JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

ASSISTANCE FOR VISION IMPAIRED USING IMAGE PROCESSING

[1]Sri Harini. S, [2]Sharal Maria Tina. V, [3]Samitha. S, [4]Dr. H. Mangalam

[1]Undergraduate Student, [2] Undergraduate Student, [3] Undergraduate Student, [4]Professor
[1] Department of Electronics and Communication Engineering,

[1] Sri Ramakrishna Engineering College, Coimbatore, India

Abstract : Vision impairments often require assistance in recognizing objects and faces in new environments. However, visually impaired individuals often rely on others for assistance. This research uses image processing to identify trained faces and detect items in front of the visually impaired person. The project uses the Festival text-to-speech library and the Open CV Python computer vision library. The visually impaired individual receives the names of the detected items or recognized people in audio format, enabling them to be independent in most circumstances. This technology aims to help individuals with vision impairments navigate in their daily lives.

Keywords: Object Detection, Face Recognition, Image Processing, Computer Vision, OpenCV.

I INTRODUCTION

People with vision impairments still struggle with indoor movement and object recognition in daily life. Ultrasonic sensors were used in earlier approaches to this challenge in order to identify obstructions. The Yolo deep learning architecture allows for the precise creation of a camera-based object detection system. To further assist those with vision impairments, text-to-speech converters can be utilized to convey audio.[1]

Therefore, utilizing the newest technology to identify nearby objects and people is the suggested option.[2]
Face recognition can be achieved through the use of the Local Binary Pattern Histogram technique. Despite variations in the ambient lighting, it is resilient in obtaining the intended features from the frames.[3]

II LITERATURE SURVEY

There has been significant advancement in the fields of object detection and facial recognition. The blind aid system consists of three main components: picture gathering, decision-making, and feature extraction. The process of gathering visual data using a camera is called image acquisition, as opposed to feature. Relevant visual features are extracted by extraction techniques so as to represent the context. After being extracted, these features are fed into an intelligent machine learning model to identify and classify objects, barriers, and landmarks. The model has been trained on relevant datasets. In order to help the user understand their surroundings and navigate safely, the decision-making portion of the system interprets the model's output and provides tactile or aural information.[4]

In order to save processing time, object detection in this paper is carried out after the RGB image has been converted to a grayscale image. Before the photos are given into the algorithm, they are normalized. The object's position is ascertained by measuring the bounding box's height and width. With the use of the Google text-to-speech API, the name of the discovered object is output as audio. The objects that have a probability higher than the threshold value are identified by the system.[1]

III PROPOSED SYSTEM

The project involves object detection and face-recognition. COCO (Common Objects in Context), a dataset is used to detect objects. In face-recognition the images containing faces of the people is captured and saved as .png image with the person's name as the file name. The capturing and saving of image are done using cv2(computer vision), python.

The OpenCV (Open-source computer vision library) supports the deep learning algorithms for object detection. Thus, the YOLO (You Only Look Once), a deep learning algorithm is used. The OpenCV supports the Haar cascade classifier algorithm, which is used for face detection and the Local

Binary Pattern Histogram Algorithm for face-recognition.

The object detection is performed by using a pre-trained model, trained with the COCO dataset by the YOLO (You Only Look Once) algorithm.

The face-recognition includes two steps, the first step is to detect the presence of face. The pre-trained model by Haar cascade classifier algorithm is used to detect the presence of face.

The second step is to recognize the face, whom it does belong. For this step, the dataset containing the face images with the person's name as the file name is used. This dataset is trained by the Local Binary Pattern Histogram (LBPH) algorithm. The trained model is saved in the YAML (Yet Another Markup Language or YAML ain't markup language) format.

The Hardware consists of Raspberry Pi 3 B+, a pi camera is connected to the raspberry pi through Camera Serial Interface and an earphone is connected to the raspberry pi through a audio jack.

The project is written in python language and the code is uploaded to the Raspberry Pi 3B+, to run the code the Thonny Python IDE (Integrated Development Environment) is used.

