

DOI: 10.13382/j.jemi.B2306373

注意力机制 CNN 结合肌电特征矩阵的手势识别研究*

赵世昊^{1,2} 周建华^{1,2} 伏云发^{1,2}

(1. 昆明理工大学信息工程与自动化学院 昆明 650500; 2. 昆明理工大学脑认知与脑机智能融合创新团队 昆明 650500)

摘要:当前基于卷积神经网络(CNN)的手势识别研究集中于增加网络深度,较少关注改善样本数据分布带来的性能提升。针对此类问题,提出一种量化表面肌电信号(sEMG)特征相关性的肌电特征矩阵(EFM)样本输入有效通道注意力(ECA)机制 CNN,用于识别 NinaproDB1 中 52 类手势。首先使用时间窗截取低通滤波后的 sEMG,计算多种信号时域特征;然后利用笛卡尔积组合并相乘不同特征,对特征相乘值进行归一化后得到 EFM。同时,引入 ECA 机制使网络关注重要的深层特征,从而提升手势分类效果。分别输入 sEMG、肌电时域特征和 EFM 到注意力机制 CNN 进行手势识别,EFM 识别准确率最高,达到了 86.39%,高于近年来手势识别研究方法精度。验证了提出方法的有效性,为多类别手势准确分类提供可行新方案。

关键词: 手势识别;肌电特征矩阵;有效通道注意力;卷积神经网络

中图分类号: TP391.4 **文献标识码:** A **国家标准学科分类代码:** 310.61

Investigation of gesture recognition using attention mechanism CNN combined electromyography feature matrix

Zhao Shihao^{1,2} Zhou Jianhua^{1,2} Fu Yunfa^{1,2}

(1. Faculty of Information Engineering and Automation, Kunming University of Science and Technology, Kunming 650500, China; 2. Brain Cognition and Brain-Computer Intelligence Integration Group, Kunming University of Science and Technology, Kunming 650500, China)

Abstract: Current research on gesture recognition based on convolutional neural network (CNN) focuses on increasing the depth of the network, and pays less attention to improve the distribution of sample data which can brought the performance improvement. Aimed at these problems, a kind of electromyography feature matrix (EFM) sample that quantifies the correlation of surface electromyography (sEMG) features is fed into the efficient channel attention (ECA) mechanism CNN, which is used to identify the 52 types gesture in NinaproDB1. Firstly, the time window is used to truncate the low-pass filtered sEMG and calculate various signal time domain features. Then, the cartesian product is used to combine and multiply different features. The EFM is obtained after normalizing the feature multiplication values. At the same time, ECA mechanism is introduced to make the network focus on the important deep features, thereby improving the effect of gesture classification. sEMG, EMG time-domain features and EFM are fed to the attention mechanism CNN respectively for gesture recognition. The recognition accuracy of EFM is the highest and reached 86.39%, which is higher than the accuracy of gesture recognition research methods in recent years. The effectiveness of the proposed method is verified, and a feasible new scheme for accurate multi-category gesture classification is provided.

Keywords: gesture recognition; electromyography feature matrix (EFM); efficient channel attention (ECA); convolutional neural network (CNN)

0 引言

手势识别是人机交互^[1]的重要载体,本质上属于模式识别任务。相比于计算机视觉^[2]与惯性信号^[3]的手势识别效果,表面肌电信号^[4](surface electromyography, sEMG)手势识别的鲁棒性更高,适用于康复医疗^[5]、假肢外骨骼^[6]等领域。sEMG 由运动单元动作电位沿肌肉纤维经脂肪和溶液传导,在皮肤表面叠加后通过肌电极采集,具有差异性、非平稳性和超前性。sEMG 手势识别的关键在于信号预处理与分类模型,二者分别用于去除 sEMG 噪声干扰和预测样本类别。其中,手势分类模型分为机器学习和深度学习两大类。机器学习^[7]需要人工制作特征集,利用支持向量机、线性判别分析和随机森林等模型分类手势,但分类效果依赖特征质量。相比于机器学习,深度学习^[8]以多层表征的形式学习数据特征,由样本数据端到端地预测手势类别,是手势分类模型的研究热点。

目前,卷积神经网络(convolutional neural network, CNN)作为深度学习分类模型被广泛应用于手势识别, Ketykó 等^[9]提出的自适应框架,识别 12 种手势的准确率为 84.7%;Ding 等^[10]用一维卷积预处理 sEMG 再进行分类,对 18 种手势的平均识别率为 78.86%;Atzori 等^[11]提出基准 sEMG 数据集 NinaproDB1^[12],首次使用 CNN 进行 sEMG 手势识别;Geng 等^[13]提出基于瞬时肌电图的 CNN,用于高密度 sEMG 手势识别效果较好。Cheng 等^[14]提取多种 sEMG 特征,生成多通道特征图输入 CNN,DB1 中 52 类手势识别平均准确率达 82.54%。然而上述研究工作主要集中于网络深度,忽视了样本数据分布改善带来的分类性能提升。同时,网络深度增加导致学习的特征信息冗杂,造成手势识别精度降低。

为优化数据间的关联性,提升不同手势样本的区别性。同时增强网络的特征学习能力,提供多类别手势准确识别的可行方案。提出一种肌电特征矩阵(electromyography feature matrix, EFM)输入有效通道注意力^[15](efficient channel attention, ECA)机制 CNN 的手势识别方法,用于 NiaproDB1 中 52 类手势识别平均准确率达 86.39%,证明本文方法有效性。

