

融合小波门控时序卷积网络的肌电手势识别研究^{*}

孙 鹏¹ 宋 杰¹ 张 克² 卫文韬¹ 李忠新¹

(1. 南京理工大学机械工程学院 南京 210094; 2. 中国兵器装备第208研究所 北京 102202)

摘 要:在肌电信号手势识别领域,表面肌电信号固有的信号变异性是导致当前深度学习模型特征提取能力不足,识别准确率降低的主要原因。针对这一问题,结合时序卷积神经网络的长期上下文建模优势与小波变换对非平稳信号的时频分析特性,提出一种新型小波门控时序卷积网络模型。首先通过小波卷积模块将输入的肌电信号进行多级离散小波分解,并将得到的各级分量分别进行一维卷积,小波域的卷积操作在实现多尺度提取时频特征的同时又自适应地增强关键的时频模式并抑制冗余分量,接着将卷积后的细节系数与近似系数进行离散逆小波变换以重构信号,最后将重构信号输入融合门控单元的时序卷积网络,利用时序卷积神经网络捕捉 sEMG 信号中的长期依赖关系,并使用门控单元对所提取的特征进行过滤。该网络结构在 Ninapro DB1 数据集上对 52 类手势分类实现了 81.85% 的准确率,相比于传统时序卷积网络,准确率提高了 4.9%,与近年本领域主流深度模型(如 DuNet, GengNet 等)相比,该方法在保持模型参数量较小前提下,准确率相对提升幅度达 4.0%~7.8%。

关键词:肌电信号;手势识别;离散小波分解;时序卷积神经网络;门控单元

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

Electromyographic signal gesture recognition method based on wavelet gated temporal convolution

Sun Peng¹ Song Jie¹ Zhang Ke² Wei Wentao¹ Li Zhongxin¹

(1. School of Mechanical Engineering, Nanjing University of Science and Technology, Nanjing 210094, China;

2. The 208th Research Institute of China Ordnance Equipment Group, Beijing 102202, China)

Abstract: This paper constructs a wavelet-gated temporal convolutional network model. First, the input electromyographic signals are subjected to multi-level discrete wavelet decomposition through a wavelet convolution module, and the components of each level are respectively subjected to one-dimensional convolution. Then, the detailed coefficients and approximate coefficients after convolution are reconstructed via discrete inverse wavelet transform. This process of multi-level decomposition, convolution, and step-by-step reconstruction enables the model to adaptively focus on key time-frequency features. The reconstructed signals are then input into a temporal convolutional network integrated with a gating unit. The proposed network structure achieves an accuracy of 81.85% for 52-class gesture classification on the Ninapro DB1 dataset, which is 4.9% higher than that of traditional temporal convolutional networks. Compared with recent mainstream deep models in this field (such as MSHilbNet, GengNet, etc.), this method achieves a relative accuracy improvement of 4.0%~7.8% while maintaining a smaller number of model parameters.

Keywords: EMG signal; Gesture recognition; DWT; TCN; GLU

0 引言

基于肌电信号(surface electromyography, SEMG)的手势识别是人机交互、智能假肢控制和康复医疗等领域的

重要研究方向^[1]。然而在肌电控制领域存在若干挑战,其中 SEMG 信号的变异性导致模型特征提取能力不足,识别能力受限是最具挑战性的问题^[2]。SEMG 信号变异性源于肌肉的时变状态与运动学特性,以及不同用户神经控制机制的差异^[3-4]。此外,电极放置位置的差异以及

电子元件的固有噪声也会导致信号变异性^[5]。

为应对这一挑战,在肌电信号手势识别领域,深度学习 方法得到广泛研究,卷积神经网络(convolutional neural network, CNN)是被广泛应用的基础方法之一, Manfredo 等^[6]首次将卷积神经网络应用于肌电信号手势识别,利用一个简单的 4 层网络结构对 NinaproDB1 数据集的 53 类手势实现了 66.59%±6.40%的准确率,与基于传统机器学习(machine learning, ML)方法取得的效果相当,但未超过 ML 最优识别率。随后 Tsiganos 等^[7]在此基础上加入 dropout 层并将平均池化替换为最大池化,实现了 70.48%的准确率,超过了 ML 方法的最高准确率。Li 等^[8]开发了基于 CNN 的端到端深度学习框架用于手势分类,其准确率显著高于传统机器学习技术。Javier 等^[9]将 SEMG 信号的功率密度图作为 CNN 输入,成功识别了 6 种手势。Geng 等^[10]则使用高密度 SEMG 图像作为 CNN 输入,对 8 种手势的识别率达到 89.30%。然而,此类基于 CNN 的模型其核心问题在于未能有效捕捉 sEMG 信号中的时序依赖关系,导致识别精度难以提升;此外,复杂的网络结构也带来了高昂的计算成本。在此背景下,时序预测模型中的时序卷积神经网络(temporal convolutional network, TCN)由于具有捕获上下文时序特征以及网络参数量较少的特点在本领域得到关注^[11]。Tsiganos 等^[12]首次将 TCN 模型应用到肌电信号手势所 识别领域,在 ninapro 数据集上取得了 76.95%的准确率。Betthausen 等^[13]证实,TCN 能够有效利用 sEMG 信号中的 时序信息,从而实现卓越的预测性能^[14]。因此,将肌电 信号解码视为序列建模问题,可提升假肢控制系统的可 靠性、响应速度和运动复杂度^[15]。Ma 等^[16]将 TCN 与长 短期记忆网络(long short-term memory, LSTM)融合,进一 步挖掘信号中的时序信息,在 8 种手势中取得了 98.89% 的准确率。Lin 等^[17]提出大规模时间卷积网络(LS- TCN),发现通过增加卷积核尺寸使之接近 SEMG“原子 段”长度时,网络预测精度最佳。以上研究均表明了 TCN 在肌电信号时序特征建模上的优势——可利用肌电信号 的历史信息捕捉时序依赖关系。然而仅依赖 TCN 中的 卷积操作会忽略信号中频域的分析,这会导致特征数据 丢失^[18]。

