

USO DE RECONHECIMENTO DE FALA NA AUTOMAÇÃO DE EQUIPAMENTOS DE GINÁSTICA

Katia Goldchleger, Rogério de Oliveira

Universidade Presbiteriana Mackenzie

Faculdade de Computação e Informática

kahgoldchleger0897@gmail.com, rogerio.oliveira@mackenzie.br

ABSTRACT

Voice automation technologies represent a potential solution to improve people's lives, enabling a wide range of applications in industry, commerce, security, home automation and even smart cities. This study consists of the implementation of a voice control solution in low-cost processing devices, enabling its application in gym equipment for greater accessibility and use of such equipment.

Key-words: Voice control, Speech command recognition

RESUMO

Tecnologias de automação de voz representam uma solução em potencial para melhoria da vida das pessoas possibilitando uma grande quantidade de aplicações na indústria, no comércio, em segurança, automação doméstica e até em cidades inteligentes. Este estudo consiste na implementação de uma solução de controle de voz, em dispositivos de processamento de baixo custo, viabilizando sua aplicação em equipamentos de ginástica para sua maior acessibilidade e uso desses equipamentos.

Palavras-chave: Controle de voz, reconhecimento da fala

1. INTRODUÇÃO

Pessoas no mundo todo vem se beneficiando de tecnologias de voz. Seu uso na telefonia digital e controle de objetos são alguns exemplos de aplicações que trazem mais conforto ou mesmo transformam a vida das pessoas. O problema de acesso de pessoas com deficiência a alguns equipamentos de ginástica, através do uso de tecnologia de voz para o seu controle, pode minimizar esse problema e possibilitar uma melhor qualidade de vida dessas pessoas. Assim, este trabalho explora o uso de tecnologias de reconhecimento de fala para aumentar a acessibilidade de equipamentos como esteiras e escadas rolantes.

A prática de exercícios e o uso de equipamentos de ginástica traz um grande benefício à vida e à saúde das pessoas com deficiência, contribuindo como parte do tratamento ou na sua melhoria de saúde em geral. Entretanto, há um grande número de deficientes para os quais o uso de equipamentos de ginástica é bastante limitado. De acordo com o IBGE (Pessoas com deficiência. IBGE. [S. l.], 2010), há somente no Brasil, cerca de 46 milhões de brasileiros com alguma dificuldade ou deficiência, cerca de 24% da população, para os quais o maior acesso a equipamentos de ginástica poderia representar uma grande melhoria de qualidade de vida.

A Figura 1 mostra o resultado de uma pesquisa de opinião realizada neste trabalho para avaliar a importância dada ao uso do reconhecimento de fala no controle dos equipamentos empregados para exercícios físicos nas academias.

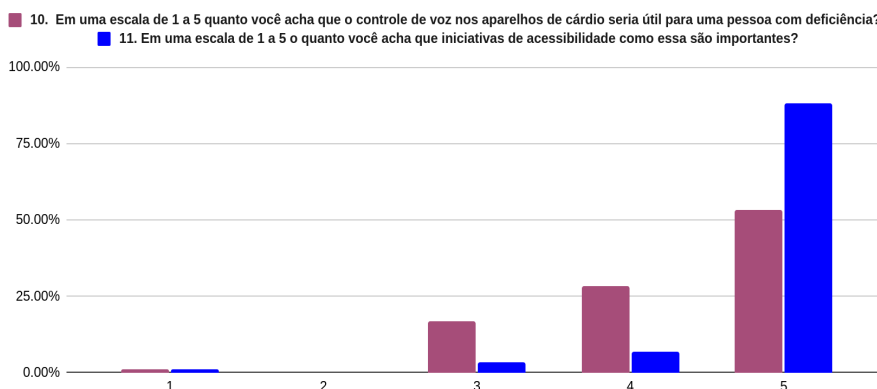


Figura 1. Pesquisa sobre a importância da automação por voz para equipamentos de ginástica entre seus usuários

A pesquisa teve 60 respondentes, dentre usuários de equipamentos de ginástica em academias, e mostra como relevante o uso de soluções como essa.

Este estudo apresenta uma solução conceitual, e minimamente funcional, do uso do reconhecimento de fala para o controle de equipamentos de ginástica, como esteiras e escadas rolantes, ou mesmo equipamentos empregados na reabilitação de pacientes, como robôs de exercícios (Gassert & Dietz, 2018), em que a pessoa pode controlar fatores

como velocidade, força ou extensão, através de simples comandos de voz, aumentando assim a sua acessibilidade, conforto de uso e autonomia. A Figura 2 ilustra como esse estudo poderia evoluir para a aplicação em equipamentos reais.



