**Second Vision - Sistema para auxílio de deficientes visuais**

*Second Vision - System to help the visually impaired*

*Second Vision - Sistema de ayuda a los discapacitados visuales*

**Gustavo Mendes Ventieri Mariano [[1]](#footnote-1)**

*ventierigustavo@gmail.com*

**Nickolas Maia de Araujo 1**

*Nickolasmaraujo@gmail.com*

**Pedro Fernandes Araújo1**

*pedrofeearaujo@gmail.com*

**Tiago Bryan Ramos de Oliveira 1**

*tiagobryanroliveira@gmail.com*

**Jeferson Roberto de Lima1**

*Jeferson.lima17@etec.sp.gov.br*

|  |  |
| --- | --- |
| **Palavras-chave:**  *Sistema.*  *Autonomia.*  *Deficiência Visual.*  *Visão Computacional.*  **Keywords:**  *System.*  *Autonomy.*  *Visually Impaired.*  *Computer Vision.*  **Palabras clave:**  *Sistema.*  *Autonomía.*  *Discapacidad Visual.*  *Visión Por Ordenador.*  **Apresentado em:**  05 dezembro, 2024  **Evento:**  7º EnGeTec  **Local do evento:**  Fatec Zona Leste  **Avaliadores:**  Avaliador 1  Avaliador 2  [Desenho com traços pretos em fundo branco e letras pretas em fundo branco  Descrição gerada automaticamente com confiança média](https://creativecommons.org/licenses/by-nc-sa/4.0/) | **Resumo:**  Este trabalho apresenta um sistema que auxilia a autonomia e inclusão de pessoas com deficiência visual. Por meio da tecnologia vestível e da visão computacional, o sistema oferece uma gama de funcionalidades que podem auxiliar a locomoção. O objetivo principal é promover a inclusão social e melhorar a qualidade de vida dessa parcela da população, que enfrenta sérios desafios e desigualdades, mesmo com os benefícios legais. A pesquisa foi baseada em estudo quali-quantitativo para identificar as principais dificuldades dos deficientes visuais. O sistema abarca um IoT para a captação de objetos possivelmente perigosos e textos em placas estáticas em um software para gerar a saída dessa detecção em áudio. Espera-se que o sistema proposto tenha um impacto positivo na vida dessas pessoas, reduzindo o número de acidentes, melhorando a saúde física, mental e autoestima, promovendo a inserção comunitária e, por fim, elevando a qualidade de vida. Conclui-se que, por meio desse trabalho, o público-alvo consiga uma melhora na condição de vida, beneficiando-se de um possível ganho de autonomia para atenuar a marginalização desse grupo.  **Abstract:**  This paper presents a system that aids the autonomy and inclusion of visually impaired people. Using wearable technology and computer vision, the system offers a range of functionalities that can help people get around. The main objective is to promote social inclusion and improve the quality of life of this section of the population, which faces serious challenges and inequalities, even with legal benefits. The research was based on a qualitative-quantitative study to identify the main difficulties faced by the visually impaired. The system includes an IoT to capture possibly dangerous objects and texts on static signs and software to generate audio output from this detection. It is hoped that the proposed system will have a positive impact on the lives of these people, reducing the number of accidents, improving their physical and mental health and self-esteem, promoting community integration, and, ultimately, raising their quality of life. It is concluded that, through this work, the target audience will achieve an improvement in their living conditions, benefiting from a possible gain in autonomy to mitigate the marginalization of this group.  **Resumen:**  Este trabajo presenta un sistema que ayuda a la autonomía e inclusión de personas con discapacidad visual. A través de tecnología wearable y visión por computador, el sistema ofrece una serie de funcionalidades que pueden ayudar a la locomoción. El objetivo principal es promover la inclusión social y mejorar la calidad de vida de este sector de la población, que se enfrenta a graves desafíos y desigualdades, incluso con beneficios legales. La investigación se basó en un estudio cualitativo-cuantitativo para identificar las principales dificultades a las que se enfrentan las personas con discapacidad visual. El sistema incluye un IOT para detectar objetos y textos posiblemente peligrosos en señales estáticas y un software para generar una salida de audio a partir de esta detección. Se espera que el sistema propuesto tenga un impacto positivo en la vida de estas personas, reduciendo el número de accidentes, mejorando su salud física y mental y su autoestima, promoviendo la integración en la comunidad y, en definitiva, aumentando su calidad de vida. Se concluye que, a través de este trabajo, el público objetivo conseguirá una mejora en sus condiciones de vida, beneficiándose de una posible ganancia de autonomía que mitigue la marginación de este colectivo. |

