

# QTFD 与 DenseNet 相结合的运动想象分类方法

金晶<sup>1,2</sup> 杨益雕<sup>1</sup> 孙浩<sup>1</sup> 王行愚<sup>1</sup>

(1. 华东理工大学能源化工过程智能制造教育部重点实验室, 上海 200237;

2. 华东理工大学深圳研究院, 广东深圳 518063)

**摘要:** 运动想象脑机接口 (Motor Imagery Brain Computer Interface, MI-BCI) 技术近年来在医疗康复、娱乐等许多领域得到了广泛的运用。然而, 如何处理非平稳的脑电信号 (Electroencephalography, EEG), 并从中提取可辨识的特征并分类仍然是主要难点之一。针对这个问题, 本研究提出了一种基于二次型时频分布 (Quadratic time-frequency distributions, QTFD) 和密接型网络 (DenseNet) 的新型 MI-EEG 分类模型。具体地, 我们首先使用 QTFD 初步提取 MI 任务相关的脑电时频特征, 并构造得到 EEG 片段的高分辨率时频表示。常用的线性时频分析方法往往会忽略部分非线性信息, 难以准确地描述 MI 信号的能量分布。与线性时频分析方法相比, QTFD 方法以二次型变换的形式将信号从时域投影到时频域, 能更好地描述信号的能量分布, 其对时间和频率的变化具有不变性, 能提供较为稳定准确的时频特征。随后, 本研究采用了轻量级网络模型 DenseNet 对时频表示的浅层和深层特征进行逐级提取并整合。DenseNet 可训练参数量较少, 适用于数据量较少的 MI-BCI 应用, 它在每层网络之间都建立了直接连接, 每一层网络都可以访问之前所有网络的特征图, 从而得到更具有区分性的特征表示。最后, 本研究在 BCI 竞赛 IV 数据集上进行了实验验证, 将提出的分类模型与各先进对比算法进行了比较。结果表明, 我们所提出的方法在使用脑电通道数更少的情况下, 获得了更好的分类性能。

**关键词:** 脑机接口; 脑电信号; 运动想象; 神经网络; 时频变换

**中图分类号:** R318 **文献标识码:** A **DOI:** 10.16798/j.issn.1003-0530.2023.08.010

**引用格式:** 金晶, 杨益雕, 孙浩, 等. QTFD 与 DenseNet 相结合的运动想象分类方法[J]. 信号处理, 2023, 39(8): 1443-1454. DOI: 10.16798/j.issn.1003-0530.2023.08.010.

**Reference format:** JIN Jing, YANG Yidiao, SUN Hao, et al. QTFD and DenseNet based motor imagery classification method[J]. Journal of Signal Processing, 2023, 39(8): 1443-1454. DOI: 10.16798/j.issn.1003-0530.2023.08.010.

## QTFD and DenseNet Based Motor Imagery Classification Method

JIN Jing<sup>1,2</sup> YANG Yidiao<sup>1</sup> SUN Hao<sup>1</sup> WANG Xingyu<sup>1</sup>

(1. Key Laboratory of Smart Manufacturing in Energy Chemical Process, Ministry of Education, East China University of Science and Technology, Shanghai 200237, China; 2. Shenzhen Research Institute of East China University of Science and Technology, Shenzhen, Guangdong 518063, China)

**Abstract:** Motor imagery (MI)-based brain-computer interfaces (BCIs) have gained immense popularity in a wide range of fields, including medical rehabilitation and entertainment. However, effectively processing non-stationary Electroencephalography (EEG) signals and extracting recognizable features from them remain major obstacles. In this study, our primary objective was to address this challenge by employing Quadratic Time-Frequency Distribution (QTFD) as an initial step to

收稿日期: 2023-05-06; 修回日期: 2023-06-02

基金项目: 国家自然科学基金项目(62176090); 科技创新 2030 重点项目 2022ZD0208900; 部分获得上海市科技重大专项 2021SHZDZX 项目资助; 部分获得“111 计划”B17017 项目资助; 国家地方科技发展引导基金项目(深圳市)(2021Szvup043); 江苏省科技计划 2022 年度专项资金项目 BE2022064-1(重点研发计划、产业前瞻和关键核心技术)

extract relevant EEG time-frequency features associated with MI tasks. By utilizing QTFD, we constructed high-resolution time-frequency representations of EEG fragments, enabling a more comprehensive analysis. Conventional linear time-frequency analysis methods often overlook nonlinear information and struggle to accurately depict the energy distribution of MI signals. In contrast, QTFD projects the signal from the time domain to the frequency domain using quadratic transformations, resulting in a more precise description of the signal's energy distribution. Moreover, QTFD exhibits invariance to changes in time and frequency, offering relatively stable and accurate time-frequency characteristics. To further enhance feature extraction, we adopted the lightweight DenseNet network model. This model progressively extracts and integrates shallow and deep features from the time-frequency representations. DenseNet, with its small number of trainable parameters, is particularly well-suited for MI-BCI applications with limited available data. A notable advantage of DenseNet is its direct interconnection between each layer of the network, allowing feature graphs from all preceding layers to be accessed. Consequently, this facilitates the generation of a more distinguishable feature representation. Finally, we conducted experimental validation on the BCI Competition IV dataset, comparing the performance of our proposed classification model with various state-of-the-art algorithms. The results conclusively demonstrate that our method achieves superior classification performance, even when utilizing a reduced number of EEG channels. These findings underscore the efficacy of our approach in effectively processing non-stationary EEG signals and extracting informative features for MI-BCI applications.

**Key words:** brain-computer interface; electroencephalography; motor imagery; neural network; time-frequency transform

## 1 引言

脑机接口(Brain Computer Interface, BCI)是一种将脑电信号转换为外部设备控制命令的新兴人机交互技术<sup>[1]</sup>,能够为用户与外界建立起沟通的桥梁,近年来受到了研究者的广泛关注。BCI技术最早应用在康复医学领域,它能够帮助中风和瘫痪患者实现与外界环境的沟通交流,辅助行动不便的患者使用轮椅或假肢等<sup>[2-3]</sup>。此外,随着信号处理方法和脑科学的不断发展,BCI也逐渐延伸应用于其他领域,如机器人技术、教育、娱乐和军事等<sup>[4-7]</sup>。

通过电极在头皮表面采集的脑电图信号(Electroencephalogram, EEG),具有时间分辨率高、无创性等优点,同时兼具设备成本低特性,已经成为BCI系统中应用最为广泛的大脑活动记录方法<sup>[8]</sup>。根据不同的激活模式,EEG可以分为慢皮层电位(Slow Cortical Potentials, SCP)<sup>[9]</sup>、事件相关电位(Event-Related Potentials, ERP)<sup>[10-11]</sup>和感觉运动节律(Sensorimotor Rhythm, SMR)<sup>[12-13]</sup>等。其中,受到许多研究者关注的运动想象(Motor Imagery, MI)BCI主要使用了SMR信号。当人体执行肢体运动任务或想象肢体运动时,大脑皮层感觉运动区域的SMR将会增加或减少,出现局部事件相关同步(Event-Related Synchronization, ERS)或事件相关去同步(Event-Related Desynchronization, ERD)现象,MI-BCI可以通过ERD/ERS的不同EEG模式来解码用

