

# 基于表面肌电信号估计多自由度手指力的 ShuffleVT 模型<sup>\*</sup>

李恩诺<sup>1</sup> 耿艳娟<sup>2</sup>

(1. 桂林理工大学机械与控制工程学院 桂林 541006; 2. 中国科学院深圳先进技术研究院 深圳 518055)

**摘 要:** 为了实现基于表面肌电信号(sEMG)的连续手指力估计,本文提出了一种融合 ShuffleNetV2 基本单元与 Vision Transformer(ViT)结构的新模型,命名为 ShuffleVT。为验证该模型的性能,采用公开数据集 NinaPro,其包含 40 名健康受试者的 sEMG 数据和 6 个自由度的手指力数据。性能评估指标为 Pearson 相关系数(CC)、均方根误差(RMSE)和决定系数( $R^2$ )。结果显示,ShuffleVT 模型的 CC、RMSE 和  $R^2$  平均值分别为  $0.92 \pm 0.05$ 、 $1.27 \pm 0.66$  和  $0.83 \pm 0.10$ ,显著优于 ShuffleNetV2、ViT、Transformer 和 LSTM 等 4 种深度学习模型。该结果展示了 ShuffleVT 模型在基于表面肌电的连续运动意图估计中的应用潜力。

**关键词:** 表面肌电信号;手指力;连续估计;深度学习

**中图分类号:** TP391;TN911.7 **文献标识码:** A **国家标准学科分类代码:** 510.4030

## ShuffleVT model for estimating multi-degree-of-freedom finger force based on surface electromyography

Li Sinuo<sup>1</sup> Geng Yanjuan<sup>2</sup>

(1. College of Mechanical and Control Engineering, Guilin University of Technology, Guilin 541006, China; 2. Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China)

**Abstract:** In this study, a novel deep learning model named ShuffleVT was proposed to estimate continuous finger force based on surface electromyography (sEMG). This model was composed of the basic units of ShuffleNetV2 and the Vision Transformer (ViT) structure. Its performance was validated on a publicly available NinaPro database, which includes sEMG data from 40 healthy subjects and finger force data from 6 degrees of freedom. Three performance metrics including Pearson correlation coefficient (CC), root mean square error (RMSE), and coefficient of determination ( $R^2$ ) were used. And another four deep learning models, ShuffleNetV2, ViT, Transformer, and LSTM were included for comparison. The results showed the averaged CC, RMSE, and  $R^2$  was  $0.92 \pm 0.05$ ,  $1.27 \pm 0.66$ , and  $0.83 \pm 0.10$ , respectively, significantly better than that computed with another four models. It indicates that the newly proposed ShuffleVT model could be potentially applied into the sEMG-driven continuous estimation of human motor intention.

**Keywords:** surface electromyography; finger force; continuous estimation; deep learning

## 0 引 言

表面肌电图(surface electromyography, sEMG)作为神经肌肉活动的直接反映,能够有效解码人类的运动意图。肌电控制技术通过记录和处理 sEMG,将运动意图转化为控制命令,使用户能够主动控制设备,提供易操作、无创且低成本的人机界面。手指力量是手部运动意图的重要体

现,通过解析 sEMG 来准确提取用户的指尖力量意图,是肌电控制技术的核心研究方向,特别是在人工假肢、康复机器人、遥操作、工业机器人等人机交互场景中。

自加拿大阿尔伯塔大学提出典型的肌电信号时域特征后,采用特征工程和模式分类技术来识别各种手势及肌肉收缩力等级是主要的研究思路<sup>[1]</sup>。随着深度学习技术的进步,近十年来,大量研究采用不依赖于特征提取的深度学习

框架<sup>[2]</sup>。例如,Chen 等<sup>[3]</sup>提出了一种结合注意力机制的多尺度多任务学习网络,该网络可以同时识别手势和力量等级,对于肢体健全的参与者,该网络在识别 4 种手势和 3 种力量等级时的准确率分别达到了 95.19%和 89.91%;而在截肢者中,相应的识别准确率分别为 90.78%和 75.96%。尽管如此,该类方法局限于离散的力量等级识别,不能提供连续变化的力度控制。在实际应用中,对连续多自由度力进行准确估计,有助于实现更自然、更准确的人机交互控制,显著提升系统的功能性和良好的用户体验。

目前,连续多自由度力估计方法大致可分为基于模型的方法和无模型方法两大类。最早提出的模型是希尔(Hill)收缩动力学模型,这类基于肌肉生理学模型的优势在于能够直观的解释运动的产生过程,并定量分析 sEMG 与肌肉运动量之间的关系,同时模型的参数可以表达肌肉骨骼系统的属性。尽管近年来许多研究已经融入了生理可解释的模块<sup>[4-6]</sup>,但人体肌肉模型的构建过程复杂,相关生理参数的测量难度大且数量众多,优化困难。此外,不同关节活动区的肌肉分布差异显著,导致在使用模型计算连续输出时,肌肉激活程度的不确定性使得模型在不同个体间适应性较差。因此,研究人员在实际控制问题中更倾向于采用无模型方法。

在无模型的方法中,常采用机器学习和深度学习技术。Wu 等<sup>[7]</sup>比较了多种机器学习方法对手指按压压力预测和不同手指之间夹紧运动分类中的表现,但该研究中所需的 8 个特征是从 sEMG 中手动提取的。鉴于 sEMG 在动态运动过程中的非线性特性,以及深度学习模型能够自动从原始数据中提取有用特征的能力(无需人工特征工程),越来越多的研究采用了深度学习方法,并展示出显著的潜力。例如,Sun 等<sup>[8]</sup>利用多流卷积神经网络(MSCNN)捕获不同尺度 sEMG 特征来实现手势识别,同时利用长短期记忆网络(long short-term memory, LSTM)对肌肉力量进行估计,并且与支持向量机、极限学习机以及线性动态模型等传统机器学习方法作对比,结果表明其在肌肉力量估计方面表现出了显著的优势。然而,由于多自由度的 sEMG 信号数据量庞大,且不同个体的肌肉结构和神经分布差异显著,这些因素使得现有深度学习方法在连续手指力估计的精度方面仍面临挑战。

