RESEARCH ARTICLE OPEN ACCESS

The Third Eye – An Assistive Technology for the Blind

Amith Kumar V A, Kailas P A, Nidhin John, Mohammed Sahl Saju Kasim
Department of Computer Science
Toc H Institute of Science & Technology
Arakkunnam
India

ABSTRACT

The Third Eye device uses the power of artificial vision (Image Processing) and machine learning by incorporating pioneering technology into a wearable platform which improves the lives of individuals who are blind, visually impaired, and have reading difficulties. This consist of a small camera that can be clipped on to the person's spectacles and which is connected to a small speaker device. It assist visually impaired and blind people in their day-to-day activities. It can read any printed or digital text, including the morning paper, text messages, and even the computer screen at their office. It can recognize the faces of their friends and family. It can help one identify products, which allows people who are blind or visually impaired to identify objects in front of them so that they need not require assistance from anyone else. Assistive technology for blind and visually impaired people includes assistive, adaptive, and rehabilitative devices that promote greater independence. It does so by helping people to perform tasks that they have difficulty performing. It helps with such tasks, mainly activities involving reading. This is not technology for the blind that enables seeing but rather technology that allows the user to be able to read texts without asking others for help. The device recognize any text from any surface, and relays the information to the user via a small speaker. This allows the user to be able to read work emails, medicine labels, and more all on their own. For people who are visually impaired, this kind of technology for the blind makes the ability to walk independently much easier.

Keywords :- OCR, Deep learning, Object Recognition, Face Recognition

I. INTRODUCTION

The total number of visually impaired people worldwide is approximately 285 million, in other words more than 3.86% of the world population. Visually impaired people have numerous difficulties with accessing printed text using existing technology. As much as 80% of the visual blindness seen across the globe, according to the World Health Organisation, is either preventable or treatable. So those who remain on the visual impairment spectrum are left needing to find other ways of dealing with the problem. But if the ocular condition can't be improved, there are, as object-recognition technology and artificial intelligence develop, more opportunities to counter its limiting effects.

The field of assistive technology (AT), commonly considered to be technology designed for individuals with some form of impairment (or the elderly people), is a vital field expanding at a swift pace. Assistive technology for the visually impaired (VI) and Blind people is concerned with "technologies, equipment, devices, apparatus, services, systems, processes and environmental modifications" that enable them to overcome various physical, social, infrastructural and accessibility barriers to independence and live active, productive and independent lives as equal members of the society. This project represents a smart device that assist a visually impaired person which effectively and efficiently detect objects and reads out printed text.

We are developing a camera based vision technology for the blinds that guides and provide assistance by detecting objects/obstacles in front of them and give out voice output about the object/obstacle through speakers. It also recognize faces of the user's friends or relatives as they approach towards them. They also provide voice based assistance to the users which include getting details about current time, location and whenever needed a help message can be sent to the caretaker's mobile phone through voice input.

Based on the aforementioned characteristics and serving purposes of camera-based solutions, it is possible to use them for detecting objects and assisting with orientation and navigation based on processed images. With them, people with visual impairments and blindness will not have as much cognitive workload as they have today. However, their current usage is limited: not every aid is intended for people with vision loss. Although camera-based vision technologies have not yet received much attention in commercial markets, academic research is making progress in developing the potential of such systems for detecting objects and supporting orientation and navigation to support independent travel for people with vision loss

II. LITERATURE SURVEY

The existing system is the one which is either a camera based or sensor based devices that performs only limited functionalities or applications like either only reading or object detection but not both at the same time.

Existing systems are:

FingerReader is an index-finger wearable device that

supports the VI in reading printed text by scanning with the finger. Features novel hardware and software that includes haptic response, video-processing algorithms and different output modalities. The finger-worn design helps focus the camera at a fixed distance and utilizes the sense of touch when scanning the surface. Additionally, the device does not have many buttons or parts in order to provide a simple interface for users and easily orient the device. The MIT finger device is only usable for text reading for blinds, it does not help in navigation or face detection.

