

# Áurea: Óculos de Inteligência Artificial de Baixo Custo para Descrição de Cenas e Assistência de Mobilidade para Indivíduos com Deficiência Visual

Vinicius Marques Santos Silva<sup>1</sup>, Charles Tim Batista Garrocho<sup>1</sup>

<sup>1</sup>Instituto Federal de Educação, Ciência e Tecnologia de Minas Gerais,  
Ouro Branco, Minas Gerais, Brasil

viniciusmrqs@outlook.com, charles.garrocho@ifmg.edu.br

**Abstract.** *Visual impairment poses daily challenges, such as difficulty navigating environments with obstacles and accessing visual information, which can limit autonomy and social integration. To address these barriers, a low-cost pair of glasses has been developed, combining artificial intelligence and an ultrasonic sensor. This device generates environmental descriptions using artificial intelligence and alerts users to nearby obstacles through a vibration system, aiming to restore some independence and promote greater social inclusion for visually impaired individuals.*

**Resumo.** *A deficiência visual impõe desafios diários, como a dificuldade de navegar em ambientes com obstáculos e acessar informações visuais, o que pode limitar a autonomia e a integração social. Para enfrentar essas barreiras, foi desenvolvido um óculos de baixo custo que combina inteligência artificial e sensor de ultrassom. Este dispositivo gera descrições do ambiente por meio de inteligência artificial e alerta sobre obstáculos próximos através de um micro motor de vibração, buscando devolver parte da independência e promover maior inclusão social para pessoas com deficiência visual.*

## 1. Introdução

Pessoas com deficiência visual enfrentam desafios constantes para se locomover de maneira segura e autônoma, especialmente em ambientes urbanos com infraestrutura inadequada. A ausência de sinalização tátil eficiente, a presença de obstáculos inesperados, como buracos, postes e veículos mal posicionados, e a falta de tecnologias acessíveis para navegação aumentam significativamente o risco de quedas e colisões [Khan and Khusro 2021]. Como consequência, muitos indivíduos evitam frequentar locais desconhecidos, reduzindo sua mobilidade e limitando suas oportunidades sociais e profissionais. Esse contexto pode contribuir para um ciclo de isolamento e dependência, afetando sua qualidade de vida e participação na sociedade [Hakobyan et al. 2013].

Além da locomoção, o acesso a informações visuais é outro desafio significativo. Sem um acompanhante ou um guia, pessoas com deficiência visual não conseguem perceber detalhes essenciais do ambiente, como a localização de escadas, portas, placas e outros elementos [Kuriakose et al. 2022]. Essa limitação reforça a necessidade de tecnologias que permitam tanto a navegação autônoma quanto a interpretação eficiente do espaço ao redor.

De acordo com a Pesquisa Nacional de Saúde (PNS) de 2019, aproximadamente 6,7% da população brasileira possui algum grau de deficiência visual, totalizando cerca de 14 milhões de pessoas. Dessas, aproximadamente 506 mil são cegas, enquanto cerca de 7 milhões possuem grande dificuldade para enxergar [IBGE 2019]. Apesar de avanços legais, como a Lei Brasileira de Inclusão (LBI) nº 13.146/2015 e a Norma ABNT NBR 9050/2015, a implementação prática de acessibilidade ainda é limitada [Brasil 2015]. Diante desse cenário, torna-se essencial o desenvolvimento de soluções inovadoras que garantam maior independência para essas pessoas.

Atualmente, existem diversas tecnologias assistivas para mobilidade de deficientes visuais. Estudos recentes [Yang 2024, Bai 2017, Li 2017] exploraram soluções vestíveis baseadas em sensores de profundidade e câmeras estéreo para navegação assistida. No entanto, muitas dessas abordagens dependem de hardware caro e volumoso, dificultando sua adoção em larga escala. Outras soluções utilizam inteligência artificial (IA) para converter imagens em áudio [Diniz and Junior 2024], mas o alto custo ainda é uma barreira para a maioria dos usuários.

Neste contexto, este trabalho propõe o *Áurea*, um dispositivo vestível de baixo custo que integra IA e sensores ultrassônicos para auxiliar deficientes visuais. O sistema combina duas funcionalidades principais: (i) detecção de obstáculos e alertas táteis progressivos; e (ii) audiodescrição em tempo real, fornecendo informações detalhadas sobre o ambiente. A tecnologia emprega modelos de visão computacional e algoritmos de segmentação para destacar elementos relevantes da cena, além de uma síntese de voz humanizada para tornar a interação mais intuitiva. Além disso, sua arquitetura foi projetada para garantir eficiência energética e operação prolongada sem necessidade de recargas frequentes.

O restante deste artigo está organizado da seguinte forma: a Seção 2 apresenta conceitos fundamentais sobre audiodescrição e os trabalhos relacionados. A Seção 3 descreve a arquitetura do sistema *Áurea* e o detalhamento de seu funcionamento. Na Seção 4, são apresentados detalhes técnicos de hardware e software da prova de conceito e os testes realizados. A Seção 5 apresenta e discute uma avaliação do *Áurea* e a análise de desempenho do dispositivo. Por fim, a Seção 6 apresenta as conclusões e sugestões para trabalhos futuros.

## **2. Background**

Esta seção apresenta conceitos essenciais para o desenvolvimento do *Áurea*, abordando a importância da audiodescrição na inclusão social. Também é apresentada e discutida uma revisão dos principais trabalhos relacionados ao *Áurea* no contexto de dispositivos assistivos para pessoas com deficiência visual.