受以上工作启发,本文首次将用于时频域分析的小 波变换与 TCN 结合起来,通过引入小波卷积模块对原始 肌电信号进行多尺度分解、卷积与重构,这一过程使模型 能从不同尺度捕获时频域特征,并通过可训练卷积核在 不同尺度系数上滑动使模型学习特定尺度下信号的局部 模式,聚焦并增强关键特征,同时抑制冗余分量。重构后 的信号虽与原始信号保持一定差异,但通过与基线模型 进行实验对比,这种差异本质是模型在端到端训练时特 征优化的过程。在时序卷积模型的基础上加入门控单

元,对每层卷积层提取到的特征进行动态筛选,从而优化 信息流动路径。实验结果表明,这种多尺度特征提取与 动态特征选择相结合的设计显著提升了模型的分类 性能。

1 数据预处理

1.1 实验数据集

本研究将采用公认的科学基准数据库 Ninapro^[19-21] 对所提架构进行评估,该数据库是目前稀疏多通道 SEMG 信号模型开发领域接受度最广的基准测试平台。 该实验选用 Ninapro DB1 开展实验,该库使用 10 个 Otto Bock MyoBock 13 E200 电极以 100 Hz 采样频率采集 10 通道无线电极 SEMG 信号。DB1 包含 27 名健全受试者, 每位受试者重复执行 52 种手指、手部及腕部动作,手势 标签及分类如表 1 所示。

表 1 手势标签与动作分类

Table 1 Gesture labeling and action classification	
标签	手势动作分类
1~12	手指动作
13~29	手或手腕部分的动作
30~52	抓握动作

数据库收录了 20 名男性和 7 名女性(年龄 28±3.4 岁, 其中 25 名右利手,2 名左利手)的肌电信号。每个动作 重复 10 次,每次持续 5 s 执行后间隔 3 s 休息。为防止实 验结果过于依赖测试集数据质量,实验遵循数据库规范 及既往研究的建议^[22-24],将 10 次动作重复按 6:2:2 的 比例随机划分为训练集、验证集和测试集,并采用五折交 叉验证。受试者 s1 的各手势类别样本数所占百分比如 图 1 所示。

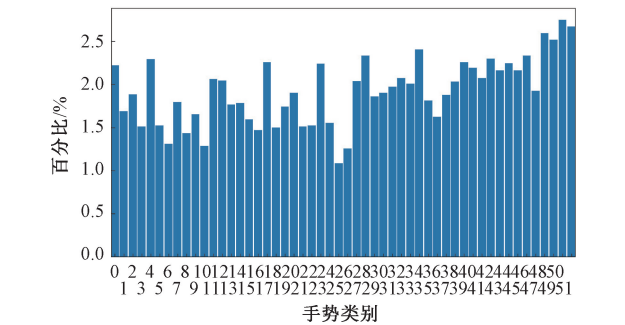


图 1 各手势类别样本数量占比

Fig. 1 Proportion of samples per gesture category

所有受试者各类别手势样本数量不平衡比率(IR) 均小于 5,变异系数(CV) 均小于 30%,样本占比呈现均 衡分布特征,符合实验设计中的类别平衡性要求^[25]。

$$CV = \frac{\sigma}{\mu} \times 100\%$$

(1)

$$IR = \frac{N_{\max}}{N_{\min}} \quad (2)$$

式中: σ 表示各类别样本数的标准差; μ 表示各类别样本数的均值; $N_{\max}$ 、 $N_{\min}$ 分别表示数据集中样本数量最多与最少的类别的样本数。

1.2 低通滤波与归一化

SEMG 信号的主要频率范围为 10~500 Hz, 通常集中在 10~150 Hz^[26], 在 Ninapro DB1 数据集中, 原始肌电信号经过 Otto Bock 电极的放大, 50 Hz 的工频滤波后再进行整流与低通滤波以提取 SEMG 的包络信号以反应肌肉激活强度随时间的变化, 频率范围在 0~25 Hz, 随后以 100 Hz 频率采样得到原始输入信号^[27]。文献[28]通过相干性分析指出 sEMG 的包络信号与肌肉力度的相关性集中在 0~3 Hz 频段, 高频成分(>5 Hz)与力度的相关性显著下降。使用截止频率为 5 Hz 的一阶低通巴特沃斯滤波器进行滤波去除高频噪声, 肌电信号经过巴特沃斯低通滤波的前后对比如图 2 所示。

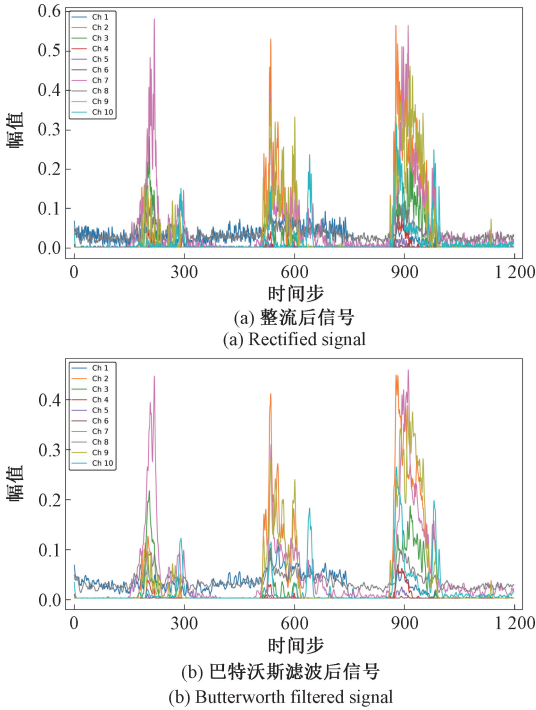


图2 肌电信号低通滤波前后对比

Fig. 2 Comparison before and after low pass filtering

SEMG 信号在时域中的有效信息多集中于零值附近^[24], 该实验采用基于 μ 律变换的非线性归一化方法, μ 律归一化通过对数特性放大小幅值传感器的输出, 同时保持大幅值传感器的比例关系以提升信号的可用性, 计算公式如式(3)所示。

$$F(X_t) = \text{sign}(X_t) \frac{\ln(1 + \mu |X_t|)}{\ln(1 + \mu)} \quad (3)$$

对于多通道肌电信号, 不同通道的信号幅度范围差异较大(图2), 为避免 μ 律归一化导致信号失真及弱通道信噪比大幅下降, 先对多通道肌电信号做通道独立 Minmax 归一化, 将幅值限定在 [0, 1]。 μ 律归一化中的参数 μ 控制着非线性压缩的程度, 其取值直接影响信号的形态与后续识别性能, μ 的典型取值范围是 100~500, 为确定最优的 μ 值, 以小波门控时序卷积网络(wavelet gated-temporal convolutional network, WG-TCN)作为分类器, 对 μ 值在常用取值 64、100、128、255 内进行了格点搜索, 当 $\mu=128$ 时, 模型在测试集上取得了最高的识别准确率 81.85%, 故选定 $\mu=128$ 。经过 μ 律归一化后的肌电信号如图 3 所示。

