

基于改进 ResNet50 的表面肌电信号手势识别^{*}

牛群峰 石磊 贾昆明 桂冉冉 董鹏豪 王莉
(河南工业大学电气工程学院 郑州 450000)

摘要: 为了提高手势动作在类别众多且相似度高的情况下的识别精度,提出了一种基于连续小波变换和残差神经网络 ResNet50 的表面肌电信号手势识别方法。首先对 Ninapro DB2 和 DB3 的原始表面肌电信号进行预处理和连续小波变换,得到 Multi-sEMG Wavelet Map 数据集,然后送入改进的 ResNet50 模型进行识别分类。实验结果表明,改进后的 ResNet50 网络模型在 Multi-sEMG Wavelet Map DB2 和 DB3 中 17 种手势动作的平均准确率分别达到了 96.40% 和 94.11%,相比 ResNet50 网络模型方法提升了 4.87% 和 5.83%。实现了手势动作在类别繁多、相似度较高的情况下的精准识别。为基于非侵入式传感器和机器学习控制的假肢手提供了新方案。

关键词: 表面肌电信号;连续小波变换;Multi-sEMG Wavelet Map;ResNet50

中图分类号: TP18;TN911.7 **文献标识码:** A **国家标准学科分类代码:** 520.2

SEMG gesture recognition based on improved ResNet50

Niu Qunfeng Shi Lei Jia Kunming Gui Ranran Dong Penghao Wang Li
(College of Electrical Engineering, Henan University of Technology, Zhengzhou 450000, China)

Abstract: To enhance the recognition accuracy of gesture actions in scenarios with numerous and highly similar categories, a gesture recognition method for surface electromyographic signals based on continuous wavelet transform and residual neural network ResNet50 is proposed. The raw surface EMG signals collected are first preprocessed and continuously wavelet transformed to obtain the Multi-sEMG Wavelet Map dataset, and then fed into the improved ResNet50 model for recognition and classification. The experimental results show that the improved ResNet50 network model achieves an average accuracy of 96.40% and 94.11% for 17 gesture actions in Ninapro DB2 and DB3, respectively, which is an improvement of 4.87% and 5.83% compared to the ResNet50 network model method. Achieved accurate recognition of gesture actions in situations with numerous and highly similar categories. A new scheme is provided for prosthetic hand based on non-invasive sensors and machine learning control.

Keywords: surface electromyography; continuous wavelet transform; Multi-sEMG Wavelet Map; ResNet50

0 引言

近年来,随着传感器技术、信号处理技术、计算机科学和康复医疗技术的发展,基于非侵入式传感器和机器学习控制的假肢手被引入和研究^[1]。表面肌电信号(surface electromyography, sEMG)是一种由肌肉纤维的电活动引起的电信号,可以提供关于肌肉活动模式、力量、时序和协调性等信息^[2]。因其采集方便和非侵入性的特点,以 sEMG 作为假肢手外部控制源被广泛应用。其中,对

sEMG 的特征提取和信号分类是基于 sEMG 的假肢手控制系统的核心。

在过去的几年里,研究者们针对 sEMG 的特征提取和信号分类的研究主要采用手动选择合适的特征以及传统的机器学习分类方法。刘闯闯等^[3]以绝对平均值、中值频率等作为 sEMG 的特征向量,采用混合粒子群算法优化的支持向量机对 5 种手语动作进行分类识别,识别率达到 96.78%。赵诗琪等^[4]提取了 4 种手势的 4 通道 sEMG 的时域、频域、时频域和非线性特征构成特征集,并在特征降

收稿日期:2023-10-16

^{*} 基金项目:河南工业大学创新基金计划(2022ZKCJ03)、河南省科技研究计划(2013000210100)项目资助

维后送入支持向量机,平均识别率达到了 99.92%。

然而,当分类任务变得复杂即分类动作增加以及动作相似度高时,手动选择的特征无法完全捕捉到肌电信号中的所有相关信息,一些微妙但有益的信号会被忽略,对分类结果影响较大。Shen 等^[5]提取 29 种手势的 10 通道 sEMG 的时域、频域特征后送入线性判别分析模型和支持向量机,平均识别率分别为 65.62% 和 63.09%。Atzori 等^[6]对 12 通道的 52 种手势采取了 5 种信号特征提取,随后送入随机森林进行分类,平均识别率为 75.27%。

随着深度学习的发展,卷积神经网络(convolutional neural network,CNN)因其减少了人工提取特征的过程,能够自主提取图片中有用的特征信息,基于 CNN 的肌电识别方法正在被越来越多的研究者关注和应用^[7]。Cheng 等^[8]基于 10 通道 sEMG 生成了多 sEMG 特征图,并提出了一个 CNN 模型对 52 种手势动作的特征图进行分类,平均识别准确率达到 82.54%。曹书豪等^[9]对 52 种手势动作的 sEMG 滤波和扩充后,采用 ResNet34 对 sEMG 图进行分类,平均识别准确率达到 82.25%。

目前大多数研究尚未考虑随着网络深度的增加,可能导致微小特征丢失的问题,以及在多通道 sEMG 中不同通道提供的特征信息对最终识别效果的贡献程度差异的问题。因此,本文首先对原始 sEMG 进行连续小波变换(continuous wavelet transform, CWT),得到 Multi-sEMG Wavelet Map 数据集,然后在 ResNet50 基础上提出了一种基于改进 ResNet50 的手势识别算法。该方法通过 CWT 为模型的输入增加了更为丰富的时空分布和时频特征的信息,避免了手动提取信号特征时,有用信息缺失的问题;而且,通过添加多尺度结构和高效通道注意力机制模块,解决了在 Multi-sEMG Wavelet Map 中提取微小特征的问题以及自适应调整通道权重的问题,成功实现了手势动作在类别繁多、相似度高的情况下的精确识别。为基于非侵入式传感器和机器学习控制的假肢手提供了新方案。

