

UNIVERSIDADE SÃO JUDAS TADEU

Gabriela Lombardi Lima

Guilherme Yoshikawa

Hugo Uchoa Dantas

Lucas Eduardo Nakayama

Lucas Lima Sousa

Óculos inteligente para deficientes visuais com visão computacional  
embarcada para detecção de objetos

Engenharia Elétrica

Prof. Dr. Carlos López Noriega

São Paulo, SP

2022

## RESUMO

A integração do deficiente visual na sociedade está diretamente ligada com a disponibilidade de pessoas dispostas a ajudar e principalmente com os recursos em sua volta. No entanto surge a necessidade da realização de recursos para facilitar o convívio do deficiente visual com a sociedade. Atualmente os deficientes visuais, com baixa ou perda total da visão em ambos os olhos, trabalham, estudam, têm seus lazeres entre outras atividades diárias, nas quais são significativas perante a sociedade, porém ainda é um grande obstáculo a ser quebrado aderindo a inclusão completa dessas pessoas no meio social. O presente estudo analisou um método que fornecesse recursos básicos suficientes para atividades diárias simples e que realmente fossem significativas para o usuário. Para tanto, um voluntário com as mesmas características descritas acima foi submetido à realização de testes para desempenho do projeto. Dentre todas as atividades básicas diárias do voluntário foi observado que a maior dificuldade é a detecção de obstáculos e objetos, pois essas são as características principais que o fazem aprender o mundo externo e se adaptar ao meio. Com essas dificuldades mais relevantes, levanta-se a questão do desenvolvimento de recursos que contemplem o auxílio das atividades com maior dificuldade do deficiente visual. Apesar das especificidades de tarefas corriqueiras, os dados obtidos nessa pesquisa evidenciam a complexidade do recurso a ser desenvolvido para suportar o usuário com a integração na sociedade.

**Palavras-chave:** Deficiente Visual; Recursos; Detecção; Obstáculos; Objetos.

## **ABSTRACT**

The integration of the visual deficient into the society is directly linked to the availability of people willing to help, and mainly to the resources around them. However, there is a need for creating resources to make easier the life of the visual deficient in the society. Currently, the visual deficient, with partial or total loss of vision in both eyes, works, studies, have his/her leisure time, among other significant daily activities, however, it is still a great obstacle to be broken, adhering to the complete inclusion of these people in the social environment. The present study analyzed a method that would provide enough basic resources for simple daily activities that were really meaningful to the user. For this, a volunteer with the same characteristics described above was submitted to tests for the performance of the project. Among all the basic daily activities of the volunteer it was observed that the greatest difficulty is the detection of obstacles and objects, because these are the main characteristics that make him learn the external world and adapt to the environment. With these most relevant difficulties, the question arises about the development of resources that contemplate the aid for the activities with more difficulty for the visual deficient individual. Despite the specificities of common tasks, the data obtained in this research shows the complexity of the resource to be developed to support the user with the integration in society.

**Keywords:** Visual deficient; Features; Detection; Obstacles; Objects.

## 1. Introdução

A deficiência pode ser definida como um distúrbio ou uma anomalia em que o estado de saúde é alterado, podendo ocasionar sintomas, ou não. De forma simplificada, deficiência é o antônimo de saúde e existem diversas formas das quais uma deficiência pode se manifestar, podendo ser encontrada como uma alteração das condições fisiológicas, intelectuais ou até moleculares (TEPPERWEIN, 2002). Dentre as deficiências, a visual será o tópico principal desenvolvido no presente trabalho.

Ao longo da história da humanidade, as pessoas com deficiência passaram por diversos desafios para conseguirem se adaptar às necessidades da sociedade. Como exemplo disso, pode-se citar na antiguidade a prática da caça e da agricultura, em que era necessário um certo tipo físico - e até mesmo psíquico - para conseguir cumprir com as atividades demandadas. Ou seja, é de fácil percepção a ideia de que, se a pessoa não possuísse os tipos físicos considerados ideais, a sociedade não o aceitaria facilmente e a sua sobrevivência se tornaria cada dia mais difícil (GUGEL,2007; SILVA,1986).

Ademais, com o passar dos anos pode-se observar fatores sociais que demonstram a exclusão dos deficientes do meio social, como ocorria na Grécia Antiga, em que Platão deixou escrito na sua obra “A República” (PLATÃO, sec IV, apud, GUGEL, 2007, p.63) o que a sociedade deveria fazer com as pessoas consideradas "disformes" - escondê-las em lugares longe dos demais ou, então, matá-las. Ato contínuo, avançando um pouco na história da humanidade, consegue-se observar descrito no livro mais vendido do mundo, a Bíblia, que a mentalidade das pessoas sobre os deficientes era extremamente dolorosa, pois isso significava que eles eram pecadores ou filhos de pecadores, e por isso, eram mal vistos. Tal visão foi, inclusive, destruída por Jesus Cristo em uma das passagens relatadas no Evangelho de João (João 9:1-3), onde um deficiente visual obteve a cura.

Todavia, é possível encontrar deficientes que atingiram altos lugares nas sociedades antigas, como Homero, que era cego e foi o responsável por duas das obras antigas mais famosas sobre a Grécia, *Ilíada* e *Odisseia* (BOAS, 2018). Assim, com o avanço das civilizações foi possível observar a preocupação da integração das pessoas nessas condições na sociedade e, inclusive, isso fica muito mais explícito quando Aristóteles proclama “Tratar os desiguais de maneira igual constitui-se em injustiça” e quando Jesus Cristo anuncia que os deficientes não eram pecadores, quebrando, também, as crenças religiosas da época.

Voltando os olhos para os deficientes visuais, por volta do século XV com o desenvolvimento científico, a filosofia cria um conceito humanitário relacionado às pessoas com deficiência visual. A partir de então, essas condições são caracterizadas como um problema de saúde e, por isso, cria-se a necessidade de absorção de pessoas dentro dessas condições na sociedade, utilizando-se de métodos de integração do indivíduo na sociedade por meio da tecnologia assistiva.

Em questões nacionais, há um departamento especializado para inclusão dessas pessoas, chamado de Comitê de Ajudas Técnicas – CAT. Essa mesma instituição declara: "Tecnologia Assistiva é uma área do conhecimento, de característica interdisciplinar, que engloba produtos, recursos, metodologias, estratégias, práticas e serviços que objetivam promover a funcionalidade, relacionada à atividade e participação de pessoas com deficiência, incapacidades ou mobilidade reduzida, visando sua autonomia, independência, qualidade de vida e inclusão social" (ATA VII - Comitê de Ajudas Técnicas (CAT) - Coordenadoria Nacional para Integração da Pessoa Portadora de Deficiência (CORDE) - Secretaria Especial dos Direitos Humanos - Presidência da República, 2006).

A visão é um dos recursos do corpo humano mais importante e significativo para o relacionamento do indivíduo com a sociedade. Através da visão é que o indivíduo se relaciona com o mundo exterior, realizando, assim, a concretização de todos os sentidos do corpo humano. Entre as dificuldades que o deficiente encontra, a maior de todas é a locomoção, com a necessidade de detectar objetos e, obviamente, a dificuldade de se enquadrar na sociedade. Nesse cenário é que se faz necessário a criação de recursos embasados em tecnologia assistiva para viabilização da facilidade de atividades diárias do deficiente. Portanto, a elaboração de um projeto que atende as necessidades de inclusão do deficiente é essencial.

Respalhando as dificuldades encontradas do usuário, mencionadas acima, foi desenvolvido um projeto onde serão utilizados alguns recursos fundamentais estudados durante o curso de engenharia elétrica. Para a elaboração do protótipo foram aplicados e utilizados determinados métodos e materiais, como o sensor ultrassônico, responsável pela detecção de obstáculos, na parte de processamento de informações foi empregado o microcontrolador ESP32, módulo I2S para conversão sonora dos dados, para a captura e processamento de imagem foi adotado o ESP-32 CAM e uma bateria de 9V de 5500mAh para alimentação do circuito elétrico. Desta forma, com o auxílio dos componentes eletrônicos citados e conhecimentos adquiridos ao longo do curso, foi desenvolvido um projeto visando o suporte aos deficientes visuais.