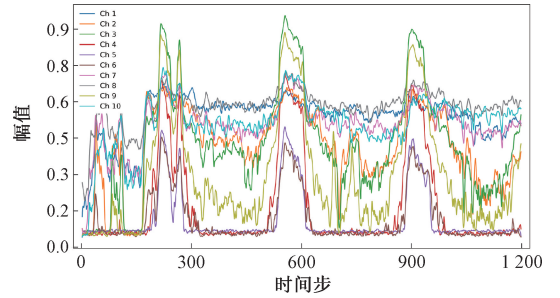


图3 μ 律归一化后肌电信号

Fig. 3 μ -Law normalized EMG signal

2 网络结构设计

2.1 小波门控时序卷积网络

通过增加小波卷积模块与门控单元来提升时序卷积网络的时频分析与特征提取能力, 网络由小波卷积模块与门控时序卷积网络两部分组成, 其总体结构如图 4 所示。

原始肌电信号首先经由小波卷积模块进行多尺度分解, 获得细节系数与近似系数; 随后分别对这两类系数进行一维卷积处理, 通过可学习的卷积操作, 使模型能够自适应地调整各频带子信号的贡献, 然后将卷积结果重构为信号表示。重构后的信号输入至融合门控单元的时序卷积网络, 以同时提取时域与频域特征, 其中每一层时序卷积均通过门控机制对特征进行筛选。最终, 将特征通过全连接层进行分类, 输出对应手势类别。门控时序卷积网络包含两个门控时序卷积模块, 第 1 个模块的卷积层的输出通道数分别是 32、64, 卷积核大小为 3, 第 2 个模块的卷积层输出通道数分别是 64、128, 卷积核大小同样为 3, 为加快模型收敛速度, 每个卷积层后加入 BN, ReLU 非线性激活函数。

为防止过拟合, 每个门控时序卷积模块后都加入 Dropout 层, 丢弃率为 0.05, 最后将最后一个模块的输出 $X^{(b)} \in R^{C \times T}$ 经过在时间步维度上求平均(average over

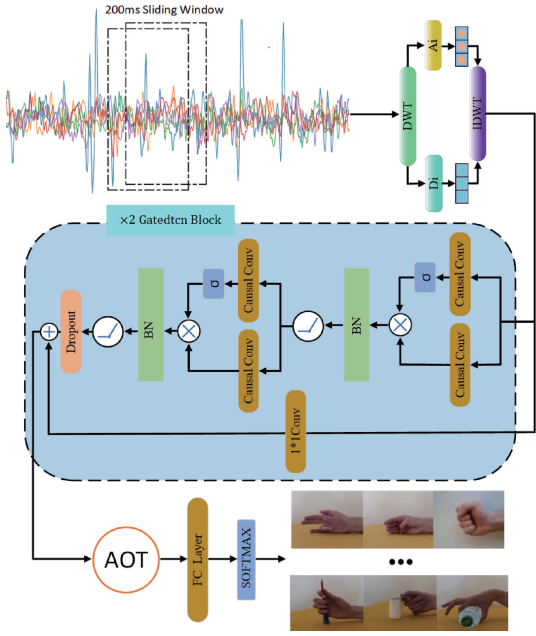


图 4 网络总体结构

Fig. 4 Network architecture

time)后再经过全连接层进行标签预测。如式(4)、(5)所示。

$$y_c^{(b)} = \sum_{i=1}^T X_{c,t}^{(b)} \tag{4}$$

$$\hat{\theta} = \text{softmax}(W_0 \times y_c^{(b)} + b_0) \tag{5}$$

式中: T 表示时间步个数; C 表示通道数; b 表示第 b 个样本; W_0 、 b_0 表示最后一个全连接层的可训练参数; $\hat{\theta}$ 表示最后各个类别的预测概率。

2.2 小波卷积模块

肌电包络信号自身蕴含多尺度生理信息^[29],离散小波变换进一步解耦该趋势内部的动态成分:低频分量与肌肉力量输出、疲劳趋势强相关,细节系数反映运动单元同步化、协调性控制或病理性震颤^[30]。此类相对高频分量已被证明与运动控制策略^[31]、及动作模式鉴别^[32]显著相关,为肌肉功能评估提供多分辨率视角。Phinyomark等^[33]在小波域对 SEMG 信号进行阈值去噪后重构,显著提升分类精度。

小波卷积模块将肌电包络信号在时频域上逐层分解以提取信号的多尺度结构,将得到的各个尺度下的低频系数 (approximation coefficients, AC) 与高频系数 (detail coefficients, DC) 分别进行卷积,以引导模型更有效地捕获肌电信号中的低频特征,同时通过可学习的卷积核实现了特征的自适应选择与增强,然后逐层进行离散逆小波变换 (inverse discrete wavelet transform, IDWT),将增强后的多尺度特征融合成富含关键时频信息的表示,使得模型更好地处理时序信号的动态变化。为方便展示,2

