

DOI:10.19651/j.cnki.emt.2518011

# 通道与空间特征协同提取的抗噪手势识别方法<sup>\*</sup>

逯英航 仇大伟 刘静 李全伟 王锡城

(山东中医药大学医学信息工程学院 济南 250355)

**摘要:** 表面肌电信号作为肌肉活动的直接反映,能够有效捕捉肌肉收缩的模式和强度信息,因此在手势识别中被广泛应用。然而,其稀疏性、非线性和噪声干扰对特征提取提出了严峻挑战。为此,本文提出了 RASTNet 模型,以 ResNet50 为主干网络,将每层最后一个 block 中的  $3 \times 3$  卷积替换为空洞空间金字塔池化模块,通过不同空洞率的空洞卷积捕获 sEMG 多尺度信息。然后在每一层的连接处加入 STConv 模块,该模块在 SCConv 模块的基础上创新性地融入了三重注意力机制,在提取精细化通道与空间特征后进一步强化三维特征融合。本研究分别在用 4 种方法进行数据增强的 ninapro DB1 和 DB5 数据集上进行实验。结果表明,RASTNet 较原模型准确率平均提升了 1.83% 和 1.57%。与 ResNeXt、Swin Transformer、CnovNeXt 等主流经典模型在拟真噪声下横向对比,其召回率、F1 分数等指标均表现更优。并且在面对最新的无噪声闭源模型时仍保持领先,展现出在复杂手势识别任务中的抗噪性。此外,RASTNet 在跨数据集的泛化性验证中表现出色,进一步增强了其在实际应用中的适用性和鲁棒性。

**关键词:** 表面肌电信号; 手势识别; 通道与空间特征提取; 抗噪

中图分类号: TN911.7 文献标识码: A 国家标准学科分类代码: 510

## Noise-resistant gesture recognition method based on the collaborative extraction of channel and spatial features

Lu Yinghang Qiu Dawei Liu Jing Li Tongwei Wang Xicheng

(College of Intelligence and Information Engineering, Shandong University of Traditional Chinese Medicine, Jinan 250355, China)

**Abstract:** Surface electromyography directly reflects muscle activity, effectively capturing muscle contraction patterns and intensity, making it widely used in gesture recognition. However, its sparsity, non-linearity, and noise interference pose significant challenges for feature extraction. To address this, we propose the RASTNet model, using ResNet50 as the backbone and replacing the  $3 \times 3$  convolution in each layer's last block with an Atrous Spatial Pyramid Pooling module to capture multi-scale information via dilated convolutions. An STConv module, incorporating a triple attention mechanism into SCConv, is added to enhance the fusion of channel and spatial features. Experiments on the NinaPro DB1 and DB5 datasets, augmented with four methods, show that RASTNet improves accuracy by 1.83% and 1.57% on average. Compared with models like ResNeXt, Swin Transformer, and CnovNeXt under simulated noise, RASTNet outperforms in recall rate, F1 score, and other metrics. It also remains superior to the latest closed-source models without noise, demonstrating robustness and noise resistance in complex gesture recognition tasks. Additionally, RASTNet shows strong generalization across datasets, enhancing its real-world applicability and robustness.

**Keywords:** sEMG; gesture recognition; channel and spatial feature extraction; noise resistance

## 0 引言

表面肌电信号(surface electromyography,sEMG)通过记录骨骼肌连续收缩过程中的生物电活动,提供了关于肌

肉收缩模式和强度的关键信息,因其非侵入性和高效捕捉人体运动意图的能力,已在智能假肢<sup>[1]</sup>、机器人<sup>[2]</sup>、人机交互<sup>[3]</sup>等领域得到广泛应用。

基于 sEMG 的手势识别问题通常被归纳为多分类任

收稿日期:2025-01-28

\* 基金项目:国家自然科学基金面上项目(82174528)、山东省研究生精品和优质课程建设项目(SDYAL21054,SDYKC2023041)、山东中医药大学青年科研创新团队项目(校科字[2024]1号)、山东中医药大学科学研究基金面上项目(KYZK2024M14)资助

务,传统方法依赖于在时域、频域或时频域中手动提取特征<sup>[4]</sup>,并利用支持向量机<sup>[5]</sup>、随机森林<sup>[6]</sup>、线性判别分析<sup>[7]</sup>等统计模型进行分类。然而,随着深度学习的发展,端到端模型逐渐取代了人工特征工程,其中卷积神经网络(CNN)在捕捉 sEMG 动态特征方面表现出色,不仅在肌电图手势识别中获得了较传统方法更高的准确率<sup>[8-13]</sup>,还出现了融合长短记忆(LSTM)的 CNN-LSTM 模型<sup>[14-15]</sup>,以及通过将自注意力机制融入到神经网络模型中以捕捉全局依赖关系的尝试。例如,Hu 等<sup>[16]</sup>提出了一种基于注意力机制的混合 CNN-RNN 网络结构,基于 NinaPro DB1 数据集实现了 84.80% 的平均手势识别准确率。与此同时,多尺度特征提取模块利用不同大小的卷积核和池化操作,有效促进了跨尺度信息的融合。例如,Shen 等<sup>[17]</sup>提出了基于多尺度模块的 sEMG 手势分类模型,在 NinaPro DB5 上实现了 79.43% 的识别准确率。Jiang 等<sup>[18]</sup>引入了基于 inception 多尺度融合卷积和 ECA 机制的 RIE 模型,在 NinaPro DB1 上实现了 88.27% 的平均准确率,超过传统 CNN 7.89%。尽管如此,sEMG 信号的稀疏性、环境噪声以及个体间生理差异使得信号中有效肌电特征与噪声在时频域高度耦合<sup>[19]</sup>,导致模型在跨场景下的稳定性和泛化能力受限。传统滤波等去噪方法在复杂应用场景下往往难以兼顾噪声抑制与关键信息保留。此外,预处理方法通常在理想的实验环境中设计,而在实际应用中噪声特性难以预测,导致模型鲁棒性不足。

为解决这一难题,近年来研究者开始探索基于单一模态的数据增强策略,通过针对性扰动提升模型对特定噪声或运动变异的鲁棒性。在噪声增强方面,Krell 等<sup>[20]</sup>研究了通过改变 EEG 头帽的电极位置来生成对应于空间畸变的增强数据的可能性。Fan 等<sup>[21]</sup>对 Myo 胳膊带的八个电极进行圆周旋转,通过改变信号空间耦合模式增强模型对肢体位姿变化的适应性,最终将智能手机解锁系统的识别准确率提升 12.6%。Tsinganos 等<sup>[22]</sup>通过加性噪声和小波分解模拟传感器噪声以及变换提取多尺度特征,通过振幅变换和 sEMG 信号仿真扩展数据集。其中,振幅变换和小波分解方法在 Ninapro DB1 和 putEMG 两个数据集提升了高达 16% 的分类精度。Mathur 等<sup>[23]</sup>通过在 EEG 信号中加入白噪声并调节 SNR 进行验证,结果表明 GSP 方法在噪声环境下优于现有的心理任务 EEG 分类方法。Oh<sup>[24]</sup>针对中风偏瘫患者动作分类实验提出了基于时间序列变换的数据增强策略。研究重点测试了旋转变换、排列变换和时间扭曲 3 种增强方法。其中,旋转变换通过空间几何扰动模拟偏瘫患者运动轨迹幅度波动与方向偏移,使得 F1 分数提升了 13.6%,为高噪声、小样本场景下的康复动作精准识别提供了新思路。Ma 等<sup>[25]</sup>通过对 NinaPro DB1 中 23 类动作加入高斯噪声以模拟交警手势在真实环境下的噪声干扰,结果表明 FFCSLT 模型使分类准确率提升了 0.75%,验证了其在噪声环境下的有效性。

上述研究表明,单一增强策略通过针对性扰动特定信号维度可显著提升模型在特定挑战下的性能。然而,现有工作仍存在两点局限:其一,多数模型依赖浅层卷积网络提取特征,难以捕获 sEMG 信号的多尺度时空依赖性;其二,增强策略与模型架构未能协同优化,传统神经网络对噪声敏感的特征融合机制易造成高频干扰传播。

基于上述背景,本文提出了一种融合空洞卷积和注意力机制的抗噪 sEMG 手势识别模型,通过多尺度特征提取和噪声增强训练策略,有效缓解了信号稀疏性和噪声干扰带来的挑战,同时在复杂环境下实现了更高的泛化能力和识别准确率<sup>[26]</sup>。

本文的其余部分组织如下:第 1 节介绍原始 sEMG 数据的预处理方法;第 2 节详细阐述了所提出的 RASTNet 模型;第 3 节通过消融实验等验证了模型各模块的有效性;第 4 节总结了本文的工作成果并讨论了未来的研究方向。

## 1 方 法

该实验服务器配置为英特尔(R)酷睿(TM)i9-12900 CPU@ 2.40 GHz, GPU 为 NVIDIA TITAN RTX 3090。所提出的方法由 pytorch 1.13.1 框架结合 Python 3.8 实现。这项研究利用了基准 Ninapro 数据集,该数据集可从官方 Ninapro 存储库<sup>[27]</sup>访问。

### 1.1 数据集

Ninapro 是一个广泛使用的表面肌电手势识别公共数据集,提供丰富的人体生理数据,广泛应用于机器学习、机器人控制、医学和神经科学等领域。本文选择了 Ninapro DB1 和 DB5 数据集,涵盖了图 1 中的 A、B、C 三种练习类型,共包含 52 种手势,具体分为手指运动、手运动、基本手腕运动以及抓握和功能性运动四大类。

### 1.2 数据预处理

如图 2 所示。首先,对两个数据集进行数据增强,然后用滑动窗口进行切片生成肌电图像<sup>[28]</sup>。最后,将两个数据集的数据打乱,按照 8:2 的比例划分为训练集和测试集。

本文利用重叠滑动窗口技术<sup>[29]</sup>将原始信号划分为序列,分割过程如图 3 所示。大多数研究表明,最佳的窗口长度在 100~300 ms 之间。基于此,采用 200 ms 的窗口长度和 50 ms 的滑动长度来实现实时控制。在这种情况下,sEMG 被转换为图像,其中通道的数量对应图像的宽度,采样点的数量对应图像的长度。