为了解决上述问题,本研究提出了一种基于 sEMG 的深度学习模型,命名为 ShuffleVT。该模型结合了 ShuffleNetV2 基本单元与 Vision Transformer (ViT) 结构,ShuffleNetV2 采用轻量级设计,并运用通道重排技术(channel shuffle)增强特征之间的信息交流以提供良好的特征提取能力,而 ViT 则通过将输入数据划分为多个固定长度的片段,并将每个片段添加位置信息,利用自注意力机制有效捕捉输入信息中的全局和局部特征。通过结合这两种架构的特点,本文期望在提升数据处理速度的同时,加强模型对复杂特征的学习能力,进而提高对用户间信号的理

解能力和解码精确度。本研究将通过与 4 种深度学习模型的对比实验验证本文提出的模型在提高连续手指力估计精度方面的可行性。

1 方法

1.1 数据库

实验采用了 NinaPro 公开数据集,该数据集旨在推动人类、机器人和假手的机器学习研究。考虑到受试者的多样性,选用 NinaPro 中的 DB2 数据库。该数据库包含了 40 位健康受试者的 50 种手部动作,受试者的年龄在 23~45 岁。实验的 sEMG 通过 Delsys Trigno 无线肌电采集系统记录(共 12 个通道),采样率为 2 kHz,电极贴放在右肢的前臂和上臂,如图 1(a)所示,其中 8 个电极(1~8)等距分布在前臂周围,两个电极(9 和 10)分别放置在指屈肌和指伸肌,最后两个电极(11 和 12)分别放置在肱二头肌和肱三头肌。此外,如图 1(b)所示<sup>[9]</sup>,受试者使用其惯用手的一个或多个手指指尖按压设备,产生九种力量模式,每种动作重复 6 次,保持 5 s,休息 3 s,这些动作涉及 6 个自由度,包括五指弯曲和拇指外展(如表 1 所示),力信号通过手指力线性传感器(FFLS)进行采集。

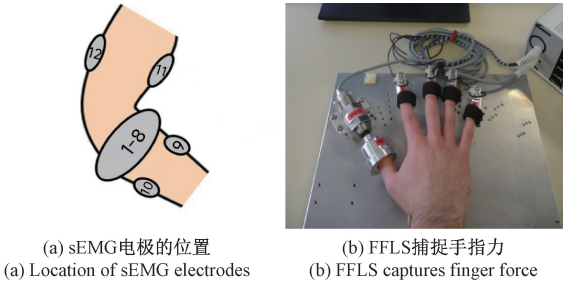


图 1 数据采集  
Fig. 1 Data acquisition

表 1 6 个自由度描述

Table 1 Description of 6 degree-of-freedom	
自由度	描述
1	小拇指弯曲
2	无名指弯曲
3	中指弯曲
4	食指弯曲
5	大拇指外展
6	大拇指弯曲

1.2 数据预处理

1) 去噪

采用截止频率为 5~450 Hz 的巴特沃斯带通滤波器对 sEMG 进行滤波,去除直流分量和高频域噪声,并采用陷波滤波器衰减电源线干扰(50 Hz)。

2) 归一化

使用均方根归一化算法将 sEMG 缩放到 $[-1, +1]$ 数

据范围,手指力数据使用最小-最大归一化(min-max normalization)算法缩放到 $[0, +1]$ 之间。

3) 样本分割

sEMG 以 2 000 Hz 采样,使用 400 ms(800 个样本)的窗口,滑动窗口增量为 10 ms(20 个样本)进行分割。为了提高计算效率,数据集以 10 倍的因子进行了子采样<sup>[10]</sup>。

4) 训练集和测试集的划分

数据根据动作的重复次数划分训练集和测试集。每种动作重复 6 次,采用交叉验证的方法,从每个动作的 6 次重复中随机选择两次作为测试集,剩余 4 次作为训练集,并对所有组合情况下的计算结果取平均值。

1.3 ShuffleVT 模型

ShuffleVT 模型的整体结构如图 2 所示。预处理后的 sEMG 首先通过 1 个  $3 \times 3$  卷积层(Conv)和 1 个最大池化层(Maxpooling),接着,信号进入由两个 ShuffleNetV2 基本单元组成的 ShuffleNetV2 Stage,最后,全连接层(Linear)将 ViT 中编码器层的输出转换为手指力回归预测值。

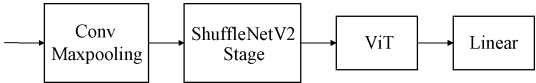


图 2 ShuffleVT 模型整体结构

Fig. 2 Overall structure of ShuffleVT model

ShuffleNetV2 基本单元和 ViT 模块主要描述如下。

1) ShuffleNetV2 基本单元

ShuffleNetV2 模型提出了一种新颖且轻量化的基本单元。它的基本模块包含两种不同结构,如图 3 所示,左边为步长(stride)=1 的结构,右边为 stride=2 的结构,这两种结构的主分支都由 3 个 Conv 构成,分别为  $1 \times 1$  Conv、 $3 \times 3$  深度可分离卷积(DWConv)和另一个  $1 \times 1$  Conv,但输入端和侧分支的构成有所不同。在步长为 1 的结构中,输入首先经过通道分割(Channel Split),侧分支直接接收来自 Channel Split 的输出,并与主分支的结果进行维度上的拼接(concat);在步长为 2 的结构中,输入被分成两路,分别经过主分支和侧分支的处理,其侧分支由  $3 \times 3$  DWConv 和  $1 \times 1$  Conv 构成。在这两种结构中,最后都通过通道混洗(Channel shuffle)进行两个分支之间的信息交流。

