

## 基于 AE 和 Transformer 的运动想象脑电信号分类研究

蒋锐<sup>1</sup>, 孙刘婷<sup>1</sup>, 王小明<sup>1</sup>, 李大鹏<sup>1</sup>, 徐友云<sup>1,2</sup>

(1. 南京邮电大学通信与信息工程学院, 江苏 南京 210003;

2. 南京邮电大学通信与网络技术国家工程研究中心, 江苏 南京 210003)

**摘要:** 基于运动想象的脑机接口系统一直是海内外研究学者的关注对象。针对传统运动想象脑电识别系统不能精准提取显著特征、分类识别准确率低等问题, 提出一种新的基于自编码器(AE, auto-encoder)降维的Transformer分类识别模型。该方法使用滤波器组共空间模式(FBCSP, filter bank common spatial pattern)对数据进行多个频段的特征提取, 并利用AE获得降维后的特征矩阵。同时借助Transformer模型的位置编码考虑全局信号特征影响并利用多头自注意力机制考虑特征矩阵的内部关联性, 提升系统分类识别效果。与传统的基于线性判别分析(LDA, linear discriminant analysis)的K-近邻(KNN, K-nearest neighbors)法分类识别系统进行对比, 实验表明AE+Transformer模型的分类识别效果优于LDA+KNN系统, 说明这种改进后的算法适用于运动想象的二分类。

**关键词:** 运动想象; 深度学习; 自编码器; 注意力模块; Transformer模型

**中图分类号:** TN911.73

**文献标志码:** A

**doi:** 10.11959/j.issn.2096-3750.2023.00310

## Research on EEG signal classification of motor imagery based on AE and Transformer

JIANG Rui<sup>1</sup>, SUN Liuting<sup>1</sup>, WANG Xiaoming<sup>1</sup>, LI Dapeng<sup>1</sup>, XU Youyun<sup>1,2</sup>

1. School of Telecommunications and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China

2. National Engineering Research Center for Communication and Network Technology,  
Nanjing University of Posts and Telecommunications, Nanjing 210003, China

**Abstract:** The motor imagery brain-computer interface has always been the focus of scholars. But traditional system cannot accurately extract significant signals and has low classification accuracy. To overcome such difficulty, a new Transformer model was proposed based on the auto-encoder (AE). The filter bank common spatial pattern (FBCSP) was used to extract the features of multiple frequency bands, and the AE was exploited to obtain the dimensionality-reduced feature matrix. Finally, it considered the influence of the global signal features by the position encoding of the Transformer model and considered the internal correlation of the feature matrix by using the multi-head self-attention mechanism. By comparison with the traditional K-nearest neighbors (KNN) system based on linear discriminant analysis (LDA), the experimental results validates that the classification effect of AE+Transformer model is better than that of LDA+KNN system. It shows that the improved algorithm is suitable for the binary classification of motor imagery.

**Key words:** motor imagery, deep learning, auto-encoder, attention module, Transformer model

### 0 引言

2013年, 美国政府公布了“脑研究计划”, 该

计划促进了脑神经科学的发展。自此脑机接口(BCI, brain-computer interface)进入研究人员的视野, 并在21世纪初得到空前的发展。BCI系统使用特殊的

收稿日期: 2022-07-08; 修回日期: 2022-11-06

通信作者: 蒋锐, j\_ray@njupt.edu.cn

基金项目: 国家自然科学基金资助项目(No.61971241, No.62071245)

Foundation Items: The National Natural Science Foundation of China (No.61971241, No.62071245)

脑电图 (EEG, electroencephalogram) 信号采集设备, 通过算法系统分类识别出 EEG 信号所带有的类人思考特征, 从而派生出对应的输出信息, 实现人脑与外部控制设备的“对话”。

目前, BCI 研究<sup>[1]</sup>主要分为几大领域, 分别是 EEG 信号采集、EEG 信号预处理<sup>[2]</sup>、信号特征提取、特征分类识别等。其中最为核心的是 EEG 信号预处理、信号特征提取和特征分类识别<sup>[3-4]</sup>。EEG 信号预处理一般针对采集 EEG 信号过程中出现的不可避免的噪声, 主要包括人体肌肉产生的肌电信号、眼球的转动产生的伪迹。常见的 EEG 信号预处理采用频率滤波或者阈值法等。采用频率滤波预处理时, 能够有效去除带内的噪声, 但是可能直接滤除部分频带外的有用信号分量, 丢失了部分特征信息。采用阈值法<sup>[5]</sup>预处理时, 会保留未达到阈值的 EEG 信号, 从而保存较为完整的特征信息。但是由于生物信号的非平稳性, 很大可能会留下藏在生物信号里面的伪迹。

对运动想象 EEG 信号进行特征提取能有效捕获具有较为显著特征的向量矩阵, 为提升分类准确率打下基础。常用的特征提取方法主要有: 功率谱密度 (PSD, power spectral density)、自适应自回归 (AAR, adaptive auto-regressive)、独立成分分析 (ICA, independent component analysis)、共空间模式 (CSP, common spatial pattern)<sup>[6]</sup>。PSD 分析法直接将时域的 EEG 信号转换至频域中, 从而直观地看到 EEG 信号的能量变化。但是该转换会损失 EEG 信号的时域信息, 较为片面<sup>[7]</sup>。AAR 模型法大多用于平稳信号模型, 虽然其计算过程较为简单, 但是 EEG 信号的不平稳性使得 AAR 算法也不是绝佳的选择<sup>[8]</sup>。ICA 算法的核心是独立分量提取, 主要针对的是信号频段间无重叠区域的 EEG 信号。该算法用于 EEG 信号上, 一般效果不甚理想<sup>[9]</sup>。CSP 算法主要针对数据矩阵的协方差矩阵进行处理。该算法找到使两个矩阵在空间中方差最大化的方向, 以分辨不同类的脑电特征。因此不需要将受试者的 EEG 信号频段分成无重叠的频段区间后再进行特征提取, 但是在单一频段具有局限性<sup>[10]</sup>。

研究 BCI 的最终目标是对采集到的 EEG 信号进行分类识别, 从而达到对人类脑信息的推测, 给身体不便但脑未受损的人群带来福音。目前, EEG 信号的分类算法<sup>[11-12]</sup>主要包括线性回归、线性判别分析 (LDA, linear discriminant analysis)、 $K$ -近邻

