International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 08 | Aug 2020 www.irjet.net p-ISSN: 2395-0072

FACIAL RECOGNITION USING DEEP LEARNING

Ayush Bijoura[1], Sritama Banerjee[1], Kirti Shekhar Pandey[1], Dr Arivoli A[2]

[1]School of Computer Science and Engineering, VIT University, Vellore, Tamil Nadu, India
[2]Assistant Professor(Sr), School of Computer Science and Engineering, VIT University, Vellore, Tamil Nadu, India

---

Abstract - *Facial expressions does not involve information between humans in face to face interactions. Facial recognition is automatic, and has a important role in human-machine interfaces, this has attracted increasing the attention of researchers from nineties. The approach towards machine learning always require a feature extraction process and always produces poor results. In this paper, we will use real time data set and compare it with the current image .We will also detect the image from the running video then we will use various detection techniques to detect the image. In this way we will come to know whether the person is authorized or not. We will also find the current mood of the person from the detected image.*

Key Words: **Deep learning, Facial Expression Recognition, Neural Network, Human machine interfaces, Real time dataset**

1. INTRODUCTION

Face recognition[1] can be defined as the process of identifying and verifying people(s) in the photograph by their face. The process comprises of detection, alignment, feature extraction, and a recognition task. This is a computer vision program which helps in finding the faces in photo(s).

Deep Learning[2]can be defined as a worldview of Machine learning, or all the more exactly one of its calculations. For the best degree, it depends on an idea of a human cerebrum and the connection of neurons. Deep learning has the ability to construct recognition[3] biometric software which is capable of confirming an individual. The characteristics that seem common to us in humans (example, color of eye) which will not make any sense for a computer which is analyzing each and every pixels in an image.

The most important thing while recognizing the face is the how much is the distance between the eyes, and what is the width and length of the nose, chin width , other things. After this, the data which is obtained are compared with those data which are available in the database, all these parameters matches, then the person is identified.

The approach towards machine learning will often require a complex feature extraction process and produce poor results. In this project data will be stored in form of image. It will be stored as csv file .We will divide the data into two sets i.e. training dataset and testing dataset. After processing the image will go for thresholding, segmentation, canny edge detection, camera detection, ROI (Region of Image) and Hard Cascading. Then we will keep the image in Open CV library. First it will detect the image and it will find whether it is the image of a human being or not. If it is human being, it will return “Image Detected”.

After all these processing the image will get converted in the form of pixels and it will compare the current image with the image or video stored in the database as dataset[5]. If it matches we will be able to find that the person is a member of particular organization. In this way it will be used for security purpose as we can find whether the person is authorized or not.

The new concept in our project is that it will detect the image from the current running video also. The image which is detected we will find the current mood of the person.

2. RELATED WORK

After processing, the image will go for thresholding[6], segmentation, Canny edge detection, camera detection, ROI (Region of Image) and Hard Cascading.

2.1 Open CV

OpenCV short for Open Source Computer Vision can be defined as a library of programming functions and are mostly used in real-time computer vision. Basically, it is library which is used for processing the image. It is mostly used to do all the operation(s) which are related to image(s).

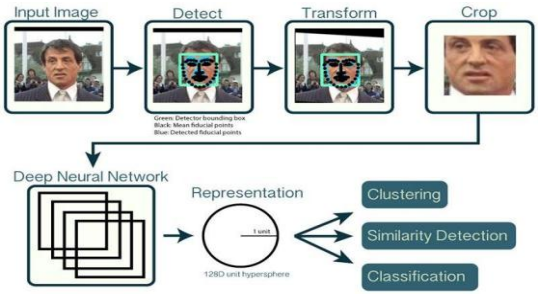

Fig 2.1:- Representation of Deep Neural Network

This process can also be done by Keras. Keras are used in deep learning to classify the images and to build a convolutional neural network (CNN). The images are seen by computer using pixels. Pixels are related to images. For an example, group of pixels may mark an edge in an image or it may define some other pattern. Convolutions are used and help to identify the images. A convolution multiplies a matrix consisting of pixels with a filter matrix and then sums up the values of the multiplication. The convolution then slides to the next pixel and same process is repeated till all the pixels in the images are covered.