Figura 2. Representação da prova de conceito e aplicação da pesquisa

Este trabalho traz assim as seguintes contribuições:

1. Relaciona os principais componentes empregados para uma solução de comandos por voz para equipamentos de ginástica/exercícios.
2. Fornece um protótipo conceitual e minimamente funcional da solução e que pode ser empregado em novos desenvolvimentos.
3. Aponta limitações e melhorias futuras da solução proposta.

O restante deste estudo está organizado do seguinte modo: a seção Referencial Teórico traz a fundamentação teórica e trata de comandos de voz, do reconhecimento de fala e da dificuldade do uso de equipamentos de ginástica por pessoas com deficiência; a seção Metodologia discute a implementação realizada, incluindo os dispositivos, bibliotecas e pacotes empregados e o código implementado; por fim, a seção Resultados e Discussão, apresenta os resultados, as limitações encontradas e uma breve pesquisa sobre a utilidade do uso desse tipo de solução.

2. REFERENCIAL TEÓRICO

O primeiro sistema de reconhecimento de voz, denominado “Audrey”, foi desenvolvido no início dos anos 50 pelos Laboratórios Bell para a compreensão de dígitos numéricos. Nos anos 60, um sistema da IBM era capaz de responder até 16 palavras e na década seguinte um sistema desenvolvido pela DARPA/Carnegie Mellon compreendia até 1000 palavras.

Os sistemas de reconhecimento de fala evoluíram do uso predominante de modelos de HMM antes dos anos 2000, para modelos de redes profundas, usados nos principais sistemas atuais (Google, Siri, Alexa). Hidden Markov Models (HMM) são modelos baseados em uma base de um conjunto de técnicas de sucesso para modelagem acústica em sistemas de reconhecimento de fala. As principais razões para esse sucesso se devem à capacidade analítica desse modelo no fenômeno da fala e sua precisão em sistemas práticos de reconhecimento de fala (Negin, N., Farbod, R, Hossein, S. (2009)).

As redes neurais profundas são sistemas de computação com nós interconectados que funcionam como os neurônios do cérebro humano, usando algoritmos, elas podem reconhecer padrões escondidos e correlações em dados brutos, agrupá-los e classificá-los, e – com o tempo – aprender e melhorar continuamente. Isso teve um papel central nos desenvolvimentos recentes na área da tecnologia, tem sido consistentemente encontrado uma poderosa abordagem de aprendizado na exploração de dados de treinamento em larga escala para construir sistemas de análise dedicados (Zhang et al. 2018), e obteve sucesso considerável em uma variedade de campos, como Automatic Speech Recognition.

Nossa solução compreende o uso de um sistema de reconhecimento de fala público através de uma API (Application Program Interface) que é programada sobre uma placa de processamento único. Esses recursos serão brevemente revisados a seguir.

2.1 Sistemas de Reconhecimento de Fala

Automatic Speech Recognition (ASR) é a tecnologia que permite aos seres humanos usar suas vozes para falar com uma interface de computador de uma forma que, em suas variações mais sofisticadas, se assemelha a uma conversa humana normal.

O principal objetivo dos sistemas ASR é a transcrição de fala humana em palavras faladas. É uma tarefa muito desafiadora porque os sinais de fala humana são altamente variáveis devido a vários atributos do falante, diferentes estilos de fala, e ruídos ambientais, dialetos, gírias, a própria língua empregada, velocidade da fala etc. (Abdel-Hamid, O. Abdel-Hamid, O., Mohamed, A., Jiang, H., Peng, L., Penn, G., and Yu, D., 2014). A explicação de como o reconhecimento de fala funciona (Figura 3) é utilizado por vários sistemas mais modernos que empregam em geral uma solução de redes neurais com o treinamento de um grande conjunto de dados por computadores em todo o mundo. Isso é utilizado por assistentes virtuais, como Alexa, Siri e Google Assistant, que são exemplos de sistemas de reconhecimento de fala.

