

PROTÓTIPO HUET: AUTOMAÇÃO RESIDENCIAL UTILIZANDO LIBRAS

Filipe Vieira Brison¹, Carlos Dias da Silva Júnior² Elias José de Rezende Freitas³

¹Instituto Federal de Minas Gerais – IFMG Campus Avançado de Itabirito, Itabirito/MG, Brasil (felipebrison@gmail.com)

^{2,3}Instituto Federal de Minas Gerais – IFMG Campus Ibirité, Ibirité/MG, Brasil

Resumo: Este trabalho descreve um protótipo de automação residencial baseado na Internet das Coisas, que se destaca por sua capacidade de reconhecer sinais de Libras para acionar dispositivos. O protótipo recebe o nome de *Huet*, em homenagem a Ernest Huet, um educador francês que incentivou a aprendizagem da língua de sinais por alunos surdos no Brasil. Atualmente, o protótipo encontra-se em estágio inicial e é composto por um nó central (Raspberry Pi 4) responsável pelo processamento de imagens capturadas por várias plataformas microcontroladas ESP32-CAM distribuídas pela residência. Ao detectar um sinal de Libras pré-programado, o nó central envia uma mensagem via WiFi para outro microcontrolador que ativa ou desativa um relé, possibilitando ligar ou desligar equipamentos, como lâmpadas ou ventiladores. Os resultados demonstram o funcionamento geral do protótipo *Huet*, utilizando um interruptor paralelo, sendo analisado o tempo e a precisão na identificação dos sinais. Além disso, este trabalho contribui com a comunidade surda, ao desenvolver tecnologia que auxilia na superação das barreiras de comunicação entre essa comunidade e a sociedade, no geral.

Palavras-chave: Internet das Coisas; Libras.; Visão computacional; Reconhecimento de mão.

INTRODUÇÃO

De acordo com a Organização Mundial da Saúde (OMS, 2020), estima-se que atualmente existem mais de 466 milhões de pessoas surdas ou com deficiência auditiva em todo o mundo, sendo que aproximadamente 34 milhões dessas pessoas são crianças (UNICEF, 2019). No Brasil, cerca de 5% da população, o que corresponde a cerca de 10 milhões de pessoas, possuem algum tipo de deficiência auditiva. Entre aqueles que têm perda auditiva total ou quase total, considerados surdos, o número no Brasil chega a 2,3 milhões, de acordo com dados do Instituto Locomotiva e da Semana da Acessibilidade Surda (BRASIL, 2021).

O Brasil tem uma das maiores comunidades surdas do mundo, sendo que essa comunidade surda é formada não só por pessoas propriamente surdas, mas também por familiares, amigos, educadores, intérpretes e outros profissionais (EIJ, 2018).

A comunicação com a sociedade é uma das principais barreiras enfrentadas pela comunidade surda. No entanto, a tecnologia desempenha um papel fundamental ao auxiliar e promover a inclusão em diversas áreas. Atualmente, existem várias tecnologias disponíveis, tais como aparelhos auditivos, legenda automática, intérpretes de Libras virtuais através de aplicativos, campainhas e alarmes para surdos, transcrição de voz para texto em dispositivos móveis

e outras soluções. Essas tecnologias têm o potencial de superar as limitações de comunicação e facilitar a interação entre surdos e a sociedade em geral.

Por exemplo, o *EchoDot*, criado pela Amazon, é um alto-falante inteligente, controlado por voz. É um dispositivo compacto e acessível, amplamente utilizado em casas inteligentes por meio da voz, cujos donos procuram uma maneira conveniente e prática de gerenciar suas tarefas diárias. O morador usa a assistente virtual Alexa para executar uma variedade de tarefas, como, por exemplo, tocar música, definir alarmes e lembretes, fazer chamadas, informar-se sobre notícias e clima, controlar dispositivos domésticos inteligentes, entre outras coisas.

