软件学报 ISSN 1000-9825, CODEN RUXUEW
Journal of Software,2019,30(10):3005−3016 [doi: 10.13328/j.cnki.jos.005782]
©中国科学院软件研究所版权所有.
E-mail: jos@iscas.ac.cn
http://www.jos.org.cn
Tel: +86-10-62562563

CNN 实现的运动想象脑电分类及人-机器人交互*

程时伟[1], 周桃春[1], 唐智川[2], 范 菁[1], 孙凌云[3], 朱安杰[1]

[1](浙江工业大学 计算机科学与技术学院,浙江 杭州 310023)

[2](浙江工业大学 设计艺术学院,浙江 杭州 310023)

[3](计算机辅助设计与图形学国家重点实验室(浙江大学),浙江 杭州 310058)

通讯作者: 范菁, E-mail: fanjing@zjut.edu.cn

摘 要: 基于脑电的脑机交互能帮助肢体运动障碍患者进行日常生活和康复训练,但是,由于脑电信号存在信噪比较低、个体差异性大等问题,导致脑电特征的提取与分类还需要进一步提高准确性和效率.因此,在减少脑电采集通道数目、增加分类数目的前提下,基于卷积神经网络对运动想象中的脑电信号进行分类.首先,基于已有方法进行探索实验,建立由3层卷积层、3层池化层和2层全连接层构成的卷积神经网络;然后针对想象左手、右手、脚的运动和静息态设计与开展了实验,获取了相关脑电数据;之后,利用脑电数据训练出基于卷积神经网络的分类模型,测试结果表明,该模型平均分类识别率达到了 82.81%,且高于已有的相关分类算法;最后,将已建立的分类模型应用于运动想象信号的在线分类,设计与开发了脑机交互应用原型系统,驱动人-机器人之间的实时交互,帮助用户利用运动想象控制仿人机器人的抬手、前进等运动状态.进一步的测试结果表明,机器人对用户控制命令的平均识别率达到了 80.31%,从而验证了所提方法可以对运动想象脑电数据进行较为精确的实时分类,可以促进脑机接口技术在人-机器人交互中的应用.

关键词: 运动想象;脑机接口;人机交互;深度学习;卷积神经网络

中图法分类号: TP242

中文引用格式: 程时伟,周桃春,唐智川,范菁,孙凌云,朱安杰.CNN 实现的运动想象脑电分类及人-机器人交互.软件学报, 2019,30(10):3005−3016. http://www.jos.org.cn/1000-9825/5782.htm

英文引用格式: Cheng SW, Zhou TC, Tang ZC, Fan J, Sun LY, Zhu AJ. CNN based motor imagery EEG classification and human-robot interaction. Ruan Jian Xue Bao/Journal of Software, 2019,30(10):3005−3016 (in Chinese). http://www.jos.org.cn/1000-9825/5782.htm

CNN Based Motor Imagery EEG Classification and Human-robot Interaction

CHENG Shi-Wei[1], ZHOU Tao-Chun[1], TANG Zhi-Chuan[2], FAN Jing[1], SUN Ling-Yun[3], ZHU An-Jie[1]

[1](School of Computer Science and Technology, Zhejiang University of Technology, Hangzhou 310023, China)

[2](School of Design, Zhejiang University of Technology, Hangzhou 310023, China)

[3](State Key Laboratory of CAD & CG (Zhejiang University), Hangzhou 310058, China)

Abstract: The electroencephalograph (EEG) driven brain-computer interaction can promote daily life and rehabilitation training for physically disabled people, however, EEG has several problems such as low signal-noise ratio, significant individual difference, and these problems result in the low accuracy and efficiency for EEG feature extraction and classification. In the context of reducing numbers of

* 基金项目: 国家重点研发计划(2016YFB1001403); 国家自然科学基金(61772468, 61702454, 61672451)

Foundation item: National Key Research & Development Program of China (2016YFB1001403); National Natural Science Foundation of China (61772468, 61702454, 61672451)

本文由“自然人机交互新进展”专题特约编辑田丰、喻纯推荐.

收稿时间: 2018-08-18; 修改时间: 2018-11-01; 采用时间: 2018-12-25; jos 在线出版时间: 2019-04-29

CNKI 网络优先出版: 2019-04-30 09:19:04, http://kns.cnki.net/kcms/detail/11.2560.TP.20190430.0918.006.html

© 中国科学院软件研究所 http://www.jos.org.cn

electrodes and increasing identified classes, this study proposed an approach to classify motor imagery (MI) EEG signal based on convolutional neural network (CNN). Firstly, based on existed approaches, experiments were conducted and the CNN was constructed with three convolution layers, three pooling layers, and two full-connection layers. Secondly, MI experiment was conducted with the imagination of left hand movement, right hand movement, foot movement, and resting state, and the MI EEG data were collected at the same time. Thirdly, the MI EEG data set were used to build the classification model based on CNN, and the experiment results indicate that the average accuracy of classification is 82.81%, which is higher than the related classification algorithms. Finally, the classification model was applied in the online classification of MI EEG, and a BCI prototype system was designed and implemented to drive the real-time human-robot interaction. The prototype system can help users to control motion states of the humanoid robot, such as raising hands, moving forward. Furthermore, the experimental results show that the average accuracy of robot controlling reaches to 80.31%, and it verifies the proposed approach not only can classify MI EEG data with high accuracy in real time, but also promote applications of human-robot interaction with BCI.

Key words: motor imagery; brain computer interface; human-computer interaction; deep learning; convolutional neural network

脑机接口(brain computer interface,简称 BCI)通过采集与提取大脑产生的脑电图信号(electroencephalogram,简称 EEG)来识别人的意图,基于 EEG 信号完成大脑与外部设备的信息传递与控制任务,进而恢复甚至增强人的身体运动和感知、认知能力[1].近年来,BCI 的相关研究得到了快速发展,但是,由于 EEG 信号具有非平稳、非线性、低信噪比等特点[2],在预处理、特征提取、多模式分类等方面尚有诸多难题有待解决,从而导致可实际应用的 BCI 系统较少.

