International Journal of Early Childhood Special Education (INT-JECS)
DOI:10.9756/INTJECSE/V14I7.68 ISSN: 1308-5581 Vol 14, Issue 07 2022

Personal Assistive System for Dumb-Blind People using Raspberry Pi and Machine Learning

Vini Antony Grace N
Associate Professor Department of Electronics and Communication Engineering,
R.M.D. Engineering College, Kavaraipettai, Thiruvallur

Chilambuchelvan A
Professor Department of Electronics and Communication Engineering,
R.M.D. Engineering College, Chennai, TamilNadu, India

Karthika
Assistant Professor Department of Electronics and Communication Engineering,
R.M.D. Engineering College, Kavaraipettai, Thiruvallur

Aarthy B, Sri Venkata Mounika Alavala
UG Scholars Department of Electronics and Communication Engineering,
R.M.D Engineering College, Chennai, TamilNadu, India.
Corresponding Author: antonygrace81@gmail.com

Abstract: - In the era of artificial intelligence lot of object detection devices are available for visually impaired people. When it comes to subject of dumb-blind people there is none of devices available in market. Personal assistive gadgets are not available to them; only communication devices such as MyVox and Sparsha are available. In this paper we are proposing a personal assistive device to help dumb-blind people using Machine learning and Python code. It consists of multiple in-built options like object detection and communication module. A machine learning objection detection model is used, that is trained using the CNN technique for object detection. It can detect multiple objects at same time and detected objects name is intimated to user by audio output. To make communication between users with other people we are using keypad by pressing the keypad we can replay the stored audio which can be used for speech conversation.

Keywords—*Raspberry Pi, Machine Learning, Text to Voice Conversion, VNC Viewer, Python*

I. INTRODUCTION

In this modern world, on an average 1.3 billion people are suffering from vision impairment. Around 36 million people of age above 50 years are blind. Most of the blind people live in India. Likewise many

people are deaf and mute. Statistics say that more than 466 million people suffer from hearing loss. In recent years, life has become more convenient due to the advancement in technology. These advancements has still not reached the community of people with physical impairment. These people are still facing a lot of issues in their day to day lives.

Few mode of communication like Braille and sign language helps them to share their thoughts, but still there's a gap left and they must rely on another person's support. This paper focuses on techniques that help them to feel independent and do their work on their own to some extent as other normal people. The device discussed in this paper uses three modules dedicated to people with visual, audio and vocal impairment. The main unit consists of Raspberry Pi along with Google API. It also has a camera, speaker, a microphone and a screen. The device helps the visually challenged people by taking a picture of the printed document with the help of the inbuilt camera. This is then converted into a digital text by the Google Vision API. The digital text is further converted into an audio file with the help of TTS (text to Speech) library. At the same time for persons with hearing loss, the recorded speech or audio is converted into text and it will be displayed on a screen for them to read. For vocally challenged people this device speaks by converting the text to speech, whatever the user types using the customized keyboard

II. LITERATURE SURVEY

The development of user-centered interfaces and technology has become increasingly important in the process of designing for individuals with disabilities. Adding an extra feature isn't enough to help visually impaired people use technology [5-7]. There are a variety of device-based hardware and software innovations available to help the visually impaired. They have functions like reading printed or writ- ten text, expanding characters on braille systems and machines Based on computer vision [3]. For blind musicians, prototypes that work with cell phones and cameras are used to aid in the analysis of photos to identify movement patterns. [8,9]. AudioMUD [4] is a multicast virtualization technology designed specifically for blind persons that includes spoken cues. The original MUDs (Multi User Dimension) are generally text based and do not contain any sort of graphical interfaces. Users generally use MUD (Multi User Dimension) style games to per- form a