户的运动意图<sup>[14-15]</sup>。与其他非侵入式BCI相比,例如P300-BCI、稳态视觉诱发电位(Steady-State Visual Evoked Potential, SSVEP)BCI,MI-BCI不需要外部刺激来诱发EEG的电位活动,不需要额外的辅助硬件设备,因此对用户和研究者而言使用更为便捷,具有更广泛的应用前景<sup>[16-18]</sup>。然而,由于EEG信号属于非平稳性信号,具有低信噪比的特性,其信号方差和均值随着时间不断变化,并且容易受到噪声和伪迹的干扰。因此,提高MI-BCI的分类性能是一个具有挑战性的问题<sup>[19]</sup>。

MI分类任务的特征提取算法可以大致分为传统机器学习和深度学习两种。在传统机器学习方面,共空间模式(Common Spatial Pattern, CSP)算法应用最为广泛<sup>[20]</sup>,该算法通过计算特定频段的空间滤波器来提取该频段的EEG频域信息,但特定频段的选择过于依赖研究者个人经验,并且不同受试者和任务对应的最佳频段往往是不同的。因此,只包含了单一特定频段信息的传统CSP方法无法构建通用的分类模型<sup>[21]</sup>。为了缓解单一频段选择带来的影响,ANG等提出了滤波器组CSP(Filter-Bank CSP, FBCSP)<sup>[22]</sup>,FBCSP使用多个空间滤波器将EEG分解为多个子频带,使用CSP算法从多个子频带中提取出更多频域相关的信息。然而,FBCSP仍然非常依赖人类脑动力学等先验知识和经验,这也是其他传统机器学习方法所面临的问题,而仅仅依靠现有的经验知识很难兼顾实际应用中的场景,无法保障模

型的鲁棒性,导致最终的应用效果较差,MI信号的识别准确率较低<sup>[23]</sup>。

近年来,深度学习方法获得了长足的发展,许多领域的前沿算法都逐渐转向了深度学习的研究。在MI-BCI中,深度学习方法将特征提取模型的参数和分类器的参数进行了整合优化,统筹改进了整个分类框架。与传统CSP方法相比,深度学习方法能够构建端到端的分类模型,能够获得更好的特征提取和分类性能,具有更好的鲁棒性。针对MI信号的特征提取和分类,研究者们提出了多种深度学习网络模型。其中,卷积神经网络(Convolutional Neural Network, CNN)在处理MI分类任务方面展现出了巨大的优势和潜力。CNN无须手动提取特征就可以直接从EEG信号时序列中学习ERD/ERS脑电模式,并从中提取层次结构<sup>[24-25]</sup>。例如,LAWHERN等人提出了一种EEGNet浅层CNN模型,它将原始EEG信号作为输入,将多个卷积滤波器以并行、串行的形式相结合,逐级提取信号的时间和空间特征,相比于传统机器学习方法获得了更优越的分类性能<sup>[26]</sup>。DAI等人提出了一种混合卷积尺度的CNN模型,采用了多个不同尺度的时间卷积核对EEG信号进行时域特征提取,并最终整合多尺度特征进行MI任务分类<sup>[27]</sup>。除了EEG时间序列外,EEG的二维表示也常作为CNN的输入来进行MI-BCI的分类任务,其中时频分布(Time-Frequency Distribution, TFD)的二维表示应用最为广泛,如频谱特征<sup>[28-30]</sup>及其二维图像<sup>[31-32]</sup>等。TFD包含了EEG的重要时频特征,常用的时频变换方法主要包括短时傅里叶变换(Short-Time Fourier Transform, STFT),小波变换(Wavelet Transform, WT)等。将EEG的时间序列转换为TFD的二维表示后,许多用于图像分类的深度学习模型也可以应用于BCI的解码过程中,拓宽了BCI的算法研究思路。但TFD与CNN相结合的MI-BCI识别框架的相关研究中,仍然存在以下问题:(1)STFT、WT等二维表示方法都是以线性时频联合函数来描述EEG频谱和能量密度与时间之间关系的,但能量是信号的二次型表示,是非线性的,所以这种线性的时频变换方法往往会忽略掉部分非线性信息<sup>[33]</sup>,难以准确地描述MI信号的能量分布。(2)现有的MI数据集规模较小,面对较深的网络模型或较多的训练参数时,存在训练数据量不足的问题,导致无法训

练得到较好的分类模型。针对以上问题,本研究提出了一种基于非线性二次型时频分布(Quadratic Time-Frequency Distributions, QTFD)和密接型网络(DenseNET)的轻量级MI分类模型。其中,QTFD使用二次型变换的方式将时域EEG信号投影到时频域,能更好地描述MI信号的能量分布,且时频分辨率较高,蕴含的时频信息更准确丰富。基于DenseNet的轻量级网络的训练参数较少,对训练数据量的需求较低,且能逐级提取信号的时频特征,并充分利用和整合浅层特征和深层特征,获得更好的特征提取效果和分类性能。实验结果表明,基于QTFD和DenseNet的MI分类模型相较于对比算法获得了更好的分类性能。

## 2 数据集与信号预处理

### 2.1 数据集描述

本研究使用BCI竞赛IV的2a公开数据集进行MI识别模型的效果验证和结果讨论,该数据集包含了9名被试者的22通道EEG信号和3通道眼电信号。实验过程中,被试者被需要执行4种不同的运动想象任务,分别为想象左手、右手、舌头和脚的运动,每个类别的任务各执行144个trial,每个被试共计576个trial的运动想象任务,运动想象范式如图1所示。伴随着滴声,一个trail开始,屏幕上出现一个固定十字;2s后指示运动想象任务类型的箭头(分别为上、下、左、右)出现并持续1.25s,提示被试者开始进行相应的运动想象任务。被试者被要求持续执行运动想象任务3s直到屏幕上的固定十字消失,随后开始静息态的休息。EEG信号的采样频率为250 Hz,经过0.5 Hz~100 Hz的带通滤波,并使用50 Hz陷波滤波器消除原始EEG信号中的50 Hz电力线路噪声干扰。

### 2.2 数据预处理

本研究的数据预处理过程主要包括三部分:通道选择、数据截取和带通滤波。首先,为了减少输

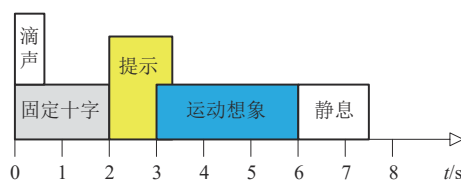


图1 运动想象范式计时方案

Fig. 1 Timing scheme of the motor imagery paradigm

入特征的维度,加快特征提取和分类的速度,本研究只保留了与ERS/ERD密切相关的 $C_3$ 、 $C_2$ 和 $C_4$ 通道,去除了其他冗余脑电通道。同时,也降低了算法的实际应用成本,减少了实验准备时间,提高了被试的使用体验。除了执行MI任务期间的3 s EEG信号之外,本研究还截取了MI任务开始之前的0.5 s EEG数据。因为许多研究指出,MI任务开始之前的数据对MI分类任务也有一定帮助<sup>[34]</sup>。因此,截取每个trial  $[-0.5, 3]$  s的EEG数据进行后续的分析。最后,使用8~30 Hz的五阶Butterworth带通滤波器获取感觉运动节律相关频段的EEG信号片段,主要包含alpha波(8~13 Hz)和beta波(13~30 Hz)频段内的信号成分。

