

DOI: 10.19650/j.cnki.cjsi.J2614967

基于多模态传感融合的人体下肢肌肉力预测及其 在外骨骼控制中的应用方法研究*

刘 宸^{1,2,3}, 包瑞来^{3,4}, 姚 杰³, 张廷丰⁴, 张 弼³

(1. 天津理工大学天津市先进机电系统设计与智能控制重点实验室 天津 300384; 2. 机电工程国家级实验教学示范中心
(天津理工大学) 天津 300384; 3. 中国科学院沈阳自动化研究所机器人与智能系统全国重点实验室 沈阳 110016;
4. 辽宁工业大学电气工程学院 锦州 121001)

摘 要: 人体肌肉力的实时准确感知是实现下肢外骨骼“按需辅助”控制的核心前提。针对外骨骼在肌肉疲劳、皮肤出汗等工作条件下基于传统表面肌电信号(sEMG)的运动意图识别易失效的局限,提出一种基于深度视觉与惯性测量单元(IMU)多模态融合的非侵入式下肢肌肉力预测及控制方法。首先,构建了5名健康受试者在平地及不同坡度上下坡等5种典型地形下的视觉与运动学多模态数据集,并基于sEMG-Hill肌肉模型离线计算股四头肌和腘绳肌内部肌肉力作为监督真值。其次,设计了一种双分支时空特征融合深层网络,通过并行通道分别提取深度图像序列的空间几何特征与惯性运动数据的时序动态,并引入自适应注意力机制进行特征智能异构加权融合,实现在实际应用阶段完全摆脱对表面肌电传感器的依赖。为实现从感知到控制的闭环转化,构建了由预测肌肉力实时驱动的自适应阻抗控制架构,将肌肉力变化连续映射为外骨骼的辅助刚度。实验结果表明,该方法预测精度显著优于单一模态方案,跨受试者平均决定系数(R^2)超过0.81,均方根误差(RMSE)优于6 N;在模拟肌肉疲劳干扰下预测曲线形态高度一致,表现出极强的稳健性。最后,在商用下肢外骨骼系统上完成平台验证,实现高时间相位同步与低横向漂移,为复杂动态环境下外骨骼高鲁棒按需辅助控制提供了新路径。

关键词: 下肢助力外骨骼;肌肉力预测;按需辅助;多模态融合

中图分类号: TH-39 TP242 **文献标识码:** A **国家标准学科分类代码:** 510.40

Research on the prediction of lower limb muscle forces based on multimodal sensor fusion and its application methods in exoskeleton control

Liu Chen^{1,2,3}, Bao Ruilai^{3,4}, Yao Jie³, Zhang Tingfeng⁴, Zhang Bi³

(1. Tianjin Key Laboratory for Advanced Mechatronic System Design and Intelligent Control, School of Mechanical Engineering, Tianjin University of Technology, Tianjin 300384, China; 2. National Demonstration Center for Experimental Mechanical and Electrical Engineering Education, Tianjin University of Technology, Tianjin 300384, China;
3. State Key Laboratory of Robotics and Intelligent Systems, Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang 110016, China; 4. College of Electrical Engineering, Liaoning University of Technology, Jinzhou 121001, China)

Abstract: Accurate real-time perception of human muscle force is a core prerequisite for achieving "assistance-on-demand" control in lower-limb exoskeletons. To address the vulnerability of traditional surface electromyography (sEMG)-based motion intention recognition to failure under challenging real-world conditions such as muscle fatigue, electrode displacement, and skin perspiration, this paper proposes a non-invasive lower-limb muscle force prediction and control method based on the multimodal fusion of depth vision and inertial measurement unit (IMU). First, a multimodal visual and kinematic dataset was constructed involving five healthy subjects across five typical daily terrains, including level ground and sloped walking at various gradients. The internal muscle forces of the quadriceps and hamstrings were calculated offline using an sEMG-driven Hill-type musculoskeletal model to serve as the supervised ground truth.

收稿日期: 2026-02-26 Received Date: 2026-02-26

* 基金项目: 国家自然科学基金面上项目(62473361)、辽宁省优秀青年科学基金项目(2025JH6/101000028)、机器人学国家重点实验室开放课题项目(2024-O17)资助

Second, a dual-branch spatiotemporal feature fusion deep network was designed, which extracts spatial-geometric features from depth image sequences and temporal dynamics from inertial motion data through parallel channels. By incorporating an adaptive attention mechanism for the intelligent, heterogeneous weighted fusion of these features, the proposed framework completely eliminates the reliance on sEMG sensors during the practical application phase, thereby bypassing signal degradation issues. To bridge the gap from perception to closed-loop control, an adaptive impedance control framework driven by real-time predicted muscle force was established, continuously mapping variations in muscle force to the assistance stiffness of the exoskeleton. Experimental results demonstrate that the proposed method significantly outperforms unimodal approaches, yielding a robust cross-subject average coefficient of determination (R^2) exceeding 0.81 and a root-mean-square error (RMSE) of less than 6 N. Furthermore, the predicted force curves maintain highly consistent profiles under simulated muscle fatigue interference, exhibiting superior robustness against physiological disturbances. Finally, platform validation on a commercial lower-limb exoskeleton system achieved high temporal phase synchronization and low lateral drift, providing a promising and viable new paradigm for the highly robust, assistance-on-demand control of exoskeletons in complex, dynamic environments.

Keywords: lower-limb assistive exoskeleton; muscle force prediction; assist-as-needed; multi-modal fusion

0 引言

随着全球人口老龄化进程的加速以及脑卒中等神经系统疾病发病率的攀升,针对下肢运动功能障碍的康复需求呈现爆发式增长。外骨骼机器人作为一种能够辅助人体运动并提供康复训练的机电一体化系统,已成为现代康复医学与机器人技术交叉领域的研究热点^[1]。现代康复理论强调,患者的主动参与对神经可塑性重塑至关重要。为此,能够依据使用者实时运动意图动态调整辅助力的“按需辅助”(assist-as-needed, AAN)控制策略,已取代传统的轨迹跟踪,成为外骨骼控制的主流^[2]。AAN策略成功的关键,在于实时、准确地估计人体肌肉力,从而使外骨骼的辅助力矩与穿戴者的主动运动意图精确同步。因此,如何构建一个在复杂动态环境下依然稳健的肌肉力感知系统,已成为当前外骨骼人机交互领域的核心科学挑战。

为了获取人体内部的肌肉力信息,现有的研究主要分为基于生物力学模型的方法^[3-4]和基于数据驱动的方法^[5-7]两大类。基于生物力学模型的方法,如经典的Hill肌肉模型,通过建立肌肉骨骼系统的数学方程来解析肌力。尽管该方法具有明确的物理意义,但其复杂的参数(如肌腱松弛长度、最佳肌纤维长度等)存在显著个体差异且难以在体精确标定,加之逆动力学求解的巨大计算开销,限制了其在线应用。相比之下,基于信号的数据驱动方法因其端到端的映射能力而备受青睐。表面肌电信号(surface electromyography, sEMG)被广泛视为在受控实验室环境下反映肌肉激活状态的“金标准”^[8-10]。然而,在实际的临床应用与日常穿戴场景中,sEMG传感器的局限性逐渐暴露。其信号不仅微弱且易受电磁噪声干扰,更严重的是,它对采集条件高度敏感:运动出汗导致的皮肤阻抗变化会引起信号基线漂移^[11-12];而在外骨骼穿戴过程中,设备的微小滑动引起的电极位置偏移,则可

能导致信号特征分布剧变,进而使预训练模型失效^[13]。这些非理想因素严重制约了基于sEMG的力预测系统在长时间、非实验室环境下的鲁棒性。

为克服表面肌电信号的固有局限,研究者正积极探索更为稳健的感知手段^[14]。一个极具前景的方向是从肢体的外部运动学反推内部驱动力。从根本上说,肢体的运动学状态(位置、速度、加速度)是内部肌肉力作用于骨骼系统的外在宏观表现。依据刚体动力学原理(如牛顿-欧拉方程),这些运动学参数与关节驱动力矩之间存在着明确的物理映射关系,这为绕开sEMG的生物电化学不确定性,从物理层面推断内部驱动力提供了根本性的理论基础。计算机视觉与惯性测量单元(inertial measurement unit, IMU)的融合,为精确捕捉完整的运动学链条提供了理想的解决方案。其中,视觉传感器以其全局视场和非接触式测量的优势,能够获取高精度的实时关节角度与肢体空间姿态,有效校准IMU在长时间积分过程中产生的累积漂移误差^[15-16];而IMU则凭借其高频采样与对动态变化的高敏感性,能够精准捕捉肢体的瞬时加速度与角速度,这是推算瞬时力矩的关键输入^[17-18]。与sEMG依赖经验映射的“黑箱”模型相比,这种基于物理约束的方法对电磁干扰、汗液积聚等外部环境变化天然免疫,其物理可解释性也使其在面对肌肉疲劳(肌肉激活效率变化但动力学关系不变)等生理状态变化时,具有更强的泛化潜力,为外骨骼机器人实现精准、可靠的按需辅助控制提供了新的技术路径^[19]。

然而,尽管前景广阔,但现有的视觉-惯性融合方法大多依赖于简单的特征拼接,未能充分挖掘跨模态数据在时序上的深度依赖关系和内在的物理一致性。当面对康复训练中常见的步态模式转变或非周期性动作时,这些简单的融合策略难以捕捉跨模态特征间的动态演变,导致预测精度下降^[20]。此外,如何将先验的动力学约束有效嵌入到深度学习网络中,以增强模型的泛化性和物理可解释性,仍是一个开放性问题。

针对上述挑战,本文提出了一种面向非结构化环境的视觉与惯性多模态肌肉力鲁棒预测方法。本研究的核心思想在于:在离线模型训练阶段,充分利用受控实验室环境下高精度 sEMG-Hill 模型提供的肌肉力作为监督真值,使深度学习网络能够精确捕捉“外部运动学特征到内部肌肉力”的动力学映射关系。而在外骨骼实际部署与运行阶段,则完全剥离对 sEMG 传感器的依赖,仅凭借具备环境鲁棒性的视觉与惯性测量单元数据实现肌肉力的在线预测。