1 材料与方 法

1.1 sEMG 数据库

Ninapro 是一个公开的多模态数据库,旨在促进对人类、机器人和假肢手以及基于机器学习的控制系统的研究^[10],是用于手势识别任务的公开数据集。

本文使用了 Ninapro 项目数据库中的第 2 个和第 3 个数据库的原始 sEMG,即 DB2 和 DB3^[11]。Ninapro DB2 和 DB3 是使用相同数据采集协议获得的两个数据库,作为本文比较对象。两个数据库唯一的区别是实验的受试者,其中 DB2 的受试者是 40 名健康的受试者,没有神经病变或上肢创伤;而 DB3 的受试者是 11 名经桡动脉截肢者。两个数据库均使用 Delsys Trigno Wireless EMG 系统,该系统具有 12 个无线电极(通道),以 2 kHz 的速率收集 49 种手势的肌肉的电活动。关于 Ninapro 数据库的更多细节见^[10]。在给出的 49 种手势中,本文分别选择 DB2

和 DB3 库中前 5 名的受试者的 17 种常见的手指和手掌动作作为研究对象。17 种手指和手掌动作分别为伸出大拇指;伸出食指和中指;无名指和小拇指并拢,其余张开;大拇指并拢,其余张开;五指张开;握拳;食指外伸,其余并拢;五指并拢;五指张开,沿食指外转;五指张开,沿食指内转;五指并拢,沿小拇指内转;五指并拢,沿小拇指外转;手腕内屈;手腕外伸;手腕朝桡骨方向偏;手腕朝尺骨方向偏;握拳伸腕。动作图片如图 1 所示。

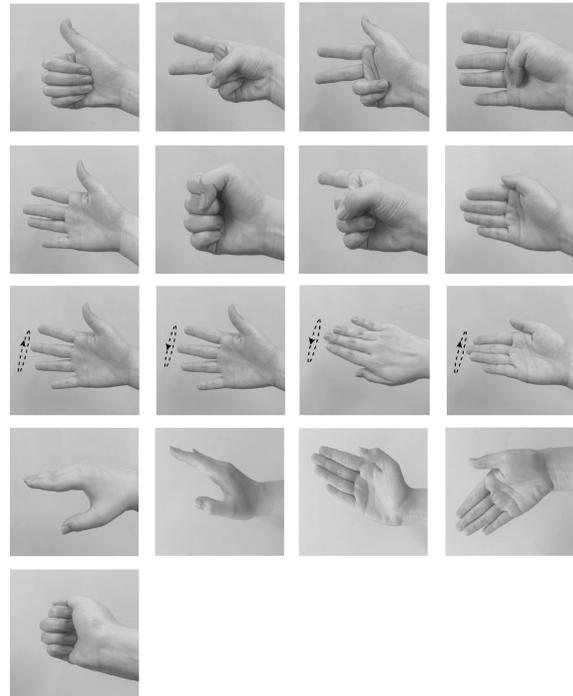


图 1 17 种手势动作
Fig. 1 17 gesture movements

1.2 sEMG 预处理

考虑到人体前臂的表面肌电信号相对较弱,再加上实验装置不同、个体差异以及设备采集过程中采集环境的影响,最终采集到的表面肌电信号中含有大量的噪声,如工频噪声、基线漂移等,严重影响了原始信号中关键信息的提取和识别^[12]。通常需要对直接从人类前臂采集的原始肌电信号进行预处理。Ninapro 数据库对表面肌电信号进行了初步预处理,包括信号全波整流、同步和重新标记。因此,只需要考虑采集设备的高频噪声。利用截断频率为 1 Hz 的一阶巴特沃斯滤波器,对每个通道的原始表面肌电信号进行低通滤波,以去除信号中隐藏的高频噪声。随后,对滤波后的 sEMG 进行 Minmax 归一化处理以便后续更好的提取信号特征。

为了实时处理表面肌电信号数据,使用窗口大小为 256 ms 的滑动窗口对处理后信号进行切片。为了引入时间变化和增加更多的样本,将每个窗口之间的步幅设置为小于窗口大小的 80 ms,从而导致连续样本之间存在一些重叠。

1.3 Multi-sEMG Wavelet Map

深度学习网络能够直接处理表面肌电信号,但 sEMG 是非平稳的,其中包含了多个频率成分和动态变化,直接处理会丢失信号中的隐藏信息^[13]。因此,本文引入了 CWT 作为表面肌电信号分析的关键工具。CWT 的独特之处在于能够提供高频分辨率,允许以局部化的方式观察信号的时域和频率演化^[14]。这一优点对于表面肌电信号的分析至关重要,因为这些信号往往包含了短暂的生理事件、频率成分的多样性以及时间上的动态变化。而且, CWT 具有自适应性,可以根据信号的不同部分的特性自动调整分辨率,能较好地处理非平稳信号^[15]。

综上,本文提出了一种基于 CWT 的表面肌电信号特征融合图,将其命名为 Multi-sEMG Wavelet Map,该图实现步骤如下。

使用 Morlet 小波对每个滑动窗口的表面肌电信号进行连续的小波变换。计算公式如下:

$$CWT(a, b) = \int_{-\infty}^{\infty} x(t) \cdot \psi^* \left(\frac{t-b}{a} \right) dt \quad (1)$$

式中: $CWT(a, b)$ 表示信号 $x(t)$ 的连续小波变换结果; $\psi^* \left(\frac{t-b}{a} \right)$ 表示小波函数的复共轭; a 表示尺度参数; b 表示平移参数。参数 a 在具体实现时,系统会分别根据信号的长度和特性来确定的。参数 b 在具体实现时,会在信号的全局范围内进行遍历,以覆盖整个信号的时间域。Morlet 小波函数为:

$$\psi(t) = \frac{1}{\sqrt{\pi f_b}} e^{j2\pi f_c t} e^{-\frac{t^2}{f_b}} \quad (2)$$

