



e-ISSN:2582-7219



INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH IN SCIENCE, ENGINEERING AND TECHNOLOGY

Volume 7, Issue 7, July 2024



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 7.521



6381 907 438



6381 907 438



ijmrset@gmail.com



www.ijmrset.com



Artificial Intelligence Applications for Visually Impaired People

Dr. Vishwanath V Murthy, Dr. Rajani Narayan, Apeksha BK, Inchara S.M

Associate Professor, Department of MCA, RNSIT, Bangalore, India

Associate Professor, Department of MCA, RNSIT, Bangalore, India

Department of MCA, RNSIT, Bangalore, India

Department of MCA, RNSIT, Bangalore, India

ABSTRACT: The paper examines how Artificial Intelligence affects the daily life of blind people. They experience a great deal of hardship that they are visually impaired. When walking alone around the city, many worry about their safety. The AI's main objective is to provide blind people with low-cost navigational assistance that mimics artificial vision by alerting them of the objects in their environment. An ultrasonic sensor is used to estimate the distance between objects for a blind person who is directed by speech and vibrations that they can hear and feel. The program can analyze spoken language, recognize written content in documents, and use voice instructions to recognize objects in the actual world. For blind persons, it could be a vital tool for motivation and communication.

KEY WORDS: visually impaired, Raspberry Pi, image recognition, object detection, Speech recognition, AI(Artificial Intelligence), Text to Speech conversion.

I. INTRODUCTION

Vision could be a complicated handle that requires various components of the human eye and brain working together. Vision upgrades pictures, counting the subtle elements of movement and colour, on a time so quick that a break within the activity is nearly never perceived[1]. The every day challenges that dazzle individuals confront extend from reading a book to strolling along the road. A healing condition called visual impedance or visual impedance may influence a person's capacity to see, particularly in their superior eye. If left untreated without corrective eye-wear, assistive devices, and medical treatment, it can make daily activities difficult[2]. Many blind people are unemployed as they can perform limited work hence resulting in lower employment rates. They usually depend on their families both financially and for mobility. Their mobility limits them from taking part in social interactions and also other activities.

Human insights is the ponder and creation of computer programs that can observe aim and take activity. Most individuals utilize their bodies, voice, and facial expressions to communicate non-verbally and express their sentiments. There are too a part of feelings present[3].

Through the creation of camera items that coordinated machine learning advancement with other commercial commodities like OCR, advanced progressions in computerized recognition, advanced cameras, and versatile computers precisely offer assistance people [4]. Individuals who are outwardly impeded can utilize voice commands. The current setup with counterfeit frameworks permits confront location of the pictures you have got clicked on in expansion to supporting the outwardly impeded with voice enlightening [5]. People are using many technologies like image recognition, speech recognition, text-to-speech conversions, obstacle detection etc along with AI to create applications, assistants, canes, glasses and other tools for helping the blind people.

II. LITERATURE SURVEY

This exposition investigates how counterfeit insights can improve the day by day lives of people who are dazzle. Once you run into unanticipated conditions that you just may not have been mindful of, you endure enormously. When voyaging alone over the city, individuals are concerned for their security [3].This consider centers on a few AI procedures created basically for applications that help dazzle individuals. Enas Elbarbary, Abdelhalim Zekry and

Hussam Elbehiery[1].They proposed an “Blind Assistant” application whose main objective is to help the blind people to recognize their surrounding objects, alerted by non-safe objects like moving car, trains etc or stationary ones like trees, walls etc.P. Vijayakumar, Mr. T.Yuvaraj, C. A Sathiya Moorthy, Makarand Upadhyaya, Pankaj Dadheech, Pravin R. Kshir- sagar,[6]. In this paper they have proposed a software which alerts the user about their surrounding objects. A blind person is guided by speech and vibrations that they can hear and feel using an ultrasonic sensor to measure the distance between things. The software can analyse text, recognise written content in documents, and identify objects in the real world using voice commands.