BLOCK DIAGRAM

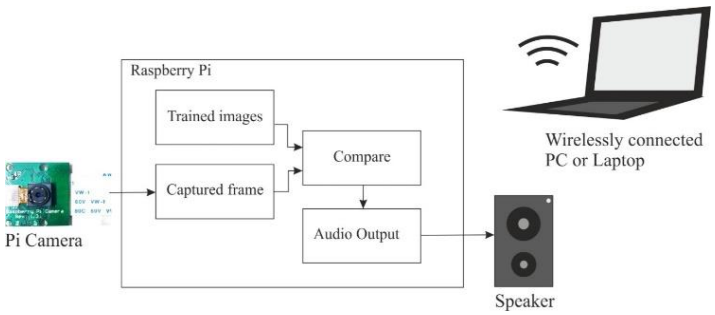

Figure 1

From the Figure 1 the block diagram of the proposed system and the connections among the hardware components can be understood

FLOW DIAGRAM

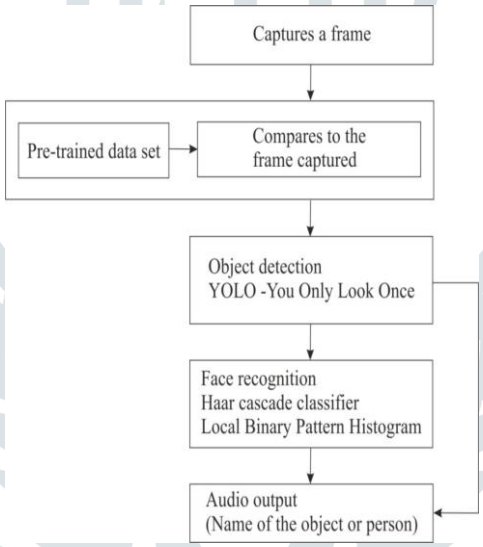

Figure 2

In the Figure 2, the work flow of the system is shown. The frames captured by camera is compared to the trained datasets to deliver an appropriate audio output.

IV METHODOLOGY

4.1 RASPBERRY P 3B+

Figure 3

The Broadcom BCM2837B0 processor-equipped Raspberry Pi single-board computer is widely used in hardware and Internet of Things (IoT) applications. Because of its price and adaptability, it's a great tool for beginners learning to program from scratch.

The Raspberry Pi is a great place for prospective programmers to start since it offers a hands-on platform for studying coding ideas and electronics principles, along with a wealth of community support and resources.

4.2 PI CAMERA

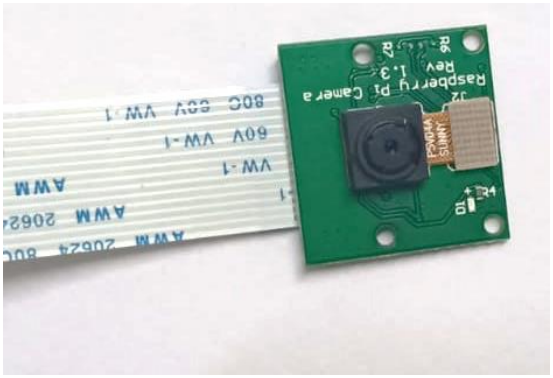

Figure 4

A Raspberry Pi board is utilized with the Pi camera. The Raspberry Pi device is used to record and capture photos in many different projects.

The input device is the Pi camera.

The frames are captured and processed further after that. Five Megapixel cameras are utilized for still photos, while a 1080p camera with a 30 frames per second maximum resolution is used to take videos.

4.3 EARPHONE

Figure 5

The output device is an earphone or speaker that provides audio output to people who are blind or visually impaired.

In addition to improving the user's awareness and making navigation and interaction with their surroundings easier, this audio feedback provides information about the surrounding environment.

4.4 Thonny python IDE:

An Integrated Development Environment (IDE) designed specifically for learning and working on Python projects is called Thonny Python. Because of its intuitive interface, which makes code development, execution, and debugging easier, it's especially good for novices.

Developers can quickly create image processing techniques using certain algorithms within this IDE. Users can concentrate on improving their programming abilities and learning the nuances of image processing by using Thonny's solid features and simplicity, which simplify the learning process.

Consequently, Thonny Python becomes the preferred option for individuals wishing to learn Python programming and take on image processing projects.

4.5 OpenCV:

With its wide range of features, OpenCV (Open Source Computer Vision Library) is a crucial library for computer vision projects. OpenCV was first developed in C++ but now supports MATLAB, Octave, Python, and Java, among other computer languages.

Its wide range of features enable developers to investigate and apply various computer vision technologies. OpenCV is a vital tool for both researchers and practitioners, facilitating a wide range of applications from image processing to machine learning.