1 EFM 样本

基于 sEMG 手势识别流程如图 1 所示。首先截取滤波后的 sEMG,生成训练集样本输入网络,迭代训练出较高分类准确率的模型。同时,生成测试集样本检测识别新样本能力最优网络模型。

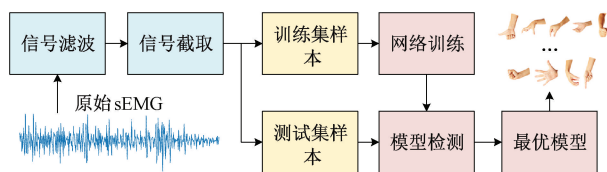


图 1 基于 sEMG 手势识别流程

Fig. 1 Gesture recognition process based on sEMG

1.1 sEMG 预处理

sEMG 是一种非平稳的生理电信号,幅值主要在 0.01 ~ 10 mV 之间,主要能量在 0 ~ 500 Hz^[1]。NinaproDB1 中 sEMG 已经去除工频干扰和基线偏移,本文只采用截止频率为 1 Hz 的 1 阶巴特沃斯滤波器对信号低通滤波^[11],最大限度地保留信号原有的特性。同时,为满足人机交互的实时性,sEMG 设备控制周期要小于 300 ms^[16]。用时长为 T 、移动时长为 S 的毫秒时间窗截取 sEMG,计算公式如下:

$$L = T \cdot F / 1\,000 \quad (1)$$

$$L_d = S \cdot F / 1\,000 \quad (2)$$

式中: L 为时间窗帧数, L_d 为滑动步长帧数, F 为信号采样率,DB1 中等于 100 Hz。得到数据样本 $x \in R^{C \times L}$ 。其中, C 表示信号采集通道数量,DB1 中等于 10。sEMG 滑动截取如图 2 所示。

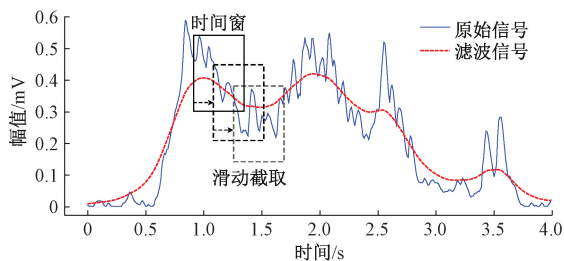


图 2 单通道滤波 sEMG 滑动截取

Fig. 2 Sliding truncation of single channel filtering sEMG

1.2 EFM 生成原理

近年来基于 CNN 的 sEMG 手势识别研究中,通常以简单预处理的 sEMG 作为输入样本,忽略了改善输入样本数据分布对分类效果的提升作用。Cheng 等^[14]虽然提取 sEMG 时域特征构造多通道特征图像,将 sEMG 数据分布映射到特征空间。但特征分布于不同通道弱化了数据之间的关联性,削弱了 CNN 对不同区域数据局部连接分析的基础。因此,本文提取 sEMG 多种时域特征进行拼接,利用笛卡尔积得到不同特征之间的组合。将特征相乘后量化 sEMG 不同特征的相关性,经 Sigmoid 函数归一化为 EFM。EFM 生成过程如图 3 所示。

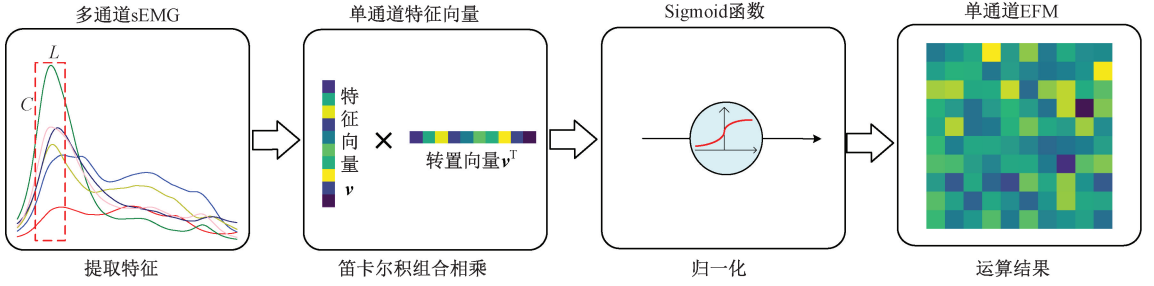


图3 EFM 生成过程

Fig. 3 EFM generating process

本文根据文献[17]中的手势分类效果,使用 L 帧时间窗截取 sEMG,计算 10 种时域特征。时域特征 $v_1 \sim v_{10}$ 定义及计算如下:

1) v_1 为均方根 (root mean square, RMS): 代表恒力与非疲劳收缩指标,计算公式如下:

$$v_1 = \sqrt{1/L \sum_{i=1}^L x_i^2} \quad (3)$$

2) v_2 为平均绝对值 (mean absolute value, MAV): 代表肌肉收缩水平指标,计算公式如下:

$$v_2 = 1/L \sum_{i=1}^L |x_i| \quad (4)$$

3) v_3 为波长 (wave length, WL): 代表 sEMG 波形 L 帧累加长度,计算公式如下:

$$v_3 = \sum_{i=1}^{L-1} |x_{i+1} - x_i| \quad (5)$$

4) v_4 为简单平方积分 (simple square integral, SSI): 代表信号能量,计算公式如下:

$$v_4 = \sum_{i=1}^L x_i^2 \quad (6)$$

5) v_5 为 sEMG 第三时间矩 (third time moment, TM3),计算公式如下:

$$v_5 = 1/L \sum_{i=1}^L x_i^3 \quad (7)$$

6) v_6 为 sEMG 第四时间矩 (fourth time moment, TM4),计算公式如下:

$$v_6 = 1/L \sum_{i=1}^L x_i^4 \quad (8)$$

7) v_7 为 sEMG 第五时间矩 (fifth time moment, TM5),计算公式如下:

$$v_7 = 1/L \sum_{i=1}^L x_i^5 \quad (9)$$

8) v_8 为积分肌电值 (integral electromyography, iEMG): 代表 sEMG 信号功率,计算公式如下:

$$v_8 = \sum_{i=1}^L |x_i| \quad (10)$$

9) v_9 为 MAV 改进 1 型 (MAV1): 提升 MAV 特征鲁

棒性,计算公式如下:

$$v_9 = 1/L \sum_{i=1}^L n_i |x_i| \quad (11)$$

$$n_i = \begin{cases} 1, & 0.25L \leq i \leq 0.75L \\ 0.5, & \text{其他} \end{cases}$$

10) v_{10} 为 MAV 改进 2 型 (MAV2): 提升 MAV 特征平滑性,计算公式如下:

$$v_{10} = 1/L \sum_{i=1}^L n_i |x_i| \quad (12)$$

$$n_i = \begin{cases} 1, & 0.25L \leq i \leq 0.75L \\ 4i/L, & i \leq 0.25L \\ 4(i-L)/L, & \text{其他} \end{cases}$$

式(3)~(12)中, x_i 为第 i 帧 sEMG 数据。拼接特征值成特征向量 $\mathbf{v}$, 转置后如式(13)所示:

$$\mathbf{v}^T = [v_1, v_2, v_3, v_4, v_5, v_6, v_7, v_8, v_9, v_{10}] \quad (13)$$

根据笛卡尔积的数学定义组合特征向量元素得到特征集合。为量化不同特征之间的相关性,将组合元素相乘计算具体值,得到 sEMG 特征相关性矩阵,过程如式(14)所示:

$$\mathbf{v} \times \mathbf{v}^T = \begin{pmatrix} v_1 v_1 & v_1 v_2 & \cdots & v_1 v_{10} \\ v_2 v_1 & v_2 v_2 & \cdots & v_2 v_{10} \\ \vdots & \vdots & \ddots & \vdots \\ v_{10} v_1 & v_{10} v_2 & \cdots & v_{10} v_{10} \end{pmatrix} \quad (14)$$

特征之间存在差异性,导致特征相关性矩阵中不同元素值相差较大。为合理化数据分布,需要归一化特征相关性矩阵得到 EFM。计算过程如下:

$$\sigma(x) = 1/(1 + e^{-x}) \quad (15)$$

$$M_F = \sigma(\alpha \cdot (\mathbf{v} \times \mathbf{v}^T)^\beta) \quad (16)$$

式(15)表示 Sigmoid 函数;式(16)表示数据归一化运算,利用 Sigmoid 函数性质将特征矩阵值缩放至 (0, 1) 区间内。为避免数据落在 Sigmoid 函数饱和区失去差异性,设置归一化参数 α, β 均为 0.5,得到 EFM 样本 M_F 。

2 注意力机制 CNN 模型

2.1 ECA 机制

得益于局部连接和权值共享功能, CNN 以较少的权重参数获得良好的大数据分析能力。提升模型分类性能需要加深网络。但是网络加深时, 提取的特征相应增多且信息复杂度上升, 导致网络的学习能力弱化, 分类准确率下降。为了充分学习数据关键信息, 引入 ECA 机制赋予各通道特征权重, 使网络关注重要的特征, 原理如图 4 所示。

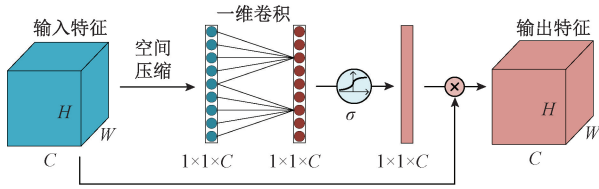


图 4 ECA 机制

Fig. 4 ECA mechanism

首先压缩输入特征空间维度 H 、 W 上的数据, 进行全局平均池化, 得到 $1 \times 1 \times C$ 初始通道权值。单通道特征空间压缩计算如下:

$$S_{1c} = (1/(H \times W)) \sum_{j=1}^H \sum_{i=1}^W X_{ij} \quad (17)$$

$$S_1 = [S_{11}, S_{12}, \dots, S_{1C}]$$

式中: S_{1c} 表示第 C 个通道的初始权值; X 为单通道输入; H 和 W 为空间维度长和宽; S_1 为拼接 C 通道初始权值得到的初始权重。

然后, 卷积计算部分通道权重, 融合为表征全部通道关联信息的输出特征, 经过 Sigmoid 函数映射为注意力权重 S_2 。映射计算如下:

$$S_2 = \sigma(f(S_1)) \quad (18)$$

式中: f 为一维卷积运算, σ 为 Sigmoid 运算。一维卷积运算时, 按照特征的通道数自适应改变卷积核大小的计算如下:

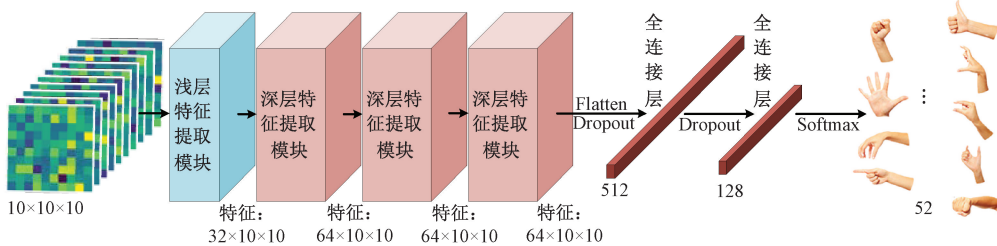
$$k = \lceil (\log_2^C + \lambda) / \gamma \rceil \quad (19)$$

式中: λ 、 γ 分别设置为 1 和 2, C 为输入特征通道数, k 为自适应一维卷积核大小, 卷积步长设置为 1。ECA 模块添加在特征输出层后, 用于筛选重要的特征。

2.2 注意力机制 CNN

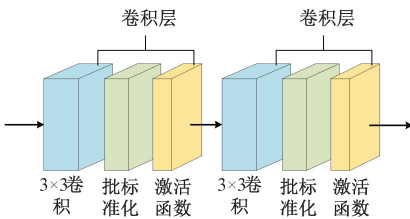
本文参考 VGGNet^[18] 结构, 堆叠特征提取模块设计分类模型。为避免池化操作信息丢失, 使用 Pointwise 卷积原理中 1×1 卷积层代替池化层, 构成局部连接层。采用重叠 3×3 卷积层代替 5×5 卷积层组成浅层特征提取模块, 减少网络权重参数的同时, 增大浅层特征感受野。在浅层学习近似全局数据分布信息, 有助于提高网络深层特征的信息综合度。手势识别模型框架如图 5 所示。

ECA 机制分别与卷积层、局部连接层结合, 构成模型深层特征提取模块。卷积层和局部连接层包含以下运算过程:



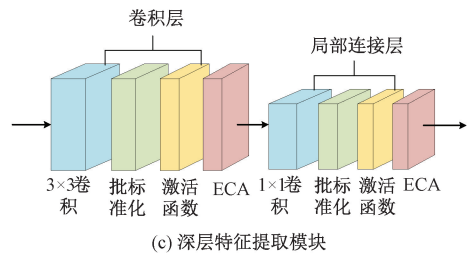
(a) 注意力机制 CNN 模型架构

(a) Architecture of attention mechanism CNN model



(b) 浅层特征提取模块

(b) Shallow feature extraction module



(c) 深层特征提取模块

(c) Deep feature extraction module

图 5 手势识别模型框架

Fig. 5 Model framework of gesture recognition

1) 卷积运算

浅层特征输入深层特征提取模块后,由卷积核提取深层特征,计算过程如下:

$$Y_{i,j} = \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} K_{m,n} \cdot X_{i+m-p,j+n-p} \quad (20)$$

式中: X 表示输入特征, i,j 分别为长和宽维度; K 表示卷积核,尺寸 $M \times N$; Y 表示输出特征,与输入特征 X 保持维度相同。

2) 批标准化

批标准化^[19] (batch normalization) 规范输入数据为标准正态分布,避免网络前向传播内部协变量漂移和梯度消失,加快模型收敛。批输入 $x = \{x_1, x_2, \dots, x_n\}$ 时,批标准化计算过程如下:

(1) 计算批输入的平均值,如式(21)所示:

$$\mu_B = \frac{1}{n} \sum_{i=1}^n x_i \quad (21)$$

(2) 计算批输入的方差,如式(22)所示:

$$\sigma_B^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \mu_B)^2 \quad (22)$$

(3) 标准化批输入,添加极小值 ε , 防止出现零除数,如式(23)所示:

$$\hat{x}_i = \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \varepsilon}} \quad (23)$$

(4) 添加标准化尺度变换因子 γ 、偏移因子 β , 避免数据为标准正态分布削弱网络学习能力,合理化数据分布,如式(24)所示:

$$BN_{\gamma,\beta}(x_i) \equiv \gamma \hat{x}_i + \beta \quad (24)$$

3) 激活函数

本文的网络激活函数选用 ReLU, 当输入大于 0 时输出不变,小于 0 时输出为 0。添加 ReLU 激活函数增强网络非线性并稀疏网络,降低过拟合风险。

提出的注意力机制 CNN 模型中,深层特征提取模块的输出平铺展开 (Flatten) 后输入全连接层。由于全连接层权重参数过多,模型容易复杂化,在全连接层中引入概率为 0.5 的随机丢弃 (Dropout)。使用 Softmax 输出层预测样本属于各种手势类别的概率。

3 实验结果与分析

3.1 实验参数设置

实验选用的 NinaproDB1 数据集包含 27 个被试 (年龄:25~31 岁,男性 20 人,女性 7 人) 所做 52 类手势,每类手势动作重复 10 次,每次持续 3 s 左右。手势动作的 sEMG 数据由 8 个等距围绕肘关节下方和 2 个位于指伸肌和指浅屈肌主活动点的 MyoBock13E200 电极采集,采样率 100 Hz。手势动作如图 6 所示。



