International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2019 IJSRCSEIT | Volume 5 | Issue 2 | ISSN : 2456-3307
DOI : 10.32628/CSEIT1952239

Object and Face Recognition Using Raspberry Pi

Divya S[1], Kowsalya K[1], Shalu Sunil[1], Subashini V[1], Suvetha V[2], Dr. Suguna N[3]
([1]UG Scholar, [2]Assistant Professor, [3]Professor)
(Department of Computer Science and Engineering, Akshaya College of Engineering and Technology, Coimbatore, India)

ABSTRACT

Speech and text is the main medium for human communication. A person needs vision to access the information in a text. However, those who have poor vision can gather information from voice. This paper proposes a camera based assistive text reading to help visually impaired person in reading the text present on the captured image. The Object can also be detected when a person enters into the frame by the mode control. The proposed idea involves text extraction from scanned image using Tesseract Optical Character Recognition (OCR) and converting the text to speech by e-Speak tool, a process which makes visually impaired persons to read the text. This is a prototype for blind people to recognize the products in real world by extracting the text on image and converting it into speech. Upon entering the camera view previously stored faces are identified and informed which can be implemented.

Keywords : OCR, Raspberry pi, frame, object, face recognition.

I. INTRODUCTION

A person needs vision to access the information in a text. This paper proposes a camera based assistive text reading to help visually impaired person in reading the text present on the captured image. The faces can also be detected when a person enter into the frame by the mode control. The proposed idea involves text extraction from scanned image using Tesseract Optical Character Recognition (OCR) and converting the text to speech by e-Speak tool, a process which makes visually impaired persons to read the text.

II. LITERATURE SURVEY

OCR (Optical Character Recognition) based reading aid

There is need to convert paper books and documents into text. OCR is still imperfect as it occasionally mis-recognizes letters and falsely identifies scanned text, leading to misspellings and linguistics errors in the OCR output text. A multilingual voice creation toolkit that supports the user in building voices for the Open source MARY TTS platform, for two state-of-the-art speech synthesis technologies: unit selection and HMM-based synthesis. For languages not yet supported by MARY TTS, the toolkit provides the necessary tools and generic reusable run-time system modules for adding support for a new language. Here, discussed the design and implementation of assistive platforms, and combine today's growing technology along with smart phone capabilities with the advantages offered by the high scalable and reactive cloud resources, making it a more economical and cost-effective way.

A solution for Reading Machine for Visually Impaired Using TTS AND OCR discussed in paper proposes a standalone reading machine that takes

Arabic input text and produces speech for visually impaired people. It also presented the components of reading machines in general using OCR TTS for visually impaired people. It is considered as an assistive technology that promotes educational purposes and knowledge for visually impaired people. Also it is not very different from the other Latin standalone machines in terms of components and design, the main difference is in terms of software readability and character recognition.

Raspberry Pi And Image Processing Based Person Recognition System For Visually Impaired People

The implementation of image processing in today's world has done a lot in various field. Visually impaired or people with low vision always face difficulties recognizing a person while social interactions and regular activities. The proposed system is designed for blind people to reduce the complexity of blind people in regular activities and engage them. This paper represents a Raspberry Pi embedded system based person's face recognition system using local binary pattern algorithm of image processing aiming visually impaired people. This device detects and recognize human faces of pre-made datasets, fetches information from MySQL database on match and spells through a microphone applying text to speech method. In the experimental case, the system gives 96% accuracy recognizing single person and 92% accuracy in the case of recognizing multiple people and in complex background.

According to an estimation there are 253 million individuals live with vision disability: 36 million are blind and 217 million have direct to extreme vision impairment. Among blind and have moderate or severe vision impairment, 81% people are aged 50 years and above. Sometimes loss of vision is related with a misfortune of freedom. Blind and visually impaired individuals often face many challenging situations in their day to day life. One of the biggest problems blind people face is difficulty with navigation Individuals who are completely blind or have disabled vision as a rule have a troublesome time navigating outside the spaces that they're usual to. They had to depend on the external support system which can be maintained by humans, guided dogs, or special electronic gadgets. Another major problem they had to face is identifying a person in a variety of social interactions.