### 2.3 实验设置

本研究的实验在配置有32GB RAM和AMD R7 3700X CPU的PC上进行,网络模型在Nvidia 2080Ti平台上进行训练和测试。采用PyTorch框架搭建CNN模型,并在MATLAB R2021a上对脑电数据进行预处理。对于每个被试,训练数据集、测试数据集和验证数据集按照8:1:1的比例进行划分,并采用了十折交叉验证方法求取平均分类准确率。

## 3 基于非线性二次型时频分布和密接型网络的轻量级MI分类模型

### 3.1 分类模型整体框架

本文提出的基于非线性二次型时频分布和密接型网络的轻量级MI分类模型结构如图2所示。首先,经8~30 Hz带通滤波的EEG信号通过QTFD初步提取时频特征,构造得到其高分辨率的时频表示,并输入所提出的密接性网络进行进一步的特征

提取和分类。在所提出的DenseNet中,输入的QTFD特征先经过一个初始卷积层,提取其浅层特征。初始卷积层包含了3个 $7 \times 7$ 的卷积核(其中 $\text{stride}=2, \text{padding}=3$ ),对应着3个EEG通道,分别对3个通道的时频特征进行初步的特征提取。随后,特征被输入密接块进行进一步的特征提取,并使用过渡层进行降维,且该过程被重复三次。其中,每个过渡层输出的特征图数量均为输入的一半,每个密接块其内部特征图尺寸分别为: $[256, 256]$ 、 $[128, 128]$ 、 $[64, 64]$ ,并在每个密接块内部特征图尺寸保持不变。在经过3个密接块和过渡层之后,网络训练提取得到的深度抽象特征被输入全连接层,最后使用softmax层进行分类。

### 3.2 二次型时频分析方法

为捕捉EEG信号随时间变化的频谱成分,需要从时频域的角度出发对EEG信号进行分析。时频分析方法可以分为两大类:一种是线性时频分析,另一种是二次型时频分析。

在时频分析过程中,信号的能量分布能揭示出信号许多重要的统计性质,对信号分析十分重要。线性时频分析方法以线性时频联合函数描述EEG频谱和能量密度随时间变化的情况,而能量是信号的非线性二次型表示,这就导致线性时频分析方法往往会忽略部分非线性信息,很难较好地阐明信号的能量分布。

二次型时频分析,也称为双线性时频分析,是一种将信号从时域投影到时频域的二次型变换,具有时移和频移不变性,在时频分析领域受到广泛的关注和研究。研究表明,QTFD在时间和频率域上具有高分辨率,能更好地描述信号的能量分布,且

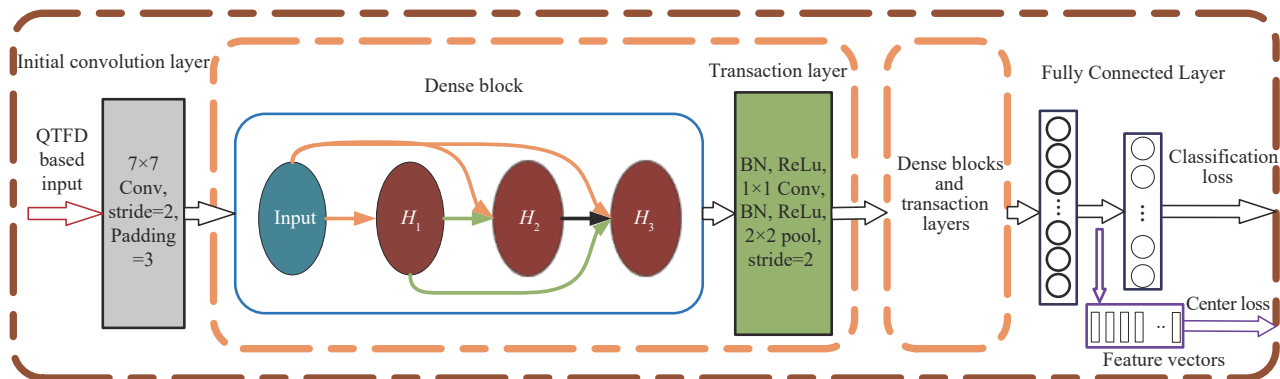


图2 本文所提出的分类模型示意图

Fig. 2 Schematic illustration of the proposed classification model

其对时间和频率的变化具有不变性,能提供较为稳定准确的时频特征。因此,QTFD 能够提供更高质量的 EEG 信号时频分布特征,具有提升 MI-BCI 分类性能的潜力。

QTFD 的构造主要由两部分组成:维纳分布(Wigner-Ville Distribution, WVD)和时频平滑核。对于时域 EEG 信号  $x(t)$ ,可以用式(1)来表示其 QTFD:

$$Q_a(t, f) = W_a(t, f) \underset{t}{*} \underset{f}{*} \gamma(t, f) \quad (1)$$

其中,  $W_a(t, f)$  代表 WVD,  $*$  代表卷积操作,  $\gamma(t, f)$  代表时频平滑核。QTFD 的构造需要  $W_a(t, f)$  和  $\gamma(t, f)$  在时域和频域各进行一次卷积操作。为了提高时频表示的分辨率,需要将时域信号  $x(t)$  转换为其解析形式  $a_x(t)$ <sup>[35]</sup>:

$$a_x(t) = x(t) + jHT\{x(t)\} \quad (2)$$

式中,  $HT\{\}$  代表希尔伯特变换<sup>[36]</sup>。解析信号  $a_x(t)$  的 WVD 在正负频率分量之间不存在交叉项,所以在负频率处能量为零,不会造成信号的信息损失。为获得更高的时频分辨率,可以用式(3)求取瞬时频率  $f_i(t)$ 。

$$f_i(t) = \frac{1}{2\pi i} \frac{d\theta(t)}{dt} \quad (3)$$

式中,  $\theta(t)$  为  $a_x(t)$  的相位。在求得解析信号和瞬时频率后,通过式(4)构造 WVD:

$$W_{a_x}(t, f) = \int_{-\infty}^{\infty} a_x(t + \frac{\tau}{2}) a_x^*(t - \frac{\tau}{2}) e^{-j2\pi f \tau} d\tau \quad (4)$$

WVD 是 QTFD 的一种基本分布,相比于该类的

其他分布,WVD 具有更多有利于信号分析的数学性质,能够提高信号时频表示的分辨率<sup>[37]</sup>。最后,将 WVD 在时频域上与时频平滑核进行卷积计算就能够得到 EEG 信号  $x(t)$  的 QTFD 表达。

此外,时频平滑核的选择对于 QTFD 的表现也是非常重要的。研究表明,WVD 中包含的交叉项可能会影响时频表示中部分信号成分<sup>[32]</sup>。针对此问题,本研究选择了一种指数时频平滑核(Exponential Time-Frequency Kernel)来降低交叉项对 QTFD 性能的影响,并增强了所构造的时频表示的可解释性<sup>[38]</sup>,如式(5)所示。

$$Z(t, f) = \exp(-\frac{t^2 f^2}{\beta^2}) \quad (5)$$

作为密接型网络的输入,QTFD 的时频采样点数均设为 512,每个通道的 QTFD 特征尺寸为(512, 512)。最后,将 3 个通道的 QTFD 特征组合为(512, 512, 3)的张量,输入 DenseNet 模型进行进一步的特征提取和分类。