2) ViT 模块

ViT 模块包含信号分割与片段嵌入、位置编码、编码器层 3 个核心部分,其结构如图 4 所示。

信号分割与片段嵌入部分的工作原理是将输入的 sEMG 划分为多个固定长度的片段,并将每个片段映射为一个向量表示。具体而言,对于一个输入维度为 Batch\_Size(B) \* Input\_Channels(C) \* Sequence\_Length(S) 的信号,本文沿着 S 维度将其切割为 T 长度的小段,这样可以从整个序列中得到  $N=S/T$  个片段。随后,这些片段被重组为  $B \times N \times C$  的形式。接下来,通过展平操作,每个片段

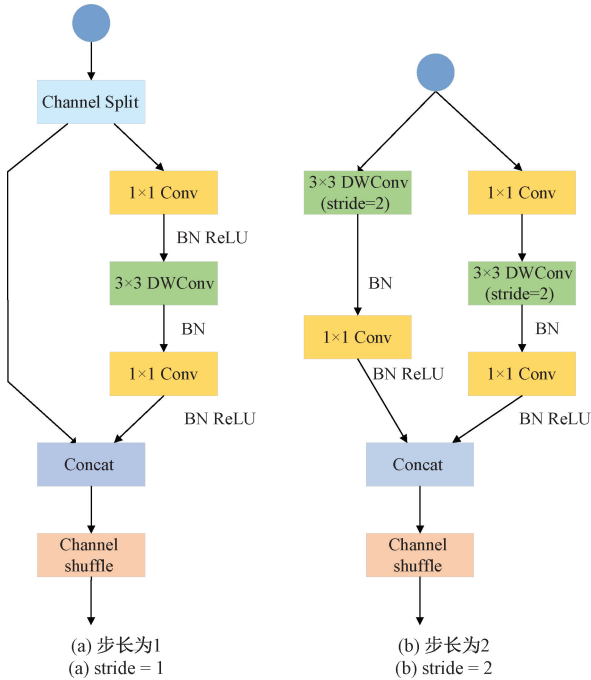


图 3 ShuffleNetV2 基本单元

Fig. 3 ShuffleNetV2 basic unit

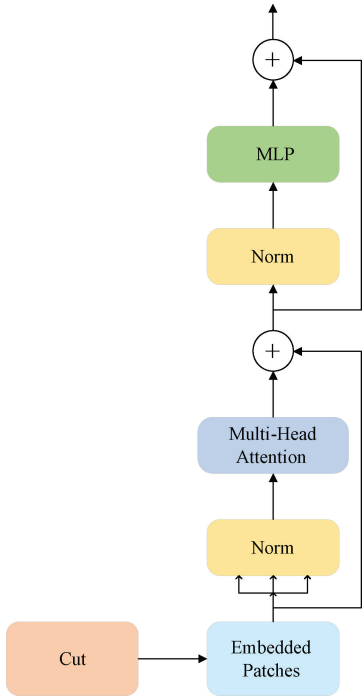


图 4 ViT 结构

Fig. 4 ViT structure

被转换为长度为  $C \times T$  的一维向量,该操作有助于保持信号的时间特性并便于后续处理。切割长度 T 的选择至关重要:如果 T 设置得过小,可能无法捕获足够的时序信息;如果 T 设置得过大,每个片段包含的信息可能过于复杂,从而增加后续学习的难度。在本模型中,选择  $T=50$ 。之

后,这些一维向量经过片段嵌入处理,将它们映射至高维空间,形成  $N \times D$  维度的二维向量集,其中  $D$  为嵌入维度,旨在深入捕获 sEMG 的动态特性及其特征信息。

由于自注意力机制本身无法区分输入信号的顺序关系,因此需要在位置编码部分为输入序列中的每个片段添加位置信息。在本模型中,位置编码通过将每个片段的嵌入向量与相应的固定或可学习的位置信息相加,从而增强模型对序列数据的理解能力。注意机制不仅在长时间序列的预测精度上具有优势,而且在语言翻译任务中训练时间也显著减少<sup>[11]</sup>。

编码器层的计算主要包括多头注意力计算以及残差连接计算。其中,多头注意力操作的计算如式(1)所示。

$$Attention(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d}}\right)V \quad (1)$$

式中:  $\text{softmax}$  为激活函数,  $d$  表示输入编码器层的维度,  $Q, K, V$  分别表示 Query、Key、Value 矩阵,这 3 个矩阵计算公式如式(2)~(4)所示。

$$Q = W_1x \quad (2)$$

$$K = W_2x \quad (3)$$

$$V = W_3x \quad (4)$$

式中:  $W_1, W_2, W_3$  分别表示 Query、Key、Value 矩阵的权重。另外,模型中残差连接计算的输出结果如式(5)所示。

$$Z = \text{MSA}(\text{LN}(Z_0)) + Z_0 \quad (5)$$

式中:  $\text{MSA}$  为多层注意力操作,  $\text{LN}$  为层归一化操作,  $Z_0$  表示编码器层的输入。编码器层的输出将被送入一个回归头进行连续回归估计任务。

#### 1.4 对比试验模型

除了新提出的 ShuffleVT 模型外,本研究还包括 ShuffleNetV2、ViT、Transformer 和 LSTM 进行对比分析。所有算法均基于 Pytorch 2.3.0 框架上进行建模,并在配备 62 GB 内存、Intel Core i9 CPU(时钟频率 3.50 GHz)和 NVIDIA GeForce RTX 4090 GPU 的环境下进行训练,使用 Adam 优化器进行优化。模型中,隐藏层均包含 256 个隐藏单元,最后连接一个线性层输出 6 个通道手指力数据。

1) ShuffleNetV2<sup>[12]</sup> 是一种轻量级的卷积神经网络(CNN),旨在提高效率和性能,同时减少计算资源的需求。在本研究中,ShuffleNetV2 整体结构包括 Stage1、Stage2、Stage3 和 Stage4,之后连接一个卷积层和全局池化层。Stage1 包含初始的卷积层和池化层,输出通道数是 24,Stage2 至 Stage4 均由重复两次的基本单元构成,输出通道数分别是 48、96、192。