通常,EEG 数据的处理和分析方法主要包括 3 个步骤:(1) 对 EEG 数据进行预处理;(2) 对预处理后的 EEG 数据进行特征提取;(3) 最后对提取的特征向量进行分类.这类方法的主要局限在于,需要依靠研究人员的先验知识进行复杂的数据预处理和特征提取,但在预处理过程中容易剔除可能隐含着有用特征的数据,而在特征提取过程中又容易忽略不同通道脑电信号间的关联特性[2],这些都会降低分类结果的准确性和可靠性.而卷积神经网络(convolutional neural network,简称 CNN)融合了无监督学习和有监督学习的优点,对高维特征向量具有较强的分类能力,能从原始数据中学习到有利于分类的特征,减少了人为因素对特征选择造成的主观性和不完备性.

因此,为了针对 EEG 信号进行便捷、精确的特征提取和分类,本文设计运动想象实验,在大幅减少脑电电极数量(传统方法需要 64 个或更多电极,本文只采用 14 个电极)、同时增加运动想象分类任务(传统方法多数只支持 2 或 3 分类,本文支持左手、右手、脚和静息态的 4 分类)的情况下,提取被试用户的 EEG 数据,并采用 CNN 方法进行数据的特征学习与分类,结果表明,本文方法比已有运动想象分类算法具有更高的识别率.进而,基于本文方法设计与开发了一个基于 BCI 的人-机器人交互系统,实验结果表明,该原型系统能快速且准确地判断出用户的运动想象指令,控制机器人完成相应的动作,进一步验证了本文方法的实用性.

1 相关工作

近年来,国内外学者提出的 EEG 信号特征提取方法主要包括 3 类:(1) 基于时域分析的方法,如采用均值、方差、概率密度函数等[1];(2) 基于频域分析的方法,如 Pfurtscheller[3,4]提出的基于 mu 节律的频域分析方法并应用在运动想象任务中、Zhou 等人[4]将小波包分解(wavelet packet decomposition,简称 WPT)和独立成分分析(independent component analysis,简称 ICA)应用在运动想象数据特征提取中;(3) 基于空域分析的方法,如 Samek 等人[5]提出将共空间模式(common spatial pattern,简称 CSP)应用在脑机接口中.另一方面,也可采用机器学习的方法,如孙会文等人[6]使用支持向量机(support vector machine,简称 SVM)方法对经过希尔伯特黄变换(Hilbert-Huang transform,简称 HHT)的 EEG 数据进行分类;张毅等人[7]使用自回归模型(autoregressive model,简称 AR)对特征提取后的 EEG 信号进行分类;刘伯强等人[8]使用神经网络的反向传播(back propagation,简称 BP)算法进行 EEG 数据分类.

CNN 是一种带有卷积结构的深度神经网络,卷积结构可以利用空间结构关系,减少需要学习的参数数量,从而提高反向传播算法的训练效率,不仅可以防止过拟合,还能降低神经网络的复杂度[9].LeCun 等人[10]采用神经

© 中国科学院软件研究所 http://www.jos.org.cn

网络误差反向传播算法进行手写数字的识别,在网络结构设计中加入下采样与权值共享,大幅减少了神经网络的参数量.此外,为了尽量保证数据的平移、尺度、畸变不变性,LeCun 等人[11]设计了局部感受野和共享权重,提出用于字符识别的卷积神经网络 LeNet-5,在银行的手写数字识别系统中取得了较好的结果.近年来,CNN 也被应用于 EEG 信号的特征提取和分析.例如,Hubert 等人[12]基于 CNN 对脑电事件相关电位(event-related potential,简称 ERP)中的 P300 成分进行分类识别,实现字符拼写,并在 BCI 竞赛的公共数据集(http://www.bbci.de/competition/iii/)上进行实验,结果表明,其识别率最高可以达到 95.5%;蔡邦宇[13]使用 CNN 模型对视觉诱发电位(visual evoked potential,简称 VEP)进行时域和空域特征分析,结果表明,平均受试者工作特征(receiver operating characteristic,简称 ROC)曲线面积比 SVM 方法提高了 4.4%;唐智川等人[14]基于 CNN 对左手和脚的运动想象脑电信号进行分类识别,平均识别率为 88.75%±3.42%,并将其应用到基于 BCI 的外骨骼应用原型系统中;王卫星等人[15]基于 CNN 进行左、右手二分类动作和单手的 3 分类动作识别,识别精度比原有方法分别提高了 4%和 8%.

上述相关研究验证了 CNN 方法在脑电信号处理中应用的可行性,但鲜有研究能在有效减少脑电电极数目的同时,还能对 4 个及以上的运动想象分类问题进行精确的识别,并最终设计与开发出相应的 BCI 原型应用系统,实现精确的人-机器人实时交互.

2 基于 CNN 的脑电信号分类方法

2.1 脑电信号采集

脑电信号的采集采用 10-20 国际标准导联放置[1],将鼻根和枕外粗隆相连接,在冠状位把鼻根、外耳孔和枕状粗隆相连接,中点为头顶(即 Cz).通过 Cz 将两个连线各分为 2 个 10%和 4 个 20%的距离.本文采用 Emotiv System 公司(http://www.pstnet.com/eprime.cfm)的 Emotiv EPOC 脑电仪,其包含 14 个电极(如图 1 所示),采样频率为 128Hz.

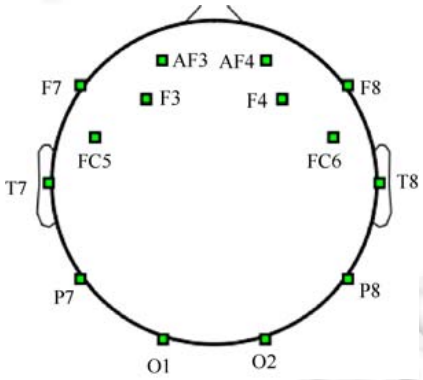

Fig.1 Electrodes configuration of Emotiv EPOC

图 1 Emotiv EPOC 脑电仪的电极位置分布

2.2 脑电实验范式

在实验过程中采用如图 2 所示的实验范式[16].