图 3 是被试 S3 在进行左手运动想象时,经 QTFD 处理后  $C_3$  通道的 EEG 信号的时频图。从图 3(b)中可以看出,8~30 Hz 频段外的 EEG 信号皆被滤除,只保留了与 MI 任务最相关的频段。此外,MI 开始后的 0.5~2.5 s 期间,从图 3(a)的三维时频图中可以明显看出 mu 节律能量有显著下降,并在运动想象结束后恢复,符合左手运动想象期间的右脑皮层感觉运动区的 ERD 脑电模式。这表明该被试在运动想象任务开始后进行了有效的想象任务,并产生了

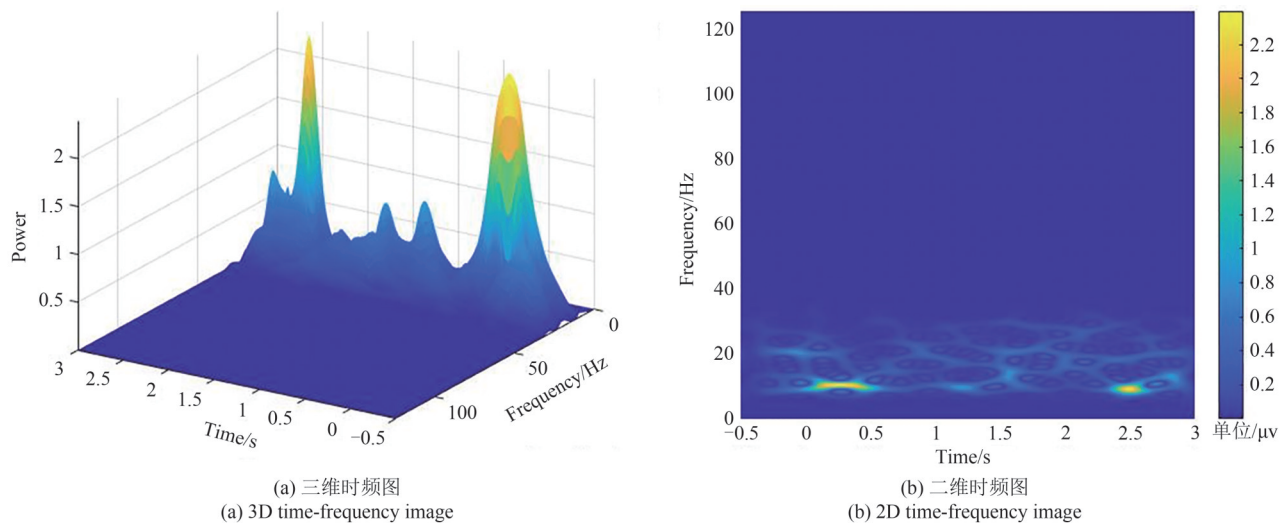


图 3 左手运动想象任务期间  $C_3$  通道的时频图

Fig. 3 Spectrogram of  $C_3$  channel during left hand motor imagery task

相对应的EEG时频特征。

### 3.3 DenseNet模型

DenseNet是一种密接型CNN。相比于传统CNN, DenseNet在每层网络之间都建立了直接连接,每一层网络都可以访问之前所有网络的特征图,从而能够更充分地利用之前层的信息,获得更好的特征提取效果<sup>[39]</sup>。DenseNet的基本结构是由一系列的“密接块(Dense Blocks)”和“过渡层(Transition Layers)”组成的。在密接块中,每个网络层输出的特征图都会被传递到之后的所有网络。在过渡层中,使用 $1 \times 1$ 的卷积操作来减少特征图的大小,从而控制模型的参数数量和计算复杂度。

如图4的上图所示,在密接块中,每个网络层以前馈的方式直接与其他网络层相连,改善了网络层之间的信息流。在数学模型中,即将第 $l$ 层接收所有之前网络层的特征图 $x_0, x_1, \dots, x_{l-1}$ 作为输入:

$$X_l = H_l([x_0, x_1, \dots, x_{l-1}]) \quad (6)$$

其中, $X_l$ 代表第 $l$ 层的特征图, $x_0, x_1, \dots, x_{l-1}$ 代表第0至 $l-1$ 层特征图的串联。 $H_l(\cdot)$ 是一个复合函数,由批归一化(Batch Normalization, BN)、线性整流函数(Rectified Linear Unit, ReLU)和一个 $3 \times 3$ 的卷积层(Convolution Layer, Conv)组成。BN将 $x_0, x_1, \dots, x_{l-1}$ 映射为方差为1、均值为0的特征图,有助于提高模型训练的速度和稳定性。ReLU函数如式(7)所示,将所有输入的值置零并保持所有正值不变,为模型引入了非线性特征表示。

$$\text{ReLU}(x) = \max(0, x) \quad (7)$$

如图4所示,在每个密接块生成特征图的过程中,特征图的尺寸不变,数量不断增加。为了详细阐述特征图数量变化过程,本研究引入了增长率 $k$ 的概念, $k$ 为每个 $H_i$ 函数新生成的特征图数量。例如,若 $H_i$ 函数生成 $k$ 个特征图,那么密接块的第 $l$ 层的特征图数量 $f_n$ 为:

$$f_n = k_0 + k \times (l - 1) \quad (8)$$

其中, $k_0$ 是当前密接块输入特征图的数量。增长率 $k$ 决定密集块中每个网络层 $H_i$ 生成的特征图数目,决定每个网络层的新信息对全局状态的贡献量。尽管每个网络层生成的特征图数量是固定的,但是每个网络层的输入是逐渐递增的。因此,在 $3 \times 3$ 卷积层之前引入了 $1 \times 1$ 的卷积作为瓶颈层,通过对输入通道进行加权求和以减少输入特征图的数量,提高模型的计算效率。最终,密接块中每个层 $H_i$ 的结构为BN-ReLU-Conv( $1 \times 1$ )-BN-ReLU-Conv( $3 \times 3$ ),如图4的下图所示。

密接块之间的层称为过渡层,由 $1 \times 1$ 卷积层和平均池化层(average pooling)组成。由于DenseNet特殊的连接方式,密接块中的特征图会逐级递增,因此需要通过过渡层进行通道数的调整和特征图的压缩。每个过渡层的结构为BN-Conv( $1 \times 1$ )-Pooling( $2 \times 2$ ),若某个密接块内的特征图大小为 $[2w, 2h]$ ,则其下个密接块内的特征图大小则为 $[w, h]$ 。过渡层可以大大减少特征图的通道数,并在空间上对特