### **2.1. Importância da Audiodescrição na Inclusão Social**

A acessibilidade é um fator fundamental para a inclusão de pessoas com deficiência visual em diferentes contextos sociais, educacionais e profissionais. A audiodescrição desempenha um papel essencial nesse processo, pois converte informações visuais em descrições verbais, permitindo que indivíduos cegos ou com baixa visão compreendam e interajam melhor com o ambiente ao seu redor [Franco and Silva 2010].

A presença da audiodescrição em diversas áreas, como teatro, cinema, museus e materiais educacionais, amplia o acesso à cultura e ao aprendizado. Além disso, essa tecnologia fortalece a autonomia dos usuários, permitindo uma experiência mais independente e interativa [e Inclusão 2016].

Com os avanços da IA, sistemas automatizados têm se tornado cada vez mais precisos na identificação de objetos, expressões faciais e elementos contextuais, possibilitando descrições mais detalhadas e dinâmicas. No entanto, a revisão humana continua sendo essencial para garantir a precisão e sensibilidade na interpretação do conteúdo [MWPT 2024].

No contexto educacional, a audiodescrição tem sido amplamente empregada para tornar materiais pedagógicos acessíveis, incluindo a descrição de gráficos, tabelas e equações matemáticas. Ferramentas modernas já permitem a conversão de imagens e diagramas em descrições textuais detalhadas, proporcionando uma experiência de aprendizado mais inclusiva [UNB 2025].

Além do impacto na educação e no acesso à informação, a audiodescrição contribui para uma sociedade mais equitativa. A democratização do acesso a conteúdos audiovisuais e informativos amplia a participação ativa de pessoas com deficiência visual em diversos espaços sociais, promovendo a igualdade de oportunidades à pessoas com deficiência visual [SHOWCASE 2023].

## 2.2. Trabalhos Relacionados

O desenvolvimento de tecnologias assistivas para pessoas com deficiência visual tem sido alvo de diversas pesquisas, resultando em soluções que combinam sensores, IA e dispositivos vestíveis [Lavric et al. 2024]. No entanto, muitas dessas tecnologias ainda enfrentam desafios relacionados ao custo, portabilidade e usabilidade. Nesta seção, discutimos as principais abordagens na área, destacando contribuições relevantes e lacunas que o presente trabalho busca endereçar.

Na literatura é possível encontrar diferentes dispositivos e sistemas que foram propostos para auxiliar a mobilidade de deficientes visuais. Dispositivos baseados em sensores ultrassônicos, como os investigados por Bai et al. [Bai 2017], demonstraram eficiência na detecção de obstáculos. Abordagens utilizando visão computacional [Li 2017] também mostraram potencial na interpretação do ambiente. Contudo, muitos desses dispositivos empregam hardware de alto custo ou possuem um design volumoso, dificultando sua adoção em larga escala. Além disso, soluções comerciais, como o *OrCam MyEye* e o *Envision Glasses*, possuem preços elevados, tornando-se inacessíveis para grande parte da população [Diniz and Junior 2024].

A revisão de Wang et al. [Wang 2023] categorizou as aplicações da IA para acessibilidade visual em duas frentes principais: (i) diagnóstico de doenças oculares por meio de aprendizado profundo e (ii) desenvolvimento de dispositivos assistivos para navegação autônoma. Os autores destacam que modelos baseados em redes CNN (*Convolutional Neural Network*) e reconhecimento de padrões têm sido amplamente utilizados para criar sistemas de audiodescrição e detecção de obstáculos. No entanto, desafios como a adaptação a condições de iluminação variáveis e a eficiência na identificação de objetos móveis ainda precisam ser aprimorados [Wang 2023].

Cortez et al. [Cortez and Almeida 2022] propuseram um protótipo de óculos inteligentes baseado no *Raspberry Pi 2*, utilizando algoritmos YOLOv3 e OCR (*Optical Character Recognition*) para detecção de objetos e reconhecimento de texto. O sistema identificava até 15 objetos diferentes e interpretava placas de sinalização em inglês, convertendo essas informações em áudio. Entretanto, limitações foram identificadas, como a dependência de iluminação adequada para o funcionamento eficiente e dificuldades no reconhecimento de fontes estilizadas [Cortez and Almeida 2022].

A pesquisa de Muhsin et al. [Muhsin et al. 2024] analisou diversos dispositivos assistivos disponíveis no mercado, categorizando-os conforme a tecnologia empregada: (i) sistemas baseados em sensores ultrassônicos, (ii) abordagens com visão computacional para reconhecimento do ambiente e (iii) soluções híbridas que integram diferentes formas de feedback, como áudio e vibração. Os autores ressaltam que, apesar dos avanços tecnológicos, muitos dispositivos apresentam desafios de usabilidade devido à complexidade da interface e ao tempo de adaptação necessário [Muhsin et al. 2024].

Além dos dispositivos vestíveis, a IA tem sido amplamente explorada para melhorar a acessibilidade. Os autores Yang et al. [Yang 2024] desenvolveram um sistema de assistência baseado em redes neurais profundas e sensores de profundidade, utilizando a câmera *Intel RealSense R200* para mapeamento do ambiente e um fone de condução óssea para feedback auditivo. Embora o sistema proposto tenha se mostrado eficiente na navegação interna e externa, o alto custo dos sensores empregados limita sua acessibilidade [Yang 2024].

Ademais, Li et al. [Li 2017] propuseram uma abordagem baseada em câmeras estéreo para auxiliar na locomoção de deficientes visuais, permitindo a identificação de objetos e estruturas ao redor do usuário. No entanto, a adaptação desse sistema para ambientes externos ainda enfrenta desafios devido às variações de iluminação e movimentação de objetos [Li 2017].