(KNN,  $K$ -nearest neighbor)、支持向量机 (SVM, support vector machine) 等。其中使用比较多的分类器是 SVM、KNN, 一般能得到较为满意的分类结果。近几年, 深度学习框架逐渐在 BCI 系统领域中占有一席之地, 其模型之一的 Transformer 模型更是吸引了大众的目光。Transformer 模型于 2018 年被提出, 是一种强大的自然语言处理模型<sup>[13]</sup>。它的本质是一种注意力机制模型, 单个模块的计算复杂度低于其他的神经网络, 且多层结构能够同时进行计算, 加快计算效率。2021 年, Tao<sup>[14]</sup>等提出一个门控 Transformer 结构, 并将 Transformer 应用到脑视觉领域和 EEG 信号领域, 通过借助门控机制, 在前馈层或者自注意力层前对输入矩阵进行归一化处理, 并在每个子层后加上门控层, 从而强化学习。但是该模型在应用到 EEG 领域时, 分类准确率较低, 后续优化有待研究。2022 年, 李畅<sup>[15]</sup>等公布了一个基于 Transformer 引导卷积神经网络 (CNN, convolutional neural network) 的脑电信号分类方法的专利。他们将原始的 EEG 数据进行预处理后, 通过 CNN 和 Transformer 交替进行局部特征和全局特征的提取。但是 CNN 的卷积层在系统中进行局部特征提取, 会导致小部分信号特征丢失。同时 CNN 的结合导致整体网络较为复杂, 系统运行时间较长, 网络学习效率较低。

本文将 Transformer 模型应用到 EEG 信号特征分类中, 利用自编码器进行数据降维。自编码器作为降维算法, 可以有效减少矩阵的维度, 提高网络运行速率; Transformer 模型通过改变其模型的部分结构, 能够适应 EEG 领域的研究需求, 简化脑电识别流程, 有效地进行 EEG 信号分类。

## 1 基于 AE 和 Transformer 的运动想象 EEG 信号分类算法

运动想象分类识别系统主要分为 3 个模块, 分别是 EEG 信号预处理、信号特征提取和特征分类识别。传统的分类识别系统常常利用频率滤波对信号进行预处理, 然后通过 CSP 算法提取信号特征向量矩阵, 最后通过经典的 SVM 算法或者 KNN 算法进行分类识别。但是传统的系统容易出现单一特征域分析思维。从特征识别而言, 传统的 KNN 算法已经不能满足二分类的需求, 引入具有强大自适应功能的深度学习框架优化分类器的思维逐渐成为热点<sup>[16]</sup>。基于上述情况, 本文提出一种 EEG 信号

分类识别系统。利用滤波器组公共空间模式(FBCSP, filter bank common spatial pattern) 算法进行特征提取, 并将输出的特征向量矩阵作为输入, 进入 AE 和 Transformer 的分类识别系统, 从而提升分类准确率。

### 1.1 滤波器组共空间模式

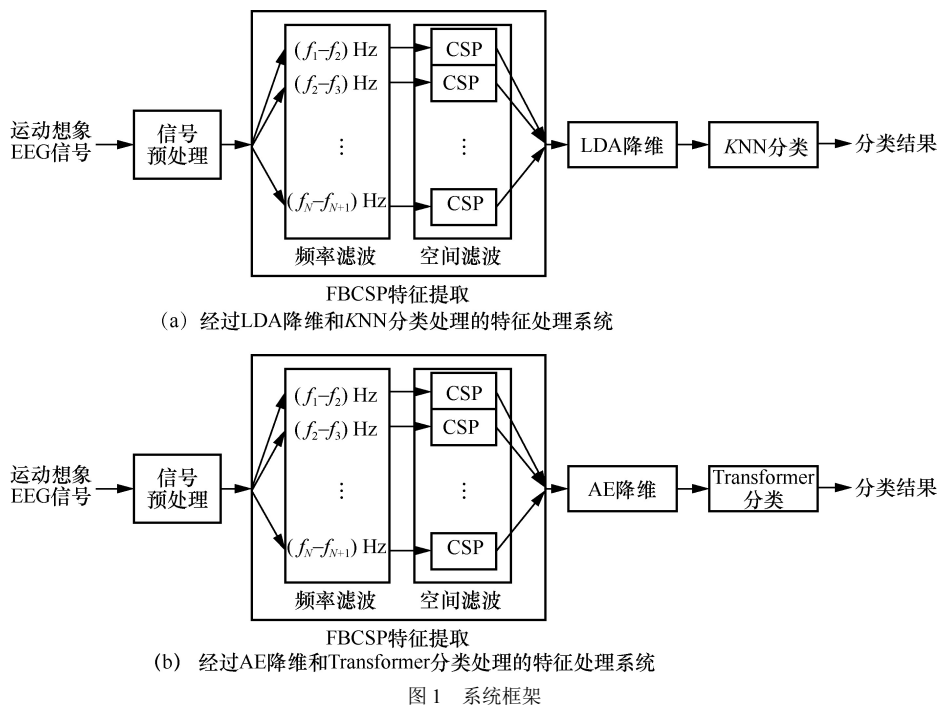
CSP 算法<sup>[17]</sup>通过 EEG 信号数据的协方差矩阵进行对角化, 在空间中找到能使两个矩阵方差达到最大的方向, 也就是将两个矩阵的差距最大化, 从而得到空间滤波器。

然而, CSP 算法只能够在某一频段中进行信号特征提取。因此, 算法性能取决于 EEG 信号的工作频段, 具有一定的局限性。FBCSP 算法则是在 CSP 算法的基础上进行改进, 将数据在整个频段上进行分割, 从而形成多个小频段的数据。然后针对每个频段的数据使用空间滤波器进行滤波, 将每个小频段的协方差矩阵差距最大化, 从而提取每个小频段上的重要

特征向量矩阵并进行特征分类识别。分频段的做法避免了选取频段从而丢失重要特征信息的情况。

利用 FBCSP 算法进行 EEG 信号的特征处理主要分为以下 4 个步骤: 频率滤波、空间滤波、特征选择和特征分类。其中, 频率滤波和空间滤波又被统一称为特征提取阶段<sup>[18-19]</sup>。系统框架如图 1 所示, 在本次运动想象分类识别实验中, 选取 FBCSP 算法的频率滤波和空间滤波作为特征提取阶段, 将特征提取后的矩阵分别经过 LDA 降维和 KNN 分类处理的传统方法系统以及 AE 降维和 Transformer 分类处理的本文方法系统, 以证明本文方法的系统性能。

将经过滤波预处理的 EEG 信号分成不同频段的信号样本, 并将训练和测试的 EEG 信号分成  $f_1 \sim f_2$ 、 $f_2 \sim f_3$ 、 $f_3 \sim f_4$ 、 $\dots$ 、 $f_N \sim f_{N+1}$  共  $N$  个频段。随后对每个频段的信号进行 CSP 特征提取, 提取相关特征矩阵, 便于后续分类识别操作。