2.2 Image Thresholding:

Image threshing is very simple, and effective way to separate an image into a frame and a layer. It is a type of image classification which separates the objects by converting the grayscale image(s) into binary image(s). To process an image, the threshold is used to divide the image into smaller sections using one color or grey color at least to define the border. The main advantage of getting a binary image first is that it reduces complexity and simplifies the process of recognition and classification

2.3 Image Segmentation

Image segmentation[8] (in terms of computer vision)can be defined as the process of dividing a digital image into different categories (pixel sets, also known as photo objects). The main purpose of segmentation is to simplify a or to transform the representation of an image into a logical and easy-to-diagnose component .A part of the art section is a way of dividing or dividing an image into parts, called parts. It is very useful for applications such as image reduction or object recognition, because in these types of applications, it is not efficient to process the entire image.

2.4 Canny edge Detection

Canny edge detection[10] is a multi-step algorithm that can detect edges with high noise simultaneously. Slide images with Gaussian filter to reduce unwanted noise and details. Canny found that the application requirements for finding edges in different viewing systems were the same. Therefore, a solution to find an edge to address these needs may work in a variety of situations. Edge Acquisition is a low cost, which means that the acquisition should take exactly as many edges shown in the image as possible. The point received by the operator should apply directly to the center edge.

The edges provided in the image should only be marked, and where possible, the sound of the image should not create false edges. To satisfy these needs Canny used a different calculation - a technique that gets the job done right. The proper function in the Canny detector is defined by the sum of the four exponential terms, but can be estimated by the first Gaussian output.

Among the most advanced edge detection methods to date, the Canny detection algorithm is one of the most described methods that provides good and reliable detection. Due to its combination of three-dimensional detection and simplifying the implementation process, it became one of the most popular algorithms to find.

2.5 Camera detection

Most home security cameras[7] these days have facial recognition, allowing you to build a database of friends and family who visit your house frequently. After that, when the camera detects a face, it decides whether or not your known database is a name.

The chip inside the camera is constantly scanning the image on its burner, nose, ears and teeth, making it 10 faces at a time before hitting the shutter. Generally, scanned faces should add at least 10 percent of the LCD view height - this requirement prevents the camera from trying to lock face-to-face in the background. You can turn off the facial features when your subject is not personal.

2.6 ROI (Region of Image)

Performance is measured according to the accuracy of the facial recognition category, as well as the accuracy of ROI detection[9] from various facial images. To evaluate performance, the proposed face algorithm was evaluated by assigning the face value N and the number of Negative Faces M.

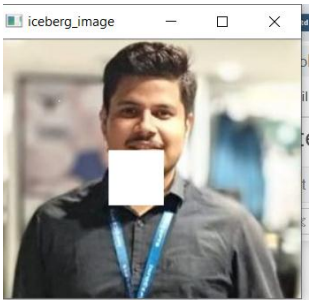

Fig 2.2 Region of Image

2.7 Hard Cascading

Skipping Cascading is a case study[11] based on a wide variety of students, using all the information gathered from the findings from the classifier provided as additional details for the next lesson at Cascade. Unlike volumes or ensembles, which are many expert systems, skiing is a multistage one.

Cascading separators are trained by looking at large "good" samples of something and counterfeit "not good" images of the same value. After the student is educated he can be placed in the photographic region and see the object being discovered. To search for an item in the entire frame, a search window can be moved across the image and scanned wherever the subject is placed. This method is widely used in image processing to locate objects and track them, especially for visual and facial recognition.

3. METHODOLOGY

Face recognition is where images from the video and still pictures are emerging as a research area having different applications in various other fields like law, commercial, etc. This system would find numerous applications like criminal identification, finding missing persons, credit/debit card verification and many more.

When the system[12] is designed, different problems are needed to be addressed. For example when image is detected the pattern of image are taken as a subject and then against either a uniform or a complex background. Then, the identification and verification of face images are done using a proper classification algorithm and then the processed results are posted using schemes which is based on model.

Fig 3.1:-Face detection process

3.1 Neural Networks

Neural networks can be used to recognize the face(s)by learning the coefficients and the calculation is done by the Eigenfaces algorithm. After this the network is then trained on the images present in the database, and then ready to identify the face images which are given to it. The principal component analysis is used to reduce the input series of the neural network, so that the running time of the application can be reduced. The parameters are shown in Fig. 1.2 and is calculated by the formula Input unit(IU)= number of picture/person. Output unit(OU)=total no of person(s) which is to be recognized.

