

RESSALVA

Atendendo solicitação do autor,
o texto completo deste documento será
disponibilizado somente a partir de
29/11/26.



UNIVERSIDADE ESTADUAL PAULISTA

"JÚLIO DE MESQUITA FILHO"

Reginaldo De Jesus Oliveira Silva Junior

**TECNOLOGIA ASSISTIVA:
RECONHECIMENTO DE OBJETOS RESIDENCIAIS E
SINTETIZAÇÃO DE VOZ UTILIZANDO INTELIGÊNCIA
ARTIFICIAL**

Sorocaba

2026

Reginaldo De Jesus Oliveira Silva Junior

**TECNOLOGIA ASSISTIVA:
RECONHECIMENTO DE OBJETOS RESIDENCIAIS E
SINTETIZAÇÃO DE VOZ UTILIZANDO INTELIGÊNCIA
ARTIFICIAL**

Dissertação apresentada como requisito para a obtenção do título de Mestre em Engenharia Elétrica, da Universidade Estadual Paulista “Júlio de Mesquita Filho”.

Área de concentração: Sistemas Eletrônicos.
Orientador: Prof. Dr. Ivando Severino Diniz

Sorocaba

2026

“Dedico esse trabalho a todos que me apoiaram para chegar até aqui”

AGRADECIMENTOS

Agradeço a Deus, acima de tudo, e aos meus familiares, especialmente à minha esposa, Aline, que me apoia e com quem compartilho sonhos há 7 anos. Aos professores Doutor Ivando Severino Diniz e Mestre Danilo Ruy Gomes, bem como aos funcionários, professores e alunos da UNESP – Sorocaba/SP.

*Não há nada mais prático do
que uma boa teoria.*

Kurt Lewin

RESUMO

A inteligência artificial (IA) tem sido empregada para ampliar acessibilidade, especialmente no apoio a pessoas com deficiência visual, por meio de soluções capazes de interpretar o ambiente e fornecer retorno por voz. Inserido nesse contexto, este trabalho apresenta um protótipo embarcado de baixo custo para reconhecimento de objetos em ambientes domésticos, operando *off-line* e em tempo real. A solução é composta por uma Raspberry Pi Zero 2 W com câmera, classificadores em cascata (*Haar*) treinados no OpenCV e um módulo de síntese de voz em pt-BR que anuncia os objetos detectados por meio de um amplificador I²S e alto-falante externo. O pipeline implementado inclui captura, pré-processamento, detecção, confirmação temporal e audiodescrição; as decisões de engenharia priorizam simplicidade, privacidade e baixa latência no dispositivo. A dissertação também apresenta uma revisão do estado da arte em visão computacional e tecnologias assistivas, situando a proposta em relação a soluções existentes e a modelos leves para execução embarcada. Em avaliações controladas e em cenários domésticos reais, o protótipo demonstrou desempenho consistente para classes comuns e latência reduzida; foram observadas limitações em baixa iluminação, objetos de contorno fino e superfícies reflexivas. Como contribuições, o trabalho oferece uma arquitetura replicável e um conjunto inicial de classificadores treinados; como trabalhos futuros, propõem-se a ampliação do conjunto de dados, ajustes de pré-processamento e a investigação de modelos compactos adicionais para robustez em campo.

Palavras-chave: Acessibilidade. Síntese de voz. Tecnologia. Inteligência artificial. Deficientes Visuais.