在数据预处理中,为了优化模型的训练效果,本文丢弃了休息标签数据。此外,两个数据集在数据结构和信号特性上也存在一定差异。通过一系列实验,筛选出最优的超参数设置进行了数据增强。如表 1 所示,分别在 DB1 和 DB5 数据集中加入了高斯噪声、脉冲噪声以及进行时序偏移和振幅变换,以便全面评估模型在不同噪声条件下的鲁棒性。两个数据集的参数差异如表 2 所示。

高斯噪声模拟了传感器固有的随机电气噪声、外部电

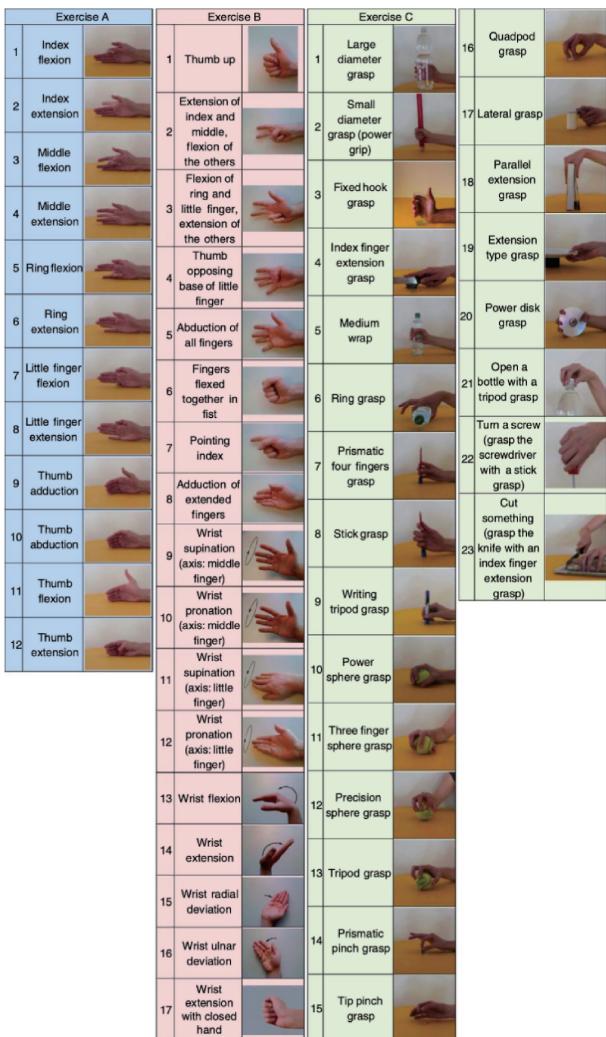


图 1 NInapro 数据集

Fig. 1 NInapro dataset

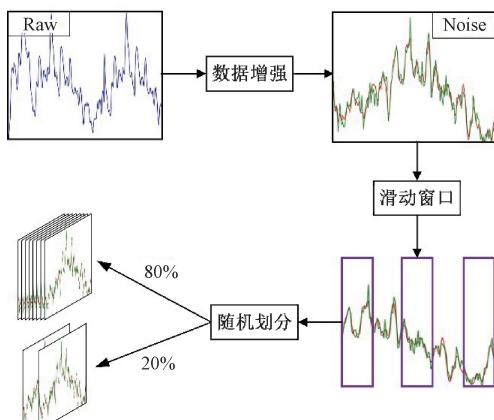


图 2 数据预处理

Fig. 2 Data pre-processing

磁干扰以及采集过程中的环境噪声,呈现为持续性且服从高斯分布的随机波动;脉冲噪声则模拟了突发的肌肉反应

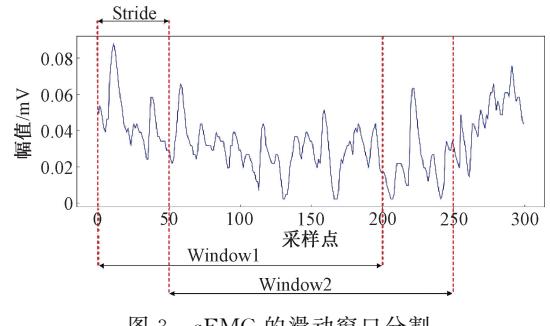


图 3 sEMG 的滑动窗口分割

Fig. 3 Sliding window segmentation of the sEMG

表 1 不同噪声水平

Table 1 Different noise levels

Dataset	Noise Type	Description
DB1	Gaussian	$m=0, v=0.05$
	Impulse	$P=0.01, a=0.8$
	Time Shifting	25
	Amplitude Scaling	0.8~1.2
DB5	Gaussian	$m=0, v=4$
	Impulse	$P=0.01, a=30$
	Time Shifting	20
	Amplitude Scaling	0.8~1.2

表 2 两个数据集的统计信息

Table 2 Statistical information of the two datasets

Dataset	DB1	DB5
Total	105 020	47 458
Training	84 016	37 966
Testing	21 004	9 492
gestures	52	52
channels	10	16
format	(200×10)	(200×16)
Hz	100	200

或电极接触不良引起的瞬时信号干扰,表现为信号中的突发性异常波动或尖峰;时序偏移模拟了肌电信号在采集过程中由于电极位置变化或传输延迟所引起的时间偏移,使模型能够学习对齐误差带来的时间不稳定性;振幅变换旨在模拟不同个体肌肉收缩强度的变化,以及由于传感器接触压力或皮肤电导率差异导致的信号幅值波动,从而增强模型对不同幅值尺度的适应能力。

定义高斯随机变量  $z$  的概率密度函数  $P(z)$  如下:

$$P(z) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(z-\mu)^2}{2\sigma^2}} \quad (1)$$

式中:  $\mu$  是数学期望,  $\sigma$  是标准差。

定义脉冲变量  $I$  的概率密度函数  $P(I)$  如下:

$$P(I_{ij} = k) = \begin{cases} 1-p, & k = 0 \\ \frac{p}{2}, & k = a \\ \frac{p}{2}, & k = -a \end{cases} \quad (2)$$

式中:  $p$  表示脉冲出现的概率,  $a$  表示脉冲的幅度。

定义时序偏移变量  $x$  的变换函数  $T(x)$  如下:

$$T(x) = x(t - \Delta t) \quad (3)$$

式中:  $x(t)$  为原始信号,  $\Delta t$  为时间偏移量。

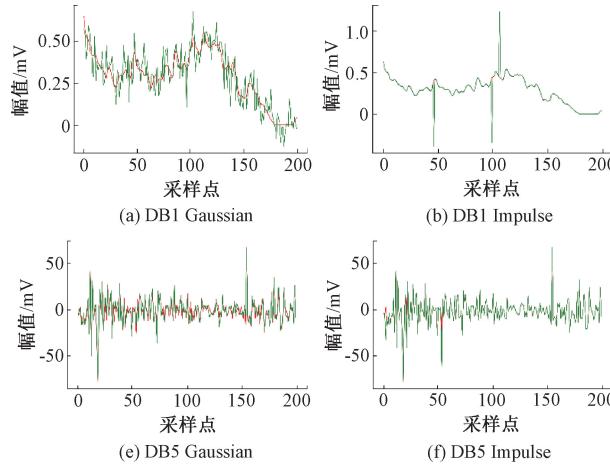


图 4 加噪音后的表面肌电图

Fig. 4 sEMG signals after adding noise

## 2 模型

为了有效提升通道与空间特征的提取能力, RASTNet 在经典的 ResNet50 架构基础上引入了空洞空间金字塔池化模块(atrous spatial pyramid pooling, ASPP)和 STConv 模块。RASTNet 的总体结构如图 5 所示, 网络被划分为分为 6 个部分。

在模型的第 1 部分(Layer 0), 输入的是 sEMG 数据。输入首先经过一个  $7 \times 7$  的大核卷积, 紧接着进行批量归一化(batch normalization, BN)、激活函数 ReLU 以及最大池化层, 以降低特征图的尺寸, 减少计算量并保留最具代表性的特征。

第 2~5 部分(Layer1~Layer4)是由不同的残差块(Conv Block, Block, AST Block)组成。每一部分的第一个和最后一个残差块是 Conv Block 和 AST Block, 中间的 Block 数量分别为 1、2、4、1。其中 Conv Block 采用了由三层卷积组成的残差结构, 每层卷积后都伴随着 BN 层和 ReLU 激活函数。这些卷积操作分别是:  $1 \times 1$  卷积(降维),  $3 \times 3$  卷积(特征提取),  $1 \times 1$  卷积(升维)。此外, Conv Block 还有一个跳跃连接, 通过  $1 \times 1$  的卷积调整输入的通道数, 使其与主路径的输出尺寸一致, 然后将两者相加。Block 与 Conv Block 不同, 它用于处理输入输出维度相同的情况。因此, Block 的跳跃连接不

需要进行额外的卷积操作, 直接将输入与输出相加。

AST Block 在 Block 的基础上将  $3 \times 3$  卷积替换为 ASPP 并加入了由 SCConv<sup>[30]</sup> 和 Triplet Attention 机制组成的 STConv。

最后一部分(Layer 5)则是对得到的特征图进行平均池化计算, 再经过全连接层得到最终的分类结果。

### 1) ASPP 模块

在本文所构建的模型中, 创新性地将原本的一个  $3 \times 3$  普通卷积替换为图 6 所示的 ASPP 模块。

$3 \times 3$  普通卷积在特征提取时, 其感受野相对固定, 仅能在单一尺度上捕捉输入特征图的信息, 对于 sEMG 这类复杂且蕴含多尺度肌肉活动特征的数据源而言, 局限性明显。

而引入的 ASPP 模块通过采用不同扩张率(rate)的空洞卷积操作来捕捉多尺度的上下文信息。在该模块中, 存在多个并行分支, 其中有不同扩张率(如 rate=1、rate=2、rate=4)的  $3 \times 3$  空洞卷积, 当扩张率为 1 时, 其作用类似于普通卷积, 能精准捕捉 sEMG 的局部细节; 而扩张率为 2 和 4 时, 随着感受野扩大, 可将提取更大空间范围内的特征。

此外, ASPP 模块中的  $1 \times 1$  卷积分支可灵活调整通道数, 确保各分支特征在融合前维度适配, 避免信息丢失, 保障后续高效地特征融合, 全局平均池化分支还能获取全局