## 1.1. Justificativa

De acordo com o Censo de 2010, cerca de 24% da população brasileira declarou ter algum grau de dificuldade em pelo menos uma das habilidades questionadas: enxergar, ouvir, caminhar ou subir degraus, ou possuir algum tipo de deficiência mental ou intelectual, representando um total de 46 milhões de pessoas. Dentre estas pessoas, cerca de 18% da população declarou possuir alguma dificuldade ao enxergar, onde 3,4% da população brasileira apresentou a deficiência total da visão (IBGE, 2010). A Pesquisa Nacional de Saúde em 2019 apontou que, cerca de 67% dos deficientes visuais se encontravam desempregados na semana de referência em que a pesquisa foi realizada, o que retrata a grande dificuldade deste grupo em se posicionar no mercado de trabalho (IBGE, 2019).

Atualmente, estima-se que 2,2 bilhões de pessoas no mundo apresentam algum tipo de deficiência visual, seja à curta ou longa distância, onde a perda de visão pode afetar pessoas de qualquer idade (ORGANIZAÇÃO MUNDIAL DA SAÚDE, 2021). Além disso, a cegueira traz consigo diversas dificuldades e limitações ao longo de toda a vida do deficiente visual, principalmente na infância, das quais destacam-se: a dificuldade de reconhecer o ambiente à sua volta, andar nas ruas e acelerar o ritmo, utilizar o computador, higiene pessoal e a prática da leitura e escrita (ASSOCIAÇÃO CATARINENSE PARA INTEGRAÇÃO AO CEGO, 2017). Sob este contexto, a acessibilidade possui como principal objetivo, garantir à pessoa com deficiência não apenas uma vida independente, mas também o exercício de seus direitos de cidadania e participação social (VAILATTI et al., 2019).

De acordo com o Art. 3º da Lei Nº 13146, a definição de acessibilidade é descrita como a possibilidade e condição de alcance para utilização, com segurança e autonomia, de espaços mobiliários, equipamentos urbanos, edificações, transportes, informação e comunicação, inclusive seus sistemas e tecnologias, bem como de outros serviços e instalações abertos ao público, de uso público ou privados de uso coletivo, tanto na zona urbana como na rural, por pessoa com deficiência ou com mobilidade reduzida (BRASIL, 2015). Aliada aos conceitos de acessibilidade, encontramos na evolução tecnológica um caminho para tornar as coisas mais fáceis para as pessoas sem deficiência, e possíveis para os deficientes (RADABAUGH, 1993, apud BERSCH, 2017).

A tecnologia assistiva corrobora com a acessibilidade possibilitando a inclusão e suporte às pessoas com deficiência funcionalizando a atividade e participação destas pessoas na sociedade a partir de soluções que façam uso de produtos, equipamentos, recursos,

metodologias, estratégias e serviços (VAILATTI et al., 2019). Dentre as tecnologias assistivas existentes voltadas ao deficiente visual, encontra-se o Braille Fácil que realiza a conversão de textos convencionais em braille, Dosvox que facilita o uso de computadores por meio de um sistema computacional e o leitor de tela para aparelhos celulares, o Talks (SECRETARIA DE CIÊNCIA, TECNOLOGIA E INOVAÇÃO DO DISTRITO FEDERAL, 2019). Seguindo a mesma linha de raciocínio dos exemplos citados, nosso projeto visa contribuir com mais uma solução voltada ao deficiente visual.

Portanto, tendo em vista a dificuldade de interação do deficiente visual com a sociedade e com o meio em que vivemos, o projeto apresentado neste trabalho visa contribuir com a vida do deficiente visual, apresentando mais uma solução direcionada à inclusão e acessibilidade com o uso da tecnologia. O principal objetivo deste projeto é proporcionar ao deficiente visual a capacidade de identificar objetos a partir de um dispositivo eletrônico, na qual emite uma notificação ao usuário sobre o objeto detectado em formato de áudio, fornecendo assim mais uma ferramenta de apoio ao deficiente visual.

## 1.2. Objetivos

Desenvolver um dispositivo eletrônico vestível sob uma estrutura de óculos que, atendendo aos princípios mínimos de ergonomia e com uma tecnologia de visão computacional embarcada, seja capaz de escanear, interpretar e comunicar, por meio de áudio, os objetos ou obstáculos detectados ao deficiente visual a uma distância de no mínimo 2 metros de distância.

1.2.1 Desenvolver o procedimento de computação gráfica aplicada na detecção de objetos em até 5 segundos desenvolvido em um óculos inteligente.

1.2.2 Identificar e informar ao usuário a presença de obstáculos a uma distância de no mínimo 2 metros.

1.2.3 Notificar ao usuário, mediante áudio, a captura obtida com uma acurácia de no mínimo 80%.

1.2.4 Apresentar um consumo de corrente total, somando-se todos os componentes, de no máximo 687,5mA, aliado à uma bateria recarregável de 5500mAh oferecendo uma autonomia de no mínimo 8 horas de uso contínuo.

## **2. Revisão bibliográfica**

### **2.1.IA e a Visão Computacional**

A visão computacional pode ser definida como a ciência responsável por proporcionar o sentido da visão à uma máquina. Desta forma, o recurso computacional é capaz de identificar o meio à sua volta, podendo extrair informações relevantes de imagens capturadas por alguns dispositivos, tais como câmeras de vídeo ou fotográficas, scanners e sensores, onde estas informações podem ser manipuladas e interpretadas (MILANO e HONORATO, 2014).

A visão computacional deu início na década de 1950, utilizando algumas das primeiras redes neurais, para realizar a detecção dos limites de objetos e classificação em categorias simples, como quadrados e círculos. Após 20 anos a visão computacional permitia interpretar textos, como manuscritos e digitados, isto ocorria com o reconhecimento óptico dos caracteres. Ao passar dos anos com o surgimento da internet, foi se avançando a tecnologia e muitas imagens foram disponibilizadas na rede mundial. Com a evolução do conjunto de dados, as máquinas começaram a ter facilidade em reconhecimento de pessoas em fotos e vídeos. Com o avanço da tecnologia muitas áreas se beneficiam da aplicação da visão computacional, um exemplo disso está no setor da segurança. Os sistemas de segurança estão fazendo o uso de reconhecimento facial, em locais públicos, privados e em dispositivos eletrônicos, como os celulares. Nas redes sociais isto é algo comum, por exemplo no aplicativo Facebook, ao publicar uma foto o sistema reconhece o conjunto de traços do rosto e sugere marcações nas fotos.

### **2.2.Tecnologia Assistiva**

Segundo Rita Bersch (2008) o termo tecnologia assistiva é ainda algo novo em nossa geração, utilizado na identificação de todo o conjunto de recursos e serviços para auxiliar, proporcionar ou ampliar habilidades funcionais de pessoas que possuem alguma deficiência e, dessa forma, proporcionar ao público alvo uma vida independente e inclusiva. O termo Tecnologia Assistiva vem da tradução Assistive Technology, onde foi oficialmente criado no ano de 1998 nos EUA, como principal elemento jurídico dentro da legislação norte-americana conhecida como Public Law 100-407, qual constitui com a outra lei ADA - American with Disabilities Act. Estes conjuntos de leis exercem os direitos dos cidadãos com deficiência nos

EUA, garantindo que os mesmos possuam uma vida mais independente, produtiva e inclusiva (BERSCH, 2005). A tecnologia assistiva vem sendo vista como uma área de grandes conhecimentos, práticas, recursos, metodologias e estratégias, que têm como objetivo aumentar a autonomia, qualidade de vida e inclusão social das pessoas que possuem alguma deficiência. Nesse pilar, incluem-se diversos profissionais de diferentes áreas, de psicologia à engenharia, que se empenham em desenvolver o que for necessário para garantir a aplicação da tecnologia assistiva em todos os lugares e pessoas em geral.