图 6 52 类手势动作

Fig. 6 Movements of 52 types gesture

每个被试的训练集与测试集样本按照手势动作重复划分,第 2、5、7 次重复为测试集,其余为训练集。实验以 N 个被试的测试集手势识别平均准确率作为评估,计算方法如式(25)所示:

$$\text{手势识别平均准确率} = \frac{1}{N} \sum_N \frac{\text{单被试正确分类样本数}}{\text{单被试分类样本总数}} \quad (25)$$

实验环境基于 PyTorch 搭建,使用 RTX3090 型号的 GPU 加速网络训练,训练轮次设置为 30,输入样本的批数据量为 1 024。使用交叉熵函数计算网络预测结果与样本真实标签之间的误差损失,计算公式如下:

$$\text{loss} = - \sum_{i=1}^n y_i \log(\hat{y}_i) \quad (26)$$

式中: y_i 为样本属于第 i 类标签手势的真实值, $\hat{y}_i$ 为样本属于第 i 类标签手势的预测值, n 表示类别数量,loss 为样本预测值与真实值误差损失。标签独热编码,方便交叉熵计算。由于每个输入样本对应一种手势标签,手势识别可以视作一种单标签多分类问题。利用反向传播对 loss 链式求导更新网络权重,采用 AdamW 优化 loss,权重衰减为 0.005,正则化权重更新防止过拟合。初始的学

习率为 0.003,训练中用余弦退火策略调整。网络训练超参数如表 1 所示。

表 1 训练超参数
Table 1 Training hyper-parameter

超参数	数值
训练轮次	30
批数据量	1 024
优化器	AdamW
初始学习率	0.003
权重衰减	0.005

3.2 参数实验

为评估不同时间窗手势识别,滑动步长设置为 50 ms 时,用不同长度时间窗截取 sEMG 生成 EFM 进行手势识别,实验结果如表 2 所示。

表 2 不同时间窗长识别准确率
Table 2 Recognition accuracy in different time windows

时间窗长/ms	手势识别准确率/%			
	被试 1	被试 9	被试 15	被试 26
100	82.49	81.36	87.51	86.05
150	83.52	82.97	88.43	86.78
200	87.84	85.30	92.47	90.95
250	86.33	83.79	90.17	88.41

由于生成样本过多,为提高实验效率,按男女比例选用 DB1 被试 1、9、15 和 26 的 sEMG 进行实验。当时时间窗长 100 和 150 ms 时,手势识别准确率较低。当时时间窗长 200 ms 时,准确率提升约 4%。当时时间窗长 250 ms 时,准确率出现下降。上述结果的原因在于时间窗长较短时,不同手势时域特征值相差较小,网络无法准确预测 EFM 样本类别。同时,时间窗长与生成样本数量成反比,时间窗较短样本较多导致网络学习的信息冗杂,时间窗较长样本较少导致网络学习的信息不足,两种情况均对网络预测产生负面影响。后续实验时间窗长设置 200 ms。

为验证不同特征通道数的手势识别效果,在深层特征提取模块设置不同数量卷积核进行手势分类。表 3 中结果可看出,卷积核数量为 64 时,不同被式手势识别的综合效果最好。卷积核数量为 128 和 256 时,虽然部分被试的准确率小幅提升,但网络中参数量剧增,模型优化困难,计算成本远高于卷积核数量为 64 的网络。卷积核数量为 32 时,提取的特征数量过少,手势分类准确率最低。后续实验卷积核数设置为 64。

3.3 网络消融实验

为验证模块化网络搭建的有效性,选用 EFM 输入

表 3 不同卷积核识别准确率

Table 3 Recognition accuracy in different convolutional kernel

卷积核数量	手势识别准确率/%			
	被试 1	被试 9	被试 15	被试 26
32	80.37	81.25	85.77	84.65
64	87.84	85.30	92.47	90.95
128	87.14	86.36	94.05	89.87
256	87.45	86.62	93.67	90.25

A、B 和 C 这 3 种网络模型,进行不同模型结构手势识别实验。A、B、C 型网络的区别在于网络深层特征提取模块层数的不同,分别为 A 中 2 层、B 中 3 层以及 C 中 4 层。同时,为验证添加 ECA 机制的网络手势识别性能提升,分别进行添加 ECA 与不添加 ECA 的网络手势识别实验,结果如图 7 所示。

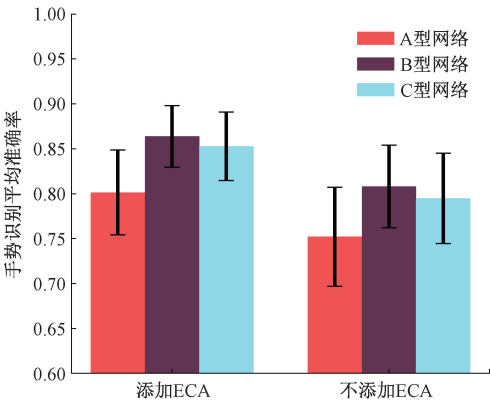


图 7 基于 EFM 的不同网络手势识别

Fig. 7 Gesture recognition in different network based on EFM

图 7 黑色插条代表不同被试手势识别标准差,识别效果越好则标准差越小。添加 ECA 机制后,网络的手势识别准确率提升,标准差减小,证明注意力机制对网络性能的提升。添加 ECA 的 B 型网络中,手势识别准确率最高,比 A 型和 C 型网络分别高出约 6% 和 1%,这说明加深网络虽然能提高准确率,但对于简单的单标签多分类手势识别任务,需要设置合适的网络深度,避免网络复杂化导致分类性能下降。不同网络结构的手势识别结果如表 4 所示。

