基于全局与局部肌电特征交互的手势识别网络

肖城钢, 闵华松[†]

(武汉科技大学信息科学与工程学院,湖北武汉 430081)

摘要:为了更有效地捕捉肌电信号中的长期动态依赖关系和局部细节信息,减少固有肌电特征信息损失对手势 分类精度的影响,本文提出一种基于全局-局部特征交互的手势识别网络GL-EMG-Net.首先,融合空洞卷积和多头 自注意力机制,设计全局特征提取模块Global-DT,提取肌电信号中的全局信息;然后,借助深度可分离卷积和注意 力机制,设计局部特征提取模块Local-SK捕捉肌电信号中不同尺度的局部细节信息,并将提取的细节信息通过反馈 机制反馈至Global-DT模块,完成局部特征与全局特征的交互;最后,将全局特征与局部特征融合后进行分类.实验 结果表明,该手势识别网络在Ninapro DB5数据集的52种手势和实际12种常见手势分类中,表现出较高的手势分类 精度和较强的鲁棒性.

关键词:表面肌电信号;手势识别;空洞卷积网络;注意力机制;特征融合

引用格式: 肖城钢, 闵华松. 基于全局与局部肌电特征交互的手势识别网络. 控制理论与应用, 2025, 42(3): 609 – 617

DOI: 10.7641/CTA.2023.30204

Gesture recognition network based on the interaction of global and local myoelectric features

XIAO Cheng-gang, MIN Hua-song[†]

(School of Information Science and Engineering, Wuhan University of Science and Technology, Wuhan Hubei 430081, China)

Abstract: In order to capture the long-term dynamic dependencies and local detail information in the electromyography (EMG) signal more effectively and reduce the impact of the loss of intrinsic EMG feature information on the gesture classification accuracy, we propose a gesture recognition network the global and local-electromyography-network (GL-EMG-Net) based on the interaction of global and local features. Firstly, the dilated convolution and multi-head self-attention mechanism are integrated to design the global feature extraction block the global-dilation transformer (Global-DT) to extract the global information in the EMG signal. Then, with the help of the depth separable convolution and attention mechanism, the local feature extraction block the local-selective kernel (Local-SK) is designed to capture the local detail information of different scales in the EMG signal, and feedback the extracted detail information to the Global-DT module through the feedback mechanism to complete the interaction between local features and global features. Finally, the global features are fused for classification. The experimental results show that the gesture recognition network shows high gesture classification accuracy and strong robustness in the 52 gestures of Ninapro DB5 dataset and 12 actual common gestures.

Key words: surface electromyography; hand gesture recognition; dilated convolutional networks; attention mechanism; feature fusion

Citation: XIAO Chenggang, MIN Huasong. Gesture recognition network based on the interaction of global and local myoelectric features. *Control Theory & Applications*, 2025, 42(3): 609 – 617

1 引言

表面肌电信号(surface electromyography, sEMG) 是伴随人体肌肉活动而产生的一种生物电信号,负载 着丰富的人体肢体运动信息^[1].因其低成本和无创性

收稿日期: 2023-04-10; 录用日期: 2023-11-03.

本文责任编委:杨辰光.

的特点,基于sEMG的人体意图识别在各个领域都有涉及,如人体假肢手^[2]、机器人控制等^[3].

虽然基于sEMG的手势识别在各领域的应用越来 越广泛,但其仍存在特征提取不充分、易受噪声干扰

[†]通信作者. E-mail: mhuasong@wust.edu.cn; Tel.: +86 13971365898.

国家自然科学基金项目(62073249),国家重点研发计划项目(2022YFB4700400)资助.

Supported by the National Natural Science Foundation of China (62073249) and the National Key Research and Development Program of China (2022YFB4700400).

和手势分类类别有限等问题^[4].通过特征工程的方法 手动提取sEMG的时频域特征会使sEMG中有效信息 丢失^[5],造成提取到的特征不能最大程度地表征肢体 运动的动作模式,从而达不到理想的手势分类效果^[6].

近年来,为了更有效提取 sEMG 中蕴含的动作信 息,相关研究人员采用深度学习的方法提取sEMG中 的固有特征. 文献[7-9]利用卷积神经网络(convolution neural network, CNN)对原始sEMG的特征进行提 取,证明CNN能够有效提取sEMG中的动作信息,但 由于CNN感受野的局限性,其无法有效提取长期动态 依赖关系和全局特征,手势类别数的增加会使识别准 确率大大降低^[10]. Liu等^[11]设计双流空洞卷积网络捕 获肌电序列中的整体变化趋势和全局信息,10种手势 动作准确率为93.69%,但其忽略了肌电信号中的局部 细节信息.上述文献仅考虑了肌电序列中长期依赖的 动态关系或局部细节信息,未同时提取这两种特征. 然而在异常检测和交通流量预测领域,发现将全局-局部特征交互融合可以让模型达到更好的效果,Yu^[12] 提出一种基于Transformer和双流时间卷积网络(temporal convolutional network, TCN)的异常检测和诊断 模型,利用空洞卷积和因果卷积分别设计全局特征提 取模块和局部特征提取模块,其检测效果提升明显且 模型的训练时间大大减少. Ren等^[13]在交通流量预测 方面同时提取交通流信号中的全局特征和局部特征, 提出一种新的全局-局部TCN模型来捕捉交通流量中 非线性的时空依赖关系,在真实数据集上,该模型表 现出优越的预测性能. 文献[12-13]表明, 综合考虑局 部特征和全局特征能使网络表现优越性能,因此本文 在肌电信号分类中,同时将肌电信号中的全局特征和 局部特征进行交互融合.