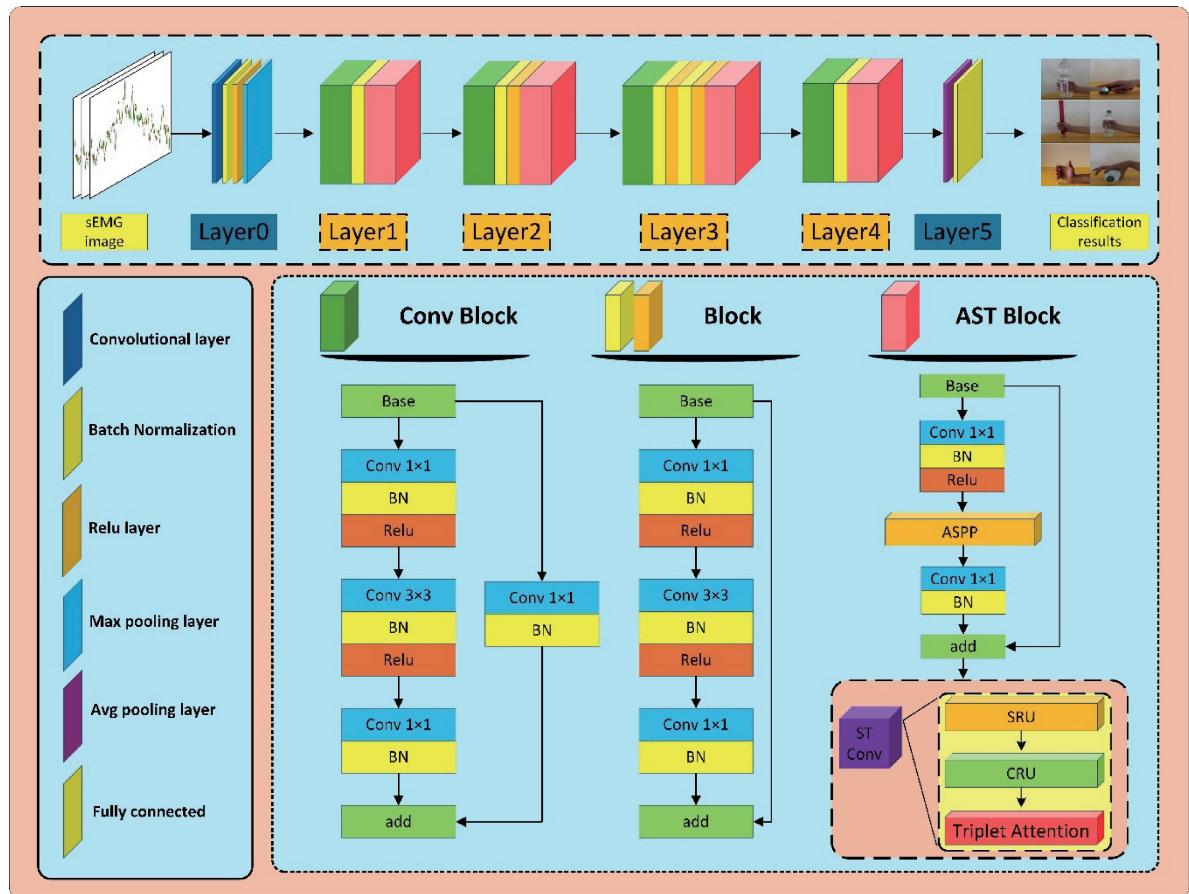


图 5 RASTNet 框架结构

Fig. 5 Framework of RASTNet

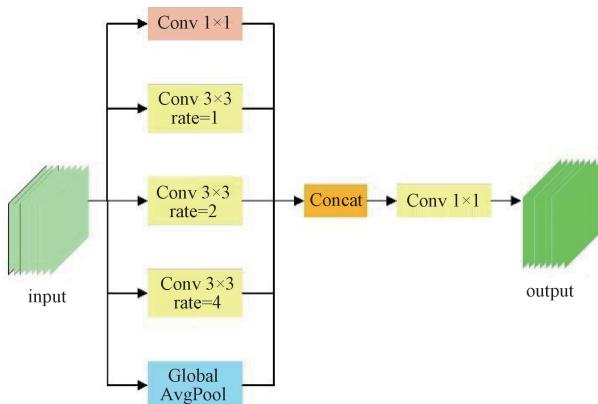


图 6 ASPP 的体系结构

Fig. 6 The architecture of ASPP

上下文信息,让模型能从宏观角度整体把握 sEMG 信号的态势。

最后,各分支特征图经拼接、 $1 \times 1$  卷积得到输出特征图。

将  $3 \times 3$  普通卷积替换为 ASPP 模块,使多尺度信息捕捉能力克服了普通卷积的单尺度局限,丰富了 sEMG 特征的多样性。

## 2) 空间重构单元

如图 7 所示,空间重构单元(SRU)包含分离和重构两部分。分离是将空间内 sEMG 信息丰富的特征图与 sEMG 信息少的特征图分离。重构是将两种特征图交叉重构相加,生成更丰富的 sEMG 信息特征图。

首先,对输入的 sEMG 特征进行标准化,具体如下:

$$X_{out} = GN(X) = \gamma \frac{X - \mu}{\sqrt{\sigma^2 + \epsilon}} + \beta \quad (5)$$

其中,GN 为组归一化,  $\gamma$  和  $\beta$  为可训练参数;  $\mu$  和  $\sigma$  分别表示输入特征的均值和标准差;  $\epsilon$  为保证除法稳定性的小的正常数。 $\gamma$  的大小反映了各个通道中的空间像素和空间信息的丰富程度。通过以下公式,可以得到归一化的各个通道的表明不同 sEMG 特征图重要性的权重。

$$W_\gamma = \{w_i\} = \frac{\gamma_i}{\sum_{j=1}^C \gamma_j}, i, j = 1, 2, \dots, C \quad (6)$$

然后,经过权重重加权后的 sEMG 特征图权重  $W_\gamma$  通过 sigmoid 函数映射到(0,1)范围内,并通过一个阈值进行筛选。将大于阈值的权重设为 1, 得到的信息权重为  $W_1$ , 小于阈值的权重设为 0, 得到的非信息权重为  $W_2$ 。随后分

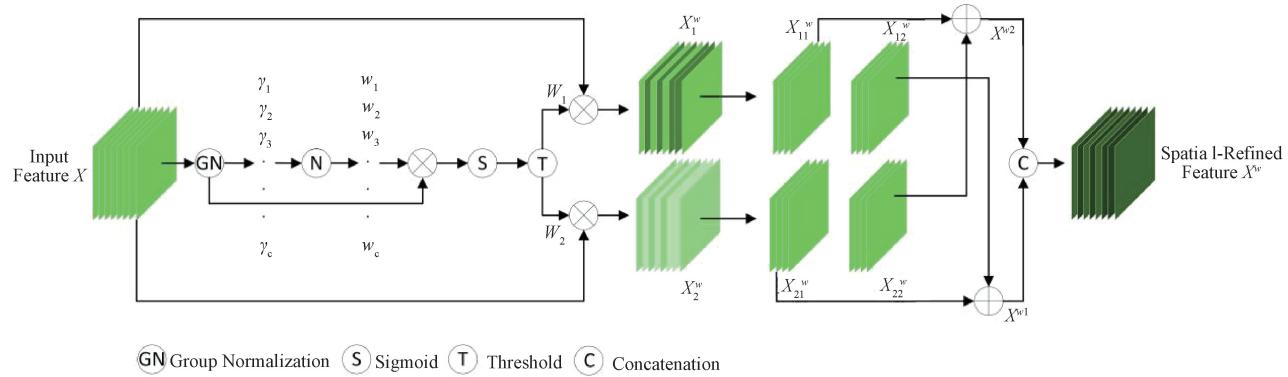


图 7 空间重构单元的体系结构  
Fig. 7 The architecture of spatial reconstruction unit

别将输入特征  $X$  乘以  $W_1$  和  $W_2$ , 得到两个加权特征: 信息丰富的特征  $X^w_1$  和信息较少的特征  $X^w_2$ 。

最后, 进行重构操作, 将信息丰富的特征与信息较少的特征交叉重构相加, 以充分结合这两类加权的不同信息特征, 并增强它们之间的信息流动。

通过分离和重构策略, SRU 能够有效地从噪声中分离出有价值的 sEMG 特征, 减少了无关信息的干扰。

### 3) 通道重构单元

如图 8 所示, 为了减少通道维度冗余, 通道重构单元 (CRU) 采用拆分-转换-融合策略。

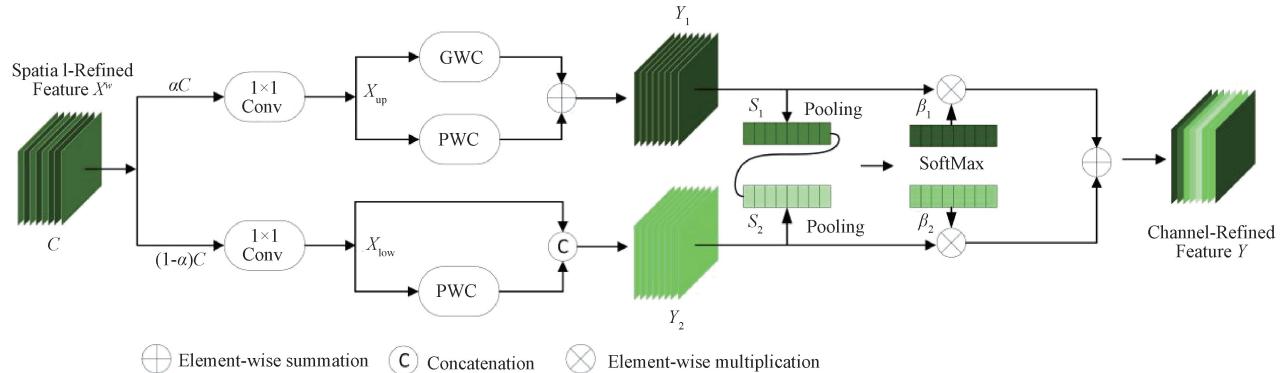


图 8 通道重构单元的体系结构  
Fig. 8 The architecture of channel reconstruction unit

拆分, 首先将通道根据拆分比例  $\alpha$  分为两部分, 分别经过  $1 \times 1$  卷积来压缩 sEMG 特征图的通道, 以提高计算效率。然后引入压缩比例  $r$  来控制特征通道。最后将空间特征  $X$  划分为上部分  $X_{up}$  和下部分  $X_{low}$ 。

