NeuroView: Óculos Sensoriais Inteligentes para Pessoas Cegas

Gabriel Costa De Jesus[1], Gabriel Lima Fratoni[1], Kailane Souza Araujo[1], Ryan Cerqueira Brandão[1]

[1]Curso de Desenvolvimento de Sistemas
SENAI - Serviço Nacional de Aprendizagem Industrial, Feira de Santana - BA - Brasil
Orientador: Prof. Ivanildo Gomes da Silva - invanildo99gomes@gmail.com

gabrielcosta.js05@gmail.com, gabriellimafratonifff@gmail.com,
kailanesouza18014@gmail.com, ryancerqueiradev@gmail.com

Abstract. This paper presents NeuroView, a low-cost wearable system designed to support blind people in mobility and environmental awareness. The current prototype combines an ESP32-CAM, an Android application, and Firebase data storage. The device captures images and sends them to the smartphone via Wi-Fi as an MJPEG stream; each frame is processed locally with the open-vocabulary object detection model YOLOE-26, running through TensorFlow Lite, and the result is returned to the user as audio feedback. The paper describes the development path, the current architecture and preliminary tests, which achieved 86% detection accuracy, an average response time of approximately 3 seconds, and a 95% Wi-Fi connection success rate.

Resumo. Este artigo apresenta o NeuroView, um sistema vestível de baixo custo desenvolvido para apoiar pessoas cegas na mobilidade e na percepção do ambiente. O protótipo atual integra ESP32-CAM, aplicativo Android e armazenamento de dados em Firebase. O dispositivo captura imagens e as envia ao smartphone por Wi-Fi em um stream MJPEG; cada quadro é processado localmente com o modelo de detecção de objetos de vocabulário aberto YOLOE-26, executado via TensorFlow Lite, e o resultado é devolvido ao usuário por áudio. O trabalho descreve a evolução do projeto, a arquitetura atual e os testes preliminares realizados, que alcançaram 86% de acurácia na detecção, tempo médio de resposta de aproximadamente 3 segundos e taxa de sucesso de conexão Wi-Fi de 95%.

Palavras-chave: Tecnologia assistiva. Internet das Coisas. Visão computacional. Aplicativo mobile. Deficiência visual.

1. Introdução

A deficiência visual ainda representa um desafio importante para a inclusão, especialmente na mobilidade urbana e no acesso à informação sobre o ambiente. Segundo a Organização Mundial da Saúde, pelo menos 2,2 bilhões de pessoas no mundo possuem algum tipo de deficiência visual (WORLD HEALTH ORGANIZATION, 2019). No Brasil, os resultados preliminares do Censo Demográfico 2022 apontam 14,4 milhões de pessoas com algum tipo de deficiência (7,3% da população de 2 anos ou mais); desse total, 7,9 milhões relataram dificuldade de enxergar, a limitação funcional mais frequente entre as investigadas (IBGE, 2025). Essa

realidade impacta diretamente a autonomia, a segurança e a qualidade de vida dessas pessoas no dia a dia.

Pessoas cegas, em particular, dependem fortemente de recursos como a bengala para se locomover com segurança. Esses recursos, no entanto, não resolvem todas as situações: a bengala não identifica obstáculos acima do nível do solo, nem informa que tipo de objeto está à frente do usuário. Essa limitação se torna ainda mais crítica em ambientes urbanos dinâmicos, onde a falta de informação sobre o entorno pode gerar insegurança, acidentes e restrição da mobilidade e da independência dessas pessoas.

Reconhecendo esse cenário, o presente artigo introduz a concepção do NeuroView, um sistema vestível de baixo custo voltado à mobilidade e à percepção do ambiente por pessoas cegas. O sistema utiliza uma câmera acoplada aos óculos e um modelo de detecção de objetos para transformar imagens do ambiente em informações por áudio, permitindo que o usuário receba uma descrição simples do que está próximo sem depender da tela do celular. O projeto envolve ainda aplicativo móvel e banco de dados em nuvem, aproximando a solução do conceito de Internet das Coisas, ao conectar hardware, comunicação sem fio e interface de usuário em um mesmo sistema (ATZORI; IERA; MORABITO, 2010), buscando criar uma alternativa acessível, de baixo custo e centrada no usuário.