### 1.2 Transformer 模型

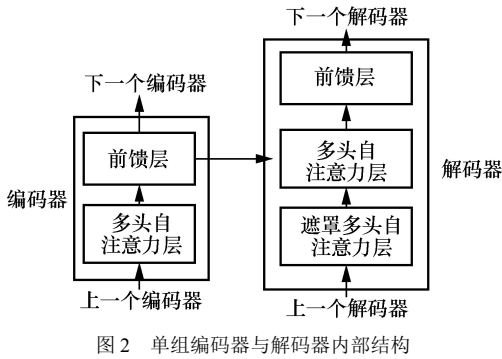
经过预处理的 EEG 信号被分成不同的频段分别进行 CSP 算法空间滤波, 经过特征选择算法捕捉较为显著的特征, 形成特征向量矩阵, 随后进行 Transformer 模型分类识别。

在输入特征信号矩阵后, Transformer 模型对序列数据嵌入了位置信息, 随后将含有位置信息的输

入数据传输到编码解码组件中。编码解码组件内部嵌套一个或多个小编码解码结构, 针对数据进行处理<sup>[20]</sup>。整个过程基于注意力机制, 该模型利用多个编/解码器进行并行计算, 缩短了网络的计算时间, 大大增加了网络的学习效率<sup>[21]</sup>。单组编码器与解码器内部结构如图 2 所示。

编码器和解码器都包括多个注意力模块。单个

编码器的模块包括一个多头自注意力层和一个前馈层，其输入是目标序列经过位置编码后的目标矩阵。单个解码器的模块包括一个遮罩多头自注意力层、一个多头自注意力层和一个前馈层。解码器的输入和编码器类似，也是经过位置编码后的目标序列。在编码时，通过多头自注意力层的输出矩阵使用残差神经网络的模式将其和目标矩阵相加并进行批量标准化，最后将结果传输至前馈层。前馈层的实质是由两层全连接层构成，并使用线性整流函数（ReLU, linear rectification function）增加非线性关系。前一个编码器内部的前馈层的输出结果一方面传输至下一个编码器的多头自注意力层，另一方面传输至该编码器对应的解码器内部多头自注意力层。解码的过程与编码类似，最后解码器的输出是经过线性变换和 Softmax 计算得到的最终分类结果。



将 Transformer 模型应用到 EEG 信号分类识别领域中，假设经过 FBCSP 算法进行信号特征提取后形成的特征向量矩阵为  $\mathbf{C} \in \mathbf{R}^{N \times d_{\text{input}}}$ ， $N$  为 FBCSP 算法的频段数目， $d_{\text{input}}$  表示 Transformer 输入特征矩阵的行数。考虑各频段特征向量时序信息对分类结果的影响，需在经过维度转换后通过位置编码将 Transformer 分类器的每个特征数据嵌入时序信息。即利用正弦和余弦<sup>[19]</sup>函数添加表征每个频段特征信号时序信息的位置编码，位置编码为

$$\text{PE}(\text{pos}, i) = \begin{cases} \sin(\text{pos} / 10000^{2i/d_{\text{model}}}) & i=0, 2, 4, \dots \\ \cos(\text{pos} / 10000^{2i/d_{\text{model}}}) & i=1, 3, 5, \dots \end{cases} \quad (1)$$

其中， $\text{pos}$  表示该数据在整个频段内的位置，且  $\text{pos} = 0, 1, 2, \dots, N-1$ ； $i$  表示当前数据在当前数据矩阵内的位置  $\text{index}$  对 2 求模后的值，即  $i = \text{index} // 2$ ， $i = 0, 1, 2, \dots, d_{\text{model}} // 2$ ； $d_{\text{model}}$  表示 Transformer 输出特征矩阵的列数。

经过位置编码后的矩阵也是多头自注意力层

的输入矩阵，表示为

$$\mathbf{X} \in \mathbf{R}^{d_{\text{input}} \times d_{\text{model}}} \quad (2)$$

其中，每一行是一维特征矩阵， $d_{\text{model}}$  是 Transformer 输出特征矩阵的列数。将输入矩阵  $\mathbf{X}$  分别与预设的权重矩阵相乘得到查询向量矩阵  $\mathbf{Q}$ 、键向量矩阵  $\mathbf{K}$  和值向量矩阵  $\mathbf{V}$ ，有

$$\begin{cases} \mathbf{Q} = \mathbf{X}\mathbf{W}^Q \\ \mathbf{K} = \mathbf{X}\mathbf{W}^K \\ \mathbf{V} = \mathbf{X}\mathbf{W}^V \end{cases} \quad (3)$$

其中，权重矩阵  $\mathbf{W}^Q$ 、 $\mathbf{W}^K$  和  $\mathbf{W}^V$  均是  $d_{\text{model}} \times d_{\text{model}}$  的线性矩阵。此时， $\mathbf{Q}$ 、 $\mathbf{K}$ 、 $\mathbf{V}$  本质仍是同一个矩阵，维数均是  $d_{\text{input}} \times d_{\text{model}}$ 。设  $H$  是多头自注意力层的头数，将  $\mathbf{Q}$ 、 $\mathbf{K}$ 、 $\mathbf{V}$  分别投影  $H$  次，每次投影采用不同的权重矩阵，设第  $i$  次投影的权重矩阵为

$$\begin{cases} \mathbf{W}_i^Q \in \mathbf{R}^{d_{\text{model}} \times d_Q} \\ \mathbf{W}_i^K \in \mathbf{R}^{d_{\text{model}} \times d_K} \\ \mathbf{W}_i^V \in \mathbf{R}^{d_{\text{model}} \times d_V} \end{cases} \quad 1 \leq i \leq H \quad (4)$$

$$d_Q = d_K = d_V = d_{\text{model}} / H \quad (5)$$

则第  $i$  次形成的一组查询向量  $\mathbf{Q}_i$ 、键向量  $\mathbf{K}_i$  和值向量  $\mathbf{V}_i$  为

$$\begin{cases} \mathbf{Q}_i = \mathbf{Q}\mathbf{W}_i^Q \\ \mathbf{K}_i = \mathbf{K}\mathbf{W}_i^K \\ \mathbf{V}_i = \mathbf{V}\mathbf{W}_i^V \end{cases} \quad 1 \leq i \leq H \quad (6)$$

基于多头自注意力层对  $\mathbf{Q}$ 、 $\mathbf{K}$ 、 $\mathbf{V}$  进行投影的本质是将向量矩阵投影到不同的表示子空间。多头自注意力层相关计算式为

$$\text{Attention}(\mathbf{Q}_i, \mathbf{K}_i, \mathbf{V}_i) = \text{softmax}(\mathbf{Q}_i \mathbf{K}_i^T / \sqrt{d_K}) \mathbf{V}_i \quad (7)$$

$$\text{head}_i = \text{Attention}(\mathbf{Q}_i, \mathbf{K}_i, \mathbf{V}_i) \quad (8)$$