为实现这一目标,本文提出了一种基于视觉与惯性信息的双分支时空特征融合网络 (dual-branch feature fusion network, DBFFNet)。该网络通过并行的专用通道分别提取深度图像序列的空间几何特征与 IMU 数据的肢体运动时序动态。特别地,通过引入自适应加权机制, DBFFNet 能够自适应地评估并加权不同模态在时间维度上的贡献,有效解决异构模态在复杂动态场景下信噪比波动和局部失效的问题,实现更为稳健的特征融合。最终,该网络旨在建立一种从外部运动表征到内部肌肉力变化的端到端高精度映射模型,以显著提升模型对动态环境及个体差异的适应能力。

本文的主要贡献可概括为 3 个方面:

1) 构建了一个多地形、长时序下肢肌肉力预测的视觉-运动学多模态数据集。该数据集涵盖 5 名受试者在 5 种典型地形工况 (平地、 $\pm 5^\circ$ 及 $\pm 10^\circ$ 坡度) 下的行走数据,为解决该领域缺乏公开标准化测试基准的问题提供了重要数据支撑,并为研究模型在复杂场景下的泛化能力奠定了基础。

2) 提出了一种双分支时空特征融合网络。该方法摒弃了对不稳定的 sEMG 信号的依赖,创新性地通过端到端的方法学习并建立了从外部运动学到内部肌肉力的鲁棒映射,有效解决了在皮肤汗液分泌、电极偏移及肌肉疲劳等非理想条件下模型鲁棒性差的难题。

3) 系统性地验证了所提方法在跨地形和抗肌肉疲劳等多个维度上的优越性与泛化能力。实验表明,模型在复杂工况下实现了强鲁棒的肌肉力预测,平均决定系数 (coefficient of determination, R^2) 达到了 0.81 以上,均方根误差 (root mean square error, RMSE) 控制在 6.0 N 以内,证明视觉-运动学融合路径是实现外骨骼精准按需辅助控制的一条可行高效的技术路线。

1 实验系统设计与多模态数据集构建

1.1 实验平台与数据采集

本研究设计了一套轻量级、非侵入式异构多模态数据采集平台,如图 1 所示。该平台旨在同步捕捉人体下肢运动时的外部运动学信息与周围环境的几何信息,其

核心设计思想在于通过视觉与惯性测量的优势互补,为肌肉力预测模型提供丰富且鲁棒的输入特征。



图 1 实验设备
Fig. 1 Experimental equipment

视觉感知模块的核心为一台主动立体红外深度相机 (Intel RealSense D435i)。相较于易受光照变化影响的 RGB 相机,深度相机能直接获取场景的三维几何信息,为模型提供了稳定的环境上下文与肢体形态特征。相机被固定于受试者胸前的可穿戴支架上,模拟外骨骼的第一人称视角。经预实验标定,相机俯仰角设定为 -25° , 以确保其视场能完整覆盖下肢运动范围。深度图像以 640×480 分辨率, 30 fps 帧率进行采集,以平衡图像特征的丰富度与后续深度网络的计算负载。运动学感知模块由一个包含 5 个高精度惯性测量单元 (IMU, Noraxon Ultium) 的无线传感器网络构成。基于生物力学原则, 1 个基准 IMU 固定于受试者腰椎中心,用于解算躯干姿态并建立以骨盆为中心的全局参考坐标系;其余 4 个 IMU 对称布置于双侧大腿及小腿外侧,用以实时获取肢体节段的运动信息并重构下肢运动链。数据通过专用无线协议以 2 000 Hz 的高采样率传输,确保了对足地冲击等高频动态事件的精确捕捉。高精度时空同步是融合异构传感器的关键。本系统采用基于硬件触发的同步方案。通过一个中央同步单元 (MyoSync, Noraxon, USA) 接收来自视觉采集上位机的 TTL 方波信号,为所有传感器 (深度相机与 IMU 网络) 提供统一的时间基准。该硬件同步机制从根源上消除了由软件延迟或时钟漂移引起的异步问题,确保了每一帧视觉数据都与相应时刻的运动学数据精确对齐,为后续的特征融合奠定了坚实基础。

1.2 数据预处理

在数据质量增强方面,针对 RGB 图像在动态光照、阴影及快速运动模糊下特征提取不稳定的问题,本研究仅利用深度图像作为视觉输入,以利用其光照不变性特

征。对于像素位置 (u, v) , 其对应的深度值可表示为:

$$Z(u, v) = \frac{f \cdot B}{d(u, v)} \quad (1)$$

其中, f 和 B 分别为相机焦距以及投影器与相机之间的基线距离, $d(u, v)$ 则表示结构光在相机像平面上的视差。

然而, 在物体表面反射率低、存在遮挡或结构光条纹匹配失败时, 视差 d 无法有效估计, 此时深度值被置为 0, 从而在深度图中形成深度空洞区域。

为解决数据空洞问题, 本文首先采用拉普拉斯插值算法对无效区域进行重构, 如式 (2) 所示, 该方法利用邻域像素的梯度信息生成平滑且具有几何连续性的深度估计结果。随后, 引入双边滤波器进行全图平滑处理, 如式 (3) 所示, 在有效抑制高频噪声的同时, 最大限度地保留肢体轮廓的边缘高频信息, 从而避免运动特征的模糊化。

在需要填充的区域 Ω , 拉普拉斯方程表示为:

$$\begin{cases} \nabla^2 I(X) = 0, & X \in \Omega \\ \nabla^2 = \frac{\partial^2}{\partial l} + \frac{\partial^2}{\partial l} \end{cases} \quad (2)$$

其中, $I(X)$ 为图像在位置 $X = (x, y)$ 处的深度值, ∇^2 为拉普拉斯算子。

双边滤波公式定义为:

$$I_f(x, y) = \frac{1}{W_p(i, j) \in \Psi} \sum I(i, j) \cdot w_s(i, j, x, y) \cdot w_r(i, j, x, y) \quad (3)$$

其中, Ψ 为以 (x, y) 为中心的滤波窗口, w_s 为空域权重, w_r 为值域权重, W_p 为归一化因子。

对于 IMU 采集的关节角度数据, 本研究在数据采集阶段引入了标准化的在线运动学预处理流程。具体而言, 基于惯性传感器测得的角速度、加速度及磁场信息, 系统通过多传感器姿态融合算法估计关节段的空间姿态:

$$\hat{R}(t) = F(w(t), a(t), m(t)) \quad (4)$$

其中, $w(t)$ 、 $a(t)$ 和 $m(t)$ 分别表示角速度、线加速度和磁场测量值。随后将姿态矩阵转换为欧拉角形式:

$$[\phi(t), \theta(t), \psi(t)]^T = \text{Euler}_{ZYX}(\hat{R}(t)) \quad (5)$$

其中, $\phi(t)$ 、 $\theta(t)$ 和 $\psi(t)$ 分别表示翻滚角 (Roll)、俯仰角 (Pitch) 和偏航角 (Yaw)。考虑到人体运动的连续性与传感器的测量噪声, 进一步引入在线低通滤波算子 $\mathcal{L}(\cdot)$ 对原始角度信号 $\theta(t)$ 进行去噪处理:

$$\tilde{\theta}(t) = \mathcal{L}(\theta(t)) \quad (6)$$

其中, $\tilde{\theta}(t)$ 表示去噪后的平滑关节角度序列。

上述预处理流程在数据采集过程中实时完成, 确保输出的关节角度序列在保证信噪比的同时, 最大程度保留运动过程中的瞬时动态特性。

1.3 数据归一化

两种模态数据在经过上述预处理后, 虽然在时间维度上实现了对齐, 但由于图像特征与运动学特征在物理量纲及数值动态范围上存在显著差异, 无法直接输入神经网络模型进行训练。为此, 基于不同模态数据的统计特性与物理约束, 本研究设计了差异化的归一化方案:

1) IMU 数据: Z-score 标准化

关节角度数据反映了肢体的动态摆动幅度, 其数值分布通常近似服从高斯分布, 且无明确的硬性物理边界。为了将数据中心化并统一尺度, 使其更适配网络对时序输入的敏感性, 本研究采用 Z-score 标准化方法。对于 t 时刻运动学特征 $\tilde{\theta}_t$ 其标准化计算为:

$$\theta_z(t) = \frac{\tilde{\theta}(t) - \mu_{\tilde{\theta}}}{\sigma_{\tilde{\theta}}} \quad (7)$$

其中, $\mu_{\tilde{\theta}}$ 和 $\sigma_{\tilde{\theta}}$ 为关节角度的均值和标准差。

2) 视觉数据: Min-Max 归一化

深度图像的像素值表示相机坐标系下的物理距离, 其有效值受到传感器探测范围的约束。为在保留深度图像相对几何结构信息的同时, 将其数值映射至神经网络的有效激活区间, 本研究对每幅深度图像的有效像素区域采用 Min-Max 线性归一化处理, 定义为:

$$\tilde{D}(x, y) = \frac{D(x, y) - D_{\min}}{D_{\max}} \quad (8)$$

其中, $D_{\min}$ 和 $D_{\max}$ 分别表示当前深度图像中有效深度像素的最小值与最大值。

1.4 基于 sEMG-Hill 模型的肌肉力标签生成

为了实现对多模态融合网络的监督训练, 获取准确的肌肉力作参考真值至关重要。鉴于直接测量人体内部肌肉力具有极高的侵入性与操作难度, 本研究采用生物力学领域广泛认可的 sEMG 驱动的 Hill 型肌肉模型^[21]来估计肌肉力, 并将其仅作为模型训练阶段的监督标签。

1) sEMG 信号采集与预处理

本研究主要关注控制膝、髋关节屈伸运动的核心肌群。本研究在受试者大腿表面选定 5 块目标肌肉进行 sEMG 采集: 包括股四头肌群 (股直肌、股内侧肌、股外侧肌) 和腓绳肌群 (半腱肌、股二头肌)。使用无线肌电信号传感系统 (Ultium EMG, Noraxon, USA) 以 2 000 Hz 的采样频率记录肌肉激活变化。表面肌电传感器和 Ag/AgCl 电极贴通过双面胶固定于皮肤表面。所有电极贴片贴放位置均依据 SENIAM^[22] 国际标准进行定位, 以保障信号的一致性与可比性。

sEMG 作为一种微弱的生物电信号, 其幅值通常处于微伏 (μV) 级别, 采集过程中易受环境电磁噪声及运动伪影的干扰。因此, 在进一步分析前必须对其进行有效的预处理, 以抑制低频漂移和高频噪声^[23-25]。基于频谱