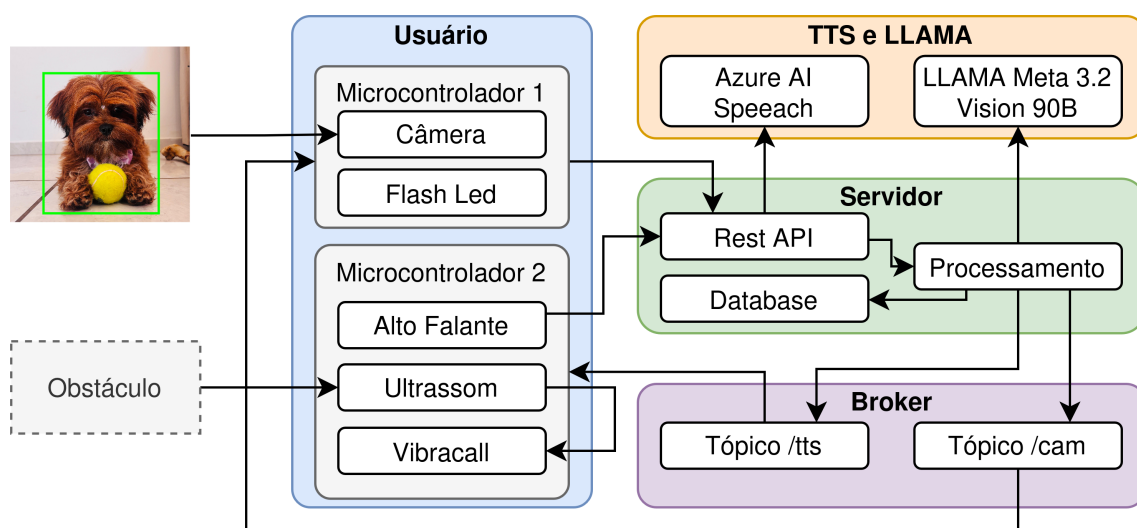
A partir dessa análise, observa-se que os principais desafios enfrentados por dispositivos assistivos incluem acessibilidade financeira, portabilidade e eficiência na identificação de objetos em diferentes condições ambientais. O presente trabalho busca contribuir para essa área por meio do desenvolvimento de um dispositivo compacto e de baixo custo, integrando sensores ultrassônicos para detecção de obstáculos e um sistema de IA para audiodescrição em tempo real. Dessa forma, pretende-se oferecer uma solução inovadora e acessível, promovendo maior autonomia para pessoas com deficiência visual.

### **3. Arquitetura e Funcionamento do Áurea**

O *Áurea* foi projetado para integrar tecnologias de IA e sensores táteis, permitindo a descrição de ambientes e a detecção de obstáculos para auxiliar pessoas com deficiência visual. Nesse contexto, a arquitetura do sistema *Áurea* se baseia em uma divisão modular, com componentes distribuídos entre o dispositivo usuário, o servidor e um *broker* MQTT (*Message Queuing Telemetry Transport*), equilibrando eficiência e acessibilidade e garantindo uma comunicação eficiente e responsiva. O processamento das imagens ocorre de forma remota, onde um modelo de visão computacional interpreta a cena e gera descrições textuais que são, posteriormente, convertidas em áudio e enviadas ao dispositivo do usuário.



A Figura 1 apresenta a arquitetura do sistema. A separação entre captura de imagens, processamento remoto e feedback tátil permite um funcionamento otimizado, reduzindo a carga computacional nos microcontroladores e aumentando a autonomia energética do dispositivo. Além disso, a utilização de protocolos leves, como MQTT, possibilita uma comunicação rápida e confiável entre os componentes, assegurando que o usuário receba as informações em tempo real. A escolha do modelo de IA LLaMA Meta 3.2 Vision 90B permitiu uma análise de imagens detalhada sem comprometer o tempo de resposta do sistema, enquanto a integração com o Azure AI Speech garantiu uma conversão de texto para áudio com entonação natural e clara.



**Figura 1. Visão Geral da Arquitetura de Sistema Áurea.**

O uso de sensores ultrassônicos para a detecção de obstáculos complementa a experiência assistiva ao fornecer alertas táteis progressivos, permitindo que o usuário perceba a proximidade de objetos sem depender exclusivamente do áudio. Essa abordagem foi adotada para garantir maior flexibilidade na adaptação do sistema a diferentes ambientes e condições de uso. Em ambientes com alto ruído ou situações onde a audição do usuário está comprometida, o feedback tátil se torna um recurso essencial para a navegação segura.

