

Ciência da Computação / Sistemas de Informação

Prototipagem de Robô Social Assistivo para Cafeicultores Idosos no Sul de Minas Utilizando Plataforma Distribuída de Baixo Custo

Gustavo Rocha Passos - 8º módulo de Engenharia de Controle e Automação, UFLA, bolsista PIBIC/FAPEMIG

Rodrigo Marques Duarte - 10º módulo de Engenharia de Controle e Automação, UFLA, iniciação científica voluntária.

André de Lima Salgado - Orientador DCC, UFLA - Orientador(a)

Resumo

Este trabalho apresenta o desenvolvimento de um robô social assistivo, fundamentado em uma arquitetura cliente-servidor distribuída de baixo custo, capaz de realizar inferência multimodal em tempo real para interação humana. O sistema foi projetado com base no paradigma arquitetural Sentir-Pensar-Agir (Sense-Think-Act - STA), implementado usando de microsserviços que operam em dois nós computacionais distintos: um servidor local de desktop, responsável pelo processamento intensivo de IA, e um dispositivo cliente, uma TV Box HTV8, que gerencia a interface com o usuário. Esta estratégia de offloading computacional permite a integração de modelos complexos, como o Whisper para reconhecimento de fala e o Vision Transformer (ViT) para análise de imagem, com agentes cognitivos orquestrados pelo framework CrewAI, contornando as limitações de hardware do dispositivo cliente. As entradas do usuário (áudio e imagem), capturadas pela TV Box, são transmitidas ao servidor para processamento e inferência contextual. As respostas geradas são então convertidas em fala por um sistema de Text-to-Speech (TTS) e reproduzidas no cliente. Adicionalmente, foi implementado no servidor um banco de dados de memória vetorial com Geração Aumentada por Recuperação (Retrieval-Augmented Generation - RAG), permitindo que o sistema contextualize suas respostas com base em interações anteriores. Técnicas de otimização, como quantização de modelos, foram empregadas para garantir a execução em menos de um minuto no hardware de servidor de baixo custo. Os resultados demonstram a viabilidade de simular interações conversacionais naturais mesmo em ambientes computacionalmente restritos, abrindo caminho para soluções assistivas aplicáveis a contextos sociais, de saúde e educacionais. Este trabalho contribui para o acesso à robótica social no Brasil ao propor uma arquitetura modular, extensível e economicamente acessível, alinhada com o estado da arte em inteligência artificial generativa. Os autores agradecem o apoio da Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG), da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e da Universidade Federal de Lavras (UFLA) pela infraestrutura e suporte para a realização deste trabalho.

Palavras-Chave: Robo Assistivo, Modelos de Linguagem, Assistente de Voz.

Instituição de Fomento: FAPEMIG

Link do pitch: <https://www.youtube.com/watch?v=gYQoGWN3Ho>