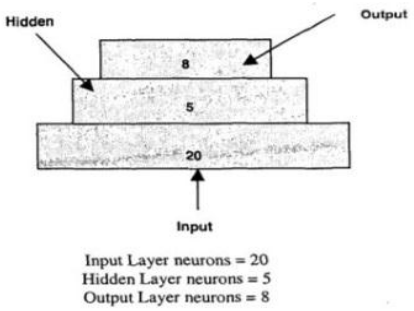

Fig 3.2 :- Network Architecture

4. SYSTEM ARCHITECTURE

4.1 Overview of EigenFace

Eigenfaces are defined as a related set of facial characteristics that a computer uses to recognize a person's face. In this technique, we are having training and test set of images, and then we will compute the eigenvectors of the covariance matrix of the training set of images. They do not directly correspond to any of the features of face like eyes, nose and mouth. When the eigenvectors are displayed, they look like a ghostly face, and they are termed eigenfaces.

4.2 EigenFace Representation

To create an eigenface, a set of images[13] which are large are collected. These are the images which are the good representation of all faces that one might occur. When we configure the first set of the images of face .These face image(s) are converted into row vector by scanning each of these face(s) image from left to right, top to bottom. And after all this all the face vectors are combined so that each row must have a face vector and the whole face distribution is made. Now , here the training set will have N images, the principal component analysis will form the basis set images set M, here, $M < N$. The reconstruction error can be reduced by increasing the eigen pictures numbers , N is always chosen less . For example, if we want to generate M a number of eigenfaces for a given training set for N face images, then we can say that each face image is to be made up of proportion(s) having all the K features or the features of eigenfaces: Face image = (24% of E_1) + (3% of E_2) + (52% of E_3) + ... + (2% E_n).

4.3 Training Data

While starting the training, here we will select only ten face image(s) of individual person which are properly aligned and having a frontal view. The extraction of face extraction is done by the face extraction unit[14].

4.4 Training the Network

We have to prepare a set of training data, for this we take ten images of the same person with varying face expressions. For training to get started we load all the training face picture to the network, here not whole picture is loaded to the network but the projection of the picture using the eigenface space.

4.5 Classification

Classification can be defined as a process of categorizing the classes in the given set of data. This classification can be performed on both the structured and unstructured data. When the data is trained, the face recognizer will be ready for recognition.

4.6 Age Recognition

Age[14] can also be recognized from face. Since, aging is a complex process which will affect the shape, size and texture of the face. The aging process is also visible in different age groups. The aging in the face is represented by growth in the face of age groups (for example, the face of 18 years old), is represented by major texture changes and small changes in the shape. These are due to the changes in the weight and thickness of the skin, people whose age groups are greater than 18 years. Therefore, the age adjustment is must which can compensate both types of aging processes.

4.7 2D Facial Feature Point Detection

Here, we are using manually marked feature points in aging[15] model construction. Therefore, in the test stage, the feature points are detected automatically. The feature points which are present on the 2D face images they are detected using the conventional Active Appearance Model (AAM).

4.8 3D Aging Model

The aging model can be defined as the aging pattern[16] which comprises of face array model from a single subject which is indexed by the given age of the person. So, by assuming that any aging pattern will be approximated by average weighted which is having aging patterns present in the training set. Here, the model construction will differ mostly, here the modeling shape and texture differ separately at different ages using the shape which is pattern space in aging and the texture which is an aging pattern space, because of the fact that the 3D shape and the texture images they are less correlated than the 2D shape and texture which is in use. Here, we will also adjust the shape of 3D which is explained below. Here, the different pattern spaces are described below.

4.9 Shape Aging Pattern

The Shape pattern space[17] will capture the variations in the the shape and the size of the face which is internal. These can be obtained from preprocessing phase and will be used for construction the shape of the pattern space.

Those people who are under the age of 18, the key effects of aging are seen in the increase in the size of carnial, in the later ages, the facial growth[18] is very small considering the height and width. Those who are under the age of 19, we are rescaling the overall size and shapes of 3D. We will perform PCA over all the 3D shapes, where, Sji, is irrespective of age j and is subject to i. All the mean is subtracted Sji into the subspace which is spanned by the columns of Vs to obtain Sj i.

$$\mathbf{s}_i^j = \mathbf{V_s}^T(\mathbf{S}_i^j - \overline{\mathbf{S}}), \quad (1)$$

It is an Ls * l vector.