分析,sEMG的有效生理信息主要集中于20~500 Hz频带。为此,本研究采用四阶Butterworth带通滤波器以消除基线漂移与高频干扰,并配合50 Hz陷波滤波器抑制工频干扰。为提取反映肌肉激活程度的包络信息,对预处理后的信号进行全波整流,并施加截止频率为6 Hz的低通滤波器,从而得到平滑的线性包络。该包络能够清晰表征运动过程中肌肉活动的动态变化趋势^[12]。最后,对包络信号进行归一化处理,将其转换为无量纲的神经肌肉激活度指标 $e(t) \in [0,1]$ 。

2) Hill模型力学解算

肌肉的实际出力不仅取决于激活度 $e(t)$,还受肌肉-肌腱单元的瞬时长度和收缩速度影响^[26-27]。如图2所示的神经肌肉骨骼模型,Hill模型精确地描述了这一复杂关系。

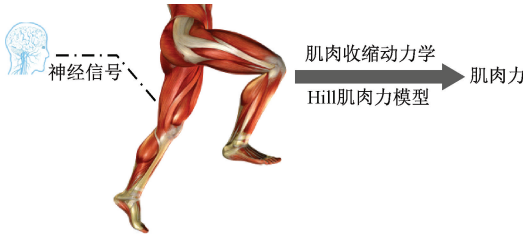


图2 下肢神经肌肉骨骼模型

Fig. 2 Lower limb neuromusculoskeletal model

为了从激活度 $e(t)$ 推导出具有生物力学意义的肌肉力 F^T ,本研究利用Hill肌肉力模型进行计算。Hill肌肉力模型是HILL通过大量的医学数据与实验应用验证的模型,其模型结构如图3所示,图3中CE表示主动收缩单元,PE表示被动弹性单元, α 表示肌肉纤维与腱的夹角, $l^M \cos \alpha$ 表示肌纤维长度, l^T 表示腱长度, l^{MT} 表示肌腱单元长度, F^M 表示肌纤维力, F^T 表示肌腱单元力。

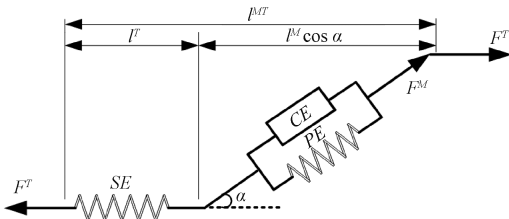


图3 Hill肌肉理论模型

Fig. 3 Hill muscle model

其数学表达形式为:

$$\begin{aligned} F^T &= F_{CE}^M + F_{PE}^M \\ F_{CE}^M &= e \cdot F_{\max} \cdot f(l) \cdot f(v) \\ F_{PE}^M &= F_{\max} \cdot f(l) \end{aligned} \quad (9)$$

其中, $f(l)$ 和 $f(v)$ 为生理参数关系,可以通过数据拟合, l 、 v 为实际生理参数均为常值, F_{CE}^M 为肌纤维收缩单

元产生的力, $F_{\max}$ 为最大等长收缩力,而 F_{PE}^M 为表征肌纤维及周围结缔组织被动产生的力。

本研究中Hill型肌肉模型参数的选取遵循生物力学领域的通用范式。在理论架构层面,采用Zajac^[28]提出的肌肉-肌腱单元无量纲化模型框架,确立了参数归一化的理论基础。在形态学参数选取上,参考Delp等^[29]建立的下肢肌肉骨骼模型数据库。该数据库涵盖了最佳肌纤维长度、肌腱松弛长度及羽状角等关键结构参数,是国际主流仿真平台OpenSim的核心数据源,具有广泛的代表性。

既往研究已证实了“通用参数+个体缩放”策略的有效性。Lloyd等^[27]的研究表明,即便在未进行肌电标定的条件下,基于文献经验参数并结合个体特征进行尺度缩放,仍能准确重建关节力矩的动态变化趋势。基于此,本研究以Delp数据库为基准,依据受试者的身高与体重数据,对最佳肌纤维长度及肌腱松弛长度进行几何相似性缩放。此外,为兼顾模型数值稳定性与个体生理差异,肌肉最大等长收缩力 $F_{\max}$ 基于生理横截面积计算,其中肌肉比张力统一设定为50 N/cm²。

2 肌肉力预测模型构建与应用

为实现外骨骼在复杂康复场景下的精准按需辅助,本研究构建了一套从多模态感知到柔顺控制的闭环人机交互框架,如图4所示。

该框架的核心挑战在于:如何有效融合物理属性与时空特性差异显著的视觉和惯性数据,构建一个从外部运动学到内部肌肉力的鲁棒映射模型;以及如何将该预测结果转化为外骨骼的实时、柔顺辅助力矩。为此,本章首先详细阐述了所提出的DBFFNet的架构设计,随后介绍了如何将网络预测的肌肉力嵌入到自适应阻抗控制策略中,最终实现对外骨骼的智能驱动。

2.1 双分支特征融合网络架构

针对多模态数据在物理属性与信息维度上的显著异构性,数据在输入时直接拼接往往会导致高维特征被低维特征掩盖,难以有效捕捉各模态的规律。为此,本研究提出了一种双分支特征融合网络架构,如图5所示。

该网络通过并行的专用处理通路,分别深度挖掘各模态的独特信息,再通过自适应加权机制进行智能融合,最终实现端到端的肌肉力预测。整体网络由视觉特征提取分支、时序特征提取分支、多模态注意力融合模块及肌肉力回归头4个核心部分构成。

1) 视觉特征提取分支

视觉分支的核心任务是将高维的深度图像序列映射为紧凑的空间几何特征向量,以表征下肢在不同步态周期中的肢体构型与环境交互状态。鉴于外骨骼控制系统

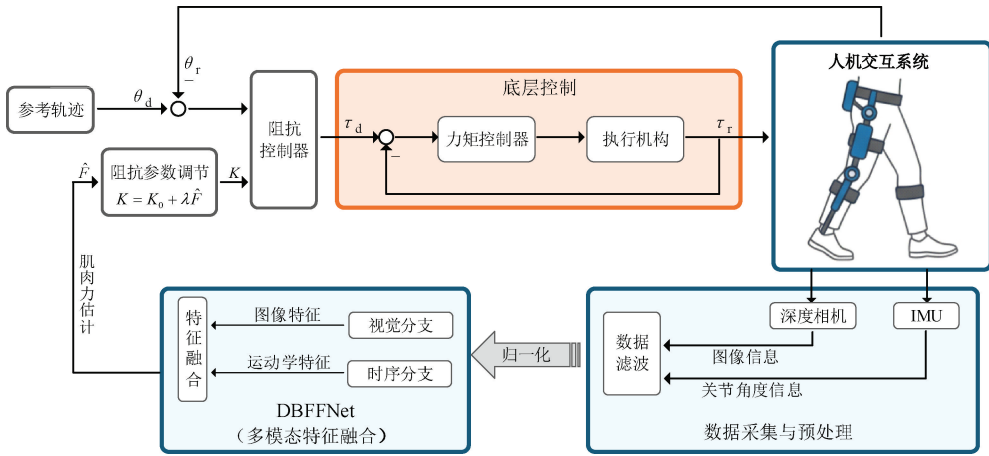


图 4 基于多模态融合的外骨骼人机控制系统架构

Fig. 4 Architecture of exoskeleton human-robot control system based on multimodal fusion

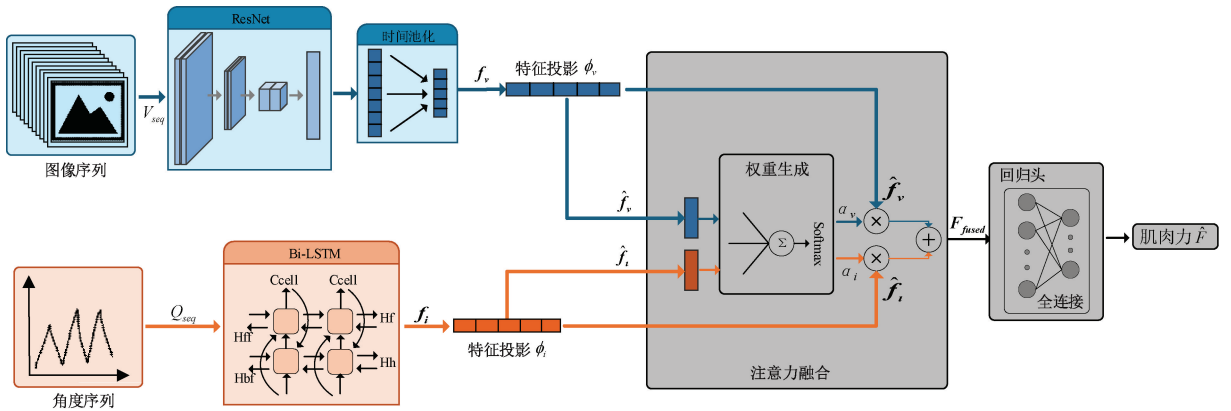


图 5 双分支特征融合网络架构

Fig. 5 Dual-branch feature fusion network architecture

对实时性的严苛要求,在平衡特征表达能力与计算推理延迟的基础上,本文构建了基于改进 ResNet 架构的轻量化视觉特征提取网络。该网络以 ResNet-18 为骨干,保留了其核心的残差连接机制^[30],有效克服了深层网络的梯度消失问题。同时,针对嵌入式计算平台的算力约束进行了结构适配,更契合外骨骼系统对低延迟响应的实际需求。数据处理流程为:首先,对深度图像序列 $V_{seq} = \{V_{t-L+1}, V_{t-L+2}, \dots, V_t\}$ 的深度像素值执行归一化处理,基于图像中心进行中心裁剪,统一调整至 224×224 的输入尺寸。随后,利用预训练的 ResNet-18 网络作为主干特征提取器对图像序列进行编码。为了本文的连续回归任务需求,本研究对原始网络结构进行了调整,移除了用于图像分类的末端全连接层 (fully connected layer, FC) 及 Softmax 层,仅保留卷积层与全局平均池化层 (global average pooling, GAP)。经过 GAP 层处理后,输出特征图被压缩为 512 维的空间特征向量 f_v ,该向量包含了当前时刻下肢的空间姿态信息与环境几何信息,为后续的