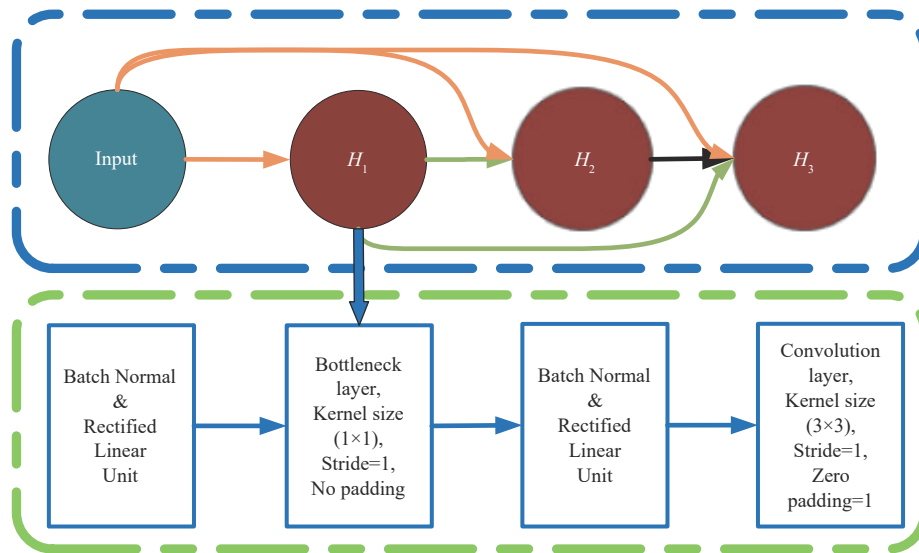


图4 密接块的网络结构及其参数

Fig. 4 The structure and parameters of dense blocks

征图进行下采样,从而显著地减少网络的参数量和计算量,提高模型的效率,并帮助后续的密接块更高效地进行特征学习。

### 3.4 损失函数

本研究使用二进制交叉熵损失(Binary Cross-entropy Loss, BCE)来最小化输出标签和目标标签之间的分类误差,如式(9)所示。同时,我们引入了L2正则损失函数,L2正则损失函数在CNN中的作用是对模型的权重进行惩罚,以防止过拟合问题。在CNN训练过程中,模型的权重会不断地被更新,当权重值过大时,模型会倾向于过度拟合,从而导致泛化能力不足,不能很好地处理新的数据。L2正则损失函数通过向损失函数中添加权重的平方和作为正则化项,使得在训练过程中,关注减低损失函数的同时,也要尽可能减小权重的大小。从而让模型在学习训练数据的同时,尽可能避免对训练数据的过度拟合。其表达式如式(10)所示。

$$\text{Loss}_{\text{BCE}} = -\frac{1}{N_{\text{batch}}} \sum_{i=1}^{N_{\text{batch}}} (t_i \log(o_i) + (1-t_i) \log(1-o_i)) \quad (9)$$

$$\text{Loss}_{L_2} = \|\omega\|_2^2 \quad (10)$$

其中 $t_i$ 和 $o_i$ 分别代表第 $i$ 个训练样本的目标标签和输出标签, $\omega$ 代表模型参数。

此外,许多研究表明中心损失函数可以使得类间距离更小,同类样本更聚集,从而提高模型识别不同类特征的能力。因此,本研究提出的分类模型中也使用了中心损失函数。

$$\text{Loss}_{\text{Center}} = \sum_{j=1}^{N_{\text{batch}}} \|f_k - v_j^m\|_2 \quad (11)$$

其中, $f_k$ 代表第 $k$ 个样本的特征向量, $m$ 代表迭代次数。 $v$ 代表中心向量,按照下式进行迭代:

$$\left\{ \begin{array}{l} \Delta v_j^k = \frac{\sum_{i=1}^{N_{\text{batch}}} \theta(t_i = j) * \|f_i - v_j^m\|_2}{1 + \sum_{i=1}^{N_{\text{batch}}} \theta(t_i = j)} \\ v_j^{m+1} = v_j^m - \rho * \Delta v_j^m \end{array} \right. \quad (12)$$

$$\theta(t_i = j) = \begin{cases} 0, & t_i \neq j \\ 1, & t_i = j \end{cases} \quad (13)$$

其中, $\Delta v$ 表示特征向量和中心向量的平均距离, $\rho$ 是中心损失的学习率。

最终,模型的总损失函数如式(14)所示:

$$\text{Loss} = \text{Loss}_{\text{BCE}} + \lambda_1 \text{Loss}_{L_2} + \lambda_2 \text{Loss}_{\text{Center}} \quad (14)$$

其中, $\lambda_1, \lambda_2$ 为损失权重参数。

### 3.5 评价指标

本研究采用了两种广泛应用于分类模型性能评估的评价指标,分别是准确率(Accuracy, Acc)和Cohen的Kappa系数。Acc指分类器正确分类的样本数占样本总数的比例,Kappa系数也叫假设性的随机一致性概率,用于衡量分类任务的一致性。其数学表达式如(15)式所示, $p_e$ 代表随机预测的Acc。

$$K = \frac{\text{Acc} - p_e}{1 - p_e} \quad (15)$$

此外,在时频分析中,由于各种算法都有其特定的参数和计算方式,会得到不同的时频分析结果,为了对比不同时频分析方法的性能优劣,且对于MI任务,alpha波(8~13 Hz)和beta波(13~30 Hz)频段的信号能量分布尤为重要,因此本文引入了频率聚集度(Concentration)Cf指标。频率聚集度是一种广泛应用于描述信号在时频域上的局部频率集中程度的评价指标<sup>[40-41]</sup>。时频分析方法本质上揭示了信号在某一时间点、某一频率下的能量聚集度,聚集度越高代表其对信号的能量分布的描述越清晰,频谱泄漏也越少。不同频率下的能量聚集度时频表示在频率 $f$ 上的聚集度可以用式(16)来计算。

$$C_f(f) = \sum_{k=1}^T |\text{TFR}(k, f)|^2 \quad (16)$$

其中,TFR表示经时频分析处理后的时频表示, $T$ 为时间采样点数。为了方便不同方法之间进行对比,使用了z-score方法进行正则化处理。

## 4 实验结果

### 4.1 基于QTFD的时频表示特征及聚集度对比

为了检验QTFD方法生成的时频表示特征的质量,验证其是否符合运动想象期间的EEG模式,本文随机选择了BCI竞赛IV2a数据集中被试4在进行左右手MI任务期间的EEG信号,将其通过QTFD方法构造得到的时频特征矩阵绘制成了时频图,以直观地展现其频谱能量随时间变化的特征。

如图5(a)所示,在进行左手MI任务期间,位于大脑皮层左侧感觉运动区的C<sub>3</sub>通道的脑电信号在0 s即实验范式发出运动想象指示后,mu节律出现了明显的能量降低,一直持续到2.5 s左右才恢复,符合ERD的特征模式。对于位于大脑皮层右侧的C<sub>4</sub>通道,MI任务期间EEG信号的mu节律和beta节律一直维持了较高的能量,属于显著的ERS现象。