式中: $\psi(t)$ 表示 Morlet 小波函数; f_b 表示中心频率,取值为 1; f_c 表示中心频率,取值为 $\frac{1}{\pi}$ 。

对于每个通道的表面肌电信号而言,经 CWT 后,会形成一个 Single-sEMG Wavelet Map。收集所有的 Single-sEMG Wavelet Map 并形成 Multi-sEMG Wavelet Map(图 2)。Multi-sEMG Wavelet Map 提供了全局的时空分布和时频特征的信息,从而传达了有关肌肉协同作用的信息,这些信息有助于精确识别手势动作。该图的分辨率用 $L \times W$ 表示,其中 L 是肌电传感器数量与小波变换尺度数量的乘积, W 是当前滑动窗口长度。这里,12 个电极 \times 62 个尺度返回 $L = 744, 256$ ms 的窗口大小返回 $W = 512$ (频率是 2 kHz)。将信号值线性映射到 $[0, 255]$,形成单通道灰度图像,随后复制 3 张灰度图像并创建 RGB 图像。

分别将 DB2 和 DB3 数据库中前 5 名的受试者的 17 种 sEMG 按照上述步骤,得到 Multi-sEMG Wavelet Map DB2 数据集和 Multi-sEMG Wavelet Map DB3 数据集(简称为 DB2 和 DB3),分别包含 26 336 张和 31 473 张图像。将数据集以 8 : 2 的比例随机划分训练集和测试集,表 1 为 DB2 和 DB3 数据集。

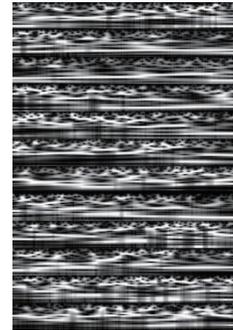


图 2 Multi-sEMG Wavelet Map
Fig. 2 Multi-sEMG Wavelet Map

表 1 DB2 和 DB3 数据集
Table 1 DB2 and DB3 dataset

手势编号	DB2		DB3	
	训练集	测试集	训练集	测试集
1	1 791	447	1 568	391
2	1 600	400	1 597	399
3	1 477	369	1 510	377
4	1 227	306	1 536	383
5	1 140	284	1 496	373
6	1 192	297	1 546	386
7	1 743	435	1 699	424
8	1 234	308	1 596	399
9	1 230	307	1 558	389
10	970	242	1 238	309
11	1 241	310	1 561	390
12	1 006	251	1 276	318
13	1 012	253	1 338	334
14	1 004	250	1 505	376
15	1 213	303	1 415	353
16	1 045	261	1 462	365
17	951	237	1 285	321
合计	21 076	5 260	25 186	6 287

2 手势识别模型

2.1 ResNet50 基线网络

ResNet50 网络是 He 等^[16]于 2015 年提出的一种 ResNet 模型。与传统的卷积神经网络不同,ResNet50 采用了残差连接的思想,即通过引入跳跃连接来解决训练深层网络时的梯度消失问题。这种跳跃连接允许信息在网络中直接传递,使得网络可以更轻松地学习到对输入的有益变换。

ResNet50 网络一共由 50 层网络构成,包含了 49 个卷积层和 1 个全连接层。其网络结构模型如图 3 所示。从图 3 可看到,Resnet50 网络结构可以分成 6 个部分,第 1 部分主要对输入的图片进行卷积、正则化、ReLU 激

活函数、最大池化的计算;第2~5部分由不同数量的残差块(residual block, Res Block)组成(图3)。其中,每个部分的第1个Res Block都是Conv Block,其余的Res Block均为Identity Block;最后一部分则是对得到的特征图进行平均池化计算和展平,再经过全连接层(fully connected, FC)得到最终的分类结果。

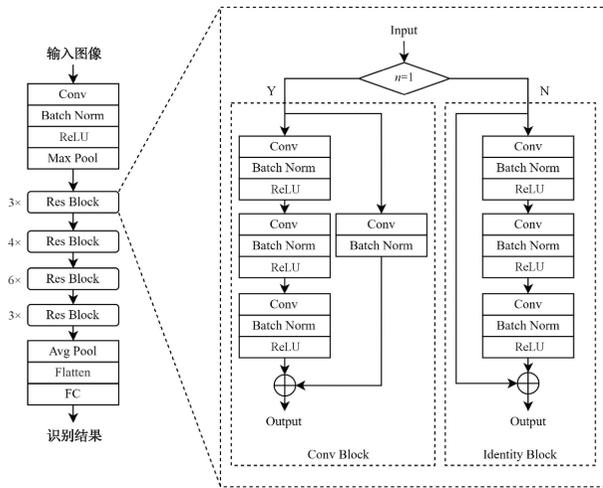


图3 ResNet50网络结构及Res Block结构
Fig. 3 ResNet50 network architecture and Res Block structure

2.2 改进 ResNet50 网络

本文针对 ResNet50 网络进行了两方面的改进:1) 本文将多尺度结构引入 ResNet50 网络模型中进行特征融合,提高网络对表面肌电信号特征融合图中的微小特征进行提取;2) 本文在模型中加入了高效通道注意力模块,为不同通道的特征信息赋予不同的权值,增强网络的特征表达能力,使得模型能够更好地学习图像中的复杂特征。改进 ResNet50 网络结构如图4所示。