级小波分解卷积重构过程如图 5 所示。

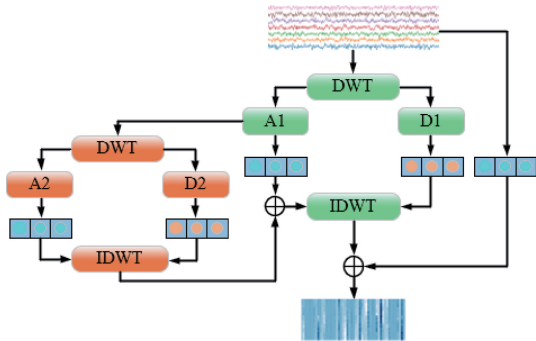


图 5 小波卷积模块示意图

Fig. 5 Schematic diagram of the wavelet convolution module

首先将输入的多通道肌电信号 X 按通道进行多级离散小波分解,得到多级近似系数 X_{Li} 与细节系数 X_{Hi} :

$$[X_{Li}, X_{Hi}] = \text{Conv}([f_L, f_H], X_{i-1}) \tag{6}$$

式中: f_L 、 f_H 分别表示小波变换中的低通与高通滤波器。

接着,对每一级别的低频与高频分量进行一维分组卷积,以此来捕获时频特征,如式(7)所示。

$$[Y_{Li}, Y_{Hi}] = \text{Conv}(W_i, (X_{Li}, X_{Hi})) \tag{7}$$

式中: Y_{Li} 、 Y_{Hi} 表示卷积后得到的低频与高频分量; W_i 表示一维卷积的权重。卷积核形状设为 $(K_i, 1)$ 。

卷积处理后的低频与高频分量将被用于重构当前层级的低频成分,然后将其与上一级卷积后的高频成分进行小波逆变换,如式(8)所示。

$$Z_i = \text{IWT}(Y_{Li} + Z_{i-1}, Y_{Hi}) \tag{8}$$

这一重构过程不仅保留了信号的时域特征,也融合了从不同尺度捕获的频域特征,增强了模型对复杂时序动态的解析能力。

其中小波基函数的选取应满足几个标准:具备对称性以减少相位畸变,具有适当的消失矩和足够的正则性,且与原始信号具有相似性^[34-36]。基于以上几点,选取了 db4、db6、sym5、coif3 等小波,当分解级数越高,最底层低频子带的带宽越窄,其边界效应和频带交叠导致的泄漏也越显著^[37],故进行 4 级以内分解,实验结果如表 2 所示。

表 2 不同小波基函数对分解准确率的影响
Table 2 Wavelet decomposition accuracy statistics

(%)				
小波基	Lv1	Lv2	Lv3	Lv4
db4	79. 17	79. 83	80. 78	80. 55
db6	80. 45	80. 92	80. 81	80. 40
sym5	79. 80	80. 83	81. 07	80. 60
coif3	79. 32	79. 95	80. 72	80. 51

选取 sym5 为小波基,分解级数为 3,卷积核大小为

3. sym5 小波的尺度函数 (具有低通特性) 与小波函数 (具有高通特性) 的波形如图 6 所示。

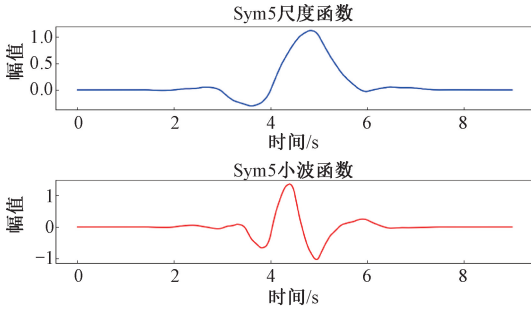


图 6 sym5 小波基
Fig. 6 sym5 wavelet basis

Ninapro DB1 数据集受试者 s1 握拳时连续 5 帧的 (时间窗口 200 ms, 窗口步长 10 ms) 原始肌电信号经过 3 级小波卷积模块重构后得到的特征图如图 7 所示。

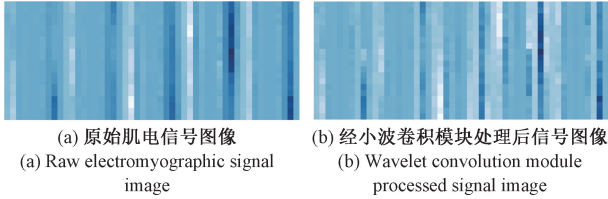


图 7 肌电信号经小波卷积模块处理前后特征图
Fig. 7 EMG feature maps before and after wavelet convolutional processing

虽然重构后的信号与原始信号相比存在差异,后续与基线模型的实验结果对比表明,这种差异本质是模型通过端到端训练主动学习的特征优化过程。

2.3 融合门控单元的时序卷积神经网络

时序卷积神经网络 (TCN) 通过前向零填充 (zero-padding) 输入序列,严格保证当前时间步的输出仅依赖于历史时间步,从而消除未来信息的潜在泄漏。卷积核的扩张使感受野能够覆盖更长的时间信息。因果卷积中填充长度 (padding) 与卷积核尺寸 k , 扩张率 d 的关系如式 (9) 所示。

$$p^l = (k^l - 1) \times d^l \quad (9)$$

式中: d^l 表示当前层的扩张率,扩张率分别为 1、2、4、8,卷积核大小为 3。

对于输入一维序列 $x \in \mathbf{R}^n$, 输出序列在位置 s 处的值经扩张因果卷积的计算方式如式 (10) 所示。

$$F(s) = \sum_{i=0}^{k-1} f(i) \times x_{s-d \cdot i} \quad (10)$$

式中: $f(i)$ 为滤波器的第 i 个权重参数; $x_{s-d \cdot i}$ 为上一层中的第 $s - d \cdot i$ 个元素。

此外残差结构缓解深层网络梯度消失问题,使得模型能够训练更多的层数。结构如图 8 所示。由于肌电信号中不同手势的激活模式差异可能非常细微,在此基础上,为了使模型能够自动学习不同时间点特征的重要性权重,屏蔽对结果影响较小的特征,引入门控单元 (gated linear unit, GLU), 如式 (11) 所示。