$\text{MultiHead}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{Concat}(\text{head}_1, \dots, \text{head}_H) \mathbf{W}^o$  (9) 其中， $\text{softmax}(\cdot)$  表示 Softmax 映射<sup>[22]</sup>，Concat 是连接函数， $\mathbf{W}^o \in \mathbf{R}^{Hd_V \times d_{\text{model}}}$  为联合连接多头自注意力层的权重矩阵。

权重矩阵  $\mathbf{W}_i^Q$ 、 $\mathbf{W}_i^K$ 、 $\mathbf{W}_i^V$  和  $\mathbf{W}^o$  的内部参数是可以学习训练的，经随机初始化后在反向传播或者反馈校验的时候不断更新修正。对于权重矩阵的参数修正，其修正思想类似于神经网络反馈修正的过程。首先确定网络的损失函数  $\text{loss}$ ，并设置网络学习率  $\eta$ ，则新的权重矩阵为<sup>[23]</sup>

$$W_{\text{new}} = W - \eta \frac{\partial \text{loss}}{\partial W} \quad (10)$$

在权重矩阵修正的过程中，损失函数 loss 一般选取均方损失或者交叉熵损失，且对其求偏导并更新权重参数的过程常会利用优化算法，如 Adam 算法和 SGD 算法，从而进行反向修正<sup>[24]</sup>。

在 EEG 信号系统中，一般作为输入矩阵  $X$  的样本数据量维度较大。为了解决该问题，本文算法采用了 AE 对特征数据进行处理，从而降低输入矩阵  $X$  的维度和数据量。

### 1.3 自编码器

AE 本质是一种 3 层神经网络结构，其框架结构如图 3 所示，整个编码器包括输入层、隐藏层和输出层。输入层将输入数据进行编码操作并传递到隐藏层，然后经过解码过程传输到输出层。隐藏层捕获了原始输入的重要特征。该编码器的实际原理类似于主成分分析 (PCA, principal component analysis) 法<sup>[25]</sup>，改变隐藏层中的权重参数，使得整个系统的输入与输出相同，从而实现自适应无监督的特征提取<sup>[26]</sup>。

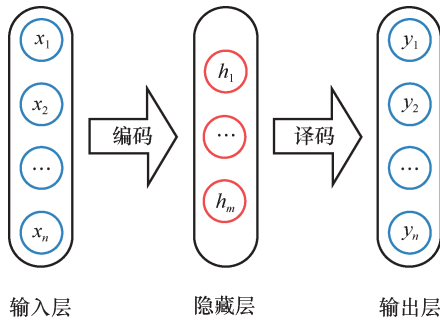


图3 AE 框架结构

设自编码器的原始输入信号  $x \in [0,1]^n$ 。在编码阶段，通过编码函数就可以得到隐藏层  $h$ 。编码函数的定义为

$$h = \text{encoder}(x) = g(W_h \cdot x + b) \quad (11)$$

其中， $W_h \in \mathbf{R}^{m \times n}$  为与输入层和隐藏层层数有关的权重矩阵， $g(\cdot)$  为激活函数， $b \in [0,1]^m$  为偏置向量。

在解码阶段，为了得到输出层  $y$ ，将隐藏层  $h$  作为解码函数的输入，有

$$y = \text{decoder}(h) = g(W_h' \cdot h + b') \quad (12)$$

其中，隐藏层和输出层的权重矩阵为  $W_h' \in \mathbf{R}^{n \times m}$ ，偏置向量为  $b' \in [0,1]^n$ 。

在模型的训练过程中，为使每个输入信号与输出信号相同，则模型的目标函数为

$$\min \sum_j |y^{(j)} - x^{(j)}| \quad (13)$$

当式(13)取得最小值时，该状态下的隐藏层输出矩阵即对输入信号进行 AE 降维后的信号矩阵。一般 EEG 信号的数据维度较高，直接使用一层 AE 的隐藏层将高维数据降到二维或三维或者是便于后续处理的较低维度会丢失一些微小特征，不能尽可能少失真地还原数据，从而导致分类准确率降低。将 AE 应用到 EEG 信号降维领域，一般叠加多个隐藏层，将上一个隐藏层的输出作为下一层的特征输入。同时降维幅度一般采用折半降维，指数地减少训练数据量，增加压缩效率。

## 2 实验与分析

在实际采集 EEG 信号过程中，由于信号幅度较小，寻常的 EEG 信号难以进行后续滤波和分类处理。但有以下几种信号的幅度较为明显：运动想象信号、稳态视觉诱发电位和 P300 信号。本文选取幅度明显的运动想象信号作为研究对象，采用经典的 BCI2003 竞赛的开源数据 III 作为基本数据。该数据集被广泛应用于运动想象分类识别系统中，从而验证模型的性能。

### 2.1 数据处理

本文采用的数据集是来自 BCI2003 竞赛的开源数据 III，数据由奥地利格拉茨技术大学 BCI 实验中心上传<sup>[27]</sup>。该数据集将一位 25 岁年轻健康的女性作为实验对象，该实验对象需要坐在屏幕正前方的椅子上。受试者在轻松安静的环境中进行实验，她需要观察屏幕上的箭头来执行相应的操作，即屏幕会随机显示向左或者向右的箭头，受试者需要在看到指示箭头之后用大脑想象相应的左右手的任务。EEG 信号采集的电极位置与实验范式如图 4 所示。

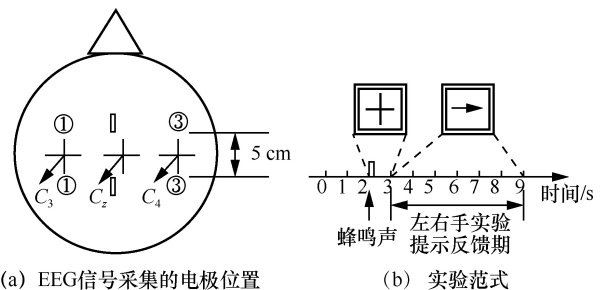


图4 EEG 信号采集的电极位置与实验范式

该数据集由 7 组实验组成，每组实验利用该实验范式进行 40 次试验，且 7 组实验均在同一天进行，中间间隔几分钟。整个数据集给出了 280 个试验数据，包括用于训练的 140 个数据和用于测试的 140 个数据。无论是训练数据还是测试数据，均各含有 70 个左手运动想象和右手运动想象，保证左右手想象概率相等。每次试验的整体时间是 9 s，如图 4(b)所示。具体的实验范式内容如下：在试验开始的 0~2 s 时间内，受试者平静且放松地坐于休闲椅上，并保持静息状态。当  $t=2$  s 时，开始出现蜂鸣声这个听觉刺激且声音由低到高，代表试验正式开始。同时受试者正前方的屏幕上开始出现十字“+”符号，整个事件时间持续为 1 s。当  $t=3$  s 时，受试者正前方的屏幕上随机出现向左或向右的指向箭头，受试者需要立刻用大脑想象左右手运动任务，直至 9 s 试验结束。

