International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-8 Issue-5S, May 2019

Real Time Solution for Prosopagnosia

Abhilash Manu, Aravind H S

Abstract— ***The paper aims at building a prototype for solving the problem of a rare neurological disorder 'Prosopagnosia'. It is also called face blindness / facial agnosia. This is a behavioral disorder of face perception that impairs the ability to detect familiar faces, which include one's own face, while other forms of visual processing and intellectual functioning remain intact. The Extensive research has indicated 1 out of 50 people may have this neurological disorder. In order to help the significant number of affected people overcome this difficulty, we have built a prototype which uses a camera to capture the image and an appropriate face recognition code using the Histogram of Oriented Gradients (HOG) approach is implemented for face detection. After detection, SVM classifier is used for classification and the name of the identified person will be displayed. Simultaneously, the conversation is being recorded and text mining is performed to extract the keywords of the conversation. The result is displayed on a suitable interface. The hardware module Raspberry Pi is used as a processor for processing the incoming image and audio data.***

Index Terms— ***Prosopagnosia, Raspberry Pi, HOG, Healthcare applications, Internet of Things.***

I. INTRODUCTION

Prosopagnosia is a neurological disorder indicated by the inability to recognize faces. Some people with severe prosopagnosia may only have difficulty recognizing a familiar face, while others may not even be able to distinguish a face from different from the object. In some cases, patients could not recognize their own face.

Many people manage this disability and also develop compensatory methods to assist them to function effectively in their daily life. While others will have a greater impact on daily functioning few of them avoid social interactions, encounter interpersonal relationship problems or damage their career, and even report episodes of depression.

The large number of people suffering from this condition drives us to help find a solution. Since there are limited advancements in the medical field, it is of at most importance to develop a real time solution for this disorder. The currently available face recognition methods and technologies are either extremely platform specific or are considered to be very "heavy" since they require extensive computational power and time.

Manuscript published on 30 May 2019.
***** Correspondence Author (s)
Abhilash Manu, Solution Architect Products, SCS Accenture Bengaluru
Dr.Aravind H S, Professor, ECE Department,JSSATE, Bengaluru

II. LITERATURE SURVEY

For Face Recognition

The face detection approach by IhorPaliy[1] uses both the Haar-like features and the CNN for the process of verification. This is a near real-time mode with the negligible false alarms rate. Li Cuimei1, etal.,[2] projected algorithm for new human face detection along with primitive Haar cascade algorithm using additional classifiers is proposed. Here skin hue histogram matching, mouth detection and eyes detection are considered for classification. At last for reducing false positive rate a mouth detection operation is used. Finally results are obtained by considering different images of people with some degree of orientation and rotation. Results shows that it works better and is effective and simple using OpenCV.

To examine the question of feature sets queries for robust visual object recognition by Navneet etal[4], were taken linear SVM based human detection as a test case. They demonstrated experimentally that grids of Histograms of Oriented Gradient (HOG) descriptors consistently outperform existing feature sets for human detection after analyzing existing edge and gradient based descriptors.

J. Kulandai Josephine Julina and T. Sree Sharmila [5] presented paper to check the given face input is a registered person or not in the database. Histogram of Oriented Gradients (HOG) technique is used for Face identification. SVM is also used along with HOG. The matching between the faces is checked with different conditions. The outcomes indicates good detection accuracy.

For speech recording

For this, a popular open source large vocabulary continuous speech recognition system, is utilized with hand-held devices. On an average, this system is 8.03 times faster compared to other systems.

Implementation

The proposed prototype will include the Raspberry Pi 3 Model B, a single board computer, to perform all the image and audio data processing. The image input to the system will be obtained from any Android employing platform such as a mobile or a tablet while the audio input to the system will be through a mini-microphone connected to the Raspberry Pi 3. Once the computation is completed, the name and text mined version of the previous conversation is displayed on the Android device. Simultaneously, the actual conversation that was recorded is played back through a loudspeaker or earphones.