2) ViT<sup>[13]</sup> 是一种基于 Transformer 架构的视觉注意力模型,原理是将图像分割成一系列的固定大小的补丁(patch),并将这些 patch 视为序列输入到 Transformer 模型中,所以其架构本质上适合处理序列数据。在本研究中,输入  $C=12, S=400$ ,沿着  $S$  维度切割为  $T=20$  长度的小段,每层中多头注意力机制的头数为 8。

3) Transformer: CNN 在提取局部特征方面表现出色,但在处理序列中的长程依赖关系时表现不佳。相比之下,Transformer 模型能够有效地捕捉序列中的长程依赖关系,从而生成更全局性的特征表示<sup>[14]</sup>。Transformer 及其变种可以并行处理输入序列中的所有元素,这大大加快了训练速度。在本研究中,每层中多头注意力机制的头数为 8。

4) LSTM<sup>[15]</sup> 是一种多对多递归神经网络(RNN),用于模拟时间序列并避免长期依赖问题。与 RNN 相比,LSTM 在处理基于回归的问题时表现更佳,尤其是在输入数据包含长时间信息的情况下<sup>[16]</sup>。这里,本文使用了 5 个堆叠的 LSTM 来增加模型的深度,并且,加入了一个 dropout 层以防止过拟合,丢弃率设定为 0.1。

#### 1.5 评估指标

为评估上述 5 种深度学习模型估计多自由度手指力的准确度,采用 Pearson 相关系数(CC)、均方根误差(RMSE)和决定系数( $R^2$ )3 个指标。此外,将预测的手指力曲线和对应的实际手指力曲线绘制在同一张图中,以可视化评估性能。

1) CC 用于衡量两个连续变量之间线性关系强度和方向,它的取值范围在  $-1 \sim 1$ , CC 值越接近 1,表示预测的手指力与实际手指力值越接近,估计精度越高。在这里,本文用它来评估预测的手指力与相应的实际手指力的拟合程度。其计算公式为:

$$CC = \frac{\sum_{i=1}^N (y_{pred} - \overline{y_{pred}})(y_{real} - \overline{y_{real}})}{\sqrt{\sum_{i=1}^N (y_{pred} - \overline{y_{pred}})^2} \sqrt{\sum_{i=1}^N (y_{real} - \overline{y_{real}})^2}} \quad (6)$$

式中:  $y_{pred}, \overline{y_{pred}}, y_{real}, \overline{y_{real}}$  分别为手指力的估计值、手指力的估计平均值、手指力的实际值和手指力的实际平均值。

2) RMSE 是用来衡量预测值与实际观测值之间差异的一种常见指标。RMSE 越小表示预测模型的预测精度越高。在这里,本文使用 RMSE 来评估手指力的估计值与实测值之间的偏差。RMSE 的计算方法为:

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (y_{pred} - y_{real})^2}{N}} \quad (7)$$

3)  $R^2$  是用来评估回归模型拟合优度的指标,取值范围为  $0 \sim 1$ ,  $R^2$  值越接近 1,估计性能越好。在这里,我们使用  $R^2$  来衡量手指力的估计值与实测值之间的变异程度解释比例。 $R^2$  的计算公式为:

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_{real} - y_{pred})^2}{\sum_{i=1}^N (y_{real} - \overline{y_{real}})^2} \quad (8)$$

#### 1.6 统计方法

为了检验上述 5 种深度学习算法之间的差异性,本研究使用了 Friedman 检验,这是一种适用于重复测量设计中多组数据比较的非参数方法,可作为单因素方差分析(one-way ANOVA)的替代方法。检验中 CC、RMSE 和  $R^2$  分别



用作因变量。如果 Friedman 检验显示存在统计显著性差异,则进一步使用威尔科克森符号秩检验(Wilcoxon signed rank test)进行事后比较,并通过 Bonferroni 校正来调整多重比较中的  $p$  值。在这项工作中,设定  $p$  值小于 0.05 为统计显著性的标准阈值。

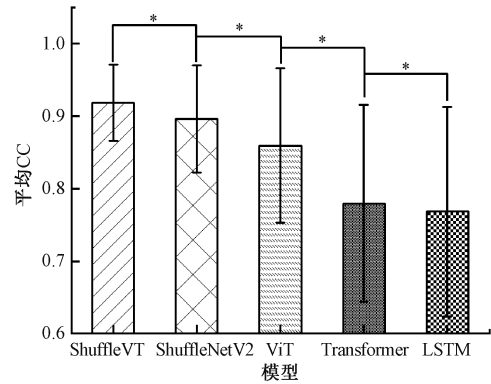
2 结果

手指力估计在 5 种深度学习模型的整体性能如图 5 所示,其中 CC、RMSE 和  $R^2$  是在所有自由度和所有受试者

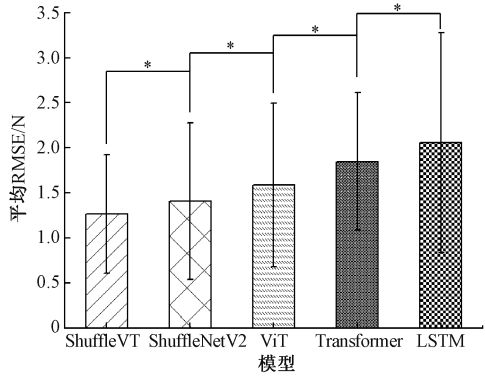
中的平均值。具体来看,ShuffleVT、ShuffleNetV2、ViT、Transformer 和 LSTM 对应的 CC 均值分别为  $0.92 \pm 0.05$ 、 $0.89 \pm 0.07$ 、 $0.86 \pm 0.11$ 、 $0.78 \pm 0.14$  和  $0.77 \pm 0.14$ 。对应的平均 RMSE 分别为  $1.27 \pm 0.66$ 、 $1.41 \pm 0.87$ 、 $1.59 \pm 0.91$ 、 $1.84 \pm 0.76$  和  $2.06 \pm 1.22$ 。平均  $R^2$  分别为  $0.83 \pm 0.10$ 、 $0.79 \pm 0.13$ 、 $0.73 \pm 0.17$ 、 $0.62 \pm 0.20$  和  $0.58 \pm 0.22$ 。结果表明,本文提出的 ShuffleVT 模型在 CC、RMSE 和  $R^2$  3 个指标上均显著优于其他 4 种深度学习模型( $p$  值  $< 0.05$ )。此外,ShuffleVT 模型在所有受试者上都表现出较低的 RMSE,这说明了该模型不仅具有较高的准确性,而且对不同的个体数据也表现出了良好的适应性和稳健性。

