



DOI:10.12404/j.issn.1671-1815.2403775

引用格式: 巩炫麟, 陶庆, 苏娜, 等. 基于 CNN-CBAM-LSTM 的稳态视觉诱发电位脑电信号识别方法[J]. 科学技术与工程, 2025, 25(10): 4175-4182.

Gong Xuanlin, Tao Qing, Su Na, et al. A recognition method for steady-state visual evoked potential EEG signals based on CNN-CBAM-LSTM[J]. Science Technology and Engineering, 2025, 25(10): 4175-4182.

电子技术、通信技术

# 基于 CNN-CBAM-LSTM 的稳态视觉诱发电位脑电信号识别方法

巩炫麟<sup>1</sup>, 陶庆<sup>2\*</sup>, 苏娜<sup>3</sup>, 马金旭<sup>1</sup>

(1. 新疆大学智能制造现代产业学院(机械工程学院), 乌鲁木齐 830017;

2. 新疆大学工程师学院, 乌鲁木齐 830017; 3. 新疆医科大学第一附属医院, 乌鲁木齐 830054)

**摘要** 在使用传统方法处理稳态视觉诱发电位(steady-state visual evoked potentials, SSVEP)的脑电信号时,特征提取的准确性和充分性存在不足,影响信号的识别准确率。为此提出了一种基于卷积神经网络(convolutional neural networks, CNN)与卷积注意力机制模块(convolutional block attention module, CBAM)和长短时记忆网络(long short-term memory, LSTM)的信号分类识别方法。以 CNN 为基础框架,通过引入注意力机制对通道及空间特征进行充分提取,加入 LSTM 提高对时序特征的提取能力,实现对 SSVEP 信号的目标识别。实验结果显示,所提方法能够充分有效的提取各级特征且识别准确率较高,相比于典型相关分析方法(canonical correlation analysis, CCA)、CNN、CBAM-LSTM、CNN-CBAM 识别准确率分别提高了 5.3%、2.95%、2.27%、1.71%,可见该模型对 SSVEP 信号的分类识别有较好的效果。

**关键词** 稳态视觉诱发电位; 卷积神经网络; 卷积注意力机制模块; 长短时记忆网络; 目标识别

中图分类号 TN911.7; 文献标志码 A

## A Recognition Method for Steady-State Visual Evoked Potential EEG Signals Based on CNN-CBAM-LSTM

GONG Xuan-lin<sup>1</sup>, TAO Qing<sup>2\*</sup>, SU Na<sup>3</sup>, MA Jin-xu<sup>1</sup>

(1. School of Intelligent Manufacturing Modern Industry (School of Mechanical Engineering), Xinjiang University, Urumqi 830017, China;

2. College of Engineering, Xinjiang University, Urumqi 830017, China;

3. The First Affiliated Hospital of Xinjiang Medical University, Urumqi 830054, China)

**[Abstract]** When traditional methods were used to evoked the potentials SSVEP (steady-state visual evoked potentials) EEG (electroencephalogram) signals, the accuracy and sufficiency of feature extraction were insufficient, which affected the recognition accuracy of signals. A novel approach was proposed which based on a CNN (convolutional neural network) integrated with a CBAM (convolutional block attention module) and a LSTM (long short-term memory network). By incorporating attention mechanisms, both channel and spatial features were effectively extracted within the CNN framework. Additionally, LSTM was introduced to enhance the extraction of temporal features, enabling accurate recognition of SSVEP signals. The experimental results show that the proposed method can effectively extract hierarchical features and achieves a high recognition accuracy. Compared to canonical correlation analysis (CCA), CNN, CBAM-LSTM, and CNN-CBAM, the proposed model improves the recognition accuracy by 5.3%, 2.95%, 2.27%, and 1.71% respectively. It can be seen that the model has a good performance in the classification and recognition of SSVEP signals.

**[Keywords]** SSVEP (steady state visual evoked potential); CNN (convolutional neural networks); convolutional attention mechanism module; LSTM (long short-term memory network); target recognition

脑机接口(brain computer interface, BCI)是一种直接将人脑活动与外部设备进行交互的技术。

它通过采集和解读大脑信号,将人的意图转化为命令或控制信号,从而实现与计算机或其他外部设备

收稿日期: 2024-05-21; 修订日期: 2025-01-13

基金项目: 新疆维吾尔自治区“天山英才”科技创新领军人才项目(2023TSYCLJ0051)

第一作者: 巩炫麟(1999—),女,汉族,新疆博乐人,硕士。研究方向:脑电信号。E-mail:905835582@qq.com。

\*通信作者: 陶庆(1978—),男,汉族,江苏常熟人,博士,教授。研究方向:先进设计与制造。E-mail:xjutao@qq.com。

的交互<sup>[1]</sup>。如今 BCI 技术已经在多个领域如医疗康复、智能感知等方面得到广泛的应用<sup>[2-4]</sup>。稳态视觉诱发电位 (steady-state visual evoked potentials, SSVEP) 是指在视觉刺激频率稳定的情况下, 大脑皮层产生的电生理响应。作为一种可靠、易于获取且具有较高识别准确性的脑电信号类型, 已经在很多领域得到了应用<sup>[5-6]</sup>。但由于其信号微弱, 容易受肌电和自发脑电等伪迹的干扰, 识别往往较为困难<sup>[7]</sup>。在 SSVEP-BCI 系统中, 如何提高目标识别准确率一直是关键。在算法研究方面, 如何对信号进行有效的特征提取、分类与解码识别是 SSVEP 信号分类研究的重点, 也是后续在线实验过程中提高控制准确率和降低控制延迟率的关键。

