Person Identification Using Face and Speech Recognition for Visually Challenged with Mask Detection

Anitha R [a,1], Rakesh Gupta G [b], Manoj V [b] and Bhargav M [b]

[a] *Associate Professor, Koneru Lakshmaiah Educational Foundation, Vaddeswaram*
[b] *Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram*

Abstract. A system and method for facilitating a visually impaired person for identifying a person. The method includes the step of storing a plurality of instructions for facilitating the visually impaired person identify the person in front of them by their face and/or voice characteristics by updating our project with Mask Detection using OpenCv and keras. It includes the step of receiving voice signals from the person present in surrounding of the visually impaired person and includes the step of capturing the pictures of a particular person and their surroundings of the visually impaired person and storing the processed data into the database or any storage devices. The data will be processed to AWS server or any local storage for processing and determining the person with the help of the database we already have. After processing and identifying the person with the help of face and voice recognition modules the name is sent to the visually impaired user's phone in the form of a text message which will be read aloud by his phone's virtual assistant.

Keywords. Visually impaired, Voice signals, Database, Voice recognition, Virtual assistant, Mask Detection.

1. Introduction

This Structure goes likely an Briliant specialist which helps in carrying the earth nearer and nearer to the environment a supported spot over living for Visually obstructed individuals. The Person is seen, and the personality of the individual is gotten the message out about for the customer, which is another case that would be an issue to recognize and there is a probability of getting cheated and this ensures security and prosperity by alerting the user. The face acknowledgment (also simply called as face Recognition) [1, 2] is the one able to recognize and see a person by their facial characteristics. The face is multidimensional and consequently requires a ton of mathematical computations. Face affirmation system is essential and huge for giving security, mug shot planning, law approval applications, customer affirmation, customer access control, etc., and is commonly used for affirmation for various applications. These all applications require a compelling Face affirmation system.

---

[1] Anitha R, Department of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswaram, India.
E-mail: anitharaju15@gmail.com.

Figure 1. Face Recognition System (Courtesy from google)

1.1. Face Recognition

In Face recognition system, a predictable picture is given to the machine or framework, and through component extraction measure that image is checked against the information or simply the database as of late put to the side in the construction. Later the highlighted pictures will be taken out from the database existed.
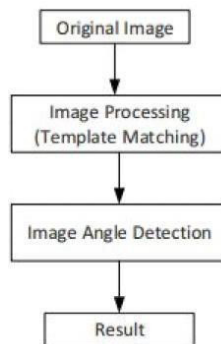

Figure 2. Workflow of Face-Recognition

In Face detection [3], a computerized picture is given to the system and through extraction in view of the highlights, that image is looked against the information base which is now put away in the framework. Features of the test picture will at that point be compared with the highlights of the information base. The least variation in the highlights will help in distinctive the profile of an individual in the test picture. Face identification was finished by utilizing "**Haar cascade classifiers**", that distinguished the appearances, and afterward the recognized countenances were being put away by utilizing python orders in an information base. Preparing was done so the photographs inside the information base were changed over to yml record for additional utilization in attestation measure. At last, assertion measure was finished by utilizing LBPH while keeping be cautious with a yml report made during setting up this cycle, if the picture of the individual before the camera created with the information base records, which construed that the capability between of two pictures didn't gather quite far, the entryway lock would open or, point of fact alert would begin.

1.2. Speech Recognition

Discourse acknowledgment has two primary factors that influence the exhibition for example voice acknowledgment and order acknowledgment. At the point when the framework was prepared with the voice of individual A, it was difficult for the framework to look at the voice orders [4–6]. In offline system development of framework v3 module was utilized for SST which means transferred from discourse to text and TTS which means text is transferred to discourse which in addition offers an inherit port for the receiver. It can basically store 80 requests, so it isn't inside and out ideal to anticipate comparative requests with various voices, so a refreshing was made through setting up the framework with standard highlight and a customary tone. To avoid disillusionment as a result of much disturbances and contortion in the ecological variables, planning is done into some degree disorderly temperament to evade dissatisfactions while working with bad quality progressing tone orders and a substandard standard mic is used to get most extreme tone.

1.3. Mask Detection

This Face Mask Detection uses Artificial Networking to identify whether the person who came near to the person wears a mask or not. The Machine can be associated with any current or new mask recognition cameras to identify individuals with or without a cover. This Mask Detection has some special properties of detecting a person whether he/she wears a mask or not in a single or Group Environment. As a further process we are going to say for visually impaired person that the person who came near to him/her wears a mask or not.