Um exemplo de tecnologia assistiva é o Avatar de Libras, desenvolvido por dois pesquisadores, na área de Engenharia Elétrica e de Computação da Unicamp (Universidade Estadual de Campinas). Ambos elaboraram um avatar em 3D, através de um personagem tridimensional, o qual reproduz os sinais a partir de palavras que são enviadas em forma de texto e traduzidas em libras em tempo real. Esse sistema pode ser também adaptado para qualquer equipamento computadorizado, como smartphones e tablets. Este projeto foi desenvolvido na acessibilidade dos deficientes auditivos, na ajuda do aprendizado em sala de aula, pois segundo uma pesquisa realizada em 2010 pelo IBGE (Instituto Brasileiro de Geografia e Estatística) cerca de 2,7 milhões de deficientes auditivos no Brasil foram alfabetizados por sinais de libras (AMARAL e MARTINO, 2012).

### **2.3.APIs e Serviços Google Cloud**

Google Cloud é uma plataforma que fornece recursos para aplicações web. Tem à sua disposição inúmeras ferramentas para os usuários gerenciarem suas aplicações na nuvem. Dentre os recursos fornecidos, podemos elencar alguns mais importantes como: APIs, máquinas virtuais, banco de dados, compartilhamento e processamento de dados, hospedagem de sites, servidores VPS, etc. Todas essas funcionalidades colaboram para a conectividade de conceitos digitais.

As APIs são bibliotecas que permitem interação com usuário final para utilização dos serviços Google Cloud. Com esses recursos disponíveis, o usuário pode integrar a conectividade na nuvem, armazenamento de dados, computação e a inteligência artificial (TRAFANIUC, 2021).

### **2.3.1. Google Cloud Vision**

No ano de 2016, o Google anunciou a liberação da sua nova API, a Google Cloud Vision. A funcionalidade da Google Cloud Vision se baseia no desenvolvimentos de aplicativos que consigam compreender o conteúdo de imagens, como identificação de objetos, frutas, entre outros, semelhante ao Google Photos. Além disso, tornando possível a identificação de expressões como alegria e surpresa, através do posicionamento dos olhos, nariz e boca. A API em REST (Transferência de Estado Representacional, refere-se a um conjunto de princípios e definições necessários para o desenvolvimento de um projeto com interfaces definidas (EQUIPE TOTVS, 2020)) contém modelos de machine learning, podendo detectar paisagens naturais e ler palavras impressas em imagens. Este recurso permite que desenvolvedores consigam barrar recursos impróprios em suas aplicações.

### **2.3.2. Google Translate**

O Google Translate é um software desenvolvido pelo Google com o simples objetivo de possibilitar a tradução automática de textos e documentos, de maneira rápida e gratuita. A ferramenta foi desenvolvida em 2006 e contava apenas com a tradução para dois idiomas: inglês e árabe, além de possuir alguns erros de gramática ao funcionar. Depois, com o decorrer dos anos, o Google Translate evoluiu e, hoje em dia, contém 130 (cento e trinta) idiomas que possibilitam uma tradução rápida, automática e confiável (ALVES, 2020).

A popularidade do software em território brasileiro é tão grande que o próprio Google indicou o país como líder em uso da ferramenta e, além disso, indicou que mais de 500 milhões de pessoas ao redor do mundo utilizam o recurso e traduzem, em média, 100 bilhões de palavras todos os dias (ALVES, 2016). Ainda, com o avanço do uso do Google Translate outras extensões foram criadas, como o Google Translate Toolkit que possibilita o uso apenas de traduções online e usa como base para seu funcionamento o banco de dados de traduções automáticas, fornecido pelo próprio Google Translate (GOMES, 2018). Por fim, é importante destacar o uso do Google Translate para fins universitários e acadêmicos. Em 2013 foi realizado um estudo entre diversos estudantes de faculdades que comprovaram o uso da ferramenta por alunos com proficiências baixa e média em um segundo idioma, e os mesmos comprovaram que o software atende as expectativas onde realiza muito bem o papel para que foi proposto (COSTA, 2013). Não obstante, a ferramenta ainda é usada para ensinar outras

línguas para crianças de diversas faixas etárias – possibilitando, assim, o aprendizado a partir da tradução de textos e, conseqüentemente, auxiliando na aprendizagem gramatical e de novas palavras (NEULS, 2015).

### **2.3.3. Google Text to Speech**

O Google Text to Speech consiste em uma ferramenta, ou melhor, uma API desenvolvida pela Google para possibilitar a conversão de textos em dados de áudios, como MP3 ou LINEAR16. Hoje em dia, o software possui em torno de 56 vozes, em 14 idiomas diferentes (IT FORUM, 2018). Basicamente, o funcionamento da ferramenta se dá pelo processo de conversão da entrada de texto – chamada de síntese – a partir de um texto bruto ou em formato SSML e, então, a conversão deste para um arquivo de áudio. Neste áudio, as vozes são de fala humana natural, assim, o texto soa como uma pessoa comum e ainda há a possibilidade de escolher a voz preferida, com a entonação e o idioma disponíveis.

Ademais, esse software de Text to Speech foi definido por Burnett e Shuang (2010) como “o processo de geração automática de fala a partir de texto ou texto anotado”. Entre os benefícios de seu uso encontra-se o aumento da acessibilidade para os usuários finais, que poderão ouvir textos em diferentes línguas, de maneira rápida e fácil, beneficiando pessoas estrangeiras que não falam a língua de um país, pessoas que não sabem ler e até mesmo idosos – que terão o conteúdo falado de maneira mais acessível. Ainda, para as grandes empresas e organizações, o Google Text to Speech providencia uma melhor experiência para o seu cliente final, criando assim um marketing mais valoroso e, ainda, possibilita outros benefícios como economia de tempo e dinheiro e uma implementação fácil dos conteúdos desejados dentro da rede de Internet, proporcionando, então, um aumento do alcance da marca da empresa.

### 3. Materiais e Métodos

O presente projeto consistiu em desenvolver um dispositivo eletrônico (óculos) com tecnologia de visão computacional embarcada para detecção de objetos. Desta forma, se fez necessário a divisão do planejamento e execução do projeto em 3 partes principais: construção física, hardware e desenvolvimento do software, com o intuito de se obter uma melhor performance durante o desenvolvimento do projeto, dividindo-se as demandas exigidas entre os participantes do grupo, além de manter uma ordem lógica, sequencial e organizada na execução do trabalho.