在传统 SSVEP 信号目标识别方法中, 通常由人工提取特征后再进行目标识别<sup>[8]</sup>。Wu 等<sup>[9]</sup>为提高信号识别准确率, 使用小波包分解 (wavelet packet decomposition, WPD) 提取信号中的时频特征, 但结果易受小波基函数选择的影响。Chen 等<sup>[10]</sup>采用快速傅里叶变换和叠加平均, 结合典型相关分析 (canonical correlation analysis, CCA) 方法提取相关的频率信息进行分析。但该方法无法提高高频率刺激下的识别精度, 不具有普适性。Antelisjm 等<sup>[11]</sup>通过测量信号中基频的功率谱峰值的突出度和宽度的比值, 来计算信号中存在的谐波 (harmonic) 成分, 再结合经验模式分解对 SSVEP 信号进行分类。然而, 信号分解后得到的模式之间不是完全正交, 影响最终的准确率。与传统方法相比, 深度学习规避了人工选择特征效率的有限性、适应性弱缺点, 从而更加有效地提取信号中的特征, 并提升目标识别性能。尽管深度学习在图像识别<sup>[12]</sup>、语音识别等领域取得了不错进展, 然而在脑机接口的应用中, 采用深度学习方法进行研究的数量相对较少。Tabra 等<sup>[13]</sup>在对运动想象的脑电信号时频图识别时, 采用的卷积神经网络, 验证分析发现该方法表现出更好的分类性能; An 等<sup>[14]</sup>在左右手运动想象分类研究中, 使用深度置信网络 (deep belief networks, DBN) 进行实验, 结果显示, 相比于传统方法如支持向量机 (support vector machine, SVM), DBN 方法具有明显更高的分类精度。

综上所述, 相比于传统的机器学习, 深度学习在处理信号时可获取更丰富的信息与更高级别的特征表达。在深度学习中, 使用 CNN 处理 SSVEP 信号在表征与学习方面取得了较好成绩。但因为 SSVEP 信号是随时间变化的, 若单一使用 CNN 来处理信号无法很好地捕捉全局特征, 从而影响识别率。

因此, 为解决上述信号处理过程中特征表征问

题, 现提出基于 CNN 与 CBAM、LSTM 的 SSVEP 信号分类识别方法。其中, LSTM 可以更好地捕捉 SSVEP 信号中的时序特征, 有助于提高信号处理的准确性和可靠性。CBAM 可对特征进行更加细化的提取, 从而有效地提升模型性能。将改进模型与现有方法对比, 期望获得较好的目标识别效果。

## 1 稳态视觉诱发电位实验

### 1.1 实验范式设计

本次实验采用基于场景动画的稳态视觉刺激诱发范式<sup>[15]</sup>。该范式采用方波调制模式对黑白反转色图片进行视觉刺激, 图片为独立的现实生活情景, 且在使用前进行灰度标准化处理。实验时屏幕背景为深灰色, 屏幕分辨率为  $1\ 920 \times 1\ 080$  像素, 刷新率为 60 Hz。刺激范式的界面设计和呈现通过 MATLAB 的 Psychophysics Toolbox 工具箱完成, 具体呈现如图 1 所示。屏幕共呈现 4 张图片, 按上下左右顺序诱发频率分别为 8、9.24、10.9、12 Hz, 分别对应手张开、手抓握、拿书、抓水瓶 4 个动作。

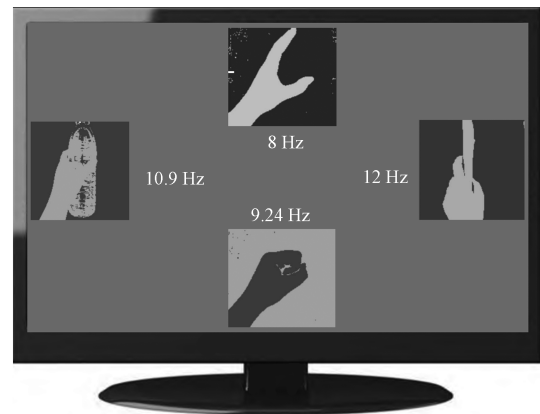


图 1 场景动画诱发范式刺激图

Fig. 1 Scene animation evoked potential paradigm stimulus image

### 1.2 实验流程

本次实验选择 5 名视力正常的健康受试者作为研究对象, 且均没有参与过类似实验。年龄在 22 ~ 26 岁, 包括 4 名男性和 1 名女性。实验在安静的房间内进行, 要求受试者距离显示器为 70 ~ 80 cm 并保持舒适的坐姿。受试者需要将注意力集中在特定目标上, 以充分利用视觉皮层中央放大效应, 同时在实验过程中尽量减少肢体和眼部运动, 以避免对脑电信号的干扰。

实验开始, 受试者有 2 s 准备时间, PC 机提示音响起后, 屏幕刺激块开始闪烁。受试者按照顺序依次注视上、下、左、右的刺激目标, 对应标签 0、1、2、3。其中 0 代表 8Hz 刺激下的信号, 1 代表