2. Fundamentação Teórica

A tecnologia assistiva busca ampliar a autonomia de pessoas com deficiência por meio de recursos, serviços e dispositivos que reduzam barreiras no cotidiano. Soluções vestíveis para pessoas cegas têm recebido atenção por acompanharem o usuário durante o deslocamento e fornecerem respostas sem ocupar as mãos. Bai et al. (2017) apresentaram óculos para navegação em ambientes internos usando sensores e retorno tátil, e Bai et al. (2019) propuseram um dispositivo vestível com câmera e aprendizado profundo para navegação de pessoas com deficiência visual.

Trabalhos mais recentes, publicados em eventos revisados por pares da Sociedade Brasileira de Computação, reforçam esse campo. Santos Silva e Garrocho (2025) apresentaram o projeto Áurea, óculos de baixo custo que combina inteligência artificial e sensor ultrassônico para descrever o ambiente e alertar sobre obstáculos por vibração. Lima Neto, Cunha e Carvalho (2020), em revisão sistemática apresentada na ERBASE, observaram que sensores de proximidade e placas de prototipagem de baixo custo são recorrentes em soluções embarcadas para locomoção de pessoas cegas, enquanto a visão computacional embarcada ainda é pouco explorada, lacuna que o NeuroView busca ocupar. Fora do meio acadêmico, óculos inteligentes comerciais como os da linha Ray-Ban Meta já vêm sendo adotados por iniciativas municipais brasileiras no atendimento a pessoas cegas, a exemplo do projeto pioneiro do município de Cocal do Sul (SC) (ND MAIS, 2026).

A visão computacional é parte central do projeto. O modelo YOLO, proposto por Redmon et al. (2016), ficou conhecido por realizar detecção em tempo real, tratando identificação e localização de objetos como um único problema. O NeuroView utiliza o YOLOE-26, variante de vocabulário aberto da família YOLO26, que dispensa a etapa de supressão não máxima (NMS) na inferência e reduz a latência em dispositivos com

poder de processamento limitado (ULTRALYTICS, 2026), reconhecendo categorias de objetos sem depender de uma lista fixa definida em treinamento, característica útil diante da variedade de objetos que o usuário pode encontrar durante o deslocamento.

A integração entre componentes também é central: o ESP32-CAM reúne câmera e comunicação sem fio em módulo compacto, adequado para protótipos de IoT (ESPRESSIF SYSTEMS, 2025).

3. Metodologia

O desenvolvimento do NeuroView aconteceu de forma incremental, em três versões: a primeira usava sensor ultrassônico e vibração para alertar sobre obstáculos; a segunda testou a detecção de objetos com YOLO pela câmera do próprio celular; a versão atual integra óculos com ESP32-CAM, aplicativo Android e banco de dados em nuvem. No hardware, o ESP32-CAM captura as imagens do ambiente, com bateria LiPo de 1800 mAh, módulo TP4056, botão, LEDs indicadores e case de proteção, como mostra a Figura 1. A Figura 2 detalha o pipeline completo de operação.



Figura 1. Protótipo físico do NeuroView. Fonte: autoria própria.

