

Sistema de Baixo Custo para Auxílio de Detecção de Obstáculos para Pessoas com Deficiência Visual

Augusto Scarduelli Prudencio
augusto.s.prudencio@gmail.com

Universidade Federal de Santa Catarina - Campus
Araranguá
Araranguá, Brasil

Jim Lau
jim.lau@ufsc.br

Universidade Federal de Santa Catarina - Campus
Araranguá
Araranguá, Brasil

Pedro Menezes Pimenta
pedro.pimenta.ufsc@gmail.com

Universidade Federal de Santa Catarina - Campus
Araranguá
Araranguá, Brasil

Antonio Carlos Sobieranski
a.sobieranski@ufsc.br

Universidade Federal de Santa Catarina - Campus
Araranguá
Araranguá, Brasil

ABSTRACT

In this work, a prototype for a low-cost application was developed aiming to assist vision impaired individuals. The prototype is a wearable device, equipped with a Raspberry Pi for image processing of the frames captured from a camera strapped to the user's chest. The feedback is provided by means of vibration motors, attached to an armband, allowing to distinguish obstacles according to its proximity. The obtained preliminary results demonstrated the feasibility of the proposed approach, collaborating with a better mobility process with low-cost setups to assist the detection of obstacles.

KEYWORDS

Visão Computacional, Tecnologia Assistiva, Baixo Custo, Raspberry Pi, Deficiente Visual, Resposta Tátil, Vestível

1 INTRODUÇÃO

De acordo com a Organização Mundial de Saúde (OMS) [8], estima-se que pelo menos 2,2 bilhões de pessoas apresentem algum tipo de deficiência visual, seja para visão de perto ou de longe. Essa condição acarreta um custo anual global de aproximadamente US\$411 bilhões de dólares.

Para o Brasil, segundo a Pesquisa Nacional de Saúde (PNS) [3] de 2019 feita pelo IBGE, pessoas com deficiência visual representam 3,4% do total da população, o que gira em torno de 7 milhões de pessoas, já as que necessitam de óculos ou outro tipo de equipamento para visão são 40,4% da população. Na PNS de 2010 [1] foi registrado que aproximadamente 528 mil pessoas são incapazes de enxergar.

O dispositivo assistivo mais comum entre pessoas com deficiência visual significativa é a bengala, oferecendo um grande auxílio ao qual estão habituadas. Segundo [4], embora soluções inovadoras tenham sido desenvolvidas desde a década de 1950, nenhuma delas alcançou uma taxa de aceitação superior à da bengala como principal auxiliar na locomoção. O custo de uma bengala convencional é aproximadamente R\$150, enquanto o modelo avançado da Bengala Inteligente WeWalk pode atingir até R\$5,4 mil.

Nos últimos anos, as soluções acadêmicas para essa área têm empregado processamento de imagem por meio de IA, demandando

considerável capacidade computacional, como destacado em [4] e [11]. Isso resulta em custos elevados para suas aplicações, devido às especificações necessárias.

Neste artigo, com o intuito de auxiliar a locomoção de indivíduos com deficiência visual severa, foi desenvolvido uma tecnologia assistiva de baixo custo para auxiliar esse grupo na detecção de objetos situados aproximadamente a 1,5 metro do usuário. Vale ressaltar que esta abordagem não substitui o uso da bengala, mas sim se destina a complementá-la.

Na Seção 2 abordamos alguns trabalhos relevantes desenvolvidos na mesma área de pesquisa. Em seguida na Seção 3 descrevemos o processo de desenvolvimento, os dispositivos utilizados e a implementação do algoritmo proposto. Na Seção 4 realizamos uma comparação entre os resultados de distância obtidos pela câmera e um *LiDAR* de celular. Por último, na Seção 5 apresentamos as conclusões provenientes do estudo e discutimos algumas ideias de trabalhos futuros.