$$h_t(X) = (X \times W + b) \otimes \sigma(X \times V + c) \quad (11)$$

式中: X 为当前层的输入; W, V 为不同的卷积核权重; b, c 为偏置项; $\otimes$ 表示 Hadamard 积; σ 表示激活函数。

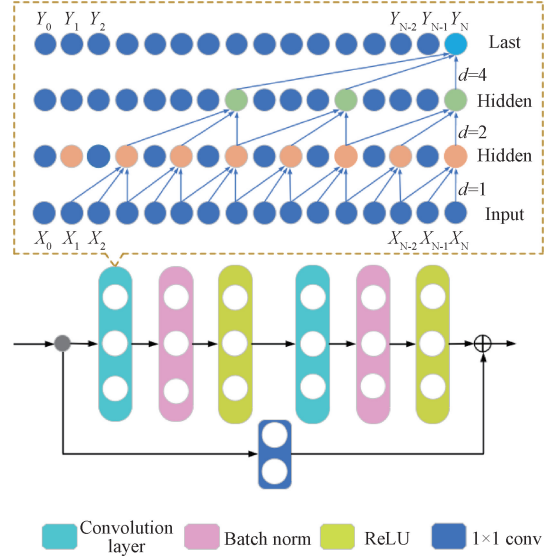


图 8 时序卷积神经网络结构示意图

Fig. 8 Schematic diagram of temporal convolutional neural network architecture

为了增强模型对序列数据中跨时间步信息交互的学习能力,并进一步强调跨时间步之间的依赖关系,如式 (12) 所示,首先对输出矩阵 $A \in \mathbf{R}^{T \times D}$ 进行转置操作,使其维度由 $[T, D]$ (时间步,通道数) 变为 $[D, T]$ (通道数,时间步),矩阵的每行代表通道在所有时间步上构成的完整时间序列,后对其使用激活函数,使其能够基于上下文为每个特征通道的时间序列生成门控信号,此后再次转置回原始维度 $[T, D]$,与另一卷积后的输出 $B \in \mathbf{R}^{D \times T}$ 进行逐元素相乘,其流程如图 9 所示。

$$Gated = \text{sigmoid}(A^T)^T \otimes B \quad (12)$$

3 实验验证

3.1 实验设置

所有实验均在配置为 Intel Core i7-12700K CPU、NVIDIA GeForce RTX 3060 GPU (12 GB 显存) 和 16 GB DDR4-3200 MHz 内存的台式机上完成。模型基于 PyTorch 2.0.1 框架实现,通过 CUDA 11.7.1 加速。为避

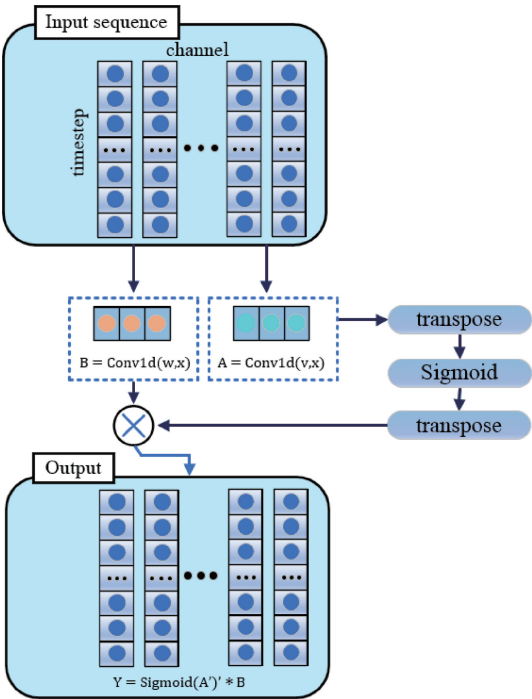


图 9 门控单元流程

Fig. 9 Flowchart of the gating unit

免实验结果依赖测试集数据质量,在每种手势的 10 次重复中,按 6 : 2 : 2 比例随机划分出训练集,验证集与独立测试集,采用 5 折交叉验证,在训练集上训练模型,验证集上进行超参数选择与模型调优,最终模型在测试集的表现作为实验结果。输入数据的滑动窗口长度为 200 ms,移动步长为 10 ms,批次大小为 512。训练共 66 个 epoch,每个 epoch 耗时约 34 s,使用交叉熵损失函数(cross entropy loss)和 Adam 优化器,初始学习率 0.01,学习率衰减策略为每 5 个 epoch 乘以 0.9。为抑制过拟合,采用 L2 正则化(权重衰减 1×10^{-6})。

3.2 消融实验

为了评估小波门控时序卷积神经网络中各模块对模型性能提升的贡献度,该实验设计了 4 组对照实验,选取准确率(accuracy)、精确率(precision)、召回率(recall)和 F1 分数作为评估指标。所有实验指标均基于 DB1 数据集所有受试者的平均值进行计算,各组模型及其实验结果如表 3 所示。

表 3 消融实验各评价指标对比

Table 3 Metrics comparison for ablation experiments

模型	准确率/%	精确率/%	召回率/%	F1 分数	参数量
TCN (baseline)	76.95	78.89	77.29	0.77	59 700
小波卷积模块+TCN	80.55	81.20	80.79	0.80	59 980
门控单元+TCN	80.45	81.25	80.72	0.81	103 668
小波门控时序卷积	81.85	82.80	82.13	0.82	103 948

实验结果表明,相较于基线 TCN 模型,单独引入小波卷积模块和门控单元分别使分类准确率提升了 3.6% 和 3.5%,虽然两种网络结构改进策略对模型性能的提升较为均衡,但小波卷积模块只增加了少量参数。其中在原始 TCN 的基础上仅加入小波卷积模块与基线模型相比,准确率提高了 3.6%,F1 分数提高了 0.03,这表明经小波卷积模块进行分解—卷积—重构后的信号虽不与原始信号保持一致,但这种差异并非信息损失,而是模型通过端到端训练实现的判别性特征增强。本文提出的小波门控时序卷积网络(WG-TCN)获得了最高准确率 81.85%,比基线模型的准确率高出了 4.9%,且 F1 分数高出 0.05。原始 TCN 与 WG-TCN 模型在以 200 ms 滑动时间窗作为输入的收敛结果如图 10 所示,其中 TCN 的验证集损失在 1.0 左右,且存在较大波动,训练集损失在 0.3,网络存在过拟合,WG-TCN 模型收敛效果较好,无过拟合现象。