转换,  $X_{up}$  作为“丰富特征提取器”被输入到上转换阶段。本文采用组卷积(group-wise convolution, GWC)和点卷积(point-wise convolution, PWC)操作替代标准卷积, 以提取高层次的代表性信息并降低计算成本。由于稀疏卷积连接, GWC 减少了参数和计算量, 但会切断通道组之间的信息流。而 PWC 补偿了信息的丢失, 并帮助信息在 sEMG 不同特征通道之间流动。在执行 GWC 和 PWC 之后, 将输出相加, 得到  $Y_1$ 。 $X_{low}$  被输入到下转换阶段, 将重用后的  $X_{low}$  特征图和经过了 PWC 之后的特征图拼接起来, 得到  $Y_2$ 。

融合, 首先将  $Y_1$  和  $Y_2$  经过全局平均大池化(global average pooling, GAP)之后使用通道自注意力操作生成特

征重要性向量  $\beta_1$  和  $\beta_2$ 。然后再将  $Y_1$  和  $Y_2$  分别乘以  $\beta_1$  和  $\beta_2$  后相加得到  $Y$ 。

通过拆分-转换-融合策略, CRU 能够有效地减少 sEMG 中通道的冗余信息, 同时增强通道之间的信息流动, 提取出更具代表性且高效的特征, 从而提升手势识别模型的性能。

### 4) Triplet Attention 模块

传统注意力机制通常在通道维度和空间维度上独立计算注意力, 并没有考虑它们之间的关系。如图 9 所示, Triplet Attention 模块通过引入 3 个分支来分别捕获输入张量的  $(C, H)$ 、 $(C, W)$  和  $(H, W)$  维数之间的依赖关系, 通过捕获空间维和输入张量的通道维之间的相互作用来解决这一缺点。

第一个分支负责处理输入张量的通道  $C$  和空间高度维度  $H$  的交互。在 sEMG 处理中, 空间高度维度 ( $H$ ) 通常代表信号随时间的变化。首先将输入张量  $x \in R^{W \times H \times C}$

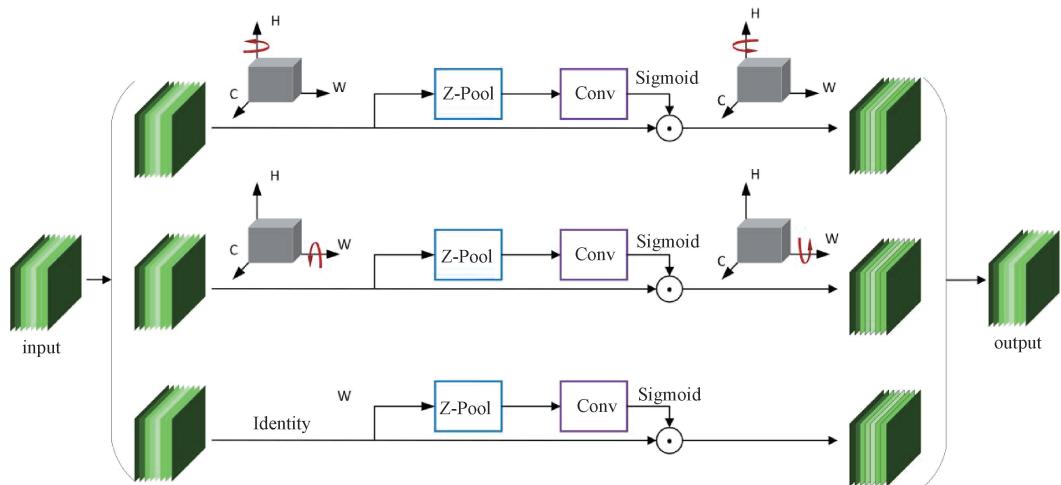


图 9 三重注意力模块的体系结构

Fig. 9 The architecture of triple attention module

沿着  $H$  轴逆时针旋转  $90^\circ$  得到  $x_1$ , 其形状为  $(W \times H \times C)$ , 并对其应用最大池化和平均池化后变为  $(2 \times H \times C)$  的  $x_2$ 。这种池化操作称为 Z-pool, 此过程有助于保留 sEMG 中的关键特征, 同时减少冗余信息, 具体公式如下。随后, 将  $x_2$  通过卷积和批归一化层, 得到形状为  $(1 \times H \times C)$  的中间输出张量。最后, 通过 Sigmoid 激活函数生成注意力权重应用于  $x_1$ , 并沿  $H$  轴顺时针旋转  $90^\circ$  以恢复到原始输入形状  $x$ 。

$$Z\_pool(x) = [MaxPool_{od}(x), AvgPool_{od}(x)] \quad (7)$$

其中,  $od$  表示维度为 0。

第二分支负责处理通道  $C$  和空间宽度维度  $W$  的交互。在 sEMG 的空间宽度维度 ( $W$ ) 上, 通常包含了不同肌肉群的活动信息。通过将输入张量沿  $W$  轴逆时针旋转  $90^\circ$ , 该分支能够有效捕捉肌肉群之间的复杂关系。

第三分支类似于 CBAM 中的空间注意力机制, 专注于捕捉输入张量在空间维度 ( $H \times W$ ) 上的交互。它帮助模型关注 sEMG 局部区域的细节特征, 尤其是在时间序列和空间维度上变化的动态模式。

最终, 三条分支的输出张量通过平均操作聚合生成带有特征融合增强的输出。总结来说, Triplet Attention 通过多层次、全方位地关注 SCCConv 输出的 sEMG 的空间和通道信息, 使模型在面对噪声和干扰时能够更加稳定地提取有价值的特征。对于输入张量  $x \in R^{W \times H \times C}$  的精炼过程可以表示为以下公式:

$$y = \frac{1}{3}(y_1 + y_2 + y_3) \quad (8)$$

$$y_1 = \overline{x_1 \sigma(\gamma_1(x_2))} \quad (9)$$

$$y_2 = \overline{x_a \sigma(\gamma_2(x_b))} \quad (10)$$

$$y_3 = \overline{x_A \sigma(\gamma_3(x_B))} \quad (11)$$

其中,  $y_1, y_2, y_3$  分别表示上述的 3 个注意力分支,  $\sigma$

表示 sigmoid 激活函数,  $\gamma_1, \gamma_2, \gamma_3$  表示三重注意力的 3 个分支中标准卷积层和归一化层输出结果。

### 3 实验和结果

#### 3.1 参数设置

经过大量实验和参数调优, 设置了一组高效的超参数以优化模型性能。训练批量大小为 64, 以实现计算稳定性和高效的 GPU 利用。最大迭代次数设为 100 次, 确保模型有充足的训练时间, 同时有效避免过拟合。选用具有动态学习率调整能力的 Adam 优化器, 以其高效的梯度更新机制加速模型收敛。初始学习率设定为 0.001, 这是经过多次实验验证的有效起点, 有助于模型在早期阶段快速收敛。此外, 引入了学习率递减策略, 每 20 个 epoch 后学习率按 0.1 倍衰减。这一策略不仅帮助模型在初期快速找到最优路径, 还在后期精细化参数调整, 从而显著提升模型的收敛稳定性和泛化能力。

#### 3.2 评价标准

为了全面衡量 RASTNet 的有效性, 实验采用准确率 (accuracy, ACC)、精确率 (precision, P)、召回率 (recall, R)、特异度 (specificity, S) 和 F1 分数 (F1 score, F1) 作为评价指标:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \quad (12)$$

$$P = \frac{TP}{TP + FP} \quad (13)$$

$$F1 = 2 \times \frac{P \times R}{P + R} \quad (14)$$

$$S = \frac{TN}{TN + FP} \quad (15)$$

$$F1 = 2 \times \frac{P \times R}{P + R} \quad (16)$$

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。

### 3.3 不同噪声条件下的实验结果

如表 3 和图 10 所示, RASTNet 在各种噪声条件下的

准确率、精确率、召回率、特异性和 F1 分数均显著优于 ResNet50。

具体来说,在 DB1 数据集上准确率平均提升 1.83%, 其中在脉冲噪声条件下, 提升最为显著, 达到 2.41%。在时序偏移条件下, RASTNet 成功地维持了高水平的特异

表 3 不同噪声条件下的实验结果

Table 3 Experimental results under different noise conditions

Dataset	Model	Noise Type	Acc/%	Precision/%	Recall/%	Specificity/%	F1/%
DB1	ResNet50	Gaussian	95.18	95.21	95.20	99.91	95.19
		Impulse	94.30	94.32	94.30	99.89	94.30
	RASTNet	Time Shifting	96.86	96.89	96.86	99.95	96.87
		Amplitude Scaling	96.49	96.49	96.49	99.94	96.49
DB5	ResNet50	Gaussian	97.15	97.00	96.99	99.94	96.98
		Impulse	96.71	96.69	96.66	99.94	96.67
		Time Shifting	98.32	98.23	98.22	99.97	98.22
		Amplitude Scaling	97.96	97.76	97.78	99.96	97.76
	RASTNet	Gaussian	90.64	90.76	90.63	99.82	90.66
		Impulse	93.66	93.68	93.65	99.88	93.66
		Time Shifting	96.36	96.39	96.40	99.93	96.39
		Amplitude Scaling	96.44	96.46	96.48	99.93	96.46

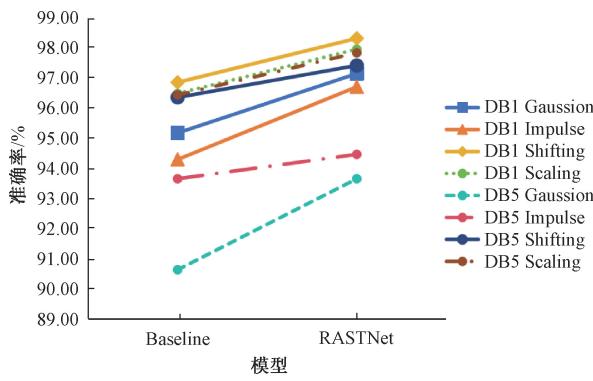


图 10 模型性能差异

Fig. 10 Performance differences of the model

性, 这表明其在识别过程中能够有效过滤噪声干扰, 从而降低误判的发生概率。在 DB5 数据集上的准确率平均提升 1.57%, 并在高斯噪声下表现出最大提升, 达到 3.02%。即使在脉冲噪声条件下提升幅度较小, 但在精确率和 F1 分数上的优化同样体现了模型在复杂噪声下的鲁棒性。这种能力确保了模型在显著背景噪声的环境中, 依然能够精准地识别出真实信号。此外, 各噪声条件下召回率的提升进一步表明, 在遭遇短暂且剧烈的信号干扰时, RASTNet 仍能可靠地捕捉到正确的手势信号, 这反映了

