

DOI: 10.19650/j.cnki.cjsi.J2007103

基于长短时记忆和卷积神经网络的手势肌电识别研究*

陈思佳, 罗志增

(杭州电子科技大学智能控制与机器人研究所 杭州 310018)

摘要:用表面肌电进行手势识别具有细节信息可选择性和抗外界干扰能力强的优势,但现有方法的适应性和识别准确性还不足。通过在卷积神经网络的基础上增加长短时记忆网络处理层,构筑手势识别模型,它能捕获手势动作过程的肌电时序特征,一定程度上减少了过拟合的现象。利用手势肌电丰富的时频域信息,提取手势肌电的小波包特征图像,并与手势肌电图像一起作为识别模型的输入数据,拓展手势识别模型中肌电信号的类别信息,同时在长短时记忆网络处理层与卷积神经网络层之间引入注意力机制,使得该模型能间接提高关键手势肌电通道的权重。实验结果证实本识别模型结合肌电两种特征输入的处理方法,与普通卷积神经网络模型以肌电图像输入的方法相比,识别准确率提升了4.25%。

关键词:肌电信号;小波包分解;长短时记忆神经网络;注意力机制;手势肌电识别

中图分类号:TP391.4 TH89 **文献标识码:**A **国家标准学科分类代码:**510.40

Research on gesture EMG recognition based on long short-term memory and convolutional neural network

Chen Sijia, Luo Zhizeng

(Intelligent Control and Robot Research Institute, Hangzhou Dianzi University, Hangzhou 310018, China)

Abstract:The gesture recognition using electromyography (EMG) has advantages of selective detail information and strong anti-interference ability. However, the adaptability and recognition accuracy of the existing methods are insufficient. By adding a long-term and short-term memory network layer on the basis of the convolutional neural network, a gesture recognition model is formulated. In this way, it can capture the EMG timing characteristics of the gesture, and the phenomenon of overfitting is reduced to a certain degree. The rich time-frequency domain information of EMG is utilized to extract the wavelet packet feature image of EMG. In addition, the input data of the recognition model are used with the EMG image to expand the category information of the EMG signal. Meanwhile, the attention mechanism is introduced between the time memory network processing layer and the convolutional neural network layer. Then, the model can indirectly increase the weights of the key gesture EMG channels. Compared with the method of ordinary convolutional neural network model using single EMG image, experimental results show that the recognition accuracy rate of the processing methods of EMG two feature inputs is improved by 4.25%.

Keywords:surface electromyography; wavelet packet; long short-term memory network; attention mechanism; gesture recognition

0 引言

表面肌电信号(surface electromyography, sEMG)是一种能够反映肌肉兴奋程度的生物电信号,目前被广泛应用于人机接口的研究。其中,基于sEMG的手势识别的研究最为广泛,可以被应用于智能假肢控制、康复训

练、机器人遥操作等。目前手势识别可以通过数据手套、图像识别、肌电信号等传感方式实现^[1]。数据手套虽然能较准确的识别手势,但装置较笨重,使用不便;基于视觉图像的手势识别,近年来也取得了不少改进,但该方法很容易受到环境如光照的影响。基于肌电的手势识别可避免遮挡等干扰,数据采集也相对简单,且随着传感器微型化、无线传输等技术的发展,便携式肌电手环得到越来

收稿日期:2020-11-10 Received Date: 2020-11-10

* 基金项目:国家自然科学基金(61671197)项目资助

越广泛的应用。肌电手势识别通过提取肌电信号中的时域和频域特征,结合常用的最近邻法(K-nearest neighbor, KNN),支持向量机(support vector machine, SVM),随机森林(random forest, RF),线性判别分析(linear discriminant analysis, LDA)等识别方法进行训练与识别,这些方法虽然也能取得相对不错的识别结果,但它们会随着识别对象的输入量和维数的增大而明显降低^[2]。利用肌电信号控制的商业化产品如肌电假肢^[3],只能识别有限的手部动作,或者基于动作顺序的控制,实用效果不理想。随着计算机算力的不断提高,深度学习被广泛的用于很多领域,除了图像识别,语音识别与翻译,自动驾驶等常见的领域之外还被用于设备的故障预测^[4],生物电信号的识别^[5]等,近年来,基于深度学习的肌电识别研究已证明比传统的机器学习方法在解决肌电识别率、普适性、可泛化能力弱等问题上有了较大的提高^[6-7]。

Atzori 等^[8]用一个四层的简单卷积网络结构发现,在 NinaPro 肌电数据集 DB1^[9]上进行 53 种手势动作的识别,实现了 66.59±6.40% 的准确率,优于传统的模式识别方法,并指出影响卷积神经网络结构识别效果的因素主要有网络结构,数据的预处理和参数优化,但并未看到作者进一步的研究成果。Tsinganos 等^[10]在 Atzori 研究基础上通过添加 dropout 层,改进了卷积神经网络模型的过拟合,取得了肌电识别率的进一步提高,但仍存在训练集的识别效果明显优于测试集。卷积神经网络的优势是减少了人工提取特征的过程,可以从原始数据中自主的学习相应的特征和规则,是一种端到端的学习过程。对于有规则的结构性数据,它有很强的特征提取能力,但对时间序列变化的特征处理能力偏弱。循环神经网络(recurrent neural network, RNN)区别于卷积神经网络的最大特点是它有记忆功能,即输出不仅仅依赖于当前的输入,也依赖于上一次输出,使之具有时间标记信息的处理能力。长短期记忆网络(long short-term memory, LSTM)是一种特殊的循环神经网络,它克服了循环神经网络梯度消失的问题,适合处理具有时间序列特征的生物电信号。Verma 等^[11]根据被试的心电图信息,采用 LSTM 方法,很成功地识别了被试者潜在的健康状况。