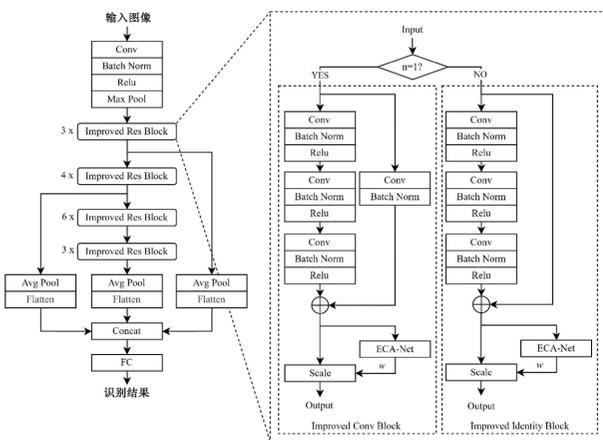


图4 改进 ResNet50 网络结构及改进 Res Block 结构
Fig. 4 Improved ResNet50 network architecture and improved Res Block structure

1) 多尺度特征融合

由于 sEMG 本身存在微弱、非平稳的特征,而且本文识别的 17 种手势动作中,部分动作相似度很高,除此之外,模型深度的增加会丢失部分特征。浅层特征图感受野较小,小目标细节信息较多,但缺少丰富的语义信息;深层特征图感受野较大,有丰富的语义信息,但包含的小目标信息较少。为防止网络模型在特征提取时忽略了一些微弱但有意义信息,本文在 ResNet50 网络中引入多尺度结构(multi-scale structure, MS)。将 ResNet50 网络的第 1 个和第 2 个 Res Block 的输出结果另起分支经过 Avg Pool 层和 Flatten 层,在通过 Concat 层进行特征融合进行输出该输出经过全连接层得到最终的分类结果。通过使用 MS 结构能对表面肌电信号特征融合图中的微小特征进行有效提取,提升网络识别准确率。

2) 高效通道注意力模块

本文输入网络的 Multi-sEMG Wavelet Map 是由 12 通道的肌电数据拼接而成,这样有助于网络模型对受试者手臂全局信息的捕获。然而,不同通道的肌电数据所提供的特征信息对最终的识别效果的贡献程度是不同的,因此本文在网络模型中加入 ECA-Net 注意力模块。

ECA-Net 通道注意力机制模块是由 Wang 等^[17]于 2020 年提出,旨在解决通道注意力机制的计算效率问题,提供了一种轻量级而有效的方式,以增强深度卷积神经网络的性能。这个模块的创新点在于引入了全局平均池化和自适应的权重调整,使得通道注意力的计算量大幅度降低,同时保持了良好的性能。不同于挤压与激励网络对通道实行降维策略来实现权值分配,ECA-Net 使用一维卷积来实现跨通道的维度交互,避免了降维。在保证性能不下降的同时又降低了复杂度。实验证明降低通道维度会对注意力机制的学习模式产生影响,其结构如图 5 所示。

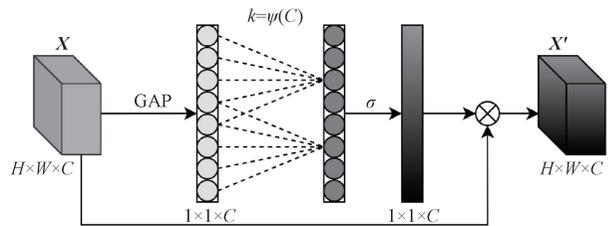


图5 ECA-Net 模块结构
Fig. 5 ECA-Net module structure diagram

首先输入特征 $X^{H \times W \times C}$ 通过全局平均池化,在空间维度对输入特征进行聚合,获得全局信息,得到 $X^{1 \times 1 \times C}$ 。全局平均池化的公式如下:

$$y = \frac{1}{H \times W} \sum_a^H \sum_b^W x_i(a, b) \quad (3)$$

式中: x_i 表示大小为 $H \times W$ 的第 i 个特征图; y 表示得到的全局特征。

其次,根据通道数 C 自适应计算出局部跨通道交互的覆盖范围 k 。自适应公式如下:

$$k = \psi(C) = \left\lfloor \frac{\log_2 C}{\gamma} + \frac{b}{\gamma} \right\rfloor_{\text{odd}} \quad (4)$$

式中: C 表示通道数量; b 和 γ 都是常数, $b = 1, \gamma = 2$; $\lfloor t \rfloor_{\text{odd}}$ 表示 t 的最邻近奇数。

然后,通过大小为 k 的一维卷积核卷积实现每个通道及其 k 个近邻通道的交互。一维卷积公式如下:

$$w = \sigma(C_1 D_k(y)) \quad (5)$$

式中: w 表示通道权重; σ 表示 Sigmoid 函数; $C_1 D$ 表示一维卷积; k 表示卷积核大小; y 为全局平均池化的结果。

本文将 ECA-Net 添加在每一个 Res Block 中,在 Conv Block 和 Identity Block 的框架基础上,在其结果输出部分额外添加一个分支进入 ECA-Net 得到通道权重 w 并将权重与原输出点积得到最终得结果。Improved Conv Block 和 Improved Identity Block 结构如图 4 所示。该策略使模型在每个阶段自适应地调整通道的权重,从而提高对各个阶段特征的感知能力。

3 实验与结果分析

3.1 实验环境

本文试验基于 Windows10 操作系统, GPU 为 GeForce GTX 3080(10 GB 显存), 处理器为 Inter(R) Core (TM) i7-12 700 K CPU@3.61 GHz, 运行内存为 64 G。实现 Multi-sEMG Wavelet Map 的开发环境使用 MATLAB2020b。模型的搭建与训练测试的开发环境使用 Pycharm, 深度学习框架采用 PyTorch, 并行计算框架使用 CUDA 11.3 版本, 使用的 Python 版本为 3.8。