其在复杂环境中的鲁棒性和可靠性。

整体而言, RASTNet 在噪声条件下的卓越表现源于其引入的 ASPP 模块和 STConv 模块, 构建了一种高度协调的特征提取机制。确保在强噪声干扰下, 模型仍能专注于有价值的信号区域。这种多层次、多维度特征聚焦能力, 使得模型在面对各中噪声时, 能够保持稳定的特征提取能力, 从而使 RASTNet 在性能表现上显著优于 ResNet50。

### 3.4 对比实验

为了进一步验证本文设计的 RASTNet 模型在基于 SEMG 手势识别领域的优越性, 将其与例如 ResNeXt、Swin Transformer 等主流经典模型和近年其他相关模型进行了对比实验。

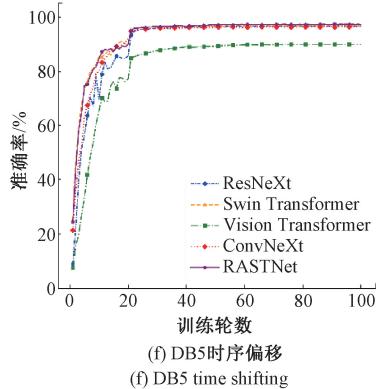
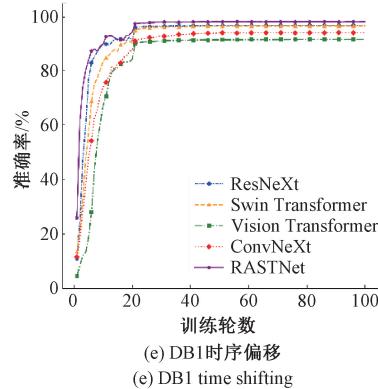
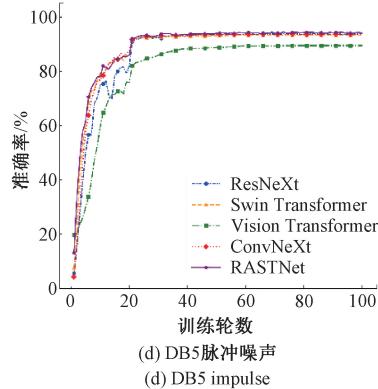
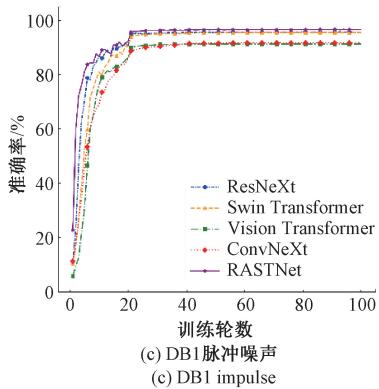
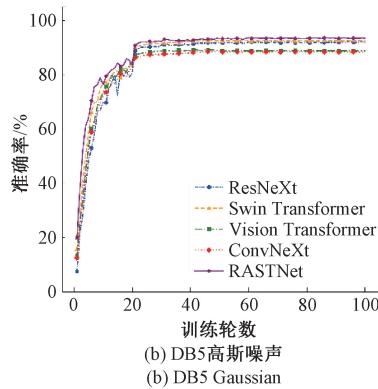
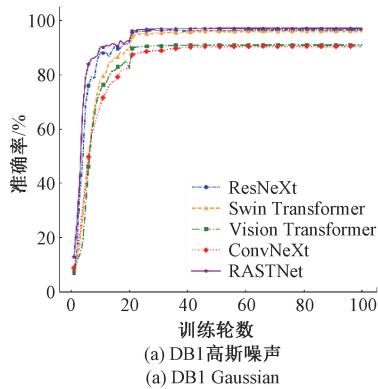
实验结果如表 4 和图 11 所示, 在 DB1 数据集的 4 种噪声条件下, RASTNet 均展现出显著性能优势。面对高斯噪声, 其准确率达 97.15%, 较 ResNeXt 与 Swin Transformer 分别提升 0.62% 和 1.20%, 且 F1 分数与精准率偏差仅为 0.02%, 验证了其对高频噪声的强鲁棒性。在脉冲噪声下, RASTNet 以 96.71% 的准确率领先 ConvNeXt 达 4.95%, 且特异性 99.94% 为所有模型最高, 体现突发噪声过滤能力。针对时序偏移, RASTNet 准确

表 4 不同方法的分类结果

Table 4 Classification results of different methods

%

Noise	Model	DB1	DB5
		Acc/Precision/Recall/Specificity/F1	Acc/Precision/Recall/Specificity/F1
Gaussian	ResNeXt	96.53/96.56/96.54/99.93/96.55	92.36/92.45/92.38/99.85/92.39
	Vision Transformer	90.98/91.01/90.99/99.82/90.99	89.22/89.25/89.24/99.79/89.27
	Swin Transformer	95.95/95.98/95.95/99.92/95.95	92.73/92.82/92.76/99.86/92.78
	ConvNeXt	90.44/90.54/90.48/99.81/90.49	88.49/88.63/88.53/99.77/88.54
	RASTNet	<b>97.15/97.00/96.99/99.94/96.98</b>	<b>93.66/93.62/93.70/99.88/93.63</b>
Impulse	ResNeXt	95.82/95.85/95.83/99.92/95.84	94.04/94.11/93.94/99.88/93.99
	Vision Transformer	91.31/91.33/91.30/99.83/91.31	89.46/89.55/89.39/99.79/89.38
	Swin Transformer	95.54/95.55/95.54/99.91/95.54	93.32/93.32/93.28/99.86/93.28
	ConvNeXt	91.76/91.82/91.78/99.84/91.78	93.69/93.70/93.65/99.87/93.66
	RASTNet	<b>96.71/96.69/96.66/99.94/96.67</b>	<b>94.47/94.50/94.43/99.89/94.43</b>
Time Shifting	ResNeXt	96.98/97.00/96.99/99.95/96.99	96.94/96.93/96.93/99.94/96.91
	Vision Transformer	91.71/91.72/91.73/99.84/91.71	89.94/89.97/89.86/99.80/89.88
	Swin Transformer	96.73/96.74/96.74/99.94/96.74	96.88/96.90/96.89/99.94/96.89
	ConvNeXt	94.24/94.28/94.26/99.89/94.26	96.30/96.30/96.28/99.93/96.28
	RASTNet	<b>98.32/98.23/98.22/99.97/98.22</b>	<b>97.42/97.42/97.39/99.95/97.39</b>
Amplitude Scaling	ResNeXt	96.52/96.52/96.52/99.95/96.52	97.20/97.23/97.24/99.95/97.23
	Vision Transformer	92.83/92.84/92.84/99.86/92.83	88.53/88.62/88.47/99.77/88.47
	Swin Transformer	96.98/96.98/96.98/99.96/96.98	94.65/94.68/94.65/99.89/94.65
	ConvNeXt	93.89/93.92/93.92/99.88/93.90	96.62/96.60/96.62/99.93/96.62
	RASTNet	<b>97.96/97.76/97.78/99.96/97.76</b>	<b>97.84/97.86/97.88/99.96/97.86</b>



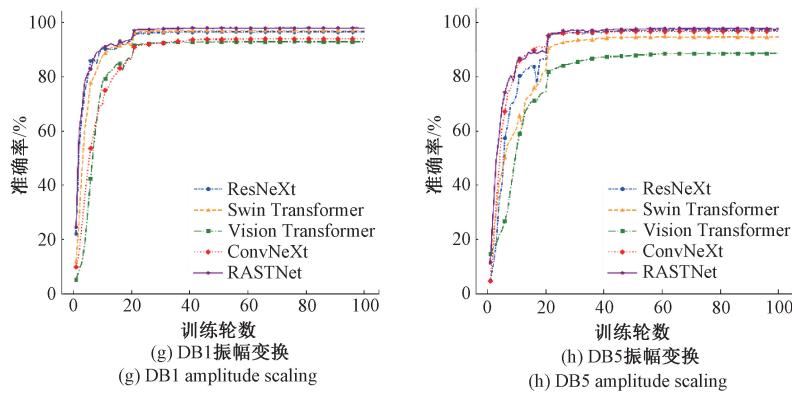


图 11 不同噪声条件下的对比实验

Fig. 11 Comparison experiments under different noise conditions

率为 98.32%，较 ResNeXt 优势达 1.34%，召回率与精准率偏差仅 0.01%，表明其时空特征融合策略可有效缓解信号时序偏移问题。对于振幅变换，RASTNet 以 97.96% 的准确率超越 Swin Transformer，F1 分数较 ConvNeXt 提升 3.86%。

在 DB5 数据集中，RASTNet 的优势进一步扩大。高斯噪声下其准确率为 93.66%，较 ResNeXt 提升 1.30%，且 F1 分数显著优于 Swin Transformer。在脉冲噪声条件下，RASTNet 以 94.47% 的准确率领先 ConvNeXt 与 Swin Transformer 分别达 0.78% 和 1.15%，召回率与精准率偏差小于 0.07%，远优于 Vision Transformer 的 0.17%。对于时序偏移，RASTNet 准确率为 97.42%，较 Swin

Transformer 提升 0.54%，且 F1 分数为同类最高。面对振幅变换，RASTNet 以 97.84% 的准确率全面领先，较 Swin Transformer 提升 3.19%，F1 分数与精准率完全一致，体现其分类决策的高置信度。

从图 11 中可以看出，RASTNet 准确率上升速度较快，能更早地达到较高水平。同时，RASTNet 在训练过程中表现出较好的稳定性，准确率曲线平稳且波动较小。总体而言，RASTNet 在准确率、收敛速度和稳定性方面均优于 ResNeXt、Swin Transformer 等主流模型。