5 个模型在 6 个自由度上的估计精度如图 6 所示。可以看出,ShuffleVT 对于每个自由度都具有较高的 CC ( $> 0.91$ )、较低的 RMSE ( $< 1.62$ ) 和较高的  $R^2$  ( $> 0.80$ ),而 ShuffleNetV2、ViT、Transformer 和 LSTM 的表现都比 ShuffleVT 差。

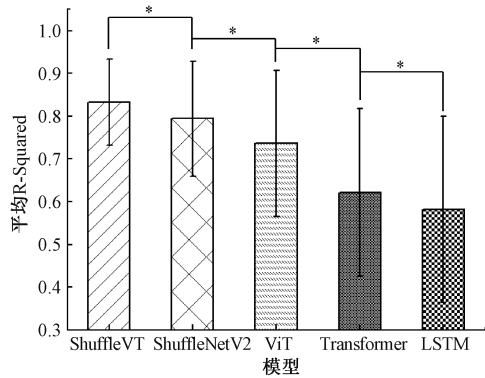
这些结果表明,本文提出的 ShuffleVT 模型具有适用于不同受试者和自由度的能力、稳定性和通用性。对每个自由度和受试者的统计分析一致表明,无论在哪个性能指标上,所提出的 ShuffleVT 模型都优于其他 4 种深度学习模型。



(a) 5个模型的平均CC值对比  
(a) Comparison of average CC values of five models



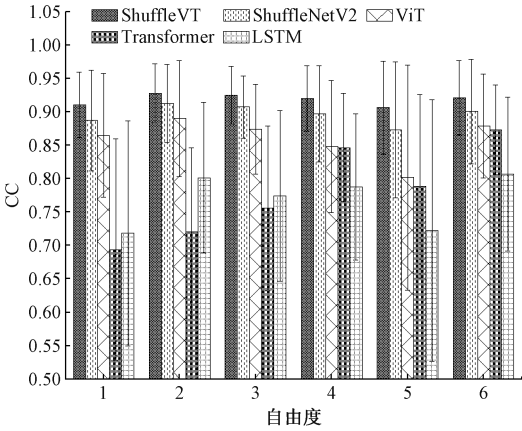
(b) 5个模型的平均RMSE值对比  
(b) Comparison of average RMSE values of five models



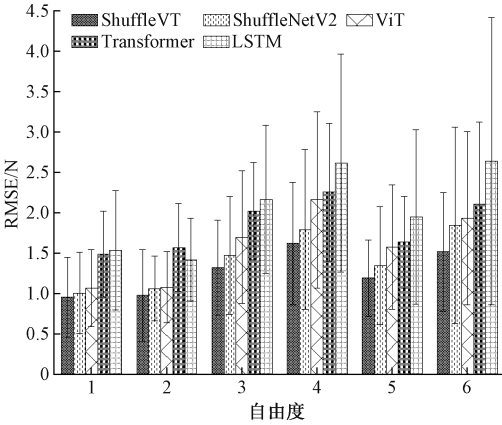
(c) 5个模型的平均R^2值对比  
(c) Comparison of average R^2 values of five models

图 5 5 个模型整体性能

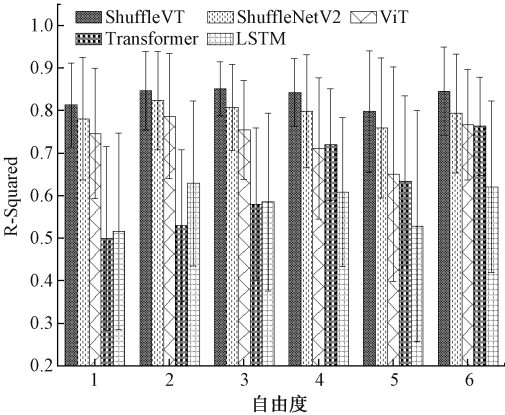
Fig. 5 Overall performance of five models



(a) 5个模型在6个自由度上的CC值对比  
(a) Comparison of CC values of five models in six degree-of-freedom



(b) 5个模型在6个自由度上的RMSE值对比  
(b) Comparison of RMSE values of five models in six degree-of-freedom



