

DOI:10.19651/j.cnki.emt.2105832

# 基于 STFT 的卷积神经网络在 MI-EEG 分类中的应用

彭禹 宋耀莲

(昆明理工大学 昆明 650050)

**摘要:** 在脑机接口技术中,针对运动想象脑电信号(MI-EEG)由于其自身的高度非平稳性导致传统的手工提取特征困难和分类准确率低的问题,在特征提取和分类这方面进行了研究和探索,设计了一种基于短时傅里叶变换(STFT)的浅层卷积神经网络(SCNN)。通过对网络模型的超参数进行调优实验,该网络模型在2008年BCI竞赛的公开数据集2b上进行测试得到了较好的分类结果,分类准确率达到80.23%,远高于没有进行STFT处理的CNN方法61.04%的准确率。在相同的测试指标下优于传统的机器学习分类方法(CSP+SVM)73.52%的分类准确率,同时也比同类型的深度学习方法(CNN-SAE, 77.60%)更具有优势。

**关键词:** 运动想象;卷积神经网络;短时傅里叶变换;深度学习

中图分类号: TP183 文献标识码: A 国家标准学科分类代码: 520.2099

## Application of convolutional neural network based on STFT in MI-EEG classification

Peng Yu Song Yaolian

(Kunming University of Science and Technology, Kunming 650050, China)

**Abstract:** In the brain-computer interface technology, MI-EEG is difficult to extract features by hand and the classification accuracy is low due to its high degree of non-stationarity. This paper conducts corresponding research and exploration to solve this problem, and designs a shallow convolutional neural network (SCNN) based on short time Fourier transform (STFT). Through tuning experiments on the hyperparameters of the network model, the model was tested on the public data set 2b of the BCI competition in 2008 and got good classification results. The classification accuracy rate reached 80.23%, which was much higher than the accuracy of the CNN model without STFT processing of 61.04%. Under the same test indicators, it is superior to the traditional machine learning classification method (CSP+SVM) with a classification accuracy of 73.52%, and it also has more advantages than the same type of deep learning method (CNN-SAE, 77.60%).

**Keywords:** motor imagery; convolutional neural network; short time Fourier transform; deep learning

## 0 引言

脑机接口(brain computer interface, BCI)是一种可以通过特定的设备来采集大脑信号,并对脑电信号进行分析解码,最后将解码得到的信息编码成计算机能够识别的指令来达到控制外接设备的技术,其中关于肢体的运动想象(motor imagery, MI)是BCI技术领域中的一个非常重要的研究课题。运动想象是指受试者的大脑产生某一个肢体动作意图,但是不会真正的去执行这个动作,常见的包括有左右手、双脚以及舌头的运动想象。这项技术对医疗康复,神经病学,心理学以及日常生活娱乐等领域都有着重要的意义。2014年,法国科学院院士尼科莱利斯设计的脑机接口

系统,在巴西世界杯上,让截肢残疾者完成了首开球,震惊了全世界。2015年,来自新加坡信息通讯科学研究所神经与生物医学技术系的Tung等<sup>[1]</sup>设计出了一套针对手臂运动康复的治疗系统,帮助中风患者进行康复治疗,并取得了不错的效果。我国在脑机接口领域虽然起步较晚,但是近几年飞速发展,仅2016年的相关专利数量就达到了207个。2020年,北京理工大学研究团队的Li等<sup>[2]</sup>开发了一款脑-控移动机器人智能系统,用户可以通过脑电信号来控制机器人的移动,同时他们设计了一个鲁棒的非线性预测控制器,提高了系统的安全性和导航性能。2021年年初,天津大学的神经工程团队的两位成员通过自主研发的“哪吒”多脑协控智能臂系统仅依靠二人的脑电信号,便让机械臂