2 TRABALHOS RELACIONADOS

Nos últimos anos, observou-se um considerável desenvolvimento de artigos na interseção entre tecnologia assistiva e visão computacional. Os artigos selecionados para este trabalho foram agrupados em duas categorias principais: Codificação e Mobilidade.

No primeiro grupo, destacam-se aqueles que exploram a codificação ou transcrição do ambiente para outros sentidos, permitindo uma interpretação mais acessível para usuários com deficiência visual. Em [12], propõe-se a codificação de uma câmera RGB em um dispositivo vestível com formato de camisa, equipado com motores de vibração. Já em [6], a transcrição para um *display* em braille é realizada para representar o contorno dos objetos na cena, facilitando sua identificação. Em [5], é desenvolvido um dispositivo que auxilia pessoas com deficiência visual a mirar um alvo em um *display*, utilizando *feedback* vibrotátil. Além disso, são testadas três técnicas (aleatória, cruciforme e radial) para determinar a mais eficiente na localização do alvo. Em [10], propõe-se uma técnica de reconhecimento de imagem que permite captar de maneira clara os sinais de trânsito de forma confiável, robusta e eficiente. Por fim, em [4], são realizados cálculos para traduzir os *pixels* RGBD do ambiente em frequências sonoras, fornecendo *feedback* por meio de um fone de condução óssea.

No grupo de iniciativas relacionadas à mobilidade, agrupam-se os trabalhos que oferecem assistência direta ou indireta na locomoção das pessoas. No estudo [9], aborda-se a manutenção do distanciamento social durante a pandemia de COVID-19, empregando algoritmos de segmentação semântica para localizar pessoas por meio de uma câmera RGB. Além disso, utilizando uma câmera de distância, identificam-se se estão a 1,5 metros de distância do indivíduo. No artigo [2], utiliza-se o Tiny-YOLOv3 e marcadores para auxiliar esse grupo a se locomover em ambientes internos. Em [16], desenvolve-se um sistema para identificar perigos com base no contexto, abrangendo ambientes tanto internos quanto externos.

Uma comparação entre tecnologias que utilizam visão computacional e aquelas que empregam BLE (*Bluetooth Low Energy*) para auxílio na locomoção em ambientes internos é explorada em [7]. No estudo de [15], os autores propõem um sistema utilizando óculos inteligentes e realidade aumentada para detectar objetos em movimento que podem representar perigos nas áreas de perda de visão periférica.

No trabalho [14], é apresentado o sistema DEEP-SEE, que utiliza algoritmos de visão computacional e redes neurais convolucionais profundas (CNNs) para detecção, rastreamento e reconhecimento de objetos em tempo real. Por fim, em [11], oferece-se suporte à locomoção de pessoas com deficiência visual e cadeirantes, utilizando deep learning em ambientes tanto internos quanto externos.

3 DESENVOLVIMENTO

Diante dos desafios enfrentados por indivíduos com severa deficiência visual em suas atividades diárias de navegação, o presente trabalho propõe um sistema inovador destinado a auxiliar na detecção de obstáculos. A essência desse sistema reside na adoção de tecnologia de última geração, que combina uma câmera de distância para capturar informações essenciais do ambiente e um processamento inteligente realizado por meio de um Raspberry Pi 3B.

A câmera de distância constitui o núcleo sensorial do sistema, capturando dados precisos sobre a disposição dos objetos no ambiente. Por meio da análise dos *pixels* da imagem, foram extraídas informações cruciais acerca das distâncias entre o usuário e os obstáculos ao redor. Essa valiosa informação é, então, encaminhada ao Raspberry Pi 3B, desempenhando um papel essencial no processamento e na tomada de decisões em tempo real.

Após a análise dos dados, o sistema proporciona um retorno imediato ao usuário através de três motores de vibração estrategicamente posicionados. Esses motores têm a função de comunicar a direção do obstáculo detectado: à esquerda, à frente ou à direita. Essa resposta tátil é fundamental para fornecer ao usuário informações vitais sobre o ambiente ao seu redor, possibilitando uma navegação mais segura e autônoma.