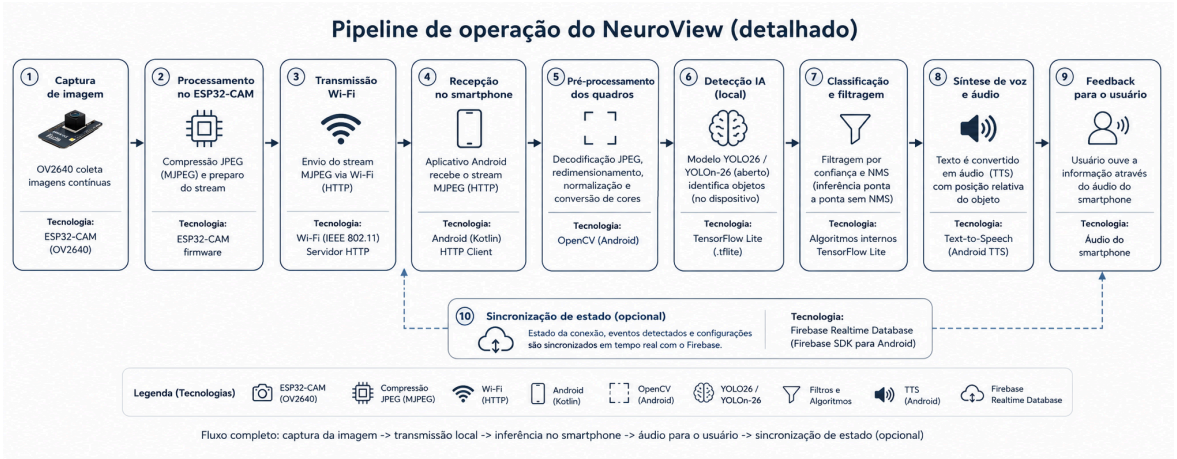


Figura 2. Pipeline de funcionamento do NeuroView: captura, transmissão, processamento por IA e feedback auditivo. Fonte: autoria própria.

3.1. Conectividade e detecção de objetos

Na versão atual, o ESP32-CAM cria sua própria rede Wi-Fi e atua como ponto de acesso, à qual o celular se conecta diretamente. Nesse modo, o usuário já acessa os dados de estado do ESP pelo aplicativo, mas o celular fica sem internet enquanto conectado à rede do dispositivo, e a sincronização com a nuvem só ocorre quando o celular volta a se conectar à rede externa.

O ESP32-CAM transmite as imagens ao aplicativo por um stream MJPEG contínuo via HTTP. A escolha do MJPEG deve-se à simplicidade de implementação no ESP32-CAM e à compatibilidade com a comunicação HTTP utilizada pelo protótipo. O reconhecimento é feito localmente no celular com o modelo YOLOE-26 (variante de vocabulário aberto do YOLO26), exportado em .tflite e executado pela API Interpreter do TensorFlow Lite. Essa escolha garante compatibilidade ampla com Android, baixo consumo de recursos e inferência sem depender de serviços externos. O processamento busca equilibrar taxa de quadros, uso de CPU e latência; nos testes, o tempo de resposta foi de aproximadamente 3 segundos (Seção 4.1).

3.2. Privacidade e interação

Todo o processamento das imagens ocorre localmente no smartphone: nenhuma imagem é enviada à nuvem, apenas dados de estado do dispositivo, conexão, bateria e falhas, sincronizados quando há internet disponível. Essa escolha reduz a exposição de dados sensíveis captados no ambiente do usuário; seus desdobramentos em usabilidade são discutidos na Seção 4.2.

No aplicativo, o usuário cadastra os óculos, abre a transmissão e acompanha a conexão (Figura 3).