为保证网络同时关注sEMG中非线性的长期依赖 动态关系和局部细节信息,使提取到的肌电特征能够 最大程度地表征肢体运动的动作模式,本文提出基于 全局与局部肌电特征交互的手势识别网络(global and local EMG classification network, GL-EMG-Net),用轻 量级的多头注意力机制和大感受野的膨胀卷积替代 循环卷积网络或长短期记忆网络(long short-term memory, LSTM)来提取肌电信号中长期依赖的序列信 息,减少网络的训练时间,使网络更快的收敛;再通过 不同大小感受野的卷积提取sEMG中不同尺度的细节 信息,然后利用通道注意力机制对提取到的局部特征 进行特征重分配,在增大肌电固有特征权重的同时, 减小纹波信号和小噪声的权重,使网络同时关注到肌 电序列中的整体运动趋势和局部细节信息,提高手势 识别准确率,增加手势分类类别数.

本文的结构安排如下:第2节介绍了肌电数据预处 理方法和数据增强方法;第3节介绍了GL-EMG-Net 手势识别网络的设计细节;第4节展示了本文使用的 肌电数据集、实验方法和实验结果;第5节总结全文, 并对肌电分类方法做出展望.

2 数据处理

2.1 表面肌电信号预处理

sEMG是一种非平稳的微电信号^[13],采集时会因为传感器内部电子元器件工作产生的纹波,肌肉活动时人体皮肤表面与电极接触过程中的相对运动和50 Hz工频信号的频率干扰,而携带很多噪声,这些噪声掺杂在肌电固有特征中,会影响手势分类精度,因此需要对这部分噪声进行滤波.另一方面,通过深度学习的方法进行手势识别需要大量的肌电样本^[4],本文通过滑动窗口的方法增加肌电数据样本.

如图 1 所示, sEMG预处理过程分为滤波、数据分割和标签标记3个步骤. 原始肌电信号X通过 20 Hz的高通滤波器 f_1 和50 Hz的陷波滤波器 f_2 后的输出信号 $Y = f_2(f_1(X));$ 输出信号 Y 经过窗口大小为W、步 长为L的滑动窗口提取后,得到一系列长度为W的肌 电序列{ Y_1, Y_2, \dots, Y_n },其中n为分割后的样本数; 最后标记分割出的肌电序列,如握拳手势的标签为0, 张掌手势的标签为1,以此类推,手腕外翻手势的标签 为N,得到肌电信号的训练样本.



Fig. 1 Flow diagram of surface EMG signal preprocessing

2.2 肌电数据增强

肌电信号的采集是一项巨大工程^[14],本文通过在 原始sEMG中加入高斯白噪声模拟环境噪声来得到新 的sEMG信号.假设原始肌电信号为X,均值为 μ ,方 差为 $\sigma(x)$,则高斯白噪声信号的方差np为

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

其中: SNR为随机信噪比 (signal to noise ratio, SNR), SNR \in (25, 35); 高斯白噪声信号的均值为0. 通过加 入不同信噪比的高斯白噪声可模拟生成具有不同环 境噪声的肌电信号, 从而大量扩充肌电数据集.

3 网络结构

3.1 GL-EMG-Net手势识别网络框架

sEMG中蕴藏的固有特征在电路噪声、生理噪声 和环境噪声的影响下很难被有效提取[4],为了获取更 多固有肌电特征,减少噪声干扰,本文同时关注肌电 信号的全局信息和不同尺度的细节信息,利用注意力 机制重分配特征图权重,使网络更加关注肌电信号的 固有特征,减少噪声和干扰信号的权重分配;同时,为 了减少模型的训练时间和防止过拟合,采用轻量级的 多头注意力机制 Multi-Head Attention 和大感受野的 膨胀卷积替代循环卷积网络或 LSTM 提取肌电信号 中长期依赖的序列信息,减少模型训练时间;最后,通 过自适应全局平均池化层替代全连接层,构建本文模 型的特征分类层进行手势类别输出,防止模型过拟 合^[15]. GL-EMG-Net 网络如图2所示. 图中, C_{in}为输 入肌电信号的维度, W为窗口大小, n为全局特征提取 模块(global-dilation transformer, Global-DT)的堆叠 次数, N为手势类别数. GL-EMG-Net由特征提取和 特征分类两部分构成. 特征提取部分由全局特征提取

