

## 基于LSTM-Transformer的脑电情感分析

王安琪<sup>1</sup>, 于超<sup>1</sup>, 陈胤玮<sup>1</sup>, 郝群<sup>2</sup>

1. 甘肃省中医药大学信息工程学院, 甘肃 兰州 730000; 2. 兰州大学第二医院信息中心, 甘肃 兰州 730001

**【摘要】**针对传统情感识别方法在处理长期依赖关系时的不足,提出一种结合长短期记忆网络(LSTM)与Transformer模块的脑电情感分析模型,称为LTNet。该模型首先对数据进行预处理,然后将提取的特征输入至LTNet。LSTM模块和Transformer模块独立对输入的序列进行建模,分别从中提取出深层的局部特征和全局特征。通过采用加权融合策略来综合这些特征,最终利用Softmax函数对情绪进行四分类。实验结果显示,在DEAP数据集上进行的五折交叉验证中,LTNet的平均识别准确率达到96.56%,相比于传统机器学习算法和其他深度学习方法提高2.74%~21.31%。

**【关键词】**脑电; 情绪识别; 深度学习; LSTM; Transformer

**【中图分类号】**R318; TP391.7

**【文献标志码】**A

**【文章编号】**1005-202X(2024)12-1550-08

## EEG emotion analysis based on LSTM-Transformer

WANG Anqi<sup>1</sup>, YU Chao<sup>1</sup>, CHEN Yinwei<sup>1</sup>, XI Qun<sup>2</sup>

1. School of Information Engineering, Gansu University of Chinese Medicine, Lanzhou 730000, China; 2. Information Center, the Second Hospital of Lanzhou University, Lanzhou 730001, China

**Abstract:** An electroencephalogram (EEG) emotion analysis model (LTNet) that combines long short-term memory (LSTM) and Transformer modules is proposed for addressing the shortcomings of traditional emotion recognition methods in dealing with long-term dependencies. After data preprocessing, the extracted features are input into LTNet. LSTM module and Transformer module model the input sequence independently, and from which deep local features and global features are extracted and then fused using a weighted fusion strategy. Finally, Softmax function is used to classify emotions into 4 categories. Experimental results show that LTNet has an average recognition accuracy of 96.56% in the 5-fold cross-validation on the DEAP dataset, which is 2.74%-21.31% higher than traditional machine learning algorithms and other deep learning methods.

**Keywords:** electroencephalogram; emotion recognition; deep learning; long short-term memory; Transformer

### 前言

情绪的识别是理解人体健康与心理状态的关键环节,尤其在心理学和计算机科学交叉领域中占据重要的地位。传统的情绪识别方法依赖于分析面部表情和语音语调等,而随着科技的发展,生理信号特别是脑电图(Electroencephalogram, EEG)信号,因其在分辨率、成本和可靠性方面的优势,被广泛应用于精确的情绪状态分析。EEG信号在多个领域展现出其独特价值,如监测飞行员和驾驶员的工作压力、辅

助情绪障碍的诊断以及对抑郁症的客观评估<sup>[1]</sup>。本文通过对EEG信号进行特征提取,利用特定算法和模型改进以提升情绪识别的准确性。研究者采用多种不同的神经网络模型结合的策略。Yin等<sup>[2]</sup>通过分析EEG信号的微分熵(Differential Entropy, DE)特征,结合图卷积神经网络(Graph Convolutional Neural Network, GCNN)和长短期记忆网络(LSTM)进行情绪状态分类,这种策略在处理受试者身份无关数据时,在情绪的效价和唤醒维度上的平均准确率分别达到84.81%和85.27%。张英杰等<sup>[3]</sup>采用DE特征并结合卷积神经网络(CNN)与LSTM,优化情绪分类的模型架构。在DEAP数据集上进行四分类实验中,该融合模型的平均分类准确率达到92.48%,展示不同神经网络架构结合应用的有效性。相较于仅使用LSTM模型<sup>[4]</sup>,融合模型在情绪分类性能上显示

**【收稿日期】**2024-07-08

**【基金项目】**甘肃省自然科学基金(20CX9JA145)

**【作者简介】**王安琪, 硕士, 研究方向: 脑电情绪识别, E-mail: 1101709211@qq.com

出显著的优势。在构建情绪识别模型时,采用多种神经网络模型结合的重要性,能够更有效地捕捉和分析情绪状态的复杂性。陈宗楠等<sup>[5]</sup>通过构建四维脑电特征数据,并应用 Swin Transformer 模型,在 SEED 脑电公开数据集上进行情绪三分类实验,跨被试者测试的准确率达到 89.63%。Zheng 等<sup>[6]</sup>提出一种直接从原始脑电信号中提取时空特征的情绪识别方法。时空对称 Transformer 模型专注于捕捉脑电信号的时空关联,这种方法在效价和唤醒的二元分类中的准确率分别为 89.86% 和 86.83%,展示了处理原始脑电信号在情绪识别中的应用潜力。