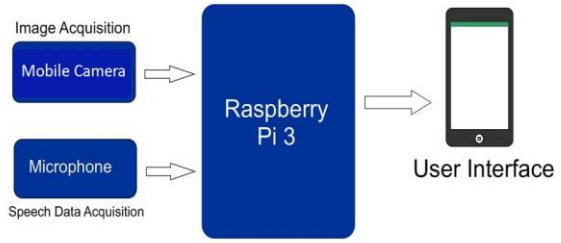

Fig.1 Block diagram of the prototype

III. HARDWARE AND SOFTWARE

1. Raspberry Pi:

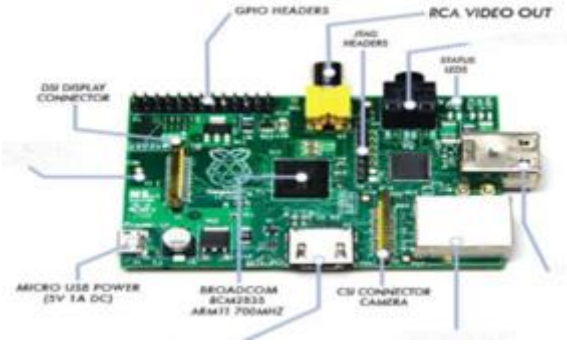

Fig.2 Raspberry Pi

Raspberry Pi is a single board computer with a compact in size, weighs around 50g and requires power rating of 5V, 700mA and it is also cost effective.
Since Raspberry It is small sized, powerful processor capable of handling complex operations, it is used for processing the image and audio data.

2. USB Microphone

A microphone, is a transducer that converts sound into an electrical signal. In the prototype USB microphone is used to record the conversation.

3. Loudspeaker (or Earphones) and Stereo Amplifier

A loudspeaker converts an electrical audio signal into a corresponding sound. This can be replaced by an earphone depending on the application. It is used to output the previously recorded conversation during the subsequent meeting with the same individual.

4. Virtual Network Computing(VNC)

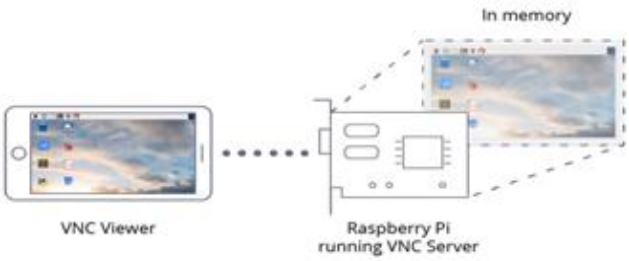

Fig.3 USB Microphone

It is a graphical desktop sharing system used to remotely control another computer. It transmits the signals from keyboard and mouse of one computer to another computer and also transmits back the graphical signals from the screen in the opposite direction. Virtual Network Computing is a platform for clients and servers. It can connect multiple clients at the same time.

Fig.4 VNC Viewer

5. IP Webcam

It converts the android phone device into a network camera with multiple viewing options. It can view the mobile camera from any web browser from around the world. Without internet access, IP Webcam can stream a video inside a Wi-Fi network. For instant global access cloud broadcasting is supported.

Face recognition methodology

STEP 1: FINDING ALL THE FACES

To find faces in an image, the image is converted to grey scale and later each pixel in the image is examined one after other. In single pixel, we need to look at the pixels that directly surrounding it. Here image is divided into 16x16 square pixels. Then the count in square is replaced with direction of arrow which are very much strong. The test image which is similar to a HOG pattern is known and is from bunch of other training faces is found.