No entanto, as pessoas surdas sentem-se limitadas com esse tipo de tecnologia. Ao pretender controlar dispositivos domésticos inteligentes, como o *EchoDot*, um usuário com esse perfil pode se deparar com possíveis dificuldades impostas pelo dispositivo para que a sua voz seja compreendida. Se ele a fala for muito baixa ou pouco clara, por exemplo, o *EchoDot* não interpreta corretamente os comandos de voz.

Nesse contexto, este trabalho apresenta o protótipo *Huet* constituído pelos dispositivos eletrônicos, *Raspberry Pi 4* e módulos ESP32CAM e ESP8266, visando à automação residencial para pessoas com deficiência auditiva. Ao em vez de utilizar a voz como outros dispositivos no mercado, o protótipo

desenvolvido permitirá que os usuários utilizem as mãos para os comandos, realizados por meio de gestos, assemelhando-se à Língua de Sinais Brasileira (Libras).

O nome do protótipo foi uma homenagem a Ernest Huet. Ele nasceu em 1820 e foi um professor e poliglota francês que ensinava a língua de sinais para as pessoas surdas. A escolha desse nome é uma forma de reconhecer alguém que fez muito pelos surdos.

METODOLOGIA

A metodologia proposta tem como base a implementação de algoritmos implementados em *Python* e em *C++* e a utilização dos seguintes dispositivos: *Raspberry Pi 4* (RPi4), ESP32-CAM e módulos ESP8266 com relé.

Uma arquitetura de comunicação sem fio foi criada, permitindo a comunicação entre o nó central (RPi4) e os nós de sensoriamento distribuídos (ESP32-CAM), além de permitir que o nó central envie informação para os nós de acionamento (ESP8266), conforme ilustrado no esquema geral da Figura 1.

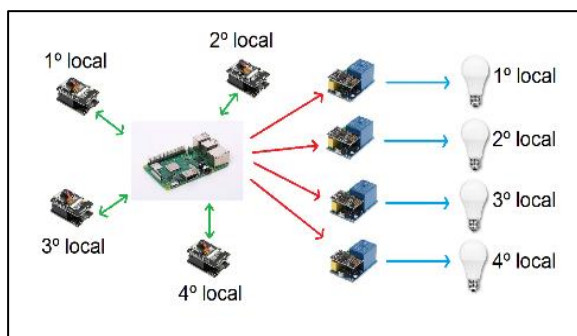


Figura 1. Arquitetura descentralizada de nós, podendo estar em cada locais diferentes na residência.

(Elaborada pelos autores, 2023).

O RPi4 foi escolhido como nó central, pois esse microcomputador possui processador quad-core de 1,5 GHz, contando com 4GB de RAM, Wi-fi e outros periféricos. Por sua vez, o nó de sensoriamento descentralizado é composto por ESP32-CAM, que é capaz de capturar uma imagem, de no máximo 2Mpixels, no ambiente em que for instalado, sem que o usuário precise ir ao local onde o nó central estiver instalado.

Além disso, o nó central recebe os dados, utilizando o protocolo HTTP, de cada nó descentralizado e envia um comando também em HTTP para os nós de acionamento, após identificar o gesto, como ilustra a Figura 2.

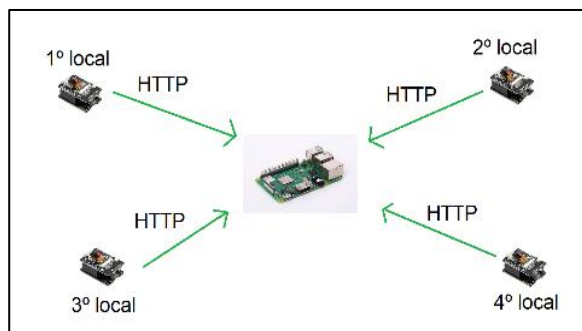


Figura 2. Nós centralizados enviam comandos em HTTP para nó central.
(Elaborada pelos autores, 2023).