多模态融合提供了空间上下文基准。

针对本研究中受试者数据样本量有限的问题,本研究采用了迁移学习策略对网络参数进行初始化,避免模型从零训练导致的过拟合及收敛困难的问题^[30-31]。具体而言,ResNet-18 骨干网络的卷积层权重被初始化为在 ImageNet 大规模数据集上预训练模型参数,利用其学到的通用底层视觉特征 (如边缘、纹理) 来加速特征提取能力的构建。在训练阶段,解冻所有卷积层参数,并采用 Adam 优化器以较小的学习率 (1×10^{-3}) 进行全局微调,从而使网络自适应地学习特定于下肢深度图像的特征表示。

2) 时序特征提取分支

与视觉分支侧重于空间上下文的捕获不同,时序特征提取分支是建立从运动学序列到肌肉响应之间的非线性映射。利用 IMU 数据的高频特性,分析肢体运动趋势、速度变化及步态相位的时序动态特征。

考虑到人体运动具有显著的时间连续性与滞后效

应,单纯的前馈网络难以处理时间依赖关系。本研究采用多层堆叠长短期记忆网络(stacked long short term memory, Stacked LSTM)作为特征提取器。LSTM单元内部独特的门控机制(遗忘门、输入门、输出门)能够有效捕捉长距离的时间依赖关系,模拟肌肉收缩过程中的机电延迟特性。该分支的输入数据定义为与视觉帧严格时间对齐的关节角度序列 $\mathbf{Q}_{seq} = \{\mathbf{q}_{t-L+1}, \mathbf{q}_{t-L+2}, \dots, \mathbf{q}_t\}$,其中 $\mathbf{q}_i \in \mathbb{R}^N$ 表示 i 时刻时关节角度, $L=66$ 为滑动时间窗口长度(对应33 ms的历史信息)。相较于原始的加速度计读数,使用解算后的关节角度 $\hat{\theta}(t)$ 作为输入具有更明确的物理意义,且降低了网络学习肢体几何约束的难度。

针对网络超参数的确定,本研究采用基于验证集的网格搜索策略^[32-33]进行参数寻优,在隐藏层维度{64, 128, 256}与堆叠层数{1, 2, 3}的组合空间内进行遍历测试。实验结果表明,当网络配置为双层堆叠结构且单层隐藏单元数为128时,模型在拟合能力与泛化性能之间取得了最佳平衡,同时有效避免了梯度消失问题。此外,为防止过拟合,在两层LSTM之间引入了比率为0.5的Dropout正则化层。

LSTM通过对关节角度序列 $\mathbf{Q}_{seq}$ 的递归更新,能够隐式地编码关节角速度与角加速度信息,从而重构出肌肉的收缩速度与爆发力特征。这种对时序动态的敏锐捕捉,使得模型能够精准识别步态切换瞬间的力矩突变,有效弥补了视觉模态帧率受限短板。最终,网络输出最后一个时间步的时序特征向量 $\mathbf{f}_i \in \mathbb{R}^{128}$,参与后续的多模态特征融合。

3) 特征融合

尽管视觉与时序分支分别提取了高维的空间与时序特征,但这两类特征向量在维度空间上存在显著差异。此外,在非结构化环境行走过程中,各模态的有效性呈现动态变化。为此,本研究在两个分支的末尾处引入一个注意力融合模块将异构特征映射至统一的语义子空间,并根据模态的实时可靠性动态分配融合权重。该模块主要包含特征对齐与权重生成两个并行计算过程,即:

(1) 特征对齐

由于视觉特征 $\mathbf{f}_v \in \mathbb{R}^{512}$ 与时序特征 $\mathbf{f}_i \in \mathbb{R}^{128}$ 维度不一致,直接拼接容易导致高维特征在后续处理中主导融合过程,且难以支持更深层次的特征交互。为了在融合前实现模态间的语义对齐,并为后续基于注意力机制的运算提供统一的度量空间,本研究引入两个独立的特征映射结构 ϕ_v 和 ϕ_i 分别将视觉特征向量 $\mathbf{f}_v$ 与时序特征向量 $\mathbf{f}_i$ 投影至同一维度。

$$\begin{cases} \hat{\mathbf{f}}_v = \phi_v(\mathbf{f}_v) = \text{ReLU}(\mathbf{W}_v \mathbf{f}_v + \mathbf{b}_v) \\ \hat{\mathbf{f}}_i = \phi_i(\mathbf{f}_i) = \text{ReLU}(\mathbf{W}_i \mathbf{f}_i + \mathbf{b}_i) \end{cases} \quad (10)$$

其中, $\hat{\mathbf{f}}_v, \hat{\mathbf{f}}_i \in \mathbb{R}^{128}$ 为对齐后的特征向量。针对映射

矩阵 $\mathbf{W}_d$ 与 $\mathbf{W}_i$ 的参数初始化,本研究采用He等^[30]初始化方法,以确保在ReLU激活函数作用下网络深层的梯度流动保持稳定,加速融合层的收敛。

(2) 加权融合

为了捕捉不同工况下各模态的相对重要性,受挤压与激励(squeeze-and-excitation, SE)网络通道注意力机制的启发^[34],本研究构建了一个基于多层感知机(multilayer perceptron, MLP)的自适应注意力权重生成模块,将对齐后的特征拼接为联合上下文向量 $\mathbf{C} = [\hat{\mathbf{f}}_v, \hat{\mathbf{f}}_i]$ 输入至MLP网络生成注意力权重。网络的结构参数配置通过基于验证集的网格搜索策略确定。本研究通过对比不同瓶颈层配置下的模型性能,最终确定隐藏层节点数为256时,模型在参数量与特征交互能力之间取得最佳平衡。

网络最终输出两个标量值,并通过Softmax函数归一化,得到自适应权重系数 $\alpha = [\alpha_v, \alpha_i]^T$,且满足 $\alpha_v + \alpha_i = 1$ 。这组系数代表了网络对于当前时刻各模态数据的置信度评估,融合特征 $\mathbf{F}_{fused}$ 通过式(11)计算。

$$\mathbf{F}_{fused} = \alpha_v \cdot \hat{\mathbf{f}}_v + \alpha_i \cdot \hat{\mathbf{f}}_i \quad (11)$$

通过这种加权求和机制,模型能够自适应地抑制受干扰模态的噪声影响,强化高置信度模态的贡献,从而显著提升复杂环境下肌肉力($\hat{F}$)预测的鲁棒性。

2.2 执行与控制

在康复外骨骼辅助训练中,如何在提供有效辅助的同时避免过度干预,是实现安全、舒适人机交互的关键问题。阻抗控制通过在力与运动之间引入虚拟的机械阻抗关系,使外骨骼机器人表现出类似生物关节的柔顺特性,已被广泛应用于人机协作与康复机器人系统中^[35]。因此,本文在多模态肌肉力感知的基础上,引入肌肉力驱动的阻抗控制策略,将高层运动意图感知结果映射至外骨骼控制参数,实现按需辅助控制。

在人-外骨骼控制系统中,阻抗控制的核心思想是将外骨骼等同于物理控制系统,人是为导纳,将外骨骼带动患者行走的动作等同于阻抗^[36]。阻抗控制的数学表达式如式(12)所示。

$$\mathbf{M}_d(\ddot{\mathbf{q}} - \ddot{\mathbf{q}}_d) + \mathbf{B}_d(\dot{\mathbf{q}} - \dot{\mathbf{q}}_d) + \mathbf{K}_d(\mathbf{q} - \mathbf{q}_d) = \boldsymbol{\tau} \quad (12)$$

其中, $\boldsymbol{\tau}$ 为带动患者按预定步态行走的期望的关节力矩, $\mathbf{M}_d, \mathbf{B}_d, \mathbf{K}_d$ 分别为惯性系数、阻尼系数和刚度系数, $\mathbf{q}$ 和 $\mathbf{q}_d$ 分别表示实际关节角和期望关节角。

为使辅助行为能够随人体主动发力水平自适应变化,本文引入由双分支特征融合网络预测得到的肌肉力 $\hat{F}$,用于对阻抗参数进行调节。具体而言,系统根据肌肉力大小对关节刚度进行连续映射:

$$\mathbf{K} = \mathbf{K}_0 + \lambda \hat{F} \quad (13)$$

其中, K_0 为基础刚度项, λ 为调节系数, 用于控制肌肉力对辅助强度的影响程度。阻尼参数 B_d 则依据系统稳定性要求与经验设计原则进行设定, 以保证在不同运动状态下的平滑交互特性。

在外骨骼被动控制中, 关节力矩控制是实现位置控制的基础。在理想情况下, 电机电流与输出扭矩成线性关系。但在实际应用中, 当电流超过额定工作范围时, 受磁饱和、驱动器限幅等因素影响, 电流与输出力矩之间的线性关系将被破坏, 进而影响控制精度与系统稳定性。

基于上述考虑, 本文采用以电机电流为控制对象的闭环控制策略, 通过调节驱动电流实现对电机输出力矩 τ_r 的间接控制。进一步地, 考虑到外骨骼系统由多级机械连杆与传动机构构成, 其动力学特性中普遍存在摩擦、间隙及结构柔性等因素, 导致系统响应过程中出现明显的相位滞后。在此情况下, 引入积分环节可能引发积分饱和或放大系统滞后效应, 从而降低控制系统的稳定性。

因此, 本文在电流控制回路中舍弃积分项, 采用比例-微分 (proportional-derivative, PD) 控制器对电机电流进行调节, 其控制律可表示为:

$$u = k_p \left[(\tau_d - \tau_r) + t_d \frac{d(\tau_d - \tau_r)}{dt} \right] \quad (14)$$

其中 k_p 和 t_d 分别是控制器的比例系数和微分系数, τ_d 是期望关节力矩。

通过上述方法, 外骨骼系统能够在人体肌肉发力水平较高时降低辅助刚度, 鼓励使用者主动参与运动; 而在肌肉输出能力下降或疲劳加重时, 系统则相应提高辅助强度, 从而实现符合康复训练需求的按需辅助行为。值得注意的是, 该控制策略并未直接依赖肌电信号, 而是基于视觉与运动学信息构建的肌肉力估计结果, 避免了肌电信号在电极移位、汗液干扰及肌肉疲劳等非理想条件下带来的不稳定性。