Microsoft has developed an artificial intelligence (AI) soft- ware called ”Seeing AI” that uses the device’s camera to recognise people and objects before providing those with visual impairments with an audible description of those objects[7]. Additionally, ”People Lens” is an AI system that gives blind or limited vision individuals additional tools to compre- hend and interact with their immediate social surroundings[8]. The Raspberry Pi prototype that Ms. Rupali, D. Dharmale, and Dr. P.V. Ingole [9] suggested extracts text from pictures taken with a webcam. After being converted to gray-scale, the coloured images are processed using Open CV and the OTSU algorithm. Then, the vertical and horizontal ratios of the transformed images are modified. Before using OTSU’s adaptive threshold method, photos are adjusted. After threshold, bounding boxes are constructed around objects and text in the photos using Open CV algorithms to create the contours. These bounding boxes are used to extract each character from the image and feed them into the OCR engine for text recognition.

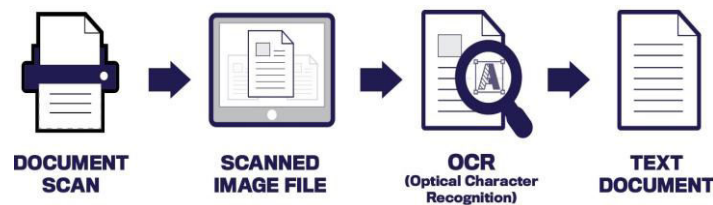


Figure. 1. Extraction of text from image using OCR.

Xiaofang Jin [10], According to a comparison of four neural networks for facial emotion detection, accuracy may not be improved by adding more layers. However, convolution neural networks’ performance can be enhanced by boosting the number of filters. A 1 h 1 convolution layer can deepen networks, boost expression, and increase non-linear representation. This study’s drawbacks include ignoring concealed emotions and having low accuracy, which will be addressed in further studies.

”Vision-based text recognition using Raspberry PI,” Na- garaja L. [12]. The study offers a camera-based framework for people with visual impairments to read text with assistance. The method isolates the target from distracting backgrounds by defining a Region of Interest (ROI) in the camera image using motion-based approaches. The moving object region is extracted using a Gaussian-based background removal method. To extract text details from the ROI, text localization and recognition are used, and the text portions are enlarged. Using a convolutional neural network (CNN) with fine-tuned filters and a 1x1 convolutional layer to boost the model’s expressive capability, the performance of the model is enhanced layer by layer. The study’s shortcomings, such as its low accuracy and disregard for emotions concealed in the face, might be addressed in future studies. Self-optical character identification software is employed for text binarization and recognition in localised text regions, and a unique gradient feature and edge pixel distribution strategy is proposed.

The “People Lens” is an open-ended AI system that gives blind or low-vision individuals additional tools to comprehend and interact with their local social settings. Blind students have recently utilised it in educational settings to support their abilities to actively communicate with classmates. The “People Lens” is an investigation into how human-AI interaction might be designed to go beyond discrete task support and offer a steady stream of dynamic data[8]. It was influenced by ethnographic study with Paralympic competitors and spectators, which documented the variety of sense-making techniques used by people with limited vision to connect with others and go around. The “People Lens” was designed to enable individuals to see details in their environment more clearly.

III. METHODOLOGY OF THE APPLICATIONS

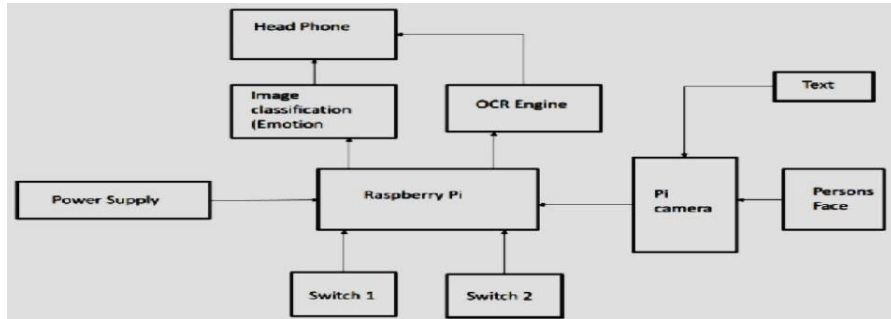


Figure. 2. System Architecture of the Application [6]