set of actions in a virtual environment with a navigable space in the presence of direction, orientation and restrictions. Because of the various sorts of interaction and text-based interface between players and the virtuality in AudioMUD with collaborative components, there is a lot of possibilities for the description of spaces and interactions. Their project focuses on the development of a server and client from scratch where the state of the world is stored in the server in such a way that when the server connects to the client, the state of virtual game is received and players can enter or exit anytime. When the blind user enters the IP address and server in the client, the player is sent into the kingdom of the human body, where they can explore the respiratory system in a random place with qualities. L. Gonzalez et al. [2] propose a technique to help the visually impaired improve their quality of life. The wearable system includes facial recognition, which uses ultrasonic sensors to generate vibration signals that indicate an obstacle, and obstacles detection, which employs the fisher faces algorithm to recognise people's faces and can identify a person through earlier system training. medication reminder to remind the user about the prescription prescribed, MP3 player as a source of amusement allowing the user to listen to music, and email reader that accesses the user's email using POP3 protocol and allows the user to listen to the email using headphones. Anusha Bhargava et al [3] propose a Raspberry Pi-based system that includes image acquisition via a webcam interface, preprocessing of the image to obtain the region of interest, template identification to detect characters and objects, converting image to text using OCR algorithm to scan the image and output a corresponding text output, saving the text data in a text file, and converting text to speech using E-speak for the blind user to hear the text. Sign language engages with people and conveys their messages mostly through manual communication such as hand movements and facial expressions. Lorenzo Monti et al [10] developed GlovePi, a wearable gadget for deaf-blind users that can recognise the people, the number and the position of individual, and facial expressions in front of the user. It mainly comprises of a gardener glove which is attached to capacitive touch sensor with raspberry Pi using a I2C interface. The Glove uses a many-to-many architecture to allow a user to register on the server, usually by sending an HTTP request, and the user is eventually added to the server, after which the server sends a latest update of all the

connected users, and thus uses peer-to-peer communication to send and receive messages.

III. TECHNOLOGIES USED

PYTHON

Python has a dynamic type system and memory management that is automated. It features a big and extensive standard library and supports several programming paradigms, including object-oriented, imperative, functional, and procedural.

Python interpreters are available for many operating systems. The reference version of Python, CPython, as well as nearly all of its variant implementations, is open-source software with a community-based development strategy. CPython is managed by the Python Software Foundation, a non-profit organisation.

Python is a multi-paradigm programming language. Many of its features allow functional programming and aspect-oriented programming (including metaprogramming and metaobjects), as well as object-oriented programming and structured programming (magic methods). Many other paradigms are supported via extensions, including design by contract and logic programming.

Python manages memory through dynamic typing and a combination of memory management and a cycle-detecting garbage collector. Late binding (dynamic name resolution) binds method and variable names while the programme is running.

Python's design offers some support for functional programming in the Lisp manner. It has functions like filter(), map(), and reduce(), as well as list comprehensions, dictionaries, and sets, as well as generator expressions. The standard library includes two modules (itertools and functools) that implement functional tools from Haskell and Standard ML. Python's creators have made it a priority to keep the language enjoyable to use. This is reflected in the language's name, which is a tribute to the British comedy group Monty Python, as well as in the language's tutorials and reference materials, which occasionally take a lighthearted approach, such as examples that use spam and eggs (from a famous Monty Python sketch) instead of the standard foo and bar.

VNC VIEWER

VNC stands for virtual network computing. This is a desktop sharing programme that allows you to operate another computer from a distance. VNC works by sending all of your keyboard and mouse actions from

one thin client machine to another large client computer. Platform independence is a feature of a virtual network computing system. This means that a client running one operating system will be unable to connect to a VNC server running a different operating system.

For many GUI-based systems, there are many different types of clients and servers. In addition, there is a virtual network computing system for Java. Some VNC apps are only compatible with the Windows operating system.

Although VNC was created by an AT&T research team, virtual network computing solutions are now widely used for both corporate and personal purposes.

Much hands on is not required to use a virtual network computing system. To connect to the computer that runs the server, all you need is a network TCP/IP connection, a VNC server, and a VNC viewer. Any thin client can be used as a VNC viewer. There are also apps that allow users to connect to computer using a smartphone.

If we can overcome the obstacles that come with using a network computing system, there is no doubt that we will reap the benefits of this technology. Aside from preserving data, the advantages include time and money savings.

We can save money on total hardware requirements. One can use VNC to run a recent programme on an older computer.

Businesses no longer require the same amount of disc memory or processor power since VNC simplifies things.

We may be able to recover any lost data on the huge client computer in the event of a disaster. Because a virtual desktop's data is stored in a secondary data centre, this is the case. This allows users to view all of the information without disrupting our normal operations.

RASPBERRY PI 3

The Raspberry Pi 3 Model B is the Raspberry Pi's third iteration. This powerful credit-card sized single board computer succeeds the original Raspberry Pi Model B+ and Raspberry Pi 2 Model B and can be utilised for a variety of applications.

The Raspberry Pi 3 Model B keeps the popular board format while adding a faster processor that is 10 times faster than the first-generation Raspberry Pi. It also has wireless LAN and Bluetooth connectivity,

making it a great option for a well-connected design.

