International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 09 | September - 2023 SJIF Rating: 8.176 ISSN: 2582-3930

VIRTUAL EYE KIT FOR BLIND PEOPLE

Shreya Badarkhe[1], Sakshi Gangdhar[2], Neha Ghogre[3] Anuradha Kasangottuwar[3]

[1]*Department of E&TC PES Modern College of Engineering Pune, India*
[2]*Department of E&TC PES Modern College of Engineering Pune, India*
[3]*Department of E&TC PES Modern College of Engineering Pune, India*

---***---

Abstract - Sight is a precious ability that grants us the privilege of perceiving the world around us. It empowers individuals to grasp their surroundings in all its splendor. However, for those who lack this gift, life can be a constant struggle. In this paper, we introduce a straightforward, cost-effective, and user-friendly virtual eye system designed to enhance the mobility and independence of individuals who are visually impaired or blind within a specific area. This innovative project aims to enable blind individuals to navigate and comprehend their environment using auditory cues. It revolves around a visual-centric concept, where essential components like a camera, a Raspberry Pi (a small computer), and earphones are integrated into a single unit. Furthermore, it harnesses the power of the internet for seamless interaction. To outline the process, the system takes image or video input (consisting of multiple frames), which is captured and analyzed by the connected camera interfaced with the Raspberry Pi or Internet of Things (IOT) technology. This analysis helps detect objects, and the corresponding audio information is then relayed to the individual through the earphones. The primary objective of this system is to elevate the quality of life for the visually impaired by leveraging cutting-edge technology. By offering a comprehensive solution, this setup empowers those with visual limitations to navigate their surroundings with increased confidence and reduced limitations.

***Key Words*:** Raspberry Pi, camera ,ultrasonic sensor

1.INTRODUCTION

On a global scale, it's been approximated that around 2.2 billion people experience some level of vision impairment or blindness. Shockingly, a substantial portion of this group, roughly 1 billion individuals, could have either avoided or still need intervention for their vision issues. With this understanding, we're embarking on the development of an application that amalgamates object recognition techniques from computer vision and image processing, along with auditory support. The ultimate aim of this application is to empower visually impaired individuals to independently identify obstacles in their path, negating the necessity for external assistance. Our brains possess the remarkable ability to swiftly recognize objects within images. However, training machines to achieve this level of object identification has traditionally demanded substantial time and an abundance of training data. Yet, thanks to the recent advancements in hardware capabilities and deep learning, the field of computer vision has greatly simplified and become more intuitive. We're ceaselessly exploring avenues to establish a "detection" or "recognition" system that rivals the prowess of the human visual system. One particularly captivating avenue that has gained significant attention is weakly supervised object localization. This approach seeks to pinpoint common objects within images using annotations that indicate the presence or absence of the target objects. In the context of our project, our emphasis lies in simultaneously uncovering and pinpointing prevalent objects in real-world images. This pursuit aligns with the same output objectives as weakly supervised object localization methodologies.

2. Literature Review

An electromagnetic sensor has been developed to enhance the autonomous mobility of individuals who are visually impaired or blind. Traditionally, those affected by visual impairments rely on white canes for navigation. However, the integration of a microwave radar onto the conventional white cane expands its functionality, alerting the user to obstacles in a wider and safer range. This advancement in Electronic Travel Aids introduces improved efficiency, resilience to interference, and a more compact design. The study also highlights recent progress, particularly in scaling down the circuit board and antennas, as part of ongoing developments in this field [1].

An assistive system has been developed to aid individuals with visual impairments in reading various texts encountered in their daily lives. The process involves several preprocessing steps, including converting the text to grayscale and then to binary form. The application of the OTSU algorithm facilitates this conversion from grayscale to binary. Subsequently, relevant sections of the processed image are extracted, followed by their interpretation through optical character recognition (OCR) software. In this context, the primary OCR method employed is MODI. The extracted text, which may encompass diverse fonts and sizes, is then independently recognized and subsequently synthesized into audible output using Text-to-Speech technology. This is achieved through the utilization of SAPI libraries. The system essentially aims to provide visually impaired individuals with the means to access and comprehend various textual materials in their environment [2].