3.2 参数设置

实验时,随机将训练集的图片顺序打乱,将输入的图片分辨率调整为 224×224 。随后按照均值 0.5 和标准差 0.5 对输入图片的每个通道进行标准化处理。实验模型的优化器选择随机梯度下降(stochastic gradient descent, SGD),其初始学习率 $lr = 0.002$ 、初始学习率因子 $lrf = 0.1$ 、动量设置为 0.9、权重衰减为 10^{-4} , 并使用 Nesterov accelerated gradient 算法加速模型参数的更新。随后,设计一个基于余弦函数的学习率调度器,会根据当前训练的最大迭代次数 epoch,更新 lrf ,使其介于 $0.1 \sim 1$,并随着训练过程线性增加至 1,以实现余弦退火学习率调度策略。损失函数选用交叉熵函数,交叉熵刻画了实际输出概率与期望输出概率分布之间的距离,交叉熵的值越小,表示模型训练过程的中学习效果越好^[18]。设置批量大小 batch size 为 32。最大迭代次数 epoch 为 80 次。每一轮迭代后,使用测试集对模型的准确性进行评估,并保留每次评估结果及模型参数。

3.3 模型评估指标

本文使用准确率(accuracy, ACC)、精确率(precision, P)、召回率(recall, R)、F1 分数作为模型性能的评估指标。

ACC 表示模型正确预测的样本数量与总样本数量之比:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \quad (6)$$

P 表示所有被分类为正例的样本中,模型正确预测为正例的样本数量与总预测为正例的样本数量之比:

$$P = \frac{TP}{TP + FP} \quad (7)$$

R 表示所有实际正例样本中,模型正确预测为正例的样本数量与总实际正例样本数量之比:

$$R = \frac{TP}{TP + FN} \quad (8)$$

F1 分数是综合考虑了 P 和 R 的指标,它是 P 和 R 的调和平均值 F1 值分数越高,表示模型整体越好。

$$F1 = 2 \times \frac{P \times R}{P + R} \quad (9)$$

式中: TP 表示被模型正确地预测为正例的样本数量; TN 表示被模型正确地预测为反例的样本数量; FP 表示被模型错误地预测为正例的样本数量; FN 表示被模型错误地预测为反例的样本数量。

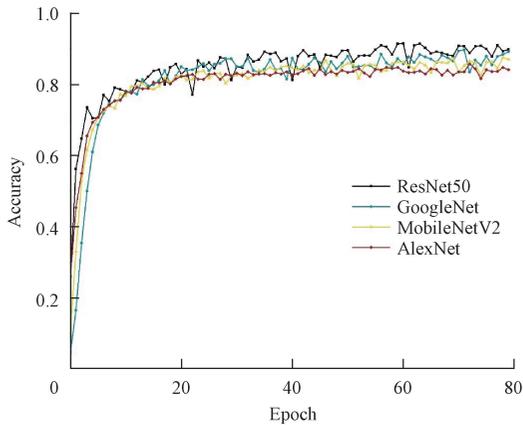
3.4 多种 CNN 基线网络的性能比较

为了证明选用 ResNet50 作为基线网络的优越性,分别选取了 AlexNet^[19]、GoogleNet^[20]、MobileNetV2^[21] 与 ResNet50 分别在 DB2 和 DB3 数据集上进行性能比较。不同 CNN 基线网络的性能指标比较如表 2 和图 6 所示。

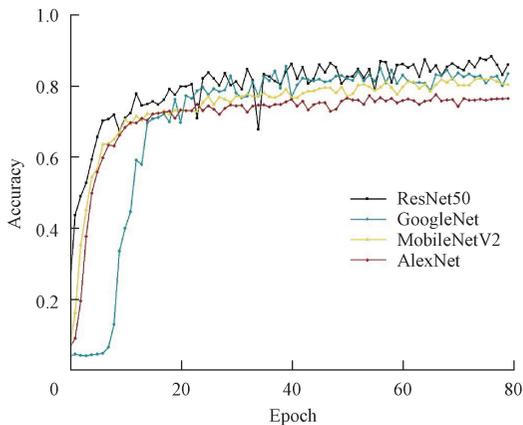
表 2 不同 CNN 基线网络的性能指标

数据集	模型	准确率	精确率	召回率	F1 分数
DB2	AlexNet	85.76	85.96	85.47	85.71
	MobileNetV2	87.54	87.83	87.28	87.55
	GoogleNet	89.71	89.64	89.36	89.50
	ResNet50	91.53	91.46	91.32	91.39
DB3	AlexNet	77.56	77.62	77.34	77.48
	MobileNetV2	82.26	82.51	82.05	82.28
	GoogleNet	85.52	85.47	85.17	85.32
	ResNet50	88.28	88.19	87.94	88.06

从表 2 和图 6 可以看出, AlexNet 表现最差,在 2 个数据集上的平均准确率为 81.66%,这是由于 AlexNet 是较早期的卷积神经网络,网络框架相对较浅,当分类任务类别复杂时表现不佳;本文选择的基线模型 ResNet50 表现最好,在两个数据集上的准确率分别为 91.53% 和 88.28%,召回率也达到了 91.39% 和 88.06%。GoogleNet 和 MobileNetV2 在两个数据集上的平均准确率分别为 87.62% 和 84.90%, GoogleNet 包含 Inception 结构造成的, Inception 结构包含多种尺度特征的融合,这使得 GoogleNet 网络能获得优秀性能,而 MobileNetV2



(a) 在DB2数据集上进行性能评估
(a) Performance evaluation on DB2 dataset



(b) 在DB3数据集上进行性能评估
(b) Performance evaluation on DB3 dataset

图6 不同CNN基线网络的准确率与循环次数曲线

Fig. 6 Accuracy and number of epochs for different CNN baseline networks

网络引入了 Inverted residual block,使得高维信息通过 ReLU 激活函数后丢失的信息更少,同时,使用 linear bottleneck 替代非线性激活,防止激活函数在低维转换时过滤太多有效信息,从而提高了分类性能。从 GoogleNet 和 MobileNetV2 网络分类性能来看,低维度的特征包含了 Multi-sEMG Wavelet Map 的细节,而低维信息的丢失会降低网络在 Multi-sEMG Wavelet Map 数据集上的性能。综上所述,本文选用的 ResNet50 作为基线模型性能最佳。