# Introduction

The visually impaired are all those with total or partial loss of sight, making up a significant portion of the Brazilian population (ALVES; DUARTE, 2008), approximately 8.9% of citizens (PNAD CONTÍNUA, 2022). As a result of this condition, they face psychological, social and economic problems, resulting in loss of self-esteem, status, occupational restrictions and reduced income (FILHO, 2012). As a result, this study aims to restore some of the autonomy of the visually impaired, provide a comprehensive form of personal prevention and improve the psychological and economic aspects of those affected by visual impairment.

Furthermore, assistive technologies do not come at an inclusive price for people with visual impairments, and due to their high cost, a portion of the population is neglected and silenced (G1, 2022). As a result, digital inclusion is needed to encourage the exercise of citizenship through technology (FILHO, 2003 apud SAMPAIO, 2006).

Given this scenario, documentation was developed based on use-case, sequence, activity and state-machine diagrams to structure the project, facilitating its effective development and maintenance, if necessary. Using an assistive technology to help the visually impaired to be autonomous in cities, based on computer vision for the analysis of static texts and potentially dangerous objects catalogued on the basis of models pre-established by the YOLOv8 artificial neural network, and qualitative and quantitative research and a bibliographic survey, software was developed to process the data detected using the Raspberry Pi 5 microcomputer integrated with the Python programming language and the OpenCV and Pytesseract libraries. The verified data is sent to the visually impaired person via a multiplatform mobile application developed in React Native and paired with the microcomputer via Bluetooth.

Subsequent chapters will present all the stages in the process of building the Second Vision - Autonomy Aid System for the Visually Impaired in the Metropolis project - comprising the theoretical framework, method, final considerations and references, with argumentative support given to the project by the company Meta on React Native technology, Van Rossum, creator of the Python programming language and Lakatos and Marconi on the research methodologies used.

1. **Theoretical Foundation**

In this section, introduce the main concepts and used technologies that underpin our theoretical foundation for the Second Vision development.

## Visually Impaired

Around 285 million people have some degree of visual impairment, of which approximately 40 million are totally blind (World Health Organization, 2023). These figures highlight the significant proportion of the global population facing this condition. Furthermore, it is essential to prioritize the health and well-being of the entire population, highlighting the importance of meeting the needs of this significant portion of society (Sustainable Development Goals, 2024).

## Computer Vision

Computer vision is conceived as a set of responsible methods and destined for machine vision when acquiring, processing, analyzing and interpreting data from the environment, allowing it to manipulate and recognize objects (FÖRSTNER; WROBEL, 2016). The artificial procedures that make computer vision suitable are image processing and pattern recognition (KLETTE, 2014). Moreover, computer vision establishes semantic descriptions of images, assigning them interpretations and classifying them into equivalent sets by identifying similar features (KHAN et al., 2018).

## Raspberry PI 5

The Raspberry Pi, created by the Raspberry Pi company in 2012, is a low-cost microcomputer and consequently began to be widely used for automation in everyday life (OLIVEIRA; NABARRO; ZANETTI, 2018).

According to figure one, The Raspberry PI 5 has a processing power of 2.4GHz, Random Access Memory (RAM) memory capacity of up to 8GB, four Universal Serial Bus (USB) ports, and connection with the internet (OFFICIAL RASPBERRY PI DOCUMENTATION, 2024).

The microcomputer will be used within the context of the project to process the object and text information coming from the camera, allowing the recognition of these detectable objects and the effective operation of the system without the intervention of cloud processing, facilitating the use of Second Vision in remote locations that do not have access to a quality signal.

## Python

The Python programming language is a general-purpose language with a focus on high readability and multi-paradigm, providing developers with greater productivity (VAN ROSSUM, 2007).

In addition, due to a programming language with a low learning curve and available for the most diverse operating systems, Python has become an essential technology for the development of large and scalable systems (LUTZ, 2001).