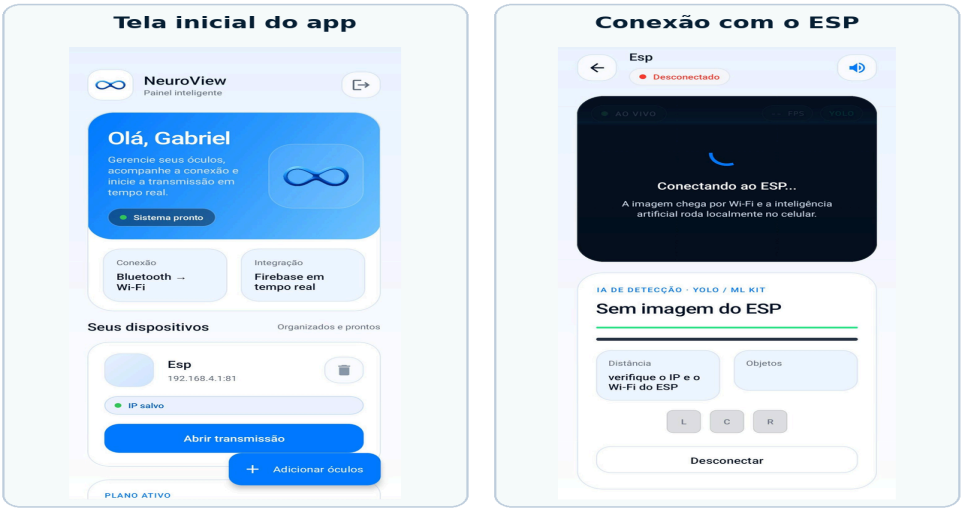


Figura 3. Telas da versão atual do aplicativo NeuroView: painel inicial e conexão com o ESP. Fonte: autoria própria.

Após a detecção de um objeto pelo YOLOE-26, o sistema gera uma mensagem curta por Text-to-Speech (TTS) do Android, informando o tipo de objeto e, quando disponível, sua posição aproximada, por exemplo, *"Pessoa à frente"* ou *"Cadeira à direita"*. Na versão atual, os avisos utilizam a mesma voz sintetizada, sem variação de tom por classe ou distância. Um mecanismo de controle de repetição evita que o mesmo

objeto seja anunciado continuamente em quadros consecutivos; a interação do usuário ocorre por comandos simples do aplicativo, sem uso de comandos de voz nesta versão.

4. Resultados

Como resultado do desenvolvimento incremental descrito na Seção 3, a equipe chegou a uma versão demonstrativa funcional, que integra os óculos e o aplicativo, com o YOLOE-26 rodando localmente no smartphone e a conectividade Wi-Fi local operacional. Essa versão foi submetida aos testes preliminares descritos a seguir.

4.1. Testes e resultados

Nos testes realizados, o sistema reconheceu corretamente diferentes categorias de objetos do ambiente (pessoas, cadeiras, plantas, computadores, entre outros). Para quantificar esse desempenho, foram avaliadas 50 imagens, das quais 43 foram identificadas corretamente, correspondendo a uma taxa de acerto de 86%. O tempo entre a captura do quadro e o início do feedback de áudio foi de aproximadamente 3 segundos. Também foram realizadas 20 tentativas de conexão entre o smartphone e o ESP32-CAM, das quais 19 foram bem-sucedidas, correspondendo a uma taxa de sucesso de 95%. Quanto à autonomia, o teste foi iniciado com a bateria completamente carregada e mantido até o desligamento do dispositivo. Nessas condições, o protótipo permaneceu em funcionamento por aproximadamente 6 horas e 40 minutos.

4.2. Privacidade e usabilidade

Como todo o processamento ocorre localmente no smartphone (Seção 3.2), o NeuroView reduz riscos de privacidade frente a soluções que enviam imagens a servidores externos. Por outro lado, a dependência de um smartphone e de uma rede Wi-Fi criada pelo próprio ESP traz desafios já identificados (Seção 3.1), somados à autonomia dos óculos, medida em aproximadamente 6h40min em teste com a bateria completamente carregada (Seção 4.1); na prática, o uso contínuo também está condicionado à bateria do smartphone, que processa o modelo localmente. Além disso, a configuração da conexão exige digitar credenciais de rede, o que pode ser difícil para usuários com pouca familiaridade técnica, reforçando a importância de aprimoramentos futuros voltados à usabilidade e ao design centrado no usuário.