9.24 Hz 刺激下的信号, 2 代表 10.9 Hz 刺激下的信号, 3 代表 12 Hz 刺激下的信号。每组实验需轮流注视 4 个刺激块, 每个刺激块的刺激时长均为 4 s, 脑电信号数据正式采集前后会设有 2 s 的休息时间。实验对每名受试者采集 3 轮数据, 每轮数据采集 10 组脑电信号。同时为避免实验时间过长引起受试者时间疲劳, 每组实验之间会短休息 1 min, 每轮之间会休息 5 min。单次采集的数据大小为  $1\ 024 \times 6 \times 4$ , 其中 1 024 为 256 Hz 下 4 s 采样点数, 6 为通道数量, 4 为分类数。每轮实验的流程安排如图 2 所示, 实验场景如图 3 所示。

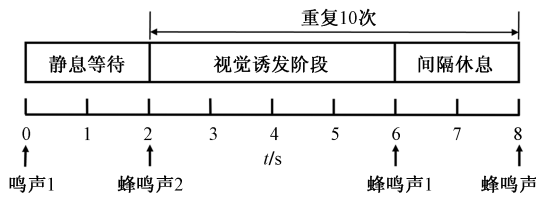


图 2 实验时序图  
Fig. 2 Experimental sequence diagram



图 3 实验场景  
Fig. 3 Experimental scene

### 1.3 信号采集与预处理

实验采用 G. tec 公司生产的 G. USBamp 低通道有线生物信号采集系统, 共采集 6 个通道信号。根据国际标准 10/20 安放法, 将电极放置在枕叶区的 Po3、Poz、Po4、O1、Oz、O2, 参考电极为 Cpz 并夹在耳乳突处, 接地电极为 Fpz 处, 具体位置如图 4 所示。采样频率为 256 Hz, 实验时需保持静息状态下阿尔法波的幅值在  $\pm 50 \mu\text{V}$  以内。

为提取出所需信息并减少噪声和干扰, 对原始数据进行预处理。同时为提高 SSVEP 信号的信噪比, 对信号进行 5 ~ 45 Hz 的带通滤波与工频陷波处理, 重参考使用平均电位。去除信号采集开始后前 24 个质量较差的采集点, 并根据标签将数据分为 4 类。随后, 将数据以通道数、采样数为结构的 2 维数据矩阵形式储存, 最后使用重叠窗对信号数据段

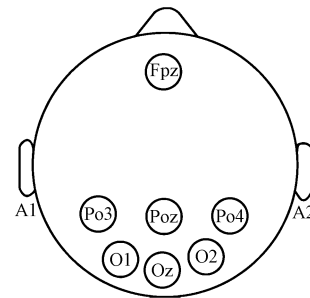


图 4 采集通道位置  
Fig. 4 Acquisition channel location

进行分割, 滑动步长为 128, 时间窗长为 256, 每个分类可得 900 条  $6 \times 256$  的 SSVEP 数据信号。

## 2 SSVEP 信号识别方法

### 2.1 卷积神经网络

CNN 的网络结构通常由卷积层、激活函数、池化层、全连接层与批归一化层构成<sup>[16]</sup>。在训练过程中, CNN 计算模型参数梯度方向的损失函数, 并按照梯度的反方向更新参数, 逐步降低损失函数的值并对网络权重进行反向调节, 不断调整网络参数来提高精度。但现有基于 CNN 对图像识别方法不适于处理 SSVEP 这种时序信号。

而一维卷积神经网络 (one-dimensional convolutional neural network, 1DCNN) 常被用于识别与提取时间序列的特征, 且与 CNN 相似同样具备平移不变性等优点, 可有效提取 SSVEP 信号中的关键特征<sup>[17-18]</sup>。因此, 本文研究选择 1DCNN 作为处理 SSVEP 信号的模型。一维卷积神经网络包括 3 个卷积层、池化层、批量归一化层 (batch normalization, BN) 以及 1 个全连接层和 1 个随机失活正则化层 (Dropout), 模型结构如图 5 所示。

采集的 SSVEP 原始信号经预处理与重叠窗口后的信号矩阵大小为  $6 \times 256$ 。信号输入后对其时

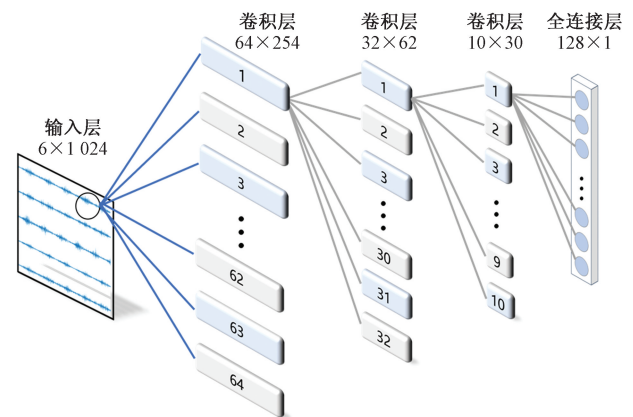


图 5 1DCNN 模型结构  
Fig. 5 1DCNN model structure

域进行一维卷积计算,计算公式为

$$h_j^l = \sum_{i=1}^n x_i^l \otimes k_{ij}^l + b_j^l \quad (1)$$

式(1)中: $x$ 为输入特征向量; $h$ 为输出特征向量; $l$ 为层数; $i$ 为卷积序数; $j$ 为元素序数; $n$ 为 $x$ 的长度; $b$ 为偏置向量; $k$ 为卷积核; $\otimes$ 代表矩阵乘法。

