International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 06 Issue: 09 | Sep 2019 www.irjet.net p-ISSN: 2395-0072

WEARABLE AI DEVICE FOR BLIND

Mugesh yadav. R[1], Dhivakar. G[2], Mohamed Rizwan. T[3], Aravindh. J[4], Vishnupriyanka. R[5]

[1,2,3,4]Dept. of Bio Medical Engineering, Rajiv Gandhi College of Engineering and Technology, Puducherry, India
[5]Assistant professor, Dept. of Bio Medical Engineering, Rajiv Gandhi College of Engineering and Technology, Puducherry, India

---

Abstract - *In this paper we proposes a method of camera based AI system for blind. This AI system we proper has three specification. Facial recognition, OCR based TTS system, object detection. All the specification are conveyed to blind through speaker. Face recognition specification works in the convolution neural network (CNN), linear discriminant analysis in which we use the above mentioned pattern and network for the training of face. Which is experienced using face libraries opencv.neural network is the problem solving for face indentification in which we use facial key point detector and gabor wavelet filtering detection systems purpose classifier to perform the object detection with high scoring points of the image are considered for detection of object. It divides the image into regions and bounding boxes for object detection.*

Key word - Optical Character Recognition (OCR), Text-To-Speech (TTS), Raspberry Pi, English text, Tesseract OCR Engine, Convolutional Neural Network(CCN), Linear discriminant analysis(LDA), Gabor filter, Feature extraction.

I. INTRODUCTION

OCR is used to convert text into ASCII form. It has wide scope in the field of research for the past few years. It gives alphabetical recognition of a printed characters or image. Most of the documents are in printed form of text such as reports, instruction and newspaper etc. Blind people cannot able to read newspaper, cannot identify currency notes, etc thus they are struggling in day to day life.

AI is the emerging technology which is still under research we use this AI for our proposed model for many security purpose and use facial recognition as biometric processes. In facial recognition we verify human face database to know the particular purpose identity.

CCN has been used for face recognition character recognition and object identification. An interestingly new approach based on a convolutional neural architecture which was presented in 2009 designed to detect highly variable face pattern. Even though there are many system lasting in object detection efficiency is less. We are trying to improve the efficiency of the system.

II. LITERATURE SURVEY

In a statistical structure research, it calculates the mathematical submission of noticed mixtures of extreme at specified roles compared with other picture. The pixel strength is categorized into first-order, second-order and higher-order research works. The grayish level co-occurrence matrix(GLCM) method is a way of getting mathematical structure. A GLCM is a matrix where the number of rows and columns is equal to the number of gray levels, G in the image. The matrix element P(m,n | Δx,Δy) is the relative frequency with two pixels, separated by a distance (Δx,Δy), occur in given neighborhood.

One pixel offset is used as reference pixel of neighbor . image pixel area is 0,0 occurs the number of times the left cell is filled.

The wearable recognition system is faster ,more responsive it should be prepared by advanced microcontroller to reduce computer like compactness.PIC is major factor offers development of system code using setup.

There are many wearable aids for are used to assist visually impaired people which divides into electronic travel aids(ETAs),electronic orientation aids(EOAs),and position locator devices(PLDs). Its all about individual visually impaired people. This survey is all about various past features and parameter f the performance its only for the past reference point of view and not critizied n those systems

III. FACIAL RECOGNITION

(A) linear discriminant analysis(lda)

LDA's main purpose is for pattern recognition, statistics which work in machine learning language to find the linear combination of the feature to same character or different one. It gathers new dimension which is high separation of the projected class and has less variance.

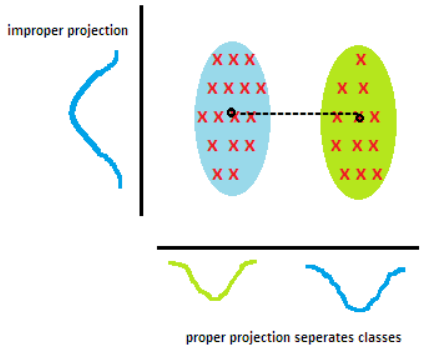

Fig -1: linear discriminant analysis.