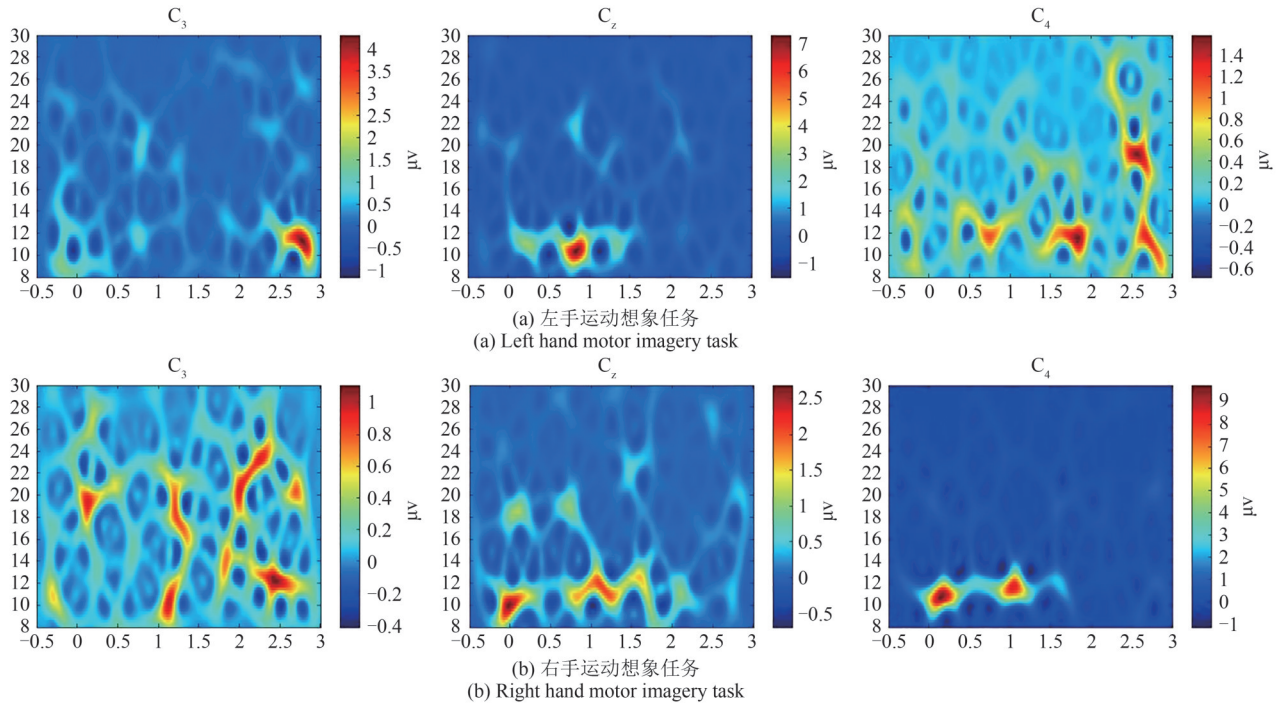


图5 左手运动想象期间基于QTFD的EEG时频表示

Fig. 5 The time-frequency representations from EEG segments by QTFD during left hand MI task

这表明,被试在收到指令后马上开始了有效的运动想象任务,并在2.5 s左右结束。

同理,从与之相对应的图5(b)中也可以发现 $C_3$ 、 $C_4$ 通道的脑电信号有明显的ERS/ERD现象。与图5(a)不同的是,该被试在这次运动想象任务中,直到任务提示开始后的1.5 s左右才进行了有效的运动想象,对应的ERD现象也产生得较晚。由此可见,MI-BCI范式容易受到被试者的个人状况和环境因素影响,其特征出现的时间、幅值均存在个体差异。因此,在特征提取过程中,其浅层的纹理、幅值特征,深层的抽象特征对于MI-BCI的分类都是不可或缺的,而DenseNet正可以很好地处理这个问题。

对于 $C_z$ 通道,由于其处于大脑皮层的中央脑回区域,也是 $C_3$ 、 $C_4$ 通道之间,其时频图表现出的特征也处于两者之间,但明显更接近产生ERD现象的侧脑区特征模式,能帮助分类器加强ERD特征的识别,对于MI-BCI分类任务也有一定的帮助。

综上所述,实验表明通过QTFD得到的时频表示特征包含了大量与MI任务相关的时频信息,证实了其特征的有效性。

相比于其他线性时频分析方法,QTFD构造的时频表示的时频分辨率高(本研究时频采样率均设

为512),蕴含更丰富的时频信息,能更好地描述信号的能量分布。此外,为了对比不同时频表示方法的EEG时频特征提取能力,本文对比了被试在进行左手MI期间 $C_3$ 、 $C_4$ 通道EEG的频率聚集度。除了本文使用的QTFD方法,选择了两种经典线性时频分析方法STFT和WT进行对比,结果如图6所示。从图中可以看出,QTFD的聚集度峰值在各频段显著高于其他两种对比方法。综合 $C_3$ 、 $C_4$ 通道来看,更高的频率聚集度差异能发现更显著的频率变化模式,使得ERD和ERS的特征模式更明显地表现在时频表征上。结果显示,相比于传统线性时频表示方法,QTFD能更好地提取MI-EEG的时频特征。

#### 4.2 分类性能及对比结果

为了验证本文所提出的方法的分类性能,本文与几种先进算法在BCI竞赛IV2a数据集上进行了对比。对比算法包括了三种机器学习算法FBCSP<sup>[22]</sup>, TS-Log<sup>[42]</sup>, NHSP<sup>[43]</sup>和一种CNN脑电分类算法EEG-Net<sup>[26]</sup>,各算法的Acc和Kappa值对比结果如表1所示。

结果显示,本文方法取得了最高的平均准确率72.39%和最高的Kappa值0.632并在6个被试(S1, S2, S5, S7, S8, S9)中都得到了最高的Acc和Kappa



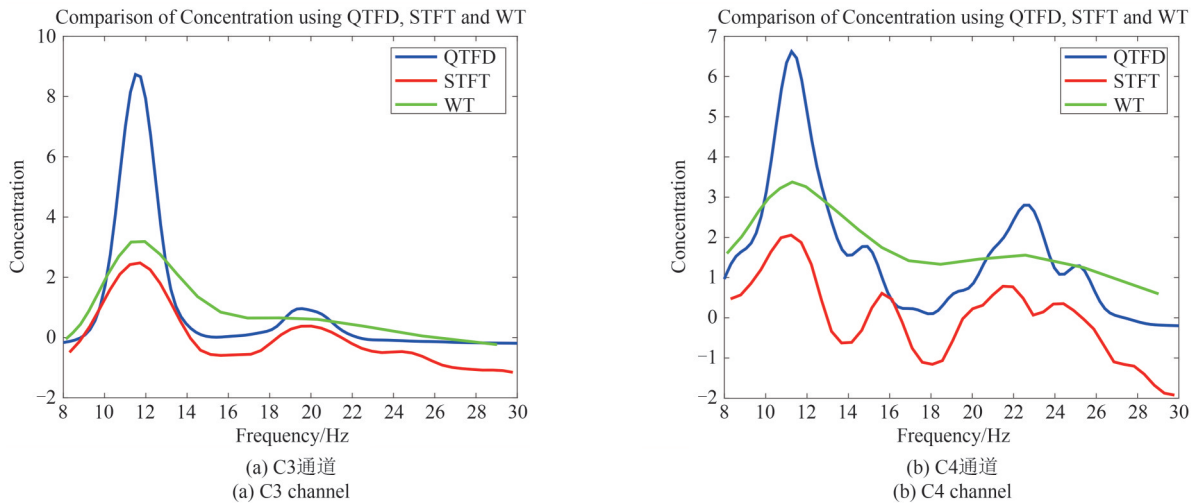


图 6 QTFD, STFT 和 WT 构造的时频表示的频率聚集度对比

Fig. 6 Comparison of frequency concentration of time-frequency representations constructed by QTFD, STFT and WT

表 1 本方法和对比算法在 BCI 竞赛 IV2a 数据集的分类准确率和 Kappa 值对比

Tab. 1 The accuracy comparison of the proposed method and the comparison algorithms in the BCI competition IV2a dataset