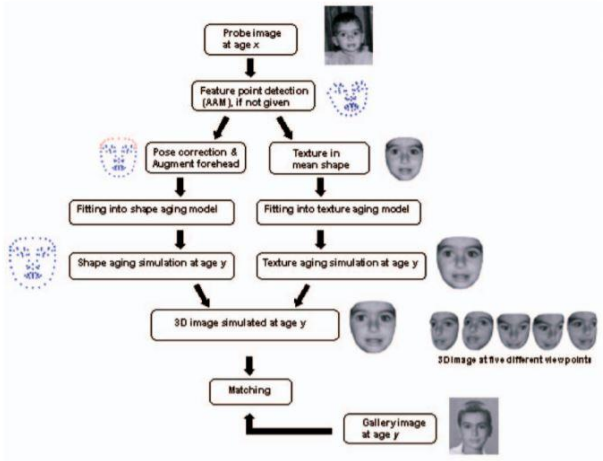

Fig 4.9:- Simulation of age which is ranging from x to y.

4.91 Texture Aging Pattern

Here Tj is the texture pattern where i is subject to age j and j can be obtained by putting the original image of the face to frontal projection which is in the mean shape of S which is followed by concatenating the pixels of image in column wise. After applying PCA on Tji , we will calculate the Vt which is transformation matrix and the texture which is projected t, j, i. We will follow the same procedure in filling which is as in the pattern shape space which will be used to construct the complete basis

for the pattern space shape by using t j i. So, the new texture which is Tj can be obtained in the same way, given j as an age and a set of weights wt, as shown below.

$$\mathbf{T}_{w_t}^{j} = \overline{\mathbf{T}} + \mathbf{V}_t \mathbf{t}_{w_t}^{j},\ 0 \le j \le m-1. \quad (2)$$

$$\mathbf{T}_{w_t}^{j} = \overline{\mathbf{T}} + \mathbf{V}_t \mathbf{t}_{w_t}^{j},\ 0 \le j \le m-1. \quad (3)$$

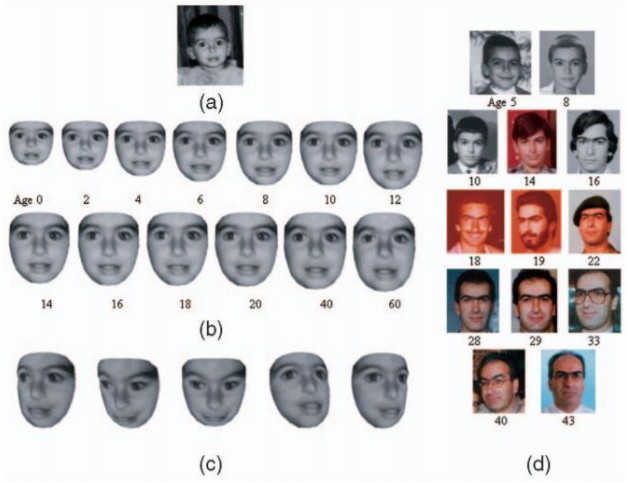

Fig.3.4 (a)At age 2 the image is given. (b) From age 0 to 60 the simulated images are shown.(c) Five different poses from the aging 20 is. (d) Original images

4.92 Emotion Recognition

Emotion[19] of a person can be recognized using his/her facial expressions. Based on his/her facial expression[20] whether the person is sad, happy, angry ,neutral , joyful. The structure is shown below.

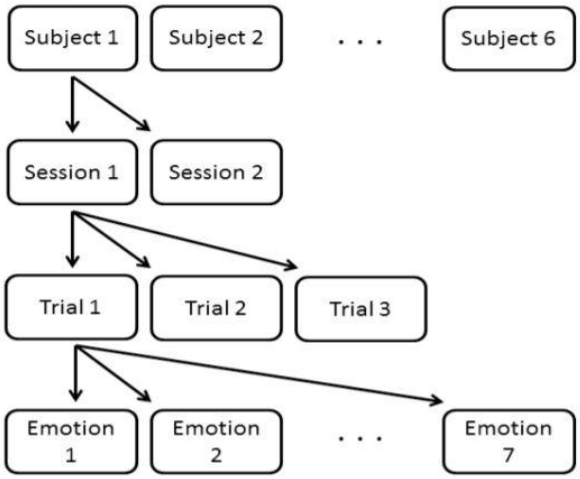

Fig4.92 Hierarchical Structure

5. CONCLUSION

By working on this project we found that when camera is capturing the face ,it is displaying the mood of person whether he/she is happy, sad, angry, neutral ,joyful . We have also used Open CV and CNN which is working properly and the algorithm which we have used has shown efficiency. The proposed model is able to recognize faces effectively but when experimented with videos, it takes more time to process. The advantage of this model is that it is able to recognize blurred images and sideways images with some traditional models which is not seen in such a scenario. The only thing is that he/she can not ssee through the glasses. In the future, this can be expanded to see people using video conferencing that will be useful in obtaining ownership of CCTV cameras that can enable the police to identify the person in no time. It can also be used in home security programs.