OBJECT DETECTION

Object detection is a computer technology that deals with finding instances of semantic items of a specific class (such as individuals, buildings, or cars) in digital photos and videos. It is related to computer vision and image processing. Detecting faces and pedestrians are two well-studied domains of object detection. Object detection can be used in a variety of computer vision applications, such as picture retrieval and video surveillance.

Object detection methods are classified as either neural network-based or non-neural approaches. Non-neural approaches require first defining features using one of the methods below, followed by classification using a technique such as support vector machine (SVM). However, neural approaches, which are often based on convolutional neural networks, are capable of doing end-to-end object detection without specifying characteristics (CNN).

IV. SYSTEM ANALYSIS

EXISTING SYSTEM

Existing technology that does object detection and face recognition to assist vision impaired persons. It makes use of headphones to create an intimate presence of object and person through audio. The solutions for deaf, dumb and blind person is not carried out in a single device. So this work has been carried out to address these two impairments in a single device.

PROPOSING SYSTEM

Proposed system utilizes matrix keypad which takes input from the user and plays the already stored audio which is given in the python code.A camera is used to get images which are then converted into audio by using machine learning and text to voice conversion. In this project we are going to propose a personal assistive device to help dumb-blind people using Machine learning and Python code. It consists of multiple in-built options like object detection and communication module. We use a machine learning objection detection model that is trained using the CNN technique for object detection. It can detect multiple objects at same time and detected objects name is intimated to user by audio output. To make communication between users with other

people we are using keypad by pressing the keypad we can replay the stored audio which can be used for speech conversation.

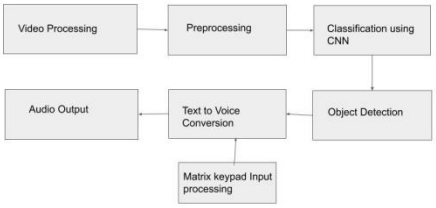

Figure 2 : Data Processing Method

V. SYSTEM DESIGN

SYSTEM ARCHITECTURE

The system architecture is implemented as follows:

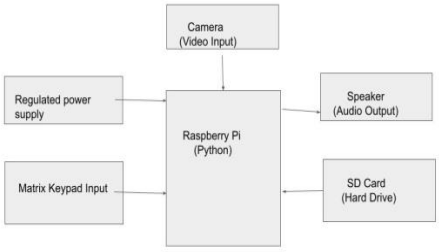

Figure 1: System Architecture

DATA PROCESSING METHOD

The data processing method is implemented as

follows:

WORKING

Raspberry Pi is the controller. Camera and Matrix Keypad are the inputs and speaker is the output.SD card is used for storage as Raspberry Pi does not have storage o board. For blind people, the camera records the video and it is pre-processed which prepares the raw data and makes it suitable for machine learning. Image classification and recognition is then done by CNN (Convolutional Neural Network). Object detection tracks the objects and accurately label them. The object is then converted into audio using text to voice conversion and we get the audio output through earphones that is the speaker. For dumb people, the input from the matrix keypad activates the particular code stored in python and it is converted into audio. Thus, the system is implemented.

MODULE DESCRIPTION

VIDEO PROCESSING

The goal of the pipeline technique is to separate and parallelize the operations that occur during processing. Decoding, processing, and encoding are the three basic activities of a typical video pipeline. At the very least, these three tasks can be completed simultaneously.

Why is the pipeline method more adaptable? The ease with which you can alter the components owing to needs is one of the pipeline's advantages. Decoding can be accomplished by encoding frames from a video file into another file. This is a regular scenario.

Instead, input from an IP camera can be transmitted through RTSP. A WebRTC connection in the browser or a mobile application can be used for output. For all combinations of input and output formats, there is a uniform architecture based on a video stream.

The computation component of the process is not always an atomic operation. It can be logically divided into sub-operations, resulting in more pipeline stages. However, it is not always reasonable to create too many components because multiprocessing latency may fade out the performance gain.

Essentially, the pipeline appears to be exactly what we require in terms of flexibility. This is due not only to the fact that we can change the input and output formats, but also to the fact that we can pause and resume processing at any time.

PREPROCESSING

Data preprocessing is an essential step in Machine Learning because the quality of data and the useful information that can be extracted from it directly affects our model's ability to learn; thus, it is critical that we preprocess our data before feeding it into our model.