O núcleo desse sistema reside na sinergia entre a tecnologia de detecção precisa de distâncias e o processamento rápido e eficaz do Raspberry Pi 3B, combinados para oferecer um auxílio confiável na detecção de obstáculos. A abordagem tátil, por meio dos motores de vibração, constitui uma interface intuitiva e eficaz para comunicar as informações necessárias ao usuário, permitindo uma maior independência e confiança em suas operações diárias de navegação.

Este sistema não apenas busca aprimorar a experiência das pessoas com severa deficiência visual, mas também representa um

avanço na aplicação da tecnologia para promover a inclusão e a autonomia. Ele viabiliza uma vida mais independente e acessível a todos, reforçando o compromisso com a promoção de um ambiente inclusivo e igualitário.

3.1 Câmera

A câmera empregada desempenha a leitura das distâncias no ambiente e é denominada MaixSense-A010, produzida pela SiPEED, com um custo aproximado de R\$210. Seu funcionamento envolve a emissão de uma onda no espectro eletromagnético e a detecção da onda refletida, com o cálculo da distância realizado por meio do tempo de “voo” (*time of flight*). Essa tecnologia é conhecida como LiDAR (*Light Detection and Ranging*).

O módulo apresenta dois botões laterais para a alteração manual de algumas configurações, como a taxa de quadros e a resolução. Além disso, conta com uma tela LCD que se revelou extremamente útil para verificar a correta reconstrução da imagem. As opções de conexão disponíveis são UART ou USB-C na parte inferior. A Figura 1 exibe uma representação visual do referido módulo.

A resolução adotada foi de 25×25, na qual regiões de 16 *pixels* da matriz 100×100 são agrupadas para alcançar uma resolução inferior. Em outras palavras, a matriz maior é dividida por quatro, de modo que cada pixel da matriz menor corresponde a uma região 4×4 da matriz 100×100. A taxa de quadros selecionada foi de 10 quadros por segundo, resultando em 10 medidas do ambiente por segundo. Essas configurações foram deliberadamente escolhidas em função de considerações de desempenho, dada a restrição de memória RAM de 1GB no Raspberry Pi. A resolução mais baixa revelou-se suficiente para obter uma imagem clara dos obstáculos no ambiente, enquanto a taxa de quadros foi otimizada para fornecer informações adequadas para uma velocidade de caminhada.



Figura 1: Câmera MaixSense-A010. Fonte: SiPEED

3.2 Resposta Tátil

Para a tradução dos dados captados pela câmera em uma resposta tátil, foram empregados pequenos módulos de motores de vibração. Esses módulos foram estrategicamente posicionados em uma pulseira improvisada para facilitar a identificação da região do pulso correspondente à vibração. Cada módulo já é equipado com um

circuito que controla a alimentação e a geração de uma frequência de vibração.

A configuração de cada módulo inclui três pinos de conexão, sendo que um par é destinado à alimentação e ao aterramento, enquanto o terceiro pino é utilizado para fornecer a frequência de vibração. A Figura 2 ilustra o módulo do motor de vibração empregado nesse contexto.



Figura 2: Módulo motor de vibração. Fonte: Autores

Optamos por utilizar uma única frequência para a resposta tátil, alcançada simplesmente inserindo um sinal de nível alto no pino de entrada. Em testes, essa abordagem demonstrou ser uma escolha eficiente, uma vez que não interferiu na identificação das origens das vibrações quando mais de um motor estava em funcionamento.

3.3 Implementação

Inicialmente, enviamos as configurações iniciais da câmera via protocolo USB. Optamos por uma resolução de 25×25 , para redução de ruído e uma taxa de 10 quadros por segundo, deste modo exigindo um menor processamento e atingindo uma resposta mais próxima em tempo real. A partir do quadro enviado pela câmera para o Raspberry Pi, realizamos os cálculos de distância na matriz quadrada de ordem 25.