基于深度学习的肌电识别研究正在被越来越多的研究者关注和应用,提升肌电识别的准确性和稳定性是基于肌电的康复理论应用于康复实践的关键。

1 理论分析

表面肌电信号是由多个运动单元发放的动作电位序列,在皮肤表面呈现的时间上和空间上综合叠加的结果。肌电信号的幅值一般和肌肉运动力度成正比,且通常超前于相应的人体关节运动 30~150 ms 产生,这也是可以

把肌电信号用于预测人体运动意图的原因之一。

一方面,肌电信号是一种时序信号,单纯的卷积神经网络并不能有效的分析时序信号的特征和历史信息。另一方面不同的人体关节运动所对应的肌肉群是不同的,比如手部肘关节处前臂的屈曲是由肱桡肌起主要作用,腕部的伸展和外展是桡侧伸腕长肌起主要作用。这时如果对多个通道的肌电信号采取相同的权重去分析显然不太合适,因为这些肌电通道中有些起了关键作用,有些作用很小甚至不相关,这样一来对后续的分类识别造成了较大的干扰。

鉴于卷积神经网络和长短期记忆网络各自的特点和优势,本文提出 LSTM-CNN 神经网络结构的手势肌电识别模型。在卷积神经网络层前嵌入 LSTM 层,以手臂不同位置多路肌电信号构成的肌电图像、小波包特征图像为输入,形成手势动作肌电时空特征的有效组合,同时在 LSTM 层之后引入注意力机制赋值肌电通道的作用权重,最终成为时间特性和结构特性兼具的手势识别模型,识别准确率得以明显提升。

2 实验

2.1 NinaPro 肌电数据集 DB1

NinaPro 肌电数据集 DB1 是一个公开的肌电实验数据集的一个子集,采集的对象由 67 个身体健康的志愿者和 11 个手部残疾的志愿者组成。所有志愿者分成 3 组,其组成如下:

第 1 组:27 个身体健康的志愿者,其中 20 个男性,7 个女性;年龄范围:28±3 岁。

第 2 组:40 个身体健康的志愿者,其中 28 个男性,12 个女性;年龄范围:30±4 岁。

第 3 组:11 个手部残疾的志愿者,性别均为男性;年龄范围:42±12 岁。

实验过程要求完成 53 个不同的手部动作(标记号 0~52),其中 0 号动作是静息,每类动作被重复 10 次。53 个标签对应的手势类别分组如表 1 所示。

表 1 DB1 肌电数据集标签分组
Table 1 Label grouping of DB1 EMG data set

标签	手势动作分类
0	静息
1~12	手指的动作
13~29	手腕部分的动作
30~52	抓取动作

本文主要采用第 1 组志愿者采集的肌电数据,该数据采集设备是 13E200 MyoBock,采集频率是 100 Hz。每

个动作均采集手部的 10 路肌电信号,其中 8 路在环绕手的前臂靠近肘关节的位置处,剩余 2 路分别是指浅屈肌和指浅伸肌。

2.2 数据预处理

1) 手势肌电的特征选择

利用 13E200 MyoBock 自带 50 Hz 的工频滤波功能,特征提取前用一阶 1 Hz 的高通巴特沃斯滤波器^[12]对实验数据进行滤波。

本文选择的肌电特征分别为 10 路肌电信号构筑的肌电图像和肌电图像经小波包转换后生成的小波包(wavelet packet, WP)特征图像(WP 特征图像),其中肌电图像的获取如下:采集 10 路肌电通道信号,其中 8 路均匀环绕地分布在前臂靠近肘关节位置处。另外两路分别是指浅屈肌和指浅伸肌。接着用滑动窗的方式如图 1(a)所示,每隔 5 个数据点进行截取,截取的窗口长度是 15 个数据点,最终形成 10×15 的肌电图像。