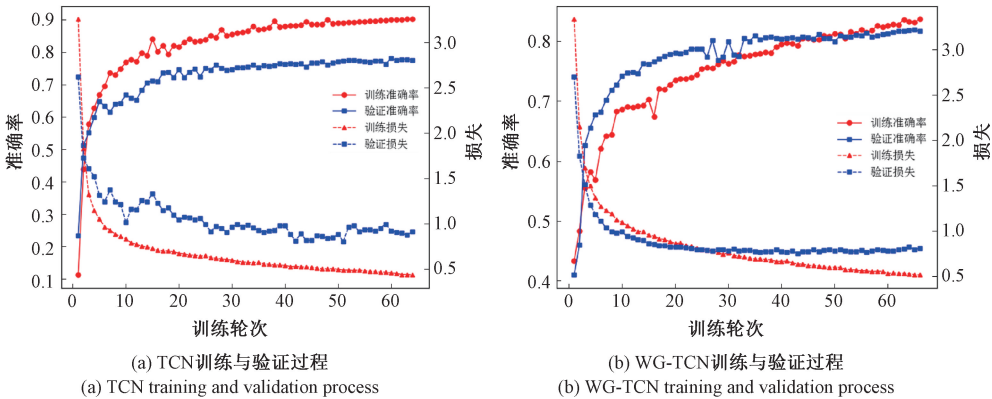


图 10 模型训练结果

Fig. 10 Model training result

两种模型在 52 类手势分类的混淆矩阵如图 11 所示,其中识别结果主要集中在对角线上,由于相邻标签手势动作相近,误分类样本主要出现在相邻类别上。

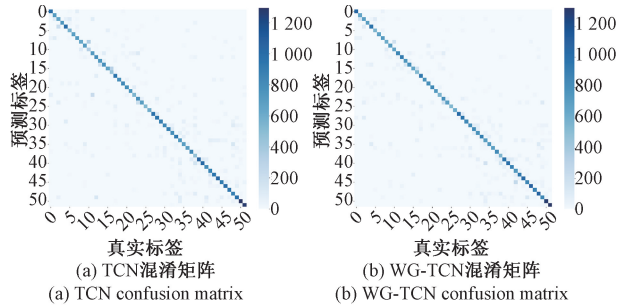


图 11 模型混淆矩阵
Fig. 11 Model confusion matrix

如图 12 所示,通过对比 WG-TCN 与 TCN 模型的特异性准确率可以发现,对于基线模型表现较差的第 6、9、11、16、35、39 和 43 类手势(准确率均低于 60%),WG-TCN 平均准确率提升了 18.6%,这充分体现了模型对复杂手势特征的增强识别能力。WG-TCN 通过特征空间的优化重构,使其更专注于提升难样本的区分能力。

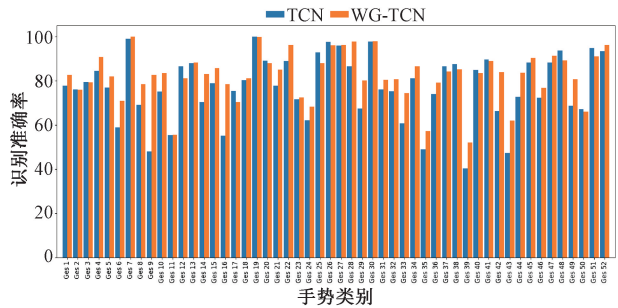


图 12 TCN 与 WG-TCN 的各类手势识别准确率对比
Fig. 12 Comparative analysis of gesture recognition accuracy between TCN and WG-TCN

3.3 本模型与其他方法对比

本研究将近年来基于深度学习的肌电信号手势识别相关模型与 WG-TCN 网络模型进行性能对比,准确率计算方式为 DB1 数据集所有受试者平均值,如表 4 所示。WG-TCN 网络模型在 DB1 数据集(包含 52 种手势)上的识别性能优于传统模型,不仅以更少的参数量实现了更高的识别准确率,同时其综合性能指标已达到基于稀疏表面肌电信号的手势识别领域的国际先进水平。虽然本模型在识别准确率上较 WeiNet 略低 2.42%,但 WeiNet 采用了多流卷积架构,需对肌电信号的每个通道执行独立卷积运算,导致其参数量激增至本模型的 70 倍。MSHilbNet 由于采用 Hilbert 空间曲线填充方法,要求输入信号的时间步长必须为 2 的整数次幂,因此在保证实

时性要求的前提下,其滑动窗口长度被限制为 160 ms。尽管 MSHilbNet 具有较小的模型规模,但本研究所提出的 WG-TCN 模型在识别准确率上显著优于 MSHilbNet。此外,与 GengNet 采用瞬时肌电图像作为卷积网络输入并通过多数投票进行分类的方法相比,WG-TCN 网络模型在显著减少参数数量的同时仍保持了较高的识别准确率。以上结果表明,通过小波卷积模块对肌电信号进行时频域分析,利用时序卷积神经网络学习其序列依赖关系,并结合门控机制优化信息流动路径,这种架构能挖掘肌电信号非平稳特性下隐含的时序信息。

表 4 sEMG 手势识别模型性能对比
Table 4 Performance comparison of sEMG gesture recognition models

模型	年份	数据集	窗口长度/	手势	准确率/%	参数量
			ms	数量		
WeiNet ^[23]	2017	DB1	200	52	84.27	7 277 625
GengNet ^[10]	2016	DB1	200	52	77.80	500 347
MSHilbNet ^[38]	2021	DB1	160	52	74.10	62 453
MST-SE-Attn ^[5]	2025	DB1	200	52	80.18	238 925
DCN ^[39]	2022	DB1	100	52	78.94	
文献[40]	2021	DB1	200	52	78.18	
2SRNN ^[41]	2019	DB1	200	52	77.8	
LSTM-CNN ^[21]	2021	DB1	150	53	75.12	
WG-TCN		DB1	200	52	81.85	103 948