A ativação dos motores de vibração ocorre por meio de *threads*. Após obtermos a matriz de distâncias, a *thread* é executada somente se houver detecção em uma distância de até 1,5 metros. Em seguida, verificamos em quais regiões da matriz há algum elemento nessa faixa de distância: esquerda (da posição 0 a 7), centro (da posição 8 a 15) ou direita (da posição 16 a 25). Por fim, ocorre o acionamento do motor correspondente à região da matriz identificada.

O desenvolvimento do código foi realizado em Python, e podemos observar o funcionamento do algoritmo no diagrama apresentado na Figura 4, que detalha todos os passos desde a captura da imagem até o processo que aciona os motores. Na Figura 3, apresentamos o protótipo vestido por uma pessoa de 181cm de altura com uma visão fronta (a) e uma visão lateral (b).

4 RESULTADOS

Para a obtenção dos resultados, inicialmente testamos o desempenho do sistema com as configurações originais: uma matriz de 100×100 para a imagem proveniente da câmera e 15 *frames* por

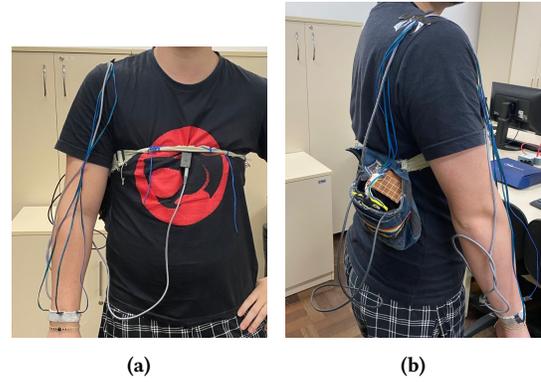


Figura 3: (a) Visão frontal e (b) Traseira do equipamento montado. Fonte: Autores

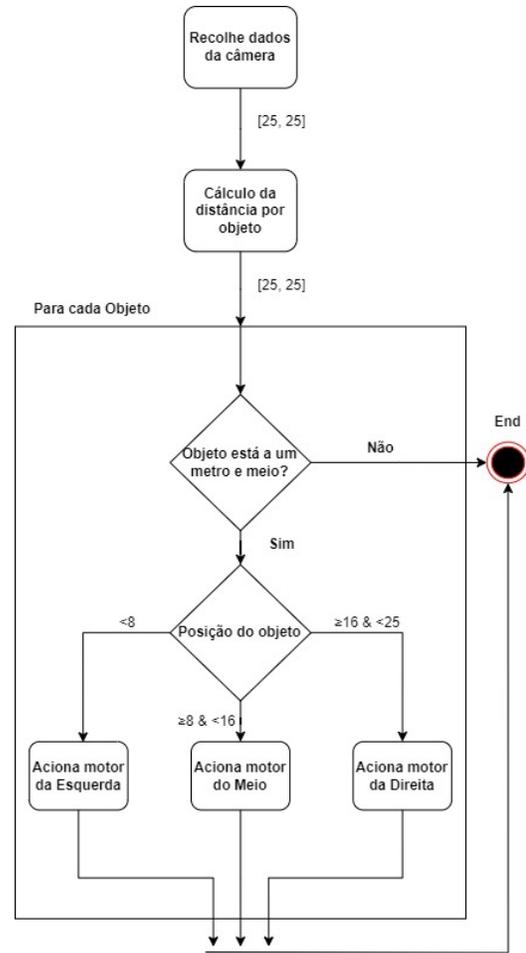


Figura 4: Diagrama do algoritmo implementado. Fonte: Autores

segundo. No entanto, nesse cenário, observamos um atraso de até 4 segundos para que o Raspberry Pi 3B pudesse processar cada *frame*,

além de receber considerável ruído, o que dificultava a utilização do algoritmo implementado. Assim, tornou-se inviável, considerando a necessidade de uma resposta em tempo real de qualidade para auxiliar pessoas com deficiência.