本文招募了 7 名被试用户(5 男 2 女,年龄在 24 岁~27 岁之间)进行实验,实验对象均是右利手,并且均是第 1 次参加本文中的脑电实验.每个被试用户需要完成 560 次基于提示的实验,每次实验持续 9s 时间,前 2s 屏幕显示“rest”字,用户休息放松,之后在屏幕中央出现一个“十”字,提示被试用户实验即将开始;从 4s~8s,屏幕上的“十”字变为随机产生的想象左手、想象右手、想象脚和静息状态的提示.为了保持提示呈现后同步获得相应的脑电信号,实验范式程序采用 Eprime 实现,在每次实验开始时,向脑电仪发送一个 marker 标记信号,并在实验结束后进行离线数据分析时,根据 marker 来分割相关脑电数据.此外,为了防止被试用户产生疲劳,影响运动想象脑电

© 中国科学院软件研究所 http://www.jos.org.cn

实验的效果,被试用户每进行 20 次实验后就休息 3 分钟.此外,为了防止被试用户在运动想象过程中产生实际的肢体动作,影响脑电数据质量,在正式实验之前对被试用户进行了运动想象训练,并在正式实验过程中剔除了产生实际肢体动作的实验数据.

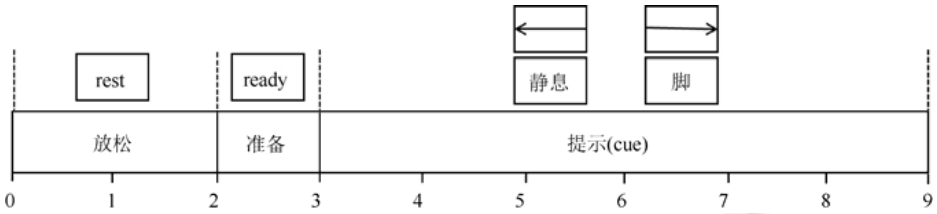

Fig.2 MI experimental paradigm

图 2 运动想象实验范式

2.3 卷积神经网络结构

本文设计的用于运动想象脑电数据处理的卷积神经网络如图 3 所示,包括 1 个输入层(input 层)、3 个卷积层(C1、C2、C3 层)、3 个池化层(S1、S2、S3 层)和 1 个输出层(output 层).

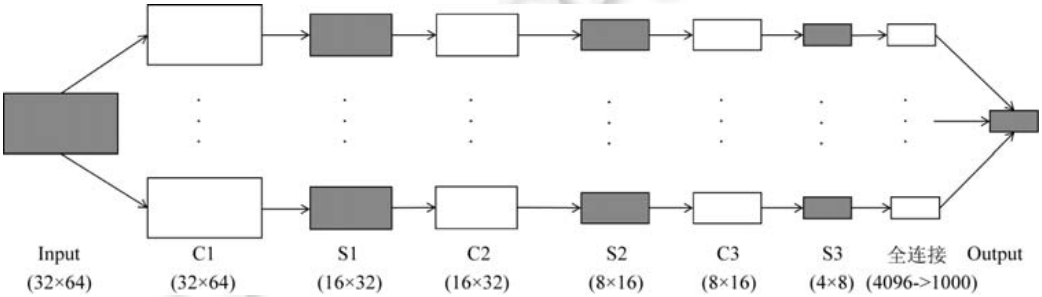

Fig.3 CNN structure for MI EEG data analysis

图 3 用于运动想象脑电数据分析的卷积神经网络结构

2.3.1 卷积层构造

卷积层主要进行特征提取,采用线性滤波器和非线性激活函数.具体地,对输入数据应用多个滤波器,并用一个输入参数进行多个类型的特征提取.基于上一层的输入,采用以下公式来提取下一层的特征:

$$x_j^l = f\left(\sum_{i\in M_j} x_i^{l-1} \times k_{ij}^l + b_j^l\right) \tag{1}$$

式(1)中,x_j^l 为卷积层 l 层第 j 个通道的输出;$f(\cdot)$为激活函数,使用修正线性单元(rectified linear unit,简称 ReLu);M_j 为选择输入的集合,x_i^{l-1} 为卷积层 l–1 层第 i 通道的输出,k_{ij}^l 为卷积核矩阵,“*”在式(1)中表示卷积符号,b_j^l 为偏置值.

在本文中,输入数据为 F3、F4、FC5 和 FC6 通道的脑电数据,数据维度为 4×512.通常,CNN 模型处理脑电数据的输入数据格式为通道数×时间点,本文初期研究通过上述数据格式进行神经网络构建与优化,但效果不佳,见表 1,平均分类识别率只能达到 50%左右.为此,本文作了进一步改进,考虑到卷积神经网络对图像数据分析具有较好的识别效果,并在处理块状数据时具有更大优势,所以本文对运动想象数据也采用类似图像的分块处理.表 1 所示,对不同的数据输入维度进行实验,结果表明,在 4 种分类状态下,32×64 的分块输入方式的平均分类识别率最高.因此,本文将原输入数据的维度设置为 32×64.

进一步地,建立起本文的卷积层结构:C1 层,输入 3 个通道,输入维度为 32×64,输出 32 个通道;C2 层,输入 32 个通道,输入维度为 16×32,输出 64 个通道;C3 层,输入 64 个通道,输入维度为 8×16,输出 128 个通道.在卷积层中,步长为 1,为了保证数据是完整、有效的信息,使用宽卷积,所以,卷积层的输出维度不变,并使用 dropout 防止数

© 中国科学院软件研究所 http://www.jos.org.cn

据的过拟合,同时采用 ReLu 代替传统神经网络中的 tanh 和 sigmod 函数,加快网络训练速度.

Table 1 Average accuracy of classification under different input dimensions (%)

表 1 不同脑电输入维度的平均分类识别率对比(%)

维度 / 被试	4×512	8×256	16×128	32×64	64×32
P1	55.00	59.64	62.14	82.50	73.21
P2	53.93	63.57	67.14	85.00	76.79
P3	56.78	60.36	64.64	79.64	74.64
P4	61.42	57.14	58.21	83.93	72.86
P5	55.71	58.21	66.07	82.86	72.86
P6	53.93	55.71	63.57	82.14	66.57
P7	55.36	59.64	64.29	83.57	67.14

2.3.2 池化层构造

池化层(pooling layer)也称为子采样层(subsample layer),用于缩减数据的规模,同时保持特征的局部不变性.实现子采样的方法有很多种,最常见的是最大值合并、平均值合并及随机合并.本文使用平均值合并进行子采样,公式如下:

$$x_j^l = f\left(\beta_j^l down(x_j^{l-1}) + b_j^l\right) \tag{2}$$

在式(2)中,$f(\cdot)$为激活函数,$down(\cdot)$为下采样函数,它基于输入特征 x_j^{l-1} 进行 $n\times n$ 的子采样,然后对子采样数据求均值. β_j^l 为 pooling 层中第 l 层第 j 个通道的权重系数, b_j^l 为采样层的偏置项.