Its open-source design encourages cooperation and creativity, which strengthens its standing as a pioneer in the computer vision industry.

4.6 Festival (library):

Festival is a free program that is widely used in many different applications and has the ability to convert text to speech. It is primarily a C++ programming language tool, but it may also be used to convert text to speech with good results.

The broad use of Festival demonstrates its efficiency and adaptability in speech synthesis applications. Festival is essential to improving accessibility and usability, whether in communication aids, educational resources, or assistive technologies.

Its open-source design encourages ongoing innovation and development, guaranteeing its applicability and usefulness in a variety of settings.

4.7 Object Detection:

The COCO (Common Objects in Context), a pre-trained dataset, is used for object detection. The YOLO (You Only Look Once) algorithm is used to train the COCO dataset.

4.8 Face recognition:

Face recognition involves two steps: the first is to detect the presence of the face in an image, and the next is to recognize the person to whom the face belongs by comparing it to the trained dataset.

To detect the presence of a face in an image, the Haar cascade classifier algorithm is used. A pre-trained Haar cascade model is used to detect faces in the image.

The Local Binary Pattern Histogram algorithm is used to recognize the faces detected in the image by comparing them to the trained dataset containing the faces of the people and their names as labels.

V RESULTS

5.1 Hardware setup:

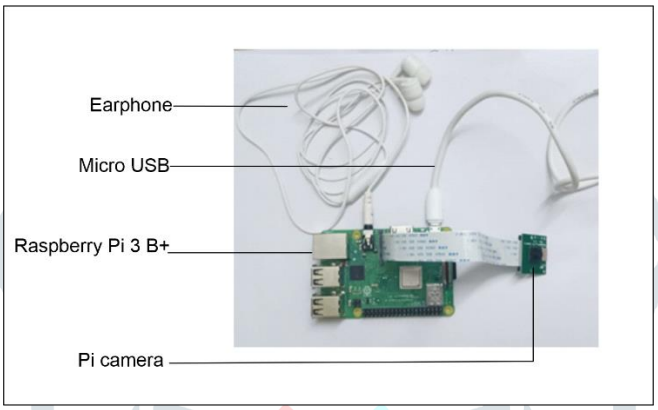

Figure 6

Serving as the brains behind the system, the Raspberry Pi makes it easier to connect the input (a Pi camera) and output (earphones or speakers).

Through the camera serial interface, the Pi camera and Raspberry Pi may communicate, allowing for the processing and capturing of images. Likewise, the audio jack on the speaker or earphone is connected to the Raspberry Pi to enable the user to get audible feedback.

Furthermore, the Raspberry Pi has remote access and control capabilities that let users monitor, configure, and communicate with the system from a laptop or personal computer. The system's adaptability and usefulness are improved by its remote accessibility.