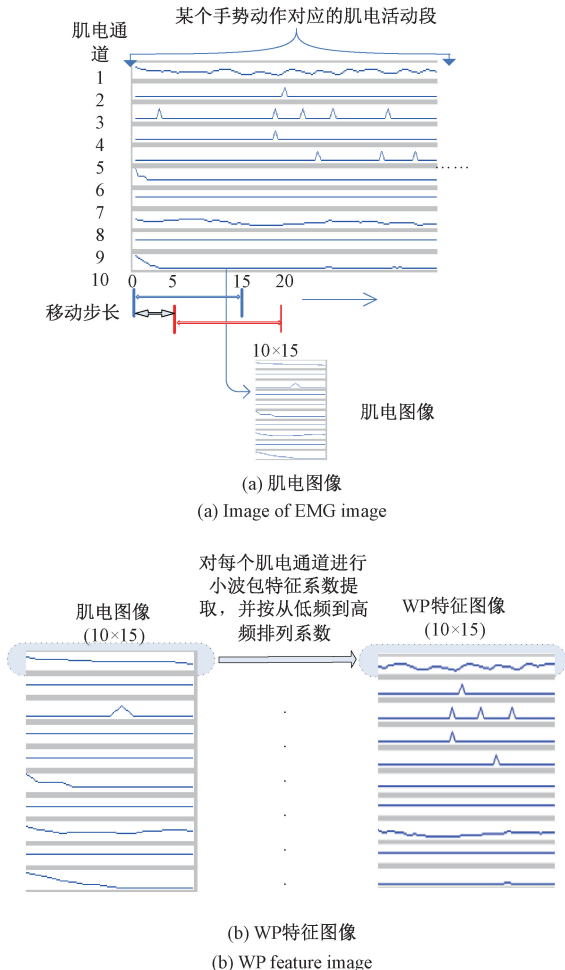


图1 肌电数据特征图像提取示意图

Fig. 1 Diagram of feature image extraction of EMG data

WP 特征图像的获取:首先是对每路肌电信号进行

小波包分解。小波分解只是对低频部分的信号进行分解,而对高频则不处理,这使得它的频率随着分解层数的增大而降级,即所获取的信号频率范围越来越窄,丢失了很多高频信息。小波包分解不仅对低频进行分解,也对高频部分进行分解。它能根据信号特性和分析要求自适应地选择相应频带与信号频谱相匹配,是一种比小波分解更为精细的分解方法。本文采用 db1 小波,进行 3 层小波包分解,如图 1(b)所示即对每个通道的肌电信号分别进行相应分解提取小波包系数,然后把把这些系数从低频到高频的顺序排列作为提取的特征,最终构成 10×15 的 WP 特征图像。

2) 标签的重定义

在 DB1 肌电数据集的采集过程,由于志愿者对动作开始和结束的反应时间会存在一定程度的差别,所以动作原始标签往往和实际采集数据的区域不是严格一一对应。制作该数据集的研究者根据广义似然比算法^[13]重新对数据集做了标签的重映射。本文所采用的标签是也是重映射后的标签。

3) 数据增强

在图像的深度学习模型中,常常采用数据增强来增加输入数据的数量,使得输入数据更加多样化,实验证明这样的预处理可以在一定程度上克服训练的过拟合问题^[14]。常用的图像数据增强的方法有对图像矩阵进行旋转,反射,平移,缩放,增加噪声等。而肌电数据相比图像数据在规模上小得多,所以增强的算法负担小。本文采用高斯白噪声,平移等方式对肌电数据进行数据增强。高斯白噪声模拟了环境的噪声影响,使肌电信号样本多样化,提升识别模型鲁棒性,添加高斯白噪声的过程如下:

假设信号为 x , x 幅值的均方值为 $\sigma(x)$ 。参数 SNR 为信噪比,它的值是从区间 (25, 45) 范围内随机生成,则高斯白噪声的方差如下:

$$np = \frac{\sigma(x)}{10^{SNR/20}} \quad (1)$$

即构筑一个均值为零方差为 np 的噪声,加入到原始的肌电信号中。而平移是指对信号提取活动段的过程中,段与段之间有部分重复,从而增加最终肌电图样本数。

3 深度学习神经网络模型

3.1 卷积神经网络模型

卷积神经网络是一种多层的监督学习神经网络,主要有局部感知,权值共享等特点,能有效减少神经网络结构中的参数,减少深层网络占用的内存量。卷积神经网络用梯度下降法最小化损失函数对网络中的权重参数逐

层反向调节,通过不断的迭代训练来提高最终的拟合精度。它包括:卷积层,池化层,全连接层。每一层有多个特征图,每个特征图通过一种卷积滤波器提取输入的一种特征,每个特征图有多个神经元。最后一层输出层是一个分类器,可以采用逻辑回归,如 Softmax 回归甚至是支持向量机对输入图像进行分类。

3.2 长短时记忆神经网络

LSTM 是一种特殊的 RNN^[15],它解决了传统 RNN 的长期依赖的问题,主要结构如图 2 所示。

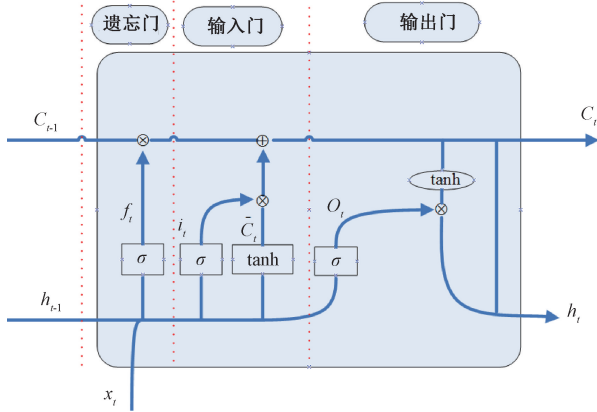


图2 LSTM神经网络拓扑

Fig. 2 Topology of the LSTM neural network

LSTM 主要由 3 部分组成,分别是遗忘门,输入门和输出门。其中遗忘门:

$$f_t = \sigma(\mathbf{W}_f \cdot [h_{t-1}, x_t] + b_f) \quad (2)$$

式中: σ 表示取 Sigmoid 函数值, $\mathbf{W}_f$ 表示权重矩阵, h_{t-1} 表示上一层 LSTM 神经网络的输出, x_t 表示输入, b_f 表示偏置量。 f_t 的元素取值范围是 0 到 1,表示遗忘的程度,0 表示全忘,1 表示全记住。