在情绪识别领域,能够有效提取 EEG 特征并用模型进行分类至关重要,这要求研究者采用合适的神经网络模型以及特征提取方法。本文引入并改进一系列算法和神经网络架构,旨在提升 EEG 情绪识别的模型性能。本文采用功率谱密度(Power Spectral Density, PSD)分析及包括 LSTM 和 Transformer 网络在内的现代神经网络架构来优化 EEG 信号处理。PSD 分析作为一种有效的信号处理手段,能够从 EEG 信号中提取频率相关特征,这对于理解情绪状态下大脑活动至关重要。脑电情绪识别涉及使用 EEG 信号来识别个体的情绪状态。这是一个时间序列问题,LSTM 在处理时间序列数据时表现出色,因为它可以有效地捕捉长期依赖关系。在脑电情绪识别中,LSTM 可以用于处理时间序列的脑电信号,以捕捉信号中的时间相关性和动态模式。另一方面,Transformer 作为一种先进的深度学习模型,也为脑电情绪识别提供一些独特的优势。虽然相对于 LSTM,Transformer 不依赖于循环结构,但它在全局依赖性建模和并行计算方面具有显著的优势。Transformer 中的多头注意力机制允许模型在不同层次上关注输入序列的不同部分。这对于理解脑电信号中的多模态信息和频率通道的重要性非常有帮助。因此,LSTM 和 Transformer 的结合可以为脑电情绪识别任务提供多层次、全方位的信息处理,从而提高模型的整体性能和泛化能力。

基于上述技术,本文提出 LTNet 混合深度神经网络模型。该模型整合 LSTM 和 Transformer 的优点,通过深度学习和模式识别显著提升情绪识别的准确度。该模型的独特结合使其在处理长期依赖关系和捕捉全局上下文方面更为强大,在情绪识别任务中实现一定程度的性能提升。

1 数据集与数据预处理

1.1 DEAP 数据集

DEAP 数据集为本研究提供丰富的多模态脑电

数据资源,该数据集记录了 32 名参与者在观看 40 段约 1 min 长度的音乐视频后,通过 32 导脑电信号和 8 导外周生理信号采集的数据。参与者根据自我评估量表对他们的情绪状态进行评分,覆盖唤醒度、效价、支配性和喜好度等多个维度,评分为 1~9 分<sup>[7]</sup>。本研究采用二维情绪模型,并将情绪状态根据评分划分为低(<5 分)与高(≥5 分)两个级别,形成 4 种不同的情绪类型:高唤醒高效价(HAHV)、低唤醒低效价(LALV)、高唤醒低效价(HALV)、低唤醒高效价(LAHV),这一分法为后续的情绪识别提供明确的分类依据。二维情绪模型如图 1 所示。

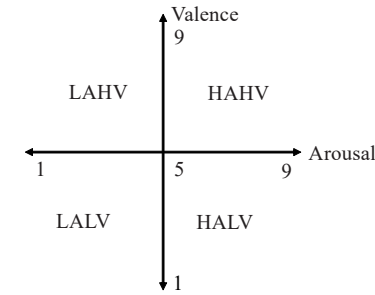


图 1 二维情绪模型  
Figure 1 Two-dimensional emotion model

1.2 数据预处理

在数据预处理部分,本文将原始数据的采样率从 512 Hz 降至 128 Hz,并通过带通滤波将信号限制在 4~45 Hz 内。此外,通过应用盲源分离法成功移除眼电伪影<sup>[8]</sup>。为了降低实验的复杂度和成本,选择直接使用数据集提供的预处理版本。DEAP 数据集中的数据格式如表 1 所示。每段数据包含 60 s 的情绪相关信号和 3 s 的基线信号。在本研究中,为了减少基线漂移对数据分析的影响,去除每段数据开始的 3 s 基线信号,使得随后的情绪状态分析更加准确。基线漂移是指长时间记录中信号的直流偏移,这种偏移可能会掩盖或扭曲情绪相关信号的真实变化<sup>[9]</sup>。接下来再使用 1 s 的时间窗口将剩余的 60 s 脑电数据分割成 60 个独立的 1 s 片段,进而生成 2 400 个样本数据。每个样本数据记录 32 通道的脑电信号,并从每个通道提取 4 个频段的功率谱特征(即 32×4)。然后,将相邻的 60 个片段中的功率谱密度沿时间方向拼接,形成脑电特征序列,为后续的情绪识别分析奠定基础。

2 模型介绍

2.1 LSTM

LSTM 是对传统循环神经网络(Recurrent Neural

表1 数据格式  
Table 1 Data formats

数据名称	数据维度	数据含义
data	40×40×8 064	视频数×通道数×数据点数
label	40×4	视频数×标签

Network, RNN)的重要改进,旨在解决处理长序列数据时RNN面临的梯度消失或梯度爆炸问题<sup>[10]</sup>。通过引入记忆单元和门控机制,LSTM能够在较长时间间隔内保留信息,从而有效地捕捉序列数据中的长期依赖关系。记忆单元是LSTM的核心,负责在网络中长期存储状态信息,使得网络能够记忆过去的信息并利用这些信息进行当前决策。此外,门控机制的设计是LSTM的一大创新,它包含3种类型的门:输入门、遗忘门、输出门<sup>[11]</sup>。输入门控制新信息的加入程度,遗忘门决定保留多少旧信息,而输出门则管理从记忆单元到输出的信息流。这些门控单元通过精细地调节信息流,使得LSTM模型在处理复杂的序列任务时,如情感分析时,表现出卓越的性能。