ABSTRACT

Artificial intelligence (AI) has been used to enhance accessibility, especially in supporting people with visual impairments through solutions capable of interpreting the environment and providing voice feedback. In this context, this work presents a low-cost embedded prototype for object recognition in domestic environments, operating offline and in real time. The proposed solution consists of a Raspberry Pi Zero 2 W with a camera, Haar cascade classifiers trained with OpenCV, and a Brazilian Portuguese text-to-speech module that announces the detected objects through an I²S amplifier and an external speaker. The implemented pipeline includes image capture, preprocessing, detection, temporal confirmation, and audio description. The engineering decisions prioritize simplicity, privacy, and low latency on the device. This dissertation also presents a review of the state of the art in computer vision and assistive technologies, positioning the proposed approach in relation to existing solutions and lightweight models suitable for embedded execution. In controlled evaluations and real domestic scenarios, the prototype demonstrated consistent performance for common object classes and reduced latency; however, limitations were observed under low-light conditions, with thin-contour objects, and reflective surfaces. As contributions, this work provides a replicable architecture and an initial set of trained classifiers. As future work, the expansion of the dataset, preprocessing improvements, and the investigation of additional compact models are proposed to increase robustness in real-world conditions.

Keywords: Accessibility. Speech synthesis. Technology. Artificial intelligence. Visually impaired people.

LISTA DE FIGURAS

Figura 1 - Haar Cascade versus CNNs.....	19
Figura 2 – Janela (matrix) 6 x 6	21
Figura 3 - Criação da matrix.....	22
Figura 4 - Pipeline de Visão computacional	25
Figura 5 - Pipeline Tradicional de Síntese de Fala (TTS)	27
Figura 6 – Fluxo de otimização de modelos de IA para dispositivos embarcados	28
.....	
Figura 7 - Tela do aplicativo Picture This	34
Figura 8 – Tela do aplicativo Rock ID: Identificar Pedras.....	35
Figura 9 - Tela do Google Lens	36
Figura 10 - Tela do Pinterest Lens.....	36
Figura 11 - Tela do aplicativo Vivino	37
Figura 12 - Tela do aplicativo TapTapSee	39
Figura 13 - Tela do aplicativo Seeing AI	40
Figura 14 - Exemplo de detecção de objetos.....	41
Figura 15 - Fluxo do sistema.....	43
Figura 16 - Placa Raspberry Pi Zero 2 W	44
Figura 17 - Raspberry Pi Camera	45
Figura 18 - MAX98357A.....	46
Figura 19 - Exemplo de imagens 24x24 pixels	47
Figura 20 - Raspberry Pi Zero 2W Pinout	48
Figura 21 - Pi Zero 2 W com a saída MAX98357A	48
Figura 22 - Imagens rotacionadas	50
Figura 23 - Exemplo de objeto estreito	53
Figura 24 - Janela 24x24 - objetos estreitos em fundos texturizados	54
Figura 25 - Data augmentation	54
Figura 26 - Objeto treinado com borda forte e fácil (geladeira).....	55
Figura 27 – Reconhecimento do objeto sofá.....	64
Figura 28 - Reconhecimento do objeto geladeira	65
Figura 29 - Reconhecimento do objeto mesa	65

LISTA DE GRÁFICOS

Gráfico 1 - Deficiência visual no mundo 2019.....	11
Gráfico 2 – Total de pessoas com deficiência no Brasil 2019.....	12
Gráfico 3 - Deficiências por tipo em 2019	13
Gráfico 4 - Efeito prático - menos estágios estabilizam mais rápido	56

LISTA DE TABELAS

Tabela 1 - Indicadores globais GSMA.....	23
Tabela 2 - Dados do treinamento.....	52
Tabela 3 - conexões Raspberry com MAX98357A	59
Tabela 4 - Métricas agregadas por cenário.....	67

LISTA DE ABREVIações E ACRônimos

ALSA – Advanced Linux Sound Architecture (Arquitetura Avançada de Som do Linux)

BGR – Blue, Green, Red (ordem de canais usada no OpenCV)

BTL – Bridge-Tied Load

CPU – Unidade Central de Processamento

CNN – Convolutional Neural Network (Rede Neural Convolucional)

CVPR – Conference on Computer Vision and Pattern Recognition (Conferência IEEE de Visão Computacional e Reconhecimento de Padrões)