输入门:

$$i_t = \sigma(\mathbf{W}_i \cdot [h_{t-1}, x_t] + b_i) \quad (3)$$

$$C_t = \tanh(\mathbf{W}_c \cdot [h_{t-1}, x_t] + b_c) \quad (4)$$

其中, i_t 表示输入状态量, C_t 表示对 i_t 的筛选。

遗忘门和输入门决定了当前神经网络层的状态信息,即

$$C_t = f_t \times C_{t-1} + i_t \times C_t \quad (5)$$

输出门:

$$o_t = \sigma(\mathbf{W}_o \cdot [h_{t-1}, x_t] + b_o) \quad (6)$$

$$h_t = o_t \times \tanh(C_t) \quad (7)$$

输出门决定了哪些信息需要作为输出。

通过遗忘门可以减弱历史肌电信息中的一些噪声干扰,输入门可以选择历史信号中起重要作用的部分,最终在输出的结果中结合了当前信息与历史肌电信息中对后续肌电信号起辅助判别作用的部分信息。

3.3 注意力机制

注意力机制^[16]已经被广泛应用到基于深度学习的自然语言处理任务中,它可以帮助模型对输入的每个部分赋予不同的权重,抽取更加关键及重要的信息,同时不会对模型的计算和存储带来更大的开销。本文中,注意力机制的作用是间接地对通道关联性强的肌电信号赋予更高的权重。计算公式如下:

$$\alpha_{is} = \frac{\exp(\text{score}(h_t, h_s))}{\sum_{s=1}^s \exp(\text{score}(h_t, h_s))} \quad (8)$$

$$H_t = \sum_s \alpha_{is} h_s \quad (9)$$

$$\text{score}(h_t, h_s) = h_t^T \mathbf{W} h_s \quad (10)$$

其中, α_{is} 表示通道的权重参数向量; h_t 表示当前的输出状态,这里用各个肌电通道的均值去表示,均值越大说明该通道在该环境下越活跃; h_s 表示最上一层的输出,这里指的是 LSTM 的输出, $\mathbf{W}$ 是权重参数矩阵,由神经网络学习获得, H 表示经过注意力机制后的输出量。

3.4 基于 LSTM 与 CNN 的深度学习模型结构

从文献[17]对 Twitter 的情感分析研究得出,对于手势肌电这样的时序信号,由于 LSTM 能起到序列信息的筛选作用,LSTM 层在前、CNN 层在后的结合更有利于发挥两者的性能。若采用 CNN-LSTM 结构,CNN 处理后,会丢失初始输入信息中的序列信息,此时 LSTM 仅起到全连接层的作用,时序信息被屏蔽。

本文采用的深度学习模型结构如图 3 所示,主体基于 keras 框架,包含 LSTM 层和 CNN 层,CNN 部分主要借鉴文献[10]的设计,包含 4 个卷积隐层和一个输出层,第 1 层和第 2 层的卷积数量都是 32,卷积大小分别是 1×10 和 3×3 ,第 3 层卷积数量是 64,大小是 5×5 ,第 4 层卷积数量是 64,大小是 5×1 ,最后是输出层卷积数量是 53 即表示一共的手势种类。其中在进入 CNN 层之前进行注意力机制的处理,针对不同的手势动作间接地赋予不同通道肌电信号相应的权重。激活函数均采用 ReLu。神经网络层的权重参数矩阵采用 Xavier 初始化方法^[18],优化器采用随机梯度下降(stochastic gradient descent, SGD)。epoch 设置 30, batch size 设为 256。为了加快神经网络的训练速度同时防止后期因学习率过大而错过最优解,设置初始学习率为 0.01,当 epoch 循环次数超过 20 次后,每经过一个 epoch 减少 1/2。

其中肌电的输入数据的大小是 10×15 ,其中 15 表示 15 个时间序列点,即活动窗的宽度,对应的时间是 150 ms,10 表示 10 个肌电采集通道。活动窗的移动步长是 5 个序列点。接着是根据该原始输入数据提取 2 层的肌电特征图像,其中第一是肌电图像,第二是肌电的 WP 特征图像。