LSTM的关键优势在于其能够捕获长期依赖性,这一点对于脑电信号的情绪识别任务至关重要。情绪状态的表征通常与时间序列数据中的动态变化密切相关,LSTM通过其独特的门控机制有效地学习这些时序特征,从而提高情绪识别的准确性。选择LSTM作为本研究的一部分是基于其在处理时间依赖性方面的强大能力,可以从脑电信号中准确提取情绪状态相关的动态特征。

2.2 Transformer模型

Transformer模型首次由Vaswani等<sup>[12]</sup>在《Attention is all you need》论文中提出,它是一种基于自注意力机制的架构,拥有并行计算能力,最初设计用于解决机器翻译问题,后来被广泛应用于各种序列建模问题,显示出其广泛的适用性。LSTM模型由于其序列数据的依赖特性,每个时间步的处理都必须等待前一时间步的完成,这限制了其并行处理能力。不同于LSTM,Transformer不含有循环连接,而是基于注意力机制构建,这使得模型能够在处理序列数据时摆脱传统的逐步处理限制,实现数据处理的高效并行。

Transformer主要有两种并行性:自注意力机制并行与多头注意力机制并行<sup>[13]</sup>。在Transformer的自注意力机制中,每个位置的输出都是通过对所有位置进行加权求和得到的。这种机制的计算是高度并行的,因为可以同时计算每个位置的注意力权重,而这些权重可以独立地应用于输入序列的所有位置。

这使得模型能够同时处理整个序列,而不需要依赖序列中前后位置的计算结果,从而提高计算效率<sup>[14]</sup>。Transformer中的多头注意力机制通过将输入进行多次投影,然后并行地计算注意力,最后将结果进行拼接。这种机制允许模型从多个子空间学习特征,每个子空间都独立地并行处理。通过增加注意力头的数量,可以提高模型对不同特征的并行处理能力,加强对序列中不同位置的关注<sup>[15]</sup>。这两种并行性使得Transformer模型能够更好地处理长序列,同时利用硬件资源进行高效的计算。这对于处理脑电信号等序列数据,特别是在情感分析任务中具有重要意义,因为这些数据通常具有较长的时间依赖性和复杂的模式。

Transformer模型除自注意力机制与多头注意力外,还有位置编码、编码器与解码器结构、残差连接以及层归一化等核心组件。位置编码用于为模型提供序列中各元素的位置信息。编码器部分处理输入信息,而解码器则负责产生输出,两者共同构成模型的输入输出机制。残差连接和层归一化的作用是优化训练过程,通过缓解梯度消失问题来加速模型的学习速度。在脑电情感分析中,Transformer模型通过其自注意力机制能够有效捕获脑电信号中的复杂模式和动态变化,使其能够在识别和分析人类情绪状态时,更加准确地理解信号中的情绪信息。多头注意力机制允许模型在处理脑电信号时,从多个角度分析情绪相关的特征,增强模型对情绪状态多维度识别的能力。Transformer能够有效地处理脑电序列数据,考虑脑电信号的时间依赖性,为情绪状态的精确识别提供必要的序列信息。残差连接和层归一化的应用,进一步优化了模型的训练过程,提高情感分析的效率和稳定性。因此,Transformer模型在脑电情感分析领域的应用,展现其处理复杂序列数据的强大潜力,为深入理解人类情绪提供一种新的技术路径。

3 模型构建

3.1 总体结构

LSTM以其循环结构能够有效地捕捉时序数据中的长期依赖关系,尤其适用于处理脑电信号的时序数据。而Transformer则通过自注意力机制能够在全局范围内并行地捕捉序列中的信息,Transformer在多头注意力机制下可以并行地处理多个维度的信息<sup>[16]</sup>。随着序列长度的增加,LSTM的训练和计算可能面临梯度消失或梯度爆炸的问题,而Transformer在处理长序列时相对更为高效,因为它的自注意力机制可以并行计算。通过融合两者,可以有效地应



对长序列问题,综合利用LSTM的时序特征捕捉和Transformer的全局信息处理,提高模型的稳定性和训练效率,减轻过拟合风险<sup>[17]</sup>。图2是LTNet算法的总体结构图。本文详细阐述4个主要阶段:EEG信号预处理、数据分割、深度特征提取以及情绪识别分类。EEG信号预处理阶段涉及将信号带通滤波至4~45 Hz并移除眼电伪影,接着将连续的EEG信号细分为单独的脑电片段,并将每个片段上的PSD计算出来。