3 实验与结果分析

为系统性地评估本文所提多模态肌肉力预测方法的有效性、鲁棒性及其在实际外骨骼系统中的应用潜力, 本章设计并开展了多组实验。

3.1 实验设计

1) 受试者与实验协议

本研究在中国科学院沈阳自动化研究所生物力学实验室进行, 并严格遵循《赫尔辛基宣言》的伦理准则^[37]。实验方案已获得中国科学院沈阳自动化研究所伦理委员会 (批复号: 2024-03) 与辽宁省人民医院伦理委员会 (批件号: 2023MSLH117) 的审核批准。研究共招募 5 名健康成年男性受试者, 其基本生理信息如表 1 所示。

表 1 受试者基本信息
Table 1 Subject's basic information

受试者编号	年龄/岁	身高/cm	体重/kg
S1	23	172	70
S2	25	181	80
S3	24	175	72
S4	23	178	78
S5	24	176	69
平均±标准差	23.8±0.8	176.4±3.4	73.8±4.9

在实验开始前, 所有受试者均已充分了解实验流程与潜在风险, 并签署了书面知情同意书。为优化数据采集条件, 减少衣物遮挡对视觉采集的干扰, 并确保表面肌电极与皮肤的良好接触, 受试者在实验过程中统一穿着定制短裤, 并对目标肌肉区域进行了备皮与酒精清洁处理。

实验在一台可精确调节坡度的 BerTec 步态分析跑台上进行, 旨在模拟日常行走中常见的地形变化。实验共设置 5 种工况: 水平地面 (0°)、小幅度上/下坡 (±5°) 以及较大幅度上/下坡 (±10°), 如图 6 所示。跑台速度恒定为 1.0 m/s。在每种工况下, 受试者首先进行 30 s 的适应性行走, 随后连续采集 120 s 的稳定步态数据。为避免肌肉疲劳对数据造成干扰, 各工况之间安排了 5 min 的静坐休息。

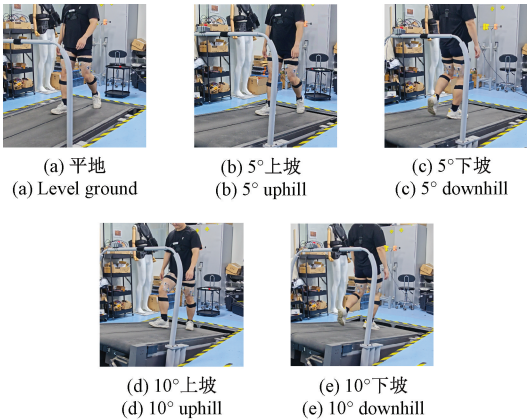


图 6 不同坡度行走实验设置

Fig. 6 Diagram of the experimental setup for walking on different slopes

2) 评价指标

为全面评估模型的预测性能, 本研究采用 R^2 和 RMSE 作为主要评价指标。决定系数衡量模型预测值与真实值之间的拟合优度, 其值越接近 1, 表明模型的解释能力越强。均方根误差衡量模型预测值与真实值之间绝对误差的大小, 单位与预测目标 (肌肉力, N) 相同, 其值

越小,表明预测精度越高。

3) 模型训练细节

本研究采用留一交叉验证策略来评估模型的泛化能力。具体而言,每次迭代中,将 1 名受试者的全部数据作为测试集,其余 4 名受试者的数据作为训练集,重复 5 次直至所有受试者均被测试一次。这种方法能有效评估模型对于未知个体的预测性能。

模型基于 PyTorch 框架实现,采用 Adam 优化器进行参数更新,初始学习率设置为 1×10^{-3} ,批处理大小为 16,共训练 100 个轮次。所有实验均在配备 NVIDIA GeForce RTX 4060 GPU 的服务器上完成。

3.2 结果与分析

1) 多模态融合有效性验证

为验证所提 Vision+IMU 多模态融合方法的优越性,本文设计与两种单一模态输入方法 (Vision-only, IMU-only) 的对比实验。在保持后端网络结构不变的情况下,仅改变输入数据源。图 7 展示了 3 种方法在不同工况下对股四头肌和腓绳肌的肌肉力预测性能 (R^2)。

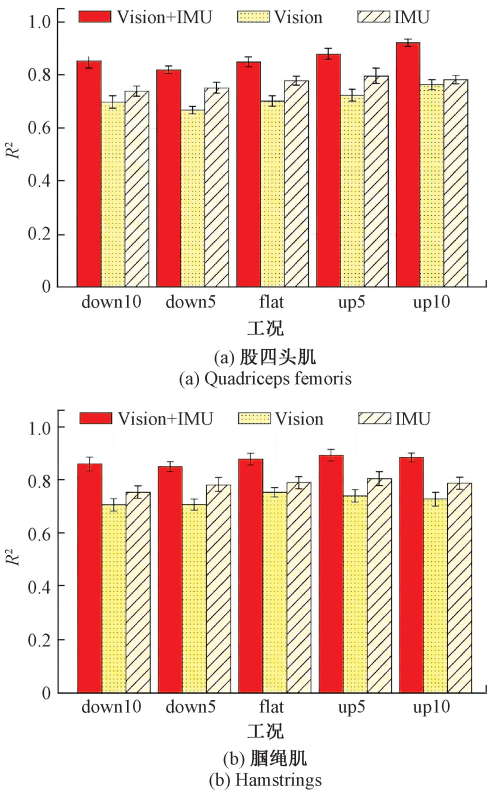


图 7 在不同工况下的跨受试者平均 R^2 对比

Fig. 7 Cross-subject average R^2 under different working conditions

从图 7 可以看出,无论对于股四头肌还是腓绳肌,本文提出的 Vision+IMU 多模态融合方法在所有工况下均表现最佳。Vision-only 模型性能最弱,其平均决定系数

R^2 在 0.65~0.76。这表明仅靠视觉信息虽能捕捉肢体姿态和环境几何特征,但缺乏对肢体动态特性的精细刻画,难以准确推断肌肉力的快速变化,尤其是在缺少地面反作用力这一关键信息时。IMU-only 模型性能居中,通过捕捉关节角度、角速度等运动学信息,IMU 数据能更好地反映肢体动力学,其平均 R^2 提升至 0.73~0.81。然而,IMU 易受积分漂移和噪声影响,在长时间或复杂动态任务中预测稳定性受限。Vision+IMU 模型性能最优,融合模型通过自适应加权机制有效结合了视觉提供的全局空间上下文信息与 IMU 提供的高频动态信息,实现了优势互补。其平均 R^2 稳定在 0.81 以上。这证明了多模态融合能够有效补偿单一模态的信息局限性,显著提升预测的准确性和稳定性。

图 8 展示了融合模型在不同工况下的 RMSE。股四头肌的预测误差在各工况下较为稳定,平均 RMSE 维持在 4.80~5.84 N。而腓绳肌的 RMSE 则呈现出明显的工况依赖性,上坡工况 (4.39~4.61 N) 显著高于平地和下坡工况 (2.58~2.88 N)。这主要是由于生物力学差异:在上坡的支撑相,腓绳肌作为主要的髋伸肌,需产生更大的爆发力来推动身体向上,导致其肌肉力峰值和动态范围远超平地行走。因此,尽管模型的拟合优度 (R^2) 依然很高 (>0.87),但 RMSE 会因基准力值的增大而相应增加。这一结果符合生物力学规律,也从侧面印证了模型的有效性。

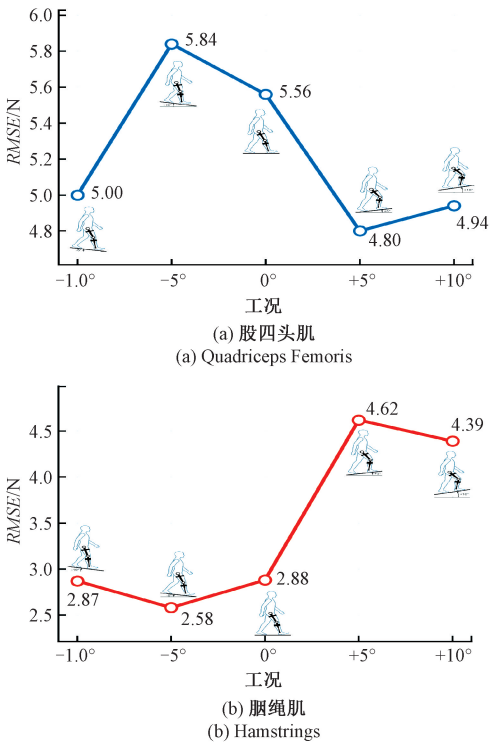


图 8 在不同工况下的跨受试者平均 RMSE

Fig. 8 Cross-subject average RMSE under different working conditions

为深入研究本文所提出的 DBFFNet 架构中双分支结构和自适应注意力融合机制的必要性,本文以腓绳肌群为例设计并构建了两个对照模型进行定量对比。

对照模型构建与实验设计:

(1) 对照模型 A(无注意力融合):基于 DBFFNet 的整体架构,但将原有的自适应注意力融合模块替换为常规的特征拼接,随后连接一个全连接层进行特征降维和融合。此模型旨在评估自适应注意力机制在模态融合中的关键作用。

(2) 对照模型 B(单分支网络):构建一个单一的 LSTM 网络。将深度图像的视觉特征和 IMU 的惯性特征在输入层进行常规拼接后,统一输入该 LSTM 网络进行时序建模和特征提取。此模型旨在验证双分支结构处理异构多模态数据的优越性。

在模型训练与评估阶段,本文沿用 3.1 节中描述的数据集、留一交叉验证策略、模型优化器及超参数设置,并采用 R^2 作为主要的性能评估指标,确保实验结果的公正性和可比性。为更直观地反映所提模型对剩余误差解释能力和预测精度的提升,本文进一步引入了相对剩余误差减少百分比作为对比度量指标,其计算公式为:

$$\Delta R^2(\%) = \frac{R^2_{DBFFNet} - R^2_{对照模型}}{1 - R^2_{对照模型}} \times 100\%$$

(15)

表 2 清晰展示了 DBFFNet 与两个对照模型在不同工况下对腓绳肌肌肉力预测的 R^2 性能对比结果。可以观察到,DBFFNet 在所有 5 种测试工况下均表现出最高的预测精度,其 R^2 值稳定在 0.85~0.91,显著优于对照模型 A 和模型 B,表明 DBFFNet 对腓绳肌肌肉力变化的解释能力极强且预测性能稳定。