表 4 不同网络手势识别结果

Table 4 The result of gesture recognition in different network (%)

网络型号	添加 ECA		不添加 ECA	
	准确率	标准差	准确率	标准差
A	80.14	4.72	75.21	5.52
B	86.39	3.44	80.81	4.60
C	85.29	3.81	79.48	5.03

ECA 机制 B 型网络输入 EFM 样本时,DB1 中 27 名被试 52 类手势识别混淆矩阵如图 8 所示。

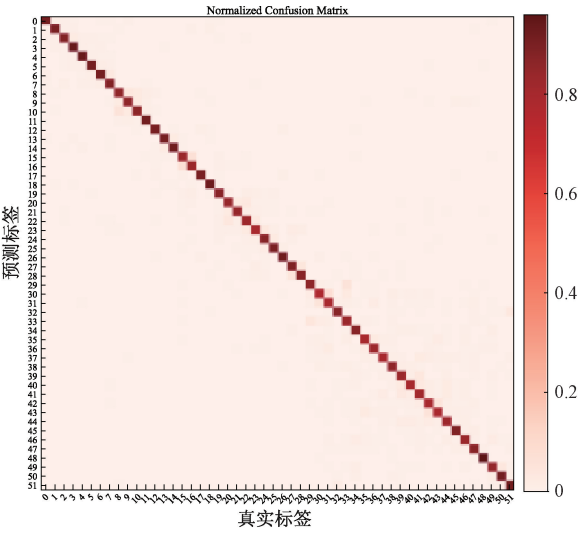


图 8 52 类手势最优识别混淆矩阵

Fig.8 Confusion matrix of best recognition of 52 types gesture

混淆矩阵中,矩阵元素代表经过归一化后网络样本预测的真实标签概率。分类样本落在对角线表示网络类别预测正确,落在其他位置则预测错误。图 8 中分类样本集中于对角线,表明本文方法的各种手势预测准确。

3.4 输入对比实验

为评估不同输入样本的手势识别结果,验证 EFM 手势识别的优势,分别用被试的 sEMG、肌电时域特征 (electromyography time-domain feature,ETF) 和 EFM 输入 ECA 机制 B 型网络进行手势识别。网络训练过程中,可

可视化 3 种输入的测试集手势平均识别准确率和误差损失的变化,如图 9 所示。

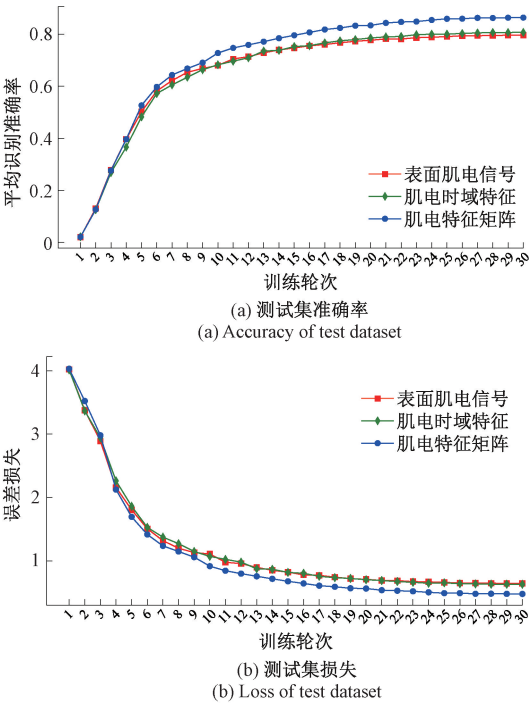


图 9 不同输入手势识别平均准确率和损失变化

Fig.9 Change in mean accuracy and loss of gesture recognition with different input

图 9 可看出 EFM 识别始终优于另外两种输入,印证了样本改进的优势。训练第 20 轮时,手势识别接近最高准确率和最低误差损失,在最后几轮中保持平稳。3 种输入的单被试手势识别结果如图 10 所示。