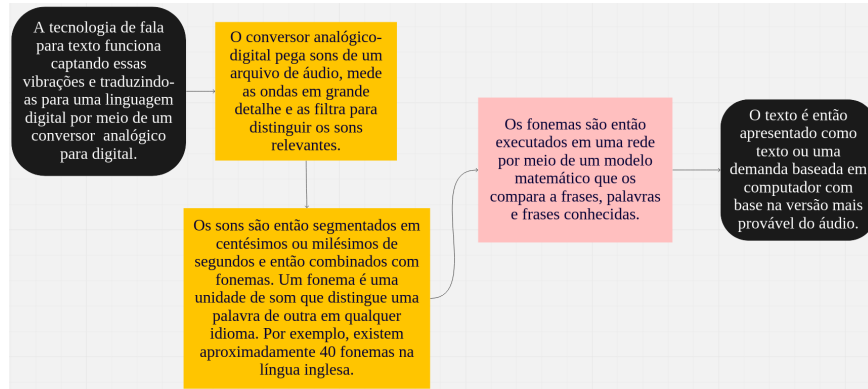


Figura 3. Fluxograma do funcionamento de reconhecimento de fala

2.2 Uso de APIs

API é uma forma de comunicação entre sistemas. Elas permitem a integração entre dois sistemas, fornecendo informações e serviços que um pode utilizar no outro, sem exigir que nenhum dos sistemas conheça os detalhes de como o software é implementado. Portanto, é uma maneira muito segura de dois aplicativos trocarem dados. APIs, são ideais para comunicação em tempo real.

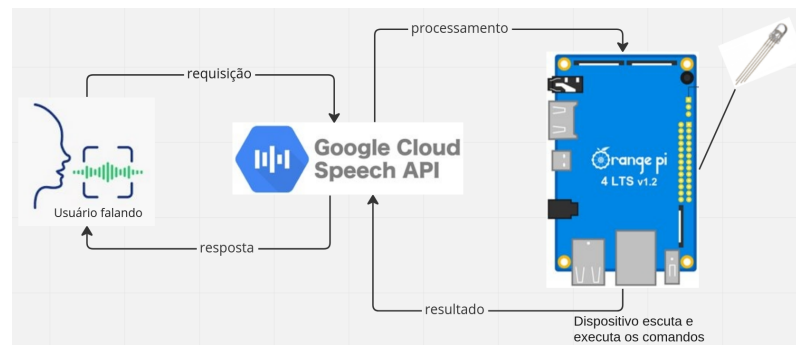


Figura 4. Funcionamento de uma API

O uso de API é extenso, pode ser usado em pagamentos; redes sociais; localização e comércio eletrônico. Porém o foco deste estudo são as APIs de reconhecimento de fala, a seguir temos alguns exemplos. Há várias APIs de mercado para o reconhecimento de fala como (CMU Sphinx; Google Speech Recognition; Google Cloud Speech API; Wit.ai; Microsoft Bing Voice Recognition; Houndify API; IBM Speech to Text e Snowboy Hotword Detection). Neste estudo nos detivemos principalmente na API Google Cloud Speech, empregada na solução final, e na PocketSphinx.

O Google Cloud Speech API pode usar modelos avançados de rede neural com a API de fala para texto. A API identifica muitos idiomas, e as funções dessa API retornam resultados de texto em tempo real. A acurácia em ambientes ruidosos é suportada pelo aprendizado de máquina. As vantagens dessa API incluem vários modelos de aprendizado

de máquina para maior precisão, reconhecimento automático de fala, transcrição de texto, informações pessoais e redução de ruído em chamadas de voz. A desvantagem é o alto custo de uso e o número limitado de dicionários personalizados.

PocketSphinx é uma biblioteca que depende de outra biblioteca chamada SphinxBase que fornece funcionalidade comum em todos os projetos CMUSphinx. Pocketsphinx com Linux, Windows, MacOS, iPhone e Android. As vantagens desta API são a possibilidade de definir uma wakeword personalizada e esta API funciona offline.

2.3 Single Board Computer

Um computador de placa única (referido como SBC a partir de agora) é um computador totalmente funcional, sistema construído em torno de uma única placa de circuito impresso. Um SBC tem um microprocessador(es), memória, entrada/saída e outros recursos exigidos de um computador com funcionamento mínimo. Ao contrário dos computadores pessoais (PC) de mesa, a maioria dos SBCs não possui slots de expansão para funções periféricas ou expansão. Como todos os componentes - processador(es), RAM e GPU, etc. são integrados em uma única placa de circuito impresso (PCB), você não pode atualizar um SBC.

Poucos SBCs são feitos para serem conectados a um backplane para expansão do sistema. SBCs entram muitas variedades, tamanhos, formas, fatores de forma e conjuntos de recursos. Devido aos avanços nas tecnologias eletrônicas e de semicondutores, os preços da maioria dos SBCs são muito baixos. Uma das características mais importantes dos SBCs é seu baixo custo. Com um preço em torno de a peça, você tem em mãos uma ferramenta de desenvolvimento adequada para novas aplicações, hacking, depuração, teste, desenvolvimento de hardware e sistemas de automação.