表 2 不同模型在腓绳肌肌肉力预测中的 R^2 表现

Table 2 R^2 values of different models in hamstring muscle force prediction

工况	DBFFNet	对照 模型 A	对照 模型 B	较模型 A 提升/%	较模型 B 提升/%
Flat	0.91	0.88	0.89	25.00	18.18
Up 5°	0.87	0.76	0.79	45.83	38.10
Up 10°	0.85	0.81	0.79	21.05	28.57
Down 5°	0.88	0.83	0.82	29.41	33.33
Down 10°	0.87	0.79	0.80	38.10	35.00

具体而言,DBFFNet 相较于对照模型 A 展现出显著的性能优势。在所有工况下,DBFFNet 的预测性能(R^2)均优于对照模型 A,尤其是在爬坡(Up 5°)和下坡(Down 10°)等动态且复杂的运动模式下,DBFFNet 的 R^2 值相较对照模型 A 分别实现了高达 45.83% 和 38.10% 的显著

提升。这有力地证明了本文所提出的自适应注意力融合机制并非简单的特征拼接,而是能够根据不同模态特征的重要性及其对预测任务的动态贡献,智能地分配权重。注意力机制使得网络在不同运动状态下能够聚焦于更具判别性的模态信息,有效抑制了冗余或噪声信息,从而显著提高了肌肉力预测的准确性与鲁棒性。

DBFFNet 相较于对照模型 B 同样表现出一致的优越性,在所有测试工况下均展现出优于对照模型 B 的性能。例如,在下坡(Down 5°)工况中,其 R^2 值相较对照模型 B 提升了 33.33%。这一结果明确指出,简单地在输入层拼接视觉和 IMU 数据并由单一网络进行处理(如模型 B 可能采用的方式),无法充分捕捉两类异构数据固有的特点和内在联系。深度图像具有丰富的空间信息,其高维特征更适合由卷积神经网络进行高效提取;而 IMU 数据则承载着连续的时序动态信息,其模式识别更适合由循环神经网络(如 LSTM)进行建模。DBFFNet 的核心创新在于其双分支结构设计,该结构能够为每一种模态定制专属的特征提取器,使其能够充分学习各自模态的深层、高维和时序表征,从而避免了因强制统一处理而导致的特征信息损失或混淆,为后续的有效融合奠定了坚实基础,最终实现了在复杂运动工况下腓绳肌肌肉力预测的显著性能提升,为精准评估和干预提供了强有力的数据支持。

综上所述,本文量化证明了 DBFFNet 网络中所采用的双分支结构对于有效处理异构多模态数据的必要性,以及自适应注意力融合机制对于实现模态间信息高效、鲁棒整合的关键作用。这些核心设计元素共同赋予了 DBFFNet 在复杂运动工况下高精度、高鲁棒性地预测人体下肢肌肉力的能力,进一步验证了本研究方法在解决传统肌肉力感知挑战方面的优越性。

2) 模型性能对比与多模态融合机制解析

为更深入地揭示 DBFFNet 模型在复杂运动中精确预测肌肉力的内在机制,并直观展示其多模态融合的优越性与自适应性,本文进一步进行了预测波形精细化对比分析和注意力权重动态分配机制分析。

图 9 展示了一个完整步态周期内腓绳肌在平地行走工况下的真实肌肉力(sEMG + Hill 模型解算值)、DBFFNet 模型的预测值以及仅使用 IMU 模态的单分支模型预测值。

从图 9 中可以清晰观察到 3 条曲线的动态变化及其相对关系:首先,真实肌肉力波形呈现出从初期低谷、逐步上升至峰值,随后平缓下降的典型模式。本文所提 DBFFNet 模型的预测结果与真实肌肉力波形展现出高度的一致性。无论是在肌肉力上升、达到峰值,还是随后的下降阶段,DBFFNet 模型均能准确跟踪真实波形的趋势、峰值大小和时间点。充分证明了 DBFFNet 模型在捕捉肌肉力细微动态和瞬时变化方面的卓越能力,其预测结

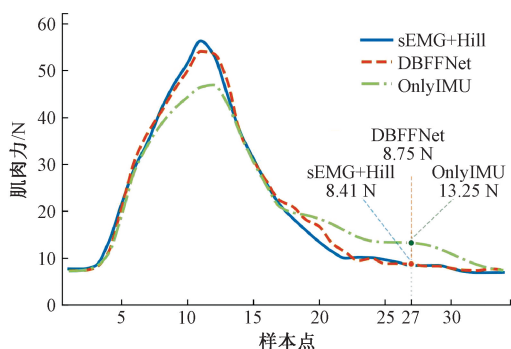


图 9 肌肉力预测波形对比

Fig. 9 Muscle force prediction waveform comparison

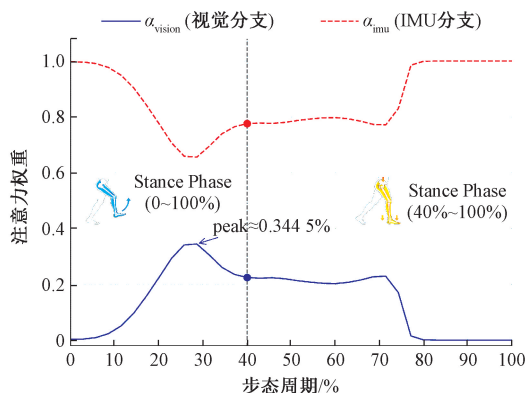


图 10 注意力权重动态变化

Fig. 10 Dynamic change diagram of attention weights

果与人体生理真实的肌肉活动模式高度吻合。

相比之下,仅使用 IMU 模态的单分支模型在预测精度上表现出明显的局限性。在肌肉力峰值区域,OnlyIMU 模型的预测值显著低于真实峰值和 DBFFNet 的预测值,存在明显的峰值低估现象。此外,其波形在整个动态变化过程中平滑度不足,未能准确捕捉肌肉力在关键阶段的快速变化,导致与真实波形存在较大偏差。同时,在肌肉力从峰值下降的后半段,OnlyIMU 模型的预测值相较于真实值和 DBFFNet 的预测值呈现出明显的相位滞后,并且在整个肌肉力衰减阶段持续性地高估了肌肉力。例如,在样本点 27 处,真实肌肉力已显著下降至约 8.41 N,OnlyIMU 模型预测肌肉力为 13.25 N,而 DBFFNet 的预测值为 8.75 N。这种滞后和持续高估表明,单一的 IMU 模态虽然能反映肢体运动,但缺乏足够的上下文信息(如与环境的交互、肢体的精确姿态等),导致其难以精准感知肌肉力由主动收缩向放松或维持较低水平转变的时刻,从而无法及时调整预测结果。

综合分析表明,DBFFNet 通过深度融合视觉与 IMU 两种模态的优势,有效弥补了单一模态的不足。视觉信息提供了肢体与环境的全局姿态、接触信息和宏观运动模式,而 IMU 信息则提供了精准的局部运动学和动力学数据。两者的结合,使得 DBFFNet 能够构建一个更全面、更鲁棒的肌肉力预测模型,从而在高动态和复杂多变的运动工况下,实现对人体下肢肌肉力更为准确、及时的预测。

图 10 展示了 DBFFNet 模型中视觉分支与 IMU 分支的注意力权重在标准化步态周期(0~100%)内的动态分配模式。图 10 中将步态周期明确划分为摆动相(Swing Phase, 0~40%)和支撑相(Stance Phase, 40%~100%),旨在便于分析模型在不同生物力学阶段的模态权重偏好。

摆动相阶段(0~40%),肢体处于腾空状态,主要完成屈伸与摆动运动。如图 10 所示,视觉分支权重在摆动

相初期(0~约 30% 步态周期)呈现显著上升趋势,并在约 30% 步态周期处达到峰值 0.3445。这表明在肢体进行高速摆动和空间定位过程中,视觉信息在精确感知肢体运动轨迹、预测其即将着地位置以及估计控制摆动所需的肌肉力方面,其贡献度显著提升。同时,IMU 分支权重在摆动相初期保持在极高水平(接近 1.0),但随着肢体进入摆动中后期(约 10%~30%),其权重逐渐下降。这显示,尽管 IMU 持续提供实时的姿态和运动学数据,但在肢体脱离地面、专注于空间运动规划时,其对预测肌肉力的相对重要性相比于视觉模态所提供的宏观空间状态和预测性信息有所降低。

进入支撑相阶段(40%~100%),肢体与地面接触,承担身体重量并提供推进力。在此阶段,IMU 分支权重迅速回升,并从约 40% 步态周期开始持续维持在接近 1.0 的极高水平,占据绝对主导地位。这有力地说明,在地面接触并承载身体负荷时,IMU 传感器所提供的高频运动学和动力学信息对于预测维持平衡、缓冲冲击以及产生推进力所需的复杂肌肉活动具有决定性作用。模型在此阶段几乎完全依赖 IMU 数据进行肌肉力估算。与之对应,视觉分支的权重在支撑相初期迅速下降,并在约 75% 步态周期后趋近于 0。这清晰表明,一旦肢体与地面建立稳定的力学接触,视觉信息在精确预测支撑与推进相关的内部肌肉力方面的相对重要性大幅降低。IMU 数据能够更直接、更实时地反映由地面反作用力和肌肉收缩产生的身体内部动力学响应,而视觉信息在此刻对内部动力学建模的贡献相对有限。

综上所述,图 10 直观展示了 DBFFNet 模型中自适应注意力机制的智能化融合策略。该机制不同于传统的固定权重融合方式,能够根据步态周期不同阶段的生物力学需求,动态评估并调整视觉与 IMU 模态的贡献权重。具体而言,在需要捕捉肢体宏观空间轨迹和实现精准目标定位的摆动相中,视觉模态权重显著提升;而在需要精

确感知肢体与地面交互、维持动态平衡以及输出大强度力量的支撑相中,IMU 模态权重则占据主导地位。这种灵活的自适应融合策略,使得 DBFFNet 能够高效地从当前最具信息量的模态中提取关键特征,有效克服了单一模态的内在局限性,从而显著提升了模型在整个步态周期内对人体下肢肌肉力预测的准确性和鲁棒性。可以认为,该自适应模态融合机制是 DBFFNet 实现卓越预测性能的核心保障。

3) 肌肉疲劳条件下的模型鲁棒性分析

为检验模型在康复训练等长时任务中,面对肌肉疲劳这一非理想因素时的鲁棒性,本文设计了疲劳诱导实验。受试者在 10° 上坡跑台上以 1.0 m/s 的速度持续行走 5 min 。分别截取实验初始阶段(非疲劳态)与末期阶段(疲劳态)的数据进行对比分析。