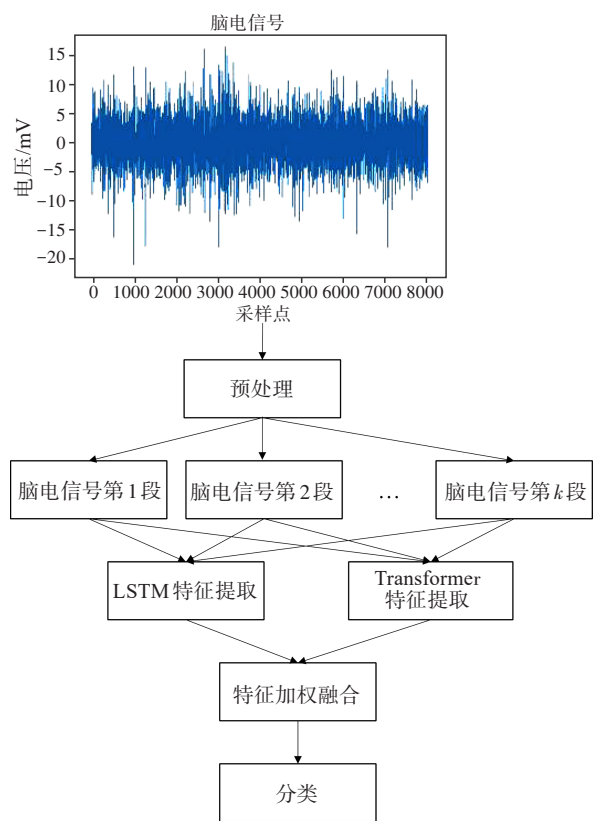


图2 LTNet算法总体结构图  
Figure 2 Overall structure of LTNet algorithm

在深度特征提取阶段,单独的脑电片段被并行输入到LSTM与Transformer模型中。本文采用权重平均的方法,它是一种经典的模型融合方法,该方法的核心思想是将多个模型的权重进行平均,以得到一个整体模型,从而达到提高模型性能和泛化能力的目的<sup>[18]</sup>。首先,独立地训练LSTM与Transformer模型,每个模型学习对数据的不同方面进行建模。在模型融合的过程中,对两个模型提取出的特征进行加权,反映其在整体性能中的贡献。这些权重可以基于不同模型在验证集上的性能表现来计算,性能较好的模型可能被分配更高的权重。然后,将每个模型的预测结果按照其相应的权重进行加权平

均。对于本文中的情绪分类问题,使用软投票的方式。最后,经过加权融合处理的特征输入Softmax分类器以获得情绪分类结果。

3.2 LSTM网络结构

本实验中的LSTM网络结构受到Alhagry等<sup>[19]</sup>启发,原始脑电信号被输入,经过分段处理,并通过两层LSTM与一层Dropout进行输出,取得良好的情绪分类效果。在这一基础上,本文与上述研究的不同之处在于将PSD特征作为模型输入进行改进,并设计一个基于LSTM的浅层结构,该结构由输入层、两个LSTM层、一个Dropout层、一个全连接层以及一个输出层组成,如图3所示。该网络的设计目的是为了处理和分析经过预处理和分割的EEG信号片段,每个片段具有128×1的输入维度。网络的第一层LSTM配置200个单元,采用ReLU激活函数,以增强模型非线性处理能力。为了抑制过拟合现象,接着第一层LSTM后设置一个Dropout层,值设定为0.2。第二层LSTM配置200个单元,采用Sigmoid激活函数,以便于处理和输出分类问题的概率。最后,数据通过全连接层进行最终的情绪状态分类。该网络结构在DEAP数据集上的表现通过五折交叉验证得以评估。实验结果显示,在训练集上,模型达到73.67%的平均准确率,证明模型在学习训练数据方面的有效性。然而,在测试集上,模型的准确率下降至65.72%,反映出模型对于未见过的数据的泛化能力存在限制。这一结果表明,在处理EEG信号时,过于复杂的网络结构可能不利于情绪分类任务的性能提升,在设计EEG信号处理模型时应考虑到网络结构的复杂度与任务性能之间的平衡。