在本文中,池化层的输入是同一网络层中卷积层的输出:S1 层,输入维度为 32×64,输出维度为 16×32;S2 层,输入维度为 16×32,输出维度为 8×16;S3 层,输入维度为 8×16,输出维度为 4×8.在上述池化层中,步长为 2,子采样为 2×2,所以,输出的两个维度同时减半.为了保证数据的完整、有效性,使用宽卷积,同时采用 ReLu 加快网络训练速度.

2.3.3 全连接层构造

将经过多层卷积和池化的运动想象 EEG 二维数据的特征图拼接为一维特征作为全连接网络的输入.全连接层 l 的输出可通过对输入加权求和并通过激活函数的响应得到,公式如下:

$$x^l = f(w^l x^{l-1} + b^l) \tag{3}$$

公式(3)中,$w^l x^{l-1}+b^l$ 称为全连接层 l 的净激活,它由前一层输出特征图 x^{l-1} 进行加权和偏置后得到.w^l 是全连接网络的权重系数,b^l 是全连接层 l 的偏置项.

3 实验结果与分析

3.1 CNN参数设置

参数设置对 CNN 模型的效果影响很大.下面对本文 CNN 模型的参数设置进行详细阐述.

(1) 批大小(batchsize):在深度学习中,一般采用随机梯度下降(stochastic gradient descent,简称 SGD),即每次在训练集中获取 batchsize 个样本进行训练.当数据集和计算量不大时,batchsize 采用全数据集,其有两个优点:第一,由全数据集确定的方向能够更好地代表样本总体,从而更准确地朝向极值所在的方向;第二,全数据集可以使用弹性传播(resilient propagation,简称 Rprop)算法[17],基于梯度符号并且有针对性的单独更新各权值,解决了多批次模型训练中,各批次训练权重的梯度值差别较大,选取一个全局的学习速率困难的问题.但是,对于较大的数据集合,上述优点却变成了缺点:第一,随着数据集和计算量的增大以及内存的限制,一次载入所有数据的可行性较低;第二,如果以 Rprop 方式迭代,由于各个 batch 之间的采样差异性,各批次梯度修正值将相互抵消,无法实现修正效果,所以需要采用均方根传播(root mean square propagation,简称 RMSProp)算法[18].考虑内存的利用率和训练完一次全数据集的迭代次数[19],结合实验经验,本文的 batch size 取值为 8,即每次取 8 个样本进行训练.

(2) 学习速率:运用梯度下降算法进行优化,在权重更新时,学习速率就是梯度项前的系数.根据 Bengio[20]对

© 中国科学院软件研究所 http://www.jos.org.cn

梯度训练的总结,如果学习速率太小,则收敛过慢;如果学习速率太大,则会导致代价函数震荡.本文基于 Bengio 提出的方法,对学习速率赋一个缺省值,取 0.01,结果发现,代价函数产生震荡,所以对学习速率进行调整,设为 0.001,最终提高了识别率.

(3) 权重和偏置:根据 LeCun 等人[21]提出的权重和偏置参数调试建议,初始化时使用截断正态分布,加入轻微噪声,打破对称性,防止零梯度问题,本文中设置正态分布标准差为 0.1.

(4) dropout:这是指在深度学习网络的训练过程中,对于神经网络单元,按照一定概率将其暂时从网络中丢弃.Hintion 提出 dropout 方法[22],在训练模型时随机地让某些隐含节点的权重不工作,这些节点暂时可以认为不是网络结构的一部分,但是它们的权重得以保留下来(暂时不更新),下次样本输入时可能重新工作.由于本文的数据样本相对较少,为了防止在训练模型时出现过拟合,采用 dropout 方法.本文在训练时卷积层和池化层的 dropout 采用 0.8,全连接层的 dropout 采用 0.5;在测试时,dropout 都为 1,更新所有权重.实验发现,在同样迭代次数下,采用 dropout 测试样本的错误率明显低于没采用 dropout 的测试样本.

3.2 实验结果与分析

实验采集了 7 个被试用户(用 P1~P7 表示)的脑电数据,每个被试采集 560 次数据,为了得到可靠稳定的模型,通过交叉验证(cross-validation)方法分离出 50%的训练样本和 50%的测试样本.权重使用截断正态分布,标准差为 0.1,学习率为 0.001.训练时卷积层和池化层的 dropout 采用 0.8,全连接层的 dropout 采用 0.5,在测试时,dropout 都为 1,脑电数据的输入层维度为 32×64.

分类问题需要对分类结果进行评价,评价指标一般包括识别率(accuracy)、精确率(precision)、召回率(recall rate)和 *F*-score,通过这些评价指标对分类模型的效果进行评估[14].4 个评价指标的定义分别如下.

(1) 识别率:对于总的测试数据集,通过训练数据训练的最优分类器对测试数据正确分类的样本数与总样本数之比:

$$accuracy=(TP+TN)/(TP+FP+TN+FN)\times 100\% \tag{4}$$

其中,*TP* 为真阳性,*TN* 为真阴性,*FP* 为假阳性,*FN* 为假阴性.

(2) 精确率:测试数据集中分类为真实正例样本数与分类为正例样本数之比:

$$precision=TP/(TP+FP)\times 100\% \tag{5}$$

(3) 召回率:分类为真实正例的样本数与所有真实正例的样本数之比:

$$recall=TP/(TP+FN)\times 100\% \tag{6}$$

(4) *F*-score:这是识别率的延伸,结合了精确率和召回率,具体计算如下:

$$F\text{-}score=2\times recall\times precision/(recall+precision) \tag{7}$$

本文从网络层数和时间序列两个方面对实验结果进行分析,具体阐述如下.