The "Enactive torch", the device sends out infrared signals which enable the user to sense nearby objects, thanks to a small buzzer attached to the wrist which alerts them if they're about to crash into a door frame or wall.

AI Glasses combines computational geometry, artificial intelligence, and ultrasound techniques, amongst other things, to create a useful aid for the visually impaired. The prototype combines glasses with stereo sound sensors and GP:S technology attached to a tablet, which can give spoken directions and recognize denominations of currency, read signs, identify colours, and other things.

III. METHODOLOGY

The third Eye's camera device clips on to a user's eyeglasses. It is connected via a Bluetooth to a small computer that fits in a pocket. The interface is minimal: the caretaker of the blind user uses the website application to register and upload faces of individuals whose faces are to be recognized by the device. The uploaded photos get stored on to the database where the system is trained to recognize the faces. So, after training as the person comes in front of the camera frame the system detects the face and spell out the corresponding names.

By default the camera opens up as soon as the application is launched and the camera remains working until the application is closed. The camera captures and sends the live sessions of video segmented images in front of the user to the server, where the image gets processed and objects are classified using deep learning algorithms. The corresponding image label is send as voice output back to the mobile application and the user gets to know about the objects in front of them.

Next for the voice assistance the user needs to hold the volume down button for 3 seconds of time and the voice assistance module gets activated and the user can now give voice inputs by saying the voice commands followed by "okay". The application process the voice command and the corresponding voice output is given to the user.

Whenever the user gets lost somewhere or when they are indeed of help, the user can ask for help through voice commands simply by activating the voice assistance module and speak out the message that needs to be sent followed by "help okay" command. The message is sent to the caretaker's registered mobile number. The caretaker can then locate the user/blind through the web portal.

For reading the printed text, the OCR (Optical Character Recognition) module needs to be activated, which can be done in the same way through voice command – "OCR okay", then the camera gets opened up and reads the printed text through camera and spells out the sentences or words that get captured in the camera frame. In order to get best results the printed text needs to be kept steady.

Following are the applications or features of the Third Eye Device:

- Reading: The device converts printed text to speech. One can read most digital text, such as that on a television or computer. It reads printed text, such as books, newspapers, food labels, restaurant menus, and even street signs.
- Identifying people: The device identifies known faces. Basically, you take an image of the face you want to be able to recognize and add a name. The device will then announce the name of that person once they enter the camera's view.
- Recognition of specific objects: In a similar way to the face recognition, there is a feature to recognize objects around the user e.g table, chair, books etc.
- Voice over assistant**:** This device helps the user to know about the time, making a phone call, current location, weather details by giving the voice as input. The system responds to users voice input and give voice output through speakers.

1. OCR (Optical Character Recognition) Module

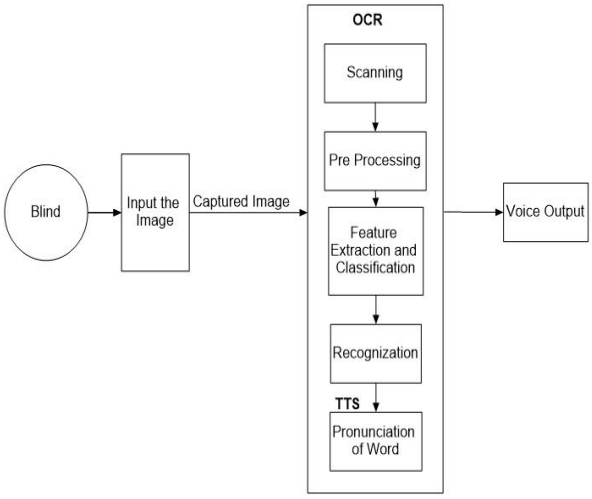

Figure 1- Block Diagram of OCR Module

This module provides a full alphanumeric recognition of printed or handwritten characters at electronic speed by simply scanning the images. Forms containing characters images can be scanned through camera and then recognition engine of the OCR system interpret the images and turn images of handwritten or printed characters into ASCII data

(machine-readable characters). This ASCII data is converted to corresponding audio output using Text To Speech synthesis technique.