O SBC é usado principalmente em aplicativos incorporados. Eles são comumente usados em aplicações de controle de processos, como sistemas robóticos complexos e aplicações com uso intensivo de informações. São as melhores e mais modernas alternativas aos microcontroladores. Existem vários dispositivos de processamento de placa únicos como o Raspberry Pi, Orange Pi. Eles se diferem principalmente pela capacidade de memória, interfaces e suporte de sistemas operacionais. (Pajankar, *et al.* (2017)).

2.4 Uso do reconhecimento de fala para pessoas com deficiência

Algumas das maneiras pelas quais o reconhecimento de fala vem sendo usado para ajudar pessoas com deficiência são chamadas telefônicas, envio de mensagens e uso de computadores sem a necessidade de uso do mouse ou teclado, o que inclui escrever texto, fazer pesquisas na web e uso dos mais diversos aplicativos como agendas de compromissos, consultar saldos e extratos, fazer pagamentos, agendar uma consulta médica ou consultar um exame. Segundo a Organização Mundial da Saúde (OMS), cerca de 15% da população global apresenta alguma forma de incapacidade e precisam enfrentar muitas barreiras e dificuldades na sua vida cotidiana (Ambewadikar, M. A., &

Baheti, M. R., 2020). Já na década de 90 o reconhecimento de fala já era visto como uma das mais promissoras tecnologias para ajudar pessoas com deficiências.

Noyes, *et al.* (1989) e outros trabalhos já apontavam para o uso do reconhecimento de fala como solução para ajudar um indivíduo com deficiência no controle do ambiente de sua vida doméstica diária como o controle de dispositivos e facilidades domésticas como televisões, telefones, portas, janelas, ar condicionado etc. ou mesmo o controle de cadeiras de rodas [Pradhan *et al.* 2018]. O reconhecimento de fala ainda permite melhorar a comunicação de pessoas com deficiências da fala, como no caso de indivíduos com paralisia cerebral (Noyes, J.M, *et al.*,1989; Ambewadikar, M. A., & Baheti, M. R., 2020). Mais recentemente o emprego de serviços em nuvem no sistemas de reconhecimento da fala proporcionou um grande avanço no uso desses sistemas, incluindo para o seu uso por pessoas com deficiência (De Russis, L., & Corno, F., 2019). De fato, assistentes de voz integrados a casas inteligentes já são uma realidade no uso doméstico para pessoas com deficiências (Pradhan, A., Mehta, K., & Findlater, L., 2018).

Do mesmo modo que controlam dispositivos domésticos podemos empregar assistentes como esses ou integrar o reconhecimento da fala para criar equipamentos de ginástica inteligentes aumentando sua acessibilidade a pessoas com deficiência. Isso é bastante importante se considerarmos os benefícios do uso dos equipamentos de ginástica para saúde, o mesmo podendo ser aplicado a equipamentos de exercícios que são empregados na reabilitação assistida (Zhou, S. H., *et. al*, 2016). Desse modo, parece ser bastante relevante buscar de IA, como o reconhecimento da fala, para a automação de equipamentos de ginástica, ou mesmo de equipamentos de reabilitação, trazendo mais acessibilidade, conforto e autonomia do uso desses equipamentos por pessoas com deficiência, bastante em linha com uma agenda do uso da IA para uma transformação e melhoria da sociedade (Wamba, S. F., *et. al* 2021; Trewin, S., *et. al*, 2021).

3. METODOLOGIA

Nossa solução é composta pelo computador de única placa Orange Pi 4 LTS, um LED (Figura 5) e o código para fazer a simulação dos comandos falados para um equipamento de ginástica a partir de operações sobre o LED, uma vez que, aumentar a velocidade de piscar um LED, ou aumentar a velocidade ou força em um equipamento de ginástica, devem seguir os mesmos princípios.

Para a implementação optamos pelo uso do computador de placa única de código aberto, uma nova geração de placas de desenvolvimento arm64, que pode executar sistemas operacionais como Android 8.1, Ubuntu e Debian. A placa de desenvolvimento Orange Pi (Orange Pi 4 LTS) usa Rockchip RK3399 ou RK3399-T SoC e possui memória LPDDR4 de 3 GB ou 4 GB e também possui microfone integrado. Não podemos usar para esse tipo de implementação em um arduíno, por causa do maior custo.