3.5 不同多尺度结构网络之间的性能评估

本文在对 ResNet50 网络的结构引入 MS 结构时,是在 Res Block 后添加额外的分支经 Avg Pool 层和 Flatten 层,在通过 Concat 层进行特征融合进行输出。会出现如下情况:1)分别从第 1~3 个 Res Block 后添加新的分支;2)分别从第 1 个和第 2 个 Res Block 后添加新的分支;3)分别从第 1 个和第 3 个 Res Block 后添加新的分支;4)分别从第 2 个和第 3 个 Res Block 后添加新的分支;

5)从第 1 个 Res Block 后添加新的分支;6)从第 2 个 Res Block 后添加新的分支;7)从第 3 个 Res Block 后添加新的分支。针对这 7 种情况,本文分别将 DB2 和 DB3 数据集送入将对应的网络模型进行识别,最终得到的识别准确率如表 3 所示。

表3 不同 CNN 基线网络的性能指标

Table 3 Performance metrics of different CNN baseline networks		networks (%)			
数据集	模型	准确率	精确率	召回率	F1 分数
DB2	ResNet50	91.53	91.46	91.32	91.39
	MS-ResNet50(1~3)	93.68	93.54	93.47	93.50
	MS-ResNet50(1&2)	94.37	94.43	94.38	94.40
	MS-ResNet50(1&3)	93.32	93.25	93.33	93.29
	MS-ResNet50(2&3)	92.73	92.68	92.53	92.60
	MS-ResNet50(1)	94.10	94.03	93.95	93.99
	MS-ResNet50(2)	93.77	93.82	93.57	93.69
	MS-ResNet50(3)	92.36	92.33	92.19	92.26
DB3	ResNet50	88.28	88.19	87.94	88.06
	MS-ResNet50(1~3)	91.07	90.98	90.83	90.90
	MS-ResNet50(1&2)	91.42	91.33	91.28	91.30
	MS-ResNet50(1&3)	90.63	90.62	90.45	90.53
	MS-ResNet50(2&3)	90.26	90.19	90.06	90.12
	MS-ResNet50(1)	91.21	91.35	91.16	91.25
	MS-ResNet50(2)	91.01	90.83	90.77	90.80
	MS-ResNet50(3)	89.91	90.03	89.75	89.89

从表 3 可以看出,引入情况 2 后,MS-ResNet50(1&2)的实验效果最好,分别在 DB2 和 DB3 的数据集上的识别准确率分别提升了 2.84% 和 3.14%,达到了 94.37% 和 91.42%。同时,从表 3 也能够看出,在越深层的 Res Block 引入多尺度结构,最终的识别率提升越低。证明在原网络基础上引入 MS 结构得到的多尺度融合特征能够提升最终的识别效果,但是,需要添加在网络浅层的 Res Block,这样最终得到的融合特征既包含了浅层特征也包含了深层特征,能够较好提升网络识别准确率。否则,最终得到的融合特征中包含了过多的深层特征,会导致特征冗余,网络无法达到最佳效果。

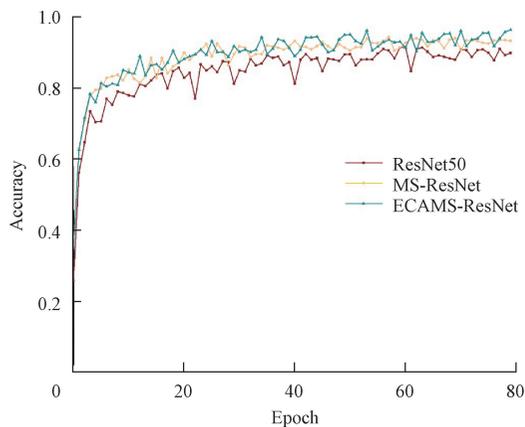
3.6 消融实验

为了验证第 2.2 节对 ResNet50 网络中的两处改进对模型性能确有提升设计了混淆实验。将加入最优 MS 结构后的网络记为 MS-ResNet50;在 MS-ResNet50 基础上加入 ECA-Net 通道注意力机制模块后的网络记为 ECAMS-ResNet50。消融实验分别在 DB2 和 DB3 数据集上进行了评估。消融实验的结果和每处改进的后网络的训练过程对比如表 4 和图 7 所示。

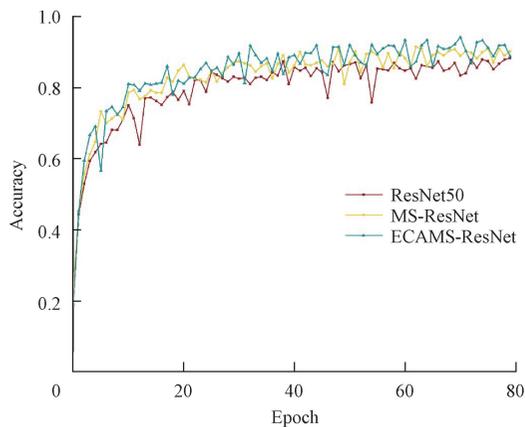
表 4 不同 CNN 基线网络的性能指标

Table 4 Performance metrics of different CNN baseline networks (%)

数据集	模型	准确率	精确率	召回率	F1 分数
DB2	ResNet50	91.53	91.46	91.32	91.39
	MS-ResNet50	94.37	94.58	94.14	94.36
	ECAMS-ResNet50	96.40	96.34	96.21	96.27
DB3	ResNet50	88.28	88.19	87.94	88.06
	MS-ResNet50	91.42	91.36	91.21	91.28
	ECAMS-ResNet50	94.11	94.08	93.89	93.98