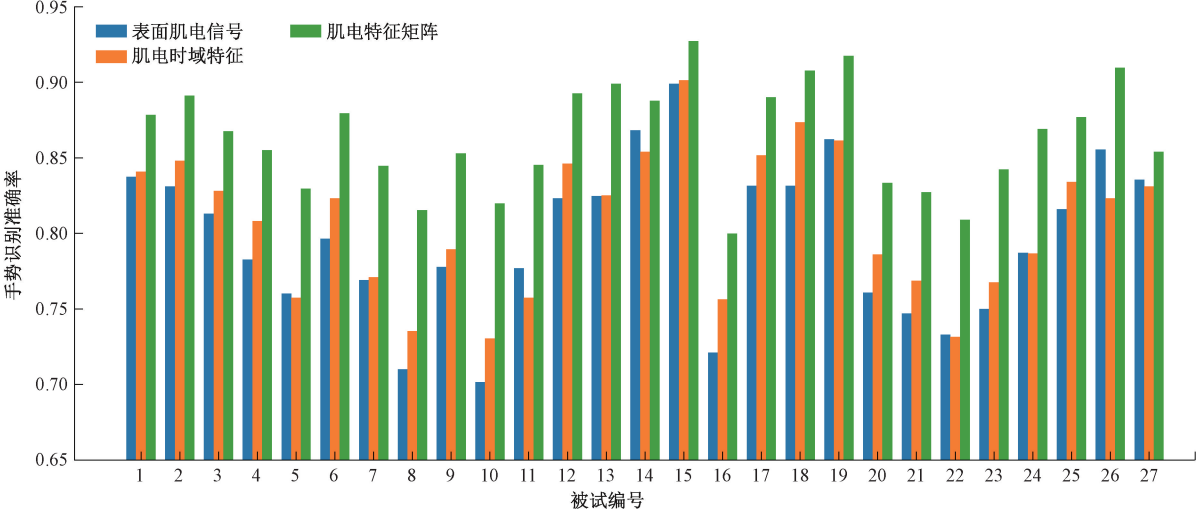


图 10 单被试不同输入手势识别准确率

Fig.10 Gesture recognition accuracy of single subject with different input

图 10 可看出单被试的 sEMG 和 ETF 手势识别结果相似,准确率均低于 EFM。并且 sEMG 和 ETF 的被试最

高和最低手势识别准确率相差 20%左右,比 EFM 高 7%左右。结果表明 EFM 提升了不同类别样本的区别性,一

定程度上缩小了 sEMG 差异性导致的不同被试间的手势识别差距。网络训练最终的 3 种输入手势识别结果如表 5 所示,EFM 的识别结果最优。

表 5 不同输入识别结果

Table 5 The result of recognition with different input		
输入样本	手势识别平均准确率/%	手势识别误差损失
sEMG	79. 64	0. 645
ETF	80. 69	0. 631
EFM	86. 39	0. 478

3. 5 与其他识别方法对比分析

本文方法与近年来用于 DB1 的手势识别研究方法进行对比,Du 等^[20]在手势识别网络中应用自适应批标准化。Tong 等^[21]提出一种长短时记忆网络与 CNN 并行的双流网络,分别提取 sEMG 时间和空间特征用于手势识别。Wang 等^[22]构造平稳小波变换(stationary wavelet packet transform,SWPT)sEMG 特征,输入 CNN 与循环神经网络的混合网络进行识别。李沿宏等^[23]将一维注意力机制引入多流卷积神经网络用于手势识别。不同方法识别结果如表 6 所示。

表 6 不同方法手势识别对比

Table 6 Comparison of gesture recognition in different methods				
网络模型	输入样本	时间窗长/ms	手势数量	平均准确率/%
DuNet ^[20]	sEMG	200	52	79. 50
TongNet ^[21]	sEMG	200	52	71. 22
WangNet ^[22]	SWPT	160	52	81. 80
ChengNet ^[14]	ETF	200	52	82. 54
LiNet ^[23]	sEMG	200	52	84. 39
ECA+CNN	EFM	200	52	86. 39

本文方法计算并量化机电特征之间的相关系数为 EFM,改善了输入样本的数据分布,增强了不同标签样本的差异性。EFM 用于手势识别时,使网络的学习能力得到提高。同时,在 CNN 中引入注意力机制赋予各通道特征不同的权重,精简网络收集的信息,提高重要信息占比。避免了信息冗杂弱化网络性能,从而提升分类效果。在 DB1 中 52 类手势识别平均准确率达到 86. 39%,结果优于其他的方法,验证了本文手势识别方法的有效性。

4 结 论

本文提出一种新颖的手势识别方法,对输入样本和网络模型进行改进。使用数据分布增强的 EFM 作为输入样本,在 CNN 中引入 ECA 机制赋予特征关键信息更高的权重。同时,设置深层特征提取模块层数模块化搭建网络,合理化分类模型深度,提高手势识别准确率。实验结果表明,本文方法在性能上优于近年来的手势识别方法,为多分类手势识别提供一种新的可行方案。

由于 sEMG 数据的被试来源有限且手部截肢患者数据较少,没有进行多被试之间的跨被试手势识别与手部截肢患者的手势识别研究。未来工作将针对跨被试与截肢患者的手势识别问题研究讨论,以实现手势识别在不同人群中的人机交互通用性。

参考文献

[1] 徐瑞,李志才,王雯婕,等. 基于肌电的人机交互控制策略及其应用与挑战[J]. 电子测量与仪器学报, 2020,34(2):1-11.

XU R, LI ZH C, WANG W J, et al. Human-computer interaction control strategies based on electromyography and their applications and challenges [J]. Journal of Electronic Measurement and Instrumentation, 2020, 34(2):1-11.

[2] 易靖国,程江华,库锡树. 视觉手势识别综述[J]. 计算机科学,2016,43(S1):103-108.

YI J G, CHENG J H, KU X SH. Review of gesture recognition based on vision [J]. Computer Science, 2016,43(S1):103-108.

[3] 肖茜,杨平,徐立波. 一种基于 MEMS 惯性传感器的手势识别方法[J]. 传感技术学报, 2013, 26(5): 611-615.

XIAO Q, YANG P, XU L B. A gesture recognition method based on MEMS IMU [J]. Chinese Journal of Sensor and Actuators,2013,26(5):611-615.

[4] LI K, ZHANG J, WANG L, et al. A review of the key technologies for sEMG-based human-robot interaction systems[J]. Biomedical Signal Processing and Control, 2020,62:102074.

[5] FANG C, HE B, WANG Y, et al. EMG-centered multisensory based technologies for pattern recognition in rehabilitation: State of the art and challenges [J]. Biosensors,2020,10(8):85.