如图 11 所示,随着疲劳的累积,由 sEMG+Hill 模型计算出的参考肌肉力曲线发生了显著变化:力的峰值出现衰减,高力输出的持续时间缩短。这反映了肌肉收缩效率下降,导致传统的基于 sEMG 的映射关系发生漂移。然而,本文提出的 DBFFNet 模型并未依赖 sEMG 信号,而是通过学习外部运动学特征到内部肌肉力的映射。由于人体的宏观运动模式在疲劳初期变化相对微小,模型依然能够准确捕捉运动特征并给出稳定的力预测,预测曲线与疲劳前的形态高度一致。这充分证明了本方法对因肌肉生理状态变化而引起的 sEMG 信号不稳定性具有天然的免疫力,表现出强大的鲁棒性。

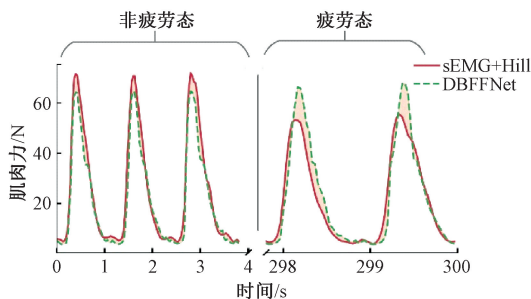


图 11 疲劳条件下的模型预测鲁棒性验证

Fig. 11 Verification of model prediction robustness under fatigue conditions

3.3 外骨骼平台应用验证

为全面验证本研究提出的多模态感知框架在实际工程场景中的应用价值,本文将该方法部署至一款集成步态识别、步态同步及云端诊断等智能功能的商用下肢外骨骼康复机器人系统(如图 12 所示)中,并开展了初步的人机交互实验^[38]。

本研究使用的商用外骨骼平台支持开发者模式,并提供开放的串行通信接口。该系统允许外部设备通过标准 USART 协议与外骨骼内部控制器进行双向数据交互,

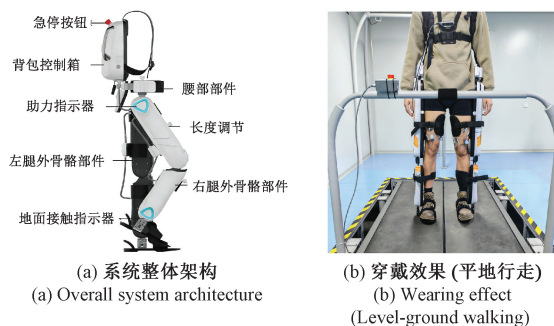


图 12 下肢外骨骼康复机器人平台

Fig. 12 Lower limb exoskeleton rehabilitation robot platform

为多模态感知算法的嵌入式部署与实时控制提供了硬件基础。为实现预测肌肉力对外骨骼的实时调节,本文构建了一套基于高性能 PC 的上位机控制系统。该系统实时采集多模态传感器的原始数据,并在线运行本文提出的 DBFFNet 模型,对穿戴者的肌肉力进行连续预测。预测结果通过串行通信接口发送至外骨骼控制器,用于在线调节阻抗控制器的刚度参数。

基于上述架构,系统能够根据预测的肌肉力水平动态调整辅助强度,实现“按需辅助”的康复策略。该策略使外骨骼输出的辅助力与使用者的运动意图保持高度同步,从而提高人机交互的自然性与协调性。

图 13 展示了平地行走实验中系统关键的运动学轨迹与动力学响应数据,旨在从整体运动稳定性和局部人机协同性两个维度,验证感知框架与闭环控制策略的有效性。

图 13(a)首先呈现了系统质心在笛卡尔空间下的位移演变轨迹,直观反映了外骨骼的行走稳定性。在前向运动维度,质心位移 P_x 表现出“宏观稳定推进”与“微观节律跟随”相结合的特征。基于位移曲线拟合的平均速度保持恒定,表明外骨骼能够辅助受试者维持稳定的行进速率,未出现明显的滞后或突变;而叠加于基准之上的微小波动 ΔP_x ,则客观反映了系统对人体自然步态周期中加减速节律的柔顺跟随能力。在侧向运动维度,质心位移 P_y 表现为以零值为中心的有界收敛震荡。其震荡幅值 ΔP_y 合理对应了行走过程中人机系统重心交替转移的生物力学特征,表明控制算法在容许必要的人体自然摆动的同时,有效防止了侧向运动幅度的失控。尽管惯性传感器长时间积分存在累积误差风险,但极低的横向漂移趋势证实了算法有效抑制了轨迹的非线性发散,保障了长时间行走中的全局方向稳定性。

为了进一步解析上述宏观稳定性的底层机理,图 13(b)与(c)分别展示了典型步态周期内髋、膝关节的辅助力矩响应特性。整体上,外骨骼输出的力矩轨迹与人体关节运动轨迹表现出极高的时间—相位同步性,验证了

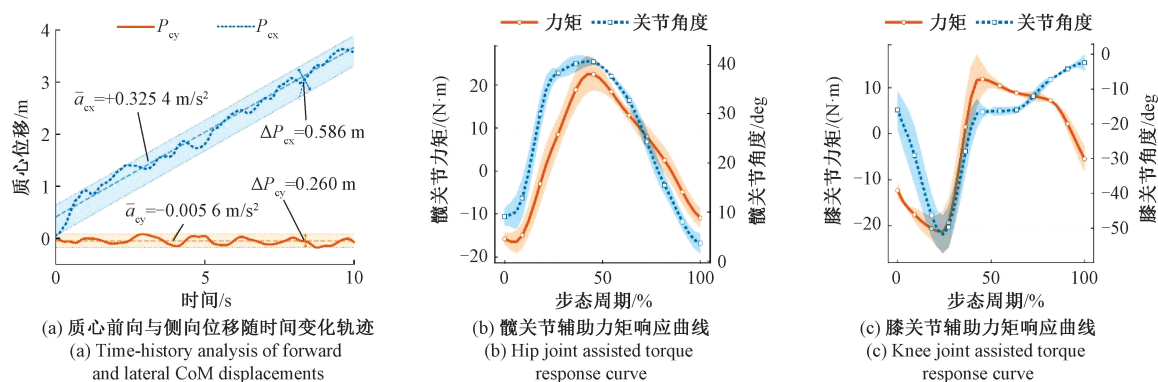


图13 外骨骼平地行走实验结果

Fig. 13 Experimental results of exoskeleton walking on level ground

控制器基于多模态感知对人体自然行走肌肉协同模式的准确解码。在髋关节层面(图13(b)),辅助力矩与关节角度呈强耦合关系。伴随屈曲角度的增加,控制器迅速响应输出正向力矩以提供前摆助力,且力矩峰值时刻与最大屈曲角度精确对应。这种“零迟滞”的跟随特性表明系统有效克服了机械结构的惯性,实现了对人体屈髋动作的精准增强。在膝关节层面(图13(c)),力矩响应呈现出复杂的双向调节特征,紧密跟随关节的屈伸节律。在步态初期的屈膝阶段,负向力矩精确配合膝关节的屈曲运动,提供必要的缓冲与支撑;而在随后的伸膝阶段,力矩迅速回升以辅助小腿摆动。这种既在支撑相提供动力学支持、又在摆动相实现柔顺跟随的控制策略,有效降低了人机交互中的对抗阻抗。

此外,图13(b)、(c)中阴影区域所代表的轨迹误差在多个步态周期内始终保持在极窄范围内。这一低离散度特征有力证明了:尽管存在地面冲击和人体运动变异性,基于肌肉力预测的闭环控制算法仍具有优异的复现性与鲁棒性,能够持续输出与使用者运动意图高度同步的动态辅助,验证了该方法在实际外骨骼助行应用中的有效性。

4 结论

本研究针对穿戴式外骨骼肌肉力感知中传统表面肌电信号鲁棒性不足的问题,提出了一种基于视觉与惯性测量单元多模态融合的创新方法。所构建的双分支特征融合网络有效克服了sEMG信号易受肌肉疲劳影响的局限性,显著提升了肌肉力预测的精度与稳定性。在模拟实际应用场景的多种坡度工况及非理想干扰条件下,该方法展现出优异的鲁棒预测性能与跨受试者泛化能力,为外骨骼系统的“按需辅助”策略提供了可靠的技术支撑。

尽管本研究取得了阶段性成果,但仍有若干问题需进一步探讨。当前模型的计算复杂度尚需优化,以适配资源受限的穿戴式设备并实现高效部署。视觉模态在极端光照变化或遮挡场景下的鲁棒性,以及IMU模态在快速、高动态运动中的精度保障,仍是未来研究的重点方向。

此外,需进一步扩展数据集的规模与多样性。当前实验样本主要集中于健康成年男性群体,虽然验证了方法在理想生理状态下的有效性,但限制了模型在更广泛人群中的泛化能力评估。未来的研究应致力于将模型推广至不同年龄层、多样化的体型特征以及轻度运动障碍患者。针对不同群体的生物力学特异性进行个性化建模与迁移学习,将是验证该方法临床适用性与普适性的关键步骤,从而为外骨骼康复机器人的精准医疗提供更坚实的技术支撑。

参考文献

- [1] 孟琳,董洪涛,侯捷,等.面向下肢康复的柔性外骨骼机器人进展研究[J].仪器仪表学报,2021,42(4):206-217.
MENG L, DONG H T, HOU J, et al. Soft exoskeleton robot facing to lower-limb rehabilitation: A narrative review [J]. Chinese Journal of Scientific Instrument, 2021, 42(4): 206-217.
- [2] SRIVASTAVA S, SHRADDHA K, PEI-CHUN K, et al. Assist-as-needed robot-aided gait training improves walking function in individuals following stroke[J]. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2015, 23(6): 956-963.
- [3] LIN J P, DIVEKAR N V, THOMAS G C, et al. Optimally biomimetic passivity-based control of a lower-limb exoskeleton over the primary activities of daily life[J]. IEEE Open Journal of Control Systems, 2022,