(a) 在DB2数据集上进行性能评估
(a) Performance evaluation on DB2 dataset



(b) 在DB3数据集上进行性能评估
(b) Performance evaluation on DB3 dataset

图 7 消融实验的准确率与循环次数曲线

Fig. 7 Accuracy and number of epochs for ablation experiments

从表 4 和图 7 可以看出,在第 1 个和第 2 个 Res Block 后引入 MS 结构并得到多尺度融合特征后,MS-ResNet50 相较于 ResNet50 识别率得到了提高,其优点在实验中已经说明。在完成网络的多尺度结构后,在每一个 Res Block 中加入 ECA-Net,得到的 ECAMS-ResNet50 网络相较于 MS-ResNet 识别准确率分别提高了 2.03% 和

2.69%,达到了 96.40% 和 94.11%。同时证明通过调整不同通道的肌电数据所提供的特征信息的权重,对提升网络性能是有益的。最终,ECAMS-ResNet50 网络相较于 ResNet50 在 DB2 和 DB3 数据集上准确率分别提升了 4.87% 和 5.83%,达到了 96.40% 和 94.11%。在两个数据集上的平均精确率、平均召回率、平均 F1 分数分别达到了 95.21%、95.05% 和 95.13%。表明本文改进的网络模型在自建的 Multi-sEMG Wavelet Map DB2 数据集和 Multi-sEMG Wavelet Map DB3 数据集中表现良好。

3.7 不同网络模型对比

为了进一步验证本文设计的网络模型在基于 sEMG 的手势识别领域的优越性,将其与近几年该领域出现的网络模型进行了比较实验。本文和近几年该领域出现的网络模型在 Ninapro DB2 和 DB3 数据集上的实验结果如表 5 所示。

表 5 不同方法的分类结果

Table 5 Classification results of different methods

数据集	方法	模型	准确率/%
DB2	文献[22]	CNN	76.34
	文献[23]	Multi-CNN	83.70
	文献[24]	ALR-CNN	87.90
	本文	ECAMS-ResNet50	96.40
DB3	文献[25]	CNN	>70
	文献[22]	CNN	73.31
	文献[24]	ALR-CNN	85.58
	本文	ECAMS-ResNet50	94.11

从表 5 可以看出,本文设计的 ECAMS-ResNet50 网络模型准确率较其他方法有较大提升,在基于 sEMG 的手势识别领域具有一定的优势。

4 结 论

为了解决基于 sEMG 的手势识别中的一系列问题—当手势动作多且动作相似度高时,手动提取的 sEMG 的时域、频域等特征会影响分类器的识别准确率;随着深度学习网络深度的增加,微小特征存在丢失问题;在多通道 sEMG 中不同通道提供的特征信息对最终识别效果的贡献程度存在差异。本文提出了一种基于 CWT 和 ECAMS-ResNet50 网络的手势识别算法。首先,分别将 Ninapro DB2 和 DB3 数据库中 5 名受试者的 17 种手势动作的原始 sEMG 经划分时间窗后进行滤波、归一化等预处理,再通过 CWT,得到 Multi-sEMG Wavelet Map DB2 数据集和 Multi-sEMG Wavelet Map DB3 数据集。随后在 ResNet50 网络基础上加入 MS 结构,使得到的融合特征既包含了浅层特征也包含了深层特征,丰富特征信息并在每一个 Res Block 中引入 ECA-Net 通道注意力机制模块来调整不同通道的肌电数据所提供的特征信息的权重加

强模型提取特征的能力。最终 ECAMS-ResNet50 网络在 Multi-sEMG Wavelet Map DB2 和 Multi-sEMG Wavelet Map DB3 数据集上的识别准确率分别为 96.40% 和 94.11%。试验结果表明,ECAMS-ResNet50 网络在两个数据集上的平均准确率相比基线网络 AlexNet、MobileNetV2、GoogleNet、ResNet50 分别提高了 13.60%、10.36%、7.64%、5.35%,在两个数据上的 F1 分数也分别达到了 96.27% 和 93.98%,模型性能提升显著。为基于非侵入式传感器和机器学习控制的假肢手提供了新方案。