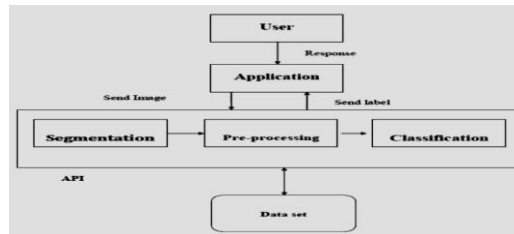


Figure. 3. Methodology of Blind Assistance Device[11]

Every application or assistant devices using AI, mainly consists of three major components namely, “Image Recognition”, “Speech Recognition” and “Text to Speech conversion”. People combine these technologies along with additional components like Raspberry Pi, EI (Emotional intelligence) etc to create the applications. Even the major companies like Microsoft, are using these components and creating assistants for the visually impaired.

A. Image Recognition

The process of identifying and detecting objects or features inside a picture is known as image recognition. It is a sort of computer vision technology that analyses and interprets digital images using a variety of algorithms and methodologies. Image recognition is used to recognize faces and identify people as one of the most popular applications in the security industry. Template matching, feature detection, and deep learning are just a few of the methods applied for image recognition.

Dealing with differences in illumination, angle, and scale is one of the difficulties in picture identification. Techniques like data augmentation and image normalization can be used to solve this.



B. Speech Recognition

Speech recognition, or voice recognition, is a technology that allows computers or devices to understand spoken language. It involves capturing audio input, converting it into digital data, and processing it using algorithms and machine learning. Speech recognition has applications in communication, automation, accessibility, and transcription services. Despite challenges, advancements in machine learning are improving accuracy. Speech recognition has transformed



human-machine interaction and is expected to continue evolving in the future.

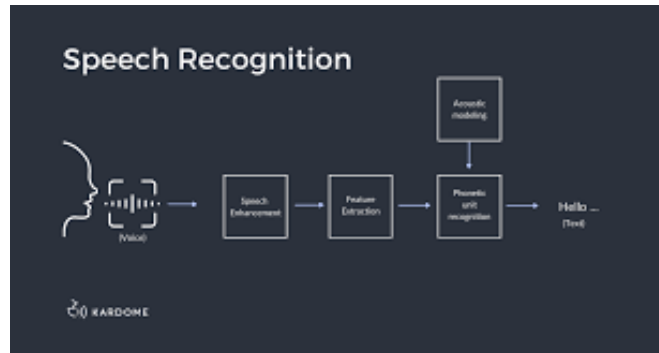


Figure. 3. Speech Recognition

C. Text-to-speech conversion

Text-to-speech (TTS) recognition is a technology that converts written text into spoken speech. It has applications in accessibility, communication, education, and entertainment. TTS systems use algorithms and models to analyze text and generate speech output. Advancements in natural language processing, deep learning, and neural TTS models have improved the accuracy and naturalness of TTS voices. Challenges remain in capturing human speech nuances, handling different languages and accents, and addressing ethical considerations. TTS technology has the potential to make written content more accessible and user-friendly.

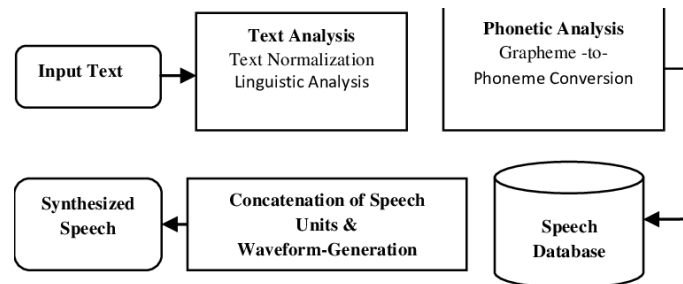


Figure. 4. Text-to-speech conversion

D EI(Emotional Intelligence)