5.2 Training the faces

5.2.1 Labelling the face to be trained

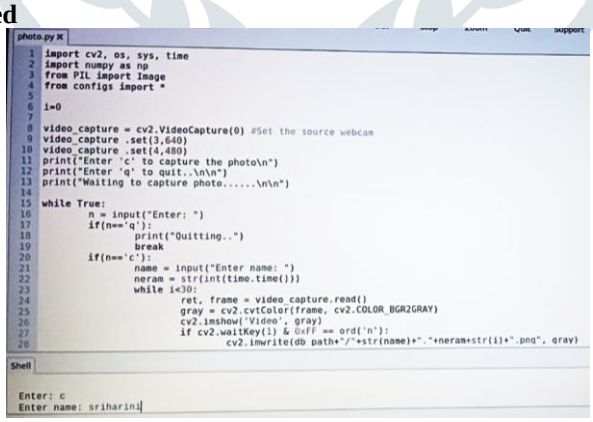

Figure 7

5.2.2 Capturing the image

Figure 8

5.2.3 Training

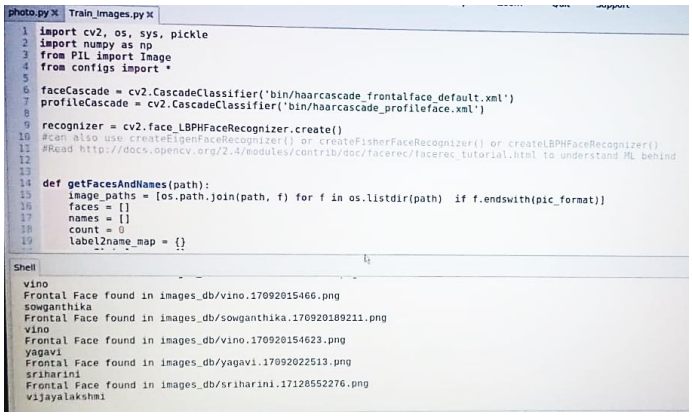

Figure 9

5.3 Output for object detection

The Detection of book can be seen in Figure 10

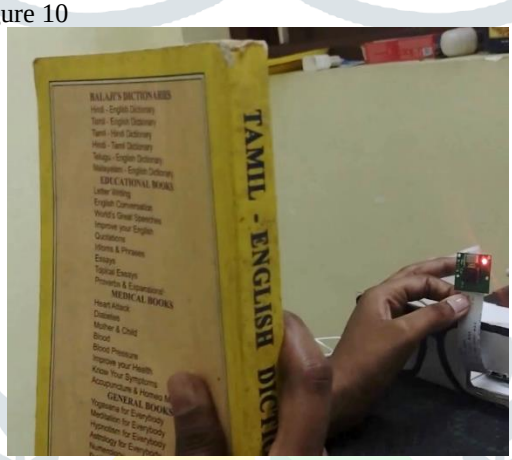

Figure 10

Output for detecting book is in the Figure 11

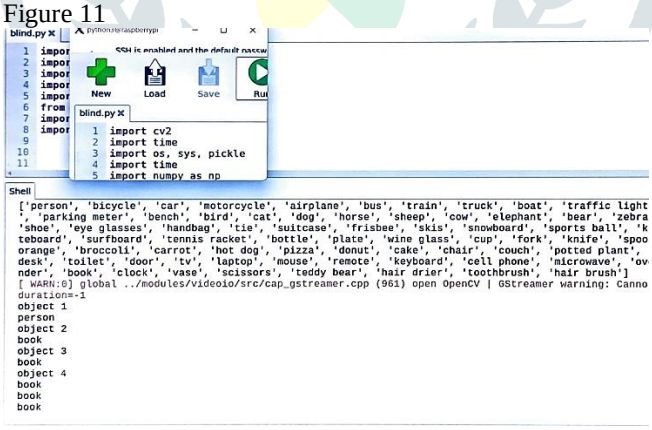

Figure 11

The Detection of cell phone can be seen in Figure 12

Figure 12

Output for detecting cell phone is in the Figure 13

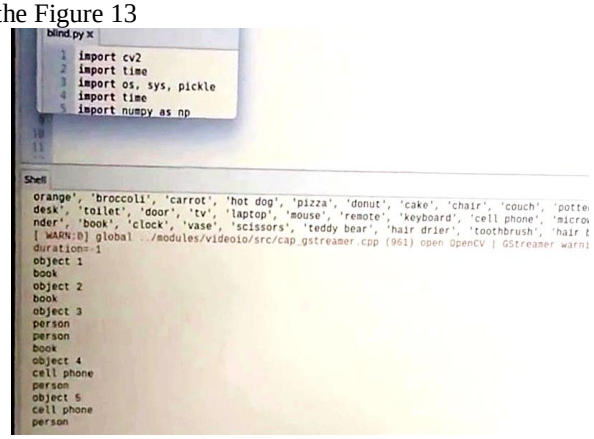

Figure 13

The Detection of stop sign can be seen in Figure 14

Figure 14

Output for detecting stop sign is in the Figure 15

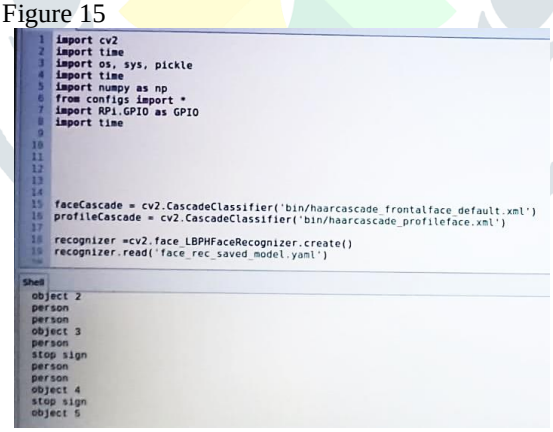

Figure 15

The Detection of bottle can be seen in Figure 16

Figure 16

Output for detecting bottle is in the Figure 17

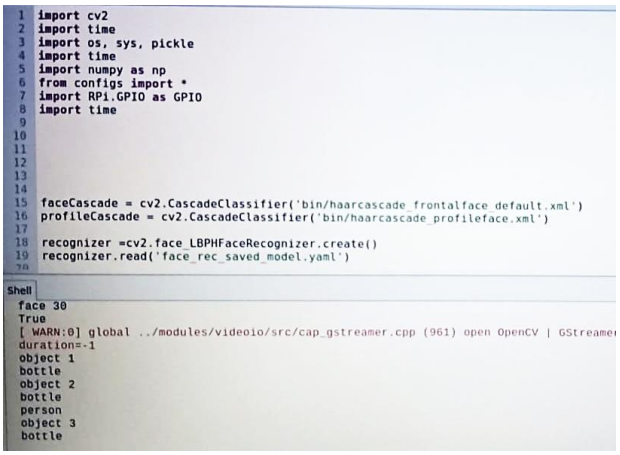

Figure 17

5.4 Face recognition

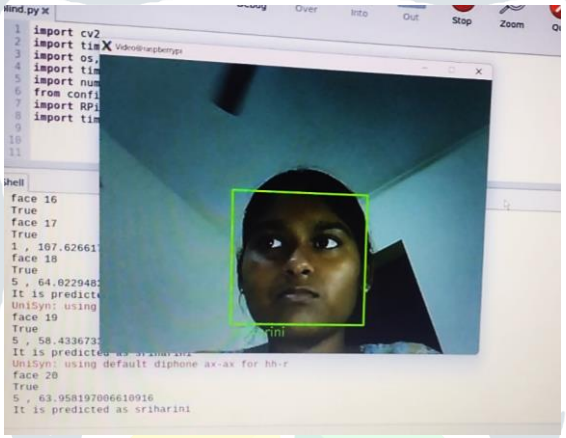

Figure 18

The output for Face recognition is in the Figure 18

VI CONCLUSION

Using a Haar cascade algorithm, a sophisticated technique for object and face detection and recognition was created to detect nearby items and faces. This invention helps people with vision impairments become more independent

This consolidation provides a flexible solution that can be used in a variety of areas while addressing earlier constraints. By integrating these features, our technology improves visually impaired people's accessibility and usefulness and gives them the confidence to effectively navigate their environment. With this development, assistive technology has advanced significantly and now offers a complete solution for those who are visually impaired to interact with their surroundings in an effective manner. Furthermore, the Haar cascade algorithm's adaptability