CSI – Camera Serial Interface

DAC – Digital-to-Analog Converter (Conversor Digital-Analógico)

F1 – F1-score (Medida-F1)

FPS – Frames Per Second (Quadros por Segundo)

GPIO – General Purpose Input/Output (Entrada/Saída de Uso Geral)

GPU – Unidade de Processamento Gráfico

IA – Inteligência Artificial

IBGE – Instituto Brasileiro de Geografia e Estatística

ICIP – International Conference on Image Processing (Conferência Internacional de Processamento de Imagens)

I²S – Inter-IC Sound (Barramento serial de áudio entre CIs)

OMS – Organização Mundial da Saúde

OpenCV – Open Source Computer Vision Library (Biblioteca de Visão Computacional de Código Aberto)

PCM – Pulse Code Modulation (Modulação por Código de Pulsos)

PNS – Pesquisa Nacional de Saúde

pt-BR – Português (Brasil)

RGB – Red, Green, Blue (canais de cor)

RGB888 – Formato de cor 8 bits por canal (24 bits)

SUS – System Usability Scale (Escala de Usabilidade do Sistema)

SVM – Support Vector Machine (Máquina de Vetores de Suporte)

TTS – Text-to-Speech (Conversão de Texto em Fala)

UEQ – User Experience Questionnaire (Questionário de Experiência do Usuário)

SUMÁRIO

1. INTRODUÇÃO	9
1.1 Acessibilidade Tecnológica.....	10
1.2 Justificativa	12
1.3 Objetivos.....	14
1.4 Organização da dissertação	15
2. DESCRIÇÃO TEÓRICA	17
2.1 Inteligência Artificial e Aprendizado de Máquina.....	17
2.2 Visão Computacional	19
2.2.1 OpenCV na prática: operações e detecção com cascatas (Haar)	20
3. REVISÃO DA LITERATURA	23
3.1 Processamento de imagem	24
3.2 Síntese da fala.....	25
3.3 Redes Neurais Artificiais.....	27
3.3.1 Arquiteturas convolucionais e evolução recente	27
3.3.2 Eficiência para dispositivos móveis e embarcados	28
3.3.3 Contraste com detectores por características manuais (Haar)	29
3.3.4 Métricas e avaliação	29
3.3.5 Interpretabilidade e confiabilidade	29
3.3.6 Relação com este trabalho	29
3.4 Classificadores Haar Cascade.....	30
3.5 Reconhecimento de padrões	32
3.5.1 Representações: de atributos manuais ao aprendizado de representações	32
3.5.2 Redução de dimensionalidade e seleção de atributos.....	32
3.5.3 Avaliação: métricas e protocolo experimental.....	33
3.5.4 Robustez no mundo real.....	33
3.6 Reconhecimento de Plantas e Insetos.....	33
3.7 Reconhecimento de Minerais.....	34

3.8	Uso Amplo e Generalista	35
3.9	Aplicações Específicas	36
3.10	Inteligência Artificial em Aplicativos Móveis	37
3.11	Tecnologias de Acessibilidade para Deficientes Visuais	38
3.12	Reconhecimento de Objetos com Visão Computacional	40
3.13	Desafios e Oportunidades	41
3.14	Conclusão da Revisão da Literatura	41
4.	MATERIAIS E MÉTODOS.....	42
4.1	Raspberry Pi Zero 2 W	43
4.2	Raspberry Pi Camera v1.3.....	44
4.3	MAX98357A8.....	45
4.4	Hardware	46
4.4.1	Protótipo	47
4.4.2	Configuração de áudio.....	49
4.5	Software.....	49
4.5.1	Python.....	49
4.5.2	Preparação dos dados.....	50
4.5.3	Geração de amostras e vetores.....	50
4.5.4	Treinamento (cascata Haar)	51
4.5.5	Integração e lógica de detecção	51
4.6	Leitura crítica dos resultados	53
4.6.1	Objetos com contornos finos (caneta, garfo, escova de dentes, chave)	53
4.6.2	Objetos retangulares/planos (porta, televisão, armário, gabinetes)	55
4.6.3	Materiais especulares ou transparentes (colher, copo, panela, liquidificador)	56
4.6.4	Buffers de treinamento (precalc).....	56
4.6.5	Número de estágios (complexidade do classificador).....	57
5.	DESENVOLVIMENTO.....	58
5.1	Concepção.....	58

