

基于 RK3588 的多模态实时分析与个性化智能健身助手

摘要

随着全民健身热潮兴起，健身动作不规范导致的运动损伤问题日益凸显。据统计，约 60% 的健身爱好者因缺乏专业指导存在动作错误，其中 30% 会因此引发肌肉拉伤、关节磨损等损伤。传统解决方案依赖健身教练或云端 AI 服务，但前者成本高昂且难以覆盖家庭场景，后者存在数据隐私风险且依赖网络稳定性。在此背景下，开发一款兼顾专业性、隐私性和实时性的本地智能健身助手成为迫切需求。针对健身动作不规范导致的运动损伤问题，本团队基于 RK3588 芯片开发了一套本地部署的智能健身助手系统。系统创新性地集成 2B 参数 Qwen2-VL 大语言模型、Silero VAD 流式语音识别模型和 Matcha-TTS 语音合成模型，通过摄像头实时采集动作数据，经本地大模型分析后，通过蓝牙设备提供语音反馈。系统支持双向交互：可主动分析动作并给出建议，也可响应用户语音询问；训练结束后自动生成总结报告。所有数据处理均在本地完成，保障用户隐私安全。

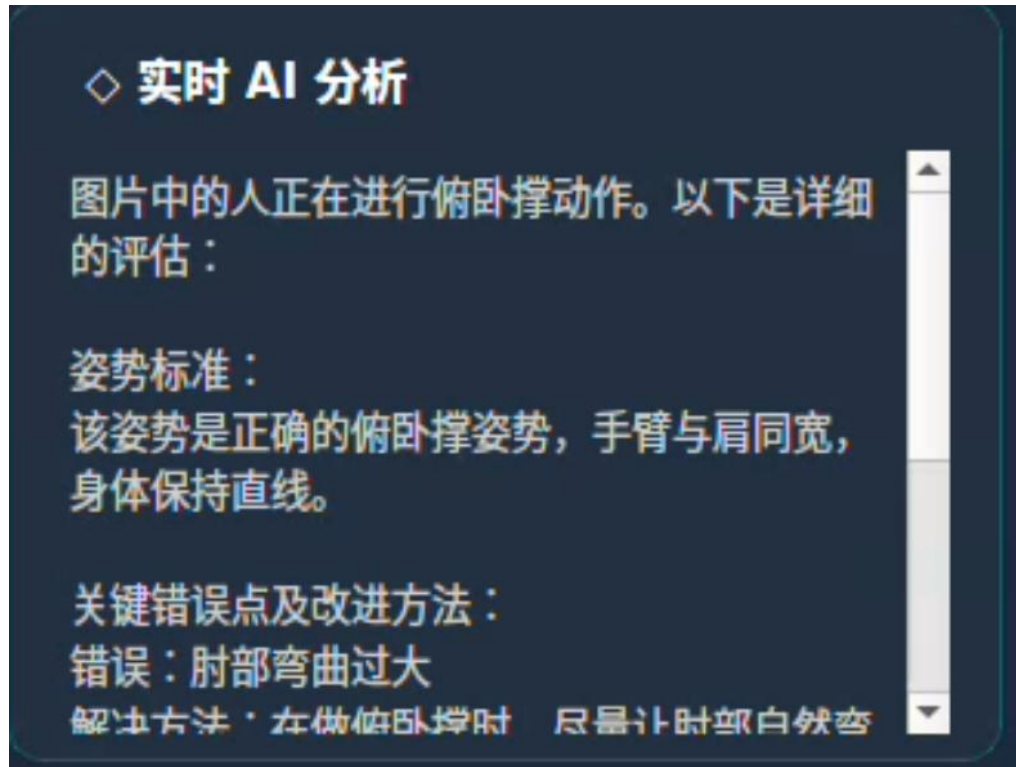
第一部分 作品概述

1.1 功能与特性

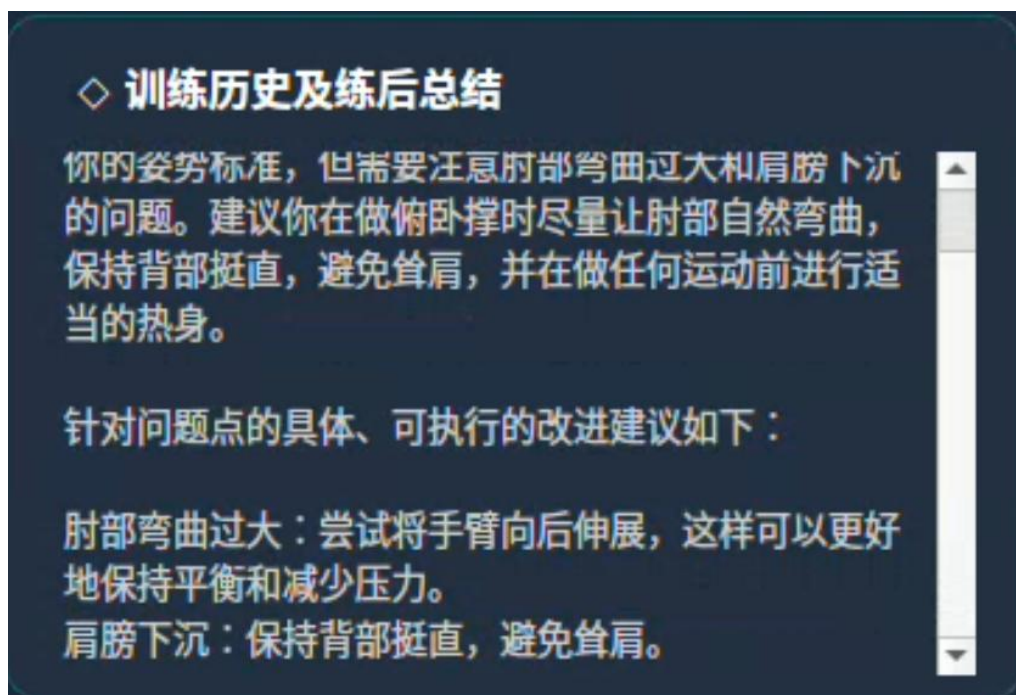
智能健身助手系统具备实时动作分析与反馈、个性化训练支持及数据管理与分析三大核心功能。系统通过传感器实时捕获用户健身动作，借助多模态大模型分析后，经蓝牙将结果反馈至用户端，助其及时纠正不规范动作；支持语音输入自定义 Prompt，满足多样化训练需求。