2. Literature Survey

Throughout the most recent ten years, security of a Home and computerization frameworks is been picking up fame due to the improvement towards science field and the associative dangers of breaking in the framework. This thing imagines a sharp secretly mastermined on the two free and arising progressions through face affirmation, for the rationale of safety, and talk certification, for the robotization purposes. Likewise, the furthest down the line work done is examined quickly and why a withdrew structure was a pressing need to fill the opening in the constant progression shows presented in the related field [3, 7].

The related study presents an acknowledgment structure, will be useful for a visually challenged individual. The hand motion recognition framework and face acknowledgment framework has been prepared by specific calculations. In the Hand signal system, Skin concealing disclosure was completed in an YCbCr concealing space, in order to identified hand lifted distortion character motivation behind the hand was used where different features which is similar as fingertips, the point in the middle of fingers are being isolated. As indicated by banner Recognized, different undertakings can be performed like turning off lights and fans. The hand signal was seen with an exactness of 95.2% was refined certification was finished with an accuracy of 92%.

Expecting cover examination doesn't particularly think about, it can change the display truly the vast majority of the refined face acknowledgment techniques. Less amounts

of facial highlights in the masked face cause challenges than other ordinary face acknowledgment procedures. In this way, the exactness pace of acknowledgment is diminishing. That is the reason the veiled face is being one of the majors concerned elements inside the space of face acknowledgment. Then again, the use of a profound learning network is more difficult on the grounds that the amount of preparing information isn't adequate to prepare the profound learning networks [8] for this application which powers to utilization of move learning. Lastly, this check cycle is solidified to perceive up-and-comers face by playing out the arrangement task inside a brought together Support Vector Machine (SVM) [9–12].

3. Existing Technologies

3.1. Be My Eyes

It is a mobile application which doesn't cost the user on installation. It associates outwardly weakened people with a volunteer through video talk for help with regular day-to-day activities. These volunteers can assist the needy people with checking the date of
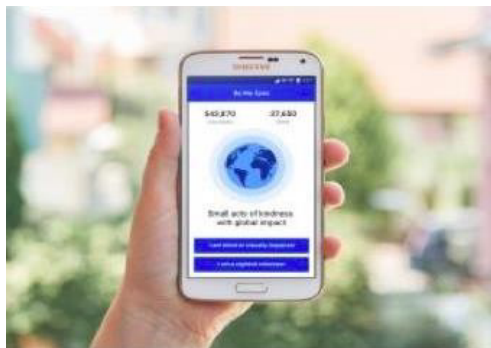

Figure 3. Be my Eyes application (Courtesy from google)

expiry on a container of milk or sort out what outfits to wear. It's an ideal method to associate the two completely different worlds and exchange knowledge from one another.

3.2. Brain Port

Brain Port uses electro-tactile method to assist blind people with direction, portability, and object identification. Accessible by solution, Brain Port involves a couple of camera-equipped shades that can be connected to a gadget that is set in the person's mouth. The gadget makes moving impressions or designs on the person's tongue that helps them in understanding or identifying the entity before them.

Figure 4. Brain Port application (Courtesy from google)

3.3. The Eye See

This was developed by undergrads, the Eye See is a protective cap model intended to help the blind people to "see" the surroundings. Also, it not only helps the users to describe the entities and people but also produces an alarming sound when the user is excessively nearer to an obstruction, making life safe and more secure.

Figure 5. The Eye see application (Courtesy from google)

4. Methodology

4.1. Speech Recognition

Talk affirmation, or talk to-message, is the limit with respect to a program or machine to recognize words communicated so anybody may hear and change over them into intelli-
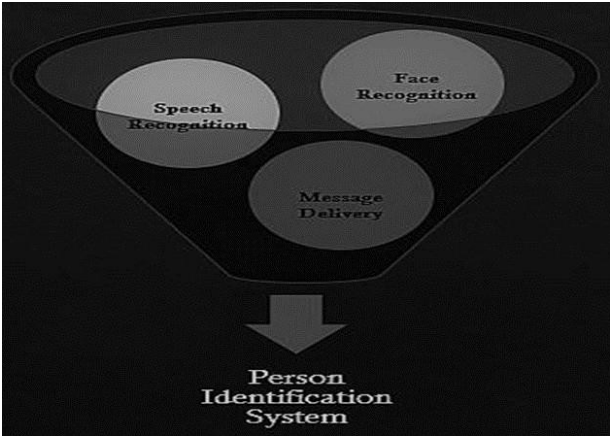
Figure 6. Design of Proposed System