参考文献

- [1] SRI-IESARANUSORN P, CHAIYAROJ A, BUEKBAN C, et al. Classification of 41 hand and wrist movements via surface electromyogram using deep neural network [J]. *Frontiers in Bioengineering and Biotechnology*, 2021 (9): 548357.
- [2] 石欣,范智瑞,张杰毅,等.基于 LMS-随机森林的肌电信号下肢动作快速分类[J]. *仪器仪表学报*, 2020, 41(6): 218-224.
- SHI X, FAN Z R, ZHANG J Y, et al. Rapid classification of lower limb movements based on LMS-random forest for electromyographic signals [J]. *Chinese Journal of Scientific Instrument*, 2020, 41(6): 218-224.
- [3] 刘闯闯,朱正伟.基于 HPSO-SVM 的多传感器手语识别方法研究[J]. *电子测量技术*, 2021, 44(10): 57-65.
- LIU CH CH, ZHU ZH W. Research on multi-sensor sign language recognition method based on HPSO-SVM [J]. *Electronic Measurement Technology*, 2021, 44(10): 57-65.
- [4] 赵诗琪,吴旭洲,张旭,等.利用表面肌电进行手势自动识别[J]. *西安交通大学学报*, 2020, 54(9): 149-156.
- ZHAO SH Q, WU X ZH, ZHANG X, et al. Automatic gesture recognition using surface electromyography [J]. *Journal of Xi'an Jiaotong University*, 2020, 54(9): 149-156.
- [5] SHEN S, GU K, CHEN X R, et al. Gesture recognition through sEMG with wearable device based on deep learning [J]. *Mobile Networks and Applications*, 2020, 25(6): 2447-2458.
- [6] 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 Neurorobotics*, 2016, DOI:10.3389/fnbot.2016.00009.
- [7] 陈思佳,罗志增.基于长短时记忆和卷积神经网络的手势肌电识别研究[J]. *仪器仪表学报*, 2021, 42(2): 162-170.
- CHEN S J, LUO ZH Z. Research on gesture electromyography recognition based on long short-term memory and convolutional neural network[J]. *Chinese Journal of Scientific Instrument*, 2021, 42(2): 162-170.
- [8] CHENG Y, LI G, YU M, et al. Gesture recognition based on surface electromyography-feature image[J]. *Concurrency and Computation: Practice and Experience* 2020, DOI:10.1002/cpe.6051.
- [9] 曹书豪,许成哲.基于深度学习模型的表面肌电信号手势动作识别算法研究[J]. *科技视界*, 2019(1): 118-120.
- CAO SH H, XU CH ZH. Research on gesture recognition algorithm of surface electromyography signals based on deep learning models[J]. *Science & Technology Vision*, 2019(1): 118-120.
- [10] ATZORI M, GIJSBERTS A, CASTELLINI C, et al. Electromyography data for non-invasive naturally-controlled robotic hand prostheses[J]. *Nature*, 2014, DOI:10.1038/sdata.2014.53.
- [11] ATZORI M, GIJSBERTS A, CASTELLINI C, et al. Effect of clinical parameters on the control of myoelectric robotic prosthetic hands[J]. *Journal of Rehabilitation Research & Development*, 2016, 53(3): 345-358.
- [12] 林锦荣,谭北海,谢胜利.基于 ADS1294 的表面肌电信号采集系统的设计[J]. *医疗卫生装备*, 2015, 36(1): 4.
- LIN J R, TAN B H, XIE S L. Design of surface electromyography signal acquisition system based on ADS1294[J]. *Medical and Health Equipment*, 2015, 36(1): 4.
- [13] 裴晓敏,宋佳强,曹江涛,等.基于 MEMD 和 TK 能量算子的肌电信号手势识别[J]. *电子测量与仪器学报*, 2021, 35(1): 82-87.
- PEI X M, SONG J Q, CAO J T, et al. Gesture recognition of surface electromyography signals based on MEMD and TK energy operator[J]. *Journal of Electronic Measurement and Instrumentation*, 2021, 35(1): 82-87.
- [14] 向小民,汪杰,卢云.基于小波时频图和 AlexNet 网络的电弧故障诊断方法[J]. *国外电子测量技术*, 2022, 41(10): 170-177.
- XIANG X M, WANG J, LU Y. Arc fault diagnosis method based on wavelet time-frequency map and AlexNet network[J]. *Foreign Electronic Measurement Technology*, 2022, 41(10): 170-177.
- [15] 孙曙光,张婷婷,王景芹,等.基于连续小波变换和

- MTL-SEResNet的断路器故障程度评估[J]. 仪器仪表学报, 2022, 43(6): 162-173.
- SUN SH G, ZHANG T T, WANG J Q, et al. Circuit breaker fault severity assessment based on continuous wavelet transform and MTL-SEResNet[J]. Chinese Journal of Scientific Instrument, 2022, 43(6): 162-173.
- [16] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition [C]. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), 2016.
- [17] WANG Q, WU B, ZHU P, et al. Supplementary material for ECA-Net: Efficient channel attention for deep convolutional neural networks[C]. Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 13-19.
- [18] 高震宇, 王安, 董浩, 等. 基于卷积神经网络的烟丝物质组成识别方法[J]. 烟草科技, 2017, 50(9): 68-75.
- GAO ZH Y, WANG A, DONG H, et al. Identification method of tobacco material composition based on convolutional neural network[J]. Tobacco Science & Technology, 2017, 50(9): 68-75.
- [19] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Image net classification with deep convolutional neural networks [C]. 25th International Conference on Neural Information Processing Systems, 2012: 1097-1105.
- [20] 章敏敏, 徐和平, 王晓洁, 等. 谷歌 TensorFlow 机器学习框架及应用[J]. 微型机与应用, 2017, 36(10): 58-60.
- ZHANG M M, XU H P, WANG X J, et al. Google tensorflow machine learning framework and its applications [J]. Journal of Microcomputer & Its Applications, 2017, 36(10): 58-60.
- [21] HOWARD A G, ZHU M, CHEN B, et al. MobileNets: Efficient convolutional neural networks for mobile vision applications [J]. ArXiv preprint, arXiv:1704.04861, 2017.
- [22] ZHAI X L, BETH J, CHAN R H M, et al. Self-recalibrating surface EMG pattern recognition for neuroprosthesis control based on convolutional neural network[J]. Frontiers in Neuroscience, 2017, DOI: 10.3389/fnins.2017.00379.
- [23] WEI W, DAI Q, WONG Y, et al. Surface electromyography-based gesture recognition by multi-view deep learning [J]. IEEE Transactions on Biomedical Engineering, 2019, 66(10): 2964-2973.
- [24] FATAYER A, GAO W P, FU Y L. sEMG-based gesture recognition using deep learning from noisy labels[J]. IEEE Journal of Biomedical and Health Informatics. 2022, 26(9): 4462-4473.
- [25] GOPAL P, GESTA A, MOHEBBI A. A systematic study on electromyography-based hand gesture recognition for assistive robots using deep learning and machine learning models [J]. Sensors, 2022, 22(10): 3650.

作者简介

牛群峰, 博士, 副教授, 博士生导师, 主要研究方向为智能仪器设备与自动化装置、人机交互。

石磊, 硕士研究生, 主要研究方向为深度学习、康复医疗机器人。

E-mail: shil970614@163.com

王莉(通信作者), 博士, 教授, 硕士生导师, 主要研究方向为人工智能、仿生传感检测技术。

E-mail: hautwangli@163.com