(1) 基于网络结构层数的结果分析

针对每个被试用户,分别构建不同层数的卷积神经网络,从而找到最好分类效果的卷积神经网络结构.受篇幅限制,这里没有列出所有不同层数的分类结果.表 2 是在相同条件下(4 种分类),分别构建 2~5 层卷积和池化结构的训练测试结果.可见,3 层的实验结果比其他层数的平均分类识别率更高,因此本文采用 3 层卷积和池化网络结构.

Table 2 Average accuracy of classification with different convolution and pooling layers (%)

表 2 不同卷积层和池化层 CNN 模型的平均分类识别率对比(%)

被试 \ 层数	2 层	3 层	4 层	5 层
P1	73.93	82.50	75.71	76.07
P2	70.71	85.00	76.79	73.57
P3	72.86	79.64	74.64	73.93
P4	76.79	83.93	76.07	75.71
P5	70.71	82.86	76.79	72.86
P6	67.14	82.14	72.86	72.86
P7	71.43	83.57	73.93	74.17

© 中国科学院软件研究所 http://www.jos.org.cn

(2) 基于时间序列的结果分析

将运动想象的脑电实验数据按时间段划分,每段时长 2s,作为输入数据.图 4 显示了在 0~6s 的 3 个时间段中,被试用户的平均分类识别率.

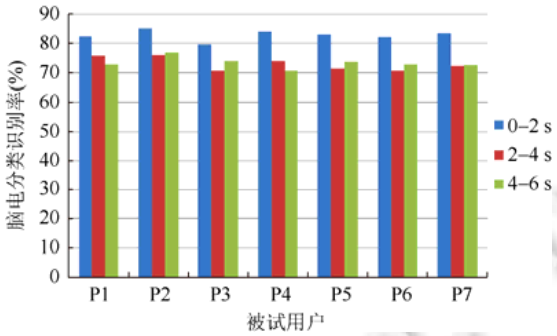

Fig.4 Classification results analysis based on temporal sequence

图 4 基于时间序列的分类结果分析

从图 4 可以看出,前 2s(第 0~第 2s)脑电数据的平均分类识别率达到最高,而最后 4s(第 2s~第 6s)脑电的数据平均分类识别率较低,说明在实验开始时,被试用户专注于运动想象实验,但是随着时间的推移,被试用户精神集中度降低,影响了脑电数据的质量,最终导致识别率降低.

因此,根据上述分析结果,选择原始输入数据维度为 32×64,卷积和池化 3 层网络,选择运动想象开始后 0~2s 的脑电数据进行分类预测.相关分类结果见表 3,对角线单元格内数值表示所有被试想象左手、右手、脚运动和静息状态时,其 EEG 信号被正确识别为相应类别的平均识别率(precision)和标准差.另外,非对角线单元格内的数值表示 EEG 信号被错误识别为其他类别的平均识别率和标准差.

Table 3 Confusion matrix of EEG classification results

表 3 脑电分类平均预测结果的混淆矩阵

		预测			
		左手运动	右手运动	脚运动	静息状态
实际	左手运动	**81.63%±6.30%**	6.53%±3.18%	7.35%±4.62%	4.49%±4.40%
	右手运动	8.37%±5.11%	**78.78%±6.03%**	9.39%±6.64%	3.47%±2.00%
	脚运动	5.51%±4.70%	3.27%±4.64%	**83.27%±5.82%**	5.92%±5.18%
	静息状态	3.47%±2.45%	3.06%±3.02%	5.92%±4.40%	**87.55%±5.20%**

此外,为了进一步验证本文方法的先进性,使用相同的实验数据,采用 4 种已有方法(SAE、CSP+SVM、softmax 和 WPT+SVM)与本文方法进行对比.这 4 种方法都需要先对脑电数据进行预处理,包括去除眼电伪迹、滤波等,再经过特征提取,最后用分类器进行分类.与已有方法相比,本文基于 CNN 方法的平均分类识别率最高,结果见表 4.

Table 4 Average accuracy of classification with different approaches

表 4 各种分类方法的平均识别率对比

分类方法	CNN(本文)	SAE	CSP+SVM	softmax	WPT+SVM
平均分类识别率	82.81%±6.14%	72.96%±7.73%	68.62%±5.98%	61.33%±7.79%	76.28%±6.31%

4 应用实例

4.1 原型系统框架设计

本文基于运动想象的人-机器人交互应用场景如图 5 所示,被试用户要求坐在屏幕前,左、右手和身体保持自然的放松姿势,并且佩戴好 Emotiv 脑电仪,固定好电极位置(参考电极为 P3 和 P4,采样频率为 128Hz),确保能

© 中国科学院软件研究所 http://www.jos.org.cn

够正常采集脑电信号.本文使用 NAO(https://www.softbankrobotics.com/emea/en/robots/nao)仿人机器人,放在被试用户正前方位置.

Fig.5 Senario of human-robot interaction based on BCI

图 5 基于 BCI 的人-机器人交互应用场景

每个被试用户需要完成 560 次基于方向箭头和文字等视觉提示(cue)的实验,其中,想象左手运动、右手运动、脚运动和静息状态(不作任何运动想象)各 140 次,每次实验之间有 3s 的时间间隔,每完成 20 次实验之后休息 5 分钟,防止被试用户过于疲劳影响实验结果.如图 6 所示,在单次运动想象结束后,采用本文前述训练出来的分类模型进行实时分类,NAO 机器人根据分类结果产生相应的动作.例如,用户进行左手运动想象,如果分类算法正确识别为左手运动,NAO 机器人则抬起左手.类似地,右手的运动想象对应于 NAO 机器人抬起右手;脚的运动想象对应于 NAO 机器人向前走动;静息态对应于 NAO 机器人保持静止,且播放语音“你现在很安静”.此外,为了防止被试用户在运动想象过程中产生实际的肢体动作,影响脑电数据质量,与第 2.2 节的处理方法一样,对被试用户进行了训练,并剔除了无效数据.