- 1: 15-28.
- [4] GASPARRI G M, LUQUE J, LERNER Z F. Proportional joint-moment control for instantaneously adaptive ankle exoskeleton assistance[J]. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2019, 27(4): 751-759.
- [5] SARTORI M, DURANDAU G, DOŠEN S, et al. Robust simultaneous myoelectric control of multiple degrees of freedom in wrist-hand prostheses by real-time neuromusculoskeletal modeling[J]. Journal of Neural Engineering, 2018, 15(6): 066026.
- [6] LOTTI N, XILOYANNIS M, DURANDAU G, et al. Adaptive model-based myoelectric control for a soft wearable arm exosuit: A new generation of wearable robot control[J]. IEEE Robotics & Automation Magazine, 2020, 27(1): 43-53.
- [7] YAO SH W, ZHUANG Y, LI ZH J, et al. Adaptive admittance control for an ankle exoskeleton using an EMG-driven musculoskeletal model[J]. Frontiers in Neurorobotics, 2018, 12:00016.
- [8] YANG CH, XI X G, CHEN S J, et al. SEMG-based multifeatures and predictive model for knee-joint-angle estimation[J]. AIP Advances, 2019, 9(9): 5120470.
- [9] DURANDAU G, RAMPELTSHAMMER W F, VAN DER KOOIJ H, et al. Neuromechanical model-based adaptive control of bilateral ankle exoskeletons: Biological joint torque and electromyogram reduction across walking conditions[J]. IEEE Transactions on Robotics, 2022, 38(3): 1380-1394.
- [10] 石欣, 敖钰民, 范智瑞, 等. 基于 sEMG 的下肢连续运动切换态实时识别方法[J]. 仪器仪表学报, 2024, 45(4): 165-174.
SHI X, AO Y M, FAN ZH R, et al. Real-time recognition method of switching states of continuous lower limb movements based on sEMG[J]. Chinese Journal of Scientific Instrument, 2024, 45(4): 165-174.
- [11] LI J W, ZHANG B, WANG ZH H, et al. Optimizing sEMG-based gesture recognition under nonideal conditions through a robust deep learning approach[J]. IEEE Transactions on Instrumentation and Measurement, 2025, 74(2): 3562991.
- [12] ZHU B, ZHANG D H, CHU Y Q, et al. SeNic: An open source dataset for sEMG-based gesture recognition in non-ideal conditions[J]. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2022, 30: 1252-1260.
- [13] 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): 1101-1110.
- [14] VAN DER KOOIJ H, VAN ASSELDONK E H F, SARTORI M, et al. AI in therapeutic and assistive exoskeletons and exosuits: Influences on performance and autonomy[J]. Science Robotics, 2025, 10(104): ad7329.
- [15] WANG J N, MATTAMALA M, KASSAB C, et al. Exosense: A vision-based scene understanding system for exoskeletons[J]. IEEE Robotics and Automation Letters, 2025, 10(4): 3510-3517.
- [16] TRICOMI E, MOSSINI M, MISSIROLI F, et al. Environment-based assistance modulation for a hip exosuit via computer vision[J]. IEEE Robotics and Automation Letters, 2023, 8(5): 2550-2557.
- [17] CHEN J, LI H, ZHU X R, et al. Research on environmental slope prediction of knee assisted exoskeleton based on multi-source signals[C]. Proceedings of 2022 Chinese Intelligent Systems Conference, 2022: 584-596.
- [18] RAMANATHAN M, LUO L C, FOO M J, et al. Heuristic vision based terrain recognition for lower limb exoskeletons[C]. Proceedings of the 16th International Convention on Rehabilitation Engineering and Assistive Technology, 2024: 25-28.
- [19] 尹相国, 张岱岩, 林明星, 等. 基于多模感知的上肢康复训练轨迹跟踪研究[J]. 仪器仪表学报, 2023, 44(2): 154-163.
YIN X G, ZHANG D Y, LIN M X, et al. Research on trajectory tracking of upper limb rehabilitation training based on multi-mode perception[J]. Chinese Journal of Scientific Instrument, 2023, 44(2): 154-163.
- [20] 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[J]. 自动化学报, 2021, 47(5): 955-969.
LI Z Y, ZHAO X G, ZHANG B, et al. Review of sEMG-based motion intent recognition methods in non-ideal conditions[J]. Acta Automatica Sinica, 2021, 47(5): 955-969.

- [21] HILL A V. The heat of shortening and the dynamic constants of muscle[J]. *Proceedings of the Royal Society B: Biological Sciences*, 1938, 126(843): 136-195.
- [22] HERMENS H J, FRERIKS B, DISSELHORST-KLUG C, et al. Development of recommendations for sEMG sensors and sensor placement procedures [J]. *Journal of Electromyography and Kinesiology*, 2000, 10(5): 361-374.
- [23] KAIFOSH P, REARDON T R, SKIPPER M. A generic non-invasive neuromotor interface for human-computer interaction[J]. *Nature*, 2025, 645(8081): 702-711.
- [24] REAZ M B I, HUSSAIN M S, MOHD-YASIN F. Techniques of EMG signal analysis: Detection, processing, classification and applications[J]. *Biological Procedures Online*, 2006, 8(1): 11-35.
- [25] CLANCY E A, MORIN E L, MERLETTI R. Sampling, noise-reduction and amplitude estimation issues in surface electromyography[J]. *Journal of Electromyography and Kinesiology*, 2002, 12(1): 1-16.
- [26] WICKIEWICZ T L, ROY R R, POWELL P L, et al. Muscle architecture of the human lower limb[J]. *Clinical Orthopaedics and Related Research*, 1983, 179(1): 275-283.
- [27] LLOYD D G, BESIER T F. An EMG-driven musculo-skeletal model to estimate muscle forces and knee joint moments in vivo[J]. *Journal of Biomechanics*, 2003, 36(6): 765-776.
- [28] ZAJAC F E. Muscle and tendon: Properties, models, scaling, and application to biomechanics and motor control[J]. *Critical Reviews in Biomedical Engineering*, 1989, 17(4): 359-411.
- [29] DELP S L, LOAN J P, HOY M G, et al. An interactive graphics-based model of the lower extremity to study orthopaedic surgical procedures[J]. *IEEE Transactions on Biomedical Engineering*, 1990, 37(8): 757-767.
- [30] HE K M, ZHANG X Y, REN SH Q, et al. Deep residual learning for image recognition[C]. 29th IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770-778.
- [31] CHOI A, HYONG KIM T, CHAE S, et al. Improved transfer learning for detecting upper-limb movement intention using mechanical sensors in an exoskeletal rehabilitation system[J]. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2024, 32: 3953-3965.
- [32] GREFF K, SRIVASTAVA R K, KOUTNÍK J, et al. LSTM: A search space odyssey[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2017, 28(10): 2222-2232.
- [33] 张菀,张泰瑀,贾民平,等. 多传感器融合和 MHA-LSTM 的电机轴承剩余寿命预测[J]. *仪器仪表学报*, 2024, 45(3): 84-93.
- ZHANG W, ZHANG T Y, JIA M P, et al. Prediction of remaining life of motor bearings using multi-sensor fusion and MHA-LSTM [J]. *Chinese Journal of Scientific Instrument*, 2024, 45(3): 84-93.
- [34] HU J, SHEN L, SUN G. Squeeze-and-excitation networks [C]. 2018 IEEE Conference on Computer Vision and Pattern Recognition, 2018: 7132-7141.
- [35] 赵东辉,王威,张紫涵,等. 基于多模态步行意图识别的助行机器人柔顺控制[J]. *仪器仪表学报*, 2022, 43(2): 205-215.
- ZHAO D H, WANG W, ZHANG Z H, et al. Compliance control of walking aid robots based on multimodal walking intention recognition [J]. *Chinese Journal of Scientific Instrument*, 2022, 43(2): 205-215.
- [36] 牛惠祺,张弼,刘丽刚,等. 人体肌肉状态疲劳监测及其在外骨骼交互控制中的应用[J]. *生物医学工程学杂志*, 2023, 40(4): 654-662.
- NIU H Q, ZHANG B, LIU L G, et al. Human muscle fatigue monitoring method and its application for exoskeleton interactive control[J]. *Journal of Biomedical Engineering*, 2023, 40(4): 654-662.
- [37] HELLMANN F, MARCEAU E, DE LA CRUZ R. 60th anniversary of the declaration of helsinki: Ethical challenges in the 10th amendment[J]. *Journal of the Royal Society of Medicine*, 2024, 117(8): 257-263.
- [38] LI X, ZHANG X, LI X, et al. BEAR-H: An intelligent bilateral exoskeletal assistive robot for smart rehabilitation[J]. *IEEE Robotics & Automation Magazine*, 2022, 29(3): 34-46.

作者简介



刘宸,2023 年于东北大学获得博士学位,现为天津理工大学讲师,主要研究方向为人机交互、医疗康复机器人。

E-mail:liuchen@email.tjut.edu.cn

Liu Chen received his Ph. D. degree from

Northeastern University in 2023. He is currently a lecturer at Tianjin University of Technology. His main research interests include human-computer interaction and medical rehabilitation robots.



包瑞来,2023 年于辽宁工业大学获得学士学位,现为辽宁工业大学硕士研究生,主要研究方向为先进控制系统及应用。
E-mail:1326406578@qq.com

Bao Ruilai received his B.Sc. degree from Liaoning University of Technology in 2023. He is currently pursuing his M.Sc. degree at Liaoning University of Technology. His main research interests include advanced control systems and applications.



姚杰,2019 年于东北大学获得硕士学位,现为中国科学院沈阳自动化研究所工程师,主要研究方向为医疗康复机器人、人机交互。
E-mail:yaojie@sia.cn

Yao Jie received his M.Sc. degree from Northeastern University in 2019. He is currently an engineer in the Shenyang Institute of Automation, Chinese Academy of

Sciences. His main research interests include medical rehabilitation robots and human-computer interaction.



张廷丰,2019 年于东北大学获得博士学位,现为辽宁工业大学副教授,主要研究方向为工业系统建模与智能控制、康复机器人控制技术。
E-mail:jzztf@sohu.com

Zhang Tingfeng received his Ph. D. degree from Northeastern University in 2019. He is currently an associate professor at Liaoning University of Technology. His main research interests include industrial system modeling and intelligent control, control technologies for rehabilitation robots.



张弼(通信作者),2017 年于东北大学获得博士学位,现为中国科学院沈阳自动化研究所研究员,主要研究方向为康复医疗机器人。
E-mail:zhangbi@sia.cn

Zhang Bi (Corresponding author) received his Ph. D. degree from Northeastern University in 2017. He is currently a researcher in the Shenyang Institute of Automation, Chinese Academy of Sciences. His main research interest includes rehabilitation medical robotics.