Conseqüentemente, reduzimos a taxa de quadros para 10 por segundo e a matriz de imagem para 25×25. Essa adaptação resultou em uma resposta suficientemente rápida para uso em tempo real e no cotidiano da pessoa, superando as limitações identificadas anteriormente.

O algoritmo e o sistema que acionam os motores de vibração são considerados triviais por não apresentarem nenhuma complexidade e serem sucintos, como apresentado na Seção 3.3. Portanto, a avaliação concentra-se na acurácia da câmera e na qualidade dos *pixels* de distância.

Para essa avaliação, conforme ilustrado na Figura 5, montamos uma estrutura com obstáculos simulados e posicionamos a câmera em um suporte à altura do peito de uma pessoa com 181 centímetros.

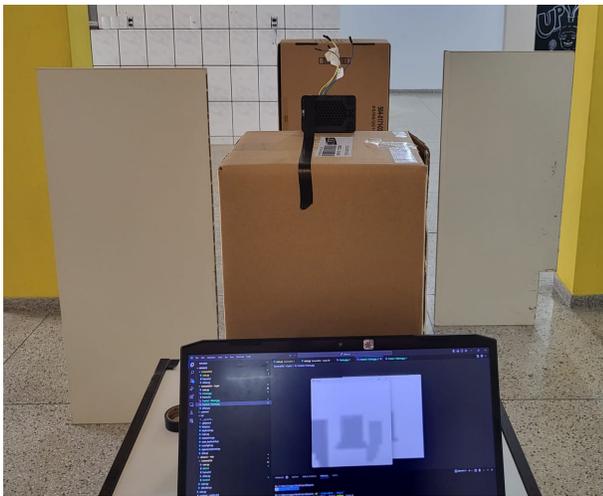


Figura 5: Configuração para aquisição de resultados da câmera. Ela foi posicionada na altura do peito de uma pessoa de 1 metro e 81 centímetros de altura. Fonte: Autores

Foram adquiridas as Figuras 6, 7, 8 e 9, sendo que as distâncias (D) dos *pixels* centrais de cada objeto foram calculadas de acordo com a Equação 1, fornecida pelo fabricante da câmera por meio do *website* [13]. A fim de avaliar os resultados, procedeu-se à divisão da imagem em três partes iguais da esquerda para a direita, considerando a posição do objeto conforme ilustrado na Figura 4. Posteriormente, os objetos principais em cada uma das partes foram isolados, e a distância de seus centros foi medida utilizando a câmera.

$$D = \left(\frac{p}{5.1}\right)^2 \quad (1)$$

Além disso, empregamos o *LiDAR* de um dispositivo celular para mensurar a distância real dos objetos na cena, com o propósito de realizar uma comparação com os resultados obtidos pela câmera, como mostrado na Tabela 1. Com o intuito de facilitar a identificação das distâncias entre os objetos, aplicamos um mapeamento de cores utilizando a função *applyColorMap* do OpenCV com o esquema

COLORMAP_TWILIGHT_SHIFTED. Desse modo, é possível observar que a imagem da Figura 6 acionaria os motores da esquerda e do centro; o da Figura 7 acionaria todos os motores; a imagem da Figura 8 acionaria apenas o motor do centro; e, na Figura 9, nenhum motor seria acionado.

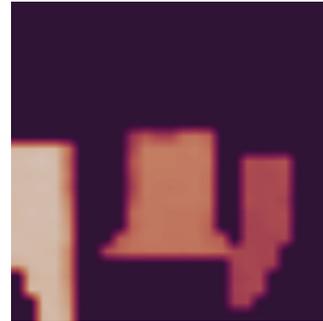


Figura 6: Imagem capturada pela câmera de distancia com objetos à 96cm, 115cm e 157cm respectivamente. Colorido depois da aquisição. Fonte: Autores



Figura 7: Imagem capturada pela câmera de distancia com objetos à 100cm, 142cm e 100cm respectivamente. Colorido depois da aquisição. Fonte: Autores