allows it to be applied in a variety of areas, expanding its possibilities beyond helping the blind. All things considered, our comprehensive approach improves the quality of life and simplifies accessibility for people with visual impairments, allowing them to live more independently and with greater ease.

VII ACKNOWLEDGEMENT

We would like to thank Dr. H. Mangalam, the project guide, the Department of Electronics and Communication Engineering professors, Sri Ramakrishna Engineering College, Vattamalaipalayam, Coimbatore, our friends and classmates, and the authors of the research and articles that are cited. Their knowledge and assistance have greatly influenced the article's content, and their contributions have greatly influenced how the topic is understood and analyzed. We also want to express our gratitude to our families for their constant support and tolerance during our academic careers.

REFERENCES

[1] Karmarkar, R. R., & Hommane, V. N. (2021). Object detection system for the blind with voice guidance. *Int. J. Eng. Appl. Sciences Technol*, *6*(2), 67-70.

[2] Aralikatti, A., Appalla, J., Kushal, S., Naveen, G. S., Lokesh, S., & Jayasri, B. S. (2020). Real-time object detection and face recognition system to assist the visually impaired. In Journal of Physics: Conference Series (Vol. 1706, Issue 1, p. 012149). IOP Publishing. https://doi.org/10.1088/1742-6596/1706/1/012149

[3] Chen, Y. C., Liao, Y. S., Shen, H. Y., Syamsudin, M., & Shen, Y. C. (2022). An Enhanced LBPH Approach to Ambient-Light-Affected Face Recognition Data in Sensor Network. *Electronics*, *12*(1), 166.

[4] Kumar, N., Sharma, S., Abraham, I. M., & Sathya Priya, S. (2022). Blind Assistance System Using Machine Learning. In Third International Conference on Image Processing and Capsule Networks (pp. 419–432). Springer International Publishing. https://doi.org/10.1007/978-3-031-12413-6_33