2. Web Application Module

The caretaker can register and login through the web browser, where the he/she can perform tasks such as view the registered devices, track the users, manage the devices of blind, and upload images of persons whose face is to be recognized.

3. Object Detection Module

The Object Recognition module provides a way to identify specific trained objects within the current image. Here we have used a pre-trained model of COCO (Common Objects in Context) dataset for the purpose. It contains around 90 trained objects which can be correctly classified and labelled.

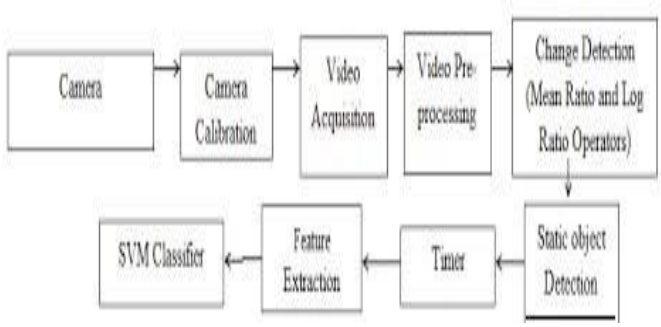

Figure 2- Block Diagram of Object Detection Module

4. Face Recognition Module

The facial recognition module is used to automatically identify people by their video images. It recognizes faces captured by camera by comparing their parameters with digital templates stored in a dedicated database and give voice outputs of names corresponding to the images. In addition to automatic identification of people based on video images, the module enables users to add or remove entries from the reference database.

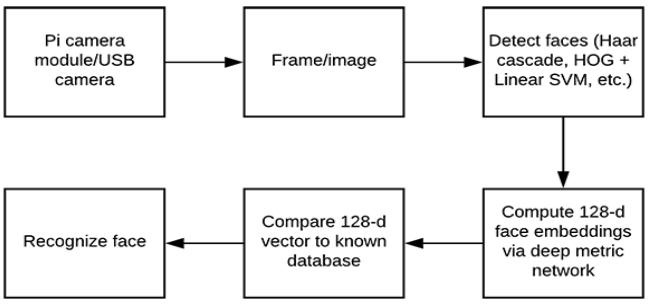

Figure 3 - Block Diagram of Face Recognition Module

5. Voice Assistance Module

The device can hear and process voice input and voice based assistance to the users. Such as, it helps the users do things like checking the current time, location and ask for help from caretakers through text messages.

IV. IMPLEMENTATION

System Requirements:
Hardware and software requirements for the installation and smooth functioning of this project could be configured based on the requirements needed by the component of the operating environment that works as front-end system here we suggest minimum configuration for the both hardware and software components.

Working off with this software is requirements concrete on system environments. It includes two phases

- Hardware requirements
- Software requirements

Software Requirements:

- Operating System: WINDOWS 7/8, Android Jelly Bean or Higher
- Front End: Python/ Android
- Back End: MySQL
- IDE Used: Jupyter Notebook / Android Studio/ Visual Studio Code/Sublime Text/SQLyog
- Deep Learning Library – TensorFlow
- Framework: Flask
- Web Browser: Internet Explorer/Google Chrome/Firefox (for Web Application)

Hardware Requirements:

- Input Device : Mouse, Keyboard, Touch Screen, Camera, Mic
- Output Device : Monitor, Speaker
- Memory : 4GB RAM or above
- Processor : i3 processor or higher

Block Level Implementation:

1. Web Application - The website is used for the caretakers to login and manage the devices of the blinds and also the new users can register through the website. The data's get stored in MySQL log database where the admin can manage the database. The caretaker can add new devices using the IMEI number of the mobile device.

2. Mobile Application - The unit runs on the blind users mobile which is connected to the server using the IP address. The application captures the image and continuously sends the image to server which process the image and send back the output. The application also provide voice assistance to the user through voice outputs.

3. Object Detection - The object detection module is implemented using python programming and tensor flow. The dataset used here is the COCO (Common Objects in Context) dataset which is a large-scale object detection, segmentation, and captioning dataset. The captured image from camera gets updated every 15 seconds in time within the server and from the image three relevant detected objects are processed and classified.