(c) 5个模型在6个自由度上的R<sup>2</sup>值对比

(c) Comparison of R<sup>2</sup> values of five models in six degree-of-freedom

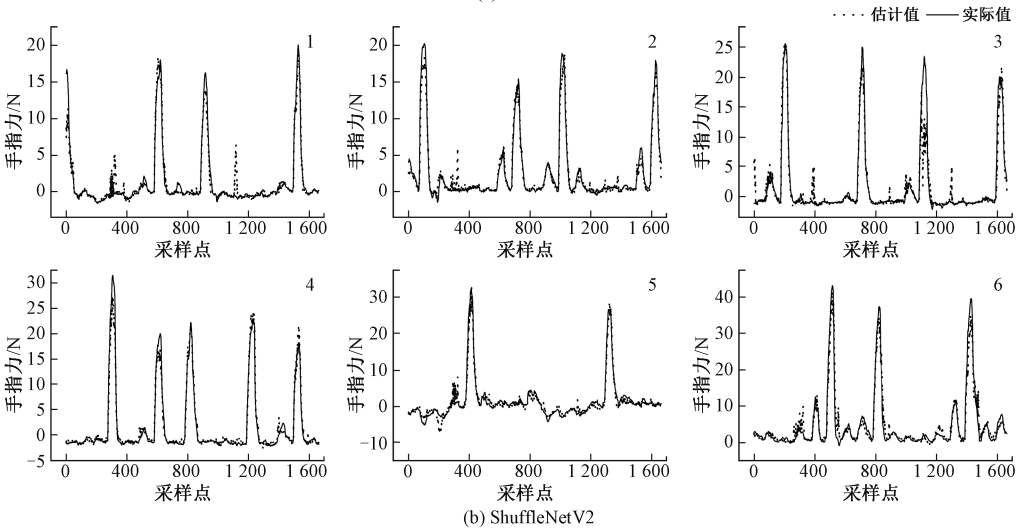
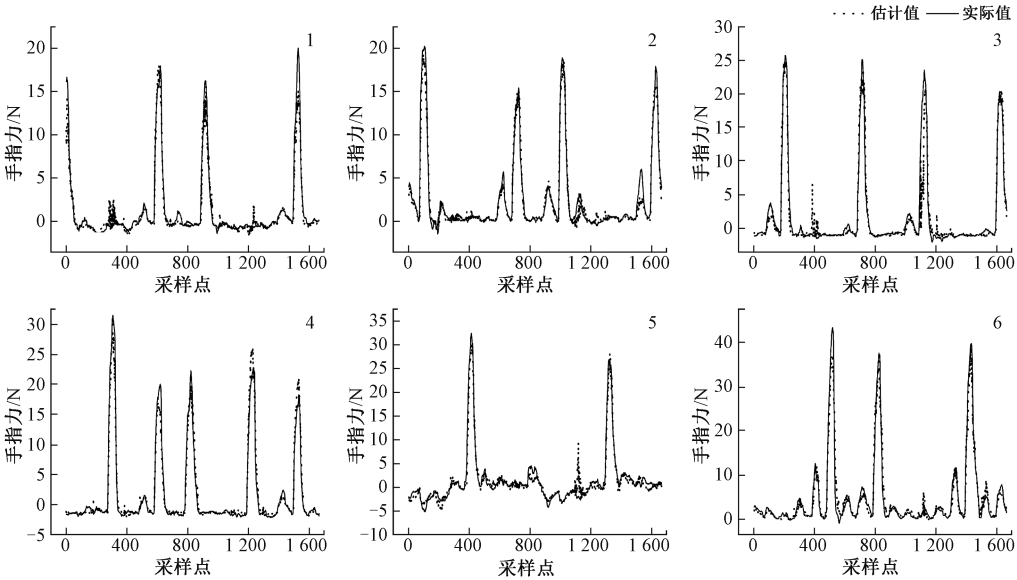
图 6 5 个模型在 6 个自由度上的性能

Fig. 6 Performance of five models on 6 degree-of-freedom

图 7 显示了各模型对受试者 33 的手指力 6 个自由度的连续预测结果,其中实线代表实际手指力曲线,点线代表模型预测曲线。显然,ShuffleVT 模型表现出色,其预测曲线与实测曲线高度相似,更接近实际值,整体拟合曲线平滑,无畸变,明显优于其余 4 种模型。ShuffleNetV2 模型也展现了准确的趋势,但是曲线中的抖动稍显明显,尤其是第 3 个自由度的曲线拟合效果稍差。相比之下,Transformer 和 LSTM 的拟合效果表现不佳,曲线抖动严重,出现的突然转折或尖锐的角度较多。

### 3 讨论

本研究利用了 40 名健康受试者的 sEMG 和 6 个自由度的手指力数据,以探讨连续估计手指力方法。本文提出了一种名为 ShuffleVT 的深度学习模型。本算法的开发是基于对现有模型的深入分析,旨在实现对多自由度手指力