收稿日期:2021-02-23

• 36 •

一笔一画地写出了一个福字。我国的脑机交互技术已从医疗领域开始往更多的领域内发展壮大,呈现出多元化的特点。

运动想象脑电信号 (MI-EEG) 的分类识别任务,主要涉及到原始数据的预处理技术,特征提取和分类方法。目前在特征提取的方法上主要有基于时域的分析方法,如概率密度函数,均值,方差等<sup>[3]</sup>。基于时频域的分析方法,如快速傅里叶变换、小波变换、希尔伯特变换等。利用时频域的分析方法来进行特征提取虽然可以从时间和频率两个维度对脑电信号特征进行表示,优于基于时域的方法,但是仍然存在忽视信号的空间分布特性的问题。Kim 等<sup>[4]</sup>提出了基于小波变换,短时傅里叶变换和希尔伯特变换的 3 种加权功率谱密度差分法,加权功率谱密度差分法考虑到了脑电的空间分布特性,先是计算所有信号通道的功率谱密度,通过检查非平稳性和类别的可分离性来选择最优的通道,在 BCI 竞赛 IV 的两个数据集上提升了分类精度,但是需要人为去设计最优通道选择的指标。近些年基于空域的分析方法,比如共空间模式 (CSP) 及其一系列的衍生方法也经常被用于 MI-EEG 信号的特征提取<sup>[5-6]</sup>,但是采用这种基于空域的方法来进行特征提取破坏了脑电信号在时间域上的变化特性。目前随着深度学习的研究热度不断上升,许多基于深度学习的脑电信号特征提取和分类方法被提出,相较于传统的机器学习方法,MI-EEG 的分类精度得到不错的提升<sup>[7]</sup>。比如 Kumar 等<sup>[8]</sup>提出,首先使用 CSP、CSSP、FBCSP 等方法来对信号进行特征提取,最后再使用深度神经网络 (DNN) 进行分类,通过实验证明 CSP-DNN 的方法具有更小的误差,且优于传统的机器学习方法。Tabar 等<sup>[9]</sup>则提出了一种 CNN-SAE 框架,通过时频分析的方法将脑电信号转换成时频图来作为深度学习模型的输入,最后的分类准确率达到 77.60%。

根据目前的研究现状,以传统的特征提取方式加上深度学习的复合型框架更适用于脑电信号的分类,因此本文设计了一种基于短时傅里叶变换 (short time Fourier transform, STFT) 的浅层卷积神经网络 (SCNN) 的网络架构。采用时频分析的方法,同时将多通道的数据融成一个多维的特征表示,作为卷积神经网络的输入。充分考虑到了信号在时间,频率和空间多个维度上的特性,更有助于分类效果的提升。为了进一步验证该方法的优点,本文还通过对传统的机器学习方法和未经过特征多维表示的卷积神经网络进行了对比。本文将在第 1 节中介绍基于 STFT 的浅层卷积神经网络 (SCNN) 的网络架构以及实验的流程;第 2 节中主要介绍实验的具体细节和实验结果分析;第 3 节结论与展望。本文的研究目的旨在通过搭建一种新型的深度学习模型来提升 BCI 系统的分类准确率,并进一步推进基于运动想象的 BCI 技术的发展。希望在不久的将来 BCI 技术能够帮助更多的周围神经系统受损的患者实现自主地肢体运动,同时进一步提升人类的生活质量。

## 1 实验原理

### 1.1 基于短时傅里叶变换的深度学习的方法

在计算机视觉,语音识别等领域,深度学习在特征提取和分类上有着比传统机器学习更明显的优势,而且已经被广泛地应用于各种领域<sup>[10-12]</sup>。但是,目前针对使用卷积神经网络的方法对 MI-EEG 进行分类的技术还不是很成熟,本文从卷积神经网络可以自主提取特征的特点出发提出了基于 STFT 的浅层卷积神经网络 (SCNN)。该网络模型在输入层上将每个通道的一维的脑电数据通过 STFT 转化成具有时间和频谱信息的二维数组,接着再将其输入到卷积神经网络的隐藏层做特征提取。图 1 展示了本文针对 MI-EEG 分类的处理流程。整个模型框架主要分为如下两个大的模块:

1) 预处理模块:在这个模块下主要是对原始信号进行 6~26 Hz 带通滤波和 ICA 处理,剔除原始信号中的眼电干扰。

2) SCNN 模块:在这个模块下首先将预处理好的脑电信号送入浅层卷积神经网络的输入层 (STFT 层) 进行短时傅里叶变换得到关于时频信息的三维矩阵,然后将包含时频信息的矩阵输入卷积层进行特征提取,最后将得到的特征向量通过全连接层进行分类。