在1DCNN中,卷积核可以看作是感受野,为减少计算量,同一卷积核在与不同局部数据进行运算时共享一组权重。计算公式为

$$\text{Cov}(x, y) = \sum_{a=0}^w F(a) \times G(x - a) \quad (2)$$

式(2)中: $w$ 为卷积核的尺寸; $a$ 为卷积核在输入宽度方向上的大小; $\text{Cov}$ 为卷积运算的输出序列; $F(a)$ 为卷积核参数向量; $G(x - a)$ 为与卷积核函数运算的局部向量矩阵。

经过卷积后对输出的特征图进行池化操作,可以减少特征图的尺寸从而降低计算量,并在一定程度上防止过拟合,提高模型的泛化能力。计算公式为

$$P = \max_w \{A^L\} \quad (3)$$

式(3)中: $A$ 为经过激活后的特征矩阵; $P$ 为最大池化后的特征; $w$ 为池化区域的宽度; $L$ 为特征图集合的宽度。

在每个卷积层后使用BN层防止模型过拟合,在BN层后,采用ReLU激活函数。它是一个非线性的激活函数,其优势在于可以将负值映射为0,从而使一部分神经元的输出为0,从而提高模型的表达能力、减少冗余信息、缓解梯度消失问题。计算公式为

$$\text{ReLU}(x) = \max(0, x) = \begin{cases} x, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (4)$$

在训练过程中,引入损失函数来度量模型性能并进行优化。本文选择深度学习中常用的交叉熵损失函数。计算公式为

$$L(\theta) = -\frac{1}{n} \left[ \sum_{i=1}^n \sum_{k=1}^K I\{y_i = k\} \lg \frac{\exp(\theta_k^T x)}{\sum_{j=1}^K \exp(\theta_j^T x)} \right] \quad (5)$$

式(5)中: $I\{y_i = k\}$ 为指示函数,当样本 $i$ 的真实类别 $y_i$ 等于类别 $k$ 时,取值为1,否则取值为0; $K$ 为类别数; $\theta$ 为模型的参数。

使用全连接层将上一层的输出展平为一个向量作为输入,再与其权重矩阵相乘。在全连接层后加入Dropout层,防止模型过拟合。使用1DCNN可有效地提取时序特征,但是1DCNN善于处理局部信息的任务,无法很好地捕捉SSVEP信号的全局特征,从而影响模型的性能。网络模型配置如表1所示。

表1 1DCNN模型参数配置

网络层	卷积核尺寸	卷积核数目	输出维度
卷积层1	3 × 1	64	64 × 254
BN1 + ReLU	64	—	64 × 254
池化层1	2 × 2	64	64 × 127
卷积层2	3 × 1	32	32 × 125
BN1 + ReLU	32	—	32 × 125
池化层2	2 × 2	32	32 × 62
卷积层3	3 × 1	10	10 × 60
BN1 + ReLU	10	—	10 × 60
池化层3	2 × 2	10	10 × 30
全连接层	—	—	128 × 1

## 2.2 CBAM

轻量级的卷积注意力模块CBAM是前馈卷积神经网络的一种注意力机制。CBAM的重要功能是过滤不相关信息、解决信息过载问题、提高任务处理的准确性。CBAM由CAM(channel attention module)和SAM(spatial attention module)两个子模块构成<sup>[19-20]</sup>,如图6所示。将卷积后得到的特征图作为CBAM模块的输入,CAM用于对特征图的通道维度进行注意力计算,生成通道注意力图。而SAM则用于对通道特征图的空间维度进行注意力计算,生成空间注意力图。通过这些操作,可以从两个维度上增强感兴趣区域,并得到推断的注意力图。然后,将通道注意力图与输入特征图进行逐元素相乘,得到的输出再与空间注意力图进行类似的逐元素相乘,实现特征图的细化,使特征处理过程更加智能和适应性更强。CBAM的网络结构如图6所示。

CAM利用通道间的关联性来确定哪些通道对于当前任务更为重要,从而提高特征的表达能力。计算如式(6)所示。主要包括全局平均池化模块、

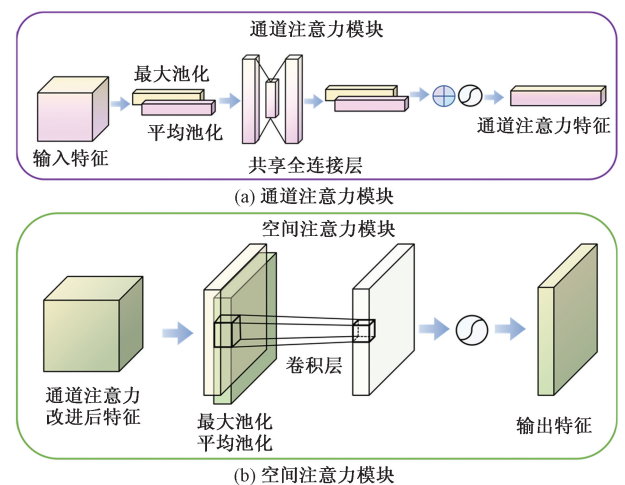


图6 CBAM网络结构

Fig. 6 CBAM network structure

全局最大池化模块和共享多层感知器 (multi-layer perceptron, MLP) 模块。共享 MLP 的输出通过逐元素求和进行融合, 然后通过 sigmoid 激活得到通道注意力图。

$$M_c(F) = \sigma \{ \text{MLP} [ \text{AvgPool}(F) ] + \text{MLP} \times [ \text{MaxPool}(F) ] \} = \sigma \{ W_1 [ W_0 (F_{\text{avg}}^c) ] + W_1 [ W_0 (F_{\text{max}}^c) ] \} \quad (6)$$