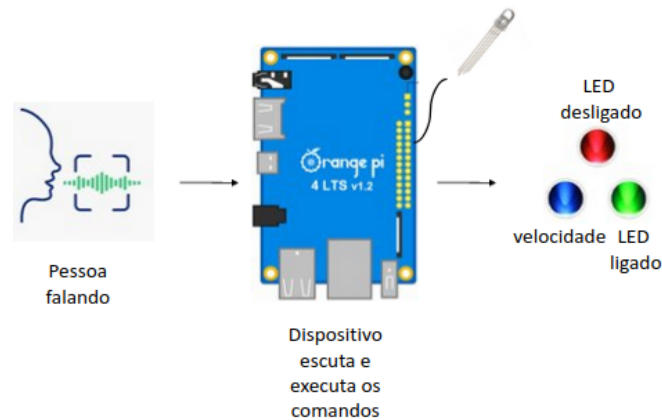


Figura 5. Representação do funcionamento da solução

A linguagem empregada foi o Python, a biblioteca empregada foi SpeechRecognition que utiliza a API da Google para fazer o reconhecimento de fala.

3.1 Implementação

A Figura 6, representa a automação do LED para o controle da velocidade. O dispositivo está sempre ouvindo o ambiente em que se encontra, quando o indivíduo fala a palavra chave mais o comando, ele executa. Assim, do mesmo modo que podemos controlar a velocidade do LED, podemos adaptar esses componentes para a interface de dispositivos como esteiras e escadas rolantes, dentre outros equipamentos, para o controle da velocidade, força ou peso.

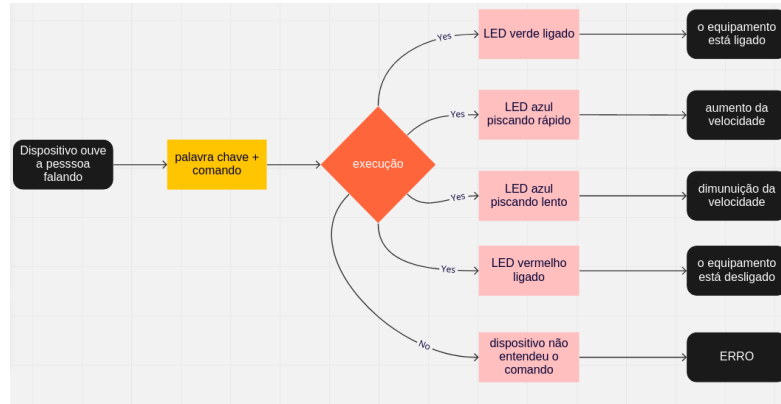


Figura 6. Fluxograma da automação do LED para o controle da velocidade

Há três bibliotecas utilizadas nesse projeto, são elas a WiringPi, SpeechRecognition e Time. O que cada uma contribui com o projeto está explicado abaixo:

- WiringPi é uma biblioteca de acesso GPIO baseada em PIN escrita em C para os dispositivos SoC BCM2835, BCM2836 e BCM2837 usados em todos os Raspberry Pi versões. Essa biblioteca é responsável pelo funcionamento do LED RGB;
- SpeechRecognition faz a realização de reconhecimento de fala, com suporte para diversos motores e APIs, online e offline. Ela utiliza a API da Google para fazer o reconhecimento de fala no projeto;
- Time é um módulo que provê várias funções relacionadas à tempo.

3.2 Código

O código do projeto encontra-se em https://github.com/TCCII/device_for_gym_equipment/. A implementação (Figura 7) foi feita pelo seguinte event loop.

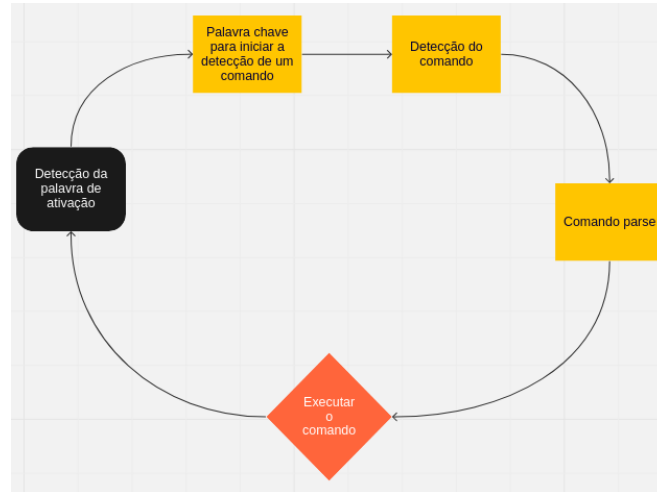


Figura 7. Fluxograma do funcionamento do código do projeto

A pessoa fala no microfone integrado no computador de placa única Orange Pi 4 LTS, ele detecta a palavra de ativação, sendo assim o processador com rede neural para 1 palavra registra e quando é falado a palavra chave que permite o dispositivo escutar os comandos que serão dados pelo indivíduo. Após a palavra chave for falada começará a detecção do comando, isso pode ser feito de várias maneiras, para esse projeto foram consideradas 3. Elas são uma rede neural treinada, Pocketsphinx e uma API externa. A rede neural não foi utilizada por causa do tempo que precisa ser empregado para treinar; o Pocketsphinx não entendia as palavras faladas na maioria das vezes; então a terceira opção foi utilizada.