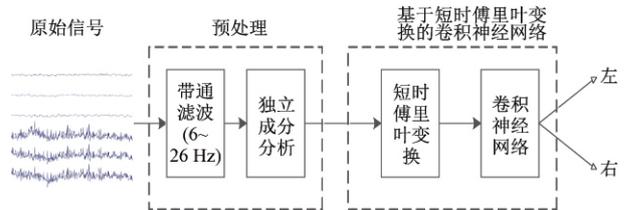


图 1 实验流程

### 1.2 SCNN 网络结构

2012 年,在 ImageNet 大赛上 Krizhevsky 等<sup>[13]</sup>提出的卷积神经网络 AlexNet,该网络以 15.2% 的错误率夺得冠军,掀起了卷积神经网络的发展热潮,AlexNet 网络模型主要有 5 个卷积层,3 个池化层以及 3 个全连接层构成<sup>[14]</sup>。本文参考了 AlexNet 的卷积层的结构特点设计了一个基于 STFT 的浅层卷积神经网络。具体的网络结构如图 2 所示。

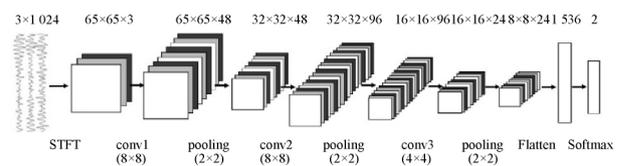


图 2 SCNN 网络结构

本文设计的网络模型的输入层即图 2 中的 STFT 层,该层的任务是将大小为  $3 \times 1024$  的三通道脑电信号进行短时傅里叶变换,并得到一个大小为  $65 \times 65 \times 3$  (time,

frequency, channels)的时频张量,最后将其输入到后面的卷积层中。STFT的定义式(1)如下:

$$STFT_s(t, \omega) = \int_{-\infty}^{+\infty} s(\tau)h(\tau - t)e^{-i\omega\tau} d\tau \quad (1)$$

$$P_{STFT}(t, \omega) = \frac{1}{2\pi} |STFT_s(t, \omega)|^2 = \frac{1}{2\pi} \left| \int_{-\infty}^{+\infty} s(\tau)h(\tau - t)e^{-i\omega\tau} d\tau \right|^2 \quad (2)$$

式中: $s(\tau)$ 表示原始信号在  $t$  时刻附近  $\tau$  时段的信号; $h(\tau - t)$ 表示时窗函数; $e^{-i\omega\tau}$ 为复指数函数; $STFT_s(t, \omega)$ 表示对原始信号进行时频变换后得到一个时频矩阵; $t, \omega$ 分别表示时间和频率; $P_{STFT}(t, \omega)$ 表示信号在  $t$  时刻的能量密度频谱。

SCNN的中间隐藏层主要由卷积层、池化层和归一化层组成。卷积层的主要功能是提取一个局部区域的特征,通常由多个卷积核(convolution kernel)和一个非线性的激活函数组成。为了充分提取到三维时频数据的局部特性,本文在卷积层使用了二维的等宽卷积,并通过如下公式对上一层的输入进行特征提取。

$$a_s^l = f(a^{l-1} \times w_s^l + b_s^l) \quad (3)$$

式中: $l, s$ 分别代表第几个卷积层,第几个卷积核; $a_s^l$ 代表第  $l$  个卷积层的第  $s$  个卷积核输出的特征映射; $a^{l-1}$ 为第  $l$  层的输入特征矩阵即上一层的输出; $w_s^l, b_s^l$ 分别代表第  $l$  层的第  $s$  个卷积核的权重参数和标量偏置; $f(\cdot)$ 表示非线性激活函数,本文使用的是修正线性单元(rectified linear unit, ReLU)。此网络模型中一共包含了 3 个卷积层,卷积层 conv1 和卷积层 conv2 的卷积核的大小均为  $8 \times 8$ ,卷积核的数目分别为 48、96,卷积层 conv3 则使用了 24 个大小为  $4 \times 4$  的卷积核。同时每个卷积层后面都增加了一个归一化层,以防止在做完卷积操作后破坏了输入数据的分布空间。