式(6)中:  $M_c$  为输出的通道注意力特征; MLP 为共享网络中的多层感知器;  $\sigma$  为 sigmoid 函数;  $F_{\text{max}}^c$  为最大池化特征;  $F_{\text{avg}}^c$  为平均池化特征; AvgPool 为平均池化; MaxPool 为最大池化。

SAM 可以自适应地调整空间区域的权重, 使得网络能够更加关注对当前任务有贡献的空间位置, 从而提升模型的性能。计算式为

$$M_s(F_1) = \sigma \{ f^{7 \times 7} \{ [ \text{AvgPool}(F_1); \text{MaxPool}(F_1) ] \} \} = \sigma \{ f^{7 \times 7} ([ F_{1\text{avg}}^s; F_{1\text{max}}^s ]) \} \quad (7)$$

式(7)中:  $F_{1\text{avg}}^s$  为经过平均池化生成的 2D 特征图,  $F_{1\text{avg}}^s \in R^{1 \times H \times W}$ ;  $F_{1\text{max}}^s$  为经过最大池化生成的 2D 特征图,  $F_{1\text{max}}^s \in R^{1 \times H \times W}$ ;  $f^{7 \times 7}$  为滤波器大小为  $7 \times 7$  的卷积运算。

CBAM 是一个轻量级的通用网络模块, 可较为容易地与 CNN 结合使用。它通过应用注意力机制来引导对池化层输出的特征图进行处理, 从而有效地提取局部特征。但是 CNN 与 CBAM 主要用于静态数据的处理, 无法充分捕捉脑电信号中的时间动态信息和长期依赖关系, 不利于 SSVEP 信号的表征。

### 2.3 LSTM

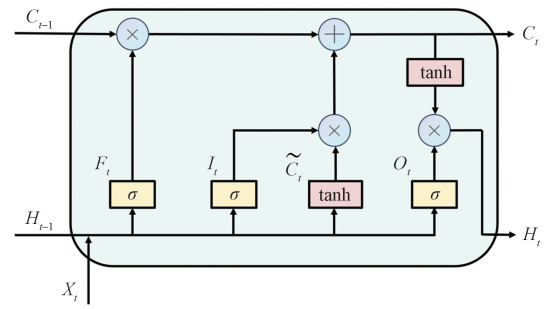
LSTM 是一种改进的循环神经网络结构, 通过引入门控机制来解决传统循环神经网络中的长期依赖问题, 从而在处理长序列时取得更好的效果<sup>[21]</sup>。输入、输出和遗忘门以及记忆单元构成了 LSTM 的大部分。这三个门都有 Sigmoid 激活机制。遗忘门调节先前记录的历史知识是否保持, 而输入、输出门分别控制神经元在任何给定时间的输入和输出信息。设置隐藏层大小为 128, 学习率 0.001。其结构如图 7 所示。

遗忘门用于控制前一时刻的单元状态中可以保留在当前时刻的信息, 其计算公式为

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad (8)$$

式(8)中:  $W_f$  为遗忘门的权重矩阵;  $h_{t-1}$  为上一时刻的输出;  $x_t$  为当前时刻的输入;  $b_f$  为遗忘门的偏置值。

输入门负责确定当前网络输入的部分是否存储到单元状态中, 其计算公式为



$C_t$  为  $0 \sim t$  时间内的所提取的关键信息;  $H_t$  为  $t$  时刻长短时记忆单元的输出;  $F_t$  为 LSTM 单元的遗忘门;  $I_t$  为输入门;  $O_t$  为输出门

图 7 LSTM 单元结构

Fig. 7 LSTM unit structure

$$i_t = \sigma(w_i[h_{t-1}, x_t] + b_i) \quad (9)$$

$$\hat{C}_t = \tanh(W_c[h_{t-1}, x_t] + b_c) \quad (10)$$

$$C_t = f_t C_{t-1} + i_t \hat{C}_t \quad (11)$$

式(11)中:  $\tanh$  为激活函数;  $b_i$  为输入门的偏置值;  $w_i$  为输入门的权重矩阵。

输出门控制着当前时刻的输出, 将上一时刻的状态和当前的输入结合起来, 生成当前时刻的输出, 其计算公式为

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \quad (12)$$

$$h_t = o_t \tanh(C_t) \quad (13)$$

式(12)中:  $W_o$  为输出门的权重矩阵;  $b_o$  为输出门的偏置值。

作为适用于序列数据的循环神经网络, LSTM 能够有效地捕捉时间序列中的长期依赖, 从而提供强大的建模能力。本文研究通过将 LSTM 与 CNN、CBAM 结合, 可以在处理具有时序性质的数据时, 更好地捕捉序列中的长期相关性, 从而提高模型的预测准确性。