Após uma imagem ser recebida pelo nó central, o RPi4 é responsável por processá-la, aplicando filtros de suavidade, de maneira a reduzir os ruídos, além de selecionar a região de interesse de detecção, utilizando a biblioteca OpenCV. Essa imagem processada é utilizada então para o reconhecimento da mão por meio da biblioteca MediaPipe. Essa biblioteca fornece funções que retornam se uma mão foi detectada e seus pontos principais.

Por meio de uma estrutura de condicionais, é possível, então, reconhecer gestos simples em Libras, por exemplo, uma letra, utilizando o algoritmo desenvolvido, que segue a lógica apresentada no fluxograma da Figura 3.

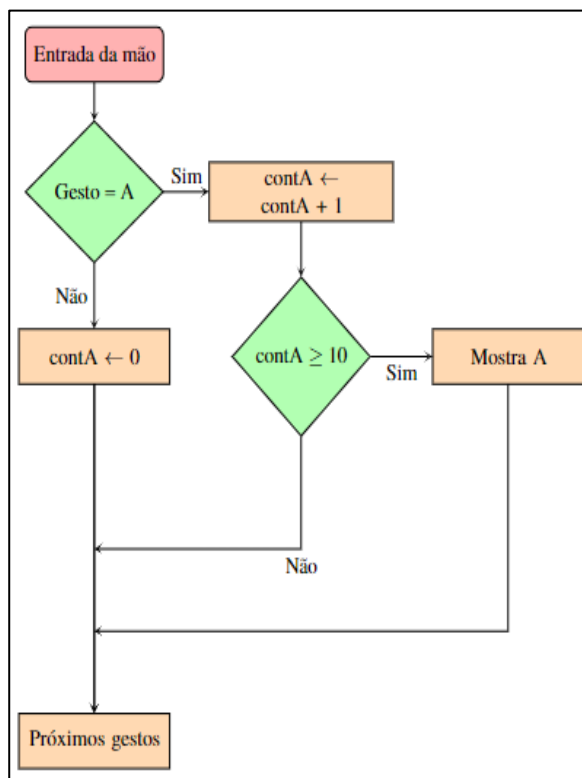


Figura 3. Fluxograma do algoritmo implementado.
(Elaborada pelos autores, 2023).

Em síntese, o funcionamento do algoritmo, pode ser descrito como: após a detecção de uma mão próxima à câmera do ESP32-CAM, na primeira decisão, se o gesto for igual à A, ou outro gesto desejado, a variável $contA$ obterá mais um, caso contrário, será zero para $contA$. Na segunda decisão, se $contA$ for maior ou igual a 10, o algoritmo mostra A, caso contrário, ainda não confirma que mostra A enquanto a mão é mantida à câmera, sinalizando a letra A. Para se ter a confirmação, contando até 10, finalizando com o reconhecimento e informação do gesto executado. A Figura 4 ilustra a ideia completa do sistema desenvolvido, da detecção da mão até o acionamento de um dispositivo, no caso, uma lâmpada.

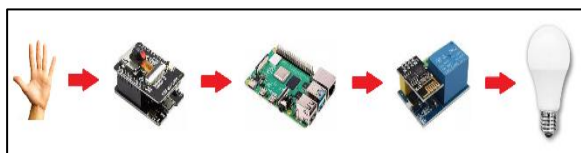


Figura 4. Sequência de ações realizadas pelo protótipo *Huet*.
(Elaborada pelos autores, 2023).

RESULTADOS E DISCUSSÃO

Na Figura 5, é apresentado o nó central, constituído pelo RPi4, que pode ser facilmente instalado em dos cômodos da casa.



Figura 5. *Raspberry Pi 4* (Rpi4) utilizado no protótipo.
(Elaborada pelos autores, 2023).

Para capturar imagens, utiliza-se o módulo ESP32-CAM, ilustrado na Figura 6, que enviará as imagens coletadas descentralizadas via Wi-Fi.



Figura 6. ESP32-CAM utilizado no protótipo.
(Elaborada pelos autores, 2023).