O comando parse tem vários tipos de comandos, os simples e os complexos. O simples é apenas uma palavra como “pare” e o complexo é um conjunto de palavras como “acelerar para 10”. A execução do comando é a representação do equipamento de ginástica que é um LED, que ao ouvir os comandos o dispositivo faz o LED ligar, piscar mais rápido ou mais devagar ou apagar. Isso representa o ligamento do equipamento, a velocidade e o desligamento. O dispositivo está sempre ouvindo, ou seja, precisa voltar para o começo, pois volta a escutar até que a palavra chave for falada novamente e faz com o programa repita os passos anteriores.

Na execução do código (Figura 8) há 4 palavras que são comandos para a representação da prova de conceito. As palavras são “start”, “faster”, “slower” e “stop”. Qualquer outra palavra falada pelo indivíduo, o programa irá indicar um erro. Há a possibilidade de selecionar o idioma que pode ser interpretado pela API. Os testes foram feitos com inglês dos Estados Unidos.

```

1  '''
2  O objetivo principal de uma instância do Recognizer é,
3  obviamente, reconhecer a fala. Cada instância vem com
4  uma variedade de configurações e funcionalidades para
5  reconhecer a fala de uma fonte de áudio.
6  '''
7  r = sr.Recognizer()
8
9  # Usa o microfone padrão como fonte de áudio
10 speech = sr.Microphone(2)
11 with speech as source:
12     '''
13     O método adjust_for_ambient_noise() lê o primeiro
14     segundo do fluxo de arquivo e calibra o reconhecedor
15     para o nível de ruído do áudio.
16     '''
17     r.adjust_for_ambient_noise(source)
18
19     # Ouve a primeira frase e extraia-a em dados de áudio
20     audio = r.listen(source, phrase_time_limit = 10)
21
22     try:
23         # reconhecer a fala usando o Google Speech Recognition
24         ans = r.recognize_google(audio, language = 'en-US')
25
26         # a fala é ininteligível
27     except sr.UnknownValueError:
28         print("O Google Speech Recognition não conseguiu entender o áudio")
29
30     if audio:
31         if ans == 'start':
32             # comando GPIO
33             tr.start()
34
35         elif ans == 'faster':
36             # comando GPIO
37             tr.faster()
38
39         elif ans == 'slower':
40             # comando GPIO
41             tr.slower()
42
43         elif ans == 'stop':
44             # comando GPIO
45             tr.stop()

```

Figura 8. Pseudo código do projeto

4. RESULTADOS E DISCUSSÃO

A Figura 9 mostra o hardware utilizado na implementação final do projeto. O computador de placa única Orange Pi 4 LTS que está conectado remotamente e o LED que simula os comandos em um equipamento de ginástica. O vídeo exibindo o funcionamento do projeto pode ser acessado em <https://www.youtube.com/watch?v=VfS8R2wF-w>.