Optical Character Recognition (OCR) refers to the process of identifying printed characters using devices that convert light into electrical signals (photoelectric devices) and computer software. This technology transforms images containing written or printed text into machine-readable text, which can then be converted into audible output. OCR finds application in various tasks such as natural language processing, machine translation, text-to-speech conversion, data entry, and text analysis.

In the context of this project, OCR is employed for character recognition within text documents. The Raspberry Pi device is used for character identification through the utilization of the Tesseract algorithm and Python programming. The resulting characters are then converted into audio output. Furthermore, OCR is extended to include pattern recognition for the purpose of Document Image Analysis (DIA). This involves utilizing structured information in a grid format within the design and development of virtual digital libraries. The project notably emphasizes the use of Python as the primary programming language [3].

The capacity and ability of vision to person is a significant factor of our life. But some individual whose unfit and have absence of vision since they are outwardly weakened. One brilliant framework which is only keen devise which become supportive for that outwardly weakened extrovert can be recognize deterrent with assistance of its daze stick. Blind individual get feel for when cell phone get vibration alarm or give some oral voice message to that individual [4].

Sight serves as a fundamental sense, with images playing a crucial role in how humans perceive their surroundings. Digital image processing, a specialized field involving the use of computers, plays a significant role in analyzing digital images. However, identifying objects within images remains challenging, especially for individuals with visual impairments. While numerous applications exist to aid in this task, there are still limitations that necessitate further enhancements. This study provides an overview and analysis of various techniques used in object identification tasks, focusing on advancements in this field. These advancements offer the potential to significantly improve the capability for visually impaired individuals to identify objects. This presents an opportunity for a meaningful substitute for visual perception among those with visual impairments [5].

3. BLOCK DIAGRAM AND DESCRIPTION

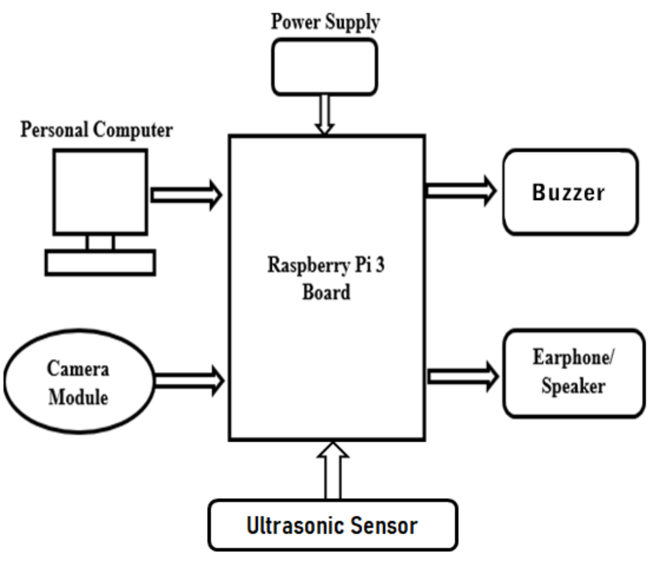


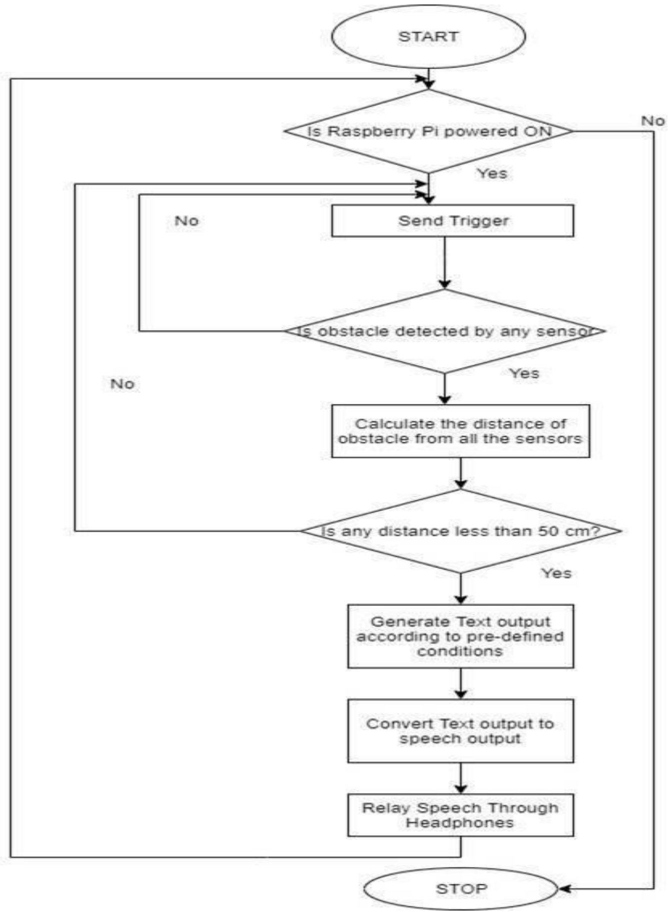
Fig -1: Block diagram for Virtual Eye Kit for Blind People