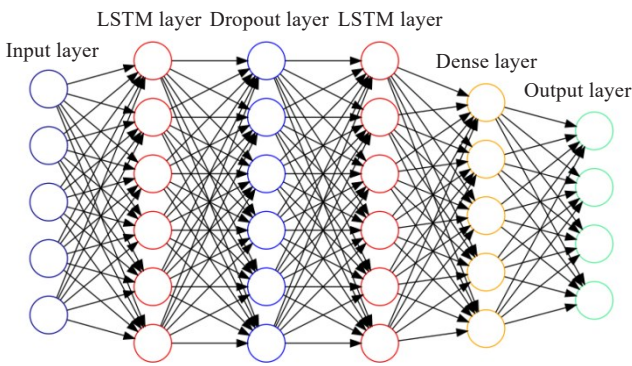


图3 LSTM网络结构图  
Figure 3 LSTM network structure

3.3 Transformer结构

Wang等<sup>[20]</sup>成功设计一种变体的Transformer结构,该结构在脑电情绪识别的二分类任务中不含解

码器,并在识别准确率方面取得较好的成绩。受到这一研究的启发,本文采用一种基于Transformer结构的相似变体,专用于脑电情绪识别任务。该变体的核心组件包括多头注意力机制、编码器、残差连接和层归一化。图4展示了Transformer结构的编码器部分。在本文的Transformer结构中,省略解码器和位置编码的部分,考虑到脑电情绪识别作为一种序列分类任务,其核心在于将编码器处理后的特征直接用于情绪状态的分类,因此解码器部分是非必需的。此外,位置编码通常在自然语言处理中区分输入序列中不同位置的单词和符号的作用。在脑电情绪识别任务中,时间序列数据已经按照时间顺序排列,每个时间步都对应着特定的时间点。因此,本研究也舍弃了位置编码模块。本文通过对Transformer模型结构的适当简化,成功利用其核心特性来处理脑电情绪识别任务,并在DEAP数据集上展示优异的性能。通过五折交叉验证的方法,本模型在训练集上的平均准确率为95.16%,而在测试集上的平均准确率为82.22%。实验结果不仅证明该模型在学习脑电信号方面的有效性,也显示其对新数据具有良好的泛化能力。

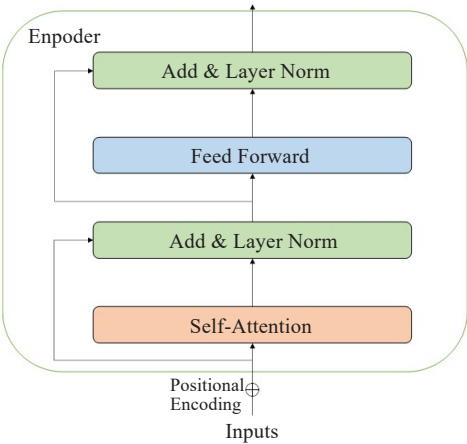


图4 Transformer 编码器部分  
Figure 4 Transformer encoder

## 4 结果与分析

### 4.1 实验设置

借助深度学习框架实现LSTM+Transformer混合模型,在Python3.8.5软件环境下进行实验,硬件设施为GeForce RTX 2080 GPU。采用Adam优化器计算交叉熵损失函数,设置学习率为0.001,采用加权融合的方法为每个子模型分配一个权重,实现500个epochs,并进行五折交叉验证实验。具体模型参数如表2所示。

表2 LSTM-Transformer 模型超参数设置  
Table 2 LSTM-Transformer model  
hyperparameter settings

超参数设置	取值
优化器	Adam
学习率	0.001
激活函数	ReLU
损失函数	交叉熵损失
Num_epochs	500
Dropout	0.2
LSTM 权重	0.5
Transformer 权重	0.5
多头注意力	4

### 4.2 评价指标

实验采用多个评价指标来综合评估模型性能,其中包括准确率(Accuracy)、精确率(Precision)、召回率(Recall)和F1分数(F1-score)。每个指标提供了对模型在不同方面表现的衡量,以便更全面地了解模型的性能。

(1)准确率表示模型预测正确的样本占总样本的比例,计算方法如下所示:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

其中,TP(True Positives)表示模型正确预测为正类别的样本数量;TN(True Negatives)表示模型正确预测为负类别的样本数量;FP(True Negatives)表示模型错误地将负类别预测为正类别的样本数量;FN(False Negatives)表示模型错误地将正类别预测为负类别的样本数量。

(2)精确率表示模型在预测为正类别的样本中有多少是真正的正类别,关注的是模型在正类别预测中的准确性,具体计算方法如下所示:

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

(3)召回率表示模型在所有真正的正类别样本中有多少被成功预测出来,关注的是模型找到了多少正类别样本,具体计算方法如下所示:

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

(4)F1分数是一种综合考虑精确率和召回率的评价指标,它通过计算这两者的调和平均值来综合考虑两者的性能。F1分数平衡了精确率和召回率,因此适用于那些需要同时关注分类器的准确性和能力捕捉正类别的任务。具体计算方法如下所示:

$$F1\_score = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

(4)

4.3 结果与讨论

本文采用五折交叉验证的方式进行模型性能评估,即将样本数分为5份,轮流选取其中的4份作为训练集,1份作为验证集,重复训练并验证模型5次,分别计算训练集的平均准确率、平均损失值与验证集的平均准确率、精确率、召回率、F1分数值。采用3种不同的模型(LSTM、Transformer、LTNet)对情绪分类任务进行训练,并在训练集上评估它们的准确率。如图5所示,随着训练次数的增加,3种模型的准确率逐渐趋于稳定状态。LTNet模型在训练过程中表现出相对较长的训练时间,而LSTM模型在此方面表现最为迅速,然而其平均准确率相对较低。相比之下,Transformer模型在耗时较少的情况下取得较高的平均准确率。这表明在情绪分类任务中,Transformer模型具备高效率 and 较优的准确率。LTNet模型在训练集中性能表现最为良好,平均准确率达到96.56%。