池化层(pooling layer)又叫子采样层(subsampling layer),其作用是对卷积层的输出进行特征选择,并同时保证特征的局部不变性,从而缩减参数的数量,通常会置于卷积层之后。主要方式有最大池化(maximum pooling)和平均池化(mean pooling),本文使用的是最大池化即在子采样区域内选择最大的活性值作为这个区域的表示,其具体公式如下:

$$y^l = \max(x_{(i,j)}^l) \quad (4)$$

式中: $x_{(i,j)}^l$ 表示当前层采样区域内所有元素的集合; $\max(\cdot)$ 表示对区域内的元素取最大值; $(i, j)$ 表示采样区域的大小,在本文中所有池化层采样区域的大小均为  $2 \times 2$ 。

最后的输出层由 1 个 Flatten 层和 1 个包含两个神经元的 Softmax 层组成,Flatten 层主要的作用是将输入为  $8 \times 8 \times 24$  大小的特征矩阵展平成 1 个一维的特征向量,并将其送入到下一层。Softmax 层则是最后的分类层,它首先将输入的特征向量通过一个权重矩阵进行加权求和,最

后通过 Softmax 激活函数的响应来得到两个分类的概率值也就是网络模型的输出值。公式如下:

$$z_i = a^{l-1} \times w_i + b_i \quad (5)$$

$$Softmax(z_i) = \frac{e^{z_i}}{\sum_{i=1}^m e^{z_i}} \quad (6)$$

式中: $z_i$ 表示通过 Softmax 层进行加权求和后的输出值; $i$ 表示该层的第  $i$  个输出节点; $a^{l-1}$ 为 Flatten 层输出的特征向量; $w_i, b_i$ 为第  $i$  个输出节点的权重参数矩阵和偏置; $m$ 代表输出节点的个数即分类的类别数。采用指数形式的 Softmax 函数不仅将分类的结果都赋予一个概率值,同时约束所有的输出节点的和为 1,还能将输出的概率数值之间的距离拉的更大,更便于分类。

## 2 实验与结果分析

### 2.1 数据集介绍

本文使用的是 BCI 竞赛 IV 2b 数据集,该数据集由奥地利格拉茨工业大学提供,一共包含了 9 个受试者 3 个电极通道(C3、Cz 和 C4)的 EEG 数据,以及 3 个通道的 EOG 数据,采样频率为 250 Hz。每个受试者共进行了 5 次运动想象实验,前两次实验屏幕不提供反馈,在后 3 次实验中屏幕会根据受试者的实验情况提供对应的反馈<sup>[15]</sup>。每次试验开始时都有一个固定十字和一个附加的短声音警告音(1 kHz, 70 ms),几秒钟后,一个视觉提示出现了 1.25 s,受试者必须在 4 s 内想象相应的手部动作<sup>[15]</sup>。图 3 所示为关于受试者每次做运动想象的具体时间安排,表 1 所示为每个受试者关于做运动想象的实验样本数量。