The Python language is used to implement the Generic Attribute Profile (GATT) server, which handles all the information coming from the camera, from its processing to its return for use by the mobile application.

## Open Source Computer Vision Library (OpenCV)

OpenCV, developed by Intel Corporation, is an open source library and multiplatform, created to facilitate the users access and the developers access to multiplatform computer vision (MARENGONI; STRINGHINI, 2009).

The aim of the OpenCV library is to provide tools for developing applications aimed at image processing, visual data analysis, data structure, camera calibration, object recognition, motion analysis (tracking) and image labeling (BRADSKI; KAEHLER, 2008).

The use of the OpenCV library allows computer vision to be implemented within the system, simplifying and facilitating image processing and detection through functions such as grayscale, which transforms the colors of the image to grayscale, preserving the light intensity data and excluding the color values.

## You Only Look Once (YOLO)

YOLO is a model for detecting and classifying objects and segmenting images in real time. Its algorithm is characterized by its fast calculation speed and the small size of the detection model (JIANG et al., 2022). By sectioning the image into grids, YOLO simultaneously calculates the predictions of the bounding boxes and classes, defining the probability of there being an object in each bounding box (HUANG; PEDOEEM; CHEN, 2018). In addition, YOLO technology has a rapid ability to identify and track vehicles, bicycles, people and other obstacles found in metropolises, exacerbating security (TERVEN; CÓRDOVA-ESPARZA; ROMERO-GONZÁLEZ, 2023).

Source: Authors (2024).

Figure 1 - Yolo working example.

The YOLO technology, as shown in Figure 1, was used within the project for the detection and recognition of potentially dangerous objects pertinent to the Second Vision context and, due to its pre-trained models with an expressive processing speed, it was used.

## Pytesseract

Tesseract is an open source optical character recognition engine that is the simplest way to convert typed, handwritten or printed text contained in images into digitally manipulable text (AKHIL, 2016). Furthermore, this mechanism is like the combination of eye and brain, as the eye receives the image as input and the brain processes and interprets the nuances of this image, but Tesseract only deals with text processing (PATEL, 2012).

Python-Tesseract (Pytesseract) is an adaptation of the open source Tesseract library to work as a feature in the Python programming language, capable of recognizing and extracting text from scanned formats using this optical procedure of Tesseract (JAYOMA, 2020).

Figure 2 - Pytesseract working example

Source: Authors (2024).

Pytesseract, as shown in figure 2, is used to detect static texts, configured to recognize only texts from the Latin alphabet and to improve the accuracy of the words detected, thus helping the visually impaired to be more autonomous in urban environments.

## React Native

React, Facebook's JavaScript library, is based on components and is responsible for structuring and pre-configuring visual elements. Its innovation lies in not having to perform micro-operations to modify the content of a component, making User Interface (UI) development possible (BODUCH, 2017).

In addition, React Native is a framework for producing multiplatform mobile applications that look “native”, structured in React (EISENMAN, 2015). In addition, React Native is developed solely in Javascript or Typescript, but the interface components are generated natively, allowing the Application Programming Interfaces (APIs) of the platforms of the respective devices to be used, thus creating a layer of abstraction (META, 2024).

The React Native framework is used throughout the development of the mobile application, responsible for reproducing the information detected from the GATT server. This technology was used to reach a wider audience due to its ability to generate cross-platform mobile applications.

# Methods

The research method used is exploratory research, which is a type of scientific investigation that aims to examine a problem situation in order to enable an understanding of the problem, as well as to formulate initial perceptions and hypotheses (LAKATOS; MARCONI, 1999). In this way, exploratory research, using a bibliographic survey, qualitative and quantitative approaches, allows us to explore the particularities experienced by the visually impaired, resulting in a greater general understanding of the needs and difficulties they face.

As for the bibliographic survey, it corresponds to scientific productions and data collection based mainly on scientific articles, books, magazines, encyclopedias and critical essays; as for the qualitative approach, it refers to the interpretation of the information of the object of study, as well as understanding their opinions and attitudes, structuring itself on the existence or non-existence of attributes; the quantitative approach, in turn, refers to statistics, numbers and percentages, which allow researchers to quantify and evaluate the object of study (LAKATOS; MARCONI, 1999).