系统还能保存训练数据与异常日志，基于数据分析为用户定制健身计划、提供练后放松建议，助力用户复盘总结，提升健身效果。在多模态大模型的部署方面，我们使用量化、模型转换方式在本地部署了多模态模型、充分利用了 RK3588 的算力优势，同时无需联网即可使用，可以保护用户健身时的隐私，大大提高了系统的稳定性与集成性。

1. 多模态实时分析：通过 USB 摄像头采集 30fps 动作视频，本地 Qwen2-VL

模型分析动作规范性。



2. 智能语音交互：集成 Silero VAD 实时语音识别，支持用户语音提问；Matcha-TTS 将分析结果转为语音通过蓝牙播放。
3. 本地隐私保护：所有模型部署在 RK3588 平台上，无需联网即可实现完整功能。



4. **训练总结功能：**自动记录训练数据，训练结束后生成包含动作评分、改进建议的总结报告。
5. **可视化界面：**7 英寸触摸屏显示实时画面、分析结果和建议内容。



1.2 应用领域

本系统适用于广泛的健身场景，包括家庭、健身房、健身工作室等。家庭用户可摆脱空间与时间限制，享受专业指导；健身房与工作室引入该系统，既能提升服务质量，又能减少教练人力成本。此外，系统也适用于健身爱好者、健身新手及康复训练人群，帮助健身爱好者突破训练瓶颈，辅助健身新手掌握正确动作，为康复训练人群提供安全、科学的训练指导，有效降低受伤风险。

1. **家庭健身：**提供私人教练级指导。
2. **健身房辅助：**作为教练补充，标准化动作教学。
3. **康复训练：**监控康复动作，预防二次损伤。
4. **运动教学：**辅助体育教学，实时动作纠正。