### 2.4 CNN-CBAM-LSTM 模型

结合上述分析, 本文提出 CNN-CBAM-LSTM 模型, 该模型由三部分组成: CNN 特征提取层、CBAM 层与 LSTM 层。整体结构如图 8 所示。原始信号样本大小为  $6 \times 1024$ , 经过预处理与重叠窗口操作后得到  $6 \times 256$  的序列输入 1DCNN 模型。模型包含三个卷积层, 可有效学习脑电数据中的高维空间特征, 卷积核大小均为  $3 \times 1$ 。卷积核数目分别为 64、32、10, 经每层卷积后的池化操作输出的特征图尺寸分别为  $64 \times 127$ 、 $32 \times 62$ 、 $10 \times 30$ 。CBAM 模块用在前两次卷积后, 将卷积池化后得到的特征图输入 CBAM 中。CBAM 从通道和空间两个维度对网络的重要性进行加权, 并对输入特征进行自适应特征细化, 细化后的特征通过多个卷积池化层进行处理, 继续进行特征提取和数据降维。特征图尺寸在经过 CBAM 模块处理前后保持不变。经过第三次卷

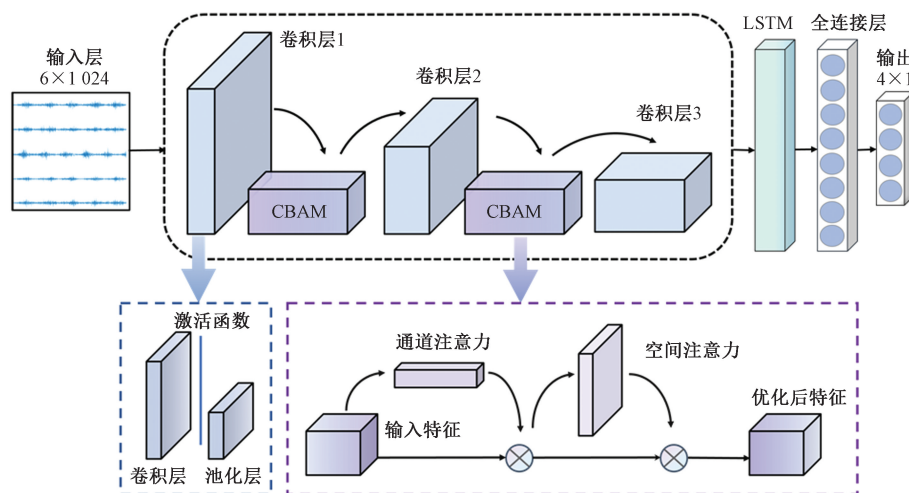


图8 CNN-CBAM-LSTM 模型总体构架

Fig. 8 Overall architecture of CNN-CBAM-LSTM model

积与池化后输出的特征图尺寸为  $10 \times 30$ , 将特征图展平成长度为 300 的一维向量输入 LSTM 中, 隐藏层大小设置为 128。LSTM 的门控机制使模型能够选择性遗忘、输入和输出信息, 从而更好地处理长序列数据。LSTM 输出长度为 128 的向量序列, 将序列转化为单个向量输入全连接层, 使用 softmax 激活函数将输出转化为概率分布, 完成对 SSVEP 信号四分类的识别。该模型实现对空间和时间信息的综合利用, 吸收了更丰富的 SSVEP 信号特征信息, 加强捕捉了序列的长期相关性, 为识别 SSVEP-BCI 任务提供了一种新的方法。

### 3 实验结果分析

#### 3.1 识别评价指标

为了评估模型的性能, 采用准确率 ( $A_{accuracy}$ )、精确率 ( $P_{recision}$ )、召回率 ( $R_{ecall}$ )、 $F_1$  ( $F_{1-score}$ ) 4 个指标来评估模型。计算公式如式 (14) ~ 式 (17) 所示。并将采集的实验数据按 7:3 分为训练数据集和测试数据集。

$$A_{accuracy} = \frac{TP + TN}{TP + TN + FN + FP} \quad (14)$$

$$P_{recision} = \frac{TP}{TP + FP} \quad (15)$$

$$R_{ecall} = \frac{TP}{TP + FN} \quad (16)$$

$$F_{1-score} = \frac{2TP}{2TP + FP + FN} \quad (17)$$

#### 3.2 信号识别效果分析

使用 CNN-CBAM-LSTM 模型训练后得到损失函数拟合曲线, 如图 9 所示。模型共经过 200 个训练周期, 在 60 个训练周期迭代后, 训练损失逐渐减小到 0.5 以下, 随着训练周期的增加, 损失在 140 个周期后

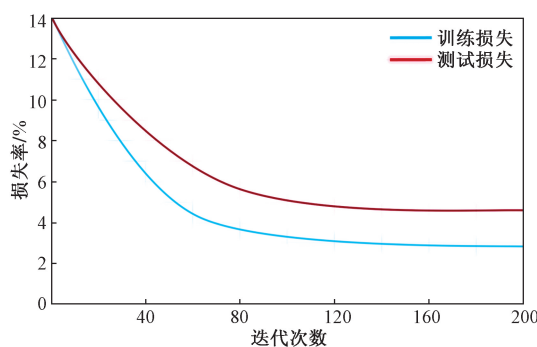


图9 CNN-CBAM-LSTM 模型损失曲线

Fig. 9 Loss curve of CNN-CBAM-LSTM model

逐渐达到稳定, 测试损失达到最低点 4.52%。

为比较不同模型的性能, 使用同组数据集将所提出的 CNN-CBAM-LSTM 方法与 CCA、CNN、CNN-CBAM、CNN-LSTM 等方法进行比较。通过识别不同的目标, 得到 5 类目标识别的概率。再计算概率与标签矩阵, 得出相应的真阳性率和假阳性率, 将得到的数据取平均值, 最终绘制出 5 种模型对应的 ROC 曲线, 如图 10 所示。由图 10 可知, CNN-CBAM-LSTM 模型曲线下的面积最大, 模型性能最好。