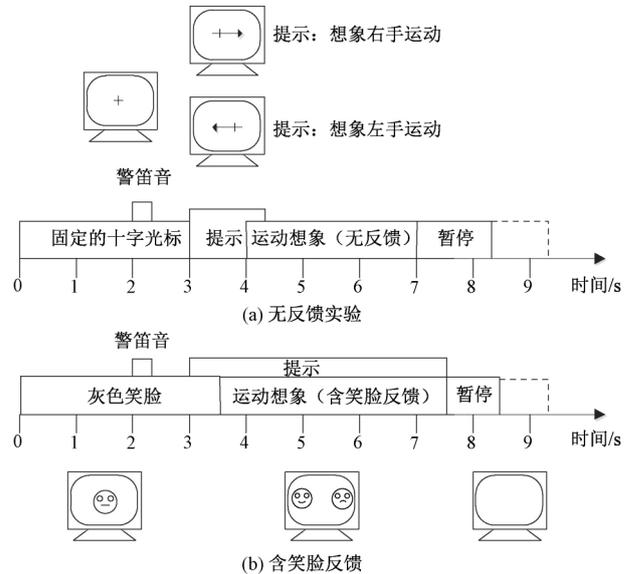


图 3 数据采集计时方案

表 1 BCI 竞赛 IV 2b 数据集

受试者	s1	s2	s3	s4	s5	s6	s7	s8	s9
样本量	720	680	720	740	740	720	720	760	720

## 2.2 数据预处理

EEG 信号在采集过程中存在其他信号的干扰,因此对原始信号进行滤波处理是一个非常重要的过程,它可以帮助神经网络更精确地提取到相关频段的特征信息。同时研究发现受试者在进行关于肢体的某个运动想象的时候,在大脑感觉运动皮层区域会出现脑电信号能量增强或者减弱的现象,这两种现象分别被称为事件相关同步化(event-related synchronization, ERS)和事件相关去同步化(event-related desynchronization, ERD), ERD/ERS 现象主要波段出现在  $\mu$  节律(8~12 Hz)和  $\beta$  节律(16~31 Hz)上<sup>[16-18]</sup>。由于手部的运动想象诱发的 ERD/ERS 现象的  $\mu$  节律和  $\beta$  节律通常比较集中出现在 10~11 Hz,以及 20~24 Hz 这两个频段范围,因此本文在进行预处理阶段选择 6~26 Hz 这个较为合理的频带对原始的 MI-EEG 信号进行带通滤波<sup>[19]</sup>。做完带通滤波后本文对其进行了独立成分分析(ICA)用以剔除眼电信号的干扰,最后对发生左右手运动想象的数据段进行事件提取,本文在发生运动想象前一秒开始提取,一共提取到了 4.1 s 的数据段。

在预处理操作后,选择 C3 和 C4 通道的数据对其进行短时傅里叶变换得到了在这个时间段内的时频功率图,如图 4 所示。可以发现当受试者进行左手的运动想象时,C3 通道的脑电数据在 3~4 s 内的 5~20 Hz 频段上的功率谱值出现升高,即发生 ERS,相反 C4 通道在同样的时间段内相同的频率段上功率谱值则出现降低,即发生了 ERD。同样在做右手的运动想象实验时,也会出现与左手类似的现象。

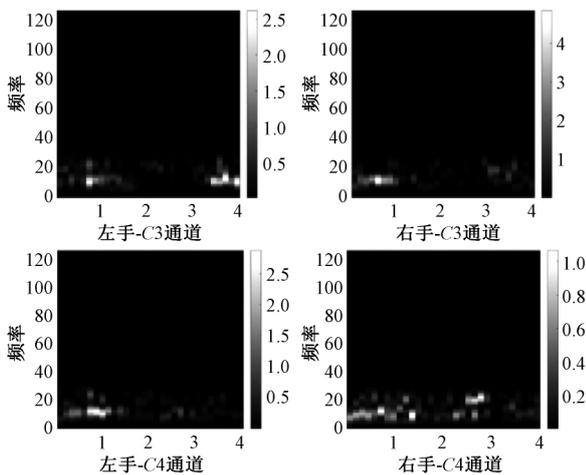


图 4 左、右手运动想象时频功率图

## 2.3 参数设置

卷积神经网络的深度,卷积核的大小,学习率等超参数对网络模型的效果有较大的影响,因此本文首先通过控制变量的方法来改变超参数,验证了不同的卷积层数和卷积核大小对模型分类效果的影响。由表 2 可知在卷积层层数相同的情况下,卷积核的大小对网络模型分类效果影响

没有太大的差距,相反卷积层的数量有着较大的影响<sup>[20]</sup>。随着卷积层数目从 2 层增加到 3 层,网络的平均分类识别有所提高,但是增加到 4 层的时候,分类准确率又下降了。这是由于本文的数据集规模比较小,当网络的深度超过一定数量时会发生严重的过拟合现象<sup>[20]</sup>。所以本文将卷积层的数量设置为 3 层,卷积核大小采用  $8 \times 8$ 。