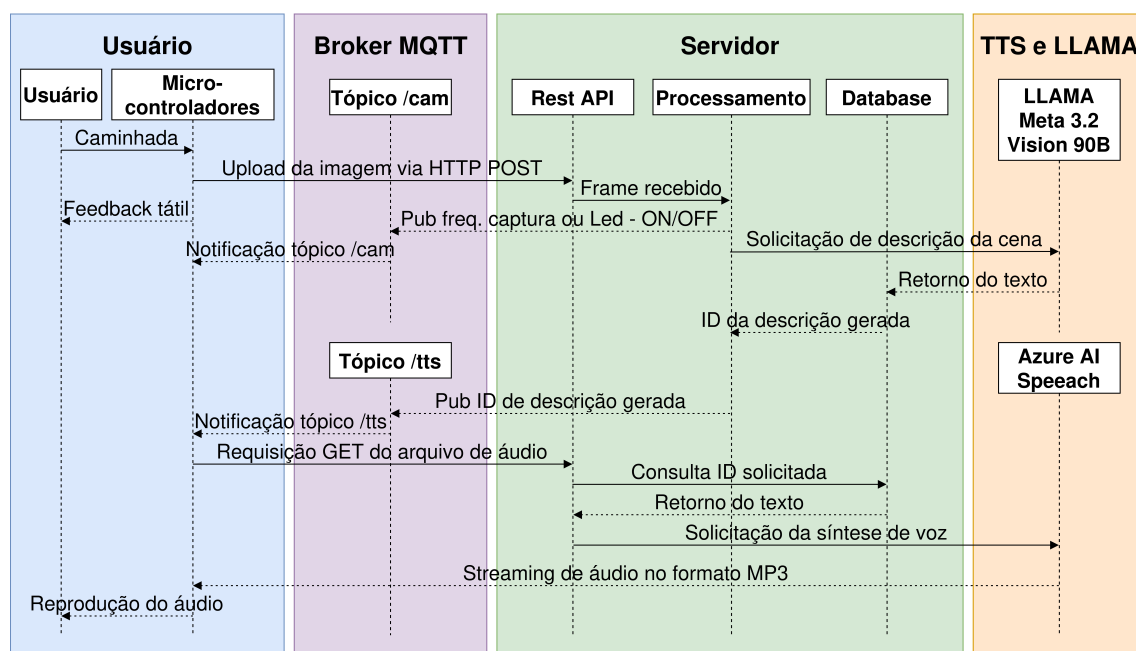
Com essa estrutura, o Áurea se diferencia de outras soluções existentes ao combinar IA para descrição de ambientes e um sistema de detecção de obstáculos acessível e de baixo custo. A arquitetura proposta permite futuras melhorias, como a integração de novos sensores, o aprimoramento dos modelos de IA para melhor interpretação de cenas complexas e a implementação de uma interface para ajustes personalizados, tornando o dispositivo ainda mais adaptável às necessidades dos usuários.

No usuário, o sistema opera com dois microcontroladores ESP32: o primeiro microcontrolador gerencia a captura de imagens e o controle do flash LED (*Light Emitting Diode*), permitindo ajustes dinâmicos na frequência de captura conforme a necessidade do ambiente. As imagens são enviadas ao servidor para processamento e posterior conversão em audiodescrição; o segundo microcontrolador é responsável pela reprodução de áudio e pelo mecanismo de detecção de obstáculos, utilizando um sensor ultrassônico para medir distâncias e um módulo *vibracall* para fornecer feedback tátil progressivo ao usuário.

A comunicação entre os componentes é intermediada por um *broker* MQTT, que facilita a troca de mensagens entre os microcontroladores e o servidor. O tópico `/cam` gerencia a captura de imagens, enquanto o tópico `/tts` é utilizado para a transmissão das descrições geradas. Esse modelo de comunicação assíncrona reduz a latência do sistema, permitindo uma resposta mais rápida e eficiente.

No servidor, a infraestrutura é composta por uma ponte de comunicação que recebe as imagens enviadas pelo primeiro microcontrolador, armazenando-as em um banco de dados para posterior análise. O processamento é realizado por um modelo de visão computacional baseado no LLaMA Meta 3.2 Vision 90B, que interpreta a cena e gera uma descrição textual detalhada. Após essa etapa, o texto é encaminhado para o serviço Azure AI Speech, responsável por converter a descrição em áudio no formato MP3. O servidor, então, publica a ID do arquivo gerado no tópico MQTT correspondente, permitindo que o segundo microcontrolador realize a requisição do áudio e o reproduza para o usuário.

O fluxo completo de funcionamento da arquitetura de sistema *Áurea* é ilustrado no diagrama de sequência apresentado na Figura 2. O processo se inicia com a captura da imagem, que é enviada por mensagem ao servidor. O processamento ocorre de forma contínua, desde a interpretação da cena até a conversão do texto em áudio. Assim que a descrição sonora está disponível, uma notificação é publicada no tópico `/tts`, permitindo que o microcontrolador 2 recupere o arquivo e inicie a reprodução do áudio.



**Figura 2. Diagrama de sequência do sistema *Áurea*.**

Paralelamente, a detecção de obstáculos funciona de maneira independente, garantindo alertas táteis sempre que um objeto for identificado à frente do usuário. Para isso, o microcontrolador 2 ativa o feedback tátil por meio de um módulo *vibracall*. O sensor de ultrassom mede continuamente a distância até os objetos à frente do usuário e ajusta a intensidade da vibração conforme a proximidade do obstáculo. Esse mecanismo garante uma resposta rápida e intuitiva para a navegação segura do usuário.