gible substance. Basic talk affirmation programming has a limited language of words and articulations, and it may perhaps recognize these if they are spoken unmistakably. More unpredictable programming can recognize basic talk, different accents, and vernaculars. Talk affirmation works using figuring's through acoustic and language showing. Acoustic exhibiting addresses the association between semantic units of talk and sound signs; language showing matches sounds with word progressions to help perceive words that sound similar
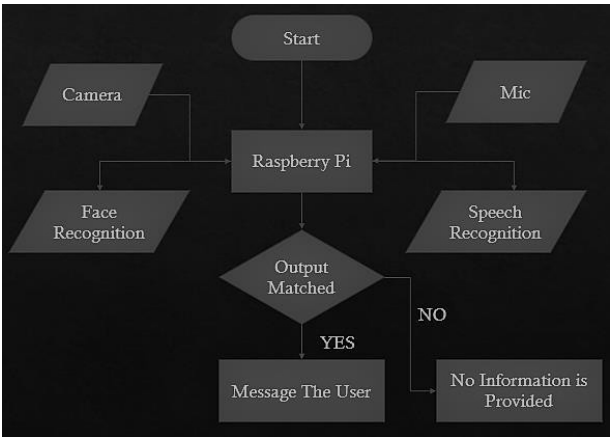

Figure 7. Workflow of Proposed System

4.2. Face Recognition System

This segment manages PCA and LDA for face acknowledgment. To begin with, we need a lot of Databases of pictures using which the figuring will be attempted. AT&T dataset pictures will be utilized. The underlying advance is to scrutinize all the dataset pictures that will be used for setting up the explanation [13]. In this section, we plate all the

mathematical advances that will be executed on the microcontroller called as Raspberry Pi 3 embedded system board to realize the face affirmation structure on it. This portion is isolated into two segments for instance Preparing stage and the Recognition stage.

4.3. Mask Detection

In this part, we will examine Face discovery or detection which has different modules that go inseparably to make the structure run suitably. The underlying stage contains snapping the photo and Detecting the face in it. Pictures can be gotten dynamically from a USB webcam related to Processor or Machine .The duty of the face acknowledgment module is to recognize whether a face is available during the continuous application. The face recognition is finished by looking at an image and finding some example that proclaims if a face is available in the picture. We built up the face cover identifier model for identifying if individual is wearing a veil. We have prepared the model utilizing Keras with network design. This Mask Detection module has some special properties of detecting a person whether he/she weared a mask or not in a single or Group Environment. As a further process we are going to say for visually impaired person that the person who came near to him/her weared a mask or not. This Machine Works efficiently when you develop mask/no mask dataset. This dataset is required to train the machine for accurate detection and Recognition.
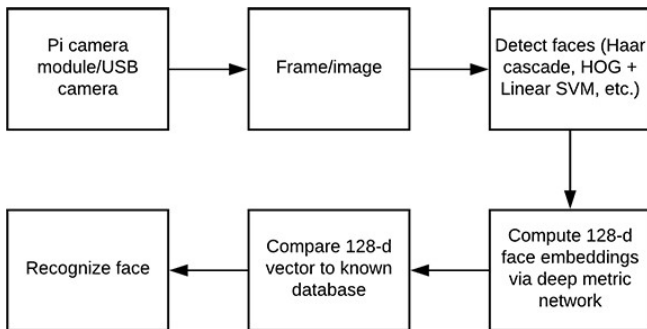

Figure 8. Design of Mask Detection.

5. Result an Analysis

Distance from the camera method is utilized to check whether an individual is moving toward the camera or disappearing. As demonstrated in the figure, decreasing distance shows the individual is moving toward the camera, and face detection can be set off. Eye line detection finds out the valley in level histogram projection. In the event that the eye line is distinguished, face detection can be applied to check whether the individual is wearing a mask or not. As it tends to be seen, in the situations where the individual isn't wearing a cover, face detection and facial part location occur appropriately. At the point when an individual is wearing a mask or face is covered with a hand or fabric, location of the face probably won't happen in specific cases, or face recognition will occur however either nose or mouth won't be recognized showing it is a cover. Eye discovery discov-
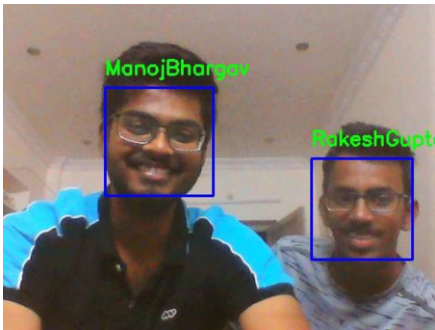
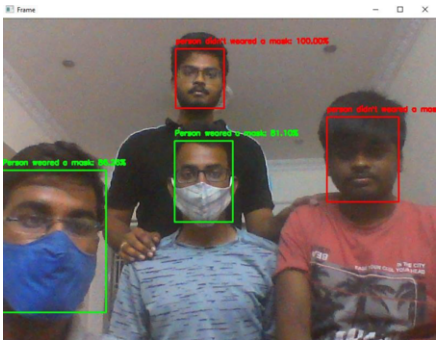**Figure 9. Sample output of Machine Recognizing two or more people**