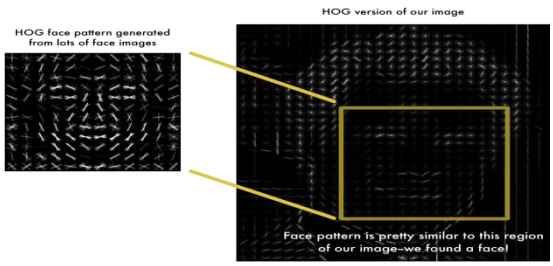

Fig.5 HOG pattern extraction

Step 2: Posing and Projecting Faces

Once the faces in a test image are isolated, the problem that a face that is turned in a direction being considered as a different face is addressed. To resolve this, each picture is wrapped such that the eyes and lips are always in the same place in the image. In total 68 particular points will be considered for identification of the face of a person. The machine learning algorithm is trained later to find these 68 particular points on any face (i.e edge of each eye, inner edge of each eyebrow, etc.) which is extracted from the algorithm.

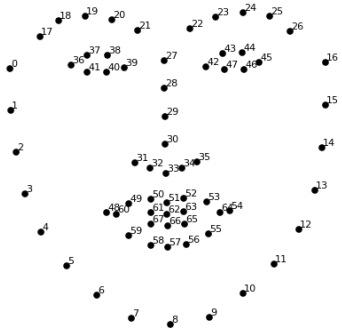

Fig.6 Sixty-eight landmarks location

Fundamental image transformations like scaling, rotation that preserves the parallel lines will be used.

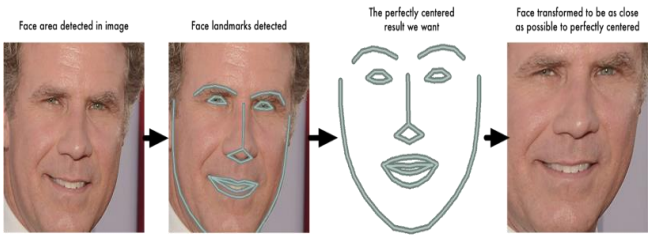

Fig.7 Projection of Faces

In the image even though face is turned, it is possible to keep eyes and mouth in the center position.

Step 3: Encoding Faces

The algorithm will able to recognize faces in milliseconds. It is needed to extract a few measurements from each face. Later, the unknown face is measured in the similar method and matched with the known face present with the closest measurements. For example, measurements of the size of ears in a face, the distance of the eyes, the length of the nose, etc.

Step 4: Finding the person's name from the encoding

It is the final step to find the person, who is having closest measurements in the existing database which is achieved by using machine learning algorithm. It uses simple linear SVM classifier. The result is the name of the person which is obtained from the classifier.

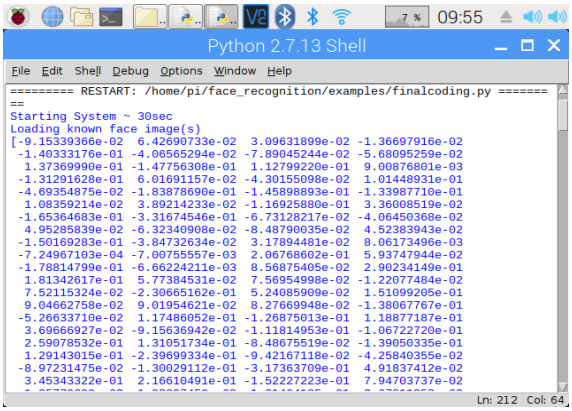

Fig.8 128 Measurements extracted

Natural Language processing (NLP)

NLP primarily orients to deals with the artificial intelligence and computer science based idea of interactions between computers and human language. it is particularly focused on processing large amounts of language data.

Speech to Text Conversion

The Google Cloud Speech-To-Text that is used here allows the functionality to perform audio (example speech) to text conversion with the help of neural networks which are provided with user-friendly API. This also supports 120 different languages and also provides the utility of pre-recorded or real-time recording of audio.

Text Mining

This is the technique of obtaining information that pertains with high quality from any given text. Here "High quality" refers to text that is of some interestingness, novelty, and relevance. This typically means that trends and patterns are observed and thus statistical patterns are derived. Additionally, it also involves structuring the data in the input text into patterns followed by evaluation and interpretation of the obtained data.