#### 3.1. Gestão do projeto

Para que o monitoramento e controle do projeto fosse executado com sucesso, foi necessário definir uma metodologia de acompanhamento de resultados e distribuição de demandas entre os componentes do grupo. Sendo assim, utilizamos alguns recursos de metodologias de gestão de projetos como o Scrum, com reuniões semanais de planejamento e distribuição de tarefas e o Kanban, classificando cada tarefa quanto ao seu andamento. Além disso, construímos algumas documentações como o cronograma e escopo do projeto para servir como referência durante todo o desenvolvimento do projeto. Utilizamos algumas ferramentas para auxílio no método de gestão do projeto, conforme descrito na tabela abaixo:

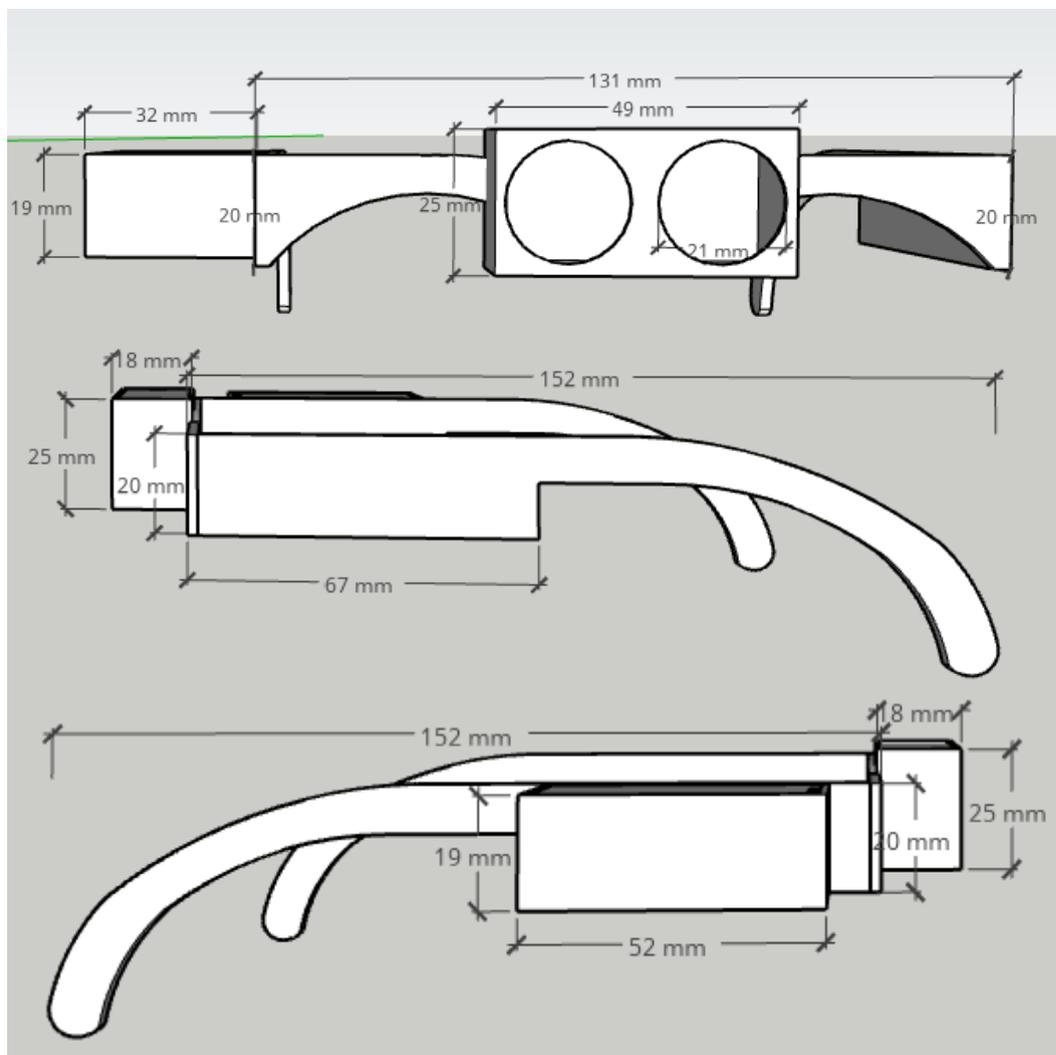
Tabela 1 - Método e ferramentas de gestão do projeto

<b>Ferramentas</b>	<b>Função</b>
Trello	Gerenciamento de atividades semanais
Google Sheets	Construção do cronograma e tabelas de registros de gastos
Google Meet	Reuniões semanais para divisões de tarefas e discussão do projeto

### 3.2.Estrutura física

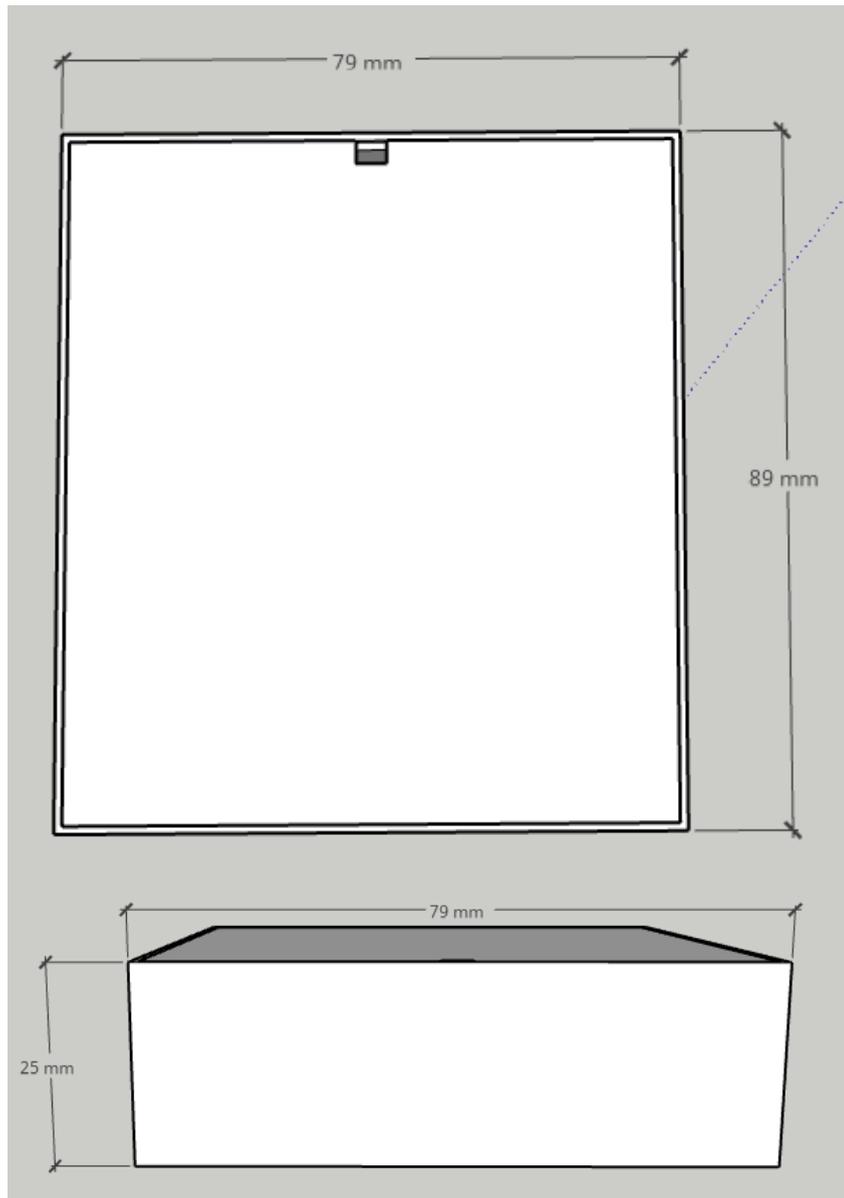
A parte física do projeto foi desenvolvida e baseada na mesma estrutura de um óculos, planejada e estruturada por meio da utilização do software Sketchup. Para a sua construção, foi realizada a impressão 3D utilizando-se o material PLA. A estrutura foi subdividida em 3 partes, composta por um esqueleto que possui o formato semelhante a um óculos convencional, aliado à duas caixas que foram acopladas nas laterais do esqueleto do óculos, para o armazenamento dos componentes. As dimensões de cada parte foram descritas na tabela abaixo:

**Figura 1** - Vistas Frontal e lateral da estrutura do óculos



*Fonte: Imagem de autoria própria*

**Figura 2** - Vistas superior e lateral do compartimento externo



*Fonte: Imagem de autoria própria*

### 3.3.Hardware

Toda a construção do hardware realizada no projeto foi baseada na utilização de placas do tipo módulo, que já possuem em sua estrutura os componentes pré interligados, montados e necessários para atender uma determinada função. A partir destes módulos, procedemos para a interligação entre os componentes de hardware, possibilitando assim a intercomunicação entre cada funcionalidade. Dentre os componentes utilizados, o único que foi desenvolvido sob demanda pelos integrantes do grupo, foi o conversor de tensão DC 5V utilizando o componente LM317. Antes da construção física do hardware, foram desenvolvidos os diagramas elétricos de cada componente e suas ligações com auxílio do software Proteus Design Suite, versão 8.9, Labcenter Electronics Ltd. Os componentes de hardware e suas especificações podem ser identificados na tabela abaixo:

Tabela 2- Lista de componentes e especificações

Componente	Especificações	Função
ESP32-CAM	<ul style="list-style-type: none"><li>- Tensão de operação: 5V;</li><li>- CPU: Xtensa® Dual-Core 32-bit LX6;</li><li>- Clock máximo: 240MHz;</li><li>- ROM: 448 Kbytes;</li><li>- RAM: 520 Kbytes SRAM;</li><li>- Flash: 4 MB PSRAM;</li><li>- Resolução da foto: 2 Megapixels;</li><li>- Wireless padrão 802.11 b/g/n;</li><li>- Conexão: Wifi 2.4Ghz (máximo de 150 Mbps);</li><li>- Modos de operação: STA/AP/STA+AP;</li><li>- Bluetooth: BLE 4.2;</li><li>- Portas GPIO: 16;</li></ul>	Captura de imagem
ESP32	<ul style="list-style-type: none"><li>- Tensão de Alimentação: 4,5 à 12,0 VDC</li><li>- CPU: Xtensa 32-Bit LX6 Dual Core;</li><li>- Clock: 80 à 240 MHz (Ajustável);</li><li>- Memória ROM: 448KB;</li><li>- Memória SRAM: 520Kb;</li><li>- Memória Flash Externa: 32-Bit de acesso e 4Mb;</li><li>- Tensão de nível lógico: 3,3VDC;</li><li>- Corrente de consumo: 80mA a 500mA</li><li>- WiFi 802.11 b/g/n: 2.4 à 2.5 GHz;</li><li>- Segurança WiFi: WPA/WPA2/WPA2-Enterprise/WPS;</li><li>- Criptografia WiFi: AES / RSA / ECC / SHA;</li><li>- Temperatura de trabalho: -40° à +85° C;</li><li>- Bluetooth 4.2 BR / EDR e BLE ( Bluetooth Low Energy);</li><li>- Tipos GPIO: Digital IO (36), ADC 12-Bits (16 canais), DAC 8-Bits (2 canais), Sensor Capacitivo (10 canais);</li><li>LNA pré-amplificador;</li></ul>	Processamento de dados
Sensor ultrassônico (HC-SR04)	<ul style="list-style-type: none"><li>- Alimentação: 5V DC</li><li>- Corrente de Operação: 2mA</li></ul>	Medição de distância e detecção de posição de

		- Ângulo de efeito: 15° - Alcance: 2cm a 4m - Precisão: 3mm	materiais
Decodificador de Áudio UDA1334A	de I2S	- Controlador: UDA1334A - Decodificador de áudio estéreo - Comunicação I2S - Tensão de operação: 3 V a 5 V - Saída P2 (fone de ouvido)	Decodificação de áudio em MP3
Bateria (Daweikala 9V USB)		- Capacidade: 5500mAh - Tensão nominal: 9V - Dimensões: 48x26x16mm - Peso: 27g - Composição: Íon de lítio	Alimentação do circuito
Conversor de tensão 5V			

A partir da análise das especificações de cada componente utilizado no projeto, foi identificado que a tensão ideal de alimentação e compatível para todo o circuito foi de 5V. No entanto, pelo fato de não existir nenhuma bateria ou fonte externa portátil neste valor, vendida ou fornecida comercialmente e que atenda aos critérios do projeto, sendo leve e de fácil manuseio, foi escolhida uma bateria de 9V recarregável. Ainda assim, para que a tensão de alimentação de 5V fosse atendida, foi construído um circuito regulador de tensão, utilizando-se o componente LM317 capaz de converter a tensão de 9V da bateria para o valor de 5V. Por fim, este circuito foi construído e soldado em uma placa ilhada de cobre, aliado aos resistores de 1kΩ e 3k1Ω e capacitores de 4,7μF e 100μF valores.

Considerando a necessidade de possíveis manutenções nos componentes, para correções e ajustes, especialmente os microcontroladores que recebem o código fonte para o funcionamento do circuito, foram construídas placas de suporte para os componentes: ESP32 CAM, ESP32 e o decodificador de áudio I2S, possibilitando uma maior facilidade no acoplamento e desacoplamento dos componentes do circuito principal. Estas placas foram construídas em placas ilhadas de cobre, elaboradas sob medida para cada componente, com pinos fêmea dispostos em 180° soldados nas mesmas. Assim, toda a ligação elétrica foi realizada entre as placas suporte e os demais componentes, deixando os microcontroladores e decodificador de áudio livres para ajustes, sem a necessidade de desconectar os demais componentes do circuito principal.

**Figura 3** - ESP32 CAM - Microcontrolador com câmera OV2640 integrada e ESP32 - Microcontrolador integrado



*Fonte: Imagem de autoria própria*

### 3.4. Desenvolvimento do Software

O desenvolvimento do software foi realizado em duas frentes principais: microcontroladores e servidor web. O funcionamento do software foi baseado na troca de dados, em que estas, realizam entre si, sendo dividido entre as seguintes etapas: captura da imagem, distância e envio ao servidor web, processamento e rotulagem da imagem, conversão da rotulagem e distância em centímetros para MP3 (MPEG 1 Layer-3) e execução do arquivo MP3 com a detecção. Na tabela abaixo podemos identificar cada recurso utilizado para o desenvolvimento do software, bem como suas principais funções no projeto:

Tabela 3 - Principais recursos utilizados no processo de detecção de objetos

<b>Recurso de software</b>	<b>Função dentro do projeto</b>
Google Vision API	Detecção de rótulos de objetos das imagens capturadas
Google Translate API	Tradução do rótulos gerados pelo Google Vision
Google Text to Speech API	Conversão da detecção em texto traduzida em áudio no formato MP3
Google Virtual Machine	Alocação de máquina virtual para execução do servidor
Microsoft Visual Code	Desenvolvimento do código em Node JS
Arduino IDE	Desenvolvimento do software gravado no ESP32 e ESP32 CAM

Os microcontroladores atuam nos dois extremos do fluxo de funcionamento, sendo o ESP32 CAM encontrado no início do fluxo, responsável pela captura da imagem, distância do objeto e envio ao servidor, e o ESP32 ao fim, responsável por acessar o servidor e executar o arquivo de áudio, com auxílio do decodificador de Áudio I2S UDA1334A. A construção e gravação do código fonte dos microcontroladores foram desenvolvidas na plataforma Arduino IDE, que, apesar de possuir uma linguagem de programação própria, possui grandes semelhanças em relação à linguagem C++. Dentro do código, foram incluídas algumas instruções para que os microcontroladores enviassem alguns retornos durante a execução, para validação do seu funcionamento por intermédio do protocolo de comunicação serial. Para comunicação entre os microcontroladores e o servidor web, foi utilizado o método HTTP POST para envio da imagem e da distância medida, que insere algumas modificações no servidor e executa as APIs de detecção de objetos. A conexão entre os microcontroladores e o servidor foi realizada a partir do IP externo do servidor alocado em uma máquina virtual.