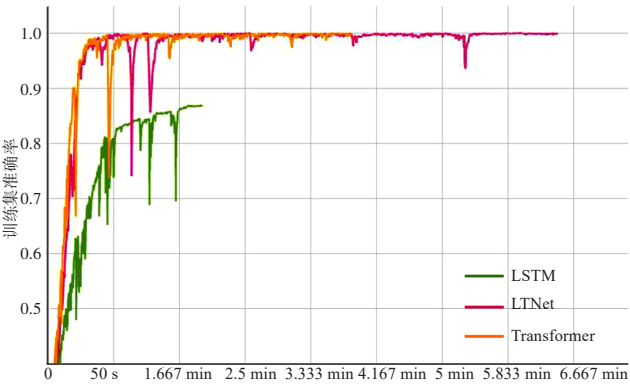


图5 模型在训练过程中的准确率变化  
Figure 5 Variation of accuracy during training

图6是3种不同模型在测试集上的准确率,值得注意的是在进行第二次交叉验证时,模型准确率显著提升,这表明模型在第一轮训练中有效地学习并捕捉到关键特征。从图中可以看出LTNet模型在第二次进行交叉验证时的准确率已经显著高于Transformer与LSTM模型,这说明LTNet模型在更深层次上理解和利用数据的信息,具备更强的泛化能力和性能优势。图7展示各模型在训练集上的损失,实验中采用交叉熵损失函数。从图中可知,在训练过程中,模型的损失逐渐减小。LSTM模型在损失方面表现最差,平均损失值达0.96。LTNet模型在损失下降速度上与Transformer模型相当,都有较快的收敛,这说明LTNet模型在训练过程中能

够有效地学习并适应数据特征,表现出良好的收敛性能,相较于LSTM模型具有更好的训练效果。表3对比3种模型在训练集和测试集上的平均准确率,在测试集上的平均精确率、召回率、F1分数值,如图8所示。实验中采用五折交叉验证,确保结果的鲁棒性。

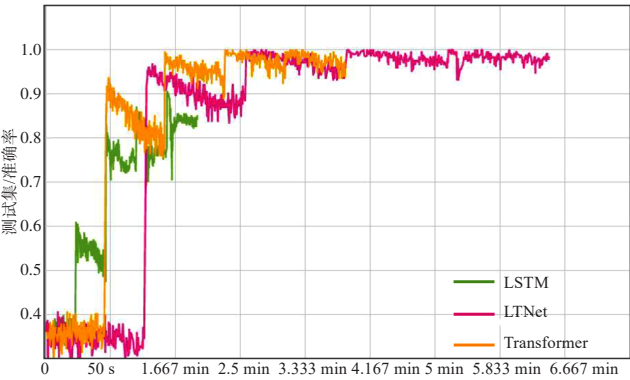


图6 模型在测试集上的准确率  
Figure 6 Accuracy of the model on the test set

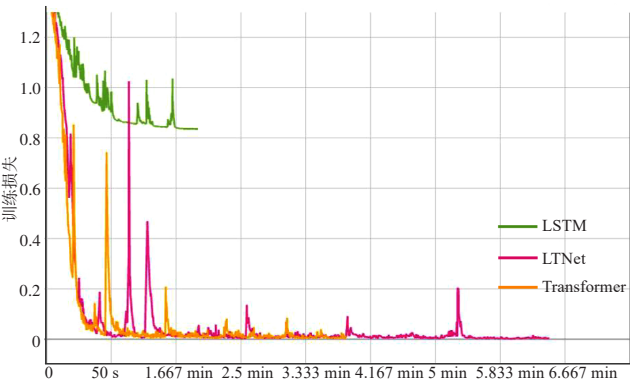


图7 模型在训练过程中的损失演变  
Figure 7 Evolution of model loss during training

对比传统的LSTM、Transformer以及提出的LTNet模型在交叉验证中的平均准确率、精确率、召回率、F1分数值和损失值。结果显示,尽管LSTM模型已经取得一定的进步,但在性能和效率方面仍有提升空间。相比之下,Transformer模型和LTNet模型展示了更为优异的性能,特别是LTNet模型在平均训练准确率上达到96.56%,平均测试损失降至0.08,这凸显了LTNet模型在处理脑电情感识别任务上的显著优势。LTNet模型在平均精确率、召回率以及F1分数方面均表现优异,超过上述两种模型,这表明该模型在综合性能上具有显著的优势。本文还与同样采用DEAP数据集的情绪识别实验准确率进行比较,



表3 训练集与测试集上的实验结果

Table 3 Experimental results on the training set and the test set

方法	训练集		测试集			
	平均准确率/%	平均准确率/%	平均精确率/%	平均召回率/%	平均F1分数值	平均损失值
LSTM	73.67	65.72	68.09	65.72	0.64	0.96
Transformer	95.16	82.22	82.59	82.22	0.82	0.12
LTNet	96.56	83.68	83.87	83.68	0.83	0.08