1.3 主要技术特点

系统融合多模态大模型、传感器技术与数据处理技术。多模态大模型实现对复杂健身动作的精准分析；蓝牙技术保障分析结果实时反馈。系统具备数据存储与深度分析能力，通过对训练数据与异常日志的挖掘，为用户提供科学健身建议，形成从动作监测、实时反馈到数据分析、计划制定的完整技术链条。

1. 端侧大模型部署：RK3588 上实现 2B 参数量大小的 Qwen2-VL 模型 INT8 量化部署。
2. 多模态融合：视觉动作分析+语音交互+文本反馈。
3. 离线运行能力：不依赖网络连接，保护用户隐私。

1.4 主要性能指标

指标项	参数值
模型推理速度	11 tokens/s
动作分析帧率	30fps
语音响应延迟	<3 秒
蓝牙传输距离	10 米
系统功耗	5W

连续工作时间	>3 小时（满电状态）
--------	-------------

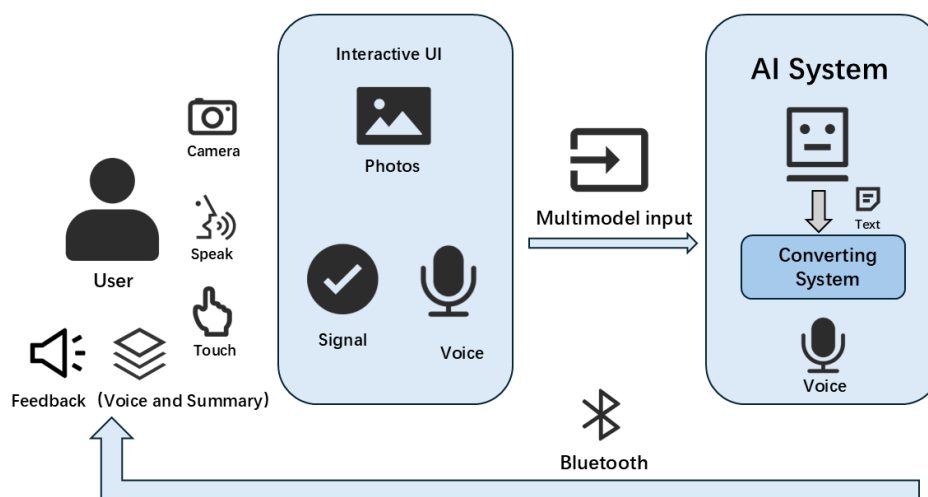
1.5 主要创新点

1. **端侧大模型应用**：在嵌入式设备上部署 2B 参数大语言模型，实现复杂动作分析。
2. **多模态交互系统**：首次将视觉分析、语音交互和本地大模型结合用于健身指导。
3. **隐私保护设计**：所有数据处理在本地完成，无需上传用户数据。
4. **实时反馈机制**：动作分析与语音反馈同步进行，提升用户体验。
5. **个性化定制服务**：通过语音自定义指令结合数据分析，提供定制化健身解决方案。

1.6 设计流程

1. **需求调研**：分析健身用户痛点，确定功能需求。
2. **硬件选型**：基于 RK3588 芯片设计核心板，集成摄像头、蓝牙模块。
3. **模型部署**：优化 Qwen2-VL、Silero VAD 和 Matcha-TTS 模型在 RK3588 上的运行。
4. **软件开发**：开发 PyQt5 图形界面，实现模型调用和交互逻辑。
5. **系统联调**：整合硬件、模型和软件，优化整体性能。
6. **测试验证**：进行功能测试和用户体验测试。

第二部分 系统组成及功能说明



2.1 整体介绍

本系统采用模块化架构，包含数据采集、本地处理模块、交互反馈交互、存储分析分析四个核心模块。硬件采集模块利用多种传感器，实现运动数据与生理参数的实时采集；数据处理模块运用多模态大模型算法，对采集数据进行深度分析与特征提取；反馈交互模块借助蓝牙通信技术，实时推送分析结果，并支持精准识别与响应语音指令；数据存储分析模块采用分布式存储，完整记录运动数据及异常事件，通过机器学习算法深入分析，生成个性化健身方案。各模块通过标准化接口协同工作，保障系统稳定运行。系统采用模块化设计，包含以下核心模块：