从表 5 的实验结果来看，RASTNet 在 DB1 数据集的 4 种噪声条件中最低的 96.71% 准确率仍显著优于 EMGHandNet、双流模型以及 Transformer 等模型，进一

表 5 不同方法的分类结果

Table 5 Classification results of different methods

Datasets	Author	Year	Class	Feature extraction	classifier	ACC/%
Ninapro DB1	Pancholi et al. <sup>[31]</sup>	2021	52	—	DLRP	91.11
	Wei et al. <sup>[32]</sup>	2021	52	—	HVPN	87.10
	Karnam et al. <sup>[33]</sup>	2022	52	CNN+Bi-LSTM	EMGHandNet	95.77
	Krishnapriya et al. <sup>[34]</sup>	2023	52	—	1D CNN	78.95
	Wang et al. <sup>[35]</sup>	2023	52	MAV, ZC, SSC, WL, and RMS	Transformer	89.43
	Zhang et al. <sup>[36]</sup>	2024	52	MAV, ZC, SSC, WL, and RMS	dual-stream model	89.66
	Shin et al. <sup>[37]</sup>	2024	52	—	4-stream DL	94.31
	Lee et al. <sup>[38]</sup>	2024	53	MAV, MAV slope 等 11 个特征	Prototype learning method	96.25
	ours	—	<b>52</b>	—	RASTNet	<b>96.71</b>
Ninapro DB5	Wei et al. <sup>[32]</sup>	2021	41	—	HVPN	88.40
	Peng et al. <sup>[39]</sup>	2022	52	MAV, RMS, VAR 等 16 个特征	EELM	77.90
	Xiong et al. <sup>[40]</sup>	2023	53	—	GLF-CNN	91.40
	Niu et al. <sup>[41]</sup>	2024	53	—	PCS-EMGNet	88.30
	Wang et al. <sup>[42]</sup>	2024	52	RMS, WL, ZC, SSC	CDEM	84.00
	Mohapatra et al. <sup>[43]</sup>	2024	52	WL, MAV, SD, RMS	SVM	84.00
	Zhang et al. <sup>[44]</sup>	2024	52	Mutil-attention	softmax	91.64
	ours	—	<b>52</b>	—	RASTNet	<b>93.66</b>

步证明了其在 sEMG 手势识别中的突出表现。在 DB5 数据集上,RASTNet 在高斯噪声条件下 93.66% 的准确率明显领先于 GLF-CNN 等传统模型。这一对比表明,即使面对噪声干扰,RASTNet 的性能下限仍超越现有方法的性能上限,印证了其在复杂手势识别任务中的强鲁棒性。

综上所述,RASTNet 在处理 sEMG 时,能够更有效地提取通道和空间特征。这主要得益于其引入 ASPP 和 STConv,改进后的模型实现了对多尺度信息的动态捕捉和特征之间的充分交互,从而提升了手势识别的准确性和泛化能力。

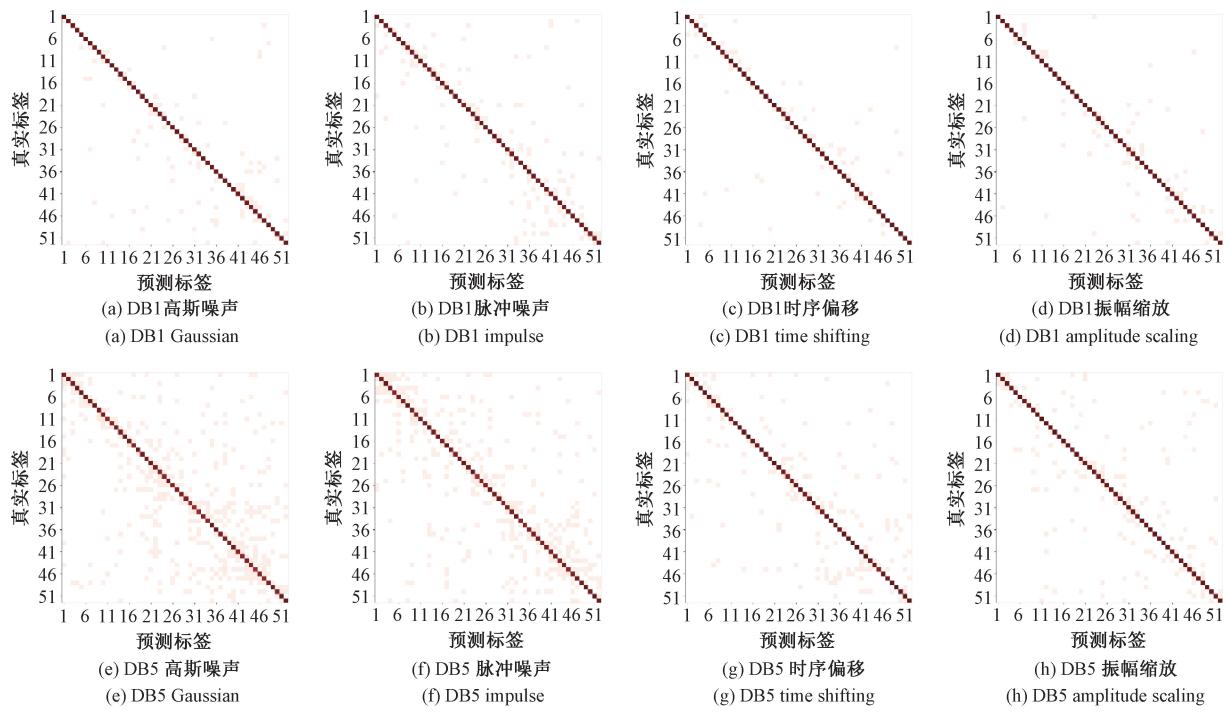


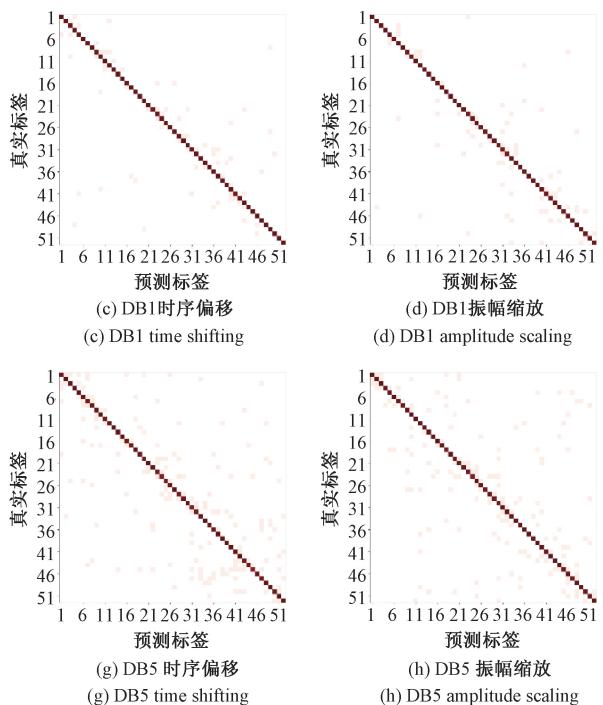
图 12 RASTNet 的混淆矩阵

Fig. 12 Confusion matrix of RASTNet

具体来说,在 DB1 各种噪声条件下,手势 32(固定钩爪)与手势 31(强力握把)以及手势 43(棱镜夹持器)与手势 44(尖端捏抓)容易混淆。在 DB5 各种噪声条件下,手势 44(尖端捏抓)、手势 32(固定钩爪)和手势 23(手腕内翻)均表现出一致的误分类趋势,手势 43(棱镜夹持器)和手势 32(固定钩爪)频繁误分类为手势 44(尖端捏抓)和手势 31(强力握把)。其中在时序偏移条件下,手势 45(四脚抓握)会误分类为手势 44(尖端捏抓)和手势 43(棱镜夹持器)。以上现象的出现可能与某些手势在形态和动作上非常相似有关,因此它们的特征在噪声条件下更容易被混淆。例如,手势 32(固定钩爪)与手势 31(强力握把)都涉及到手指的弯曲和握紧,手势 43(棱镜夹持器)与手势 44(尖端捏抓)也可能表现出相似的手指位置和运动轨迹,从而导致分类误差的发生。该分析证明了受试者内手势分类的总体高

### 3.5 RASTNet 混淆矩阵

如图 12 所示,由本文的方法生成的混淆矩阵全面展示了模型在各种噪声条件下的数据集预测结果和真实手势类别之间的对应关系。图中附带的渐变彩条用于标识样本数量,颜色越深代表真实标签与预测标签组合对应的数量越多,反之数量越少。整体来看,混淆矩阵的对角线模式表明模型在大多数手势类别上的分类表现较为准确。然而,细致分析后可以发现,某些特定手势较其他手势更容易引起分类混淆,给模型带来了更大的识别挑战。



准确性,也揭示了模型在复杂的噪声环境下对于个别手势分类的局限性,为后续的模型优化提供了方向和依据。

### 3.6 拟真噪声条件下的深度实验

在前述多噪声研究基础上,本章节聚焦高斯噪声。一方面,高斯噪声的正态分布特性能够逼真地模拟环境中的噪声;另一方面,高斯噪声带来的“适度”挑战有利于验证 RASTNet 模型所引入模块的有效性。随后,将围绕高斯噪声,用对比、消融等方法深挖模型在 DB1、DB5 数据集的表现潜力。

#### 1) 不同手势分类上的表现差异

为了更直观地展示 Baseline 与 RASTNet 模型在各类手势上的性能差异,本文设计了堆积柱状图。图 13 中阴影柱状表示原模型对每个手势的准确率,斜线柱状则展示了 RASTNet 模型相较于原模型的准确率变化。通过这种