A Figura 7 apresenta o momento em que uma lâmpada é acionada após a detecção correta de uma sequência de gestos corretas.



Figura 7. ESP8266, Relé e Lâmpada, utilizados no protótipo.
(Elaborada pelos autores, 2023).

Assim, os componentes básicos do protótipo podem ser vistos juntos na Figura 8, sendo utilizado uma fonte de alimentação de 5V para cada dispositivo e uma conexão elétrica para o equipamento (lâmpada). Cada dispositivo poderia estar localizado em cômodos diferentes da residência, por exemplo, na sala, na cozinha, nos quartos.

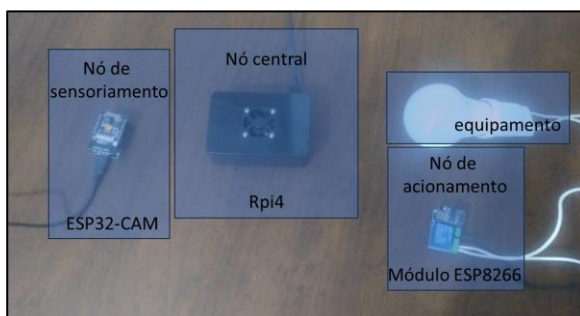


Figura 8. Esquema geral do protótipo *Huet*.
(Elaborada pelos autores, 2023).

A Figura 9 apresenta o resultado da detecção da mão, sendo indicado os pontos principais de detecção, por meio das bibliotecas OpenCV e MediaPipe.



Figura 9. OpenCV e MediaPipe.
(Elaborada pelos autores, 2023).

Dessa forma, cada letra em Libras pode ser detectada utilizando as posições relativas desses pontos em tempo real. A Figura 10 apresenta o resultado desse reconhecimento da letra A, sendo mostrada no canto superior esquerdo da imagem final, após ela ser identificada.

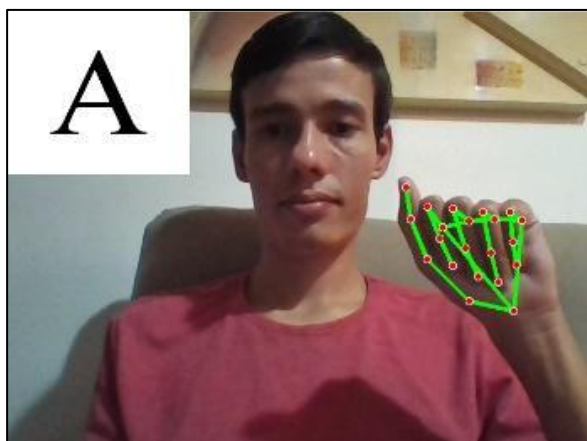


Figura 10. Identificação da letra A.
(Elaborada pelos autores, 2023).

Para a realização dos próximos testes, foram escolhidos os gestos em Libras para as letras A, L, M e V para corresponderem, respectivamente, a Alexa (inicialmente utilizada para iniciar o procedimento de detecção), lâmpada, motor e ventilador, enquanto os números 1, 2 e 3 representam a identificação/localização de cada equipamento.

Assim, cada símbolo em Libras foi sinalizado para a câmera 10 vezes. Após a conclusão do teste, foram registrados o tempo médio de identificação da letra (Tm), o desvio padrão do tempo (Dp), a porcentagem de acertos e a posição do símbolo nas 10 tentativas, conforme apresentado na Figura 11. Nota-se que houve 100% de acertos na detecção de todos os gestos.

Resultados de cada letra:				Sequência
Símbolo	Tm	Dp	Acerto (%)	
A	1.25	0.08	100	['A','A','A','A','A','A','A','A','A','A']
L	1.32	0.15	100	
M	1.61	0.23	100	
V	1.30	0.22	100	
1	1.41	0.43	100	['1','1','1','1','1','1','1','1','1','1']
2	1.39	0.22	100	
3	2.56	2.16	100	

Figura 15. Resultado da detecção de sete gestos em Libras.