The project operates by taking video or image input. This input is subjected to analysis through a camera connected to the Raspberry Pi Board. The camera captures the image, and subsequently, object detection is performed on this image. The result is conveyed through audio output. To facilitate the process, the input video is broken down into multiple frames. Key frames are then chosen and transformed into grayscale images. Preprocessing steps are applied to recognize boundaries in the image, independent of the gray levels present. Detecting objects or text in an image demands strategic placement within both lighter and darker regions of the image. To achieve this, an advanced and efficient algorithm is employed. This algorithm ensures optimal object/text detection. Additionally, a feature extraction process is applied to proficiently classify or identify objects/text in the image. Once the textual output is generated, it undergoes conversion into speech output. This is facilitated by leveraging Python libraries such as pyttsx3 and gtts, or by utilizing APIs such as Microsoft API, Google API, or even the Google Assistant SDK. The speech output, serving as audio information, is then relayed to the user through a pair of earphones.

4. SOFTWARE DESIGN

The project's initial input comprises videos or images. Using a camera connected to the Raspberry Pi Board, the system captures and evaluates these images. The primary goal is to identify objects within the images and provide corresponding audio output. For efficient processing, the video input is divided into multiple frames. From these frames, key frames are chosen for further analysis. These selected frames are then transformed into grayscale images. Through pre-processing techniques, boundaries within the images are highlighted, irrespective of the varying shades of gray present. Detecting objects or text within images demands strategic placement within both brighter and darker regions of the image. In essence, the project revolves around capturing, processing, and interpreting visual information to enable audio-based recognition of objects and text. You Only Look Once (YOLO) represents a Convolutional Neural Network (CNN) designed to execute real-time object detection tasks. YOLO's remarkable advantage lies in its speed, which outpaces other network architectures while maintaining accuracy. This algorithm can be implemented using deep learning libraries like Keras or OpenCV. CNNs function as systems capable of processing input images structured as data arrays. They then discern patterns existing between these images. In the case of YOLO and similar convolutional neural network algorithms, these systems evaluate regions within an image based on their similarity to predefined classes. Regions with higher scores are identified as positive detections for the most closely related class. Once the textual output is generated, it undergoes transformation into speech output. This conversion process employs Python libraries such as pyttsx3 and gtts, or it may involve various APIs like Microsoft API, Google API, or even the Google Assistant SDK for converting text output into audible speech. For instance, gTTS, a Python library and command-line tool, interfaces with Google Translate's text-to-speech API. It allows the conversion of text from documents into spoken MP3 files. This library offers flexible preprocessing and tokenization capabilities and supports automatic retrieval of supported languages for text-to-speech conversion. The speech output is sent as an audio to earphones which are worn by the user. Bluetooth earphones or wired earphones both can be used for getting this output or audio to be listened by the user.

1.FLOWCHART



5. RESULT AND DISCUSSIONS

The above expected result is achieved by object detection. Object detection is a computer vision technique for locating instances of objects in images or videos. Accuracy of the detected object is also mentioned in the bounding box. Voice output works through TTS (text to speech). The proposed system detects an object around them and sends feedback in the form of speech via earphones.