4 结 论

为了解决深度学习网络模型由于肌电信号的变异性而导致的识别准确率低的问题,本文基于 TCN 能够有效捕捉肌电信号时序特性的优势,创新性地引入了小波卷积模块与门控机制,从而提升了网络对肌电信号时频特征的提取能力。原始肌电信号中低频成分关联肌肉疲劳与运动单元募集,高频成分关联运动单元放电速率与快速发力。为了捕获并增强肌电信号中不同频带的对应的特异性生理活动,利用小波卷积模块对肌电信号进行处理,小波卷积模块本身是一种可学习自适应时频分析系统,首先对肌电信号高低频进行解耦,然后通过可学习一维卷积核分解后得到的低频与高频分量进行卷积,引导模型自适应的学习与任务相关的时频特征,同时抑制不相关或噪声成分,并对卷积后的信号进行重构。接着将自适应增强后的信号送入融合门控单元的时序卷积网络以捕获重构信号中的上下文时序关系,并使用门控单元对时序卷积网络中每层卷积层提取到的特征进行动态筛选,优化信息流动路径。

WG-TCN 采用 200 ms 滑动窗口作为输入,在 52 类手势分类任务中实现了 85.82% 的准确率。小波离散卷

积模块引入一维卷积核,通过增加少量参数(180 个可训练参数),取得了相较于基线模型 3.6% 的准确率的提升,并且得益于时序卷积网络中因果卷积与膨胀卷积的特性,WG-TCN 总体参数量远低于当前主流模型。实验结果表明,与当前主流肌电信号手势识别模型相比,WG-TCN 在保持较低模型复杂度的前提下,仍能获得具有竞争力的分类性能。

未来工作将重点评估本模型在自采数据集上的识别准确率,并系统考察其在跨受试者场景中的泛化性能,最终完成在嵌入式设备上的部署,提升实际应用价值。

参考文献

- [1] YOO J W, LEE D R, SIM Y J, et al. Effects of innovative virtual reality game and EMG biofeedback on neuromotor control in cerebral palsy [J]. *Bio-medical materials and engineering*, 2014, 24(6):3613-3618.
- [2] SHIN J, MIAH A S M, KONNAI S, et al. Hand gesture recognition using sEMG signals with a multi-stream time-varying feature enhancement approach [J]. *Scientific Reports*, 2015, 14(1):1-13.
- [3] 赵世昊,周建华,伏云发. 注意力机制 CNN 结合肌电特征矩阵的手势识别研究[J]. *电子测量与仪器学报*, 2023, 37(6):59-67.
ZHAO SH H, ZHOU J H, FU Y F. Research on gesture recognition based on attention mechanism CNN combined with EMG feature matrix [J]. *Journal of Electronic Measurement and Instrumentation*, 2023, 37(6):59-67.
- [4] 李思诺,耿艳娟. 基于表面肌电信号估计多自由度手指力的 ShuffleVT 模型[J]. *电子测量技术*, 2025, 48(5):102-110.
LI S N, GENG Y J. ShuffleVT model for estimating multi-degree-of-freedom finger force based on surface electromyography signals [J]. *Electronic Measurement Technology*, 2025, 48(5):102-110.
- [5] SHIN J, MIAH A S M, KONNAI S, et al. Electromyography-Based gesture recognition; hierarchical feature extraction for enhanced spatial-temporal dynamics [J]. 2025. DOI: 10. 1109/ACCESS. 2025. 3569899.
- [6] MANFREDO A, MATTEO C, MÜLLER HENNING. 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, 10. DOI: 10. 3389/fnbot. 2016. 00009.
- [7] TSINGANOS P, CORNELIS B, CORNELIS J, et al. Deep learning in EMG-based gesture recognition [J]. 2019. DOI:10. 5220/0006960201070114.
- [8] LI J, JIANG X, FAN J, et al. Deep end-to-end transfer learning for robust inter-subject and inter-day hand gesture recognition using surface EMG [J]. *Biomedical Signal Processing and Control*, 2025, 100(PartC). DOI: 10. 1016/j. bspc. 2024. 106892.
- [9] JAVIER ORLANDO PINZÓN-ARENAS, ROBINSON JIMÉNEZ-MORENO, HERRERA-BENAVIDES J E. Convolutional neural network for hand gesture recognition using 8 different EMG signals [J]. *Faculty of Engineering, Nueva Granada Military University*, 2019, DOI:10. 1109/STSIVA. 2019. 8730272.
- [10] GENG W, DU Y, JIN W, et al. Gesture recognition by instantaneous surface EMG images [J]. *Scientific Reports*, 2016, 6:36571. DOI:10. 1038/srep36571.
- [11] BAI S, KOLTER J Z, KOLTUN V. Trellis networks for sequence modeling [J]. 2018. DOI: 10. 48550/arXiv. 1803. 01271.
- [12] TSINGANOS P, CORNELIS B, CORNELIS J, et al. Improved gesture recognition based on sEMG signals and TCN[C]. *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2019.
- [13] BETTHAUSER J L, KRALL J T, BANNOWSKY S G, et al. Stable responsive EMG sequence prediction and adaptive reinforcement with temporal convolutional networks [J]. *IEEE Transactions on Biomedical Engineering*, 2020, 67(6):1707-1717.
- [14] YANG D, GU Y, THAKOR N V, et al. Improving the functionality, robustness, and adaptability of myoelectric control for dexterous motion restoration[J]. *Experimental Brain Research*, 2019, 237(2):291-311.
- [15] AO X, SHE J, WANG F, et al. Enhanced gesture recognition via interpretable temporal convolutional networks utilizing spatio-temporal sEMG features[C]. 2024 *IEEE Conference on Pervasive and Intelligent Computing (PICom)*, 2024.
- [16] MA W, SONG G, ZENG Q, et al. FFCSLT: A deep learning model for traffic police hand gesture recognition using surface electromyographic signals [J]. *IEEE Sensors Journal*, 2024, DOI: 10. 1109/JSEN. 2024. 3371588.
- [17] LIN C. sEMG-based continuous estimation of finger kinematics via large-scale temporal convolutional network[J]. *Applied Sciences*, 2021, DOI:10. 3390/app11104678.
- [18] LIU T, BAI D, MA L, et al. Complex surface electromyography signal gesture recognition based on multipathway featured scale convolutional neural network[J].