10-20 国际标准导联系统统一规定了 EEG 信号采集系统的采集通道位置，并为其标号。其中，对运动想象反应敏感的通道是  $C_3$ 、 $C_z$  和  $C_4$ 。该实验从这 3 个采样通道采集脑电数据，并使用 Ag / AgCl 干电极进行记录。该数据集的采样频率是 128 Hz，且整个采集数据被输入带宽区间为 [0.5, 30] Hz 的带通滤波器进行滤波处理，然后被保存。

2.2 实验过程

本实验按照 EEG 信号分类识别的过程，将数据集数据输入系统并将数据传入频率滤波器进行预处理，再将预处理后的数据送入 FBCSP

算法进行分频段处理从而捕获信号，最后再将特征矩阵通过降维算法传输至 Transformer 模型进行分类识别。EEG 信号特征提取分类网络框架如图 5 所示。

为了去除信号中的噪声，提高模型的分类准确率，将运动想象信号进行严格的预处理。针对运动想象 EEG 信号有较明显的信号差别现象主要发生在 8~50 Hz 的频段范围内，因此选取通带为 8~50 Hz 的带通滤波器，滤除眼电信号、肌电信号和系统噪声。由数据集的实验范式可知， $t=3\sim 9$  s 为运动想象任务时间， $t=3\sim 3.5$  s 为大脑反应时间， $t=7\sim 9$  s 为运动想象结束时间。所以在预处理阶段主要有两件事，首先将数据进行 8~50 Hz 带通滤波，滤除噪声；其次选择特征比较明显的 3.5~7 s 时间段内的 EEG 信号数据作为后续实验的基础原数据。

在特征提取阶段，采用 FBCSP 算法，使用八阶巴特沃斯滤波器将预处理后的 EEG 信号分段。根据第 1.1 节的内容可知，将运动想象在 8~48 Hz 频段上的信号分成 8~12 Hz、12~16 Hz、16~20 Hz、20~24 Hz、24~28 Hz、28~32 Hz、32~36 Hz、36~40 Hz、40~44 Hz 和 44~48 Hz 共 10 个频段，再对每个小频段的信号进行 CSP 特征提取从而得到特征矩阵，便于后续特征分类识别。

随后将特征向量矩阵送入分类器 Transformer 模型，并使用 Adam 优化器动态调整参数的学习率，使得参数趋于平稳，最后得到分类结果。在特征分类过程中，由于特征向量矩阵维度较高，系统需要

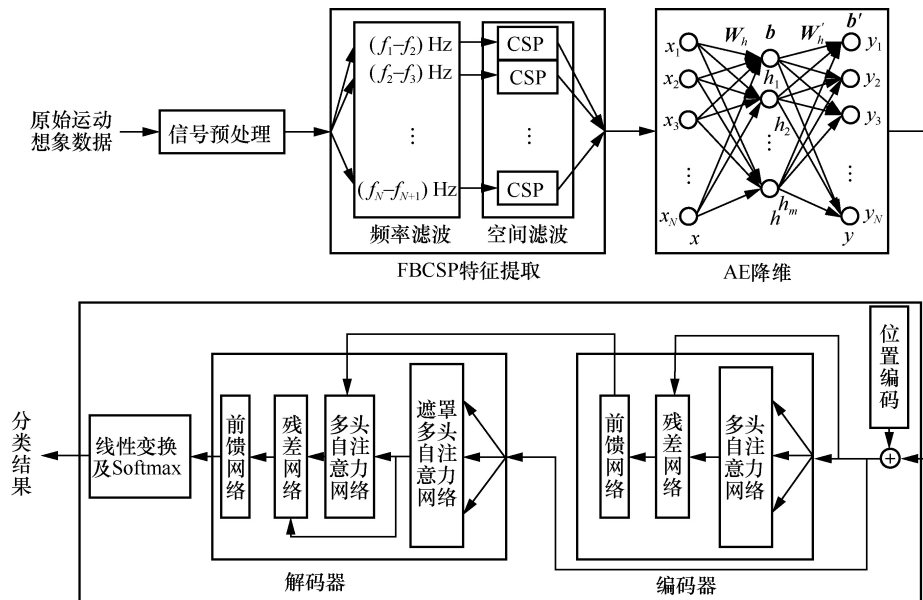


图 5 EEG 信号特征提取分类网络框架

捕捉输入数据最显著的特征从而降低数据量，所以引入 AE。为了减少训练时间，采用 Multitask 模式，即利用一个网络同时训练 AE 和 Transformer 分类器，并行计算减少了网络运行时间。

在 Transformer 模型的系统中，部分参数设置如下：整个实验使用 Torch 框架。使用 Adam 优化器，其学习率一般设置为  $10^{-3}$ ，权重衰减系数设置为  $2.5 \times 10^{-2}$ <sup>[28]</sup>，丢弃率 dropout 一般设置在 0.2~0.5。经过反复实验，当丢弃率 dropout 设置为 0.25 且实验 epoch 设置为 200 时，系统的分类识别率达到最大。

输入维度  $d_{input}=6$ ，输出特征矩阵的维度  $d_{model}=512$ ， $Q$ 、 $K$  和  $V$  的维度  $d_k=64$ ，则将多头的数量  $H$  设置为  $8$ <sup>[28]</sup>。最后特征向量矩阵通过含有两层全连接层和一个激活函数的前馈神经网络送入含有 Softmax 函数的输出层，整个网络输出是包含每类出现概率的二维输出矩阵，其表示为 EEG 信号经过系统后预测为当前类的概率。

在 AE 和 Transformer 模型的系统中，一些参数设置如下：在 AE 降维阶段，将特征矩阵经过 512 维、256 维，最后将数据压缩降至 128 维，整个过程完成从输入层到隐藏层的变换。然后，解码器利用较低维度的数据重构输入数据，将隐藏层的数据按照上述的维度参数逆向升维，使输出层结果与输入层结果相同。则可以说明隐藏层的输出矩阵是输入层的 128 维低维矩阵，将其作为 Transformer 模型的输入。在 Transformer 解码的线性变换阶段，采用了两个全连接层，将 128 维数据经过 64 维降至二维作为系统输出，便于后续显示。

在构建网络的过程中，利用损失函数计算每次迭代的前向计算结果和真实值的差距，并利用损失函数的倒数，沿着梯度最小的方向反向传播，修正前向计算的权重值。本次实验主要将联合均方和交叉熵整体作为损失函数，不断对权值进行修正。同时整个实验将特征向量矩阵加入 LDA+KNN 算法，并比较两种系统的分类结果。