模块 Global-DT 和局部特征提取模块 (local-selective kernel, Local-SK)组成, Global-DT利用多头注意力机 制的长距离捕捉能力关注肌电序列中长距离的动态 关系,同时通过N层空洞卷积 Dilated Conv 使第N层 空洞卷积的等效感受野完全覆盖肌电系列的长度,以 关注肌电序列的整体运动趋势,最后将肌电序列中的 长距离动态关系和整体运动趋势进行融合,得到肌电 序列中的全局特征; Local-SK利用3个不同感受野的 可分离卷积提取肌电序列中不同尺度的细节信息,充 分关注肌电序列中的小噪声信号,然后通过通道注意 力机制重分配特征权重,使网络更加关注肌电信号的 固有特征,最后将得到的局部特征图与全局特征图进 行融合.为了使全局特征和局部特征充分交互融合, 通过反馈机制将局部特征图与输入的肌电信号融合 后,再通过全局特征提取模块提取肌电信号中融合局 部特征的全局特征;特征分类部分通过自适应全局平 均池化层替代全连接层构建分类层进行手势类别输 出,能够使模型不易过拟合,且使网络输入的肌电序 列的窗口大小不受限制[15].



Fig. 2 GL-EMG-Net gesture recognition network frame diagram

3.2 全局特征提取模块

Transformer 的多头注意力机制能够弥补CNN不能捕捉长期依赖的序列特征的不足^[16-18],而空洞卷积可在不大量增加网络参数的情况下增加网络的感受野^[8,12],因此,本文利用多头注意力机制和空洞卷积设计全局特征提取模块Global-DT.此模块的第1条支路通过一维卷积Conv1d添加肌电序列的非线性特征的同时将肌电序列的通道数扩展至32维,再进行权重

归一化(Weight Norm)加快网络收敛,最后通过多头注 意力机制充分关注非线性肌电特征图中的长期动态 依赖关系;第2条支路为n层空洞卷积,为了保证空洞 卷积能够有效提取肌电序列曲线的整体运动趋势,使 第n层空洞卷积的等效感受野完全覆盖输入肌电序列, 之后经过权重归一化,再通过修正线性单元(rectified linear unit, ReLU)激活函数和Dropout防止模型过拟合. 第1条路径的多头自注意力机制如图3所示,假设 输入信号为 $X = [x_1 \ x_2 \ \cdots \ x_n]$,其中n表示输入信息的长度.首先,通过线性变换得到3个矩阵

$$\begin{cases} \boldsymbol{Q} = W_q \boldsymbol{X}, \\ \boldsymbol{K} = W_k \boldsymbol{X}, \\ \boldsymbol{V} = W_v \boldsymbol{X}, \end{cases}$$
(2)

其中 W_q , W_k , W_v 分别为3个矩阵的权重. 然后利用点积和Concat拼接得到多头注意力的输出为

Attention_i(
$$\boldsymbol{Q}, \boldsymbol{K}, \boldsymbol{V}$$
) =
Softmax $\left(\frac{\boldsymbol{Q}W_i^Q \cdot \left(\boldsymbol{K}W_i^K\right)^{\mathrm{T}}}{\sqrt{d_k}}\right) \cdot \boldsymbol{V}W_i^V$, (3)

MultiHeadAttention(Q, K, V) =

$$\operatorname{Concat}_{i}^{h} \{\operatorname{Attention}_{i}(\boldsymbol{Q}, \boldsymbol{K}, \boldsymbol{V})\},$$
(4)

式(4)中: h为输入序列的组数, Concat为通道维度拼接函数.最后将多头注意力的输出通过线性层Linear,保证输出与输入维度相等.这种方法能够保证多头注意力机制关注到整个输入序列的整体特征.



图 3 多头自注意力模块结构图

Fig. 3 Multi-head self-attention module structure diagram

第2条路径主要由n层空洞卷积组成,其网络结构 如图4所示.





图4中,空洞卷积通过膨胀因子Dilation和层数的 叠加使第n层卷积操作的等效感受野膨胀变大,为了 使第n层空洞卷积网络的等效感受野完全覆盖整个输 入肌电序列,以获取到肌电序列的全局信息,需要保证第*n*层空洞卷积的等效感受野不小于肌电序列的输入长度^[12].空洞卷积感受野*F*_n的计算公式如下:

$$F_n = F_{n-1} + \left((f_n - 1) \times \prod_{j=1}^{n-1} s_j \right), \ n \ge 2,$$
 (5)

$$f_n = k + (d - 1) \times (k - 1),$$
 (6)

$$d = b^{n-1},\tag{7}$$

其中: F_n 表示第n层感受野大小, F_{n-1} 表示第n - 1层 感受野大小, f_n 表示第n层的空洞卷积的等效卷积核 大小, S_j 表示第j层卷积操作的步长(本文S = 1), k为 卷积核大小(本文k = 3), d为膨胀因子, b为空洞卷积 膨胀因子的扩展基(本文b = 2).