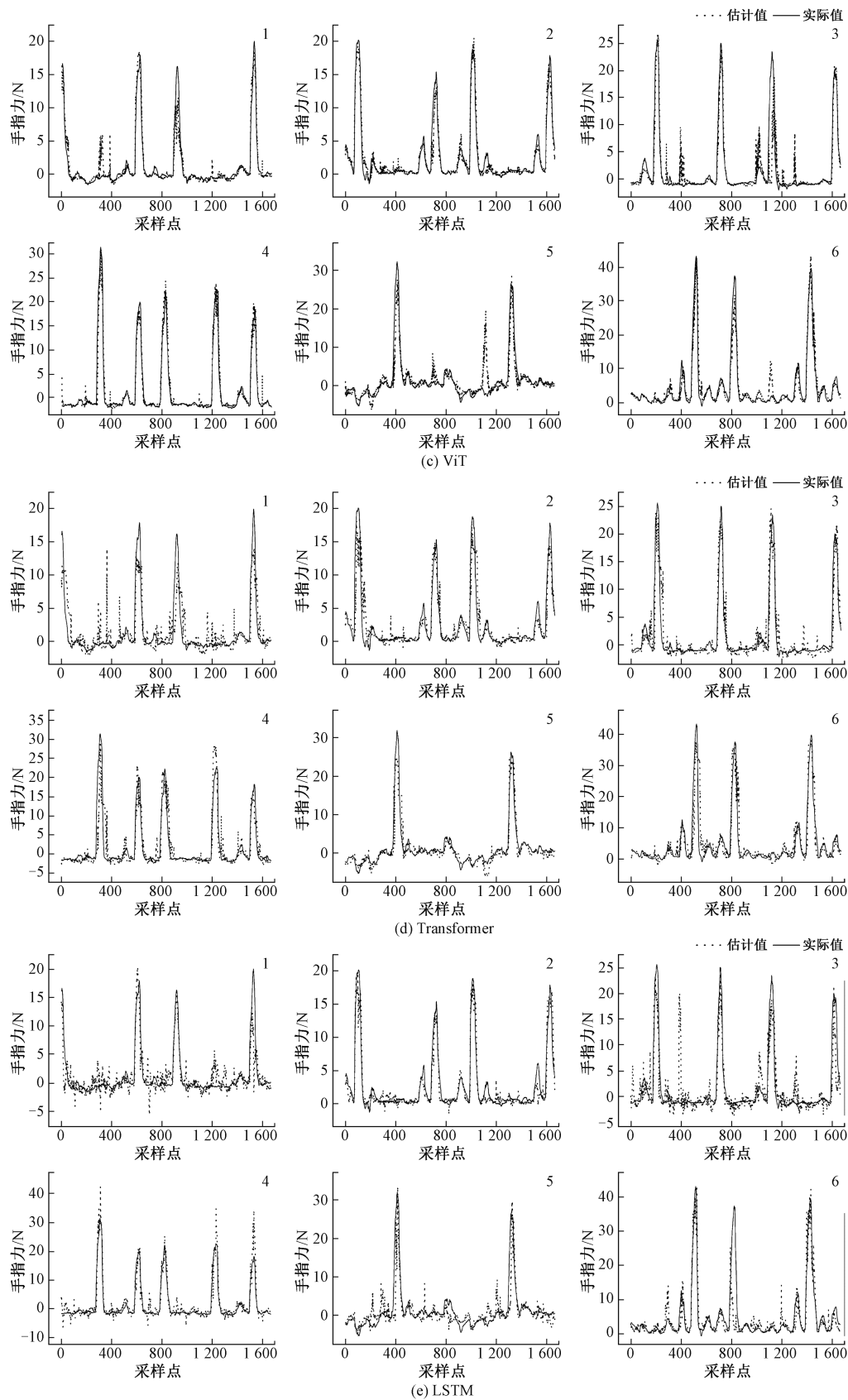


图 7 5 个模型拟合图

Fig. 7 Fitting diagram of five models

的准确估计。

为了评估 ShuffleVT 模型的性能,本研究还选择了 4 种深度学习方法 ShuffleNetV2、ViT、Transformer 和 LSTM 进行对比。为了准确描述每个模型的表现,研究中计算了实际手指力和预测手指力之间的 CC、RMSE 和  $R^2$  指标。图 5、6 的结果显示,新提出的 ShuffleVT 模型不仅是在整体性能上,而且在针对所有受试者和所有自由度的单独评估中都表现出更高的准确性和更好的稳定性,显著优于 ShuffleNetV2、ViT、Transformer 和 LSTM。其中,ShuffleNetV2 的表现仅次于新提出的模型,而 Transformer 和 LSTM 的表现最差。ShuffleNetV2 通过逐通道的组合操作和轻量级的卷积设计,在保持相对较低的计算复杂度的同时,提供了局部细节的特征提取能力,但是,由于模型大小和计算开销的限制,导致其精度稍显不足,并且对输入数据中的局部结构和特征具有较高的依赖性。ViT 将输入划分为多个固定片段,通过自注意力机制实现全局感知,能够捕捉全局和长程依赖关系,然而,由于需要维护大量的注意力矩阵,导致计算复杂度较高,并且在处理局部信息时表现略有不佳。本文新提出的 ShuffleVT 模型将这两种结构结合,不仅提高了计算效率,还可以在处理片段信息时,同时考虑到局部细节和全局信息,从而提高模型的全局理解能力和预测精确度。

图 7 通过对比预测曲线和实际曲线,进一步验证了本文提出的 ShuffleVT 方法的优势。可以发现,ShuffleNetV2、ViT、Transformer 和 LSTM 4 种深度学习方法模型都能明确地捕捉手指力大小的振幅变化,但平稳性方面存在明显不足,曲线抖动更明显,尤其是在第 3 和第 5 自由度上(其中,第 3 自由度代表中指的弯曲,第 5 自由度代表大拇指的外展),这两个动作相较于其他动作更复杂,这可能是导致估计精度下降的关键因素。相比之下,ShuffleVT 在 6 个自由度均提供了更连续、稳定且准确的手指力估计曲线,与实际曲线高度一致,进一步印证了 ShuffleVT 模型的卓越表现。

在这项工作中,还发现算法的优越性与其性能指标之间存在着密切联系。例如,整体性能方面,5 个模型的表现排名依次是 ShuffleVT、ShuffleNetV2、ViT、Transformer、LSTM(如图 5 所示),然而,在 6 个自由度的性能评估中(如图 6 所示),尽管 LSTM 在第 1、第 2 和第 3 自由度的 CC 和  $R^2$  指标优于 Transformer,但在 RMSE 结果中,LSTM 在第 1 和第 3 自由度上的表现却不及 Transformer。在所有受试者的评估结果中,本文也发现了相似的估计精度冲突结果。CC 和 RMSE 分别从波形相似度和误差幅度两个不同角度评估曲线的拟合程度,因此可能导致结果不一致。所以,在这项工作中,关于 Transformer 和 LSTM 在不同指标下表现的不一致性,突显了使用多种性能指标进行综合评价的重要性。

这项工作中也存在一些局限性。首先,本文提出的

ShuffleVT 模型仅评估了手指 6 个自由度的力,所考虑的自由度数量较为有限。其次,研究使用的 Ninapro 数据库采用稀疏电极,不能提供足够的时间或空间信息来预测更精细的运动信号变化,相比之下,放置在前臂和手部的柔性可拉伸电极能够获取高密度肌电信号,并且由于能与肌肉更好地贴合,它们能够记录更精确的肌肉活动<sup>[17-18]</sup>,因此在未来的工作中,将探索这种电极在验证模型效果中的潜力。再者,本文使用了 40 名受试者的数据进行实验,未来还需要进一步探讨在减少训练受试者数量的情况下,模型的性能是否会受到影响,从而提高模型的泛化能力。此外,为了实现假肢的高度灵巧控制,后续研究将采用运动类型分类与比例力回归相结合的控制方案。最后,本研究仅考虑了五种深度学习方法的离线性能,尚不确定这些方法是否同样适用于在线实时应用<sup>[19]</sup>。在实际场景中,电极位置变化、肌肉收缩力变化和手臂位置等因素都可能会影响 sEMG 信号模式<sup>[20-21]</sup>,因此,未来的研究需要加快算法的计算速度,并增强其对上述因素的鲁棒性,以验证其在实时应用中的有效性。