5. Conclusão

O NeuroView aborda os desafios de mobilidade e percepção do ambiente enfrentados por pessoas cegas, com foco estratégico na acessibilidade. O processamento local das imagens no smartphone, sem depender de servidores externos, concilia inteligência artificial e privacidade do usuário. Ao complementar recursos tradicionais como a bengala com uma camada de percepção auditiva do ambiente, o sistema pode oferecer mais segurança, independência e qualidade de vida a pessoas cegas. Integrando hardware acessível, processamento local e retorno por áudio em um único sistema, o NeuroView tem potencial para mudar a forma como pessoas cegas percebem e se deslocam pelo ambiente, contribuindo para tecnologias assistivas mais inclusivas.

Referências

ATZORI, Luigi; IERA, Antonio; MORABITO, Giacomo. The Internet of Things: a survey. Computer Networks, v. 54, n. 15, p. 2787-2805, 2010. DOI: 10.1016/j.comnet.2010.05.010.

BAI, Jinqiang; LIAN, Shiguo; LIU, Zhaoxiang; WANG, Kai; LIU, Dijun. Smart guiding glasses for visually impaired people in indoor environment. arXiv preprint, arXiv:1709.09359, 2017. Disponível em: https://arxiv.org/abs/1709.09359. Acesso em: 29 jun. 2026.

BAI, Jinqiang; LIU, Zhaoxiang; LIN, Yimin; LI, Ye; LIAN, Shiguo; LIU, Dijun. Wearable travel aid for environment perception and navigation of visually impaired people. arXiv preprint, arXiv:1904.13037, 2019. Disponível em: https://arxiv.org/abs/1904.13037. Acesso em: 29 jun. 2026.

ESPRESSIF SYSTEMS. ESP32 Series Datasheet. Version 5.2. Shanghai: Espressif Systems, 2025. Disponível em: https://www.espressif.com/sites/default/files/documentation/esp32_datasheet_en.pdf. Acesso em: 29 jun. 2026.

IBGE. Censo Demográfico 2022: pessoas com deficiência e pessoas diagnosticadas com transtorno do espectro autista: resultados preliminares da amostra. Rio de Janeiro: IBGE, 2025.

LIMA NETO, Ademar; CUNHA, Mônica da; CARVALHO, Lukas. Uma revisão sistemática sobre tecnologias assistivas voltadas para auxiliar a locomoção de deficientes visuais em ambiente externo utilizando soluções embarcadas. In: Escola Regional de Computação Bahia, Alagoas e Sergipe (ERBASE), 20., 2020, Arapiraca/AL. Anais [...]. Porto Alegre: SBC, 2020. p. 89-98.

ND MAIS. 'Liberdade que nunca tivemos': cidade de SC será a primeira do Brasil a oferecer óculos inteligentes da Meta. ND Mais, Criciúma, 2026. Disponível em: https://ndmais.com.br/tecnologia/cidade-sc-e-pioneira-ao-oferecer-oculos-inteligentes/. Acesso em: 14 ago. 2026.

REDMON, Joseph; DIVVALA, Santosh; GIRSHICK, Ross; FARHADI, Ali. You only look once: unified, real-time object detection. In: IEEE Conference on Computer Vision and Pattern Recognition, 2016. Proceedings [...]. Las Vegas: IEEE, 2016. p. 779-788. Disponível em: https://arxiv.org/abs/1506.02640. Acesso em: 29 jun. 2026.

SANTOS SILVA, Vinicius Marques; BATISTA GARROCHO, Charles Tim. Áurea: Óculos de Inteligência Artificial de Baixo Custo para Descrição de Cenas e Assistência de Mobilidade para Indivíduos com Deficiência Visual. In: Workshop de Computação Urbana (CoUrb), 9., 2025, Natal/RN. Anais [...]. Porto Alegre: SBC, 2025. p. 85-98. DOI: 10.5753/courb.2025.8751.

ULTRALYTICS. Ultralytics YOLO26: Unified Real-Time End-to-End Vision Models. 2026. Disponível em: https://arxiv.org/abs/2606.03748.

WORLD HEALTH ORGANIZATION. World report on vision. Geneva: World Health Organization, 2019. Disponível em: https://www.who.int/publications/i/item/9789241516570. Acesso em: 29 jun. 2026.