以下为空洞卷积最小网络层数*n*的计算过程:由于*S* = 1,所以式(5)可简化为

$$F_n = F_{n-1} + f_n - 1, \ n \ge 2,$$
 (8)

联系式(6)-(8)可得

$$F_n - F_{n-1} = (k-1)b^{n-1}, \ n \ge 2,$$
 (9)

式(9)为空洞卷积感受野随层数n变化的递推公式,通过计算可得空洞卷积感受野F_n的通项公式为

$$F_n = (k-1)\frac{b^n - 1}{b-1} + 1, \ n \ge 2, \tag{10}$$

最后根据第n层空洞卷积的感受野不小于输入的肌电 序列长度,可求解所需的最小层数n为

$$n = \lceil \log_b(\frac{(W-1)(b-1)}{k-1} + 1) \rceil, \qquad (11)$$

其中: [·]表示向上取整, W为输入肌电序列的长度.

全局特征提取模块,通过多头注意力的长距离捕捉能力和大感受野的空洞卷积替代循环卷积网络或 LSTM捕捉肌电序列的长期动态依赖关系和整体运动 趋势,保证网络充分提取肌电序列的全局信息,减少 模型训练时间,使模型更快地收敛^[19].

3.3 局部特征提取模块

原始sEMG中掺杂着很多不同尺度的噪声,利用 深度学习的方法,通过注意力机制给肌电固有特征分 配更大权重,同时抑制噪声特征的权重可减少噪声影 响.基于此,设计Local-SK模块,该模块首先通过3个 卷积核大小分别为3,5和7的深度可分离卷积提取肌 电序列中不同尺度的细节信息^[20-21];然后通过权重归 一化保证 Local-SK 模块提取的局部特征,在后期与 Global-DT提取的全局特征在融合时保持特征尺度一 致,以防止大特征同化小特征;最后通过Relu激活函 数和Dropout层防止网络过拟合.

图5为本文使用的三分支通道注意力SK模块结构 图,该模块首先通过多流卷积和不同大小感受野的卷 积操作提取不同尺度细节信息的特征

$$U_1 = \text{Weight} - \text{Norm}(\text{Conv}3(X)),$$
 (12)

$$\boldsymbol{U}_2 = \text{Weight} - \text{Norm}(\text{Conv5}(\boldsymbol{X})), \quad (13)$$

$$U_3 = \text{Weight} - \text{Norm}(\text{Conv7}(X)),$$
 (14)

其中Conv3表示卷积核大小为3的深度可分离的卷积 操作. 然后通过 element-wise summation 融合获得特 征图 $U = U_1 + U_2 + U_3$, 再经过全局平均池化层 F_{gp} 得到特征图S, 经过全连接层 F_{fc} 得到压缩后的特征Z, 即

$$\boldsymbol{S}_{W} = F_{\rm gp}(\boldsymbol{U}) = \frac{1}{C} \sum_{i=1}^{C} \boldsymbol{U}(i), \qquad (15)$$

其中:
$$S_W$$
为仅包含肌电序列长度元素的列向量,下标 W 为肌电序列长度; W_S 为全局特征 S 的权重矩阵.最后经过Softmax函数得到对应流的注意力权重

 $\boldsymbol{Z} = F_{\rm fc}(\boldsymbol{S}_W) = \delta({\rm BN}(\boldsymbol{W}_S)),$

$$a = \frac{\mathrm{e}^{A_{\mathrm{W}}\boldsymbol{Z}}}{\mathrm{e}^{A_{\mathrm{W}}\boldsymbol{Z}} + \mathrm{e}^{B_{\mathrm{W}}\boldsymbol{Z}} + \mathrm{e}^{C_{\mathrm{W}}\boldsymbol{Z}}},\tag{17}$$

$$b = \frac{\mathrm{e}^{B_{\mathrm{W}}\boldsymbol{Z}}}{\mathrm{e}^{A_{\mathrm{W}}\boldsymbol{Z}} + \mathrm{e}^{B_{\mathrm{W}}\boldsymbol{Z}} + \mathrm{e}^{C_{\mathrm{W}}\boldsymbol{Z}}},$$
(18)

$$= \frac{\mathrm{e}^{\mathrm{e}^{\mathrm{A}}}}{\mathrm{e}^{\mathrm{A}}\mathrm{w}^{\mathbf{Z}}} + \mathrm{e}^{\mathrm{B}}\mathrm{w}^{\mathbf{Z}}} + \mathrm{e}^{\mathrm{C}}\mathrm{w}^{\mathbf{Z}}},$$
(19)

其中A_W, B_W和C_W是网络训练得到的特征矩阵.

c



Fig. 5 Schematic diagram of the three-branch SKConv module

在得到各卷积流的注意力权重后,将各流的特征 图**U**₁,**U**₂和**U**₃点乘对应流的权重后,再相加融合,得 到最终的特征映射输出V,即

$$\boldsymbol{V} = \boldsymbol{a} \cdot \boldsymbol{U}_1 + \boldsymbol{b} \cdot \boldsymbol{U}_2 + \boldsymbol{c} \cdot \boldsymbol{U}_3, \quad (20)$$

其中a, b, c的和为1.