### 2.3 实验结果分析

评估一个模型的优劣可以有多重标准，在 BCI 中的应用与其他应用中是一样的，主要评估模型的分类准确率。混淆矩阵见表 1，经过实验仿真，得到不同算法的混淆矩阵结果，测试结果见表 2。

表 1 混淆矩阵

真实情况	预测结果		含义
	正例	反例	
正例	TP(将正例正确预测为正例)	FN(将正例错误预测为负例)	TP+FN 表示实际数据集中正样本的数量
反例	FP(将负例错误预测为正例)	TN(将负例正确预测为负例)	FP+TN 表示实际数据集中负样本的数量
含义	TP+FP 表示预测的正类样本数	FN+TN 表示预测的负类样本数	TP+FN+FP+TN 表示样本总数

表 2 混淆矩阵测试结果

对比项	TP	FN	FP	TN
KNN	59	11	7	63
LDA+KNN	59	11	5	65
Transformer	60	10	5	65
AE+Transformer	60	10	2	68

在本次实验中，使用以下几个评估标准。

#### 2.3.1 分类准确率与 Kappa 系数分析

根据表 1 中的混淆矩阵，可以得到一种性能评估标准——分类准确率 ACC。分类准确率本质是最终经过系统识别预测正确的 EEG 样本数和总 EEG 样本数的比值，表示为

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

定义错误率  $err = 1 - ACC$ ，同时定义当每一类的初始数目相同时， $ACC_0 = 1/N$ 。其中， $N$  表示运动想象任务的类别数目。由于本次实验数据集是二分类，所以  $N$  的值为 2，从而  $ACC_0 = 0.5$ 。

另一个性能评估方法是 Kappa 系数，Kappa 系数是在分类准确率的基础之上进行定义的，计算式为

$$k = \frac{ACC - ACC_0}{ACC_0} \quad (15)$$

Kappa 系数的值域是  $[-1, 1]$ 。根据式(15)可知，与具体的每类样本个数和类别无关。若  $k = 0$  意味着  $ACC = ACC_0$ ，即表示最后的结果对于每个样本分类的概率都是 0.5。若  $k = 1$  意味着是理想状态，每个样本的最后预测均是正确的。

利用 KNN、LDA+KNN、Transformer 模型和 AE+Transformer 模型 4 个系统对数据集进行二分类识别，其分类准确率和 Kappa 系数分别如图 6 和图 7 所示。

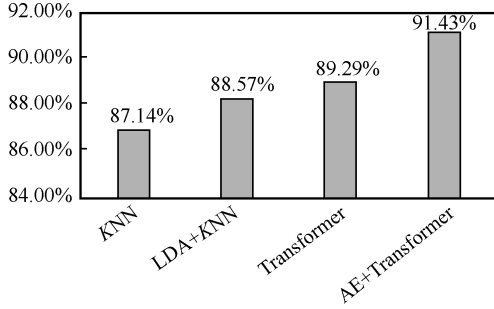


图 6 不同系统的分类准确率

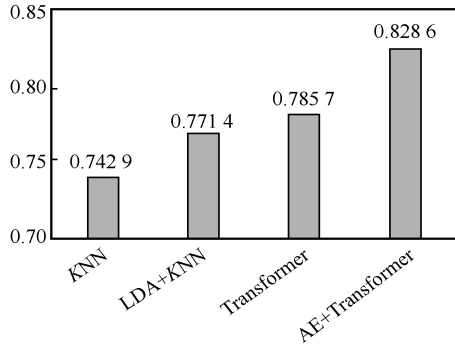


图 7 不同系统的 Kappa 系数

综合图 6 和图 7 可以看出，在 BCI 竞赛 2003 数据 III 中，Transformer 模型得到的分类准确率为 89.29%，比同等条件下的 KNN 算法提高了 2.15 个百分点。同时 Transformer 模型的 Kappa 系数为 0.7857，比 KNN 算法提高了 0.0428，可见 Transformer 模型整体结果优于 KNN 算法。由于 Transformer 模块内部结构复杂度较高，所以提前使用 AE 对经过 FBCSP 算法的特征向量进行降维处理。对比同样进行降维的 LDA+KNN 系统，该模型的准确率提高了 2.86 个百分点，Kappa 系数也提高了 0.0572。结果表明在 EEG 信号分类识别系统中，降维算法也能提高分类的准确率，从而优化整个系统。

### 2.3.2 F1-score 分析

F1-score 作为二分类的评估标准，可以通过表 1 求得分类的精确率 (precision) 和召回率 (recall)，从而被求得。

$$\text{precision} = \frac{TP}{TP + FP} \quad (16)$$

$$\text{recall} = \frac{TP}{TP + FN} \quad (17)$$

则 F1-score 为

$$\text{F1-score} = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \quad (18)$$

联合表 1 可知，precision 的实质是在所有预测为正类样本中实际真正为正类的比例；recall 的实质是在实际数据集所有的正样本中预测为正类的比例；F1-score 的实质是 precision 和 recall 的调和平均数，其值域为 [0,1]。在兼顾 precision 和 recall 的值时，F1-score 越靠近 1，代表该模型的性能越好。4 个系统的 F1-score 如图 8 所示。

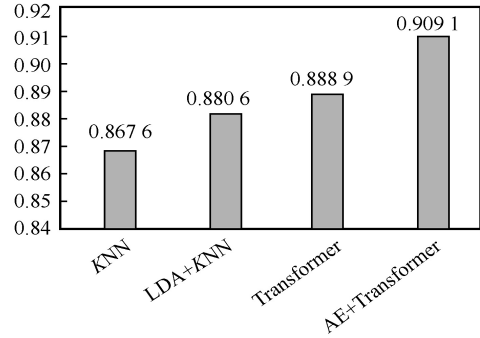


图 8 4 个系统的 F1-score

从图 8 可以看出，加入降维算法的系统比原始系统在 F1-score 方面有了较大的提升。同时对于加入降维算法的 LDA+KNN 分类系统来说，F1-score 为 0.8806，比同样加入 AE 降维的 Transformer 系统低了 0.0285，辅助 ACC 说明了 AE+Transformer 分类系统能够提升分类精度。

同时将 Transformer 系统与基于 CNN 的 Transformer 系统<sup>[15]</sup>进行对比，结果见表 3。

表 3 CNN + Transformer 与 Transformer 对比

	ACC	Kappa	F1-score	运行时间/s
CNN+Transformer	88.57%	0.7714	0.8873	32.1102
Transformer	89.29%	0.7857	0.8889	21.2441

对比发现，CNN 的卷积层可以在系统中有效地进行局部特征提取，但同时小部分信号特征也会因此而丢失，从而导致算法性能的下降。除此之外，由于 CNN 的存在使 Transformer 网络趋于复杂，增加分类识别系统的运行时间。因此，本文算法通过对全局特征进行分类，可以显著缩短系统的运行时间，提高算法实时性，并且进一步提升了系统的分类准确率。