混淆矩阵如图 11 所示, 为 CNN-CBAM-LSTM 模型在测试数据上的分类效果。矩阵对角线上的数字表示正确分类的样本数量, 而其他格中的数字表示错误的分类样本量。从图 12 中可以观察到, 该模型对第二分类效果相对较差, 在第一和第四分类上表现良好, 综合来看, 该模型能够有效地对 4 种信号进行分类。

5 种模型的评估指标结果如图 12 所示, 可以观察到: CNN-CBAM-LSTM 模型的识别准确率比 CNN-CBAM 高 1.71%, 比 CNN-LSTM 高 2.27%, 比

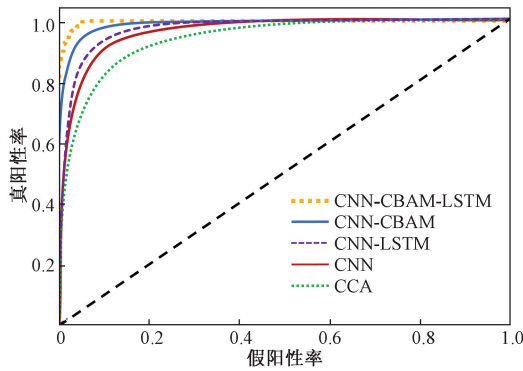


图 10 5 种模型的 ROC 曲线  
Fig. 10 ROC curves of 5 models

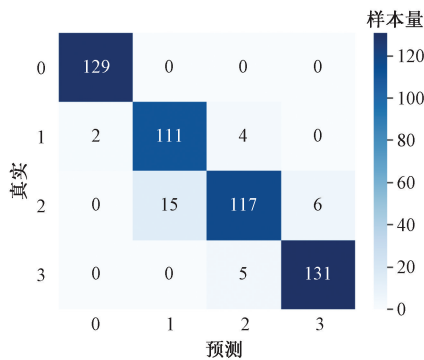


图 11 CNN-CBAM-LSTM 模型的混淆矩阵  
Fig. 11 Confusion matrix of CNN-CBAM-LSTM model

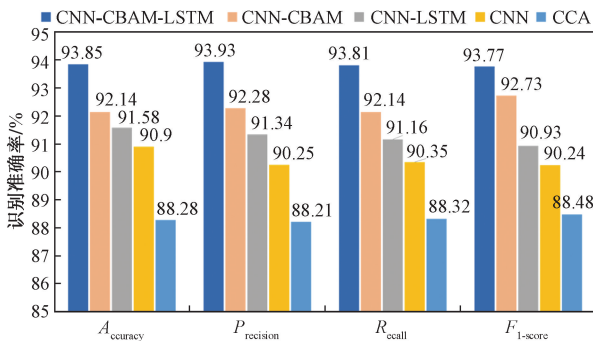


图 12 模型识别性能对比  
Fig. 12 Model recognition performance comparison

CNN 高 2.95%, 比 CCA 高 5.3%。在其余 3 种指标中, CNN-CBAM-LSTM 模型比 CNN-CBAM 模型高出 1.65%、1.67% 和 1.04%; 比 CNN-LSTM 模型高出 2.59%、2.65% 和 2.84%; 比 CNN 模型高出 3.68%、3.46% 和 3.53%; 比 CCA 模型高出 5.72%、5.49% 和 5.29%。通过对比可以发现, CNN-CBAM-LSTM 模型在 SSVEP 信号的识别任务中更为适用。

### 4 结论

提出了基于的 CNN 与 CBAM、LSTM 信号识别方法, 结合深度学习领域中使用较为广泛的注意力

机制与长短时记忆网络, 克服传统目标识别方法在处理 SSVEP 信号时的局限性, 有较高的识别准确率与自适应能力。通过在 CNN 中加入 CBAM 模块, 使 CNN 捕获更多 SSVEP 信号中的时空特征, 提高模型对于 SSVEP 信号中关键信息的识别和提取能力。引入 LSTM, 更好地捕捉了 SSVEP 信号中的时序特征, 提高了信号处理的准确性与可靠性, 实现 SSVEP 信号端到端的目标识别。最后在制作的数据集中进行验证, 实验结果表明, 相较于传统学习方法 CCA 准确率提升 5.84%, 对比基础的深度学习也有 3% 的优势。在四分类中平均准确率达到 92.5%、93.6%、90.1%、93.4%。与现有研究相比, 本文方法对于场景动画刺激范式识别有一定优势, 可进一步提高信号的识别准确率, 为脑电信号处理领域带来了新的思路和方法。使得智能设备在人机交互过程中可更加准确的识别目标, 增加用户体验有助于产品的发展。