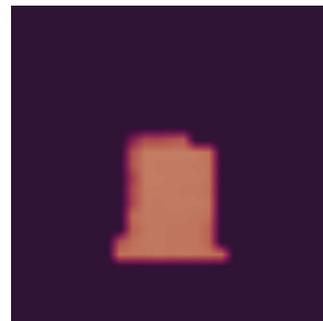


Figura 8: Imagem capturada pela câmera de distancia com um objeto à 115cm. Colorido depois da aquisição. Fonte: Autores

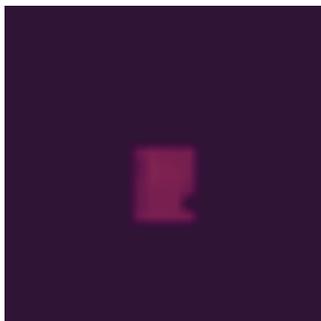


Figura 9: Imagem capturada pela câmera de distancia com um objeto à 180cm. Colorido depois da aquisição. Fonte: Autores

Figura	Distância da câmera dos objetos na parte da esquerda, centro e direita da imagem	Distância medida pelo LiDAR
Figura 6	113.7cm, 117.7cm, 158.4cm	96cm, 115cm, 157cm
Figura 7	91.2cm, 144.7cm, 92.4cm	100cm, 142, 100cm
Figura 8	119.1cm	115cm
Figura 9	186.1cm	180cm

Tabela 1: Comparação das distancias medidas por LiDAR e pelas medidas pela câmera de distância

5 CONCLUSÃO

Neste trabalho foi desenvolvido um protótipo como uma solução de baixo custo para a detecção de obstáculos a um metro e meio de distância, utilizando uma câmera acoplada ao peito de uma pessoa com deficiência visual. Nos resultados apresentados verificamos o funcionamento e qualidade da câmera e no desenvolvimento foi explicado o funcionamento do algoritmo de processamento de imagem em tempo real, e ativação da resposta tátil do sistema embarcado.

Esta solução destaca-se pela praticidade, exigindo menos equipamento e poder computacional, além de alcançar uma resposta em tempo real. Dessa forma, é possível mantê-la de maneira minimalista, contribuindo para a mobilidade ao dispensar a necessidade de um computador para processamento de imagens pesadas com IA.

Para trabalhos futuros tem-se em vista a realização de testes com grupos de pessoas com alguma deficiência visual, em ambiente controlado, para coleta e comparação dos dados obtidos. Dessa forma, será possível desenvolver funcionalidades que melhor se adequem às necessidades dos usuários. E também melhorias no retorno tátil como variação de vibração de acordo com a distância do objeto.