(B) CONVOLUTIONAL NEURAL NETWORK

LPG is used to extract the feature of human face as a result it fetch input to the deep convolutional neural network. It involves layer by layer by training network. The core idea is to classify data of a sample at some distance. Three convolutional network are used in this network structure and in full connected layer two pooled layer are used.

(C) convolutional layer

CNN is the deep class of neural network visual image are analyzed most commonly it is multi colour perception. The convolution layer is the main block in CNN. It goes deeper when compared to other networks. The input and output of the convolution neural network must be of same size to a convolutional kernel of the convolutional layer. Convolutional kernels has the content of the convolutional layer. The convolution kernel has 5× 5 size.

(D) gabor filter

The texture analysis using linear filter. It specific frequency content of the image in a linear manner of localized point analysis. It has the strong response of structure of the image. It has the receptive field properties of neurons.

(E) pooling layer

Pooling layer is called as descending or decreasing layer. The pooling layer is classified into input layer, hidden layer and output layer. It reduces the spatial size of the representation and the amount of computation of the network. It is based on discretization process.

Max pooling is maximum value of pixels in particularly in small area.

IV. TEXT-TO-SPEECH

Our method has two modules: the first main process of acquiring image from web cam and converting It into text format using OCR(optical character recognition .the second process is language processing and signal processing converts text to speech using TTS (text-to-speech)synthesizer.

Steps:

- Image acquisition by cam
- Getting the image into axial panel for creating graphical user interface(GUI)
- Pre- processing the image (contrast leveling, RGB to gray image, adapation Accordingly)
- Then preprocessed image is converted to words that is text using OCR
- Then the text converts to speech which creates a MP3 file which automatically reads using speaker

It also analysis using edge detectors and neural networks.

V. OBJECT DETECTION

Object detection is one of the most challenging task. The object detection is done using feature extraction.

(F) feature extraction

Instead of raw pixel data usage, feature extraction provides easier classification. The image can be captured through the camera. Then image recognition can be done through features extracted from the data types of image determines the types of feature extracted. Now the feature extracted are in alpha numerical representation which can be further processed in feature extraction mainly detects the important area of the image and it is actual representation of an image.

VI. HARDWARE REQUIREMENT

The main hardware requirements used for connecting this system are:

- Raspberry pi 3 B+
- USB Web camera
- Speaker
- LCD

i. raspberry pi 3 b+

Raspberry pi 3 B+ is main component used in this system. It is the final version of raspberry pi 3.

Specifications:

- 1GB RAM
- 1.4GHz 64-bit ARMv8.
- CSI port for connecting a camera.
- Micro SD port for loading operating system and storing data.

Fig - 2: Raspberry pi

ii. usb web camera

USB web camera is used to capture the person's image for face recognition and for TTS system which captures text images.

- 24 mp camera.
- CMOS Light sensor.
- 10x zoom.

Fig -3: Usb web cam

iii. speaker

Speaker is used to receive ouput in audio format.

iv. lcd screen

LCD screen is used to control the interface between the specification in the system.

specifications:

- SPI interface.
- Colour LCD display.
- Pixel resolution 128x128.

Fig -4: Lcd screen

VII. SYSTEM DESCRIPTION

The project aims at the methods involved for detection in log-gabor filters. In this system facial changes due to variation in expression and pose can be easily obtained. The whole face image are applied by global feature extraction.

In this framework two main modules are involved. The first stage involves images can be picked up by web camera and then the images can be converted into text document with the help of optical character recognition. The second stage involves converting the digital images or text document into text to speech synthesis(TTS). With the help of the speaker the audio message can be clearly identified.

Figure(5), shows the block diagram for facial recognition detection and text to speech synthesis.