Histogram of Oriented Gradients (HOG)

For detection of object, HOG is used. The first stages of this approach use a global image normalization equalization technique aimed at reducing the effects of illumination. The image gradients belonging to the first order are then computed in the subsequent stage. Contour, silhouette, texture details are the primary information that is obtained from the second stage. In addition, further resistance to illumination is also provided. The next stage is used to obtain encodings that are responsive to local image content, but also provide resistance to changes to in appearance or pose. The normalization, i.e. local groups of cells and contrast, is performed in the fourth stage, before passing the result to the next stage. The fifth and final stage involves in the collection of all the obtained HOG descriptors from every overlapping dense blocks grid in the detection window to a combined feature vector. This will be used as a window classifier.

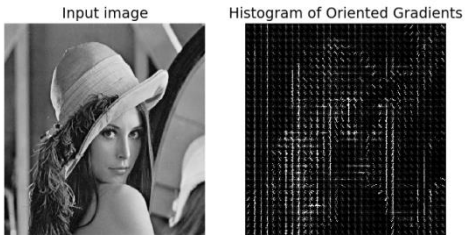

Fig.9 Test image and the corresponding HOG pattern

Support Vector Machines (SVM)

SVM utilizes machine learning as managed learning models working along with algorithms that operate on data that is used for classification and regression analysis. In reference to training examples, each referred to one category or another, the SVM algorithm that is used for training, assigns new examples to any of the already referenced categories. The SVM model functions by mapping each example to a point in space, where these examples belonging to different categories are differentiated by a clear distance, placed as far away as possible.

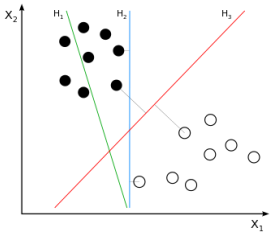

Fig.10 Hyperplanes for Classification

IV. TEST CASES AND RESULTS

Test Case 1: Similar Faces

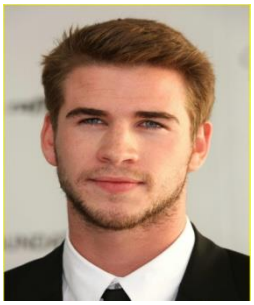

Fig.11 Trained Image 1 Fig.12 Trained Image 2

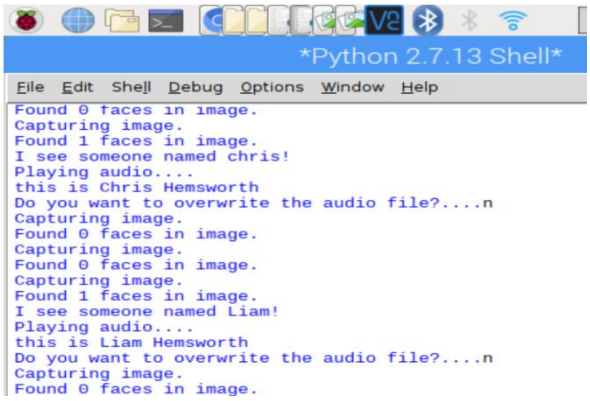

Fig.13 Case 1 Results displayed on the console

Result: It is observed that the prototype is cable of distinguishing between the two similar faces. Even though the resemblance between the two faces is extremely high, the prototype is capable of differentiating the two individuals.