4. Text Detection - The alphanumeric recognition of printed or handwritten characters at electronic speed by simply scanning the images. Forms containing characters images can be scanned through camera and then recognition engine of the OCR system interpret the images and turn images of handwritten or printed characters into ASCII data (machine-readable characters). This ASCII data is converted to corresponding audio output using Text To Speech synthesis technique.

5. Face Recognition – This unit is also implemented using python programming and tensor flow concepts. The face detection is processed using the Haar Cascade algorithm and the detected face is searched for the known persons face in database. Every person's face has many different peaks and valleys that make up facial features. These peaks and valleys are called landmarks represented as nodal points. The human face has about 80 nodal points, which is used for the searching procedure and if any matched face is found the corresponding voice output is given.

6. Voice Assistance – This unit can understand voice commands from the user and complete tasks for a user. This module gets activated by holding down the volume down button on the device. Then the user can input voice commands (time, location, OCR, help etc.) followed by the "okay" command. The voice command gets processed and the corresponding output is given to the user via voice outputs through speakers.

V. FUTURE WORK

Technology created reading machines, talking books, and computers that translate Braille. More and more smart machines have expressive and receptive speech capabilities. These devices become smaller and less expensive with each passing year. There is not a comparable set of breakthroughs with navigational technologies. Communications technologies were easier to invent and consequently were brought to market. Navigational technologies are complex and are not yet small enough or inexpensive enough for the mass market. Technologies available for blind navigation are insufficiently developed, adapted, and marketed.

For our project we desire to make the device as compactable as it gets for the easiness for users to carry along with the glasses, GPS navigation can be added in order to receive location and navigation for the user as well as to locate the blind person to their relatives or friends, night vision camera can improved high image processing and thus making face recognition and text reading more advanced during night and in low light areas, by adding ultrasonic sensors more navigational facilities can be adopted for object detection (like calculating the distance from object and warning the user), much prior improved work would be making the device cost efficient . More and more voice assistance task can be added to the device making the users to do their task more independently without seeking assistance from anyone else.

VI. CONCLUSION

In this project, we investigate the need from blind and visually impaired people. Based on the impetus of the CNN(Convolutional Neural Network), we develop a blind visualization system that helps blind people better explore the surrounding environment. A portable and real time solution is provided in the project. It is an advanced wearable assistive technology that provides independence by allowing access to visual information, conveyed by audio, on a tiny camera which can be attached to any pair of eyeglasses. The work is mainly focused for individuals who are visually impaired, partially sighted, blind, print disabilities, or have other disabilities. Using Artificial Intelligence it reads text, recognizes faces, and identifies objects. We have described a prototype system to read printed text for assisting the blind people. Off-the-shelf OCR is used to perform word recognition on the localized text regions and transform into audio output for blind users. It also provide voice assistance to the users which helps them in knowing details about current time, location or send help messages to their caretaker.

REFERENCE

[1] Patrik KAMENCAY, Miroslav BENCO,Tomas MIZDOS, Roman RADIL, "A New Method for Face Recognition Using Convolutional Neural Network", University of Zilina, Univerzitna 8215/1, 010 26 Zilina, Slovakia

[2] Thai Hoang Le, "Applying Artificial Neural Networks for Face Recognition" Department of Computer Science, Ho Chi Minh University of Science, Ho Chi Minh City 70000, Vietnam, 19 July 2011

[3] MEENA, D. and R. SHARAN. An approach to face detection and recognition. In: International Conference on Recent Advances and Innovations in Engineering (ICRAIE). Jaipur: IEEE, 2016

[3] Aashish & Vijayalakshmi , Comparison Of Viola-Jones and Kanade-Lucas-Tomasi Face Detection Algorithms, *Orient. J. Comp. Sci. & Technol.*, Vol. **10**(1), 151-159 (2017)

[4] Inseong Kim, Joon Hyung Shim and Jinkyu Yang (2016) *Face Detection*, Stanford University, *International Journal of Engineering Research and Applications*, **6**, 1, pp145-150

[5]http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html.

[6] https://www.bioid.com/About/BioID-Face-Database