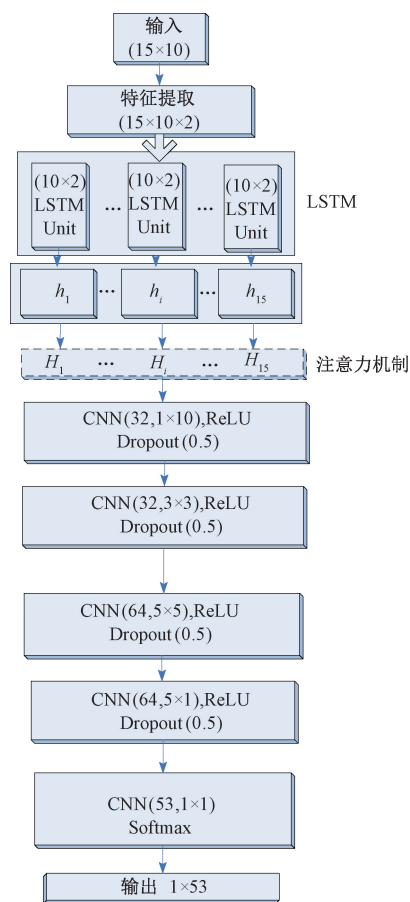


图3 深度学习模型结构

Fig. 3 Deep learning model structure

输入 LSTM 层的时间步长是 15, 输入的单元是 10×2 的向量。LSTM 层输出是对应整个时间步长的全部序列。注意力机制层的输入就是 LSTM 的输出, 此层的主要目的是根据能肌电通道的均值和 LSTM 的输出去获取相应的神经网络权重参数, 最终获取含有通道权重信息的输出。

3.5 实验方法与步骤

为了对比本文算法的效果, 在相同的深度学习神经

网络 LSTM-CNN 模型结构下, 分别以肌电图像, 肌电图像和小波包特征图像的组合输入到模型中进行对比。

识别结果中除了单帧肌电数据的识别准确率, 还提出了多数投票识别策略, 它的决策结果是建立在单帧肌电数据识别结果的基础上。即判断某个完整手势活动数据的动作类别, 先把该段肌电数据按移动窗口的方式截取并用深度学习模型识别出每帧肌电数据的手势类别, 最终把这些识别比例最多的手势当做该部分肌电数据的识别结果。为了能和已有的研究形成对比, 本文和文献[10]一样把每种手势重复动作中的第 2, 5, 7 次的肌电数据当做测试集, 其余的 7 次当做训练集。同时为了能在有限的肌电数据中找到合适的模型参数, 本文采用交叉验证的方式在训练集中进行训练优化, 即把剩余每种手势重复动作的第 1, 3, 4, 6, 8, 9, 10 次的肌电数据, 依次用其中的某次肌电数据当做验证数据, 其余当做训练数据。

针对某个手势动作其中一次的肌电数据, 首先按照图 1 所示获得该肌电数据的所有特征图像, 接着用深度学习识别模型识别出对应的手势类别, 并与该次真实的手势动作比较从而判断出是否正确。该手势动作 7 次的正确率就是该手势动作对应的多数投票识别准确率。相比直接的单帧肌电特征图像数据的准确率, 活动段的投票准确率更具有实际的应用价值。

4 结果与分析

4.1 基于肌电图像的识别结果

图 4 是肌电图像的神经网络识别模型的训练集 loss 和测试集 loss, 可以看出在 epoch 达到 15 之后 loss 均趋于稳定。训练集的 loss 最终在 1 以下, 而测试集的 loss 在 1.5 左右, 这说明该模型存在一定程度的过拟合现象。且测试集的 loss 并不十分稳定, 在 1.5 左右存在一定程度的抖动。训练集的识别准确率是 0.841 2, 测试集的识别准确率是 0.731 2, 测试集的投票准确率是 0.942 1, 如表 2 中所示。

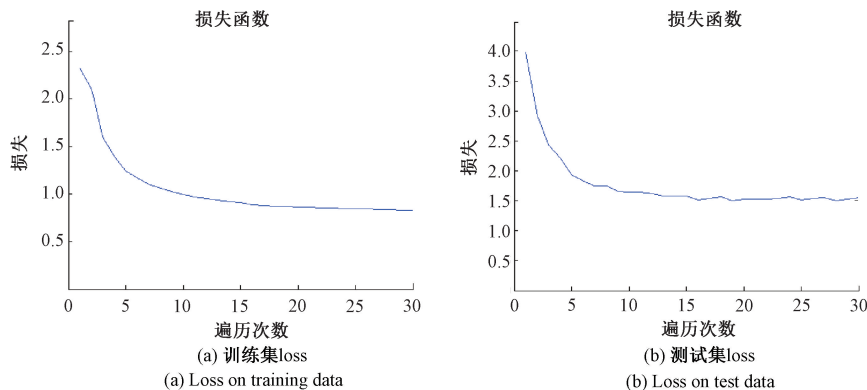


图4 基于肌电图像 loss 结果

Fig. 4 Loss results based on EMG image

图5是单帧肌电数据与多数投票表决预测的混淆矩阵,由图5(a)、(b)相结合可以看出大部分的数据点在从左上至右下这条对角线附近,手势动作的误分类主要发生在相邻标签手势的动作的情况下,而大部分相邻标签