表 2 卷积核大小与卷积层层数对分类准确率的影响

层数	%			
	$4 \times 4$	$6 \times 6$	$7 \times 7$	$8 \times 8$
1	65.31	65.10	66.23	68.07
2	74.06	75.25	74.02	76.55
3	79.75	79.15	78.12	80.02

由于本文的数据样本比较少,在进行网络训练的过程中随着训练次数的增加发生了数据过拟合的情况。针对数据过拟合的情况,本文在后面两个卷积层中分别添加的 Dropout 层,并将 Dropout 设置为 0.5。Dropout 是一种防止神经网络在训练过程中,过度拟合训练数据的方法<sup>[21]</sup>。它在模型训练过程中对某一层所有神经元的输出,按照指定的概率进行随机丢弃,使其节点的权重参数在本次的训练中不进行更新。在添加了 Dropout 层后,该网络在测试集上的表现有了明显的提升。

## 2.4 结果分析

针对本文基于运动想象分类问题,采用了分类准确率(*accuracy*)作为分类效果的评价指标。分类准确率指的是对于整个测试集的数据,通过训练集来训练得到的最优分类器在测试数据时能够正确分类的样本数与总的样本数的比率,如式(7)所示。

$$accuracy = \frac{(TP + TN)}{(TP + FP + TN + FN)} \times 100\% \quad (7)$$

其中,  $TP$  为真阳性,  $TN$  为真阴性,  $FP$  为假阳性,  $FN$  为假阴性。

考虑到输入层对脑电信号进行短时傅里叶变换时输出的数据维度受时间窗以及 *nfft* 参数的影响,进而影响到卷积层提取特征的效果。因此本文对比了几种不同的数据输入维度对模型分类准确率的影响大小,图 5 所示为 9 个受试者的数据在不同输入维度下的准确率。从图 5 中可以发现,几种数据维度中将大小为  $65 \times 65 \times 3$  的数据作为卷积层的输入,得到的分类效果更好。

为了验证 SCNN 网络模型具有的优势,本文使用相同的数据集对传统的特征提取方法和分类方法(CSP+SVM)以及未进行 STFT 的 CNN 进行分类测试并进行对比,结果如表 3 所示,从表中可以发现本文的 SCNN 比传统的分类方法更具有优势。同时本文与文献[9]中使用一个卷积层和多层的 SAE(CNN+SAE)方法进行比较,本文将 STFT 整合进神经网络的分类方法也具有一定的优势。

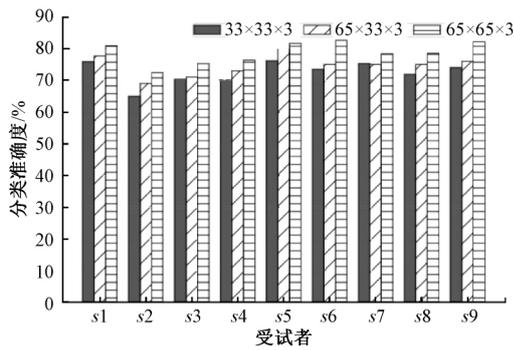


图 5 输入数据维度分析

表 3 本文方法与其他分类方法的准确率比较

方法	CNN	SCNN	CNN+SAE <sup>[9]</sup>	CSP+SVM
准确率/%	61.04	80.23	77.60	73.52

### 3 结 论

本文先将一维的运动想象数据通过短时傅里叶变换得到具有时间和频率信息的时频矩阵后,再输入到卷积层中进行特征提取,和传统的机器学习和统计学方法相比,不依赖于手工的特征工程,从而大大增加了提取特征和分类的效率。同时本文将短时傅里叶变换整合到了网络模型的输入层,网络模型更具有完整性,进一步地提升了模型训练的效率,克服了由于 MI-EEG 信号的高度非平稳性导致特征提取困难的问题。通过实验分析表明本文的 SCNN 与传统方法和同类的深度学习方法相比,在 MI-EEG 分类识别上具有更好的性能表现。但是本文的网络模型泛化能力不足,后期将从迁移学习的角度出发来改进网络模型以提升模型的泛化能力。