In today's world many researchers have discovered some technological devices and tools to aid the blinds and prosopagnosia disordered individuals. A device Option is utilized by a blind person to read the printed page straightforwardly. It comprises of a camera, which is moved over the lines of print, and a small display comprising of lines of modest vibrating pegs which appear the shape of the letter specifically beneath the camera connected with internet via Raspberry Pi 3's built in Wi-Fi system that links the device fetching data from MySQL on recognizing a person.

Image processing system is divided into three parts: Detect face and store in database with preprocessing, Training images from dataset, Classify and recognize face using LBP algorithm. Based on the match, information is fetched from MySQL database, the text to speech method using python's library pyttsx is applied which gives a voice output from earphone. As a visually impaired person can recognize and get short information people in front of him.

OCR based automatic book reader for the visually impaired people using raspberry pi

Optical character recognition (OCR) is the identification of printed characters using photoelectric devices and computer software. It coverts images of typed, handwritten or printed text into machine encoded text from scanned document or from subtitle text superimposed on an image. In

this research these images are converted into audio output. OCR is used in machine process such as cognitive computing, machine translation, text to speech, key data and text mining. It is mainly used in the field of research in Character recognition, Artificial intelligence and computer vision. In this research, as the recognition process is done using OCR the character code in text files are processed using Raspberry Pi device on which it recognizes character using tesseract algorithm and python programming and audio output is listened. To use OCR for pattern recognition to perform Document image analysis (DIA) we use information in grid format in virtual digital library's design and construction. This research mainly focuses on the OCR based automatic book reader for the visually impaired using Raspberry PI. Raspberry PI features a Broadcom system on a chip (SOC) which includes ARM compatible CPU and an on chip graphics processing unit GPU. It promotes Python programming as main programming language with support for BBC BASIC.

Optical Character Recognition or OCR is the text recognition system that allows hard copies of written or printed text to be rendered into editable, soft copy versions. It is the translation of optically scanned bitmaps of printed or written text into digitally editable data files. An OCR facilitates the conversion of geometric source object into a digitally representable character in ASCII or Unicode scheme of digital character representation.

OCRs are of two types: for recognizing printed characters and for hand written text OCR PROCESS Character recognition is a sub-field of pattern recognition in which images of characters from a text image are recognized and as a result of recognition respective character codes are returned, these when rendered give the text in the image. The problem of character recognition is the problem of automatic recognition of raster images as being letters, digits or some other symbol and it is like any other problem in computer vision.

III. PROPOSED SYSTEM

The proposed idea involves text extraction from scanned image using Tesseract Optical Character Recognition (OCR) and converting the text to speech by e-Speak tool, a process which makes visually impaired persons to read the text.

The main purpose of this model is to help blind persons by guiding them using this system design. It recognizes the face, signs, obstacles, humans such as known and unknown persons will be identified using face and text recognition features. It gives the scanned and recognized images in the form of audio output to help and guide the blind person. It is specially designed to blind navigation purpose.

In this proposed system text recognition is done by Open Computer Vision (Open CV), a library of functions used for implementing image processing techniques. Image processing is a technique of using mathematical operations in image, any form of inputs such as image, a series of images, or a video can be used for processing. An image or a set of characteristics or parameters related to image is the output of image processing. Image processing has various applications like computer graphics, scanning, facial recognition, text recognition etc. Various features of text like its font, font size, alignment, background etc. influences in its recognition. Number plate recognition is a fair example for text extraction.

Text extraction from an image is carried out by OCR. It is a method of conversion of images of writings on a label, printed books, sign boards etc. to text only. OCR helps to create reading devices for visually impaired persons and technologies involving telegraphy. The binary image is converted to text by Tesseract library in OCR engine that detects the outline, slope, pitches, white spaces and joint letters. It also checks the quality of the recognized text. In this system the conversion of text to voice output is

by e-Speak algorithm. The e-Speak is a Text-To-Speech (TTS) system which converts text into speech. The artificial production of human speech is known as speech synthesis. The speech synthesizer can be implemented in a software or a hardware product. The platform used for this purpose is known as a speech synthesizer. The storage of entire words or sentences allows for high-quality output in specific usage domains. A synthesizer can incorporate the model of a vocal tract and other human voice characteristics.