5.2 Implementação	59
5.3 Aquisição e pré-processamento.....	59
5.4 Detecção por lotes	59
5.5 Confirmação e desenho.....	60
5.6 Fila de fala e síntese em pt-BR.....	60
5.6.1 Pilha de voz	60
5.7 Observações de engenharia.....	60
5.8 Cenários e procedimentos	61
5.9 Métricas e coleta.....	61
5.9.1 Cálculo das Métricas de Desempenho	61
5.10 Resultados qualitativos esperados (à luz do treinamento).....	63
6. RESULTADOS.....	64
6.1 Desempenho Técnico	66
6.1.1 Ambiente controlado.....	66
6.1.2 Ambiente doméstico real	67
6.2 Observações de uso (qualitativas).....	67
6.3 Limitações Identificadas.....	67
7. CONCLUSÃO	69
7.1 Limitações observadas	69
7.2 Considerações finais.....	69
8. BIBLIOGRAFIA	71

1. INTRODUÇÃO

A tecnologia tem desempenhado um papel fundamental na mudança de muitos aspectos da vida cotidiana, criando soluções que aumentam a eficiência, a acessibilidade e a participação. Essas inovações incluem o uso de inteligência artificial (IA), que permite o desenvolvimento inteligente e coordenado de aplicações. Segundo Russell e Norvig (2020), a inteligência artificial inclui diversas tecnologias e ferramentas que permitem aos sistemas processar informações e tomar decisões de forma autônoma, substituindo aspectos como visão computacional, reconhecimento de padrões e aprendizado de máquina.

No Brasil, o uso da inteligência artificial em soluções acessíveis tornou-se essencial. Alves Et Al. (2019), enfatizam que a integração de tecnologias inteligentes e dispositivos físicos pode ser uma ferramenta eficaz para a inclusão social, especialmente para pessoas com deficiência, aumentando as suas oportunidades de interagir com o meio ambiente e facilitando as atividades diárias. Neste contexto, este trabalho propõe o desenvolvimento de um protótipo embarcado/sistema embarcado baseado em inteligência artificial voltado para a identificação e identificação de objetos no ambiente interno.

O protótipo emprega técnicas de aprendizado de máquina para treinar modelos que processam as imagens capturadas pela câmera, reconhecendo e classificando objetos na cena. Um diferencial do sistema é a conversão de texto em fala, responsável pela audiodescrição dos itens identificados, o que traz benefícios diretos — especialmente a pessoas com deficiência visual — ao favorecer maior autonomia de deslocamento, orientação e interação com o ambiente.

A pesquisa abrange todo o ciclo de desenvolvimento: concepção, curadoria/rotulagem de dados, treinamento e validação de modelos, integração com dispositivos móveis e testes de eficácia em cenários controlados e reais. À luz da literatura, a capacidade de aprender redes neurais profundas é essencial para robustez e eficiência; como apontam Goodfellow, Bengio e Courville (2016), tais redes viabilizam aplicações móveis eficazes mesmo em situações complexas, orientando as escolhas metodológicas e as futuras extensões deste trabalho.

Os resultados obtidos demonstram que o *software* atinge os objetivos propostos e consegue identificar corretamente objetos comuns em situações internas. Além disso, este trabalho apresenta a possibilidade de melhoria da solução, mostra a possibilidade de

expansão do sistema com novas situações e desenvolvimentos que podem ser desenvolvidos posteriormente.