#### 参考文献

- [1] TUNG S W, GUAN C, ANG K, et al. A measurement of motor recovery for motor imagery-based BCI using EEG coherence analysis[C]. 2015 10th International Conference on Information, Communications and Signal Processing (ICICS), IEEE, 2015: 1-5.
- [2] LI H, BI L, YI J. Sliding-mode nonlinear predictive control of brain-controlled mobile robots [J]. IEEE Transactions on Cybernetics, 2020: 1-13.
- [3] RAO R P N. Brain-computer interfacing: An introduction[M]. Cambridge University Press, 2013: 109-148.
- [4] KIM C, SUN J, LIU D, et al. An effective feature extraction method by power spectral density of EEG signal for 2-class motor imagery-based BCI[J]. Medical & Biological Engineering & Computing, 2018, 56(9): 1645-1658.
- [5] 马满振,郭理彬,苏奎峰.基于改进 CSP 算法的运动想

象脑电信号分类方法[J].计算机与现代化,2017(11): 23-28.

- [6] ANG 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): 1-9.
- [7] SAKHAVI S, GUAN C, YAN S. Learning temporal information for brain-computer interface using convolutional neural networks[J]. IEEE Transactions on Neural Networks and Learning Systems, 2018, 29(11): 5619-5629.
- [8] KUMAR S, SHARMA A, MAMUN K, et al. A deep learning approach for motor imagery EEG signal classification[C]. 2016 3rd Asia-Pacific World Congress on Computer Science and Engineering (APWC on CSE), IEEE, 2016: 34-39.
- [9] TABAR Y R, HALICI U. A novel deep learning approach for classification of EEG motor imagery signals [J]. Journal of Neural Engineering, 2017, 14(1): 016003.
- [10] 张顺,龚怡宏,王进军.深度卷积神经网络的发展及其在计算机视觉领域的应用[J].计算机学报,2019, 42(3):453-482.
- [11] 韩文静,李海峰,阮华斌,等.语音情感识别研究进展综述[J].软件学报,2014,25(1):37-50.
- [12] 吴睿曦,肖秦琨.基于深度网络和数据增强的多物体图像识别[J].国外电子测量技术,2019,38(5):86-90.
- [13] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks [J]. Advances in Neural Information Processing Systems, 2012, 25: 1097-1105.
- [14] 余琼芳,黄高路,杨艺,等.基于 AlexNet 深度学习网络的串联故障电弧检测方法[J].电子测量与仪器学报, 2019,33(3):145-152.
- [15] LEEB R, BRUNNER C, MÜLLER-PUTZ G R, et al. BCI competition 2008-graz data set B [J]. Graz University of Technology, Austria, 2008: 1-6.
- [16] PFURTSCHELLER G, DA SILVA F H L. Event-related EEG/MEG synchronization and desynchronization: Basic principles [J]. Clinical Neurophysiology, 1999, 110(11): 1842-1857.
- [17] PFURTSCHELLER G, NEUPER C, FLOTZINGER D, et al. EEG-based discrimination between imagination of right and left hand movement[J]. Electroencephalography and Clinical Neurophysiology, 1997, 103(6): 642-651.
- [18] 张力新,张珊珊,王坤,等.运动相关思维诱发脑电信息解码与应用综述[J].仪器仪表学报,2019,40(1):1-11.
- [19] 明东,王坤,何峰,等.想象动作诱发生理信息检测及其

应用研究:回顾与展望[J].仪器仪表学报,2014,35(9):1921-1931.

[20] 周飞燕,金林鹏,董军.卷积神经网络研究综述[J].计算机学报,2017,40(6):1229-1251.

[21] HINTON G E, SRIVASTAVA N, KRIZHEVSKY A, et al. Improving neural networks by preventing co-adaptation of feature detectors[J]. Computer Science, 2012, 3(4):212-223.

#### 作者简介

彭禹,硕士研究生,主要从事脑机接口(BCI)、深度学习方面的研究。

E-mail: 1822084857@qq.com

宋耀莲,博士,副教授,主要从事脑机接口(BCI)、无线通信、通信系统方面的研究。

E-mail: 39217149@qq.com