(Elaborada pelos autores, 2023).

Foram realizados diversos testes com o protótipo, utilizando uma lâmpada como equipamento conectado ao relé, podendo ser alterado para o equipamento desejado, apenas modificando o relé, de maneira a suportar a corrente do equipamento.

Para testar o protótipo de maneira geral, realizou-se gestos da mão, em Libras, próximo à câmera do ESP32-CAM. Dessa maneira, foi verificada a transmissão adequada da imagem para o Rpi4, possibilitando o processamento do algoritmo que reconhece os gestos, por meio dos valores dos pontos da mão calibrados previamente. Se todas as comparações forem verdadeiras, um dado é enviado para ESP8266 via HTTP, de maneira a acionar um relé, que, por fim, ligará ou desligará o equipamento desejado.

Por fim, as letras H e L foram escolhidas para significarem *Huet* e lâmpada, respectivamente. Ao sinalizar a letra H, aciona-se o protótipo *Huet*, e, ao sinalizar a letra L, é enviado o comando para que *Huet* ligue ou desligue uma lâmpada, sendo a lâmpada desejada selecionada pelo gesto de um número. Alguns dos testes de funcionamento geral do protótipo *Huet*, com a utilização de um interruptor paralelo para acionamento de uma lâmpada, podem ser vistos no vídeo: https://youtu.be/2Dlko_wzaCg.

CONCLUSÃO

Este trabalho propôs uma alternativa para superar a barreira de comunicação entre um usuário surdo e um dispositivo assistente, como o *EchoDot* ou a *Alexa*, no contexto da Automação Residencial. O protótipo *Huet* apresenta semelhanças com esses dispositivos, porém a diferença principal está na forma de comunicação, já que o mesmo utiliza uma câmera para capturar os comandos realizados por gestos de uma pessoa surda.

Os testes realizados demonstram a viabilidade da solução proposta, tendo em vista a identificação correta de alguns sinais de Libras e o tempo de resposta obtido.

Além disso, este trabalho representa uma contribuição para promover a inclusão social dos surdos por meio da tecnologia, permitindo a acessibilidade por meio do dispositivo assistente desenvolvido. Como continuidade do trabalho, espera-se adicionar novos comandos, realizar os testes com mais equipamentos,

desenvolvendo uma versão mais próxima do uso comercial.

REFERÊNCIAS

BRASIL, A. Oms estima 2,5 bilhões de pessoas com problemas auditivos em 2050. Agência Brasil, 3 2021. Disponível em: <<https://agenciabrasil.ebc.com.br/saude/noticia/2021-03/oms-estima-25-bilhoes-de-pessoas-com-problemas-auditivos-em-2050>>.

MINISTÉRIO DA SAÚDE. Surdez - Causas. 2023. Acesso em 18 mar 2023. Disponível em: <<https://bvsms.saude.gov.br/surdez-3/#:~:text=%E2%80%9320a%20surdez%20de%20c%C3%B3clea%20ou,problemas%20metab%C3%B3licos%2C%20tumores.>>>.

EIJI, H. Comunidades Surdas. 2018. Acesso em 06 mai 2023. Disponível em: <<https://culturasurda.net/comunidades-surdas/>>.

UNICEF. 17,5 milhões de crianças não estão matriculadas na educação infantil. 2019. Acesso em 18 mar 2023. Disponível em: <<https://www.unicef.org/brazil/comunicados-de-imprensa/175-milhoes-de-criancas-nao-estao-matriculadas-na-educacao-infantil>>.

OMS. OMS estima 2,5 bilhões de pessoas com problemas auditivos em 2050. 2020. Acesso em 18 mar 2023. Disponível em: <<https://news.un.org/pt/story/2020/03/1705931#:~:text=Em%20todo%20o%20mundo%2C%20466,de%20audi%C3%A7%C3%A3o%20nesse%20mesmo%20n%C3%ADvel.>>>.