Emotional intelligence (EI) is the capacity to understand, manage, and regulate one’s own emotions as well as those of others. It includes self-awareness, social awareness, emotional regulation, and interpersonal skills. EI is essential for personal and social well-being, and it has implications in many areas of life, including relationships and professional success. Emotional intelligence development entails self-reflection, mindfulness, seeking feedback, and learning effective communication and conflict resolution skills. Generally talking, three skills—emotional sensitivity, affectability, the capacity to channel sentiments and apply them to exercises like reasoning and problem-solving, and the capacity to oversee emotions—include both self-control and the capacity to calm others down[3]. Individuals with high enthusiastic insights are self-driven, persevering, and spurred by individual objective instead of outside powers like finance or reputation[6].



Figure. 5. Components of EI

E. Artificial Intelligence

Artificial intelligence (AI) alludes to the reenactment of human intelligence in innovative technological systems created to carry out assignments that customarily call for human intelligence. There are two essential sorts of AI: general or strong AI, which has human-like capabilities, and narrow or weak AI, which is confined to particular errands. Artificial intelligence (AI) frameworks depend on strategies like machine learning (ML), deep learning, and natural language processing (NLP). AI has significant implications across various industries, including healthcare, finance, transportation, manufacturing, and entertainment. Ethical considerations, such as bias, privacy, accountability, and societal impact, are important in AI development. AI is expected to continue to advance and revolutionize many aspects of society.

F. Natural Language Processing (NLP)

The field of computer science known as "natural language processing" (NLP) is more particularly from the AI field that is concerned with providing computers the capacity to comprehend written and spoken words in a manner similar to that of humans. It is the combination of Computer Science, AI and Human Intelligence. The goal of NLP is to make it possible for computers to comprehend, analyse, and produce human language. Text encoding, POS tagging, NER, sentiment analysis, machine translation, and text synthesis are among the tasks involved. NLP has to deal with issues like language uncertainty and ethical dilemmas. NLP has applicability across many industries and is predicted to develop further, revolutionizing human-computer interaction.

G. Raspberry Pi

Raspberry Pi is a small, affordable, and versatile single-board computer that was initially developed for educational purposes. Additionally, it provides a collection of GP-I/O (general purpose input/output) pins that enable you to explore the Internet of Things and manage electronic devices for IoT. It offers a wide range of applications, including home automation, media centers, robotics, and more. It operates on a Linux-based operating system and supports numerous programming languages. All skill levels of users may use it because to its compact size, low power consumption, and active community support. Also, it has a sizable user and developer community that collaborate and exchange materials, and it has been incorporated into many curricula in schools. Ultimately, Raspberry Pi has developed into a potent tool that has transformed technology and computing.

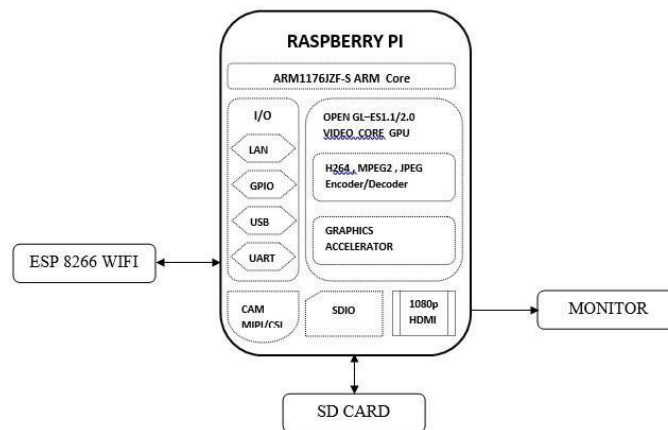


Figure. 6. Architecture of Raspberry Pi

IV. ADVANTAGES AND DISADVANTAGES

Even though high-end technologies are used to build these applications there are few pros and cons to them.