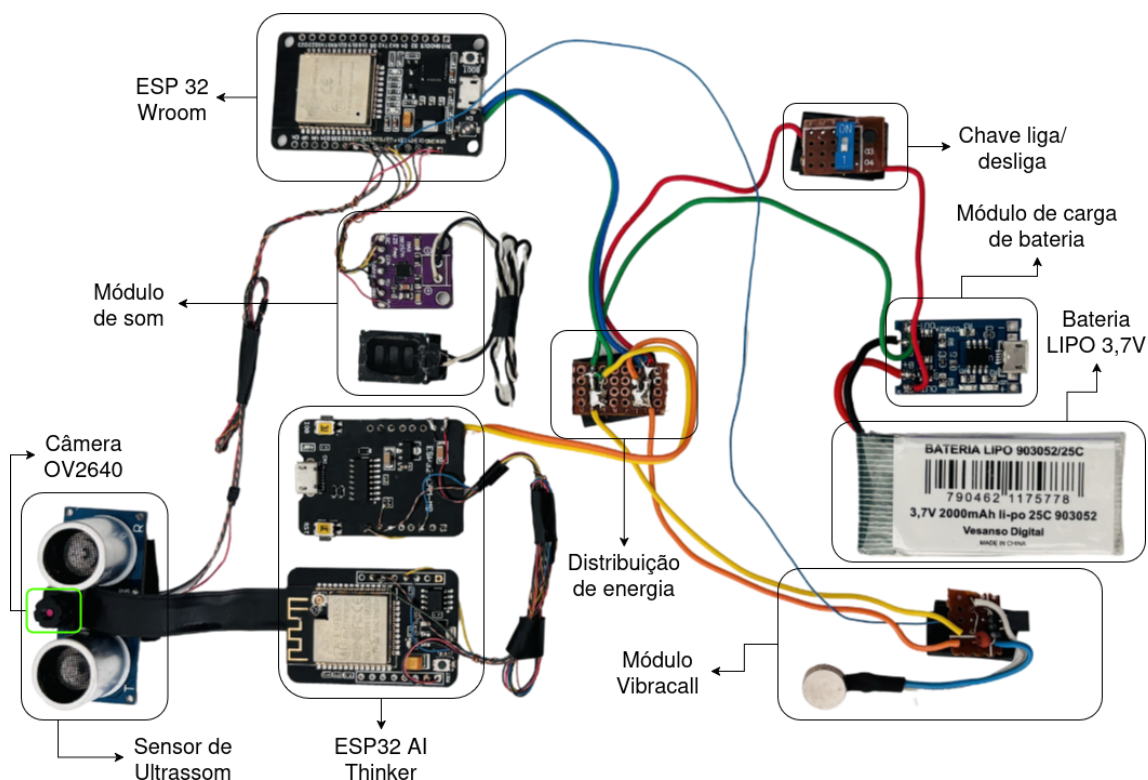
## 4. Prova de Conceito

A prova de conceito do *Áurea* foi desenvolvida com o objetivo de validar a integração dos componentes de hardware e software, assegurando a funcionalidade do sistema de audiodescrição e detecção de obstáculos em tempo real. O projeto foi estruturado para garantir que a solução fosse de baixo custo, eficiente e acessível, mantendo um equilíbrio entre processamento local e remoto.

Nesse contexto, a seleção das peças e materiais do protótipo foi guiada por três critérios principais: custo reduzido, eficiência no consumo de energia e desempenho adequado para as tarefas de visão computacional e navegação assistida. Ademais, a estrutura do protótipo foi planejada em duas camadas principais, com um primeiro microcontrolador ESP32 AI Thinker dedicado à captura de imagens e um segundo microcontrolador ESP32 Wroom responsável pela reprodução de áudio e pelo feedback tátil.

A escolha do ESP32 AI Thinker se deu pela presença do módulo de câmera OV2640, que permite a captura de imagens com qualidade adequada para processamento remoto. Já o ESP32 Wroom foi selecionado por sua capacidade de gerenciamento eficiente dos sensores ultrassônicos e do módulo *vibracall*, garantindo alertas táteis progressivos conforme a proximidade de obstáculos.

O circuito da prova de conceito, apresentado na Figura 3, foi montado visando a otimização da distribuição dos componentes e a minimização do consumo energético. A alimentação do sistema é feita por uma bateria LiPo 3,7V 2000mAh, gerenciada pelo módulo de carga TP4056. A comunicação entre os microcontroladores e o servidor ocorre via Wi-Fi, permitindo uma troca de informações rápida e eficiente.



**Figura 3. Circuito da prova de conceito com identificação dos componentes.**

#### 4.1. Levantamento de Requisitos com Usuários

Embora não tenha sido conduzido um levantamento formal e metodologicamente estruturado de requisitos com pessoas cegas, o desenvolvimento do dispositivo contou com a colaboração direta de uma pessoa com deficiência visual ao longo do processo. A interação contínua com esse usuário durante as fases de prototipação e testes possibilitou a coleta de feedbacks valiosos, que orientaram melhorias na usabilidade, sugestões de funcionalidades e ajustes na implementação de recursos. Essa contribuição prática foi essencial para alinhar o projeto às necessidades reais do público-alvo. Futuramente, pretende-se realizar um processo de cocriação mais amplo, envolvendo múltiplos usuários com deficiência visual, de forma sistemática e participativa.

#### 4.2. Estimativa de Custo dos Componentes

Componente	Custo aproximado (R\$)
ESP32 AI Thinker	44,72
ESP32 Wroom	39,99
Módulo de som	17,79
Sensor ultrassônico	12,26
Câmera OV2640	22,80
Módulo vibracall	4,21
Bateria + módulo de carga	129,90
Diversos (cabos, suporte)	15,00

**Tabela 1. Estimativa de custo por componente do dispositivo Áurea.**

*Os valores apresentados refletem os custos reais dos componentes utilizados na construção do protótipo, podendo variar conforme a disponibilidade e volatilidade do mercado, sendo que soluções comerciais similares, como o OrCam MyEye, podem ultrapassar R\$14.000,00, o que torna o Áurea uma alternativa acessível e viável para populações de baixa renda.*

O desenvolvimento do protótipo do *Áurea* seguiu uma abordagem iterativa, validando individualmente cada módulo antes da integração final do sistema. Inicialmente, todos os componentes foram montados em uma *protoboard*, permitindo a realização de testes isolados para garantir o funcionamento correto de cada sensor, atuador e microcontrolador. Essa etapa foi essencial para identificar possíveis interferências eletromagnéticas, otimizar conexões e ajustar parâmetros de calibração.

Com os módulos devidamente testados e aprovados, iniciou-se a soldagem dos circuitos para garantir maior estabilidade e confiabilidade no funcionamento do dispositivo. A montagem foi organizada considerando a distribuição de corrente e a minimização de ruídos nos sinais de comunicação.

Para a alimentação dos componentes (placas, sensores, atuadores, etc), foram utilizados fios de cobre 30 AWG (*American Wire Gauge*), que oferecem baixa resistência elétrica e flexibilidade para conexões em espaços reduzidos. Já para a transmissão de sinais entre componentes que não demandam altas correntes, optou-se pelo uso de fios esmaltados de alta resistência mecânica, com isolamento adequado para evitar curtos-circuitos e interferências externas.