1. 数据采集模块：USB 摄像头+麦克风阵列。
2. 本地处理模块：RK3588 芯片运行 Qwen2-VL/VAD/TTS 模型。
3. 交互反馈模块：7 英寸显示屏+蓝牙音频输出。
4. 存储分析模块：本地存储训练数据，生成分析报告。

2.2 硬件系统介绍

2.2.1 硬件整体介绍：

1. **核心板**：飞凌 ELF RK3588 开发板
2. **摄像头**：亚博 1080p 120 帧高清 USB 免驱摄像头
3. **音频系统**：poly20 蓝牙音箱+麦克风阵列

4. 显示系统：飞凌 7 英寸 MIPI 液晶屏

2.3 软件系统介绍

2.3.1 软件整体介绍



2.3.2 软件各模块介绍

1. 模型层：

Qwen2-VL 模型：动作分析与建议生成

Silero VAD：流式语音识别

Matcha-TTS：高质量语音合成

2. 应用层：

PyQt5 图形界面

多线程处理框架

蓝牙音频管理

3. 驱动层：

摄像头驱动

音频设备驱动

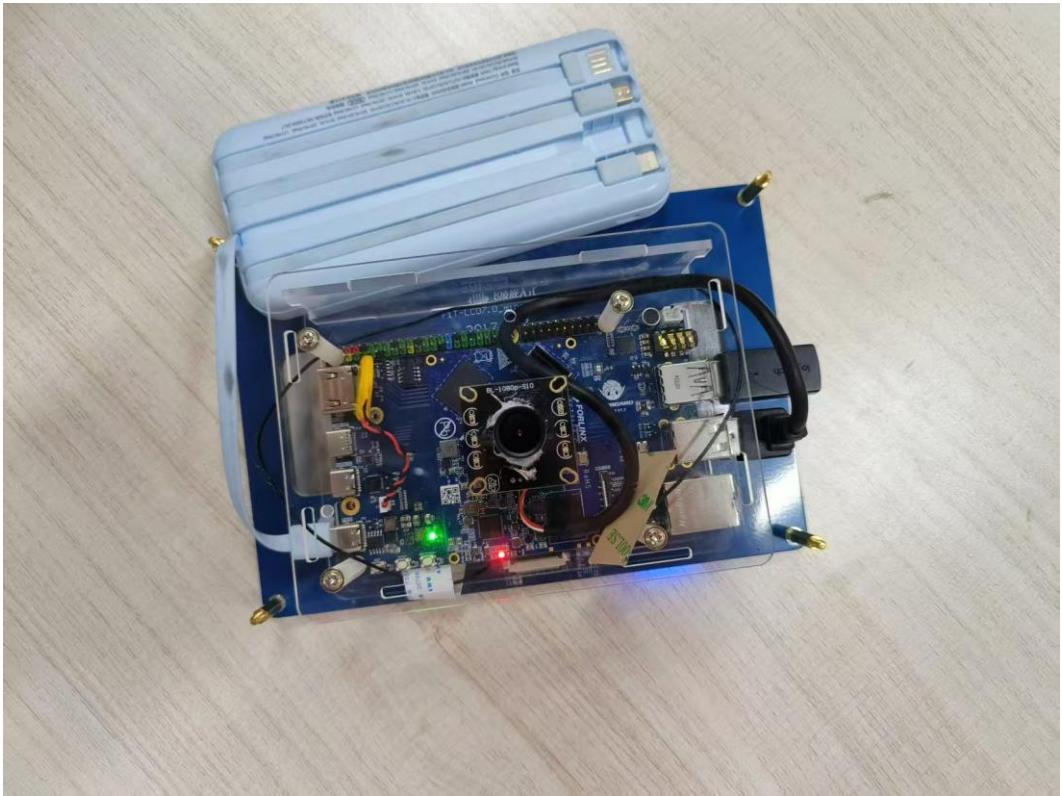
触摸屏驱动

第三部分 完成情况及性能参数

本团队成功开发了一套基于 RK3588 芯片的智能健身助手系统。系统具备高效的动作分析能力、流畅的语音交互体验和强大的本地隐私保护功能，能够有效帮助用户纠正健身动作，降低运动损伤风险，同时为用户提供个性化的健身建议和训练总结。所有功能均在本地完成，无需联网，确保用户数据安全。