O servidor Web foi desenvolvido em Node JS utilizando-se a plataforma Microsoft Visual Code. Este servidor é o responsável pelo recebimento das imagens capturadas pelo ESP32 CAM, execução das APIs e criação do arquivo de áudio de acordo com os resultados obtidos pelas APIs. Seu método de funcionamento consiste em: receber a imagem capturada e sua distância, executar a API Google Vision para detecção de rótulos, executar a API Google Translate para traduzir os resultados obtidos para o português e transformar os resultados traduzidos em um arquivo MP3. A instância da máquina virtual utilizada possui as seguintes características:

Tabela 4 - Especificações da máquina virtual

<b>Característica</b>	<b>Descrição</b>
Nome da instância	smart-glass-server
Nome da máquina	e2-micro
Memória RAM	1 GB
Plataforma de CPU	AMD Rome
Armazenamento	10 GB
Sistema Operacional	debian-10-buster-v20220118 Debian, Debian GNU/Linux, 10 (buster), amd64 built on , supports Shielded VM features
Zona de alocação da instância	us-east1-b

## 4. Resultados e discussões

### 4.1. Estrutura física do protótipo

Nas etapas iniciais do projeto foi adotada como premissa a alocação de todos os componentes eletrônicos na estrutura do óculos, visando assim facilitar o modo de uso do usuário durante a utilização do óculos. Com base nesta premissa, foi desenvolvida uma armação e estrutura de caixas em impressão 3D, utilizando-se o material poliácido láctico (PLA).

No entanto, após os testes de alocação de componentes nas caixas e fixação na armação do óculos, foi notado um grande peso e desconforto durante a utilização do protótipo, maiores do que o esperado. A partir deste peso excessivo identificado e exercido na região do rosto do usuário, foram identificados fatores que poderiam comprometer o bom funcionamento do projeto, que se encontram descritos na seguinte tabela:

Tabela 5 - Principais problemas encontrados no protótipo e suas causas

<b>Fator</b>	<b>Causa</b>
Problemas com alinhamento do sensor e câmera durante a detecção de objetos	Peso excessivo e mal distribuído na armação
Desconforto e baixa ergonomia	Posição desfavorável das caixas nas laterais do protótipo causando um desconforto nas laterais do rosto do usuário
Baixa fixação do óculos no rosto	Peso excessivo e mal distribuído causando o escorregamento do óculos durante inclinações com a cabeça ou movimentos bruscos

Sendo assim, tendo em vista tais problemas encontrados foram elaboradas adaptações na estrutura, separando-se os componentes da armação, com exceção do sensor ultrassônico e ESP32 CAM que dependem do posicionamento na estrutura do protótipo para a medição das distâncias e captura da imagem do objeto. Após os ajustes na estrutura, obtivemos os seguintes resultados:

Tabela 6 - Dimensões da estrutura física ajustada

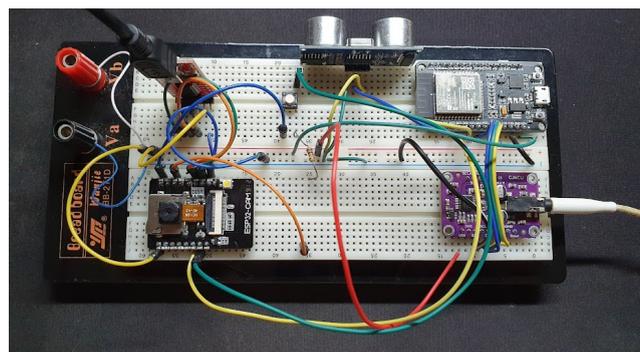
Componente	Dimensões
Armação do óculos	131mm (L) X 22m (A) X 153mm (C)
Caixa lateral	41mm (L) X 27mm (A) X 52mm (C)
Compartimento externo	81mm (L) X 25mm (A) X 91mm (C)

Com os ajustes efetivados, foi observado que, a partir de testes funcionais e vestindo-se o óculos, os fatores mencionados na tabela 5 foram solucionados, eliminando ou amenizando as possíveis situações que poderiam prejudicar o bom funcionamento do protótipo.

#### 4.2. Consumo elétrico

Para executar as devidas medições de consumo elétrico de todo o circuito, todos os componentes foram interligados em uma protoboard e medidos ainda na fase de desenvolvimento do hardware do projeto. A disposição de cada componente na protoboard pode ser identificada na imagem abaixo:

**Figura 4** - Circuito montado na Protoboard



*Fonte: Imagem de autoria própria*

Partindo-se desta mesma disposição, foram realizadas medições de consumo de corrente contínua e níveis de tensão elétrica contínua. Para a realização destas medições, foi utilizado o instrumento Multímetro Digital da marca Hikari (UNICOBA IMPORTAÇÃO E EXPORTAÇÃO LTDA), modelo HM-2030. A relação de valores especificados confrontados com os valores medidos foram resumidos e anexados na tabela 7 abaixo:

Tabela 7 - Medições de corrente e tensão

Componente	$I_{rms}$ medida[mA]	Tensão DC [V]
ESP32 CAM	97,1	5,06
ESP32	63,2	5,13
Decodificador de Áudio I2S UDA1334A	76,0	5,16
Sensor Ultrassônico	30,0	5,19
<b>Total</b>	<b>266,3</b>	<b>N/A</b>

Figura 5 - Multímetro Digital Hikari - HM 2030



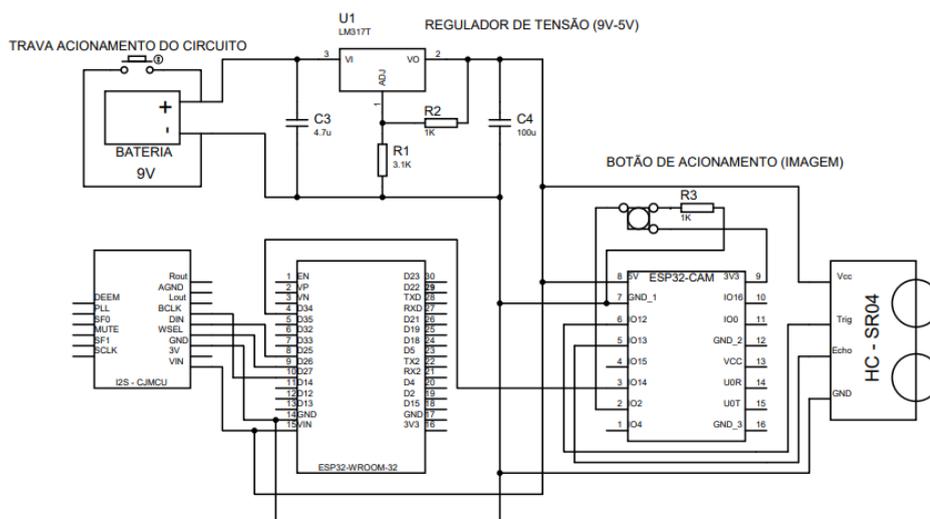
*Fonte: Imagem de autoria própria*

Somando-se os consumos de corrente de cada componente, chegamos a um valor de consumo total de 286,3mA. Considerando que a fonte de energia utilizada para a alimentação, se trata de uma bateria de 9V e 5500mAh, estima-se que o protótipo deva ter uma autonomia de aproximadamente 20 horas de uso, valor este que pode variar de acordo com a frequência de utilização e capturas de imagem.

### 4.3. Ligação elétrica

O esquema de ligação elétrica realizado entre os componentes do projeto pode ser observado no esquema elétrico abaixo:

**Figura 6 - Circuito elétrico geral**

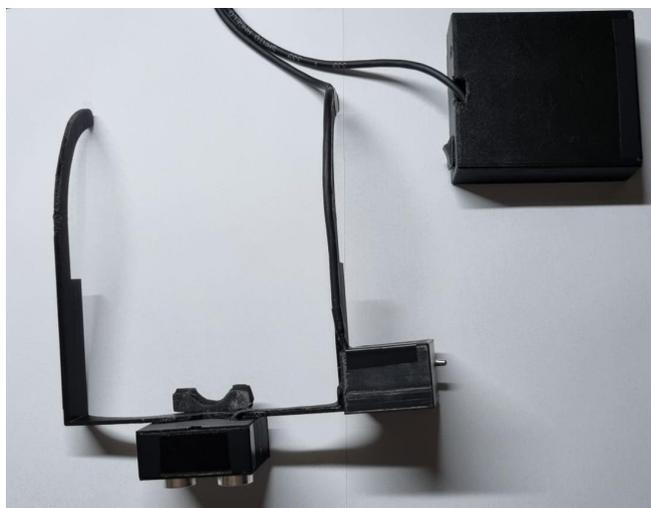


*Fonte: Imagem de autoria própria*

Devido à divisão dos componentes quanto à sua alocação no protótipo, foram definidas 2 regiões principais de fixação das placas: a primeira na estrutura do óculos e a segunda em um compartimento, em formato de caixa, posicionado externamente ao óculos.