局部特征提取模块引入不同感受野的小卷积操作 和注意力机制,保证提取到肌电信号不同尺度的细节 信息的同时,通过注意力机制减少这些细节信息中噪 声的权重,使提取到的细节信息包含更多的肌电信号 固有特征.除此之外,在小卷积操作后进行了权重归 一化,保证提取的细节信息在和全局信息进行融合时, 处于同一个数量级,防止大特征同化小特征,并且也 加快了网络的收敛速度.

4 实验与结果分析

4.1 实验设置

为了验证GL-EMG-Net模型的有效性,在Ninapro DB5数据集上进行评估.Ninapro数据集^[22]是目前最 具权威性且被广泛使用的肌电数据集,共包含10个受 试者的52种手部动作,传感器采用两个采样频率为 200 Hz的8通道Myo肌电臂环.

如图6所示, DB5采集的手势类别包括12种指部常见动作、17种手指等距等张手势或腕部运动手势和23种抓取和功能性动作. 在采集过程中, 要求每位测试者的每个手势动作重复6次, 每重复一次动作持续5 s 然后休息3 s, 以防止肌肉疲劳, 假设以300 ms的提取窗提取肌电数据样本, 则每个人的每个动作可提取100个样本, 10个受试者的52种手势共可提取约52000个数据样本.

基于Ninapro DB5数据集和实际的Myo臂环采集的数据,分别设置以下4组实验:

1) 消融实验: 基于Ninapro DB5数据集, 在300 ms 滑动窗口下进行消融实验, 验证Global-DT, Local-SK 和反馈机制的有效性;

2) 过 拟 合 情 况 和 收 敛 速 度 测 试: 基于Ninapro DB5数据集,在300 ms滑动窗口下,将GL-EMG-Net与 LSTM-CNN^[23]和DenseNet^[24]进行对比,比较其损失 曲线和准确率曲线,分析网络的收敛速度和过拟合情 况;

3) 不同滑动窗口下的精度测试: 基于 Ninapro DB5数据集, 在100 ms, 200 ms和300 ms滑动窗口(步 长为50 ms, 100 ms和100 ms)下, 将GL-EMG-Net与最

新的手势识别网络、LSTM-CNN和DenseNet对比,验 证本文模型的先进性;

4) 实际应用实验: 基于实际的Myo臂环, 采集人体 手臂的12种手势动作真实肌电数据, 在300 ms滑动窗 口下, 通过将GL-EMG-Net模型的识别精度与LSTM-CNN和DenseNet模型进行对比, 验证该模型在实际实 验的分类性能.



(c) 23种抓取和功能性动作
 图 6 Ninapro DB5的52种手势图
 Fig. 6 52 gesture graphs of the Ninapro DB5

验证平台:软件版本为Ubuntu16.03操作系统, Pytorch1.11.0深度学习框架;硬件配置为Inter Xeon (R) E5-2678 v3处理器,GeForce RTX 3080显卡和 32 GB显存.sEMG数据集的划分比例为:训练集70%, 测试集20%,交叉验证集10%;优化器为SGD,学习率 设置为0.001,损失函数选择交叉熵函数,batch-size设 置为256,训练轮数设置为75轮.评价指标采用验证集 的平均准确率.

4.2 实验结果分析

4.2.1 消融实验

为验证本文提出的手势识别网络各模块的有效性, 设计表1所示的5个实验,其中"√"表示使用了该方法 或模块,"—"表示未使用.实验1为三层全卷积网络 (fully convolutional networks, FCN)的基准网络,实验 2,3和4分别为引入Global-DT模块、Local-SK模块和 同时引入两种模块的网络,实验5为具有局部特征反 馈机制的网络.

表1中,Baseline为三层FCN网络,在Baseline基础 上加入全局注意力模块Global-DT后,手势识别准确 率由85.63%提升至87.87%;在Baseline网络上加入局 部注意力模块Local-SK后,手势识别准确率由85.63% 提升至89.02%;在Baseline网络上同时加上Global-DT 和Local-SK模块后,手势识别准确率在Baseline和仅 加 Global-DT 或仅加 Local-SK 的基础上分别提升了 4.588%, 2.351%和1.194%;继续加入局部特征的反馈 机制后,提升了0.466%.由实验结果可知, Global-DT 和Local-SK模块在共同使用时,比单独使用时的准确 率都要高,充分证明了Global-DT, Local-SK模块和反 馈机制的有效性.