### 2.3.3 ROC 曲线和 AUC 分析

受试者操作特征 (ROC, receiver operating characteristic) 曲线是判断二分类的分类器性能优劣的方法之一，其中真阳率 (TPR, true positive rate)

和假阳率 (FPR, false positive rate) 的计算式为

$$TPR = \frac{TP}{TP + FN} \quad (19)$$

$$FPR = \frac{FP}{FP + TN} \quad (20)$$

系统实验一共有两个输入,一个是数据集中真实的样本,另一个为通过系统输出的每个样本判为正类的概率。将概率由大至小排序并设为阈值,大于该阈值则预测样本为正类,小于该阈值则预测样本为负类。计算每个样本的 TPR 值和 FPR 值,并绘制曲线。在 ROC 曲线图中,每个坐标点的横坐标是对应的 FPR 值,纵坐标是对应的 TPR 值。横坐标 FPR 越大,则表示预测正类中实际负类越多;纵坐标 TPR 越大,则表示预测正类中实际正类越多。

曲线下面积 (AUC, area under the curve) 作为衡量优劣的标准之一,其实质是 ROC 曲线下的面积,其值域为 $[0,1]$ <sup>[29]</sup>。当 AUC 的值越来越靠近 1 时,ROC 曲线的整个趋势越来越上凸,即越来越靠近点(0,1)。

不同算法的 ROC 曲线如图 9 所示。从整个 ROC 曲线趋势来看,ROC 曲线越靠近坐标轴的左上方(0,1)点,则算法性能越好<sup>[30]</sup>。本文所提 AE + Transformer 算法明显具备最优的算法性能。除此之外,AUC 值越大说明分类器性能越优。本文所提 AE + Transformer 算法的 AUC 值不但比传统的 KNN 系统提升了 0.02~0.03,而且比 CNN + Transformer 算法提高了 0.02。因此,无论是从 ROC 曲线趋势的定性对比,还是通过 AUC 数值的定量分析,本文算法均表现出分类识别的优越性,可以有效应用到运动想象 EEG 信号分类中。

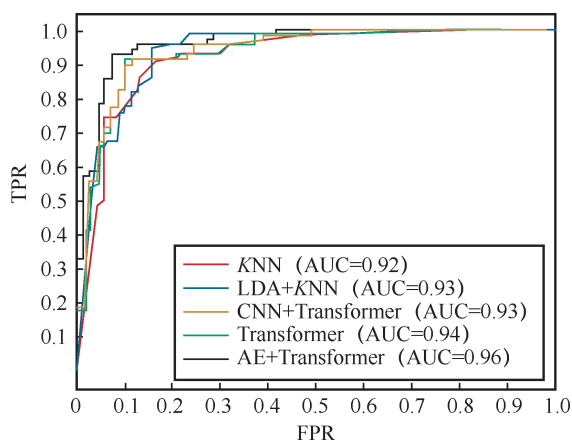


图 9 不同算法的 ROC 曲线

### 3 结束语

本文提出了一种 AE 降维与 Transformer 模型分类识别的运动想象分类算法,其实质就是将 Transformer 模型运用到 EEG 信号处理中,利用 AE 将高维度的 EEG 信号特征数据进行降维,使用 Transformer 模型将具有时间特质的 EEG 信号特征矩阵嵌入位置信息,通过编码与解码,实现 EEG 信号的特征分类。为了验证该算法是否可行,将 BCI 竞赛 2003 数据 III 作为待测数据,采用数据预处理对数据进行频率滤波,并使用 FBCSP 分段提取特征矩阵,最后进行 LDA+KNN 分类和 AE+Transformer 分类识别处理。同时采用分类准确率、Kappa 系数及相对应的 ROC 曲线对不同算法的优劣进行评估。从实验结果可以看出,本文提出将 AE 与 Transformer 模型结合应用到 EEG 信号领域,与 LDA+KNN 系统相比,准确率提高了 2.86 个百分点,Kappa 系数提高了 0.057 2。同时借助 F1-score 值、ROC 曲线和 AUC 值也辅助说明了该结果,也能看出 AE+Transformer 算法在运动想象二分类中的优势。

在未来的研究中,笔者将进一步优化该算法中的 Transformer 模型,提高其与运动想象 EEG 信号的适配度。同时将当前模型的适应领域由运动想象二分类识别领域扩大到运动想象四分类甚至其他实验范式产生的 EEG 信号中,从而扩大算法模型的影响。

#### 参考文献:

- [1] WANG S F, LEE Y H, SHIAH Y J, et al. Time-frequency analysis of EEGs recorded during meditation[C]//Proceedings of 2011 1st International Conference on Robot, Vision and Signal Processing. Piscataway: IEEE Press, 2011, 11(10): 73-76.
- [2] RUSNAC A L, GRIGORE O. EEG preprocessing methods for BCI imagined speech signals[C]//Proceedings of 2021 International Conference on e-Health and Bioengineering (EHB). Piscataway: IEEE Press, 2021: 1-4.
- [3] MÜHL C, CHANEL G, ALLISON B, et al. Third workshop on affective brain-computer interfaces (ABCI 2013): introduction[C]//Proceedings of 2013 Humaine Association Conference on Affective Computing and Intelligent Interaction. Piscataway: IEEE Press, 2013: 821.
- [4] WEN D, LIANG B B, ZHOU Y H, et al. The current research of combining multi-modal brain-computer interfaces with virtual reality[J]. IEEE Journal of Biomedical and Health Informatics, 2021, 25(9): 3278-3287.
- [5] KAUATI-SAITOÉ, SILVEIRA G F M, DA-SILVA P J G, et al. Classi-