Test Case 2: Variable Lighting Conditions

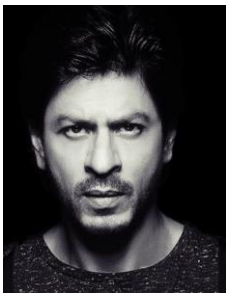

Fig.14 Trained Image Fig.15 Test Image

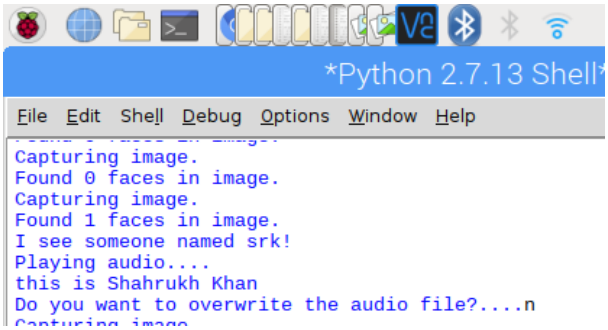

Fig.16 Case2 Results displayed on the console

Result: The developed system is able to recognize a face when it is trained with a colorful and illuminated image (Fig.13) and is tested with a low light image of the same person (Fig.14).

Test Case 3: Facial features are not discernable

Fig.17 Case 3 Trained Image Fig. 18 Case 3 Test Image

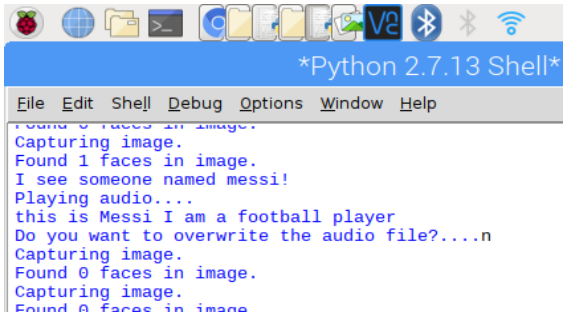

Fig.19 Case 3 Results displayed on the console

Result: The prototype is able to recognize a face when tested with an image of the same person, but with significant features of the face being blocked. (By spectacle and beard). Here, the system is trained with image (Fig.16) where all the facial features are clearly visible and is then tested with image (Fig.17) where eyes, cheeks and chin are not clearly visible.

Test Case 4: Delay Measurement

This test case is to measure the amount of initializing delay that is incurred in the system as the database size grows by the addition of multiple faces.

Retrieval Number: E10530585S19/2019©BEIESP
DOI:10.35940/ijeat.E1053.0585S19
Journal Website: www.ijeat.org

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

Size of Database	Delay
6 images	Trial 1: 1:26 mins Trial 2: 1:23 mins
10 images	Trial 1: 5:48 mins Trial 2: 5:52 mins

Table 1: Delay measurement

As observed, there is an exponential increase in the time taken to train the images as the number of images grows. But, the primary advantage of this system is that, this operation needs to be performed only in the initial stage, i.e. when the system is powered on. No such delays are experienced during the real time operation.

Test Case 5: Multiple faces

In this test case the system is testing for its accuracy to detect multiple faces in an image.

Fig.19 Case 5 Multiple faces

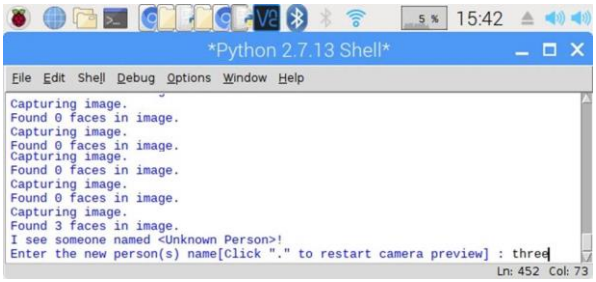

Fig.20 Case 5 Results displayed on the console

V. CONCLUSION

The developed time prototype is an optimal solution for Prosopagnosia. It is tested for a wide range of cases to outline the versatility and efficiency of the prototype. The feasibility of the prototype in terms of speed and accuracy is an ideal approach to assist the affected patients. Since Raspberry Pi is small sized, powerful processor capable of handling complex operations, it is used for processing the image and audio data. The Pi module is connected to a smart phone through a wireless LAN connection to enhance the mobility of the system. Hence, the result will be displayed to the patient on the smart phone with all the computations taking place on the Raspberry Pi processor. Miniaturization can be further enforced to add to the already incorporated portability aspects.

