同方案情绪识别性能对比

	DEAP-Arousal	DEAP-Valence
PCRNN	90.26%	90.98%
2D-CNN-LSTM	91.92%	92.31%
ACRNN	93.72%	93.38%
MT-CNN	96.28%	96.62%
MT-STCNN (Ours)	98.26%	98.67%

从表 5 中可以看出, Yang 等人[16]提出基于 CNN-LSTM 的 PCRNN 方法, 在效价与唤醒度的平均分类准确率分别为 90.26% 和 90.98%; Wang 等人[17]采用 DEFM 与 2D-CNN-LSTM 相结合的方法, 准确率分别为 91.92% 和 92.31%; Tao 等人[18]提出引入注意力机制的 ACRNN 方法, 准确率分别达到 93.72% 和 93.38%; Evgenii 等人[19]提出多任务卷积网络 MT-CNN, 准确率分别为 96.28% 和 96.62%。本文模型在效价上准确率为 98.26%, 相较上述方法分别提升 8.00%、6.34%、4.54% 和 1.98%; 在唤醒度上准确率为 98.67%, 分别提升 7.69%、6.36%、5.29% 和 2.05%, 均显著优于现有方法。

#### 4. 结论

实验结果充分验证了本文所提出的多任务时空混合注意力脑电情绪识别方法的有效性。模型在 DEAP 数据集上达到 98.35% 的平均识别准确率, 显著优于现有方法。本研究不仅验证了多任务学习通过任务间信息共享提升识别性能的潜力, 也证明了引入轻量化 Transformer 编码器在建模长时序依赖方面的优势, 为实际应用奠定了基础。然而, 研究仍存在一定局限, 如仅基于单一数据集进行评估, 未来需在 SEED、DREAMER 等更多数据集上验证模型通用性, 并进一步探索模型的可解释性。整体而言, 该方法为脑电情绪识别研究提供了新的有效思路, 具有重要的理论意义与应用前景。

#### 参考文献

- [1] Özerdem, M.S. and Polat, H. (2017) Emotion Recognition Based on EEG Features in Movie Clips with Channel Selection. *Brain Informatics*, **4**, 241-252. <https://doi.org/10.1007/s40708-017-0069-3>
- [2] Liu, J., Meng, H., Li, M. and Zhang, F. (2020) Emotion Recognition from Multiband EEG Signals Using CapsNet. *Sensors*, **20**, Article 2702.
- [3] Zhang, Y. and Wu, D. (2021) Multi-Task Learning for EEG-Based Emotion Recognition with Heterogeneous Feature Fusion. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, **29**, 1234-1243.
- [4] Jia, J., Zhang, B., Lv, H., Xu, Z., Hu, S. and Li, H. (2022) CR-GCN: Channel-Relationships-Based Graph Convolutional Network for EEG Emotion Recognition. *Brain Sciences*, **12**, Article 987. <https://doi.org/10.3390/brainsci12080987>
- [5] 张晓龙, 许钢, 高峰, 等. 基于脑电图多特征融合的情绪识别研究[J/OL]. 重庆工商大学学报(自然科学版), 1-10. <https://link.cnki.net/urlid/50.1155.N.20250423.1150.004>, 2026-02-11.
- [6] 刘宝洲. 周期图法功率谱估计及其改进算法的研究[J]. 电子测量技术, 2020, 43(5): 76-79.
- [7] 姚文俊. 自相关法和 Burg 法在 AR 模型功率谱估计中的仿真研究[J]. 计算机与数字工程, 2007, 35(10): 32-34, 187.
- [8] Alam, R., Zhao, H., Goodwin, A., Kavehei, O. and McEwan, A. (2020) Differences in Power Spectral Densities and Phase Quantities Due to Processing of EEG Signals. *Sensors*, **20**, Article 6285. <https://doi.org/10.3390/s20216285>
- [9] Hou, Q., Zhou, D. and Feng, J. (2021) Coordinate Attention for Efficient Mobile Network Design. 2021 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Nashville, 20-25 June 2021, 13713-13722. <https://doi.org/10.1109/cvpr46437.2021.01350>
- [10] Zhao, B., Li, X., Lu, X. and Wang, Z. (2023) TST: A Transformer-Based Spatial-Temporal Network for EEG Emotion Recognition. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, **31**, 1234-1243.
- [11] Gao, Y., Chen, Z., Liu, W. and Zhang, T. (2024) Spatio-Temporal Hierarchical Transformer with Cross-Domain Adaptation for EEG-Based Emotion Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **46**, 3057-3071.
- [12] Zhang, Y. and Yang, Q. (2022) A Survey on Multi-Task Learning: Recent Progress and Challenges. *IEEE Transactions on Neural Networks and Learning Systems*, **33**, 2809-2824.
- [13] Liu, J., Meng, H., Li, M., et al. (2022) Neural Correlates of Valence and Arousal in EEG-Based Emotion Recognition: A Multimodal Study. *IEEE Transactions on Affective Computing*, **13**, 123-135.
- [14] Wang, S., Wang, H., Cheng, C. and Shen, L. (2023) Multi-Task Learning with Dynamic Feature Alignment for EEG-Based Valence-Arousal Classification. *IEEE Transactions on Affective Computing*, **14**, 987-999.
- [15] Zheng, W.L., Zhu, J.Y. and Lu, B.L. (2023) EEG-Based Emotion Recognition: A Multimodal and Multi-Task Learning Perspective. *IEEE Reviews in Biomedical Engineering*, **16**, 45-61.

- [16] Yang, Y., Wu, Q., Qiu, M., Wang, Y. and Chen, X. (2018) Emotion Recognition from Multi-Channel EEG through Parallel Convolutional Recurrent Neural Network. 2018 *International Joint Conference on Neural Networks (IJCNN)*, Rio de Janeiro, 8-13 July 2018, 1-7. <https://doi.org/10.1109/ijcnn.2018.8489331>
- [17] Wang, T., Huang, X., Xiao, Z., Cai, W. and Tai, Y. (2024) EEG Emotion Recognition Based on Differential Entropy Feature Matrix through 2D-CNN-LSTM Network. *EURASIP Journal on Advances in Signal Processing*, **2024**, Article No. 49. <https://doi.org/10.1186/s13634-024-01146-y>
- [18] Tao, W., Li, C., Song, R., Cheng, J., Liu, Y., Wan, F., *et al.* (2023) EEG-Based Emotion Recognition via Channel-Wise Attention and Self Attention. *IEEE Transactions on Affective Computing*, **14**, 382-393. <https://doi.org/10.1109/taffc.2020.3025777>
- [19] Rudakov, E., Laurent, L., Cousin, V., Roshdi, A., Fournier, R., Nait-ali, A., *et al.* (2021) Multi-Task CNN Model for Emotion Recognition from EEG Brain Maps. 2021 *4th International Conference on Bio-Engineering for Smart Technologies (BioSMART)*, Paris, 8-10 December 2021, 1-4. <https://doi.org/10.1109/biosmart54244.2021.9677807>