- fication of motor tasks from EEG signals comparing preprocessing techniques[C]//XXVI Brazilian Congress on Biomedical Engineering, 2019: 109-113.
- [6] WANG K, ZHAI D H, XIA Y Q. Motor imagination EEG recognition algorithm based on DWT, CSP and extreme learning machine[C]//Proceedings of 2019 Chinese Control Conference (CCC). Piscataway: IEEE Press, 2019: 4590-4595.
- [7] 王行愚, 金晶, 张宇, 等. 脑控: 基于脑-机接口的人机融合控制[J]. 自动化学报, 2013, 39(3): 208-221.  
WANG X Y, JIN J, ZHANG Y, et al. Brain control: human-computer integration control based on brain-computer interface[J]. Acta Automatica Sinica, 2013, 39(3): 208-221.
- [8] SHI T W, REN L, CUI W H. Feature extraction of brain-computer interface electroencephalogram based on motor imagery[J]. IEEE Sensors Journal, 2020, 20(20): 11787-11794.
- [9] FEDOSOV N, LEVADNIY I, DMITRIEV A, et al. Independent component analysis for different movements detection in BCI application based on sensorimotor rhythms[C]//Proceedings of 2020 Ural Symposium on Biomedical Engineering, Radio electronics and Information Technology (USBREIT). Piscataway: IEEE Press, 2020: 69-72.
- [10] SELIM S, TANTAWI M M, SHEDEED H A, et al. A CSP\AM-BA-SVM approach for motor imagery BCI system[J]. IEEE Access, 2018, 6(10): 49192-49208.
- [11] JIA X W, LI K, LI X Y, et al. A novel semi-supervised deep learning framework for affective state recognition on EEG signals[C]//Proceedings of 2014 IEEE International Conference on Bioinformatics and Bioengineering. Piscataway: IEEE Press, 2014: 30-37.
- [12] LI X Y, JIA X W, XUN G X, et al. Improving EEG feature learning via synchronized facial video[C]//Proceedings of 2015 IEEE International Conference on Big Data(Big Data). Piscataway: IEEE Press, 2015: 843-848.
- [13] 孙嘉瑶. 深度学习注意力机制在脑电信号分类中的应用[D]. 深圳: 中国科学院大学(中国科学院深圳先进技术研究院), 2021.  
SUN J Y. Application of deep learning attention mechanism in EEG classification[D]. Shenzhen: Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, 2021.
- [14] TAO Y Z, SUN T, MUHAMED A, et al. Gated transformer for decoding human brain EEG signals[J]. Annual International Conference of the IEEE Engineering in Medicine and Biology Society IEEE Engineering in Medicine and Biology Society Annual International Conference, 2021: 125-130.
- [15] 李畅, 蔡国超, 黄晓阳, 等. 基于 Transformer 引导卷积神经网络的脑电信号分类方法: CN114564991A[P]. 2022-05-31.  
LI C, CAI G C, HUANG X Y, et al. EEG (electroencephalogram) signal classification method based on Transform guided convolutional neural network: CN114564991A[P]. 2022-05-31.
- [16] LU N, LI T F, REN X D, et al. A deep learning scheme for motor imagery classification based on restricted boltzmann machines[J]. IEEE Transactions on Neural Systems and Rehabilitation Engineering: a Publication of the IEEE Engineering in Medicine and Biology Society, 2017, 25(6): 566-576.
- [17] ZHANG R, LI X P, WANG Y W, et al. Using brain network features to increase the classification accuracy of MI-BCI inefficiency subject[J]. IEEE Access, 2019, 7: 74490-74499.
- [18] DAS R, LOPEZ P S, AHMED KHAN M, et al. FBCSP and adaptive boosting for multiclass motor imagery BCI data classification: a machine learning approach[C]//Proceedings of 2020 IEEE International Conference on Systems, Man, and Cybernetics. Piscataway: IEEE Press, 2020: 1275-1279.
- [19] CHRISTENSEN S M, HOLM N S, PUTHUSSERYPADY S. An improved five class MI based BCI scheme for drone control using filter bank CSP[C]//Proceedings of 2019 7th International Winter Conference on Brain-Computer Interface (BCI). Piscataway: IEEE Press, 2019: 1-6.
- [20] HUNG J W, LIN J R, ZHUANG L Y. The evaluation study of the deep learning model transformer in speech translation[C]//Proceedings of 2021 7th International Conference on Applied System Innovation (ICASI). Piscataway: IEEE Press, 2021: 30-33.
- [21] SALTZ P, LIN S Y, CHENG S C, et al. Dementia detection using transformer-based deep learning and natural language processing models[C]//Proceedings of 2021 IEEE 9th International Conference on Healthcare Informatics. Piscataway: IEEE Press, 2021: 509-510.
- [22] 徐嘉杰, 卢兆军, 袁飞, 等. 基于深度神经网络和 SoftMax 分类器的台区负荷分类识别方法[J]. 电气自动化, 2021, 43(6): 102-104, 114.  
XU J J, LU Z J, YUAN F, et al. Classification and identification method of station load based on deep neural network and SoftMax classifier[J]. Electrical Automation, 2021, 43(6): 102-104, 114.
- [23] 邓建国. 有监督学习中的损失函数及图像标注应用[D]. 太原: 太原科技大学, 2020.  
DENG J G. Loss function in supervised learning and its application in image annotation[D]. Taiyuan: Taiyuan University of Science and Technology, 2020.
- [24] ZHANG Z J. Improved Adam optimizer for deep neural networks[C]//Proceedings of 2018 IEEE/ACM 26th International Symposium on Quality of Service (IWQoS). Piscataway: IEEE Press, 2018: 1-2.
- [25] CHENG L W, LI D L, YU G J, et al. A motor imagery EEG feature extraction method based on energy principal component analysis and deep belief networks[J]. IEEE Access, 2020, 8(8): 21453-21472.
- [26] HOSSEINI M P, POMPILI D, ELISEVICH K, et al. Optimized deep learning for EEG big data and seizure prediction BCI via Internet of Things[J]. IEEE Transactions on Big Data, 2017, 3(4): 392-404.
- [27] SHAHLAEI F, BAGH N, SHALIGRAM A D, et al. Classification of motor imagery tasks using inter trial variance in the brain computer interface[C]//Proceedings of 2018 IEEE International Symposium on Medical Measurements and Applications (MeMeA). Piscataway: IEEE Press, 2018: 1-6.
- [28] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[EB]. 2017.
- [29] 李嘉莹, 赵丽, 边琰, 等. 基于 LDA 和 KNN 的下肢运动想象脑电信号分类研究[J]. 国外电子测量技术, 2021, 40(1): 9-14.  
LI J Y, ZHAO L, BIAN Y, et al. Classification of lower limb motor imagination signals based on LDA and KNN[J]. Foreign Electronic

Measurement Technology, 2021, 40(1): 9-14.

[30] 兰成辉, 李江天, 李敏, 等. 基于皮电和肌电的驾驶疲劳判别阈值研究[J]. 科技风, 2021, 1(5): 171-172.

LAN C H, LI J T, LI M, et al. Research on driving fatigue discrimination threshold based on skin and emg[J]. Technology Wind, 2021, 1(5): 171-172.

[作者简介]



蒋锐 (1985- ), 男, 博士, 南京邮电大学副教授, 主要研究方向为雷达信号处理、移动通信系统、脑波信号处理等。



王小明 (1986- ), 男, 博士, 南京邮电大学副教授, 主要研究方向为通信资源协调分配、深度学习等。



李大鹏 (1982- ), 男, 博士, 南京邮电大学教授, 主要研究方向为移动通信系统等。



孙刘婷 (1997- ), 女, 南京邮电大学通信与信息工程学院硕士生, 主要研究方向为脑波信号处理等。



徐友云 (1966- ), 男, 博士, 南京邮电大学教授, 主要研究方向为移动通信系统、5G 通信等。