In the Second Vision project, the bibliographic survey allows for the analysis and review of productions related to the problems and technologies used in the construction of the project, enabling researchers to become proficient in the topics in question. With regard to the qualitative approach, it focuses on the specific perception of the needs, experiences and challenges of the visually impaired in urban environments, in line with understanding the main obstacles they face in metropolises and how a system could help improve their autonomy. As far as the quantitative approach is concerned, it is intended to measure the main objects found in metropolises that can compromise the autonomy of the visually impaired, analogously to considering statistical analyses of the effectiveness of recognizing potentially dangerous objects and reading the texts on static signs, as well as the system's performance when carrying out these functions. As a result, it is stated that by integrating a bibliographic survey with a qualitative and quantitative approach, a greater understanding of the problems and needs faced by the visually impaired can be achieved, as well as the development of solutions that can contribute to user-centered accessibility.

# Results and Discussions

The project achieves a capacity to promote inclusion and accessibility in terms of urban autonomy for the visually impaired, characterizing an innovative idea with its due pertinence when pragmatizing an equity, it is manifested that the system manages to satisfactorily identify the objects, despite the limitations in their variety, in addition to the wearable attribute making it simple to use, the application shaped in User Experience (UX) added to its structure with support for screen reading technologies that guides and allows the visually impaired to move the cell phone, demonstrates its aptitude in dealing with the particularities of the persona objectified.

In addition, as an artificial intelligence (AI) model is used that has parameters limited to the needs of YOLO, not adapting to the full potential that Second Vision can achieve, this limits the ability to effectively guide the visually impaired geographically, so there are margins for a significant improvement in the nature of the system, to make it more capable in this objective through, for example, an AI model of its own and trained specifically for the purpose presented.

It can therefore be seen that the wearable attribute seems vulnerable to the climate, specifically the rain factor, so a case or structure that takes these externalities into account could improve the practicality of the system.

Finally, as time goes by, technology is exponentially evolving and becoming part of people's lives, and this integration of technology for the purpose of providing social justice is extremely pleasing. By helping with this, the system contributes to a better world that does not make people with disabilities invisible; on the contrary, it reinforces that they are also entitled to rights and deserve dignity.

## Second Vision

Similarly to the current chapter, images and detailed descriptions that elucidate the final result of the project are shown below. This information is intended to provide a clear and comprehensive understanding of the results achieved with the Second Vision device and mobile application.

Figure 3 shows the Second Vision device, made up of the Raspberry PI 5 microcomputer connected to the Uninterruptible Power Supply (UPS), which acts as the power supply for the Raspberry PI 5 and the voltage, current and power regulator for the battery; In addition, the case created through modeling and a 3D printer to attach the Second Vision - IoT device together with the vest to make the system wearable; and the camera to detect and recognize potentially dangerous objects and read text on Pessoa posando para foto

Descrição gerada automaticamentestatic signs.

Figure 3 - Image of the Second Vision device

Source: Authors (2024).

Figure 4 shows the finished wearable prototype of the Second Vision system, encompassed by a stylized case adapted to protect the physical component and coupled to a wearable support that can be adjusted to different bodies, so that it brings fluidity and comfort to its use by the visually impaired, allowing them full autonomous locomotion as users of the system.

Source: Authors (2024).

Figure 4 - Second Vision Wearable System

Figure 5 shows the main screen of the mobile application, displayed after the visually impaired person has paired their smartphone with the Raspberry PI via Bluetooth. This screen provides information about the system, such as the amount of battery remaining, the system status, the speech interval time of the detections and the detection mode currently running, followed by a brief summary of how the chosen mode operates.

Source: Authors (2024).

Figure 5 - Main Application Screen

# Final Considerations

The Second Vision project is a very important step towards increasing the safety and freedom of the visually impaired, as we have managed to create a tool that allows users greater autonomy and facilitates the process of getting around in metropolises with the help of an object recognition and static text reading system.

With the help of data to implement a more assertive system, we were able to demonstrate a working model using the methods presented that show the possibility of innovative approaches to a social problem.

Although the project managed to achieve its initial objectives, there is still room for expansion and improvement. The greater viability of the system could be improved by including new features, such as recognizing more complex situations and training more objects for detection using deep learning.

In short, Second Vision is not just a technological solution; it is a step towards a more accessible world where everyone can enjoy greater freedom and safety. We hope that this work will inspire new projects and research in this area, broadening the discussion on inclusion and innovation.