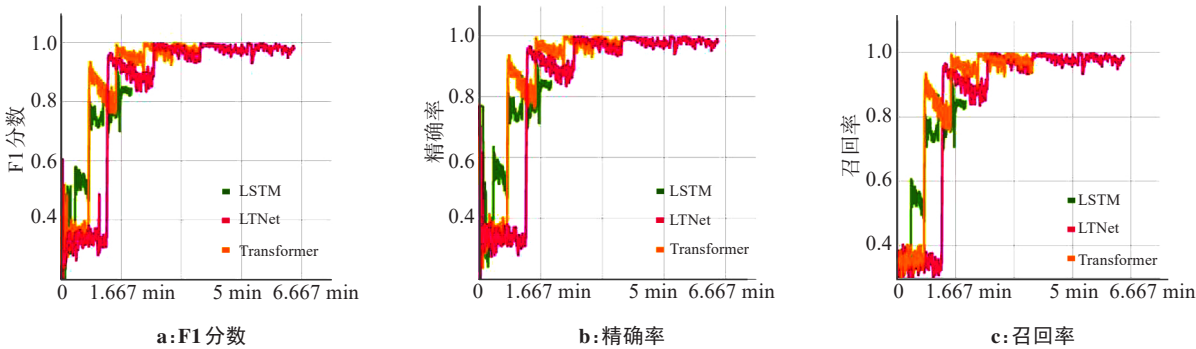


图8 测试集中模型平均精确率、召回率、F1分数值的对比

Figure 8 Comparison of average precision, recall rate and F1-score of the models on the test set

对比结果如表4所示,LTNet模型准确率相较于同类研究的脑电情绪识别实验准确率高2.74%~21.31%。

表4 与同类研究结果对比

Table 4 Comparison with similar research results

对比研究	方法	准确率/%	分类数
文献[2]	GCNN+LSTM	85.27	2
文献[3]	CNN+LSTM	92.48	4
文献[4]	Bi-LSTM	77.28	2
文献[6]	STS-Transformer	88.20	2
文献[21]	GLFANet	92.92	2
文献[22]	SPD+SVM	86.71	4
文献[23]	DA-CapsNet	75.25	2
文献[24]	MS-FTSCNN	93.82	2
文献[25]	mRMR+SVM	76.82	2
本文方法	LSTM+Transformer	96.56	4

5 结 语

本文提出一种结合 LSTM 和 Transformer 的脑电情感分析模型,名为 LTNet。该模型同时考虑脑电信号中的局部特征与全局特征,充分利用模型技术提取更多有效判别特征。对于情绪识别模型,不仅需

要考虑序列信号中的连续数据点之间的依赖关系,还需要关注输入序列的不同部分。实验证明,相比于单独使用 LSTM 或 Transformer 模型,LTNet 模型在情绪识别的平均准确率、精确率、召回率和 F1 分数等性能方面都表现更优,尤其是在准确率方面得到显著的 提升,相较于独立使用的 LSTM 模型和 Transformer 模型,LTNet 模型的准确率提高 1.40%~22.89%。此外,与其他采用 DEAP 数据集进行脑电情感识别的相关研究相比,准确率提高 2.74%~21.31%,这进一步证明 LTNet 模型在脑电情感识别任务中具有良好的识别性能和鲁棒性。这些结果证明 LTNet 模型作为一种有效工具,在理解和分析人类情感方面具有潜在的应用价值。

【参考文献】

[1] 吕宝粮,张亚倩,郑伟龙.情感脑机接口研究综述[J].智能科学与技术学报,2021,3(1):34-48.  
Lü BL, Zhang YQ, Zheng WL. A survey of affective brain-computer interface[J]. Chinese Journal of Intelligent Science and Technology, 2021, 3(1): 34-48.

[2] Yin YQ, Zheng XW, Hu B, et al. EEG emotion recognition using fusion model of graph convolutional neural networks and LSTM [J]. Appl Soft Comput, 2021, 100: 106954.

[3] 张英杰,谢云.基于CNN-LSTM的脑电情感四分类研究[J].科学技术与工程,2023,23(24):10437-10444.  
Zhang YJ, Xie Y. Four classification of EEG emotion based on CNN-LSTM[J]. Science Technology and Engineering, 2023, 23(24): 10437-10444.

[4] 刘帅,王磊,丁旭涛.基于Bi-LSTM的脑电情绪识别[J].山东大学学报(工学版),2020,50(4):35-39.