Através deste estudo, pretendemos não só contribuir para o desenvolvimento da utilização da inteligência artificial em aplicações móveis, mas também apoiar a criação de soluções tecnológicas acessíveis e apoiar a inclusão e independência digital entre diferentes públicos.

7. CONCLUSÃO

Este trabalho apresentou o desenvolvimento de um protótipo físico embarcado, capaz de reconhecer objetos domésticos em tempo real e *off-line* e de anunciar os resultados por síntese de voz em português. A proposta foi guiada por requisitos práticos de acessibilidade: baixo custo, autonomia (sem *internet*) e latência pequena o suficiente para tornar o recurso útil no cotidiano de pessoas com deficiência visual.

Do ponto de vista técnico, a opção por classificadores em cascata (*Haar-like*), janela 24×24 px, processamento em lotes e confirmação temporal (20 quadros consecutivos) mostrou-se adequada ao *hardware* alvo. Em ambiente controlado, o sistema alcançou 92% de precisão e $F1 = 0,89$; no ambiente doméstico real, com variação de luz e *clutter*, a precisão agregada situou-se em ~85%. Esses resultados confirmam a viabilidade da abordagem escolhida e indicam um equilíbrio razoável entre precisão e revocação em cenários de uso realistas.

8. BIBLIOGRAFIA

RUSSELL, S., NORVIG. **Artificial Intelligence: A Modern Approach**. Prentice Hall. 2020. Disponível em: <https://aima.cs.berkeley.edu/global-index.html> . Acesso em: 01 Dez. 2024.

GOODFELLOW, I., BENGIO, Y., & COURVILLE, A. **Deep Learning**. MIT Press. 2016. Disponível em: <https://www.deeplearningbook.org/> . Acesso em: 03 Dez. 2024.

ALVES, R. M., SILVA, P. R., & OLIVEIRA, T. F. **Tecnologias inteligentes e inclusão social: aplicações de IA para pessoas com deficiência**. *Revista Brasileira de Tecnologia Aplicada*. 2019. Disponível em: <https://revistas.ufpr.br/extensao/article/download/35315/21992/130920> . Acesso em: 08 Dez. 2024.

COSTA, M. A., & LIMA, J. P. **Desafios na implementação de sistemas assistivos baseados em IA no Brasil**. *Anais do Congresso Brasileiro de Computação*. 2020. Disponível em: <https://sol.sbc.org.br/index.php/sbie/article/download/12911/12765> . Acesso em: 08 Dez. 2024.

SZELISKI, R. **Computer Vision: Algorithms and Applications**. Springer. 2021. Disponível em: https://vim.ustc.edu.cn/_upload/article/files/d4/87/71e9467745a5a7b8e80e94007d1b/4cd69b21-85d3-43ba-9935-fd9ae33da82b.pdf . Acesso em: 08 Dez. 2024.

LECUN, Y.; BENGIO, Y.; HINTON, G. **Deep learning**. *Nature*, **521(7553)**, 436–444. 2015. Disponível em: <https://www.nature.com/articles/nature14539> . Acesso em: 02 set. 2025.

GONZALEZ, R. C.; WOODS, R. E. **Digital Image Processing**. 4. ed. Pearson. 2018. Disponível em: <https://www.pearson.com/en-us/subject-catalog/p/digital-image-processing/P200000003169/> . Acesso em: 02 set. 2025.

ORGANIZAÇÃO MUNDIAL DA SAÚDE (OMS). **World Report on Vision**. Genebra: OMS. 2019. Disponível em: <https://www.who.int/publications/i/item/9789241516570> . Acesso em: 02 set. 2025.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). **Estatísticas de Pessoas com Deficiência 2019**. Rio de Janeiro: IBGE. 2019. Disponível em: <https://www.ibge.gov.br/> . Acesso em: 02 set. 2025.