被试者	FBCSP	TS-Log	NHSP	EEGNet	本文方法
Metrics	Acc(Kappa)	Acc(Kappa)	Acc(Kappa)	Acc(Kappa)	Acc (Kappa)
S1	0.757(0.676)	0.778(0.704)	0.740(0.653)	0.719(0.625)	<b>0.7885(0.718)</b>
S2	0.563(0.417)	0.438(0.250)	0.350(0.133)	0.510(0.347)	<b>0.6673(0.556)</b>
S3	0.809(0.745)	0.799(0.732)	0.760(0.680)	0.792(0.722)	0.8214(0.762)
S4	0.611(0.481)	0.569(0.421)	0.530(0.373)	0.580(0.440)	<b>0.6491(0.532)</b>
S5	0.549(0.398)	0.480(0.306)	0.380(0.173)	<b>0.649(0.531)</b>	0.6088(0.478)
S6	0.455(0.273)	0.476(0.301)	0.310(0.080)	<b>0.510(0.347)</b>	0.4324(0.243)
S7	0.830(0.773)	0.785(0.713)	0.840(0.787)	0.663(0.551)	<b>0.8725(0.830)</b>
S8	0.816(0.755)	0.798(0.731)	0.740(0.653)	0.743(0.658)	<b>0.8364(0.782)</b>
S9	0.705(0.606)	0.813(0.750)	0.740(0.653)	0.726(0.634)	<b>0.8393(0.785)</b>
<b>Mean</b>	0.677(0.569)	0.659(0.545)	0.599(0.465)	0.655(0.540)	<b>0.7239(0.632)</b>
<b>Std.</b>	0.1829	0.2190	0.2067	0.1019	0.1452(0.194)
<b>p-value</b>	0.0183	0.0405	0.0035	0.0498	

值。4 个对比算法均采用了 22 个通道的全脑电数据,而本文方法仅选择了 3 个与运动想象任务最相关的脑电通道。在 3 个 EEG 通道的情况下,本文方法仍获得了最佳的分类性能。其中,传统机器学习算法 FBCSP、TS-Log 以及 NHSP 分别获得了 0.677、0.659、0.599 的分类准确率, CNN 脑电分类算法 EEGNet 获得了 0.655 的分类准确率,对比 4 个对比算法,本方法分别提高了 11.1%、16.0%、20.9%、10.59%。经过显著性检验, p 值分别为 0.0183、0.0405、0.0035、0.0498,均小于 0.05,验证了本方

法的性能显著优于对比算法。综上所述,本文方法相比于对比算法具有显著的优越性。

为了清晰地展现本文方法对于 4 种运动想象任务的分类性能,本文绘制了来自 9 名被试的平均混淆矩阵,如图 7 所示。

从图 7 中可以看出,本文方法在左右手分类任务上的 Acc 较高,各错分率也较低;而脚和舌头的 Acc 相对较低,相互之间的错分率也较高。这可能与所选择的 3 个脑电通道所覆盖的大脑皮层的区域有关。C<sub>3</sub>、C<sub>4</sub>和 C<sub>z</sub>通道分别位于头皮上左、右和正中

Confusion matrix

Left hand	80.89	9.14	6.43	3.54
Right hand	10.54	78.56	5.36	5.54
Foot	7.37	9.48	64.47	18.68
Tongue	12.74	8.76	12.87	65.63
	Left hand	Right hand	Foot	Tongue

图7 混淆矩阵

Fig. 7 Confusion matrix

央的位置,可以覆盖大脑运动皮层的大部分区域,包括肢体运动区和舌咽肌运动区。具体来说, $C_3$ 和 $C_4$ 通道位于头皮上的左右侧,可以反映大脑皮层上包括手臂和腿部的感觉运动区。 $C_z$ 通道位于头皮的正中央位置,可以反映大脑皮层上的感觉运动区,包括面部和咽喉肌肉运动区。而脚和舌头的感觉运动区分别位于大脑皮层的中央沟回下部和上部,与 $C_3$ 、 $C_4$ 和 $C_z$ 通道所对应的区域有一定的重叠,但并不能被完全覆盖。所以导致了在进行脚和舌头的分类任务时的准确率相对较低。但即便如此,本方法的分类性能仍显著高于其他对比算法,且仅使用了3个通道,减少了EEG信号采集和处理的复杂性,加快了特征提取和分类的速度,并提高了MI-BCI部署的便捷性,为移动BCI的发展和提供了潜在的可能性。

## 5 结论

本文构建了一种基于QTFD和DenseNet的MI-EEG分类模型,该模型利用QTFD将得到的EEG数据转换为高分辨率的时频表示特征,从原始EEG信号中提取高质量的时频信息,然后使用DenseNet充分提取并整合其浅层纹理特征和深层抽象特征,最后使用softmax层进行分类。实验结果表明,对比所提方法与传统机器学习算法以及神经网络分类算法,本文方法在仅使用3个EEG通道的情况下,在Acc和Kappa值方面都取得了最好的性能,证明了本文方法的有效性,具有开发移动BCI的巨大潜力。

## 参考文献

- [1] WOLPAW J R, BIRBAUMER N, HEETDERKS W J, et al. Brain-computer interface technology: a review of the first international meeting[J]. IEEE Transactions on Rehabilitation Engineering, 2000, 8(2): 164-173.
- [2] CHAUDHARY U, BIRBAUMER N, RAMOS-Murguialday A. Brain-computer interfaces for communication and rehabilitation[J]. Nature Reviews Neurology, 2016, 12(9): 513-525.
- [3] PANDARINATH C, NUYUJUKIAN P, BLABE C H, et al. High performance communication by people with paralysis using an intracortical brain-computer interface[J]. Elife, 2017, 6: e18554.
- [4] HE Y, EGUREN D, AZORÍN J M, et al. Brain-machine interfaces for controlling lower-limb powered robotic systems[J]. Journal of Neural Engineering, 2018, 15(2): 021004.
- [5] HOOSHYAR D, MALVA L, YANG Y, et al. An adaptive educational computer game: Effects on students' knowledge and learning attitude in computational thinking[J]. Computers in Human Behavior, 2021, 114: 106575.
- [6] VAN de Laar B, GÜRKÖK H, BOS D P O, et al. Experiencing BCI control in a popular computer game[J]. IEEE Transactions on Computational Intelligence and AI in Games, 2013, 5(2): 176-184.
- [7] GARCIA-Moreno F M, BERMUDEZ-Edo M, GARRIDO J L, et al. Reducing response time in motor imagery using a headband and deep learning[J]. Sensors, 2020, 20(23): 6730.
- [8] JIN J, XIAO R, DALY I, et al. Internal feature selection method of CSP based on L1-norm and Dempster-Shafer theory[J]. IEEE Transactions on Neural Networks and Learning Systems, 2020, 32(11): 4814-4825.
- [9] SHAKEEL A, NAVID M S, ANWAR M N, et al. A review of techniques for detection of movement intention using movement-related cortical potentials[J]. Computational and Mathematical Methods in Medicine, 2015, 2015.
- [10] LI S, JIN J, DALY I, et al. Feature selection method based on Menger curvature and LDA theory for a P300 brain-computer interface[J]. Journal of Neural Engineering, 2022, 18(6): 066050.
- [11] SHIBASAKI H, HALLETT M. What is the Bereitschaftspotential?[J]. Clinical Neurophysiology, 2006, 117(11): 2341-2356.