A disposição dos fios e a fixação dos componentes foram projetadas para manter a estrutura compacta e ergonômica dos óculos, garantindo que o dispositivo pudesse ser integrado de maneira discreta ao suporte dos óculos. Apesar de ser compacta, a organização interna dos componentes foi planejada para facilitar futuras manutenções e substituições de componentes, permitindo ajustes na arquitetura sem comprometer a integridade do sistema.

A Figura 4 apresenta diferentes ângulos do dispositivo finalizado, evidenciando a estrutura compacta e a disposição dos sensores. O design do protótipo priorizou a ergonomia e o conforto do usuário, assegurando que o dispositivo fosse leve e de fácil adaptação para o uso diário.



**Figura 4. Diferentes ângulos de visão do dispositivo *Áurea* e disposição dos sensores.**

A fase final da montagem incluiu testes de integração, onde o dispositivo foi submetido a diferentes cenários de uso, validando sua robustez e desempenho em situações reais. Durante esses testes, foram identificadas melhorias na disposição dos sensores e ajustes na calibração do feedback tátil, assegurando um funcionamento confiável antes da etapa de testes com usuários.

A Figura 5 ilustra como o dispositivo *Áurea* se apresenta quando utilizado pelo usuário, devidamente acoplado ao suporte dos óculos. Observa-se que o design compacto permite o encaixe lateral do módulo, garantindo discrição e conforto durante o uso. A visão frontal evidencia a posição dos sensores, enquanto a visão lateral destaca a organização interna dos componentes e o aproveitamento do espaço disponível.



**Figura 5. Imagens demonstrando como o dispositivo *Áurea* fica posicionado no rosto do usuário, acoplado ao suporte dos óculos.**

### 4.3. Funcionamento

O fluxo operacional segue a arquitetura previamente apresentada na Seção 3. O ESP32 AI Thinker captura imagens do ambiente em intervalos ajustáveis entre 150ms e 15.000ms, enviando-as via passagem de mensagem para o servidor. No servidor, as imagens são processadas pelo modelo de visão computacional LLaMA Meta 3.2 Vision 90B, que interpreta a cena e gera descrições textuais detalhadas.

As descrições geradas são posteriormente convertidas em áudio pelo serviço Azure AI Speech, garantindo uma entonação natural e clara. Assim que o áudio está disponível, o servidor publica a ID do arquivo gerado no *broker* MQTT, permitindo que o ESP32 Wroom faça a requisição e reproduza a descrição sonora ao usuário.

Simultaneamente, a detecção de obstáculos opera de maneira independente. O sensor ultrassônico mede continuamente a distância dos objetos à frente do usuário e ajusta a intensidade do feedback tátil conforme a proximidade. O *vibracall*, posicionado estrategicamente para garantir uma resposta sensorial eficiente, proporciona vibrações de intensidade progressiva à medida que o usuário se aproxima de um obstáculo. Essa funcionalidade complementa a audiodescrição, oferecendo um sistema híbrido de percepção espacial.

A comunicação entre os componentes é gerenciada pelo protocolo MQTT, que possibilita uma troca de mensagens ágil e eficiente. O tópico `/cam` controla o envio de imagens e a ativação do flash LED, enquanto o tópico `/tts` gerencia a reprodução do áudio gerado. O uso desse modelo de comunicação assíncrona minimiza a latência e evita congestionamento na rede, garantindo um funcionamento contínuo e responsivo.

### 4.4. Formulação e Otimização do Prompt para Geração de Audiodescrição

A audiodescrição é um dos elementos centrais do *Áurea*, permitindo que usuários cegos ou com baixa visão recebam informações relevantes sobre o ambiente ao seu redor. Para orientar o modelo LLaMA Meta 3.2 Vision 90B na geração dessas descrições, foi desenvolvido um *prompt* específico, visando garantir clareza, objetividade e eficiência nas respostas.

O *prompt* utilizado é o seguinte:

Você é responsável por descrever a cena para um usuário com deficiência visual. Forneça uma descrição clara, objetiva e detalhada, sem informações falsas ou suposições. Se houver objetos destacados, concentre-se neles. Utilize texto simples, evite emojis e caracteres especiais. Não mencione "imagem" ou sua marcação. Responda em português com até 150 caracteres.

A limitação da resposta a 150 caracteres foi definida para otimizar o consumo de tokens e garantir descrições rápidas, com geração média entre 300ms e 500ms. Isso equilibra concisão e detalhamento, essencial para o uso em tempo real.

Durante os testes, observou-se que modelos de IA tendem a produzir frases genéricas e redundantes. Para evitar esse comportamento, o *prompt* foi refinado, removendo instruções ambíguas e reforçando o uso da língua portuguesa, o que ajudou a melhorar a naturalidade e precisão das respostas.

A formulação do *prompt* também considerou a necessidade de versatilidade, já que o sistema pode ser utilizado em diversos cenários. O modelo foi treinado para reconhecer diferentes tipos de objetos, como placas, móveis, pessoas e obstáculos.

Essa estratégia resultou em uma experiência de audiodescrição mais eficiente e adaptada às necessidades reais dos usuários, reforçando o potencial do *Áurea* como uma solução assistiva acessível.