Every time we create an ML model, we must go through a preprocessing phase. As a result, the ML model we're going to create can be trained correctly on the data.

DEEP LEARNING OBJECT DETECTION

When Ross Girshick (now at Microsoft Research) published Fast R-CNN a year later, this approach quickly evolved into a

purer deep learning approach. It used Selective Search to generate object proposals, similar to R-CNN, but instead of extracting them all independently and using SVM classifiers, it applied the CNN to the entire image and then used Region of Interest (RoI) Pooling on the feature map with a final feed forward network for classification and regression.

This method was not only faster, but it also allowed the model to be end-to-end differentiable and easier to train thanks to the RoI Pooling layer and completely connected layers. The model's major flaw was that it still relied on Selective Search (or any other region proposal technique) for inference, which made it a bottleneck.

TEXT TO VOICE CONVERSION

The python Pyttsx library was used to return the output in the form of speech. For example, a chair was discovered at 21 degrees left, a bottle was discovered ahead, and so on.

DEPLOYING IN THE HARDWARE

We installed this model in the hardware, which consists of a Raspberry Pi, a camera, and a speaker, after testing it in software and achieving positive results. The output in the hardware was not very accurate in the beginning, but by continuously training the system, we were able to produce a perfect hardware model to display. As a result, the hardware project was completed effectively.

VI. RESULT

The hardware has been installed, as well as all of the necessary equipment. VNC Viewer has been installed and launched in order to connect to the Raspberry Pi.

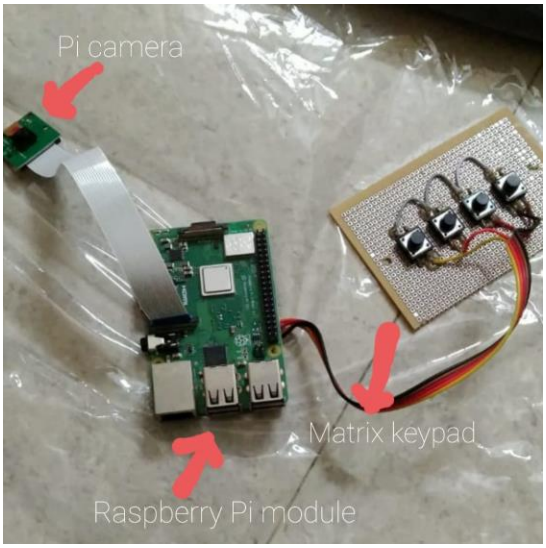

Figure 3 : Hardware Setup

By connecting them on the same wifi network, the Raspberry Pi and the VNC Viewer share the same IP address.VNC Viewer uses code written in the Python

console to call code written in the Python console.

Importing the code from Python allows the object detection model to be successfully executed.

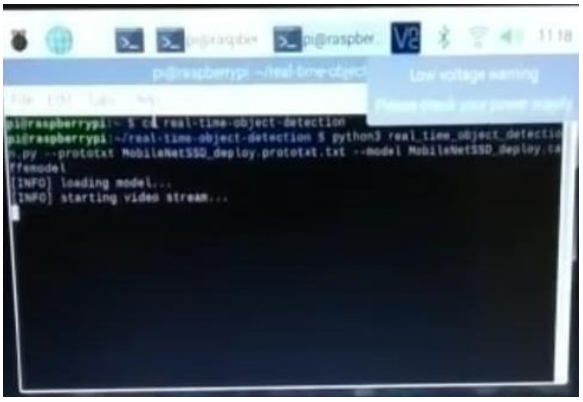

Figure 4 : Successful Execution of Object Detection

The camera captures video, after which it is processed. After that, the video is trained using the CNN algorithm and object detection, which identifies the item in terms of text and compares the accuracy with a percentage.

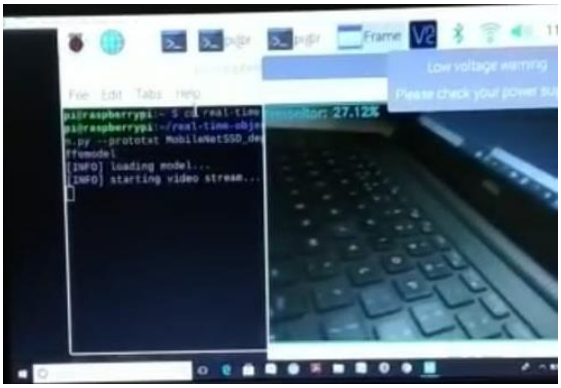

Figure 5 : Object identified as TV Monitor

The text is then transformed to audio using python code for text to voice conversion. The audio is received through the Raspberry Pi module's speaker, which is a pair of earbuds inserted into the audio port.