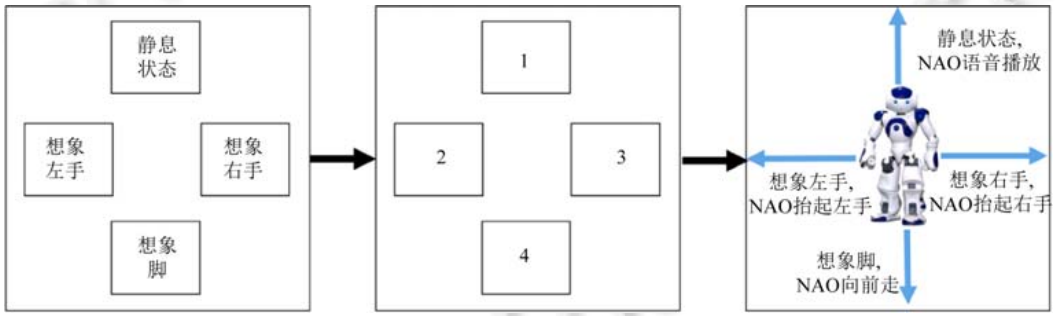

Fig.6 The mapping bettwen MI classification results and robot actions

图 6 运动想象分类结果与机器人交互动作的映射关系

该系统的框架如图 7 所示,被试用户根据屏幕上随机出现的向左、向右箭头,以及“脚”和“静息”的文字开始进行运动想象.当被试用户开始进行运动想象时,Emotiv 开始采集脑电数据,通过 TCP/IP 协议将获取的脑电数据发送给系统主机进行预处理,然后利用前文训练的分类模型对脑电数据进行分类,再将分类结果转换为控制命令,并通过 TCP/IP 发送给 NAO 机器人,驱动机器人的电机和扬声器等组件完成相应的操作.同时,在屏幕上给出分类结果的视觉反馈(如“向左”这样的文字).

© 中国科学院软件研究所 http://www.jos.org.cn

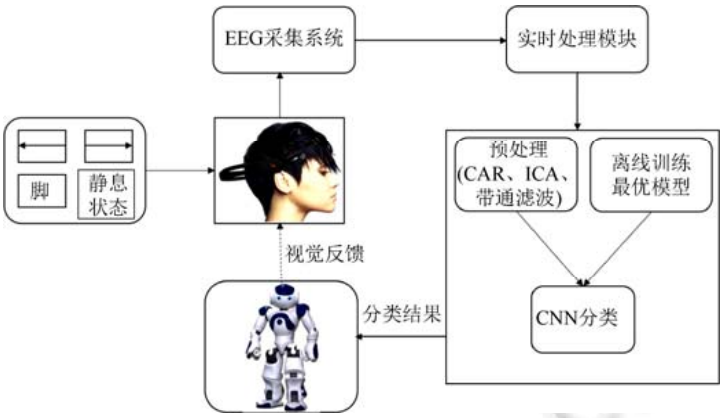

Fig.7 Framework of the prototype system

图 7 原型系统框架

4.2 原型系统开发与实验

4.2.1 系统功能

原型系统主要包括脑电数据采集、脑电数据分析和机器人控制这 3 个主要功能模块,采用 Python 编程语言实现.(1) 脑电数据采集模块:主要采集和分发系统所需的脑电数据,基于 socket 实现,开发工具包采用 pycharm;(2) 数据分析模块:主要进行原始数据的预处理、数据训练以及分类模型的求解和存储,其中的数值分析采用 Python 的科学计算工具包 numpy,并采用了 keras 和 Tensorflow 框架进行分类算法的开发;(3) 机器人控制模块:主要接收数据分析模块输出的分类结果,并对 NAO 机器人进行相应的控制,具体开发采用 NAO 的 pynaoqi-python SDK 开发包.此外,系统图形用户界面的开发基于 Python GUI 的 Tkiner 编程工具,它可以良好地运行在绝大多数平台上,实现跨平台的效果.

4.2.2 实验结果分析

在用户与机器人的交互过程中,针对 4 类运动想象的精确率、召回率和 *F*-score 这 3 个评估指标进行分析和对比,见表 5.采用分类方法×运动想象类别的方差分析(ANOVA)评估两者对分类结果的影响.置信水平为 95%.ANOVA 结果显示,分类方法对分类结果具有显著影响($p<0.05$),而运动想象的类别则对分类结果不具有显著影响($p>0.05$).进一步地,采用 Tukey Post Hoc 检验对分类方法进行两两比较,本文基于 CNN 模型分类方法的 3 个指标(精确率、召回率和 *F*-score)与其他 4 种方法相比均具有显著差异性($p<0.05$),且本文 3 个指标的值更大,因此,本文方法具有更好的指标性能,进一步验证了其在实际交互应用中的有效性和先进性.

4.2.3 讨 论

如表 6 所示,进一步将本文方法与已有文献的方法进行了对比(由于已有文献没有进行与机器人交互等应用的在线测试,故此处只取本文第 3 节的结果进行对比).分析发现,参考文献[18]采用了 64 通道、1 000Hz 采样频率的 Active Two 脑电仪采集脑电信号,并使用 CNN 对想象左手和想象脚进行分类,最终所有被试用户的平均识别率(accuracy)达到 89.51%±2.95%;参考文献[24]采用 60 通道、采样频率为 250Hz 的 Neuroscan 脑电仪,使用 CNN 方法对 4 类运动想象信号进行分类,最终分类识别率达到 85.04%±4.26%.虽然这两项研究比本文的平均识别率稍高,但脑电信号采集通道数目是本文的 4 倍左右(本文通道数只有 14 通道),而本文实验操作方便且设备价格便宜.此外,参考文献[18]的分类数目只有 2 类,而本文有 4 类.类似地,参考文献[23]采用了 64 通道、160Hz 采样频率的 Biosemi 脑电仪,使用 CNN 对 9 类运动想象进行分类,但效果不佳,平均分类识别率只有 45%.

基于上述分析,从肢体障碍患者使用的便携性、运动想象分类数目、分类识别率、设备价格和实验过程复杂度等方面进行了综合考虑,分析结果表明,本文方法具有更好的实用性和可行性.