## 5. Resultados

A avaliação do protótipo do *Áurea* teve como objetivo validar sua eficácia na identificação de elementos do ambiente, geração de descrições sonoras e fornecimento de feedback tátil em tempo real. Para isso, foram realizados testes em diferentes cenários, abrangendo tanto ambientes internos quanto externos, sob diversas condições de iluminação, conectividade e disposição de obstáculos. Um dos testes contou com a participação de uma pessoa cega, que utilizou o dispositivo dentro de casa e também ao caminhar em áreas externas, como ruas e calçadas. Durante os testes, o usuário fixou o dispositivo no rosto e forneceu feedbacks detalhados sobre sua experiência de uso.

Os resultados demonstraram que o processamento remoto das imagens, utilizando o modelo de visão computacional LLaMA Meta 3.2 Vision 90B, proporcionou alta fidelidade na identificação de objetos. O sistema reconheceu com precisão elementos como pessoas, veículos, placas e mobiliário urbano, gerando descrições detalhadas que foram convertidas em áudio com baixa latência. Esse processamento remoto permitiu reduzir significativamente a carga computacional no microcontrolador embarcado, garantindo um desempenho eficiente mesmo em hardware de baixo custo.

A detecção de obstáculos foi implementada por meio de um sensor ultrassônico, que alertava o usuário com vibrações sempre que um objeto fosse identificado dentro do raio de alcance do sensor. Os testes indicaram um bom desempenho na maioria dos cenários, mas evidenciaram limitações na identificação de superfícies inclinadas, transparentes ou muito finas, devido à forma como o ultrassom reflete nesses materiais. O usuário cego relatou dificuldades ao caminhar na rua, especialmente na detecção de certos tipos de obstáculos, como degraus e postes de sinalização muito finos. Essas observações destacam a necessidade de melhorias na escolha e disposição dos sensores para tornar a solução mais confiável e precisa.

A autonomia energética do dispositivo foi avaliada por meio de testes de longa duração, nos quais o *Áurea* operou continuamente por aproximadamente 12 horas utilizando uma bateria LiPo de 2000mAh. Esse tempo de operação é considerado satisfatório, pois está alinhado com a média de autonomia de smartphones modernos, permitindo que o dispositivo seja utilizado ao longo de um dia inteiro sem necessidade de recarga constante.

Além da análise técnica, um grupo de usuários foi convidado a testar o dispositivo em um ambiente controlado. Os participantes relataram que o sistema de audiodescrição forneceu informações relevantes e de fácil compreensão, melhorando significativamente a percepção do ambiente. O feedback tátil foi considerado especialmente útil em locais com alto nível de ruído, onde o áudio poderia ser comprometido. O usuário cego que participou dos testes destacou que a vibração foi essencial para sua locomoção em locais

movimentados, mas sugeriu que o dispositivo pudesse identificar também obstáculos situados fora do campo de visão direta de uma pessoa em pé, com a cabeça erguida. Esse ponto reforça a necessidade de aprimoramentos no alcance da detecção para proporcionar uma experiência ainda mais segura e eficiente.

Os resultados obtidos demonstram que o *Áurea* é uma solução promissora para aumentar a acessibilidade de pessoas com deficiência visual. O sistema apresentou um desempenho satisfatório na identificação e descrição do ambiente, enquanto o feedback tátil adicionou uma camada extra de segurança para a locomoção do usuário. No entanto, desafios como a latência da comunicação, a precisão da detecção de obstáculos e a autonomia da bateria ainda necessitam de aprimoramento para tornar o dispositivo mais confiável e adaptável a diferentes cenários de uso.

### 5.1. Viabilidade e Desafios

A prova de conceito validou a viabilidade técnica da solução. O processamento remoto das imagens permitiu uma análise detalhada da cena sem sobrecarregar os microcontroladores embarcados, enquanto a conversão de texto para áudio ocorreu com baixa latência, garantindo uma experiência auditiva fluida. O sistema de feedback tátil complementou a audiodescrição, proporcionando uma camada adicional de segurança na locomoção do usuário.

O desempenho computacional do *Áurea* mostrou-se satisfatório, uma vez que as operações mais exigentes ocorrem no servidor remoto. No entanto, há espaço para aprimoramentos do ponto de vista de software, como a adoção de estratégias mais eficientes de manipulação de *threads*, separação de responsabilidades entre tarefas e melhorias na arquitetura geral do sistema.

Durante os testes, observou-se que o dispositivo aquece até uma temperatura média compatível com a operação normal de microcontroladores. Embora esse aquecimento seja perceptível ao toque, ele não foi identificado como desconfortável, pois o case que abriga a parte eletrônica é fabricado em material plástico, dificultando a condução térmica direta para a pele. Ainda assim, estratégias para otimização da dissipação de calor são importantes para preservar a vida útil da bateria e dos componentes internos.

Os testes também revelaram desafios relacionados à dependência de uma conexão Wi-Fi estável para o envio e processamento das imagens. Ambientes com sinal fraco ou intermitente comprometem a operação em tempo real. Além disso, a detecção de obstáculos por ultrassom apresentou limitações com certos tipos de superfícies, como vidros, objetos finos e pisos inclinados. O usuário cego participante relatou dificuldade na detecção de pequenos obstáculos, destacando a necessidade de revisão na escolha e posicionamento dos sensores.

A modularidade do sistema foi um ponto positivo, permitindo futuras expansões como a integração de novos sensores e o aprimoramento dos modelos de IA. Com essas melhorias, o *Áurea* poderá oferecer uma experiência assistiva ainda mais robusta e adaptável às necessidades dos usuários.

### 5.2. Melhorias Futuras

O *Áurea* pode ser aprimorado com sensores de profundidade mais precisos, como ToF ou LiDAR, substituindo o ultrassônico atual e ampliando a detecção de obstáculos, inclusive



fora do campo de visão. Também é recomendada a adoção de um modelo híbrido de processamento, com parte da análise feita localmente, reduzindo a dependência de Wi-Fi e melhorando a latência.