ORGANIZAÇÃO DAS NAÇÕES UNIDAS (ONU). **Convenção sobre os Direitos das Pessoas com Deficiência**. Nova Iorque: ONU. 2006. Disponível em: <https://www.un.org/development/desa/disabilities/convention-on-the-rights-of-persons-with-disabilities.html> . Acesso em: 02 set. 2025.

WÄLDCHEN, J.; MÄDER, P. **Plant species identification using computer vision techniques: A systematic literature review**. *Ecological Informatics*, 48, 257–273. 2018. Disponível em: <https://doi.org/10.1016/j.ecoinf.2018.10.005> . Acesso em: 02 set. 2025.

BONNEY, R.; et al. **Next steps for citizen science**. *Science*, 343(6178), 1436–1437. 2014. Disponível em: <https://doi.org/10.1126/science.1251554> . Acesso em: 02 set. 2025.

DENG, J.; DONG, W.; SOCHER, R.; LI, L.-J.; LI, K.; FEI-FEI, L. **ImageNet: A large-scale hierarchical image database**. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2009. Disponível em: http://www.image-net.org/papers/imagenet_cvpr09.pdf . Acesso em: 02 set. 2025.

GSMA. **The Mobile Economy 2022**. GSMA Intelligence. 2022. Disponível em: <https://www.gsma.com/mobileeconomy/> . Acesso em: 02 set. 2025.

GOOGLE. **Google Lens hits 10 billion searches milestone**. 2022. Disponível em: <https://blog.google/products/search/google-lens-10-billion/> . Acesso em: 02 set. 2025.

VIOLA, P.; JONES, M. **Rapid object detection using a boosted cascade of simple features**. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2001. Disponível em: <https://ieeexplore.ieee.org/document/990517> . Acesso em: 02 set. 2025.

LIENHART, R.; MAYDT, J. **An extended set of Haar-like features for rapid object detection**. *Proceedings of the International Conference on Image Processing (ICIP)*. 2002. Disponível em: <https://ieeexplore.ieee.org/document/1038171> . Acesso em: 02 set. 2025.

SZELISKI, R. **Computer Vision: Algorithms and Applications**. 2. ed. Springer. 2021. Disponível em: <https://szeliski.org/Book/> . Acesso em: 02 set. 2025.

ORGANIZAÇÃO MUNDIAL DA SAÚDE (OMS); FUNDO DAS NAÇÕES UNIDAS PARA A INFÂNCIA (UNICEF). **Global Report on Assistive Technology**. 2022. Disponível em: <https://www.who.int/publications/i/item/9789240049451> . Acesso em: 05 set. 2025.

GSMA INTELLIGENCE. **Closing the Usage Gap in Brazil**. 2023. Disponível em: <https://www.gsma.com/about-us/regions/latin-america/wp-content/uploads/2023/03/EN-We-Care-Usage-Brazil-v8.pdf> . Acesso em: 05 set. 2025.

RASPBERRY PI FOUNDATION. **Camera Module documentation**. 2024. Disponível em: <https://www.raspberrypi.com/documentation/accessories/camera.html> . Acesso em: 08 set. 2025.

DATASHEETS RASPBERRY PI. **Raspberry Pi Zero 2 W**. 2024. Disponível em: <https://datasheets.raspberrypi.com/rpizero2/raspberry-pi-zero-2-w-product-brief.pdf>. Acesso em: 09 set. 2025.

HAN, SONG; HUIZI, MAO; DALLY, J.WILLIAM. **Very Deep Convolutional Networks for Large-Scale Image Recognition**. 2014. Disponível em: <https://arxiv.org/abs/1409.1556>. Acesso em: 09 set. 2025.

ARXIV. **Deep Compression: Compressing Deep Neural Networks with Pruning, Trained Quantization and Huffman Coding**. 2015. Disponível em: <https://arxiv.org/abs/1510.00149>. Acesso em: 09 set. 2025.