When a button on the matrix keypad is pressed, the previously saved python commands are activated. When button 1 is hit, the audio jack receives the voice output "Please assist me cross the road." Once button 2 is hit, the voice output is "I am hungry," for example.

Thus, for blind people, video is processed and trained by object recognition and converted into audio via which they are guided, and the stored command is transformed into audio when the button in the matrix keypad is hit, which aids communication for the dumb.

VII. CONCLUSION

This article has resulted in the creation of an unprecedented prototype to assist the visually and verbally impaired. This initiative is not only about empowering and enabling the differently abled, but it is also small and low-cost. By removing braille books and the time spent studying them, the overall cost has been reduced. It is a less

expensive solution because the device's components are all cost-effective and efficient. This device is portable, versatile, and convenient thanks to the latest and most popular technologies. The technology proposed in this research has the potential to greatly assist the differently abled in overcoming a number of obstacles. The device could be developed more compact and wearable to make it easier for the user to utilise in the future.

REFERENCES

[1] N. K., S. P. and S. K., Assistive Device for Blind, Deaf and Dumb People using Raspberry-pi, Imperial Journal of Interdisciplinary Research (IJIR), 3(6), 2017 [Online]. Available: https://www.onlinejournal.in/IJIRV3I6/048.pdf.

[2] L. González-Delgado, L. Serpa-Andrade, K. Calle-Urgiléz, A. Guzhñay-Lucero, V. Robles-Bikbaev and M. Mena-Salcedo, "A low-cost wearable support system for visually disabled people," 2016 IEEE International Autumn Meeting on Power, Electronics and Computing (ROPEC), Ixtapa, 2016, pp. 1-5. doi: 10.1109/ROPEC.2016.7830606

[3] Anushka Bhargava, Karthik V. Nath, Pritish Sachdeva MonilSamel (2015), International Journal of Current Engineering and Technology, E-ISSN 2277– 4106, P-ISSN 2347– 5161

[4] J. Sanchez and T. Hassler, "AudioMUD: A Multiuser Virtual Environment for Blind People," in IEEE Transactions on Neural Systems and Rehabilitation Engineering, 15(1), pp. 16-22, March 2007. doi: 10.1109/TNSRE.2007.891404

[5] M. Lumbreras and J. Sánchez, “Interactive 3-D sound hyper stories for blind children,” in Proc. ACM-CHI ’99, Pittsburgh, PA, 1999, pp. 318– 325.

[6] R. McCrindle and D. Symons, “Audio space invaders,” in Proc. ICDVRAT 2000, Alghero, Sadinia, Italy, Sep. 23–25, 2000, pp. 59–65.

[7] T. Westin, “Game accessibility case study: Terraformers-Real-time 3-D graphic game,” in Proc. ICDVRAT 2004, Oxford, UK, 2004, pp. 120–128.

[8] Y. H. Lee and G. Medioni, “Rgb-d camera based wearable navigation system for the visually impaired,” Computer Vision

and Image Understanding, vol. 149, pp. 3–20, 2016

[9] J. Bajo, M. A. Sanchez, V. Alonso, R. Berj ´ on, J. A. Fraile, and J. M. ´ Corchado, "A distributed architecture for facilitating the integration of blind musicians in symphonic orchestras," Expert Systems with Applications, 37(12), pp. 8508–8515, 2010.

[10] L. Monti and G. Delnevo, "On improving GlovePi: Towards a many-to-many communication among deaf-blind users," 2018 15th IEEE Annual Consumer Communications & Networking Conference (CCNC), Las Vegas, NV, 2018, pp. 1-5.doi: 10.1109/CCNC.2018.8319236

[11] R. Rastogi, S. Mittal and S. Agarwal, "A novel approach for communication among Blind, Deaf and Dumb people," 2015 2nd International Conference on Computing for Sustainable Global Development (INDIACom), New Delhi, 2015, pp. 605-610.

[12] A. M. Hassan, A. H. Bushra, O. A. Hamed and L. M. Ahmed, "Designing a verbal deaf talker system using mouth gestures," 2018 International Conference on Computer, Control, Electrical, and Electronics Engineering (ICC EEE), Khartoum, 2018, pp. 1-4. doi: 10.1109/ICCCEEE.2018.8515838.