© 中国科学院软件研究所 http://www.jos.org.cn

Table 5 Precision, recall rate and *F*-score of MI classification with different approaches across all subjects

表 5 所有被试用户在不同方法下运动想象分类的精确率、召回率和 *F*-score

		精确率			召回率			*F*-score		
CNN	左手运动	0.804 8	±	0.058 0	0.585 5	±	0.047 5	0.675 2	±	0.024 3
	右手运动	0.806 7	±	0.052 8	0.527 2	±	0.047 8	0.636 2	±	0.039 9
	脚运动	0.776 8	±	0.038 5	0.604 1	±	0.031 7	0.679 2	±	0.028 8
	静息状态	0.822 7	±	0.064 1	0.594 2	±	0.042 2	0.687 4	±	0.021 4
SAE	左手运动	0.693 6	±	0.064 9	0.466 8	±	0.057 5	0.557 3	±	0.059 7
	右手运动	0.723 2	±	0.080 9	0.444 5	±	0.059 5	0.547 4	±	0.052 7
	脚运动	0.746 0	±	0.061 8	0.443 8	±	0.045 3	0.554 2	±	0.037 4
	静息状态	0.742 7	±	0.044 4	0.510 9	±	0.062 1	0.602 8	±	0.042 8
CSP+SVM	左手运动	0.654 5	±	0.029 5	0.396 0	±	0.040 0	0.492 5	±	0.035 2
	右手运动	0.632 5	±	0.061 3	0.377 1	±	0.028 6	0.471 6	±	0.032 6
	脚运动	0.666 6	±	0.061 2	0.394 3	±	0.037 3	0.494 1	±	0.036 8
	静息状态	0.688 6	±	0.041 8	0.381 8	±	0.046 8	0.490 3	±	0.046 2
softmax	左手运动	0.581 3	±	0.027 5	0.332 3	±	0.031 0	0.422 6	±	0.030 9
	右手运动	0.571 9	±	0.036 3	0.294 3	±	0.027 6	0.387 7	±	0.025 3
	脚运动	0.599 2	±	0.066 1	0.321 9	±	0.039 9	0.417 1	±	0.040 7
	静息状态	0.598 3	±	0.058 5	0.331 5	±	0.059 8	0.424 7	±	0.057 4
WPT+SVM	左手运动	0.693 4	±	0.034 4	0.515 2	±	0.029 1	0.590 6	±	0.026 0
	右手运动	0.703 1	±	0.038 9	0.505 4	±	0.041 6	0.586 9	±	0.033 1
	脚运动	0.758 5	±	0.037 0	0.497 3	±	0.046 7	0.600 1	±	0.042 2
	静息状态	0.796 4	±	0.027 5	0.474 5	±	0.045 3	0.592 9	±	0.030 4

Table 6 Comparison between proposed approach of this paper and others

表 6 本文方法与已有文献的对比

方法	实验设备	被试人数和实验次数	运动想象类型	分类识别率(%)
参考文献[18]	Active Two 64 通道脑电仪,采样频率为 1 000Hz	4 人,每人 560 次实验	想象动左手、想象动脚(2 类)	89.51±2.95
参考文献[23]	Biosemi 64 通道脑电仪,采样频率为 160Hz	109 人,每人进行 14 次实验	静息态、动左手、动右手、想象动左手、想象动右手、动双手、动双脚、想象动双手、想象动双脚(9 类)	45
参考文献[24]	Neuroscan 60 通道脑电仪,采样频率为 250Hz	3 人,每人进行 240 次实验	想象动左手、想象动右手、想象动脚、想象动舌头(4 类)	85.04±4.26
本文	Emotiv 14 通道脑电仪,采样频率为 128Hz	7 人,每人进行 560 次实验	想象动左手、想象动右手、想象动脚和静息态(4 类)	82.81±6.14

进一步地,与其他分类方法(SAE、CSP+SVM、softmax 和 WPT+SVM)相比较,本文采用的 CNN 方法虽然在识别率上具有一定的优势,但也存在一些不足:(1) 由于 CNN 网络的复杂性以及需要调用较多的参数,需要使用较大的数据集进行训练,否则容易导致过拟合,因此本文对每个被试用户进行了 560 次实验,而这容易导致被试用户产生疲劳,从而会影响实验效果.(2) 计算成本较高,因为模型训练需要耗费较多时间、调参过程较复杂等.所以,本文方法在实际应用中还应该进一步提高模型训练效率,简化计算过程,改善用户体验感受.

5 结论与展望

本文将 CNN 用于运动想象脑电信号的分类识别,相关实验结果表明了本文方法能够从原始脑电数据中自动学习特征,并能对 4 种运动想象状态进行较精确的分类,平均分类识别率达到了 82.81%.

此外,本文也存在一定的局限性,下一步将从以下几个方面继续开展研究:(1) 完善脑电实验范式,采集更复杂的脑电数据,将 CNN 模型用于识别更复杂的脑电信号;(2) 进一步优化 CNN 的网络结构,调整合适参数,并结合其他深度学习方法,如深度置信网络(deep belief network,简称 DBN)来进一步提高分类识别率;(3) 面对实际应用中日益涌现的海量、复杂脑电数据,考虑使用 spark 架构实现分布式处理,进一步提高脑机交互应用的实时性.

© 中国科学院软件研究所 http://www.jos.org.cn

References:

[1] Rajesh PNR. Brain-computer interfacing: An introduction. In: Proc. of the Invasive BCIs. 2013. 109–148.

[2] He Q, Du S, Zhang YY, *et al*. Classification of motor imagery based on signgle-channel frame and multi-channel frame. Chinese Journal of Scientific Instrument, 2018,39(9):20–29 (in Chinese with English abstract).

[3] Pfurtscheller G, Brunner C, Schlogl A, *et al*. Mu rhythm (de)synchronization and EEG single-trial classification of different motor imagery tasks. Neuroimage, 2006,31(1):153−159.

[4] Zhou ZX, Wan BK. Wavelet packet-based independent component analysis for feature extraction from motor imagery EEG of complex movements. Clinical Neurophysiology, 2012,123(9):1779−1788.

[5] Samek W, Vidaurre C, Müller KR, *et al*. Stationary common spatial patterns for brain-computer interfacing. Journal of Neural Engineering, 2012,9(2):026013.

[6] Sun HW, Fu YF, Xiong X, *et al*. Study on EEG pattern recognition based on HHT motor imagery. Acta Automatica Sinica, 2015,41(9):1686−1692 (in Chinese with English abstract).