Figure 10. Mask Detection for multiple People Simultaneously

ers eyes and afterward triggers face identification. At the point when an individual isn't wearing a veil, eyes will be distinguished and face discovery can be applied as demonstrated in Figure. In the cases, when an individual is wearing a cover, eye location returns valid yet face identification returns bogus demonstrating it is a mask. The false identi-

Figure 11. Accuracy of a Machine

fication [13] rate is greatest in the eye line recognition calculation and afterward in eye discovery. This is on the grounds that eye recognition and eye line discovery recognize little parts in the picture and for pictures with helpless goals; this location won't be precise promoting bogus identifications. The execution season of facial part recognition is greatest contrasted with any remaining strides as it manages face identification followed by face parts location which is an unpredictable calculation. To examine the exhibition of these four stages, four recordings were tried; with veil and without cover, and precision was determined as (Number of True identifications/Expected detection) and % of exact-

ness was determined. For distance from the camera step, the precision of the individual moving toward the camera is registered

6. Conclusion

Hence, by using the IOT and implementing its use cases like Security, Media and Automation, we can provide a physically challenged person with identifying a person who is standing Infront of him. Meanwhile in the process of updating our process it is also going to find whether a person is wears a mask or not which plays a major role in every days life. This helps in his security and helps him make sure that he is not cheated on by some random people and this would make a lot of change to people who are challenged and also helps in social distancing.

References

[1] Sahani M, Nanda C, Sahu AK, Pattnaik B. Web-based online embedded door access control and home security system based on face recognition. In2015 International Conference on Circuits, Power and Computing Technologies (ICCPCT) 2015 Mar 19 (pp. 1-6). IEEE.

[2] Zhang H, Qu Z, Yuan L, Li G. A face recognition method based on LBP feature for CNN. In2017 IEEE 2nd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC) 2017 Mar 25 (pp. 544-547). IEEE.

[3] Berbar MA, Kelash HM, Kandeel AA. Faces and facial features detection in color images. InGeometric Modeling and Imaging–New Trends (GMAI'06) 2006 Jul 5 (pp. 209-214). IEEE.

[4] Michaely AH, Zhang X, Simko G, Parada C, Aleksic P. Keyword spotting for Google assistant using contextual speech recognition. In2017 IEEE Automatic Speech Recognition and Understanding Workshop (ASRU) 2017 Dec 16 (pp. 272-278). IEEE.

[5] Kumar M, Shimi SL. Voice recognition based home automation system for paralyzed people. International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE). 2015 Oct;4(10):2508-15.

[6] Kannan K, Selvakumar J. Arduino based voice controlled robot. International Research Journal of Engineering and Technology (IRJET). 2015 Mar;2(01):49-55.

[7] Cuimei L, Zhiliang Q, Nan J, Jianhua W. Human face detection algorithm via Haar cascade classifier combined with three additional classifiers. In2017 13th IEEE International Conference on Electronic Measurement & Instruments (ICEMI) 2017 Oct 20 (pp. 483-487). IEEE.

[8] Wang M. Research on artificial intelligence-based mountain mineral composition detection and big data chemistry education. Arabian Journal of Geosciences. 2021 Aug;14(15):1-9.

[9] Naveena M, HemanthaKumar G, Prakasha M, Nagabhushan P. Partial face recognition by template matching. In2015 International Conference on Emerging Research in Electronics, Computer Science and Technology (ICERECT) 2015 Dec 17 (pp. 319-323). IEEE.

[10] Raut PP, Borkar NR, Student ME, Gawai SK. Techniques and Implementation of Face Spoof Recognition: Perspectives and Prospects. International Journal of Engineering Science and Computing. 2018 Jan;8(1):345-9.

[11] Kumar T, Singh KV, Malik S. Artificial neural network in face detection. International Journal of Computer Applications. 2011 Jan;14(3):5-7.

[12] Kim T, Park H, Hong SH, Chung Y. Integrated system of face recognition and sound localization for a smart door phone. IEEE Transactions on consumer Electronics. 2013 Oct 15;59(3):598-603.

[13] Xu N. Application of remote sensing image classification based on adaptive Gaussian mixture model in analysis of mountain environment features. Arabian Journal of Geosciences. 2021 Aug;14(15):1-4.