[6] 陈艳丽,金纪东,陈国涛,等. 手部姿势与挥动速度识别的人机交互[J]. 电子测量与仪器学报, 2014, 28(1):56-61.

CHEN Y L, JIN J D, CHEN G T, et al. Human-computer interaction of hand posture and waving speed recognition[J]. Journal of Electronic Measurement and Instrumentation, 2014,28(1):56-61.

[7] 刘光达,董梦坤,张守伟,等. 基于 KPCA-SVM 的表面肌电信号疲劳分类研究[J]. 电子测量与仪器学报, 2021,35(10):1-8.

LIU G D, DONG M K, ZHANG SH W, et al. Research on fatigue classification of surface EMG signal based on KPCA and SVM [J]. Journal of Electronic Measurement and Instrumentation,2021,35(10):1-8.

[8] LECUN Y, BENGIO Y, HINTON G. Deep learning [J].

- Nature, 2015, 521(7553):436-444.
- [9] KETYKÓ I, KOVACS F, VARGA K Z. Domain adaptation for sEMG-based gesture recognition with recurrent neural networks [C]. 2019 International Joint Conference on Neural Networks (IJCNN), Budapest, Hungary, 2019:1-7.
- [10] DING Z, YANG C F, TIAN Z H, et al. sEMG-based gesture recognition with convolution neural networks [J]. Sustainability, 2018, 10(6):1865-1877.
- [11] ATZORI M, COGNOLATO M, MULLER 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):1-10.
- [12] ATZORI M, GIJSBERTS A, CASTELLINI C, et al. Electromyography data for non-invasive naturally-controlled robotic hand prostheses [J]. Scientific Data, 2014, 1(1):1-13.
- [13] GENG W D, DU Y, JIN W G, et al. Gesture recognition by instantaneous surface EMG images [J]. Scientific Reports, 2016, 6(1):1-8.
- [14] CHENG Y, LI G, YU M, et al. Gesture recognition based on surface electromyography-feature image [J]. Concurrency and Computation: Practice and Experience, 2021, 33(6):e6051.
- [15] WANG Q, WU B, ZHU P, et al. ECA-Net: Efficient channel attention for deep convolutional neural networks [C]. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020:11531-11539.
- [16] HU Y, WONG Y, WEI W, et al. A novel attention-based hybrid CNN-RNN architecture for sEMG-based gesture recognition [J]. PloS One, 2018, 13(10):e0206049.
- [17] PHINYOMARK A, PHUKPATTARANONT P, LIMSAKUL C. Feature reduction and selection for EMG signal classification [J]. Expert Systems with Applications, 2012, 39(8):7420-7431.
- [18] SHARMA S, SINGH S. Vision-based hand gesture recognition using deep learning for the interpretation of sign language [J]. Expert Systems with Applications, 2021, 182:115657.
- [19] IOFFE S, SZEGEDY C. Batch normalization: Accelerating deep network training by reducing internal covariate shift [C]. Proceedings of the 32nd International Conference on Machine Learning, Lille, France, 2015:448-456.
- [20] DU Y, JIN W, WEI W, et al. Surface EMG-based inter-session gesture recognition enhanced by deep domain adaptation [J]. Sensors, 2017, 17(3):458.
- [21] TONG R, ZHANG Y, CHEN H, et al. Learn the temporal-spatial feature of sEMG via dual-flow network [J]. International Journal of Humanoid Robotics, 2019, 16(4):515-524.
- [22] WANG L, FU J, ZHENG B, et al. Research on sEMG-based gesture recognition using the attention-based LSTM-CNN with stationary wavelet packet transform [C]. 2022 4th International Conference on Advances in Computer Technology, Information Science and Communications (CTISC), Suzhou, China, 2022:1-6.
- [23] 李沿宏, 江茜, 邹可, 等. 融合注意力机制的多流卷积肌电手势识别网络 [J]. 计算机应用研究, 2021, 38(11):3258-3263.
- LI Y H, JIANG Q, ZOU K, et al. Multi-stream convolutional myoelectric gesture recognition networks fusing attentional mechanisms [J]. Application Research of Computer, 2021, 30(11):3258-3263.

作者简介



赵世昊, 昆明理工大学硕士研究生, 主要研究方向为人机交互、表面肌电手势识别。
E-mail: 1137160540@qq.com



Zhao Shihao is now a M. Sc. candidate in Kunming University of Science and Technology of China. His main research interests include human-computer interaction, gesture recognition of surface electromyography.

周建华 (通信作者), 2008 年获得云南大学硕士学位, 现为昆明理工大学高级实验师、硕士生导师, 主要研究方向为人机交互、脑机智能融合。
E-mail: 742028837@qq.com

Zhou Jianhua (Corresponding author) received his M. Sc. degree from Yunnan University in 2008. He is now a senior experimentalist and master's supervisor at Kunming University of Technology. His main research interests include human-computer interaction and brain-computer intelligence fusion.



伏云发, 2012 年获得中国科学院沈阳自动化研究所博士学位, 现为昆明理工大学教授、硕士生导师、博士生导师, 主要研究方向为脑机接口、脑机智能融合、人机交互。
E-mail: fyf@ynu.edu.cn

Fu Yunfa received his Ph. D. degree from Shenyang Institute of Automation, Chinese Academy of Sciences in 2012. He is now a professor, master's supervisor and doctoral supervisor at Kunming University of Science and Technology. His main research interests include brain-computer interface, brain-computer intelligence fusion, human-computer interaction.