- IEEE Transactions on Instrumentation and Measurement, 2024, DOI:10.1109/TIM.2024.3485448.
- [19] ATZORI M, GIJSBERTS A, KUZBORSKI I, et al. Characterization of a benchmark database for myoelectric movement classification [J]. IEEE Trans Neural Syst Rehabil Eng, 2015, 23(1):73-83.
- [20] 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):6.
- LI Y H, JIANG Q, ZOU K, et al. Multi-stream convolutional EMG gesture recognition network fused with attention mechanism [J]. Application Research of Computers, 2021, 38(11):6.
- [21] 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 EMG recognition based on long short-term memory and convolutional neural network [J]. Chinese Journal of Scientific Instrument, 2021, 42(2):162-170.
- [22] 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.
- [23] WEI W, WONG Y, DU Y, et al. A multi-stream convolutional neural network for sEMG-based gesture recognition in muscle-computer interface [J]. Pattern Recognition Letters, 2017, DOI: 10.1016/j.patrec.2017.12.005.
- [24] RAHIMIAN E, ZABIHI S, ATASHZAR S F, et al. Xceptiontime: Independent time-window xceptiontime architecture for hand gesture classification [J]. IEEE, 2020, DOI:10.1109/ICASSP40776.2020.9054586.
- [25] JOSEPHS D, DRAKE C, HEROY A, et al. sEMG gesture recognition with a simple model of attention, 2020, DOI:10.48550/arXiv.2006.03645.
- [26] WEI W, REN L, ZHOU M, et al. Improving unimodal sEMG-based pattern recognition through multimodal generative adversarial learning [J]. IEEE Transactions on Instrumentation and Measurement, 2025, DOI:10.1109/TIM.2025.3556214.
- [27] ATZORI M, GIJSBERTS A, HEYNEN S, et al. Building the ninapro database: A resource for the biorobotics community [C]. 2012 4th IEEE RAS & EMBS International Conference on Biomedical Robotics and Biomechatronics (BioRob). IEEE, 2012, DOI:10.1109/BioRob.2012.6290287.
- [28] HOGAN N, MANN R W. Myoelectric signal processing: optimal estimation applied to electromyography-Part I: Derivation of the optimal myoprocessor [J]. IEEE Transactions on Bio-medical Engineering, 1980, 27(7):396.
- [29] BRAD S, SHERRY S. The extraction of neural strategies from the surface EMG. [J]. Journal of Applied Physiology, 2004, 96(4):1486-1495.
- [30] SHIAVI, R. Electromyography: Physiology, engineering, and noninvasive applications [Book Review] [J]. Engineering in Medicine & Biology Magazine IEEE, 2006, 25(6):11-12.
- [31] TSCHARNER V V. Intensity analysis in time-frequency space of surface myoelectric signals by wavelets of specified resolution [J]. J Electromyogr Kinesiol, 2000, 10(6):433-445.
- [32] TARPIN-BERNARD, LAURILLAU, SERVIERE, et al. Feature extraction of the first difference of EMG time series for EMG pattern recognition [J]. Computer Methods and Programs in Biomedicine: An International Journal Devoted to the Development, Implementation and Exchange of Computing Methodology and Software Systems in Biomedical Research and Medical Practice, 2014.
- [33] PHINYOMARK A, PHUKPATTARANONT P, LIMSAKUL C. Wavelet-based denoising algorithm for robust emg pattern recognition [J]. Fluctuation & Noise Letters, 2011, 10(2):157-167.
- [34] WANG T, LU C, SUN Y, et al. Automatic ECG classification using continuous wavelet transform and convolutional neural network [J]. Entropy, 2021, 23(1):119.
- [35] SUN B, CHEN W, TIAN Y. The pattern recognition of surface EMG based on wavelet transform and BP neural network [J]. Journal of Fuyang Teachers College, 2011, DOI:10.1109/icbbe.2011.5780259.
- [36] KAKOTY N M, HAZARIKA S M, GAN J Q. EMG feature set selection through linear relationship for grasp recognition [J]. Journal of Medical & Biological Engineering, 2016, 36(6):1-8.
- [37] WANG L, FU J, CHEN H, et al. Hand gesture recognition using smooth wavelet packet transformation and hybrid CNN based on surface EMG and accelerometer signal [C]. 2023.
- [38] TSINGANOS P, CORNELIS B, CORNELIS J, et al.

- Hilbert sEMG data scanning for hand gesture recognition based on deep learning [J]. Neural Computing & Applications, 2021, DOI: 10.1007/s00521-020-05128-7.
- [39] WANG H, ZHANG Y, LIU C, et al. sEMG based hand gesture recognition with deformable convolutional network[J]. International Journal of Machine Learning and Cybernetics, 2022, DOI:10.1007/s13042-021-01482-7.
- [40] 郑楠,李玉荣,张文轩,等. 基于特定任务肌肉协同的用户无关肌电手势识别研究[J]. 仪器仪表学报, 2021,42(9):253-261.
- ZHENG N, LI Y R, ZHANG W X, et al. Research on user-independent EMG gesture recognition based on task-specific muscle synergy[J]. Chinese Journal of Scientific Instrument,2021,42(9):253-261.
- [41] ISTVÁN K, FERENC K, KRISZTIÁN Z V, et al. Domain adaptation for sEMG-based gesture recognition with recurrent neural networks[J]. IEEE, 2019, DOI: 10.1109/IJCNN.2019.8852018.

作者简介



孙鹏,南京理工大学硕士研究生,主要研究方向为手势识别。

E-mail:17368950686@163.com



Sun Peng is currently a M. Sc. candidate from Nanjing University of Science and Technology. His Main Research interest includes Gesture Recognition.

李忠新(通信作者),分别在 2000 年和 2004 年获得南京理工大学硕士,博士学位,南京理工大学机械工程学院副教授,硕士生导师,主要研究方向为手势控制、图像处理与分析。

E-mail:njustlx@163.com

Li Zhongxin (Corresponding author) received his M. Sc. degree in 2000 and Ph. D. in 2004 from Nanjing University of Science and Technology. He is currently an associate professor and M. Sc. supervisor at the School of Mechanical Engineering, Nanjing University of Science and Technology. His main research interests include gesture control, image processing and analysis.