Uma funcionalidade futura relevante é a personalização do grau de detalhamento da audiodescrição, permitindo ao usuário escolher entre descrições mais concisas em ambientes familiares e mais completas em locais desconhecidos.

Outras melhorias incluem ajustes configuráveis na intensidade da vibração, uso de sensores inerciais para otimizar o consumo energético, e integração com assistentes virtuais e Bluetooth, facilitando o controle por dispositivos móveis. Tais avanços reforçam o potencial do *Áurea* como uma solução assistiva versátil e acessível.

## 6. Considerações Finais

Este trabalho apresentou o desenvolvimento e a validação do *Áurea*, um dispositivo assistivo de baixo custo voltado à mobilidade e percepção espacial de pessoas com deficiência visual. A solução combina sensores ultrassônicos para detecção de obstáculos e um modelo de IA para geração de audiodescrição, integrando feedback tátil e sonoro de forma complementar.

Os testes demonstraram que o sistema fornece descrições do ambiente de forma detalhada, com boa fluidez e tempo de resposta adequado. A avaliação com usuários, incluindo uma pessoa cega em contextos reais de uso, indicou que o dispositivo melhora a percepção do ambiente e contribui para a segurança na locomoção. A autonomia de 12 horas foi considerada satisfatória, e o feedback tátil mostrou-se especialmente útil em ambientes com grande concentração de obstáculos posicionados na altura do rosto, como galhos, placas e estruturas suspensas.

Foram identificados pontos de melhoria, como a substituição do sensor ultrassônico por sensores de profundidade mais precisos, a adoção de processamento híbrido (local e remoto) para reduzir a latência, e a personalização da intensidade da vibração. Tais ajustes visam aprimorar a usabilidade e a adaptabilidade do sistema.

Com arquitetura modular, o *Áurea* permite futuras expansões, como integração com comandos por voz, ajustes via aplicativo móvel e aprendizado contínuo. Dessa forma, o dispositivo se posiciona como uma alternativa acessível e promissora para promover autonomia e inclusão de pessoas com deficiência visual.

## Referências

- Bai, e. a. (2017). Wearable assistive device using depth camera and ultrasonic sensors for obstacle detection. *IEEE Transactions on Human-Machine Systems*, 47(6):1021–1032.
- Brasil (2015). *Lei Brasileira de Inclusão da Pessoa com Deficiência (Estatuto da Pessoa com Deficiência)*.
- Cortez, L. and Almeida, R. (2022). Low-cost smart glasses for blind individuals using raspberry pi 2. In *Proceedings of the International Conference on Emerging Technologies*, pages 189–196.

- Diniz, B. C. and Junior, M. M. (2024). Pesquisa de tecnologias para avaliar a viabilidade e métodos para a criação de um protótipo de Óculos para interação de deficientes visuais com o mundo. *Revista de Engenharia e Tecnologia*, 17(1).
- e Inclusão, F. A. (2016). A importância da audiodescrição na inclusão das pessoas com deficiência visual. *Revista de Acessibilidade e Inclusão*, 8(3):112–130. Acesso em: 14 mar. 2025.
- Franco, E. and Silva, F. (2010). A audiodescrição e o uso dos seus princípios nas descrições informais. *Revista Brasileira de Tradução e Interpretação*, 5(2):45–60. Acesso em: 14 mar. 2025.
- Hakobyan, L., Lumsden, J., O’Sullivan, D., and Bartlett, H. (2013). Mobile assistive technologies for the visually impaired. *Survey of ophthalmology*, 58(6):513–528.
- IBGE (2019). *Pessoas com Deficiência e as Desigualdades Sociais no Brasil*.
- Khan, A. and Khusro, S. (2021). An insight into smartphone-based assistive solutions for visually impaired and blind people: issues, challenges and opportunities. *Universal Access in the Information Society*, 20(2):265–298.
- Kuriakose, B., Shrestha, R., and Sandnes, F. E. (2022). Tools and technologies for blind and visually impaired navigation support: a review. *IETE Technical Review*, 39(1):3–18.
- Lavric, A., Beguni, C., Zadobrischi, E., Căilean, A.-M., and Avătămăniței, S.-A. (2024). A comprehensive survey on emerging assistive technologies for visually impaired persons: lighting the path with visible light communications and artificial intelligence innovations. *Sensors*, 24(15):4834.
- Li, e. a. (2017). A wearable obstacle stereo feedback system for indoor navigation of visually impaired individuals. *Sensors*, 17(8):1834.
- Muhsin, Z. J., Qahwaji, R., Ghanchi, F., and Al-Tae, M. (2024). Review of substitutive assistive tools and technologies for people with visual impairments: Recent advancements and prospects. *Journal on Multimodal User Interfaces*, 18:135–156.
- MWPT (2024). Ia avança nas descrições de imagens, mas ainda não substitui a revisão humana. *Revista de Inteligência Artificial e Sociedade*, 10(4):58–72. Acesso em: 14 mar. 2025.
- SHOWCASE (2023). Ia reforça a qualidade da acessibilidade audiovisual. *Revista de Tecnologia e Inclusão Digital*, 12(1):35–50. Acesso em: 14 mar. 2025.
- UNB (2025). Unb lança ferramenta de audiodescrição para equações e gráficos baseados em ia. *Jornal de Educação Inclusiva*, 15(2):78–95. Acesso em: 14 mar. 2025.
- Wang, e. a. (2023). Applications of artificial intelligence for assisting visually impaired individuals: A comprehensive review. *Journal of Assistive Technologies*, 15:50–68.
- Yang, e. a. (2024). A navigational framework based on deep neural networks and depth sensory segmentation for assisting visually impaired people. *Journal on Multimodal User Interfaces*, 18:135–156.