REFERÊNCIAS

- [1] Instituto Brasileiro de Geografia e Estatística IBGE. 2011. *Sinopse do Censo Demográfico 2010*. Instituto Brasileiro de Geografia e Estatística - IBGE, Rio de Janeiro. <http://www.ibge.gov.br/home/estatistica/populacao/censo2010/sinopse.pdf>
- [2] Mostafa Elgendy, Cecilia Sik-Lanyi, and Arpad Kelemen. 2021. A Novel Marker Detection System for People with Visual Impairment Using the Improved Tiny-YOLOv3 Model. *Computer Methods and Programs in Biomedicine* 205 (06 2021), 106112. <https://doi.org/10.1016/j.cmpb.2021.106112>
- [3] IBGE Instituto Brasileiro de Geografia e Estatística. 2019. *Painel de Indicadores de Saúde – Pesquisa Nacional de Saúde*. Retrieved 07 de Dezembro de 2023 from <https://www.pns.icict.fiocruz.br/painel-de-indicadores-mobile-desktop/>
- [4] Johan Isaksson-Daun, Tomas Jansson, and Johan Nilsson. 2020. Audomni: Super-Scale Sensory Supplementation to Increase the Mobility of Blind and Low-Vision Individuals – A Pilot Study. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 28 (05 2020), 1187–1197. <https://doi.org/10.1109/TNSRE.2020.2985626>
- [5] Kibum Kim, Xiangshi Ren, Seungmoon Choi, and Hong Z. Tan. 2016. Assisting people with visual impairments in aiming at a target on a large wall-mounted display. *International Journal of Human-Computer Studies* 86 (2016), 109–120. <https://doi.org/10.1016/j.ijhcs.2015.10.002>
- [6] Seondae Kim, Yeongil Ryu, Jinsoo Cho, and Eun-Seok Ryu. 2019. Towards Tangible Vision for the Visually Impaired through 2D Multiarray Braille Display. *Sensors* 19, 23 (2019). <https://doi.org/10.3390/s19235319>
- [7] Jayakanth Kunthoth, AbdelGhani Karkar, Somaya Al-ma’adeed, and Asma Al-Attayah. 2019. Comparative analysis of computer-vision and BLE technology based indoor navigation systems for people with visual impairments. *International journal of health geographics* 18 (12 2019), 29. <https://doi.org/10.1186/s12942-019-0193-9>
- [8] Sebastian Liste. 2023. *Blindness and vision impairment*. Retrieved 07 de Dezembro de 2023 from <https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment>
- [9] Manuel Martinez, Kailun Yang, Angela Constantinescu, and Rainer Stiefelhagen. 2020. Helping the Blind to Get through COVID-19: Social Distancing Assistant Using Real-Time Semantic Segmentation on RGB-D Video. *Sensors* 20, 18 (2020). <https://doi.org/10.3390/s20185202>
- [10] Sergio Mascetti, Dragan Ahmetovic, Andrea Gerino, Cristian Bernareggi, Mario Busso, and Alessandro Rizzi. 2016. Robust traffic lights detection on mobile devices for pedestrians with visual impairment. *Computer Vision and Image Understanding* 148 (2016), 123–135. <https://doi.org/10.1016/j.cviu.2015.11.017> Special issue on Assistive Computer Vision and Robotics - "Assistive Solutions for Mobility, Communication and HMI".
- [11] Elhassan Mohamed, Konstantinos Sirlantzis, and Gareth Howells. 2021. Indoor/Outdoor Semantic Segmentation Using Deep Learning for Visually Impaired Wheelchair Users. *IEEE Access* 9 (2021), 147914–147932. <https://doi.org/10.1109/ACCESS.2021.3123952>
- [12] Ricardo Ribani, Rodrigo Freitas Lima, and Mauricio Marengoni. 2019. Toward Encoding Vision-to-Touch With Convolutional Neural Networks. *IEEE Access* 7 (2019), 161175–161183. <https://doi.org/10.1109/ACCESS.2019.2951614>
- [13] Sipeed. 2023. MaixSense-A010 Development. https://wiki.sipeed.com/hardware/en/maixsense/maixsense-a010/at_command_en.html#Image-Packet-Description Acessado dia 09/03/2024.
- [14] Ruxandra Tapu, Bogdan Mocanu, and Titus Zaharia. 2017. DEEP-SEE: Joint Object Detection, Tracking and Recognition with Application to Visually Impaired Navigational Assistance. *Sensors* 17, 11 (2017). <https://doi.org/10.3390/s17112473>
- [15] Ola Younis, Waleed Al-Nuaimy, Mohammad H. Alomari, and Fiona Rowe. 2019. A Hazard Detection and Tracking System for People with Peripheral Vision Loss using Smart Glasses and Augmented Reality. *International Journal of Advanced Computer Science and Applications* 10, 2 (2019). <https://doi.org/10.14569/IJACSA.2019.0100201>
- [16] Ola Younis, Waleed Al-Nuaimy, Fiona Rowe, and Mohammad H. Alomari. 2019. A Smart Context-Aware Hazard Attention System to Help People with Peripheral Vision Loss. *Sensors* 19, 7 (2019). <https://doi.org/10.3390/s19071630>