**References**

BRADSKI, Gary; KAEHLER, Adrian. Learning OpenCV: Computer vision with the OpenCV library. " O'Reilly Media, Inc.", 2008.

BODUCH, Adam. React and react native. Packt Publishing Ltd, 2017.

FÖRSTNER, Wolfgang; WROBEL, Bernhard P. Photogrammetric computer vision. Springer International Publishing Switzerland, 2016.

HUANG, Rachel; PEDOEEM, Jonathan; CHEN, Cuixian. YOLO-LITE: a real-time object detection algorithm optimized for non-GPU computers. In: 2018 IEEE international conference on big data (big data). IEEE, 2018. p. 2503-2510.

JIANG, Peiyuan et al. A Review of Yolo algorithm developments. Procedia computer science, v. 199, p. 1066-1073, 2022.

KHAN, Salman et al. A guide to convolutional neural networks for computer vision. 2018.

KLETTE, Reinhard. Concise computer vision. London: Springer, 2014.

LUTZ, Mark. Programming python. " O'Reilly Media, Inc.", 2001.

META - React Native, 2024. Disponível em: https://reactnative.dev/docs/environment-setup. Acesso em: 11 setembro 2024

EISENMAN, Bonnie. Learning react native: Building native mobile apps with JavaScript. " O'Reilly Media, Inc.", 2015.

AKHIL, S. An overview of tesseract OCR engine. In: A seminar report. Department of Computer Science and Engineering National Institute of Technology, Calicut Monsoon. 2016.

PATEL, Chirag; PATEL, Atul; PATEL, Dharmendra. Optical character recognition by open source OCR tool tesseract: A case study. International journal of computer applications, v. 55, n. 10, p. 50-56, 2012.

JAYOMA, Jaymer M.; MOYON, Elbert S.; MORALES, Edsel Matt O. OCR based document archiving and indexing using PyTesseract: A record management system for dswd caraga, Philippines. In: 2020 IEEE 12th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM). IEEE, 2020. p. 1-6.

OLIVEIRA; Cláudio Luis Vieira; NABARRO, Cristina Becker Matos; ZANETTI, Humberto Augusto Piovesana. Raspberry Pi Descomplicado. 1. ed. São Paulo: Saraiva Educação S.A., 2018. 224 p.

MARENGONI, Maurício; STRINGHINI, Stringhini. Tutorial: Introdução à visão computacional usando opencv. Revista de Informática Teórica e Aplicada, v. 16, n. 1, p. 125-160, 2009.

MOLINO, Monica; CORTESE, Claudio G.; GHISLIERI, Chiara. Technology acceptance and leadership 4.0: A quali-quantitative study. International Journal of Environmental Research and Public Health, v. 18, n. 20, p. 10845, 2021.

TERVEN, Juan; CÓRDOVA-ESPARZA, Diana-Margarita; ROMERO-GONZÁLEZ, Julio-Alejandro. A comprehensive review of yolo architectures in computer vision: From yolov1 to yolov8 and yolo-nas. Machine Learning and Knowledge Extraction, v. 5, n. 4, p. 1680-1716, 2023.

UNITED NATIONS. The 17 Sustainable Development Goals. Disponível em: <https://sdgs.un.org/goals>. Acesso em: 11 de setembro de 2024.

VAN ROSSUM, Guido et al. Python programming language. In: USENIX annual technical conference. 2007. p. 1-36.

STOCKEMER, Daniel; STOCKEMER, Glaeser; GLAESER, J. Quantitative methods for the social sciences. Cham, Switzerland: Springer International Publishing, 2019.

WORLD HEALTH ORGANIZATION. Blindness and vision impairment. Disponível em: <https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment>. Acesso em: 11 de setembro de 2024.

RASPBERRY PI, Raspberry Pi Hardware. Disponível em: <https://www.raspberrypi.com/documentation/computers/raspberry-pi.html>. Acesso em: 11 de setembro de 2024.

“The author(s) of the work declare(s) that no Artificial Intelligence (AI) tool/service was used during the preparation of the manuscript, and that all the text was produced and is the responsibility of the authors.”

1. Etec da Zona Leste [↑](#footnote-ref-1)