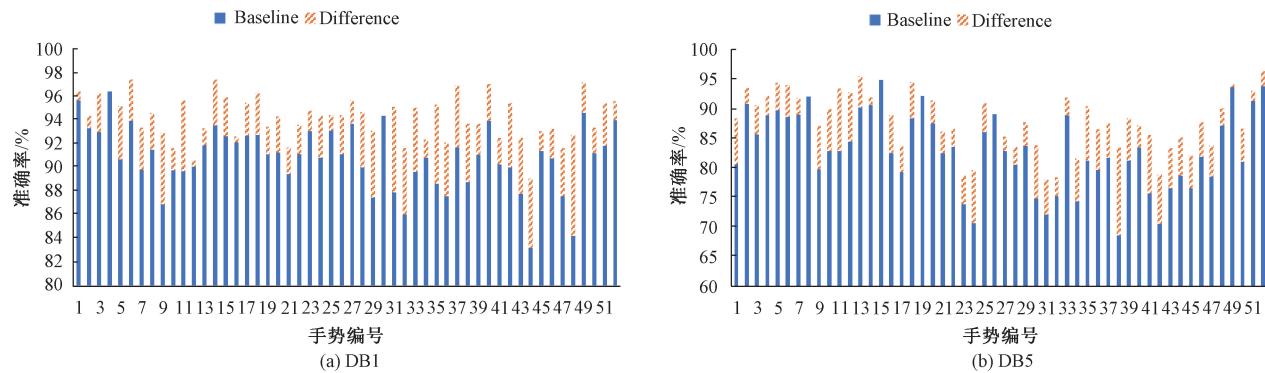


图 13 Baseline 和 RASTNet 在不同手势分类上的表现差异

Fig. 13 Performance difference between Baseline and RASTNet on different gesture classifications

方式,能够直观地看到 RASTNet 模型在哪些手势上取得了明显的准确率提升。

通过图 13(a)中的对比可以看出,RASTNet 模型在 DB1 数据集中的手势 9、11、29、31、48 等取得了显著的提升,准确率提高了 3%以上,其余手势的准确率也都有不同程度的提升。如图 13(b)所示,在 DB5 数据集中的整体表现仍以提升为主。

该堆积柱状图提供了一种高效的方式来分析不同手势在两种模型下的表现差异,进一步揭示了 RASTNet 模型在手势识别中的优越性。通过这些分析,不仅验证了模型在通道与空间协同提取特征方面卓越的能力,还凸显了其在抗噪性能上的显著优势,进一步证明了模型的鲁棒性和可靠性。

## 2) 消融实验

为了验证所提 RASTNet 模型中所引入模块的有效性,更清晰地展现各模块在通道特征提取与空间特征协同方面的实际贡献。本文进行了消融实验,Baseline 为 ResNet50 网络,Baseline + ASPP 为将 ASPP 模块替换  $3 \times 3$  卷积,Baseline + SC 为在 Baseline 模型中引入 SCConv 模块,Baseline + TA 为引入 Triplet Attention 模块,Baseline + STConv 为引入了融合 SCConv 模块和 Triplet Attention 模块的 STConv 模块。Baseline + STConv+ASPP 为结合 ASPP 模块和 STConv 模块的最终模型。

从表 6 的实验结果中可以看出,相比 Baseline 模型,Baseline+ASPP 的准确率提升至 96.58%,较 Baseline 提升 1.40%,表明原来的  $3 \times 3$  卷积作为单一尺度难以完整地表征信号特征,而使用 ASPP 通过不同空洞率全面获取不同尺度下的信号特征,为后续模型处理提供更为丰富的信息基础,进而显著促进模型性能的提升。此外,Baseline+SC 和 Baseline+TA 均带来了性能提升。Baseline+SC 的准确率提高了 1.18%。这一提升验证了 SCConv 模块的有效性,其中 SRU 在空间中将 sEMG 中的噪声和有用信息进行区分,仅保留重要的 sEMG 特征。同时,CRU 减少

了通道间的信息冗余并加强特征之间的信息流动,实现了更精确的特征提取。Baseline+TA 提高了 1.05%,这是因为 Triplet Attention 在多维空间上深入挖掘 sEMG 中不同特征之间的依赖关系,增强了模型对关键特征的关注能力。虽然单独引入 SCConv 模块和 Triplet Attention 模块都取得了一定的性能提升,但 Baseline+STConv 进一步将准确率提升至 96.83%,相比 Baseline 模型提高了 1.65%。这一结果表明,将这两种模块结合使用的效果优于单独使用其中任何一个模块。这种效果源于 SCConv 首先对 sEMG 常规特征进行了通道与空间的协同提取,为 Triplet Attention 提供了一个精细化的空间和通道基础,随后 Triplet Attention 则在更高层次上针对性地加强了特征融合,通过关注跨通道、宽度和高度的特征交互,进一步提升特征的表现力。最后,Baseline + STConv+ASPP 的准确率提升到了 97.15%,相较于 Baseline 模型提升了 1.97%,精确率、召回率和 F1 分数均提升了 1.79%,特异性提升了 0.03%。综上所述,ASPP、SCConv 和 Triplet Attention 三个模块相互配合、层层递进,从多尺度信息获取、精细化基础构建到高层次特征融合强化,全方位提升了 RASTNet 模型在处理复杂 sEMG 信号时的性能,使其在手势识别任务中展现出卓越的优势。

从表 7 可以看出,Baseline + ASPP、Baseline + SC、Baseline+ TA 和 Baseline + STConv 以及最终模型在 Ninapro DB5 数据集上均实现了性能提升。其中,RASTNet 模型的准确率提升至 93.66%,相比于 Baseline 提升了 3.02%。精确率,召回率,特异性和 F1 分数分别提升了 2.86%,3.07%,0.06%,2.97%。这说明 RASTNet 模型在处理不同数据集时,依然能够保持稳定的性能提升。

## 3.7 跨数据集泛化性验证

为了验证模型的泛化能力,本研究在 SIA\_delsys\_16\_movements\_data 数据集上进行实验。该数据集通过 Delsys 系统采集,使用 6 个电极记录前臂和手部关键肌肉的肌电信号,采样率为 2 000 Hz。数据集包含 4 名受试者

表 6 在数据集 Ninapro DB1 的测试集指标

Table 6 Test set metrics on dataset Ninapro DB1

No.	Model	Acc	Precision	Recall	Specificity	F1	$\Delta ACC$
1	Baseline	95.18	95.21	95.20	99.91	95.19	—
2	Baseline+ASPP	96.58	96.61	96.59	99.93	96.59	$\uparrow 1.40$
3	Baseline+SC	96.36	96.36	96.33	99.93	96.33	$\uparrow 1.18$
4	Baseline+TA	96.23	96.26	96.25	99.93	96.25	$\uparrow 1.05$
5	Baseline+STConv	96.83	96.80	96.77	99.94	96.78	$\uparrow 1.65$
6	Baseline+ASPP+STConv	<b>97.15</b>	<b>97.00</b>	<b>96.99</b>	<b>99.94</b>	<b>96.98</b>	$\uparrow 1.97$

表 7 在数据集 Ninapro DB5 的测试集指标

Table 7 Test set metrics on dataset Ninapro DB5

No.	Model	Acc	Precision	Recall	Specificity	F1	$\Delta ACC$
1	Baseline	90.64	90.76	90.63	99.82	90.66	—
2	Baseline+ASPP	92.30	92.29	92.28	99.85	92.26	$\uparrow 1.66$
3	Baseline+SC	91.05	91.12	91.03	99.82	91.05	$\uparrow 0.41$
4	Baseline+TA	91.36	91.43	91.35	99.83	91.35	$\uparrow 0.72$
5	Baseline+STConv	91.75	91.81	91.78	99.84	91.77	$\uparrow 1.11$
6	Baseline+ASPP+STConv	<b>93.66</b>	<b>93.62</b>	<b>93.70</b>	<b>99.88</b>	<b>93.63</b>	$\uparrow 3.02$

执行伸掌、握拳、腕外展、腕内收、捏小指等 16 种手部精细动作,每个动作重复 3 次,每次持续 5 s。本文采用相同的滑动窗口方法对 SIA\_delsys\_16\_movements\_data 数据集进行处理,共生成 31 451 张肌电图,并按照 8:2 的比例划分为 25 161 张训练集和 6 290 张测试集。最终,模型在该数据集上的测试准确率达到 92.27%。其精确率、召回率、特异度和 F1 分数分别为 92.25%、92.23%、98.89%、92.21%。从图 14 和 15 中可以看出,在识别伸掌、握拳等简单手势时准确率接近 100%,个别精细手势虽存在一定程度的混淆,但仍保持着较高的分类性能。相较于 NinaPro DB1 和 DB5,SIA\_delsys\_16\_movements\_data 在电极布局、信号采集设备及肌电特征方面存在显著差异。RASTNet 在此数据集上的高准确率验证了其跨数据集的泛化能力,表明其在通道与空间特征协同提取方面的优势,使其能适应不同数据分布,具备较强的鲁棒性和泛化性。

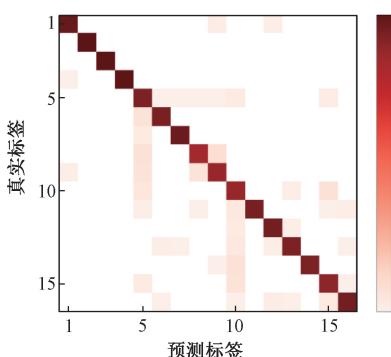


图 14 RASTNet 在 SIA 数据集上的混淆矩阵

Fig. 14 Confusion matrix of RASTNet on the SIA dataset

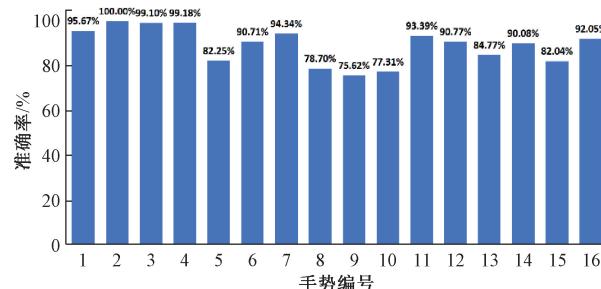


图 15 SIA 数据集手势准确率

Fig. 15 Gesture accuracy on the SIA dataset

## 4 结论

本研究针对 sEMG 信号的稀疏性、非线性及噪声干扰对特征提取的影响,提出了 RASTNet 模型。在高斯噪声、脉冲噪声、时序偏移和振幅变换四种噪声条件下的实验表明,RASTNet 始终保持优异的识别性能,展现了出色的抗噪能力。消融实验进一步验证了 ASPP 在多尺度信息捕捉中的作用以及 STConv 在抑制噪声和优化特征表达方面的有效性。此外,RASTNet 在跨数据集实验中同样表现稳定,证明了其良好的泛化能力。