3.1 整体介绍

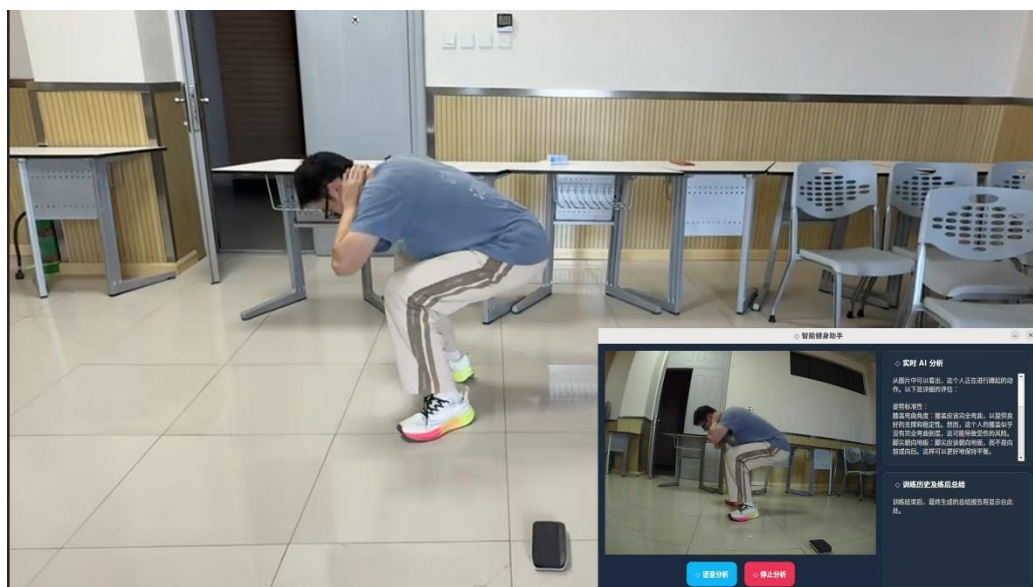




3.2 工程成果

1. 实现本地大模型部署：我们通过模型量化、模型转换，设计模型推理框架等方法在边缘计算板上部署了 Qwen2-VL 模型，并且实现了模型的流式识别
2. 开发蓝牙实时透传功能：将流式语音识别模型识别的语音转成文字后通过管道传给大模型，并将大模型的输出用 TTS 模型转成语音后发送给蓝牙音箱并输出。
3. 开发完整交互界面：我们设计了一个较为美观的 Qt 界面，可以让用户直观地看到训练成果，并且配备了可触摸的屏幕，使用户的使用更加方便。

3.3 特性成果



第四部分 总结

4.1 可扩展之处

1. 功能扩展：增加更多健身项目分析，引入社交分享功能。
2. 模型升级：支持 7B 参数模型部署，提升分析精度。
3. 生态建设：开放 API 接口，接入更多智能健身设备。
4. 多语言支持：扩展支持英语、日语等多语言交互。

4.2 心得体会

本次嵌入式芯片与系统设计竞赛是我们首次参加此大赛。在选题方面，由于我们有大模型开发与部署的相关经验，考虑到应用与 AI 技术的融合，比较了比赛各类开发板对于 AI 应用的可开发性与计算性能，我们选择了瑞芯微赛题。由于队伍的大部分人都有健身的习惯，我们深刻地认识到在健身的时候得到正确的指导，从而避免受伤的重要性。然而，健身教练的聘请费用十分昂贵，因此我们选择了开发一个边缘的智能设备，辅助我们进行健身，在锻炼的时候进行指导。同时，机器需要有可以与人交互的功能，需要充分理解用户的需求，并及时反馈到用户端。同时健身的场景下，健身者手里往往有器械，因此需要远程的透传来反馈信息。考虑到这一点，我们使用了蓝