6. REFERENCES

[1] Parkhi, Omkar M., Andrea Vedaldi, and Andrew Zisserman. "Deep face recognition." (2015).

[2] Abudarham, Naphtali, Lior Shkiller, and Galit Yovel. "Critical features for face recognition." Cognition 182 (2019): 73-83.

[3] Masi, Iacopo, et al. "Face-specific data augmentation for unconstrained face recognition." International Journal of Computer Vision 127.6-7 (2019): 642-667.

[4] Huang, Chen, et al. "Deep imbalanced learning for face recognition and attribute prediction." IEEE transactions on pattern analysis and machine intelligence (2019).

[5] Deng, Jiankang, et al. "Lightweight face recognition challenge." Proceedings of the IEEE International Conference on Computer Vision Workshops. 2019.

[6] Hau, Cheuk Yu, and Yvonne Sadovy de Mitcheson. "A facial recognition tool and legislative changes for improved enforcement of the CITES Appendix II listing of the humphead wrasse, Cheilinus undulatus." Aquatic Conservation: Marine and Freshwater Ecosystems 29.12 (2019): 2071-2091.

[7] Bowyer, Kevin W. "Face recognition technology: security versus privacy." IEEE Technology and society magazine 23.1 (2004): 9-19.

[8] Mitra, Debasree, Parantapa Sarkar, and Payel Roy. "Face Recognition by City-Block Distance Classifier in Supervised Machine Learning." IJRAR DOI: http://doi. one/10.1729/Journal 21653 (2019).

[9]Siddiqui, Muhammad Farhan, et al. "Face Detection and Recognition System for Enhancing Security Measures Using Artificial Intelligence System." INDIAN JOURNAL OF SCIENCE AND TECHNOLOGY 13.09 (2020): 1057-1064.

[10]Connolly, Hannah L., Andrew W. Young, and Gary J. Lewis. "Recognition of facial expression and identity in part reflects a common ability, independent of general intelligence and visual short-term memory." Cognition and Emotion 33.6 (2019): 1119-1128.

[11]Dandıl, Emre, and Rıdvan Özdemir. "Real-time Facial Emotion Classification Using Deep Learning." Data Science and Applications 2.1 (2019): 13-17.

[12]Riaz, Sidra, et al. "Age-invariant face recognition using gender specific 3D aging modeling." Multimedia Tools and Applications 78.17 (2019): 25163-25183.

[13] Sawant, Manisha M., and Kishor M. Bhurchandi. "Age invariant face recognition: a survey on facial aging databases, techniques and effect of aging." Artificial Intelligence Review 52.2 (2019): 981-1008.

[14]Mortezaie, Zahra, and Hamid Hassanpour. "ASurvey ON AGE-INVARIANT FACE RECOGNITION METHODS." Jordanian Journal of Computers and Information Technology (JJCIT) 5.02 (2019).

[15]Naik, Manoj Kumar, and Aneesh Wunnava. "Classical 2D Face Recognition: A Survey on Methods, Face Databases, and Performance Evaluation." Advances in Intelligent Computing and Communication. Springer, Singapore, 2020. 375-383.

[16]Ekmen, Beste, and Hazım Kemal Ekenel. "From 2D to 3D real-time expression transfer for facial animation." Multimedia Tools and Applications 78.9 (2019): 12519-12535.

[17] Singh, Yash, et al. "Method for 3D modelling based on structure from motion processing of sparse 2D images." U.S. Patent No. 10,198,858. 5 Feb. 2019.

[18] Krishnapriya, K. S., et al. "Characterizing the variability in face recognition accuracy relative to race." arXiv preprint arXiv:1904.07325 (2019).

[19] Srivastava, Yash, Vaishnav Murali, and Shiv Ram Dubey. "Hard-mining loss based convolutional neural network for face recognition." arXiv preprint arXiv:1908.09747 (2019).

[20] Sawant, Manisha M., and Kishor M. Bhurchandi. "Age invariant face recognition: a survey on facial aging databases, techniques and effect of aging." Artificial Intelligence Review 52.2 (2019): 981-1008.