### 参 考 文 献

- [1] Sebastián P Z, José L M L, Kelvin O C, et al. Review of steady state visually evoked potential brain-computer interface applications: technological analysis and classification[J]. Journal of Engineering and Applied Sciences, 2019, 15(2): 659-678.
- [2] 于淑月, 李想, 于功敬, 等. 脑机接口技术的发展与展望[J]. 计算机测量与控制, 2019, 27(10): 5-12.  
Yu Shuyue, Li Xiang, Yu Gongjin, et al. Development and prospect of brain computer interface technology [J]. Computer Measurement & Control, 2019, 27(10): 5-12.
- [3] Chamola V, Vineet A, Nayyar A, et al. Brain-computer interface-based humanoid control: a review[J]. Sensors, 2020, 20(13). DOI: 10.3390/s20133620.
- [4] 葛松, 徐晶晶, 赖舜男, 等. 脑机接口: 现状, 问题与展望[J]. 生物化学与生物物理进展, 2020, 47(12): 1227-1249.  
Ge Song, Xu Jingjing, Lai Shunnan, et al. Brain computer interface: current situation, problems and prospects[J]. Progress in Biochemistry and Biophysics, 2020, 47(12): 1227-1249.
- [5] Vialatte F B, Maurice M, Dauwels J, et al. Steady-state visually evoked potentials: focus on essential paradigms and future perspectives[J]. Progress in Neurobiology, 2010, 90(4): 418-438.
- [6] 张学军, 杨京儒. 基于 PSD 特征的 FBCCA 脑电信号识别方法[J]. 科学技术与工程, 2024, 24(4): 1411-1417.  
Zhang Xuejun, Yang Jingru. FBCCA SSVEP EEG signal recognition method based on PSD[J]. Science Technology and Engineering, 2024, 24(4): 1411-1417.
- [7] 张黎明, 张小栋, 陆竹风, 等. 用于稳态视觉诱发电位特征频率提取的同步压缩短时傅里叶变换方法[J]. 西安交通大学学报, 2017, 51(2): 20-26.  
Zhang Liming, Zhang Xiaodong, Lu Zhufeng, et al. A synchrosqueezing short-time fourier transform method for characteristic frequency extraction of steady state visual evoked potential[J]. Journal of Xi'an Jiaotong University, 2017, 51(2): 20-26.
- [8] 杜光景, 谢俊, 张玉彬, 等. 用于稳态视觉诱发电位脑机接口

- 目标识别的深度学习[J]. 西安交通大学学报, 2019, 53(11): 42-48.
- Du Guangjing, Xie Jun, Zhang Yubin, et al. A deep learning method for target recognition in steady-state visual evoked potential-based brain-computer interface[J]. Journal of Xi'an Jiaotong University, 2019, 53(11): 42-48.
- [9] Wu T, Yan G Z, Yang B H, et al. EEG feature extraction in brain computer interface based on wavelet packet decomposition[J]. Chinese Journal of Scientific Instrument, 2008, 41(6): 618-625.
- [10] Chen G, Song D D, Liao L J. A multi-channel SSVEP-based brain-computer interface using a canonical correlation analysis in the frequency domain[J]. Advances in Intelligent Systems and Computing, 2014, 215: 603-616.
- [11] Antelism J M, Riveraca C A, Galvise E, et al. Detection of SSVEP based on empirical mode decomposition and power spectrum peaks analysis[J]. Biocybernetics and Biomedical Engineering, 2020, 40(3): 1010-1021.
- [12] Zhang Z X, Geiger J, Pohjalainen J, et al. Deep learning for environmentally robust speech recognition: an overview of recent developments[J]. ACM Transactions on Intelligent Systems & Technology, 2017, 9(5): 1-28.
- [13] Taber Y R, Halici U. A novel deep learning approach for classification of EEG motor imagery signals[J]. Journal of Neural Engineering, 2016, 14(1). DOI: 10.1088/1741-2560/14/1/016003.
- [14] An X, Kuang D P, Guo X J, et al. A deep learning method for classification of EEG data based on motor imagery[C]//10<sup>th</sup> International Conference on Intelligent Computing. Berlin: Springer, 2014: 203-210.
- [15] 李新, 吴迎年, 李睿. 基于稳态视觉诱发电位的脑电信号分类算法比较[J]. 科学技术与工程, 2021, 21(19): 8106-8112.
- Li Xin, Wu Yingnian, Li Rui. Comparison of electroencephalogram signal classification algorithms based on steady state visual evoked potential[J]. Science Technology and Engineering, 2021, 21(19): 8106-8112.
- [16] 李睿, 张小栋, 张黎明, 等. 面向假肢的场景动画稳态视觉诱发脑控方法[J]. 西安交通大学学报, 2017, 51(1): 115-121.
- Li Rui, Zhang Xiaodong, Zhang Liming, et al. Brain-controlled prosthesis manipulation based on scene graph-SSVEP[J]. Journal of Xi'an Jiaotong University, 2017, 51(1): 115-121.
- [17] Lecun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1999, 86(11): 2278-2324.
- [18] 崔石玉, 朱志宇. 基于参数迁移和一维卷积神经网络的海水泵故障诊断[J]. 振动与冲击, 2021, 40(24): 180-189.
- Cui Shiyu, Zhu Zhiyu. Seawater pump fault diagnosis based on parameter transfer and one-dimensional convolutional neural network[J]. Journal of Vibration and Shock, 2021, 40(24): 180-189.
- [19] Liang Z, Wang L, Tao M L, et al. Attention mechanism based resnet network for automatic modulation classification[C]//IEEE Globecom Workshops. Madrid: IEEE, 2021: 1-6.
- [20] Wang L L, Yao W J, Chen C, et al. Driving behavior recognition algorithm combining attention mechanism and lightweight network[J]. Entropy, 2022, 24(7). DOI: 10.3390/e24070984.
- [21] Yu Y, Si X S, Hu C H, et al. A review of recurrent neural networks: LSTM cells and network architectures[J]. Neural Computation, 2019, 31(7): 1235-1270.