6. CONCLUSION AND FUTURE SCOPE

The desired outcome mentioned earlier is accomplished through a process known as object detection. This technique, rooted in computer vision, involves identifying specific objects within images or videos. Notably, the accuracy of the detected object is indicated through a bounding box. For auditory communication, a Text-to-Speech (TTS) mechanism is employed. In practice, the system identifies objects present in the user's vicinity and conveys relevant information audibly through earphones.

This endeavor is dedicated to individuals with visual impairments, granting them access to the vibrant and captivating world they are unable to see. Our initiative is aimed at enhancing their quality of life. Through this project, we enable users to identify objects in their immediate surroundings, thus offering a means to comprehend their environment more fully. Our team is committed to ongoing research and development to refine this solution. By continually enriching the algorithm with additional data, we anticipate enhancing both the accuracy and the algorithm's capability to identify a wider range of objects. This iterative improvement process is expected to bolster the model's accuracy and expand its object recognition capabilities. The benefits of this object recognition system extend beyond aiding the visually impaired. It has versatile applications in domains such as surveillance systems, facial recognition, anomaly detection, and character recognition. As we move forward, our innovation has the potential to influence diverse fields.

REFERENCES

1. Cardillo, E.; Di Mattia, V.; Manfredi, G.; Russo, P.; De Leo, A.; Caddemi, A.; Cerri, G. An Electromagnetic Sensor Prototype to Assist Visually Impaired and Blind People in Autonomous Walking. IEEE Sens. J. 2018, 18, 2568–25761. DOI= https://doi.org/10.1109/JSEN.2018.2795046.
2. Mrs.Shilpa Reddy K , Mounika S.K,Pooja K, Sahana N "Text to Speech for the Visually Impaired" International Research Journal of Computer Science (IRJCS) ,Issue 05, Volume 4 -May 2017
3. Mallapa D.Gurav1, Shruti S. Salimath, Shruti B. Hatti, Vijayalaxmi I. Byakod, ShivaleelaKanade "A Reading aid for the Blind People using OCR and Open CV" Journal 2017.
4. SuchitaWankhade,MrunaliBichukale,Shruti Desai,Shraddha Kamthe, Archana Borate "Smart Stick for Blind People with Live Video Feed",Journals 2017.
5. Miss. Kirti P. Bhure, Mrs. J. D. Dhande "Object Detection Methodologies for Blind People" Journal, 2017.
6. A. Aladrn, G. Lpez-Nicols, L. Puig, and J. J. Guerrero, "Navigation assistance for the visually impaired using RGB-D sensor with range expansion," IEEE Syst. J., vol. 10, no. 3, pp. 922–932, Sep. 2016.
7. S. Bhatlawande, M. Mahadevappa, J. Mukherjee, M. Biswas, D. Das, and S. Gupta, "Design, development, and clinical evaluation of the electronic mobility cane for vision rehabilitation," IEEE Trans. Neural Syst. Rehabil.Eng., vol. 22, no. 6, pp. 1148–1159, 2014.
8. B. B. Blasch, R. G. Long, and N. Griffin-Shirley, "National evaluation of electronic travel aids for blind and visually impaired individuals: Implications for design," in Proc. RESNA 12th Annu. Conf., 1989, pp. 133–134.
9. H. Schafer, A. Hach, M. Proetzsch, and K. Berns, "3D obstacle detection and avoidance in vegetated off-road terrain," in IEEE International Conference on Robotics and Automation, 2008, pp. 923–928.
10. H. Takizawa, S. Yamaguchi, M. Aoyagi, N. Ezaki, and S. Mizuno, "Kinect cane: Object recognition aids for the visually

impaired," in The 6th International Conference on Human System Interaction (HSI), June 2013, pp. 473–478.

11. Z. Wang, H. Liu, X. Wang, and Y. Qian, "Segment and label indoor scene based on RGB-D for the visually impaired," in MultiMediaModeling, ser. Lecture Notes in Computer Science, C. Gurrin, F. Hopfgartner, W. Hurst, H. Johansen, H. Lee, and N. O'Connor, Eds., vol. 8325. Springer International Publishing, 2014, pp. 449–460.