A. Advantages

- The major advantage of these applications and tools is that they help the visually impaired people to walk safely, know about their surroundings and also communicate better in the society.
- The image recognition helps the user to identify the objects and if there are any non-safe objects it gives a warning and alerts them too.
- The speech recognition using neural networks will be used for emotional intelligence and helps the user understand what the other person is feeling, through their speech, as they can't look at them.



- The text-to-speech conversion helps the user to read books, banners, signs etc. It converts the text to speech and the user listens to it.

B. Disadvantages

- The major disadvantage of this is that the technologies are not much reliable. No matter how advanced the technology is it can only mimic the human brain but cannot match it.
- The Image recognition is not accurate i.e., the distant objects or the same object from different angles can be misinterpreted causing improper information. It is dangerous if the image recognition is not proper while the user is walking.
- The speech recognition might not be able to interpret the speech with accents or whispers and even those who talk very fast.
- The text-to-speech conversion may not be able to scan the handwritten texts and also the converted speech may be not clear which may misguide the user.

V. CONCLUSION

In conclusion, the development of artificial intelligence (AI) technology is transforming the lives of individuals who require care, such as the elderly, visually impaired individuals, and those with disorders like autism. AI has the potential to enable machines to understand human senses, exhibit empathy, and assist in tasks such as counseling, object detection, and communication. The advancements in AI hold promise in positively impacting the lives of individuals with diverse needs in the future. The paper concludes about the AI-based systems that benefits visually impaired individuals by helping them understand their surroundings. The systems provide auditory cues about what is in front of them, helping them avoid collisions and "see" their environment. It also assists blind individuals in choosing the correct path for movement and navigation, and guides deaf individuals along unobstructed paths. The systems are capable of detecting various objects and can enhance interaction and daily activities for individuals with visual impairments.

REFERENCES

1. Enas Elbarbary, Abdelhalim Zekry and Hussam Elbehiery, "The Blind Assistant, AI helping visually impaired people," LAP LAMBERT Academic Publishing, 2016.
2. "Vision Impairment and Blindness — Examination-Based Studies — Information on Data Sources — Vision and Eye Health Surveillance System — Vision Health Initiative (VHI) — CDC". www.cdc.gov. 2021- 08-10.
3. Padma Shneha1 and Prathyusha Reddy, International Journal of Latest Trends in Engineering and Technology 031-036 (2018).
4. Raju Shrestha & FrodeEikaSandnes ,IETE Technical Review(2020).
5. Pruthvi S, and PushyapSurajNihal, International Journal of Engineering and Advanced Technology (IJEAT) 8 (5S), (2019).
6. P. Vijayakumar, Mr. T.Yuvaraj, C. A Sathiya Moorthy, Makarand Upadhyaya, Pankaj Dadheech, Pravin R. Kshirsagar, AIP Conference Proceedings · May 2022.
7. Shah, Saqib. "Microsoft's Seeing AI app for the blind now reads handwriting". Endgadget. Retrieved 6 June 2019.
8. Cecily Morrison, Ed Cutrell, Martin Grayson, Geert Roumen, Rita Marques, Anja Thieme, Alex Taylor, Abigail Sellen Interactions — June 2021 , Vol 28(3): pp. 10-13.
9. Ms. Rupali, D Dharmale, Dr. P.V. Ingole, "Text Detection and Recognition with Speech Output for Visually Challenged Person",vol. 5, Issue 1, January 2016.
10. Xiaofang Jin, Ying Xu "Research on Facial Expression Recognition Based on Deep Learning" Authorized Licensed Use Limited To: University of Canberra. Downloaded On June 07,2020 at 15:35:56 UTC From IEEE Xplore.
11. Safnaz, Shreya Shetty, Nayana Manju Jogi, Prasad Rathode and Mrs.Deeksha K R, International Journal of Engineering Research & Technology (IJERT) ISSN: 2278-0181, Published by, www.ijert.org, NCCDS - 2021 Conference Proceedings, Special Issue 2021.
12. Nagaraja, L., et al. "Vision based text recognition using raspberry PI." National Conference on Power Systems, Industrial Automation (NCPSIA 2015).



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com