BLOCK DIAGRAM:

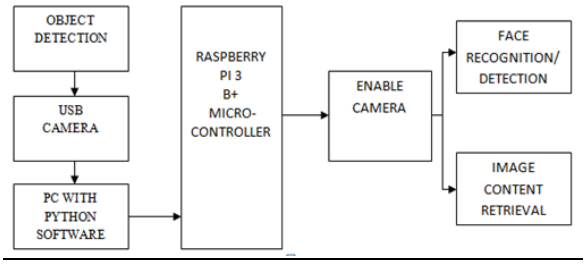

Fig -5: Block diagram

a) **text to speech synthesis**:

Speech synthesis is used in computer system. Speech synthesis is first implemented in software and then it is implemented in hardware. A text to speech converts the text into speech. There are about 16 synthesized speech can be created. For high quality output the storage of entire words can be of specific usage domain.

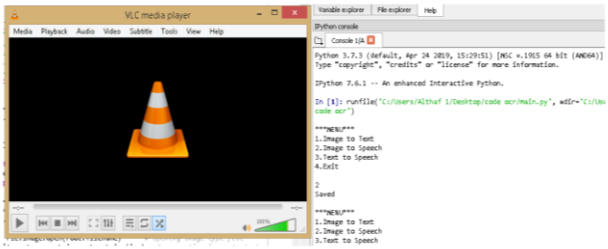

Fig -6: Output of TTS, audio file generated.

b) face detection:

In the face detection the camera will easily find the face. In an image the face can be easily identified.

Facial recognition can be easily extracted, cropped and resized and then it is converted to gray scale .Gray scale which is best to describe an image.

c) optical character recognition:

OCR follows a modification of an architecture. OCR first stage is preprocessing. The stages of preprocessing is image recognition. In the first part of object detection the text, graphics and the other parts are separated. Feature extraction is used to create the image, scaling and then the line thickness of the image. The post-preprocessor is used to improve accuracy of detection of an image.

Fig -7: Stimulation using raspberry pi(face recognition and TTS)

VIII. RESULT AND DISCUSSION

In the detection of the facial image related to the user it first compares the input image for facial with the facial image which have been required for authentication. The comparison is between 1x1. From the dataset to find the user the face matching can be done. The comparison is between 1xN.In this project we developed a device that converts text images to speech. By using a facial recognition method can be done. It is then implemented by using Raspberry Pi. The series of image pre-processing is done and it only contains the text, the background colour will be removed. We are using two methods to convert the images to speech. They are TTS system and OCR methods has been used. The audio output is heard through the speaker.

IX. CONCLUSION

This system helps the visually impaired individual helping them identify faces from the trained face. This frame work is portable and easily carried by the visually challenged people. An individual can able to read the text by using the TTS engine using optical character recognition.

X. FUTURE SCOPE

This system can be controlled using motion gesture. In future we can also use this using like alexa, which works in voice command. More compact format of the device.

XI. ACKNOWLEDGEMENT

The authors would likely to thanks Ms. R. Vishnu priyanka, Assistant professor, Department of Bio-medical engineering, at Rajiv Gandhi College of engineering and technology for her kind support, constant encouragement and development towards the project.

XII. REFERENCE

[1] P. Mohanaiah, P. Sathyanarayana and L. GuruKumar, "Image Tex-ture Feature Extraction Using GLCM Approach", International journal of scientific and research publications, Vol. 3, Issue 5, May 2019

[2] D.Velmurugan, M.S.Sonam, S.Umamaheswari, S.Parthasarathy, K.R.Arun[2016]. A Smart Reader for Visually Impaired People Using Raspberry PI. International Journal of Engineering Science and Computing IJESC Volume 6 Issue No. 3.

[3] K Nirmala Kumari, Meghana Reddy J [2016]. Image Text to Speech Conversion Using OCR Technique in Raspberry Pi. International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering Vol. 5, Issue 5, May 2016.

[4] Nagaraja L, Nagarjun R S, Nishanth M Anand, Nithin D, Veena S Murthy [2015]. Vision based Text Recognition using Raspberry Pi. International Journal of Computer Applications (0975 – 8887) National Conference on Power Systems & Industrial Automation.

[5] Minjun wang, zhihui wang, jinlin li. Deep convolutional neural network applies to face recognition in small and medium databases [2017]. International conference on systems and informatics.

[6] Jun ou, xiao-Ba Bai, Yun Pei, Liang Ma, Wei Liu [2010]. Automated facial expressin recognition using gabor filter and expression analysis.