[7] Zhang Y, Yang L, Li M, *et al*. EEG recognition of motor imagery based on AR and SVM. Journal of Huazhong University of Science and Technology (Natural Science Edition), 2011,(S2):103−106 (in Chinese with English abstract).

[8] Liu BQ, Zhang ZW, Liu ZG, *et al*. An algorithm research of EEG classification on motor imagery EEG. Journal of Biomedical Engineering Research, 2007,(1):20–23, 32 (in Chinese with English abstract).

[9] Chang L, Deng XM, Zhou MQ, *et al*. Convolutional neural networks in image understanding. Acta Automatica Sinica, 2016,42(9):1300−1312 (in Chinese with English abstract).

[10] LeCun Y, Boser B, Denker JS, Henderson D, Howard RE, Hubbard W, Jackel LD. Backpropagation applied to hand-written zip code recognition. Neural Computation, 1989,1(4):541−551.

[11] LeCun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. Proc. of the IEEE, 1998,86(11): 2278−2324.

[12] Cecotti H, Gräser A. Convolutional neural networks for P300 detection with application to brain-computer interfaces. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2011,33(3):433–445.

[13] Cai BY. Temporal-spatial analysis on event-related potential in face recognition and its application in rapid face image retrieval by single-trail detection [Ph.D. Thesis]. Hangzhou: Zhejiang University, 2015 (in Chinese with English abstract).

[14] Tang ZQ, Zhang KJ, Li C, *et al*. Motor imagery classification based on deep convolutional neural network and its application in exoskeleton controlled by EEG. Chinese Journal of Computers, 2017,40(6):1367−1378 (in Chinese with English abstract).

[15] Wang WX, Sun SQ, Li C, *et al*. Recognition of upper limb motion intention of EEG signal based on convolutional neural network. Journal of Zhejiang University (Engineering Science), 2017,51(7):1381−1389 (in Chinese with English abstract).

[16] Graimann B, Huggins JE, Schlogl A, *et al*. Detection of movement-related patterns in ongoing single-channel electrocorticogram. IEEE Trans. on Neural Systems & Rehabilitation Engineering, 2003,11(3):276–281.

[17] Riedmiller M, Braun H. A direct adaptive method for faster backpropagation learning: The RPROP algorithm. In: Proc. of the IEEE Int'l Conf. on Neural Networks. 1993. 586–591.

[18] Tieleman T, Hinton G. Lecture 6.5—RmsProp: Divide the gradient by a running average of its recent magnitude. 2018. https://www.coursera.org/learn/neural-networks

[19] Abdulnabi NZT, Altun O. Batch size for training convolutional neural networks for sentence classi ication. Journal of Advances in Technology and Engineering Studies, 2016,2(5):156−163.

[20] Bengio Y. Practical recommendations for gradient-based training of deep architectures. In: Neural Networks: Tricks of the Trade. Berlin, Heidelberg: Springer-Verlag, 2012. 437–478.

[21] LeCun YA, Bottou L, Orr GB, *et al*. Efficient backprop. In: Neural Networks: Tricks of the Trade. Berlin, Heidelberg: Springer-Verlag, 2012. 9–48.

[22] Hinton GE, Srivastava N, Krizhevsky A, *et al*. Improving neural networks by preventing co-adaptation of feature detectors. Computer Science, 2012,3(4):212–223.

[23] Dai RM. The motor imagery EEG classification based on deep learning [Ph.D. Thesis]. Beijing: Beijing Institute of Technology, 2015 (in Chinese with English abstract).

© 中国科学院软件研究所 http://www.jos.org.cn

[24] Fan ML. Feature extraction and classification of motor imagery EEG signals based on convolutional neural network [Ph.D. Thesis]. Zhengzhou: Zhengzhou University, 2017 (in Chinese with English abstract).

附中文参考文献:

[2] 何群,杜硕,张园园,等.融合单通道框架及多通道框架的运动想象分类.仪器仪表学报,2018,39(9):20–29.
[6] 孙会文,伏云发,熊馨,等.基于 HHT 运动想象脑电模式识别研究.自动化学报,2015,41(9):1686–1692.
[7] 张毅,杨柳,李敏,等.基于 AR 和 SVM 的运动想象脑电信号识别.华中科技大学学报(自然科学版),2011,(S2):103–106.
[8] 刘伯强,张振旺,刘忠国,等.一种运动想象脑电分类算法的研究.生物医学工程研究,2007,(1):20–23,32.
[9] 常亮,邓小明,周明全,等.图像理解中的卷积神经网络.自动化学报,2016,42(9):1300–1312.
[13] 蔡邦宇.人脸识别中单次 ERP 时空特征分析及其快速检索的应用[博士学位论文].杭州:浙江大学,2015.
[14] 唐智川,张克俊,李超,等.基于深度卷积神经网络的运动想象分类及其在脑控外骨骼中的应用.计算机学报,2017,40(6):1367–1378.
[15] 王卫星,孙守迁,李超,等.基于卷积神经网络的脑电信号上肢运动意图识别.浙江大学学报(工学版),2017,51(7):1381–1389.
[23] 戴若梦.基于深度学习的运动想象脑电分类[博士学位论文].北京:北京理工大学,2015.
[24] 范明莉.基于卷积神经网络的运动想象脑电信号特征提取与分类[博士学位论文].郑州:郑州大学,2017.

程时伟(1981－),男,湖北黄石人,博士,教授,博士生导师,CCF 高级会员,主要研究领域为人机交互,普适计算,协同计算.

范菁(1969－),女,博士,教授,博士生导师,CCF 杰出会员,主要研究领域为服务计算,虚拟现实,人机交互.

周桃春(1992－),男,硕士,主要研究领域为人机交互.

孙凌云(1981－)男,博士,教授,博士生导师,CCF 专业会员,主要研究领域为人工智能,设计智能,信息与交互设计.

唐智川(1987－),男,博士,副教授,CCF 专业会员,主要研究领域为脑机接口,人机交互,人机工程,智能设计.

朱安杰(1994－),男,学士,主要研究领域为人机交互.

© 中国科学院软件研究所 http://www.jos.org.cn