LIU, WEI; ANGUELOV, GRAGOMIR; ERHAN, DUMITRU. **SSD: Single Shot MultiBox Detector**. 2015. Disponível em: <https://arxiv.org/abs/1512.02325>. Acesso em: 09 set. 2025.

TENSOR FLOW. **Post-training quantization**. 2022. Disponível em: https://www.tensorflow.org/model_optimization/guide/quantization/post_training. Acesso em: 09 set. 2025.

VIOLA, PAUL; JONES, MICHAEL. **Rapid Object Detection using a Boosted Cascade of Simple Features.** 2001. Disponível em: <https://www.cs.cmu.edu/~efros/courses/LBMV07/Papers/viola-cvpr-01.pdf>. Acesso em: 09 set. 2025.

SELVARAJU, RAMPRASAATH; PARIKH, DEVI; BATRA DHRUV. **Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization.** 2017. Disponível em: https://openaccess.thecvf.com/content_ICCV_2017/papers/Selvaraju_Grad-CAM_Visual_Explanations_ICCV_2017_paper.pdf. Acesso em: 09 set. 2025.

SZELISKI, RICHARD. **Computer Vision: Algorithms and Applications, 2nd ed.** 2022. Disponível em: <https://szeliski.org/Book/>. Acesso em: 09 set. 2025.

BISHOP, M. CRISTOPHER. **Pattern Recognition and Machine Learning.** 2006. Disponível em: <https://github.com/Benlau93/Data-Science-Curriculum/blob/master/Bishop-Pattern-Recognition-and-Machine-Learning-2006.pdf> . Acesso em: 09 set. 2025.

DAVIS, JESSE; GOADRICH, MARK. **The Relationship Between Precision-Recall and ROC Curves.** 2006. Disponível em: <https://ftp.cs.wisc.edu/machine-learning/shavlik-group/davis.icml06.pdf> . Acesso em: 09 set. 2025.

ALPAYDIN, ETHEM. **Introduction to Machine Learning.** 4 ed. Cambridge, MA. MIT Press, 2020.

MITCHELL, TOM. **Machine Learning.** New York: McGraw-Hill, 1997.

POWERS, D. M. W. **Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness & Correlation. Journal of Machine Learning Technologies (Tech. Report).** 2011. Disponível em: <https://arxiv.org/abs/2010.16061> . Acesso em: 15 set. 2025.

WORLD HEALTH ORGANIZATION. **World Report on Disability. WHO.** 2011. Disponível em: <https://www.who.int/publications/i/item/9789241564182> . Acesso em: 15 set. 2025.

PRADHAN, A.; ET AL. **Voice Interfaces for Accessibility: A Research Agenda. ACM Interactions (blog/essay).** 2018. Disponível em: <https://interactions.acm.org/archive/view/may-june-2018/voice-interfaces-for-accessibility> . Acesso em: 15 set. 2025.

VALENTINI-BOTINHAO, C.; ET AL. **Noisy Speech Database for Training Speech Enhancement Algorithms and TTS Models. University of Edinburgh.** 2016. Disponível em: <https://datashare.ed.ac.uk/handle/10283/2791> . Acesso em: 15 set. 2025.

LOIZOU, P. C. **Speech Enhancement: Theory and Practice. 2ª ed. CRC Press.** 2013. Disponível em: <https://www.routledge.com/Speech-Enhancement-Theory-and-Practice/Loizou/p/book/9781138073372> . Acesso em: 15 set. 2025.

VAN ROSSUM, G.; DRAKE, F. L. **Python Tutorial. Python Software Foundation.** 2009. Disponível em: <https://docs.python.org/3/tutorial/> . Acesso em: 15 set. 2025.

PYTHON SOFTWARE FOUNDATION. **Python 3 Standard Library. PSF.** 2024. Disponível em: <https://docs.python.org/3/library/> . Acesso em: 15 set. 2025.