表1 消融实验

Table 1 Ablation experiment

实验	Baseline	Global-DT	Local-SK	反馈机制	准确率/%
1	\checkmark		_		85.63
2	\checkmark	\checkmark	_	_	87.87
3	\checkmark	—	\checkmark	_	89.02
4	\checkmark	\checkmark	\checkmark		90.22
5	\checkmark	\checkmark	\checkmark	\checkmark	90.68

4.2.2 过拟合情况和收敛速度测试结果与分析

图7为GL-EMG-Net, LSTM-CNN和DenseNet损失曲线和准确率曲线.





Fig. 7 Comparison chart of three types of network training set loss and test set accuracy

图7中LSTM-CNN结构与文献[23]的模型相似, 本文的LSTM的隐藏层数为3,隐藏层节点数为128; DenseNet的结构与文献[24]的模型相似,本文的Dense Block为3个;图中的蓝色实线表示本文网络的训练 集损失,蓝色虚线表示本文网络的交叉验证集损失; 红色实线表示LSTM-CNN训练集损失,红色虚线表示 LSTM-CNN交叉验证集损失;绿色实线表示Dense-Net训练集损失,绿色虚线表示Dense-Net交叉验证集 的损失.

由图7(a)可知, GL-EMG-Net相比LSTM-CNN和 DenseNet在训练集和交叉验证集上的Loss都为最小, GL-EMG-Net网络未出现过拟合;由图7(b)可知, GL-EMG-Net经过一轮训练后准确率达到75.49%,在第26 轮训练后准确率趋于稳定,最终准确率达到90.68%; LSTM-CNN模型经过一轮训练后准确率为65.92%, 经过45轮训练后准确率稳定为84.14%; Densenet模型 第1轮训练后准确率为70.92%,在训练30轮左右准确 率稳定在87.48%.

为了进一步比较 GL-EMG-Net 模型与 DenseNet, LSTM-CNN 的收敛速度,设定预期分类精度为84%, 在300 ms肌电提取窗下,统计各模型交叉验证集连 续3次到达预期精度的收敛轮数和训练时间,各模型 分别做5次实验,取收敛轮数和训练时间的平均值,实 验统计结果如表2所示.

表 2 3种模型的5次测试的平均收敛速度对比

 Table 2 Comparison of the average convergence speeds

 of the five tests of the three models

	GL-EMG-Net	LSTM-CNN	DenseNet
训练一轮平 均时间/s	291.523	293.775	367.812
达到指定精度 平均训练轮数	4	45	12
平均总训 练时间/s	1 166.092	13 219.875	4413.774

由表2可知, GL-EMG-Net, LSTM-CNN和 Dense Net的准确率达到84%时,分别平均需要 1166.092 s, 13219.875 s和4413.774 s的训练时间, GL-EMG-Net的训练时间最少, 收敛速度最快.

综上可知, GL-EMG-Net的Loss收敛最快, 且最终的Loss最小, 同时在交叉验证集上第1轮训练后准确率达到75.49%, 最终的准确率为90.68%. 说明GL-EMG-Net网络在加快网络收敛的同时, 也增加了手势识别的准确率, 更不易发生过拟合.

4.2.3 不同滑动窗口下的精度测试结果与分析

表3为在100 ms, 200 ms和300 ms滑动窗口提取窗下GL-EMG-Net的手势识别结果与最新的手势识别网络^[8,21-22,25]、LSTM-CNN和DenseNet进行不同滑动窗口识别精度的对比结果,为了说明GL-EMG-Net能够识别更多的手势类别数,统计了各种方法对应的手

势分类类别数. 表3中"—"表示该方法对应文献中未 做这种大小滑动窗的实验.

表 3 本文网络与其他方法的分类准确率对比

 Table 3 The classification accuracy of this network and other methods is compared

	工业来回来	准确率/%		
力法	于努尖别级	100 ms	200 ms	300 ms
ShenNet ^[21]	40	_	72.09	76.02
PanyawutNet ^[25]	41	81.37	84.25	
XuNet ^[8]	51			87.42
EsaaNet ^[22]	17		86.50	
LSTM-CNN	52	76.99	80.50	84.14
DenseNet	52	80.72	84.13	87.48
GL-EMG-Net	52	83.55	87.70	90.68

表3中, 文献[21]基于一种CNN和堆叠集成学习的 网络, 对Ninapro DB5数据集的40种手势进行分类, 当 滑动窗口为200 ms时, 分类准确率为72.09%, 当滑动 窗口为300 ms时, 分类准确率为76.02%; 文献[25]通 过三层全连接层和采用Dropout层搭建基于深度学习 的手势分类网络对41种手势进行分类, 当滑动窗口 为100 ms时, 分类准确率为81.37%, 当滑动窗口为 200 ms时, 分类准确率为81.37%, 当滑动窗口为 200 ms时, 分类准确率为84.25%; 文献[8]提出一种新 的注意力卷积网络SE-CNN对51种手势进行分类, 在 滑动窗口为300 ms时, 分类准确率为87.42%; 文献 [22]将长期的肌电信号分割为多个短期信号, 从短期 信号中提取5种时域特征, 利用提取的特征在 Nina-Pro DB5 数据集上对17种手势进行分类, 在滑动窗口 为200 ms 时, 分类准确率为86.5%.