- Liu S, Wang L, Ding XT. Emotional EEG recognition based on Bi-LSTM[J]. Journal of Shandong University (Engineering Science), 2020, 50(4): 35-39.
- [5] 陈宗楠, 金家瑞, 潘家辉. 基于Swin Transformer的四维脑电情绪识别[J]. 计算机技术与发展, 2023, 33(12): 178-184.  
Chen ZN, Jin JR, Pan JH. Swin Transformer-based 4-D EEG emotion recognition[J]. Computer Technology and Development, 2023, 33(12): 178-184.
- [6] Zheng W, Pan B. A spatiotemporal symmetrical transformer structure for EEG emotion recognition[J]. Biomed Signal Process Control, 2024, 87(Part B): 105487.
- [7] Koelstra S, Muhl C, Soleymani M, et al. DEAP: a database for emotion analysis; using physiological signals[J]. IEEE Trans Affect Comput, 2012, 3(1): 18-31.
- [8] 李锦瑶, 杜肖兵, 朱志亮, 等. 脑电情绪识别的深度学习研究综述[J]. 软件学报, 2023, 34(1): 255-276.  
Li JY, Du XB, Zhu ZL, et al. Deep learning for EEG-based emotion recognition: a survey[J]. Journal of Software, 2023, 34(1): 255-276.
- [9] Ahmed MZ, Sinha N, Ghaderpour E, et al. A novel baseline removal paradigm for subject-independent features in emotion classification using EEG[J]. Bioengineering (Basel), 2023, 10(1): 54.
- [10] 许超凡, 肖文栋, 曹征涛. 基于一维卷积特征与手工特征融合的集成超限学习机心跳分类方法[J]. 工程科学学报, 2021, 43(9): 1224-1232.  
Xu YF, Xiao WD, Cao ZT. Ensemble extreme learning machine approach for heartbeat classification by fusing 1D convolutional and handcrafted features[J]. Chinese Journal of Engineering, 2021, 43(9): 1224-1232.
- [11] 阚威, 李云. 基于LSTM的脑电情绪识别模型[J]. 南京大学学报: 自然科学, 2019, 55(1): 110-116.  
Kan W, Li Y. Emotion recognition from EEG signals by using LSTM recurrent neural networks[J]. Journal of Nanjing University (Natural Science), 2019, 55(1): 110-116.
- [12] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[C]// Proceedings of the 31st International Conference on Neural Information Processing Systems. Red Hook, NY, USA: Curran Associates Inc., 2017: 6000-6010.
- [13] Shi XB, Li BJ, Wang WL, et al. Classification algorithm for electroencephalogram-based motor imagery using hybrid neural network with spatio-temporal convolution and multi-head attention mechanism[J]. Neuroscience, 2023, 527: 64-73.
- [14] Shen X, Han DZ, Guo ZH, et al. Local self-attention in transformer for visual question answering[J]. Appl Intell, 2023, 53(13): 16706-16723.
- [15] Meng FF, Wang CA. Sentiment analysis with adaptive multi-head attention in transformer[EB/OL]. (2024-03-11). <https://arxiv.org/abs/2310.14505>.
- [16] Ebert N, Stricker D, Wasenmüller O. PLG-ViT: vision transformer with parallel local and global self-attention[J]. Sensors (Basel), 2023, 23(7): 3447.
- [17] Andayani F, Theng LB, Tsun MT, et al. Hybrid LSTM-transformer model for emotion recognition from speech audio files[J]. IEEE Access, 2022, 10: 36018-36027.
- [18] Li WS, Peng Y, Zhang M, et al. Deep model fusion: a survey[EB/OL]. (2023-09-27). <https://arxiv.org/abs/2309.15698>.
- [19] Alhagry S, Fahmy AA, El-Khoribi RA. Emotion recognition based on EEG using LSTM recurrent neural network[J]. Int J Adv Comput Sci Appl, 2017, 8(10): 355-358.
- [20] Wang Z, Wang YX, Hu CF, et al. Transformers for EEG-based emotion recognition: a hierarchical spatial information learning model[J]. IEEE Sens J, 2022, 22(5): 4359-4368.
- [21] Liu SQ, Zhao YY, An YL, et al. GLFANet: a global to local feature aggregation network for EEG emotion recognition[J]. Biomed Signal Process Control, 2023, 85: 104799.
- [22] Gao YY, Sun XY, Meng M, et al. EEG emotion recognition based on enhanced SPD matrix and manifold dimensionality reduction[J]. Comput Biol Med, 2022, 146: 105606.
- [23] Liu SQ, Wang ZY, An YL, et al. DA-CapsNet: a multi-branch capsule network based on adversarial domain adaption for cross-subject EEG emotion recognition[J]. Knowl Based Syst, 2024, 283: 111137.
- [24] Li FF, Hao KR, Wei B, et al. MS-FTSCNN: an EEG emotion recognition method from the combination of multi-domain features[J]. Biomed Signal Process Control, 2024, 88(Part A): 105690.
- [25] 胡梓煊, 张光旭, 张圣杰, 等. 基于mRMR特征优选与SVM的脑电信号的情绪识别[J]. 自动化应用, 2023, 64(21): 8-11.  
Hu ZX, Zhang GX, Zhang SJ, et al. Emotion recognition of EEG signals based on mRMR feature optimization and SVM[J]. Automation Application, 2023, 64(21): 8-11.

(编辑:陈丽霞)