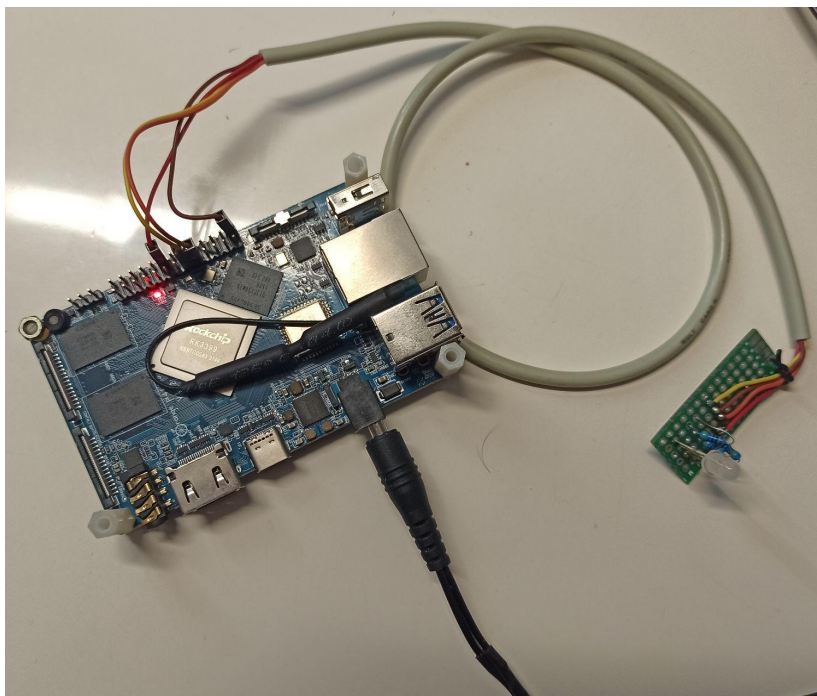


Figura 9. Hardware do projeto

Os testes realizados tiveram 90% de acerto, a cada 10 tentativas de falar um dos 4 comandos, 9 foram acertos; quando é falado um comando com mais de uma palavra, mesmo contendo uma das 4 palavras que fazem parte conjunto de comandos, dá erro.

Essa solução tem limites, pois há deficiências que esse estudo não atende. Para pessoas que são mudas, surdas que não conseguem articular normalmente e cadeirantes não faz muito sentido o uso dessa tecnologia. Nem todos os equipamentos conseguirão ter a implementação do dispositivo. Os ruídos em uma academia (som de uma televisão, pessoas conversando, se exercitando) está presente e isso pode impedir a compreensão dos comandos falados, o que pode atrapalhar o funcionamento do dispositivo.

Este estudo fornece um modelo conceitual, mas funcional, que permite facilmente evoluir para uma solução em dispositivos reais, sendo viável a baixo custo, entre R\$200 e R\$350. A pesquisa foi realizada com 60 pessoas que são usuários de equipamentos de ginástica em academias e tem interesse de ter esse tipo de tecnologia disponível para melhorar os treinos delas. Portanto, para que soluções como essa sejam melhores, elas devem ser avaliadas e desenvolvidas. Em trabalhos futuros parece oportuno o

desenvolvimento dessa solução em equipamentos virtuais antes de sua implementação física e pensamos em avaliar o uso da mesma solução para aplicações em Metaverso, tema que, até onde pudemos verificar, ainda é pouco explorado. Também deve-se buscar o uso do emprego de APIs de reconhecimento de fala com a língua nacional.

5. REFERÊNCIAS

TCC Paulo Santos. Riu UFAM edu, 2022. Disponível em: <https://riu.ufam.edu.br/bitstream/prefix/6144/2/TCC_PauloSantos.pdf>. Acesso em: 15 ago. 2022

Google Cloud Speech API. Stackshare.io. Disponível em: <<https://stackshare.io/google-cloud-speech-api>>. Acesso em: 15 ago. 2022

Speech Recognition. IBM. Disponível em: <<https://www.ibm.com/br-pt/cloud/learn/speech-recognition>>. Acesso em: 15 ago. 2022

Guide to speech recognition technology. Summalinguae. Disponível em: <<https://summalinguae.com/language-technology/guide-to-speech-recognition-technology/>>. Acesso em: 20 out. 2022

Speech Recognition. Library for performing speech recognition, with support for several engines and APIs, online and offline. [S. l.], 5 dez. 2017. Disponível em: <<https://pypi.org/project/SpeechRecognition/>>. Acesso em: 21 out. 2022

Amos, D. The ultimate guide to speech recognition with python. [S. l.]: Real Python, 7 jan. 2022. Disponível em: <<https://realpython.com/python-speech-recognition/#how-speech-recognition-works-an-overview>>. Acesso em: 24 out. 2022

Baig, Faisal, Saira Beg, and Muhammad Fahad Khan. "Controlling home appliances remotely through voice command.". Acesso em: 03 nov. 2022

Lv, Xiaoling, Minglu Zhang, and Hui Li. "Robot control based on voice command." 2008 IEEE International Conference on Automation and Logistics. IEEE, 2008. Acesso em: 03 nov. 2022

Library time. docs python. Disponível em: <<https://docs.python.org/pt-br/3.9/library/time.html>>. Acesso em: 10 nov. 2022

Wiringpi. wiringpi. Disponível em: <<http://wiringpi.com/>>. Acesso em: 10 nov. 2022

What is speech to text. AWS Amazon. Disponível em: <<https://aws.amazon.com/what-is/speech-to-text/>>. Acesso em: 10 nov. 2022