未来研究将继续优化模型结构,提高计算效率,并探索更多针对 sEMG 信号特点的抗噪策略,以增强模型的适应性和实际应用价值。

## 参考文献

- [1] XU X, DENG H, ZHANG Y, et al. Continuous grasping force estimation with surface EMG based on

- Huxley-type musculoskeletal model [J]. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2023, 31: 346-355.
- [2] KIM J, KIM M, KIM K. Development of a wearable HCI controller through sEMG & IMU sensor fusion[C]. 2016 13th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI). IEEE, 2016: 83-87.
- [3] SHIN S, TAFRESHI R, LANGARI R. Robustness of using dynamic motions and template matching to the limb position effect in myoelectric classification[J]. Journal of Dynamic Systems, Measurement, and Control, 2016, 138(11): 111009.
- [4] 周国良, 张道辉, 郭小萍. 基于 TiCNN-DRSN 模型的 sEMG 手势识别算法的研究[J]. 电子测量技术, 2024, 47(6): 190-196.
- ZHOU G L, ZHANG D H, GUO X P. Research on the sEMG gesture recognition algorithm based on the TiCNN-DRSN model [J]. Electronic Measurement Technology, 2024, 47(6): 190-196.
- [5] OSKOEI M A, HU H. Support vector machine-based classification scheme for myoelectric control applied to upper limb [J]. IEEE Transactions on Biomedical Engineering, 2008, 55(8): 1956-1965.
- [6] ATZORI M, GIJSBERTS A, CASTELLINI C, et al. Electromyography data for non-invasive naturally-controlled robotic hand prostheses[J]. Scientific Data, 2014, 1(1): 140053.
- [7] TKACH D, HUANG H, KUIKEN T A. Study of stability of time-domain features for electromyographic pattern recognition[J]. Journal of Neuroengineering and Rehabilitation, 2010, 7: 21.
- [8] PARK K H, LEE S W. Movement intention decoding based on deep learning for multiuser myoelectric interfaces [C]. 2016 4th International Winter Conference on Brain-computer Interface (BCI). IEEE, 2016.
- [9] GENG W, DU Y, JIN W, et al. Gesture recognition by instantaneous surface EMG images[J]. Scientific Reports, 2016, 6(1): 36571.
- [10] SOROUSHMOJDEHI R, JAVADZADEH S, PEDROCCHI A, et al. Transfer learning in hand movement intention detection based on surface electromyography signals [J]. Frontiers in Neuroscience, 2022, 16: 977328.
- [11] ZHAI X, JELFS B, CHAN R H M, et al. Self-recalibrating surface EMG pattern recognition for neuroprosthetic control based on convolutional neural network [J]. Frontiers in Neuroscience, 2017, 11: 379.
- [12] 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.
- [13] 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, 2019, 119: 131-138.
- [14] QUIVIRA F, KOIKE-AKINO T, WANG Y, et al. Translating sEMG signals to continuous hand poses using recurrent neural networks [C]. 2018 IEEE EMBS International Conference on Biomedical & Health Informatics(BHI). IEEE, 2018: 166-169.
- [15] BAI D, LIU T, HAN X, et al. Multi-channel sEMG signal gesture recognition based on improved CNN-LSTM hybrid models[C]. 2021 IEEE International Conference on Intelligence and Safety for Robotics(ISR). IEEE, 2021: 111-116.
- [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] 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: 2447-2458.
- [18] JIANG B, WU H, XIA Q, et al. Gesture recognition using sEMG based on multi-scale fusion convolution and channel attention[J]. SSRN Electronic Journal, DOI:10.2139/ssrn.4379823.
- [19] 李振江, 魏德健, 冯妍妍, 等. 表面肌电手部动作识别的研究进展[J]. 计算机工程与应用, 2024, 60(3): 29-43.  
LI ZH J, WEI D J, FENG Y Y, et al. Research progress on surface electromyography-based hand gesture recognition [J]. Computer Engineering and Applications, 2024, 60(3): 29-43.
- [20] KRELL M M, SEELAND A, KIM S K. Data augmentation for brain-computer interfaces: Analysis on event-related potentials data[J]. ArXiv preprint arXiv:1801.02730, 2018.
- [21] FAN B, LIU X, SU X, et al. Emgauth: An emg-based smartphone unlocking system using siamese network[C]. 2020 IEEE International Conference on Pervasive Computing and Communications(PerCom). IEEE, 2020: 1-10.
- [22] TSINGANOS P, CORNELIS B, CORNELIS J, et al. Data augmentation of surface electromyography for hand

- gesture recognition[J]. Sensors, 2020, 20(17): 4892.
- [23] MATHUR P, CHAKKA V K. Graph signal processing based cross-subject mental task classification using multi-channel EEG signals [J]. IEEE Sensors Journal, 2022, 22(8): 7971-7978.
- [24] OH Y. Data augmentation techniques for accurate action classification in stroke patients with hemiparesis [J]. Sensors, 2024, 24(5): 1618.
- [25] MA W, SONG G, ZENG Q T, et al. FFCSLT: A deep learning model for traffic police hand gesture recognition using surface electromyographic signals[J]. IEEE Sensors Journal, 2024, 24(8):13640-13655.
- [26] 隋修武,高俊杰,梁天翼,等.基于表面肌电信号及肌肉疲劳的上肢肌力预测[J].电子测量技术,2024,47(5):181-187.
- SUI X W, GAO J J, LIANG T Y, et al. Upper limb muscle strength prediction based on surface electromyography and muscle fatigue[J]. Electronic Measurement Technology, 2024, 47(5):181-187.
- [27] ATZORI M, MÜLLER H. The Ninapro database: A resource for sEMG naturally controlled robotic hand prosthetics [C]. 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society(EMBC). IEEE, 2015: 7151-7154.
- [28] 张娜,张明进,王晓冬,等.基于表面肌电信号的手指关节角度估计方法[J].电子测量与仪器学报,2023,37(8):60-70.
- ZHANG N, ZHANG M J, WANG X D, et al. A method for estimating finger joint angles based on surface electromyography signals [J]. Journal of Electronic Measurement and Instrumentation, 2023, 37(8): 60-70.
- [29] COELHO A L V, LIMA C A M. Assessing fractal dimension methods as feature extractors for EMG signal classification[J]. Engineering Applications of Artificial Intelligence, 2014, 36: 81-98.
- [30] LI J, WEN Y, HE L. Scconv: Spatial and channel reconstruction convolution for feature redundancy[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2023: 6153-6162.
- [31] PANCHOLI S, JOSHI A M, JOSHI D. A robust and accurate deep learning based pattern recognition framework for upper limb prosthesis using semg[J]. ArXiv preprint arXiv:2106.02463, 2021.
- [32] WEI W T, HONG H, WU X L. A hierarchical view pooling network for multichannel surface electromyography-based gesture recognition [J]. Computational Intelligence and Neuroscience, 2021, 2021(1): 6591035.
- [33] KARNAM N K, DUBEY S R, TURLAPATY A C, et al. EMGHandNet: A hybrid CNN and Bi-LSTM architecture for hand activity classification using surface EMG signals [J]. Biocybernetics and Biomedical Engineering, 2022, 42(1): 325-340.
- [34] KRISHNAPRIYA S, SAHOO J P, ARI S. Surface electromyography based hand gesture signal classification using 1D CNN[C]. 2023 International Conference on Intelligent Systems, Advanced Computing and Communication (ISACC). IEEE, 2023: 1-6.
- [35] WANG Z, YAO J, XU M, et al. Transformer-based network with temporal depthwise convolutions for sEMG recognition [J]. Pattern Recognition, 2024, 145: 109967.
- [36] ZHANG K, BADESCA F J, LIU Y, et al. Dual stream long short-term memory feature fusion classifier for surface electromyography gesture recognition [J]. Sensors, 2024, 24(11): 3631.
- [37] 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, 2024, 14(1): 22061.
- [38] LEE H, JIANG M, YANG J, et al. Unveiling EMG semantics: A prototype-learning approach to generalizable gesture classification [J]. Journal of Neural Engineering, 2024, 21(3): 036031.
- [39] PENG F L, CHEN C, LYU D Y, et al. Gesture recognition by ensemble extreme learning machine based on surface electromyography signals [J]. Frontiers in Human Neuroscience, 2022, 16: 911204.
- [40] XIONG B P, CHEN W SH, NIU Y X, et al. A global and local feature fused CNN architecture for the sEMG-based hand gesture recognition[J]. Computers in Biology and Medicine, 2023, 166: 107497.
- [41] NIU Y X, CHEN W SH, ZENG H, et al. Optimizing semg gesture recognition: Leveraging channel selection and feature compression for improved accuracy and computational efficiency [J]. Applied Sciences, 2024, 14(8): 3389.
- [42] WANG Z H, WAN H Y, MENG L, et al. Optimization of inter-subject sEMG-based hand gesture recognition tasks using unsupervised domain adaptation techniques [J]. Biomedical Signal Processing and Control, 2024, 92: 106086.
- [43] MOHAPATRA N R, SAHOO J P, SAHOO G K. Hand gesture classification using surface electromyography

- signals based on fusion of time domain features[C]. 2024 International Conference on Recent Advances in Electrical, Electronics, Ubiquitous Communication, and Computational Intelligence(RAEEUCCI). IEEE, 2024: 1-6.
- [44] ZHANG ZH, SHEN Q M, WANG Y Y. Electromyographic hand gesture recognition using convolutional neural network with multi-attention[J]. Biomedical Signal Processing and Control, 2024, 91: 105935.

### 作者简介

逯英航,硕士研究生,主要研究方向为深度学习与动作识别。

E-mail:luyinghang2023@163.com

仇大伟,副教授,主要研究方向为机器学习、医学图像处理。

E-mail:13954125302@163.com

刘静(通信作者),教授,主要研究方向为机器学习、医学图像处理。

E-mail:liuj\_jn@163.com

李全伟,硕士研究生,主要研究方向为深度学习与行为识别。

E-mail:1918360034@qq.com

王锡城,硕士研究生,主要研究方向为深度学习与医学图像处理。

E-mail:253285497@qq.com