对应的手势动作过程比较相近或类似。比如标签11和12都是只有大拇指在动,只是它的动作方向不同,这也基本符合实验的预期。

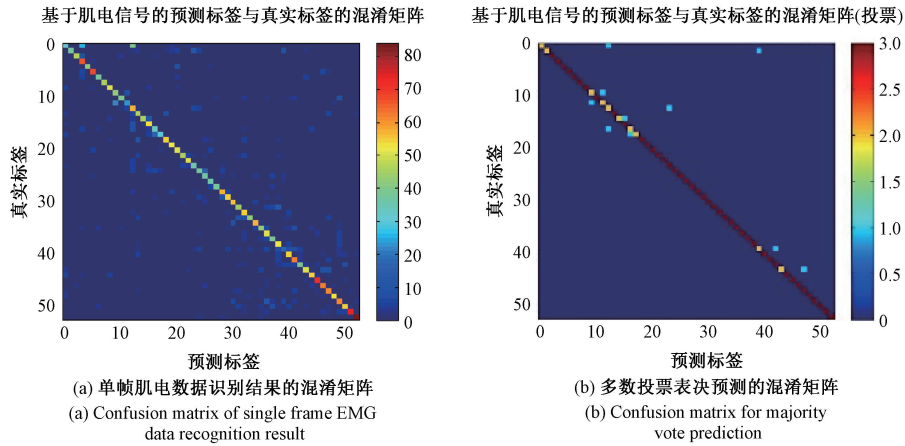


图5 单帧肌电数据与多数投票表决预测的混淆矩阵

Fig. 5 Confusion matrix of single frame EMG data and majority vote prediction

图6是单帧肌电数据在一个完整手势动作内中各个时间点所对应的识别准确率结果,可知在整个手势动作中,起始和结束部分的单帧肌电数据的识别准确率相对于中间部分的信号较低。这也反映到实际中就是手势动作中间段的肌电信号往往能量较大且明显,所以可以比较明显的区分不同活动段,使最终识别准确率比较高。

4.2 基于由肌电图像和 WP 特征图像的实验结果

本文结合肌电图像和 WP 特征图像,并以此作为输入数据输入到深度学习识别模型 LSTM-CNN 中,结果如图7所示,loss 值分别在训练集和测试集中的30次迭代分布,可以看到相比肌电图像的识别结果有较为明显的提升,两种不同输入特征手势识别方法的准确率如表2所示。

肌电图像在一个完整手势动作期间不同时刻下的准确率

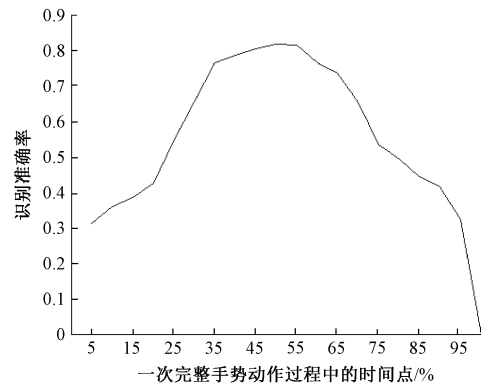


图6 一个完整手势动作内各个时间段所对应的肌电图像的识别准确率

Fig. 6 The electromyogram corresponding to each time period in a complete gesture action

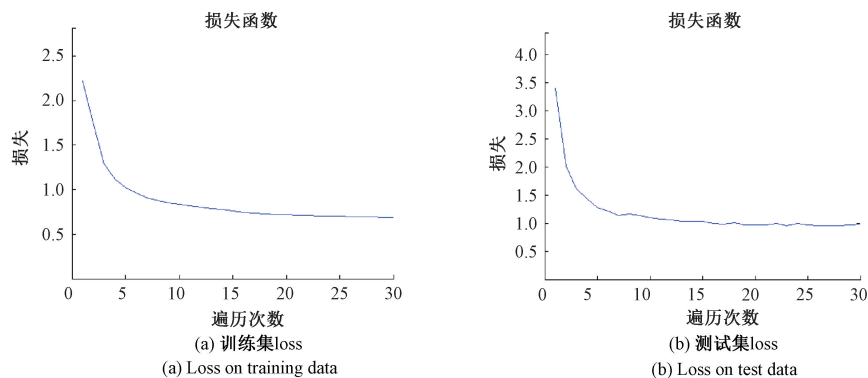


图7 基于肌电图像+WP 特征图像组合的 loss 图

Fig. 7 Loss map based on combination of EMG image and WP feature image

表2 手势动作分别识别准确率的比较
Table 2 Comparison of the recognition accuracy of gesture actions

特征/识别模型	训练集 准确率	测试集 准确率	投票 准确率
(单一原始肌电特征, CNN) ^[10]	0.812 1	0.720 6	0.930 6
单一原始肌电特征 LSTM-CNN	0.841 2	0.731 2	0.942 1
本方法	0.851 2	0.751 2	0.957 6

增加的 WP 特征图像与肌电图像组合构成了两层特征信息,每层都是对同一个对象的不同方面的描述。层的细节信息越丰富且层之间的信息冗余越少最终所形成的组合特征就越能定位所描述对象的分类区别。由表2可以看出增加了小波包特征图像之后,识别模型的识别准确率有一定程度的提升。

在同样的输入特征中,采用 LSTM-CNN 模型的识别准确率高于 CNN 模型, LSTM 层的作用不仅提取了单一

肌电图的时序特征并蕴含了此之前的系列肌电数据之间的历史信息,即进入 CNN 模型的输入信息中多了肌电数据之间的相互关系特征,最终使得识别准确率有所提高,另一方面加入了注意力机制后,神经网络模型能够更加灵活的分辨出不同的手势动作对应的各部分肌肉的活跃程度以及相互之间的关联度是不同的。