VI. FUTURE SCOPE

With advancements in image processing, real time object detection can also be incorporated in this approach to provide assistance to patients who are visually impaired. The whole prototype can be made completely wireless as portable device become more capable of performing more complicated computational tasks.

ACKNOWLEDGEMENT

We are hereby grateful for the support extended by Dr.Omkar.S.N (Principal Scientist Indian Institute of Science) and the effort by the students of JSS Academy of Tech Edn Anusha U, Arijith Amby S, Ganesh Patidar and J Karthik in making this prototype build a successful one.

REFERENCES

1. IhorPaliy, VolodymyrDovgan, OgnianBoumbarov, Stanislav Panev, Anatoly Sachenko, YuriyKurylyak, Diana Zagorodnya, "Fast and robust face detection and tracking framework", Intelligent Data Acquisition and Advanced Computing Systems (IDAACS) 2011 IEEE 6th International Conference on, vol. 1, pp. 430-434, 2011.
2. Jiajian Zeng, Siyuan Liu, Xi Li, DebbahAbderrahmane Mahdi, Fei Wu, Gang Wang, "Deep Context-Sensitive Facial Landmark Detection with Tree-Structured Modeling", Image Processing IEEE Transactions on, vol. 27, pp. 2096-2107, 2018, ISSN 1057-7149.
3. Weiming Yang, Meirong Zhao, Yinguo Huang, Yelong Zheng, "Adaptive Online Learning Based Robust Visual Tracking", Access IEEE, vol. 6, pp. 14790-14798, 2018, ISSN 2169-3536.
4. Alexander Schmidt, Antoine Deleforge, Walter Kellermann, "Ego-noise reduction using a motor data-guided multichannel dictionary", Intelligent Robots and Systems (IROS) 2016 IEEE/RSJ International Conference on, pp. 1281-1286, 2016, ISSN 2153-0866.
5. XinJin, Peng Yuan, Xiaodong Li, Chenggen Song, Shiming Ge, Geng Zhao, Yingya Chen, "Efficient privacy preserving Viola-Jones type object detection via random base image representation", Multimedia and Expo (ICME) 2017 IEEE International Conference on, pp. 673-678, 2017, ISSN 1945-788X.
6. Volker Eiselein, HajerFradi, Ivo Keller, Thomas Sikora, Jean-Luc Dugelay, "Enhancing human detection using crowd density measures and an adaptive correction filter", Advanced Video and Signal Based Surveillance (AVSS) 2013 10th IEEE International Conference
7. K Zhang Z Zhang Z. Li "Joint face detection and alignment using multitask cascaded convolutional networks [J]" IEEE signal Processing Letters vol. 23 no. 10 pp. 1499-1503 2016.
8. P Voila M J. Jones "Rapid object detection using a boosted cascade of simple features [C]" Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (ICCVPR 2001) vol. 1 pp. 1-511-1-518 8-14 Dec. 2001.
9. Dalal, N. and Triggs, B., "Histograms of Oriented Gradients for Human Detection," IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005, San Diego, CA, USA.

AUTHORS PROFILE

Mr Abhilash Manu , is a AI lead at Industry Innovation Hub, Accenture and He completed his B E in Instrumentation and Electronics from JSSATE, Bangalore. His fields of interest are Artificial Intelligence and Communication. He has published many papers in reputed journals in the field Artificial Intelligence.

Dr. H S Aravinda, is a professor and Head of the department of ECE in JSSATE, Bangalore. He has obtained his Ph D degree from Visvesvaraya Technological University, Belagavi, India. His fields of interest are Signal Processing and Communication. He has published many papers in reputed journals in the field of signal processing. Presently, he is guiding six research scholars.