Os componentes sensor ultrassônico e ESP32 CAM foram alocados na primeira região, devido à necessidade de realizar a captura e medição do objeto a ser identificado. Sendo assim, cada componente foi alocado na armação do óculos e ligados eletricamente conforme a figura abaixo:

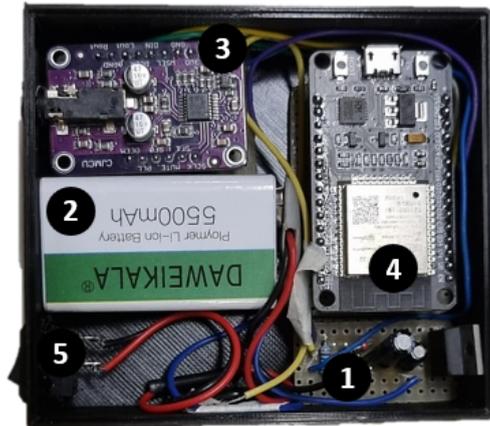
**Figura 7 - Protótipo**



*Fonte: Imagem de autoria própria*

Os demais componentes, como a bateria, ESP32, placa conversora de tensão 5V e decodificador de áudio I2S, foram posicionados e conectados no compartimento externo. Além disso, foram realizadas furações no compartimento possibilitando a exposição de uma chave para ligar ou desligar todo o circuito, e do conector de áudio. A disposição dos componentes pode ser identificada na imagem abaixo:

**Figura 8** - Disposições dos componentes alocados no compartimento externo



1. Regulador de tensão
2. Bateria Recarregável
3. Módulo I2S
4. ESP32
5. Chave On/Off

*Fonte: Imagem de autoria própria*

Por fim, para possibilitar a conexão entre as duas regiões foi confeccionado um chicote duplo de cabos flat com 6 e 7 vias. Levando-se em conta o comprimento do cabo de 1,12 metros, para validar o bom funcionamento do circuito e evitar que o cabo afetasse o desempenho do circuito, foram realizadas medições quanto à sua resistência elétrica. O resultado obtido foi de  $0,8 \Omega$ , valor este que não apresenta riscos ao funcionamento do circuito devido ao seu baixo valor em relação ao circuito como um todo. O resultado da medição, assim como as demais características do cabo podem ser identificadas na tabela 8

#### 4.4. Detecção de objetos e distância

Por fim, com a conclusão de todas as etapas de desenvolvimento, foram executados testes para a validação do funcionamento do projeto. Como objeto de amostra para a detecção e identificação, foi utilizado um quebra cabeça mecânico, conhecido popularmente como cubo mágico, por apresentar um aspecto colorido que se destaca em um ambiente uniforme, e que consiste em sua estrutura uma forma geométrica simples, conforme ilustrado na figura 9.

**Figura 9** - Objeto utilizado nos testes de identificação



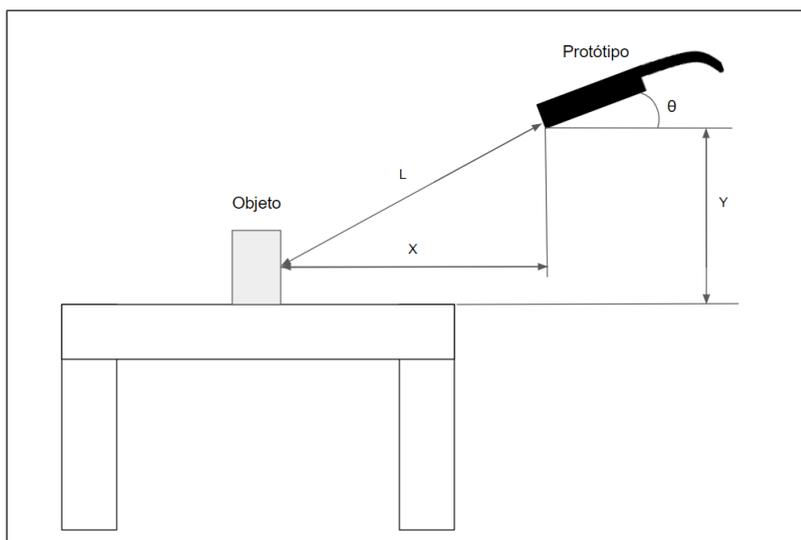
*Fonte: Imagem de autoria própria*

Com a escolha do objeto, passamos para a definição do ambiente e procedimentos adotados nos testes. O teste consistiu em posicionar o objeto sobre uma superfície plana de cor uniforme, como uma mesa de madeira, situado a uma distância  $X$  do usuário do óculos e a uma altura  $Y$  em relação à superfície plana onde se situa o objeto. Além disso, foram coletados e registrados os ângulos de inclinação do óculos em cada teste. A configuração de teste pode ser identificada na figura 11 e seus resultados e parâmetros adotados, na tabela 9.

Analisando-se a relação entre os parâmetros de distância e posição e os resultados obtidos, identificamos que a precisão da identificação realizada pelo óculos se mostra inversamente proporcional em relação à distância  $X$ . Essa queda da precisão ocorre devido ao enquadramento do objeto na imagem, perda de detalhes e precisão da imagem em pixels que se intensificam conforme o aumento da distância entre o usuário e o objeto. Além disso, a variável do ambiente pode também desfavorecer a identificação do objeto, uma vez que ao ampliar o campo de fotografia, outros objetos do ambiente podem ser enquadrados na imagem capturada pela câmera, o que pode gerar um fator de aleatoriedade na escolha do objeto para a detecção realizada pela API Google Vision.

Em linhas gerais, o protótipo se mostrou eficiente na identificação do objeto utilizado nos testes em curta distância (20 e 50 cm), na medição da distância  $L$ , na qual se manteve abaixo de 20% de erro, executando todo o processo de software e hardware descritos nos tópicos anteriores e desta forma, alcançando com sucesso a proposta e os objetivos descritos no início do projeto.

**Figura 10 - Objeto para o teste de identificação**



*Fonte: Imagem de autoria própria*

Tabela 8 - Parâmetros e resultados obtidos nos testes

Teste	X	Y	L	$\theta$	Resultado da identificação	Distância L medida pelo sensor ultrassônico
1	20cm	75cm	58cm	60°	<ul style="list-style-type: none"><li>• Cubo de Rubik</li><li>• Quebra-cabeça mecânico</li><li>• Retângulo</li></ul>	48cm
2	50cm	75cm	80cm	45°	<ul style="list-style-type: none"><li>• Cubo de Rubik</li><li>• Retângulo</li><li>• Madeira</li></ul>	70cm
3	100cm	75cm	118cm	35°	<ul style="list-style-type: none"><li>• Mesa</li><li>• Móveis</li><li>• Madeira</li></ul>	120cm

## **5. Considerações Finais**

Este projeto tem como objetivo promover a inclusão de pessoas com deficiência visual por meio de tecnologia assistiva, permitindo a capacidade de identificar objetos. Foram explorados os recursos fornecidos durante a graduação de engenharia elétrica para o desenvolvimento do protótipo, ao qual atendeu a eficiência almejada definida no escopo do projeto.