- [12] MÜLLER-PUTZ G R, SCHWARZ A, PEREIRA J, et al. From classic motor imagery to complex movement intention decoding: the noninvasive Graz-BCI approach [J]. *Progress in Brain Research*, 2016, 228: 39-70.
- [13] FARINA D, JENSEN W, AKAY M. *Introduction to Neural Engineering for Motor Rehabilitation* [M]. John Wiley & Sons, 2013: 1999.
- [14] ZHANG Y, NAM C S, ZHOU G, et al. Temporally constrained sparse group spatial patterns for motor imagery BCI [J]. *IEEE Transactions on Cybernetics*, 2018, 49(9): 3322-3332.
- [15] PFURTSCHELLER G, NEUPER C. Motor imagery and direct brain-computer communication [J]. *Proceedings of the IEEE*, 2001, 89(7): 1123-1134.
- [16] ZHAO X, JIN J, XU R, et al. A regional smoothing block sparse Bayesian learning method with temporal correlation for channel selection in P300 speller [J]. *Frontiers in Human Neuroscience*, 2022, 16: 875851.
- [17] JIN J, LI S, DALY I, et al. The study of generic model set for reducing calibration time in P300-based brain-computer interface [J]. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2019, 28(1): 3-12.
- [18] JIN J, WANG Z, XU R, et al. Robust similarity measurement based on a novel time filter for SSVEPs detection [J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2021: 1-10.
- [19] HAUMANN N T, PARKKONEN L, KLIUCHKO M, et al. Comparing the performance of popular MEG/EEG artifact correction methods in an evoked-response study [J]. *Computational Intelligence and Neuroscience*, 2016: 1-10.
- [20] BLANKERTZ B, TOMIOKA R, LEMM S, et al. Optimizing spatial filters for robust EEG single-trial analysis [J]. *IEEE Signal Processing Magazine*, 2007, 25(1): 41-56.
- [21] MIAO Y, JIN J, DALY I, et al. Learning common time-frequency-spatial patterns for motor imagery classification [J]. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2021, 29: 699-707.
- [22] ANG K K, CHIN Z Y, WANG C, et al. Filter bank common spatial pattern algorithm on BCI competition IV datasets 2a and 2b [J]. *Frontiers in Neuroscience*, 2012, 6: 39.
- [23] ZHANG C, KIM Y K, ESKANDARIAN A. EEG-inception: an accurate and robust end-to-end neural network for EEG-based motor imagery classification [J]. *Journal of Neural Engineering*, 2021, 18(4): 046014.
- [24] ZHAO X, ZHANG H, ZHU G, et al. A multi-branch 3D convolutional neural network for EEG-based motor imagery classification [J]. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2019, 27(10): 2164-2177.
- [25] SCHIRRMEISTER R T, SPRINGENBERG J T, FIEDERER L D J, et al. Deep learning with convolutional neural networks for EEG decoding and visualization [J]. *Human Brain Mapping*, 2017, 38(11): 5391-5420.
- [26] LAWHERN V J, SOLON A J, WAYTOWICH N R, et al. EEGNet: a compact convolutional neural network for EEG-based brain-computer interfaces [J]. *Journal of Neural Engineering*, 2018, 15(5): 056013.
- [27] DAI G, ZHOU J, HUANG J, et al. HS-CNN: a CNN with hybrid convolution scale for EEG motor imagery classification [J]. *Journal of Neural Engineering*, 2020, 17(1): 016025.
- [28] ZHANG X, WU D. On the vulnerability of CNN classifiers in EEG-based BCIs [J]. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2019, 27(5): 814-825.
- [29] JIN J, LIU C, DALY I, et al. Bispectrum-based channel selection for motor imagery based brain-computer interfacing [J]. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2020, 28(10): 2153-2163.
- [30] ZHOU M, TIAN C, CAO R, et al. Epileptic seizure detection based on EEG signals and CNN [J]. *Frontiers in Neuroinformatics*, 2018, 12: 95.
- [31] WEI Z, ZOU J, ZHANG J, et al. Automatic epileptic EEG detection using convolutional neural network with improvements in time-domain [J]. *Biomedical Signal Processing and Control*, 2019, 53: 101551.
- [32] ZENG M, ZHANG X, ZHAO C, et al. GRP-DNet: A gray recurrence plot-based densely connected convolutional network for classification of epileptiform EEG [J]. *Journal of Neuroscience Methods*, 2021, 347: 108953.
- [33] O'TOOLE J. *Discrete quadratic time-frequency distributions: definition, computation, and a newborn electroencephalogram application* [D]. Brisbane, Australia: The University of Queensland, 2009.
- [34] LIU C, JIN J, DALY I, et al. SincNet-based hybrid neural network for motor imagery EEG decoding [J]. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2022, 30: 540-549.
- [35] O'TOOLE J M, MESBAH M, BOASHASH B. A new discrete analytic signal for reducing aliasing in the discrete Wigner-Ville distribution [J]. *IEEE Transactions on Sig-*

- nal Processing, 2008, 56(11): 5427-5434.
- [36] EMARA H M, ELWEKEIL M, TAHA T E, et al. Hilbert transform and statistical analysis for channel selection and epileptic seizure prediction [J]. Wireless Personal Communications, 2021, F116: 3371-3395.
- [37] TOOLE J M O, BOASHASH B. Fast and memory-efficient algorithms for computing quadratic time-frequency distributions [J]. Applied and Computational Harmonic Analysis, 2013, 35(2): 350-358.
- [38] ALAZRAI R, HOMOUD R, ALWANNI H, et al. EEG-based emotion recognition using quadratic time-frequency distribution [J]. Sensors, 2018, 18(8): 2739.
- [39] HUANG G, LIU Z, VAN Der Maaten L, et al. Densely connected convolutional networks [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 4700-4708.
- [40] OROVIĆ I, STANKOVIĆ S, BEKO M. On the use of Gini coefficient for measuring time-frequency distribution concentration and parameters selection [J]. Mathematical Problems in Engineering, 2022: 1-8.
- [41] WANG Q, MENG C, WANG C. High-concentration time-frequency analysis for multi-component nonstationary signals based on combined multi-window Gabor transform [J]. Engineering Computations, 2022, 39(4): 1234-1273.
- [42] RIYAD M, KHALIL M, ADIB A. MI-EEGNET: A novel convolutional neural network for motor imagery classification [J]. Journal of Neuroscience Methods, 2021, 353: 109037.
- [43] KAM T E, SUK H I, LEE S W. Non-homogeneous spatial filter optimization for Electroencephalogram (EEG)-based

motor imagery classification [J]. Neurocomputing, 2013, 108: 58-68.

#### 作者简介



**金 晶(通讯作者)** 男,1981年生,上海人。华东理工大学信息科学与工程学院教授,博士生导师,主要研究方向为脑机交互技术、脑卒中康复技术、外骨骼控制、智能感知系统。

E-mail: jinjingat@gmail.com



**杨益雕** 男,1997年生,江西上饶人。华东理工大学信息科学与工程学院,控制科学与工程专业硕士研究生,主要研究方向为脑机接口技术、深度学习。

E-mail: y30200995@mail.ecust.edu.cn



**孙 浩** 男,1995年生,吉林人。华东理工大学信息科学与工程学院,控制科学与工程专业博士研究生,主要研究方向为运动想象脑信息解码、深度学习。

E-mail: y20180078@mail.ecust.edu.cn



**王行愚** 男,1944年生,上海人。华东理工大学信息科学与工程学院教授,博士生导师,主要研究方向为控制理论及应用、智能控制与智能系统、脑控及其应用、系统工程。

E-mail: xywang@ecust.edu.cn