图8是基于肌电图像和 WP 特征图像组合,单帧肌电输入数据下各个手臂动作的识别准确率。每个动作组里均有识别率比较高和比较低的手势动作,基于肌电图像和基于肌电图像+WP 特征图像的组合特征对各类型的手势识别率变化趋势大致相近,在以测试集识别准确率作为主要评价指标的条件下,肌电图像+WP 特征图像组合的总体识别率比文献[10]提高了4.25%。最低识别率动作都是第39动作类,即单手握网球的动作,该手势动作的变化幅度较小,造成区分度相对较小。另外由结果可知组合特征对动作类49和50的识别结果均有较大的提升。

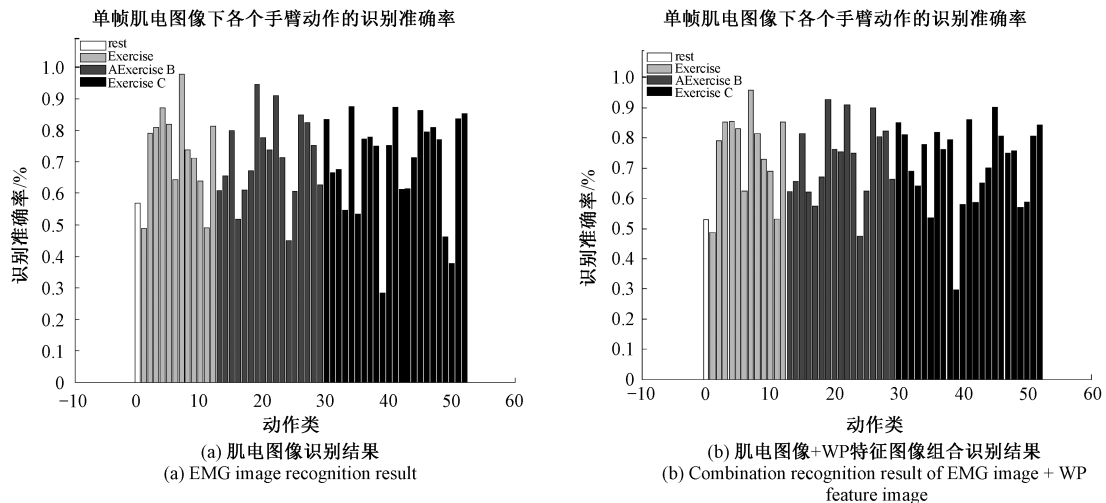


图8 基于分别由肌电图像和肌电图像+WP 特征图像组合的识别结果

Fig. 8 Based on the recognition results of the combination of EMG image and EMG image + WP feature image

5 实际应用效果和分析

MYO 腕带是加拿大创业公司 Thalmic Labs 推出的创新性臂环,它可以探测用户的肌肉产生的电活动,通过低功耗的蓝牙其他电子产品进行无线连接。MYO 8 路肌电通道的采集能力以及便携性使得被很多研究肌电的实验室所采纳使用,另一方面虽然它在开放肌电数据的同时也内置了5种固定手势的识别用于一般用户的使用,由于实际使用中的识别准确率并不是很高,用户体验感也不理想,最终导致 MYO 缺乏广泛的受众而没能普及。

针对本文所提出的方法,用 MYO 采集3种手势,分

别是静息状态,手爪的神拳和握拳。用 MYO 分别采集3位志愿者的右臂肌电数据,志愿者年龄分别是24,24,30,性别均为男性,其余操作基本和公开数据集 DB1 相似。最终用本文的模型所达到的测试集平均识别率是0.91,普通的 CNN 识别模型的测试集平均识别率是0.88。可以看出本文提出的模型在识别手势方面有较为明显的提高。

6 结 论

利用长短时记忆网络善于提取信号时序特征的优点,结合卷积神经网络在图像数据特征上的泛化能力,提

出基于 LSTM-CNN 的串行神经网络结构并在两者之间加入注意力机制,作为手势肌电的识别模型。在模型的输入数据方面,提出手势肌电图像+WP 特征图像作为网络输入的 2 层输入特征信息。多种特征信息输入与串行 LSTM-CNN 网络,构成了能协同肌电信号的结构和时序特征,而注意力机制的作用是间接地强化了在手势动作中有重要作用的肌电通道信号,最终使得该肌电识别模型具有良好识别能力。实验结果表明本文所提神经网络手势肌电识别模型有利于减少过拟合的现象,并且对肌电图像的识别准确率比普通的卷积神经网络模型^[10]提高了 4.25% 的识别准确率。同时也对用 MYO 采集的数据进行了验证,结果表明本文提出的方案能有效的提高模型的手势识别率。

本文在肌电手势识别模型的输入数据组成和网络结构方面做了优化和改进,取得了有益的结果。更深入的研究还可以探讨窗口长度与识别结果的关系,根据识别的手势动作寻找最佳的窗口长度,另一方面还可以研究如何减少输入数据中的信息冗余等方面。