牙的方式把智能健身助手的健身建议反馈给用户，从而使使用者在健身的时候可以较为便捷，不用将设备佩戴在身上。

在开发时，官方提供的教程文档并不能全部满足我们的开发需求，例如大模型需要流式地识别输入内容并输出结果，但官方给的例程只有简单的单次输入。为了解决这类问题，我们从源码入手，结合瑞芯微提供的开发工具，依据我们的需求完成了视觉由于之大模型的流式输入与输出。在开发蓝牙功能时，我们首先根据官方文档单独更新了内核，让开发板在安装蓝牙 wifi 模块后可以成功地连接蓝牙和 wifi，前没有在虚拟机中配置镜像文件的经验，这确实也是一个不小的难题。接着我们安装了 `espeak`，这是一个常用的文字转语音包（TTS），但我们马上发现这个 TTS 包的语音过于生硬，这与我们想用智能健身助手向用户生动地提出健身建议的想法相违背，于是我们很快找到了使用模型训练的 TTS 模型，并把它部署在 RK3588 上。除此之外，我们还想让用户能够与我们的智能健身助手对话，让智能健身助手能给用户更多的陪伴感，于是我们在 RK3588 上部署了流式语音识别模型。随后我们发现在 `root` 用户下并不能使用蓝牙播放 `.wav` 格式的音频文件，在查阅相关资料后我们安装了 `pulseaudio` 并将蓝牙音箱设置成 `a2dp` 格式后解决了这个问题。

通过这次的开发任务，我们深刻认识到了在开发时永远会遇到新的未知的挑战。然而，无论遇到什么挑战，最重要的是冷静地处理它，同时需要有强大的信息搜集能力，通过各种资源解决问题。同时也要时刻明确自己开发的初衷，尽量让开发的产品能够真实地满足生活中遇到的实际问题。在 RK3588 上部署 AI 模型并开发一个智能健身助手的确是一次有挑战的任务，但也让我们收获颇丰，期待下一次在开发板上开发更有意思有挑战的项目。

第五部分 参考文献

- [1]. Strömel, Konstantin R., et al. "Narrating fitness: Leveraging large language models for reflective fitness tracker data interpretation." *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. 2024.
- [2]. Vahdati, Monica, et al. "A Multi-Agent Digital Twin Framework for AI-Driven Fitness Coaching." *Proceedings of the 2025 ACM International Conference on Interactive Media Experiences*. 2025.
- [3]. Kotte, Hitesh, et al. "Fitsight: Tracking and feedback engine for personalized fitness training." *Proceedings of the 32nd ACM Conference on User Modeling, Adaptation and Personalization*. 2024.
- [4]. Kouaho, Whitney-Jocelyn, and Daniel A. Epstein. "Investigating Perspectives of and Experiences with Low Cost Commercial Fitness Wearables." *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 8.4 (2024): 1-22.
- [5]. Wang, Penghao, et al. "Afitness: Fitness monitoring on smart devices via acoustic motion images." *ACM Transactions on Sensor Networks* 20.4 (2024): 1-24.
- [6]. Tilahun, Abeye Tewodros, et al. "Bluetooth Connect: Messaging App for Offline Communication Over Wide Range Using Mesh Networking or Range Extenders." (2023).
- [7]. Rajeshwari, S., R. Ushasree, and C. Fong Kim. "Digital Notice Board using Bluetooth and Arduino." *INTI Journal* 2024 (2024).
- [8]. Huang, Albert, and Larry Rudolph. "Bluetooth for programmers." Massachusetts Institute of Technology, Cambridge (2005).
- [9]. Kumar, C. Bala, Paul J. Kline, and Timothy J. Thompson. *Bluetooth application programming with the Java APIs*. Morgan Kaufmann, 2004.
- [10]. Huang, Albert S., and Larry Rudolph. *Bluetooth essentials for programmers*. Cambridge University Press, 2007.