Com base no escopo do projeto apresentado, bem como os estudos realizados sobre inteligência artificial e visão computacional, tecnologia assistiva, APIs e serviços Google Cloud, Google Cloud Vision, Google Translate, Google Text to Speech e análise e desenvolvimento de circuitos elétricos é que se encerra este trabalho apresentado.

As tecnologias apresentadas unitariamente promovem recursos que agregam valor ao projeto executado. A soma de todas as tecnologias favorecem o engajamento da eficiência do projeto, taxados como objetivos principais. Entre eles, o principal objetivo é, contudo, a tecnologia assistiva aos deficientes visuais.

Os resultados obtidos neste projeto, evidenciam a complexidade dessa realidade, com diversos fatores que influenciam a apropriação das tecnologias para inclusão dessas pessoas na sociedade. De uma forma genérica, houve um propósito a ser alcançado como finalidade principal na execução do projeto.

Salienta-se que através da utilização das tecnologias assistivas, potencializa-se a autonomia do usuário, o que se deu o desenvolvimento do projeto. Finaliza-se o mesmo, abrindo novas portas e possibilidades para novos desafios e melhorias na vida do usuário.

## 6. Referências

RODNEY NASCIMENTO E CLEBER AKIRA NAKANDAKARE. **CPQD**. Sistemas embarcados: a chave para a inteligência das “coisas”. *[S.l.]*. CPQD, 2019. Disponível em: <https://www.cpqd.com.br/internet-das-coisas-iot-backup/sistemas-embarcados-a-chave-para-a-inteligencia-das-coisas/>. Acesso em: 25 set. 2021.

GRUPO BINÁRIO INTEGRAÇÃO E SERVIÇOS. **Grupo Binário Integração e Serviços**. IoT é uma grande oportunidade para acessibilidade. *[S.l.]*. Grupo Binário Integração e Serviços, 2018. Disponível em: <https://www.binarionet.com.br/iot-e-uma-grande-oportunidade-para-acessibilidade/>. Acesso em: 29 set. 2021.

CASA CIVIL. **gov.br**. Mais de 82% dos domicílios brasileiros têm acesso à internet. *[S.l.]*. GOVBR, 2021. Disponível em: <https://www.gov.br/casacivil/pt-br/assuntos/noticias/2021/abril/mais-de-82-dos-domicilios-brasileiros-tem-acesso-a-internet#:~:text=A%20Pesquisa%20Nacional%20por%20Amostra,nacionais%20possuem%20acesso%20%C3%A0%20internet.&text=Nos%20domic%C3%ADlios%20urbanos%2C%20a%20utiliza%C3%A7%C3%A3o,%2C7%25%2C%20em%202019..> Acesso em: 5 out. 2021.

GOOGLE. **cloud.google**. Documentação do Cloud Vision. *[S.l.]*. cloud.google, 2016. Disponível em: <https://cloud.google.com/vision/docs>. Acesso em: 8 out. 2021.

FERNANDA DEL CLARO. **IMG FAE EDU**. O avanço tecnológico no mundo econômico. Curitiba: Vitrine da Conjuntura, 2009. Disponível em: <https://img.fae.edu/galeria/getImage/45/5423643835714016.pdf>. Acesso em: 20 out. 2021.

BRUNO SPRICIGO. **Pollux, Part of Accenture**. São Paulo: Pollux, Part of Accenture, 2016. Disponível em: Spricigo, B. (2018, 26 de março). Resumo sobre Indústria 4.0: entenda rapidamente os conceitos e benefícios. Pollux. <https://pollux.com.br/blog/resumo-sobre-industria-4-0-entenda-rapidamente-os-conceitos-e-beneficios/>. Acesso em: 2 nov. 2021.

FRANCESCO AZZOLA. **SWA**. São Paulo: SWA, 2021. Disponível em:  
<https://www.survivingwithandroid.com/esp32-cam-object-detection-tensorflow-js/>. Acesso em: 4 out. 2021.

OPENJS FOUNDATION. **nodejs**. API reference documentation. [S.l.]. nodejs, 2015. Disponível em: <https://nodejs.org/en/docs/>. Acesso em: 4 nov. 2021.

JUNE TAO CHING. **towardsdatascience**. How to manage files in Google Drive with Python. [S.l.]. towardsdatascience, 2019. Disponível em:  
<https://towardsdatascience.com/how-to-manage-files-in-google-drive-with-python-d26471d91ecd>. Acesso em: 2 nov. 2021.

MARCOS ZAMBONI. **etelg**. Proteus VSM. São Paulo: Etelg, 2011. Disponível em:  
[http://www.etelg.com.br/downloads/eletronica/apostilas/Proteus\\_ISIS\\_ZAMBONI.pd](http://www.etelg.com.br/downloads/eletronica/apostilas/Proteus_ISIS_ZAMBONI.pd). Acesso em: 8 out. 2021.

CALZAVARA, João Vitor *et al.* SAÚDE E ACESSIBILIDADE: A REALIDADE DOS DEFICIENTES VISUAIS NO BRASIL. Disponível em:  
<https://proceedings.science/cied/papers/saude-e-acessibilidade--a-realidade-dos-deficientes-visuais-no-brasil->. Acesso em: 13 mar. 2022.

OLIVEIRA, Erival da Silva *et al.* Direito das pessoas com deficiência para provas de concursos, 2019

SECRETARIA GERAL. LEI nº 13.146, de 6 de julho de 2015. Disponível em:  
[http://www.planalto.gov.br/ccivil\\_03/\\_Ato2015-2018/2015/Lei/L13146.htm](http://www.planalto.gov.br/ccivil_03/_Ato2015-2018/2015/Lei/L13146.htm). Acesso em: 13 mar. 2022.

INTRODUÇÃO À TECNOLOGIA ASSISTIVA. INTRODUÇÃO À TECNOLOGIA ASSISTIVA, 2017. Disponível em:  
[https://www.assistiva.com.br/Introducao\\_Tecnologia\\_Assistiva.pdf](https://www.assistiva.com.br/Introducao_Tecnologia_Assistiva.pdf). Acesso em: 13 mar. 2022.

EDUCA, IBGE. PESSOAS COM DEFICIÊNCIA. *In*: PESSOAS COM DEFICIÊNCIA, 1 fev. 2019. Disponível em:  
<https://educa.ibge.gov.br/jovens/conheca-o-brasil/populacao/20551-pessoas-com-deficiencia.h>

tml. Acesso em: 13 mar. 2022.

IBGE. Tabela 8209 - Pessoas com deficiência visual, por condição em relação à força de trabalho na semana de referência e situação do domicílio (2019). Disponível em:

<https://sidra.ibge.gov.br/tabela/8209#resultado>. Acesso em: 13 mar. 2022.

OMS, Organização Mundial da Saúde. Blindness and vision impairment, 14 out. 2021.

Disponível em:

<https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment>. Acesso em: 16 mar. 2022.

SALOMÃO, Solange R. Visual impairment and blindness: an overview of prevalence and causes in Brazil. Anais da Academia Brasileira de Ciências, 27 abr. 2009. Disponível em:

<https://www.scielo.br/j/aabc/a/ZrtkFzklVXFQt5RBdT4R6Ks/?format=pdf&lang=en>. Acesso em: 16 mar. 2022.

LOPES, Daiane Duarte. Psicologia e a pessoa com deficiência, 2018.

FLORIPA, ACIC. 5 dificuldades que crianças cegas enfrentam, 30 jan. 2017. Disponível em:

<https://medium.com/@acicfloripa/5-dificuldades-que-crian%C3%A7as-cegas-enfrentam-89938aefbe81>. Acesso em: 16 mar. 2022.

SECTI, Secretaria de Ciência, Tecnologia e Inovação. Conheça os benefícios da tecnologia assistiva, 2 set. 2019. Disponível em:

<https://www.secti.df.gov.br/conheca-os-beneficios-da-tecnologia-assistiva/>. Acesso em: 20 mar. 2022.