综上,当滑动窗口为100 ms时,GL-EMG-Net相比 PanyawutNet,LSTM-CNN和DenseNet,其准确率分 別提高2.18%,6.56%和2.83%;当滑动窗口大小设为 200 ms时,相比于ShenNet,PanyawutNet,EsaaNet, LSTM-CNN和DenseNet,其准确率分别提高15.61%, 3.45%,1.2%,7.2%和3.57%;当滑动窗口为300 ms时, 相比ShenNet,XuNet,LSTM-CNN和DenseNet,其准 确率分别提高14.66%,3.26%,6.54%和3.2%.另外,对 于手势类别数的比较,GL-EMG-Net模型的手势类别 数最高,为52种手势,其识别准确率仍然最高,可见该 网络更适合多手势的识别.因此,GL-EMG-Net网络 优于目前大部分手势识别网络.

4.2.4 实际应用实验结果与分析

为了进一步验证本文提出的手势分类网络GL-EMG-Net的可靠性,基于Myo臂环的手势识别系统在 线验证12种手势分类效果,手势识别系统如图8所示.

图8中Myo肌电臂环采集人体小臂实时sEMG,通 过蓝牙传输至PC端; PC端接收到信号后,经过预处理 送到离线训练好的网络模型中分类;最后将分类结果转换为五指手的控制指令,实时控制五指手运动.用 Myo臂环离线采集3位年龄24±3的男性志愿者的右 小臂的指部、腕部和抓握等12种动作的sEMG数据, 300 ms 滑动窗预处理后,分别用LSTM-CNN 模型、 DenseNet 模型和本文提出的GL-EMG-Net 模型离线 训练手势分类模型.其中,12种手势动作如图9所示.



图 8 基于Myo臂环的手势识别系统

Fig. 8 Gesture recognition system based on Myo armband





图10(a), (b)和(c)分别为LSTM-CNN, DenseNet和 GL-EMG-Net的测试集结果, 其中每行表示预测手势, 每列表示真实手势, 斜对角为各类手势的正确识别结 果, 颜色越深表示识别精度越高.

由图10可知,3种模型的平均准确率分别为 89.67%,92.33%和97.83%,GL-EMG-Net模型的混淆 矩阵中除了捏卡片动作的准确率为85%,其他手势识 别率都大于95%,12种手势的平均准确率为97.83%. 对于GL-EMG-Net模型,由于捏卡片动作、捏食指动 作和捏中指动作的发力点相似,所以16%的捏卡片动 作被识别成了捏食指动作和捏中指动作,从而造成了 捏卡片动作的识别准确率为85%.但是针对捏无名 指、捏食指、捏中指等精细的指部动作,其识别准确率 仍可以保持在95%以上.



el experiments

5 结论

本文提出的网络模型GL-EMG-Net通过全局特征 提取模块Global-DT和局部特征提取模块Local-SK的 相互配合,提取肌电信号中的全局特征和局部特征, 分别在实验环境和实际手势识别应用中验证了GL-EMG-Net网络的性能,在肌电特征提取过程中减少了 肌电信号的损失,提高了手势识别准确率,增加了手 势识别类别数.

本文旨在优化网络对肌电信号动作特征的提取能力,虽然在实验环境和实际环境下该网络都能表现出 很好的分类性能,但本文的网络模型还存在一些不足: 1)该网络主要提取肌电信号的时域特征,进一步可提 取肌电信号中的频域信息和时频域信息,通过时域和 频域信息的融合进一步提高手势分类的精度和抗干 扰能力;2)该网络在全局特征和局部调整的融合上采 用了直接相加的方式,后期可研究注意力特征融合机 制来进行全局-局部特征的融合,以进一步的提高手 势分类精度.

参考文献:

- OZDEMIR M A, KISA D H, GUREN O, et al. EMG based hand gesture recognition using deep learning. *Medical Technologies Congress* (*TIPTEKNO*). Antalya, Turkey: IEEE, 2020: 1 – 4.
- [2] TAM S, BOUKADOUM M, CAMPEAU-LECOURS A, et al. Intuitive real-time control strategy for high-density myoelectric hand prosthesis using deep and transfer learning. *Scientific Reports*, 2021, 11(1): 11275.
- [3] YU Y, LI J, SOLOMON S A, et al. All-printed soft human-machine interface for robotic physicochemical sensing. *Science Robotics*, 2022, 7(67): eabn0495.
- [4] LI Ziyou, ZHAO Xingang, ZHANG Bi, et al. Review of sEMG-based motion intent recognition methods in non-ideal conditions. *Acta Automatica Sinica*, 2021, 47(5): 955 969.
 (李自由, 赵新刚, 张弼, 等. 基于表面肌电的意图识别方法在非理想 条件下的研究进展. 自动化学报, 2021, 47(5): 955 969.)
- [5] NASRI S, ORTS-ESCOLANO S, GOMEZ-DONOSO F, et al. Inferring static hand poses from a low-cost non-intrusive sEMG sensor. *Sensors*, 2019, 19(2): 371.
- [6] XU P, LI F, WANG H. A novel concatenate feature fusion RCN-N architecture for sEMG-based hand gesture recognition. *PloS One*, 2022, 17(1): e0262810.
- [7] CHEN L, FU J, WU Y, et al. Hand gesture recognition using compact CNN via surface electromyography signals. *Sensors*, 2020, 20(3): 672.
- [8] XU Z, YU J, XIANG W, et al. A novel SE-CNN attention architecture for sEMG-based hand gesture recognition. *CMES-Computer Modeling in Engineering & Sciences*, 2023, 134(1): 157 – 177.
- [9] SHEN S, GU K, CHEN X R, et al. Movements classification of multichannel sEMG based on CNN and stacking ensemble learning. *IEEE Access*, 2019, 7: 137489 – 137500.
- [10] BAI D, LIU T, HAN X, et al. Application research on optimization algorithm of sEMG gesture recognition based on light CNN+LSTM

model. Cyborg and Bionic Systems, 2021, DOI: 10.34133/2021/979 4610.

- [11] LIU X, WANG J, HAN T, et al. Real-time control of intelligent prosthetic hand based on the improved TCN. *Applied Bionics and Biomechanics*, 2022, 1: 6488599.
- [12] YU L. DTAAD: Dual tcn-attention networks for anomaly detection in multivariate time series data. ArXiv Prints, 2024, arXiv: 2302.10753.
- [13] REN Y, ZHAO D, LUO D, et al. Global-local temporal convolutional network for traffic flow prediction. *IEEE Transactions on Intelligent Transportation Systems*, 2020, 23(2): 1578 – 1584.
- [14] CHEN C, GUO W, MA C, et al. sEMG-based continuous estimation of finger kinematics via large-scale temporal convolutional network. *Applied Sciences*, 2021, 11(10): 4678.
- [15] RAHIMIAN E, ZABIHI S, FAROKH A S, et al. XceptionTime: A novel deep architecture based on depthwise separable convolutions for hand gesture classification. *ArXiv Prints*, 2019, arXiv: 1911.038 03.
- [16] MONTAZERIN M, ZABIHI S, RAHIMIAN E, et al. ViT-HGR: Vision transformer-based hand gesture recognition from high density surface EMG signals. *The 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society*. IEEE, 2022: 5115 – 5119.
- [17] SHEN S, WANG X, MAO F, et al. Movements classification through sEMG with convolutional vision transformer and stacking ensemble learning. *IEEE Sensors Journal*, 2022, 22(13): 13318 – 13325.
- [18] SALIMANS T, KINGMA D P. Weight normalization: A simple reparameterization to accelerate training of deep neural networks. *ArXiv Preprint*, 2016, arXiv: 1602.07868.
- [19] LI X, WANG W, HU X, et al. Selective kernel networks. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, CA, USA: IEEE, 2019: 510 – 519.
- [20] HAN K, WANG Y, CHEN H, et al. A survey on vision transformer. *IEEE Transactions on Pattern Analysis & Machine Intelli*gence, 2023, 45(1): 87 – 110.
- [21] SHEN S, GU K, CHEN X R, et al. Gesture recognition through sEMG with wearable device based on deep learning. *Mobile Networks and Applications*, 2020, 25(6): 2447 – 2458.
- [22] ESAA R R, JABER H A, AMEER A A. Hand movements classification based on Myo armband signals. *The 4th International Conference on Electrical, Control and Instrumentation Engineering.* KualaLumpur, Malaysia: IEEE, 2022: 1 – 5.
- [23] WU Y, ZHENG B, ZHAO Y. Dynamic gesture recognition based on LSTM-CNN. 2018 Chinese Automation Congress. Xi'an, China: IEEE, 2018: 2446 – 2450.
- [24] ZHONG T, LI D, WANG J, et al. Fusion learning for sEMG recognition of multiple upper-limb rehabilitation movements. *Sensors*, 2021, 21(16): 5385.
- [25] SRI-IESRNUSOM P, CHAIYAROJ A, BUEKBAN C, et al. Classification of 41 hand and wrist movements via surface electromyogram using deep neural network. *Frontiers in Bioengineering and Biotechnology*, 2021, 9: 548357.

作者简介:

肖城钢 硕士研究生,主要研究方向为模式识别和基于表面肌电 信号的人机交互, E-mail: 980061174@qq.com;

闵华松 博士, 教授, 博士生导师, 主要研究方向为智能机器人与 嵌入式系统, E-mail: mhuasong@wust.edu.cn.

617