OPENCV TEAM. **OpenCV Documentation. OpenCV.org.** 2024. Disponível em: <https://docs.opencv.org/> . Acesso em: 15 set. 2025.

GOOGLE RESEARCH; TENSORFLOW TEAM. **TensorFlow: An End-to-End Open Source Machine Learning Platform. Google.** 2024. Disponível em: <https://www.tensorflow.org/> . Acesso em: 15 set. 2025.

PYTORCH FOUNDATION. **PyTorch Documentation. The Linux Foundation.** 2024. Disponível em: <https://pytorch.org/> . Acesso em: 15 set. 2025.

NUMPY DEVELOPERS. **NumPy: The Fundamental Package for Scientific Computing with Python. NumPy.** 2024. Disponível em: <https://numpy.org/> . Acesso em: 15 set. 2025.
RASPBERRY PI LTD. **Picamera2 Documentation. Raspberry Pi.** 2024. Disponível em: <https://datasheets.raspberrypi.com/camera/picamera2-manual.pdf> . Acesso em: 15 set. 2025.

ESPEAK-NG CONTRIBUTORS. **eSpeak NG Text-to-Speech. eSpeak-NG.** 2024. Disponível em: <https://espeak-ng.org/> . Acesso em: 15 set. 2025.

PYTTSX3 DEVELOPERS. **pyttsx3: Text-to-Speech Conversion Library in Python. Read the Docs.** 2024. Disponível em: <https://pyttsx3.readthedocs.io/> . Acesso em: 15 set. 2025.

AHMED, M.; NANDHINI, R. **Assistive Technology for Visually Impaired Individuals: A Systematic Literature Review. Journal of Assistive Technologies.** 2024. Disponível em: https://www.researchgate.net/publication/378254054_Assistive_Technology_for_Visually_Impaired_Individuals_A_Systematic_Literature_Review_SLR . Acesso em: 09 out. 2025.

BAMDAD, M.; SCARAMUZZA, D.; DARVISHY, A. **Simultaneous Localization and Mapping (SLAM) for Visually Impaired Navigation: A Systematic Review.** 2022. Disponível em: <https://arxiv.org/abs/2212.04745> . Acesso em: 09 out. 2025.

ORTIZ-ESCOBAR, M. C.; PEREIRA, M.; SILVA, R.; LOBO, M. **Assessing User-Centred Design in Assistive Technologies for the Visually Impaired. Disability and Rehabilitation: Assistive Technology, v.18, n.5, p. 465-478.** 2023. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/37744430/> . Acesso em: 09 out. 2025.

PATIL, R.; KUMAR, A.; RAO, V. **Advancements in Smart Wearable Mobility Aids for the Visually Impaired: A Review. Sensors, v.24, n.24, 7986.** 2024. Disponível em: <https://www.mdpi.com/1424-8220/24/24/7986> . Acesso em: 09 out. 2025.

SINGH, S.; KHAN, N.; JAIN, A.; CHANDRA, R. **Towards Assisting Visually Impaired Individuals: A Review on Current Assistive Devices. Biomedical Signal Processing and Control, v.78, 103929.** 2022. Disponível em: <https://www.sciencedirect.com/science/article/pii/S2590137022001583> . Acesso em: 09 out. 2025.

YAO, L.; ZHOU, H.; HU, Z. **A Review of Vision-Based Assistive Systems for Visually Impaired. arXiv preprint.** 2025. Disponível em: <https://arxiv.org/abs/2505.14298> . Acesso em: 09 out. 2025.

EVERINGHAM, M.; et al. **The Pascal Visual Object Classes (VOC) Challenge.** *International Journal of Computer Vision (IJCV)*, v.88, n.2, p.303–338. 2010.

LIN, T.-Y.; MAIRE, M.; BELONGIE, S.; et al. **Microsoft COCO: Common Objects in Context.** *European Conference on Computer Vision (ECCV)*, p.740–755. 2014.