Pessoas com deficiência. IBGE. [S. l.], 2010. Disponível em: <<https://educa.ibge.gov.br/jovens/conheca-o-brasil/populacao/20551-pessoas-com-deficiencia.html>>. Acesso em: 10 nov. 2022

O que é um single board computer. tecncoblog. Disponível em: <<https://tecncoblog.net/responde/o-que-e-um-single-board-computer/>>. Acesso em: 11 nov. 2022

API conceito e exemplos. takenet. Disponível em: <<https://www.take.net/blog/tecnologia/api-conceito-e-exemplos/>>. Acesso em: 11 nov. 2022

Juang, B. H., & Rabiner, L. R. (2005). Automatic speech recognition—a brief history of the technology development. Georgia Institute of Technology. Atlanta Rutgers University and the University of California. Santa Barbara, 1, 67. Acesso em: 12 nov. 2022

De Russis, L., & Corno, F. (2019). On the impact of dysarthric speech on contemporary ASR cloud platforms. *Journal of Reliable Intelligent Environments*, 5(3), 163-172. Acesso em: 12 nov. 2022

Noyes, J. M., Haigh, R., & Starr, A. F. (1989). Automatic speech recognition for disabled people. *Applied Ergonomics*, 20(4), 293–298. Acesso em: 12 nov. 2022

Ambewadikar, M. A., & Baheti, M. R. (2020, October). Review on Speech Recognition System for Disabled People Using Automatic Speech Recognition (ASR). In 2020 International Conference on Smart Innovations in Design, Environment, Management, Planning and Computing (ICSIDEMPC) (pp. 31-34). IEEE. Acesso em: 12 nov. 2022

Zhou, S. H., Fong, J., Crocher, V., Tan, Y., Oetomo, D., & Mareels, I. (2016). Learning control in robot-assisted rehabilitation of motor skills—a review. *Journal of Control and Decision*, 3(1), 19-43. Acesso em: 12 nov. 2022

Wamba, S. F., Bawack, R. E., Guthrie, C., Queiroz, M. M., & Carillo, K. D. A. (2021). Are we preparing for a good AI society? A bibliometric review and research agenda. *Technological Forecasting and Social Change*, 164, 120482. Acesso em: 12 nov. 2022

Trewin, S., Basson, S., Muller, M., Branham, S., Treviranus, J., Gruen, D., ... & Manser, E. (2019). Considerations for AI fairness for people with disabilities. *AI Matters*, 5(3), 40-63. Acesso em: 12 nov. 2022

Pradhan, A., Mehta, K., & Findlater, L. (2018, April). "Accessibility Came by Accident" Use of Voice-Controlled Intelligent Personal Assistants by People with Disabilities. In *Proceedings of the 2018 CHI Conference on human factors in computing systems* (pp. 1-13). Acesso em: 12 nov. 2022

Abdel-Hamid, O., Mohamed, A., Jiang, H., Peng, L., Penn, G., and Yu, D. (2014). Conventional Neural Network for Speech Recognition. *ACM Transaction on Audio Speech, and Language Processing*, 22 (10), PP. 2329-2339. Acesso em: 15 nov. 2022

Ghahramani, Z. (2001). An Introduction to Hidden Markov Models and Bayesian Networks. *International Journal of Pattern Recognition and Artificial Intelligence*, 15 (1), PP. 9-42. Acesso em: 15 nov. 2022

Pajankar, A. (2017). Introduction to Single Board Computers and Raspberry Pi. Raspberry Pi Image Processing Programming, 1–24. Acesso em: 15 nov. 2022

Zhang, Z., Geiger, J., Pohjalainen, J., Mousa, A. E.-D., Jin, W., & Schuller, B. (2018). Deep Learning for Environmentally Robust Speech Recognition. ACM Transactions on Intelligent Systems and Technology, 9(5), 1–28. Acesso em: 15 nov. 2022

Vajpai, J.; Bora, A. Industrial applications of automatic speech recognition systems. Ijera, v. 6, n. 3, p. 88-95, 2016. Disponível em: <https://www.ijera.com/papers/Vol6_issue3/Part%20-%201/O6301088095.pdf>. Acesso em: 15 nov. 2022

Gassert, R., & Dietz, V. (2018). Rehabilitation robots for the treatment of sensorimotor deficits: a neurophysiological perspective. Journal of neuroengineering and rehabilitation, 15(1), 1-15. Acesso em: 15 nov. 2022

Negin, N., Farbod, R., Hossein, S. (2009). A novel approach to HMM-based speech recognition systems using particle swarm optimization. sciencedirect. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0895717710001597>>. Acesso em: 16 nov.