This paper aims to build an efficient camera based assistive text reading device. The idea involves text extraction from image taken by a camera installed on a spectacle. The extracted text is then converted to audio signals and to voice output. It is also used to detect a person's face in the frame. This is carried out by using Raspberry pi where the portability is the main aim, which is achieved by providing a battery backup.

Conversion of image to text using OCR tool

Tesseract is an open source-OCR engine. It assumes that its input is a binary image with optional polygonal text region defined. The first step is a connected component analysis in which outline of the components is stored. By the inspection of the nesting of outlines, it is easy to detect inverse text and recognize it as early as black on white text. At this stage, outlines are gathered together, purely by nesting, into blobs. Blobs are organized into text lines, and the lines and regions are analyzed for fixed pitch or proportional text. Slope across the line is used to find text lines. These lines are broken into words differently according to the kind of character spacing. Fixed pitch text is chopped immediately by character cells. The cells are checked for joined letters and if it is found then it is separated. Quality of recognized text is verified. If clarity is not enough the text is passed to associator.

Conversion of text to voice using e-Speak

Normal text to speech conversion is done using e-Speak which is a TTS system. The artificial production of human speech is known as speech synthesis. Speech computer or speech synthesizer is used for this purpose and can be implemented in software or hardware products. The storage of entire words or sentences for specific user domains, allows for high-quality output. To create a completely "synthetic" voice output a synthesizer can be used to incorporate a model of the vocal tract and other human voice characteristics.

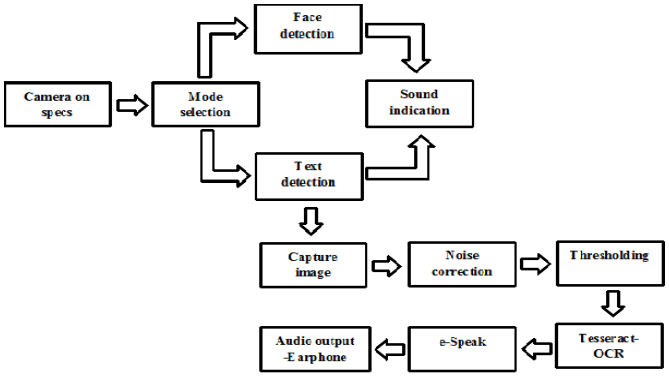
A TTS (or "engine") is composed of two parts a front-end and a back-end. The front-end has two major tasks, the normalization and phonetic transcription of text. Normalization, pre-processing, or tokenization of text is the conversion of text containing symbols like abbreviations and numbers into equivalent written-out words. The front-end then assigns phonetic transcription to each word. The prosodic units like clauses, sentences, and phrases are marked and divided. Text-to-phoneme conversion is the process of assigning phonetic transcriptions to words. The output from the front-end is a symbolic linguistic representation from the Phonetic transcriptions and the prosody information. The back-end performs the function of a synthesizer. The symbolic linguistic representation to sound conversion is achieved using this backend. The most attractive feature of a speech synthesis system are naturalness and intelligibility. The output sounds like human speech which describes the naturalness, the output is ease with the intelligibility of understanding. Speech synthesis systems usually try to maximize both natural and intelligibility which are the characteristics of an ideal speech synthesizer.

Block Diagram

The ac voltage, typically 220V rms, is connected to a transformer, which steps that ac voltage down to the

level of the desired dc output. A diode rectifier then provides a full-wave rectified voltage that is initially filtered by a simple capacitor filter to produce a dc voltage. This resulting dc voltage usually has some ripple or ac voltage variation.