4 结 论

本文提出了一种名为 ShuffleVT 的深度学习算法,旨在从 sEMG 中连续且准确地估计手指力,并在 Ninapro 数据库进行了验证。与其他现有的 4 种深度学习方法(ShuffleNetV2、ViT、Transformer 和 LSTM)相比,ShuffleVT 在 3 个性能指标以及预测手指力曲线和实际手指力曲线的拟合度上均表现出显著的准确性和稳定性优势。这些结果表明,这种高精度的预测模型为未来开发直观、灵巧的多自由度假肢手控制系统提供了技术支持,有望在未来的人机交互与合作中发挥重要作用。

参考文献

[1] 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 sEMG gesture recognition algorithm based on TiCNN-DRSN model[J]. Electric Measurement Technology, 2024, 47(6): 190-196.

[2] XIONG D, ZHANG D, ZHAO X, et al. Deep learning for EMG-based human-machine interaction: A review[J]. IEEE/CAA Journal of Automatica Sinica, 2021, 8(3): 512-533.

[3] CHEN Z, YU Y, WANG L, et al. AGMMLN: An attention-guided multiscale multitask learning network for simultaneous gesture and force level recognition [J]. IEEE Sensors Journal, 2024, DOI: 10.1109/JSEN.2024.3416198.

[4] XI X G, JIANG W J, HUA X, et al. Simultaneo-us



- and continuous estimation of joint angles based on surface electromyography state-space model[J]. IEEE Sensors Journal, 2021, 21(6): 89-99.
- [5] LIU H, TAO J, LYU P, et al. Human-robot cooperative control based on sEMG for the upper limb exoskeleton robot [J]. Robotics and Autonomous Systems, 2020, 125: 103350.
- [6] 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. Prediction of upper extremity muscle strength based on surface EMG signal and muscle fatigue [J]. Electronic Measurement Technology, 2024, 47(5): 181-187.
- [7] WU Y, LIANG S, MA Y, et al. Prediction and classification of sEMG-based pinch force between different fingers[J]. Expert Systems with Applications, 2024, 237: 121635.
- [8] SUN X, LIU Y, NIU H. Continuous gesture recognition and force estimation using sEMG signal[J]. IEEE Access, 2023, DOI: 10. 1109/ACCESS. 2023.3323586.
- [9] KOIVA R, HILSENBECK B, CASTELLINI C. FFLS: An accurate linear device for measuring synergistic finger contractions [C]. 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2012.
- [10] GIJSBERTS A, ATZORI M, CASTELLINI C, et al. Movement error rate for evaluation of machine learning methods for sEMG-based hand movement classification [J]. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2014, 22(4): 735-744.
- [11] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[J]. Advances in neural information processing systems, 2017, DOI: 10.48550/ARXIV.1706.03762.
- [12] MA N, ZHANG X, ZHENG H T, et al. Shufflenet v2: Practical guidelines for efficient cnn architecture design[C]. European Conference on Computer Vision (ECCV), 2018: 116-131.
- [13] ALEXEY D. An image is worth  $16 \times 16$  words: Transformers for image recognition at scale[J]. ArXiv preprint arXiv: 201011929, 2020.
- [14] 李赵春, 周永照, 冯卫奔, 等. 基于 Transformer 模型的手势脑电信号分类识别[J]. 科学技术与工程, 2023, 23(5): 2044-2050.
- LI ZH CH, ZHOU Y ZH, FENG W B, et al. Gesture action electroencephalogram classification and recognition based on transformer model[J]. Science Technology and Engineering, 2023, 23(5): 2044-2050.
- [15] GRAVES A. Long short-term memory [J]. Supervised Sequence Labelling with Recurrent Neural Networks, 2012: 37-45, DOI: 10.1007/978-3-642-24797-2\_4.
- [16] YU Y, SI X S, HU C H, et al. A review of recurrent neural networks: LSTM cells and network architectures[J]. Neural Computation, 2019, 31(7): 1235-1270.
- [17] YAN X Z, LIU Z Y, ZHANG Q H, et al. Quadruple H-bonding cross-linked supramolecular polymeric materials as substrates for stretchable, antitearing, and self-healable thin film electrodes[J]. Journal of the American Chemical Society, 2018, 140(15): 5280-5289.
- [18] CHENG L, LI J, GUO A Y, et al. Recent advances in flexible noninvasive electrodes for surface electromyography acquisition [J]. Npj Flexible Electronics, 2023, 7(1): 39.
- [19] NAWFEL J L, ENGLEHART K B, SCHEME E J. A multivariate approach to predicting myoelectric control usability[J]. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2021, 29: 1312-1327.
- [20] PATEL G K, CASTELLINI C, HAHNE J M, et al. A classification method for myoelectric control of hand prostheses inspired by muscle coordination[J]. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2018, 26(9): 45-55.
- [21] LI Z Y, ZHAO X G, LIU G J, et al. Electrode shifts estimation and adaptive correction for improving robustness of sEMG-based recognition [J]. IEEE Journal of Biomedical and Health Informatics, 2021, 25(4): 1-10.

## 作者简介

李思诺, 硕士研究生, 主要研究方向为多模感知。

E-mail: 2681541101@qq.com

耿艳娟(通信作者), 博士, 副研究员, 主要研究方向为运动感觉功能康复技术与机制。

E-mail: yj.geng@siat.ac.cn