参考文献

- [1] YOU B, WANG H L, HUANG L. The system of s-EMG recognition for prosthetic hand control [C]. Strategic Technology (IF-OST), Ulsan: IEEE, 2010, DOI: 10.1109/IFOST.2010.5667998.
- [2] ZHANG J, LING C, LI S. Human movements classification using multi-channel surface EMG signals and deep learning technique [C]. 2019 International Conference on Cyberworlds (CW). 2019, DOI: 10.1109/CW.2019.00051.
- [3] CHRISTIAN P, STEFAN S, LEONHARD D. Results of an internet survey of myoelectric prosthetic hand users [J]. Prosthetics and Orthotics International, 2007, 31 (4): 362-370.
- [4] 陈志强, 陈旭东, José, 等. 深度学习在设备故障预测与健康管理中的应用 [J]. 仪器仪表学报, 2019, 40 (9): 209-229.
CHEN ZH Q, CHEN XU D, JOSÉ, et al. Application of deep learning in equipment prognostics and health management [J]. Chinese Journal of Scientific Instrument, 2019, 40 (9): 209-229.
- [5] ULYSSE C, CHEIKH L, ALEXANDRE C, et al. Transfer learning for SEMG hand gestures recognition using convolutional neural networks [C]. 2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC), 2017, 1663-1668.
- [6] JAFARZADEH M, HUSSEY D C, TADESSE Y. Deep learning approach to control of prosthetic hands with electromyography signals [C]. 2019 IEEE International Symposium on Measurement and Control in Robotics (ISMCR), 2019, DOI: 10.1109/ISMCR47492.2019.8955725.
- [7] CHUNG E A, BENALCÁZAR M E. Real-Time hand gesture recognition model using deep learning techniques and EMG signals [C]. 2019 27th European Signal Processing Conference (EUSIPCO), IEEE, 2019: 1-5.
- [8] ATZORI M, COGNOLATO M, MÜLLER H. Deep learning with convolutional neural networks applied to electromyography data: A resource for the classification of movements for prosthetic hands [J]. Frontiers in Neurobotics, 2016, 10: 9.
- [9] ATZORI M, GIJSBERTS A, KUZBORSKI I, et al. Characterization of a benchmark database for myoelectric movement classification [J]. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2015, 23 (1): 73-83.
- [10] TSINGANOS P, CORNELIS B, CORNELIS J, et al. Deep learning in EMG-based gesture recognition [C]. PhyCS, 2018: 107-114.
- [11] VERMA H, KUMAR S. An accurate missing data prediction method using LSTM based deep learning for health care [C]. Proceedings of the 20th International Conference on Distributed Computing and Networking, 2019: 371-376.
- [12] GENG W, DU Y, JIN W, et al. Gesture recognition by instantaneous surface EMG images [J]. Scientific Reports, 2016, 6 (1): 1-8.
- [13] KUZBORSKI I, GIJSBERTS A, CAPUTO B. On the challenge of classifying 52 hand movements from surface electromyography [C]. 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, IEEE, 2012: 4931-4937.
- [14] ZHANG Y D, DONG Z, CHEN X, et al. Image based fruit category classification by 13-layer deep convolutional neural network and data augmentation [J]. Multimedia Tools and Applications, 2019, 78 (3): 3613-3632.
- [15] SEPP HOCHREITER, JÜRGEN S. Long short-term memory [J]. Neural Computation, 1997, 9 (8): 1735-1780.
- [16] LUONG M T, PHAM H, MANNING C D. Effective approaches to attention-based neural machine translation [J].

- Computer Science, 2015, DOI:10.18653/v1/D15-1166.
- [17] SOSA P M. Twitter sentiment analysis using combined LSTM-CNN models[J]. Eprint Arxiv, 2017: 1-9.
- [18] GLOROT X, BENGIO Y. Understanding the difficulty of training deep feedforward neural networks [C]. Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics, JMLR Workshop and Conference Proceedings, 2010: 249-256.

作者简介



陈思佳, 2013 年于杭州电子科技大学获得学士学位, 2018 年于杭州电子科技大学获得硕士学位, 现为杭州电子科技大学在读博士生, 主要研究方向为生物电信息的处理与识别, 机器人控制及其应用。

E-mail: 710442015@qq.com

Chen Sijia received his B. Sc. degree and M. Sc. degree both from Hangzhou Dianzi University in 2013 and 2018, respectively. He is currently a Ph. D. candidate at Hangzhou Dianzi

University. His main research interests include bioelectric information processing and recognition, robot control and its application.



罗志增 (通信作者), 1985 年于电子科技大学获得学士学位, 1989 年于杭州电子科技大学获得硕士学位, 1998 年于浙江大学获得博士学位, 现为杭州电子科技大学博士生导师。主要研究方向: 信息采集与多信息融合、生物医学信号处理、智能机器人等。

E-mail: luo@hdu.edu.cn

Luo Zhizeng (Corresponding author) received his B. Sc. degree from University of Electronic Science and Technology of China in 1985, received his M. Sc. degree from Hangzhou Dianzi University in 1989, and received his Ph. D. degree from Zhejiang University in 1998. He is currently a Ph. D. advisor at Hangzhou Dianzi University. His main research interests include information collection and multi-information fusion, biomedical signal processing, intelligent robots, etc.