A regulator circuit removes the ripples and also remains the same dc value even if the input dc voltage varies, or the load connected to the output dc voltage changes. This voltage regulation is usually obtained using one of the popular voltage regulator IC units.



IV. RESULTS AND DISCUSSION

The obtained output images after pre-processing are displayed below. The original image that was captured using the Pi Camera. Display the pre-processing done in each stage. And finally represents the image which is given as input to the OCR displays the text obtained at the output of the OCR engine. It is evident that the result is not completely accurate. This is because of the less resolution of the camera used. Better results can be obtained if the camera used is a High definition camera.

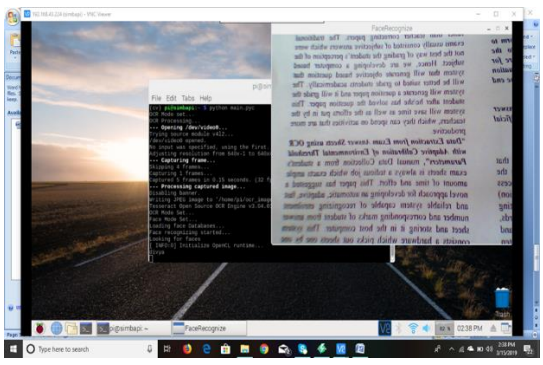


Fig.1. Text Recognition

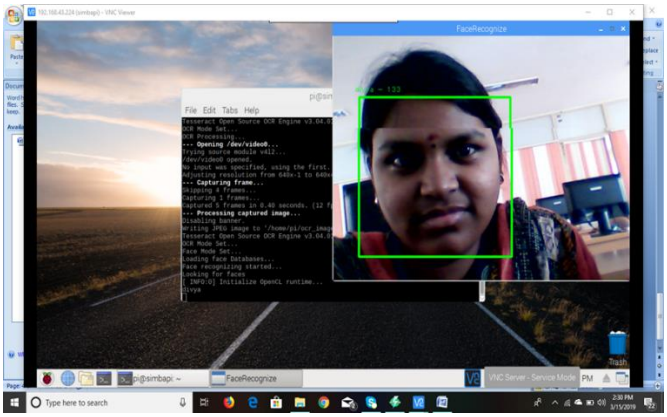


Fig.2. Face Recognition.

V. REFERENCES

[1]. Ms.Rupali, D Dharmale, Dr. P.V. Ingole, "Text Detection and Recognition with SpeechOutput for Visually Challenged Person",vol. 5, Issue 1, January 2016

[2]. Nagaraja, L., et al. "Vision based text recognition using raspberry PI." NationalConference on Power Systems ,Industrial Automation (NCPSIA 2015).

[3]. Rajkumar N, Anand M.G, Barathiraja N, "Portable Camera Based Product LabelReading For Blind People.",IJETT, Vol. 10 Number 11 - Apr 2014

[4]. Boris Epshtein, Eyal Ofek, Yonatan Wexler, "Detecting Text in Natural Scenes withStroke Width Transform."

[5]. Ezaki, Nobuo, et al. "Improved text-detection methods for a camera-based text readingsystem for visually impaired persons." Eighth International Conference on Document Analysis and Recognition (ICDAR'05). IEEE, 2005

[6]. Anusha A. et al., "Product Reading for Visually impaired Persons," International Research Journal of Engineering and Technology, Pp-ISSN: 2395-0072. Volume: 02 Issue: 04 | July-2015.

[7]. Van Erp J.B.F and van Veen, H.A.A.C,"A mobile cloud pedestrian crossing for the blind," IEEE, pp 405-408,vol.212,2017.

[8]. Prajakta S Rathod, "script to speech conversion using artificial neural network", IEEE,P ISSN: 2375-1282,2017

Cite this article as :

Divya S, Kowsalya K, Shalu Sunil, Subashini V, Suvetha V, Dr. Suguna N, "Object and Face Recognition Using Raspberry Pi", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 5 Issue 2, pp. 674-679, March-April 2019.
Journal URL : http://ijsrcseit.com/CSEIT1952239