FACE RECOGNIZATION AND IDENTIFICATION USING MACHINE LEARNING SYSTEM

Project Guide : Prof.R.R.Kolte

Department of Information Technology , KDK College of Engineering.

Ruchita Chandekar , Aman Chavare , Maithili Mate , Tejaswini Raut , Snehal Meshram

Department of Information Technology , KDK College of Engineering.

---***---

Abstract :- Face recognition (FR) has been at the crux of several novel breakthroughs over the past two decades and has steadily proffered several cross-domain applications that range from mainstream commercial software to critical law enforcement applications. Recent groundbreaking developments in Big Data analysis, Cloud Computing, Social Networks and Machine learning have vastly transformed the conventional view of how several formidable problems in Computer Vision can be tackled. Hence in this paper, we will provide a thorough survey of the concepts of Cloud Computing, Big Data, Social networks and Machine Learning from a contemporary perspective of FR, and proffer a framework for a novel FR approach based on the Extreme Learning Machines technique to perform the task of Face Tagging for Social Networks operating on Big Data. Identifying a person with an image has been popularised through the mass media. For face detection, Haar-Cascades were used and for face recognition,. The methodology is described including flow charts for each stage of the system. Next, the results are shown including plots and screen-shots followed by a discussion of encountered challenges. The report is concluded with the authors opinion on the project and possible applications.

Keywords : Face recognition, Face detection, Web-cam,

I . INTRODUCTION

Most of the facial network algorithems are based onmachin learning.Machine learning is very vast subject. Machine learning is basically based on the learning from available data and artifacts such as available sales records, available search trends, available usage behavior, pictures and images, sounds, text etc. There are broadly 3 different types of machine learning algorithms - supervised learning, unsupervised learning and reinforced learning. In the case of supervised learning, learning of the system is based on pre-classified data labels. A set of data labels is already classified by humans. Systems learns from this data and create a model(based on algorithm used). Then this model is tested and refined using the test data. In unsupervised learning, system is not fed with any specially classified data labels. Instead system classifies the available data sets based on the patterns it is able to find (again based on the applied set of algorithms) .

Facial recognition algorithms are based on identify facial features by extracting landmarks, or features, from an image of the subject's face. These features are then used to search for other images with matching features.Face recognition has become a widespread topic of study in recent times since rise in plea designed for security along with the swift development of mobile devices. Face recognition is efficient to various applications such asentrance control, personality authentication, security systems, surveillance systems unlocking of smart phones and social media networks etc. Entrance control includes offices, computers, phones, ATMs, etc.

Ever since the event of 9/11 there has a more apprehensive importance on developing security systems to make sure the safety of citizens, specifically in places such as airports, companies, border crossings where identification and verification is essential. And finally, facial recognition has raised in social media applications on platforms which is Face book which suggest users to tag friends who have been identified in images. It is clear that, there are numerous applications for facial recognition systems. In this paper face recognition for human faces are employed through Viola Jones algorithm for face identification , LBPH in favor of feature extraction , Euclidean distance classifier is for face recognition. In general the steps employed to achieve this are the following: creating a dataset, face acquisition, feature , extraction and finally classification. The entire work is done openCV using python.

II . LITERATURE SURVEY

Nawaf Hazim Barnouti [1] in the year 2016 proposed an automatic face recognition system basedon exterior-based methods. Viola-Jones method is used to detect and gather faces in each database. Square Euclidean Distance is used to determine the distance between two images, which leads to find image similarity.

Ningthoujam Sunita Devi [2] presented a methodology for face recognition based on information theory approach of coding and decoding the face image.

on Oracle Research Laboratory (ORL) face   Proposed methodology is combination of two stages – Feature extraction using principle component analysis and Int. J. Sci. Res. in Computer Science and Engineering Vol.6(5), Oct. 2018, E-ISSN: 2320-7639 © 2018, IJSRCSE All Rights Reserved 7 recognition using the feed forward back propagation Neural Network.

Robust Real-Time Objection Detection, 2001 [3] By Paul Viola and Michael J. Jones. This seems to be the first article where Viola-Jones presents the coherent set of ideas that constitute the fundamentals of their face detection algorithm. This algorithm only finds frontal upright faces, but is in 2003 presented in a variant that also detects profile and rotated views [4]. The „Methods" chapter will elaborate more on the basic version of this algorithm.

III . PROPOSED SYSTEM

Below are the methodology and descriptions of the applications used for data gathering, face detection, training and face recognition. The project was coded in Python using a mixture of IDLE and PYCharm IDEs.
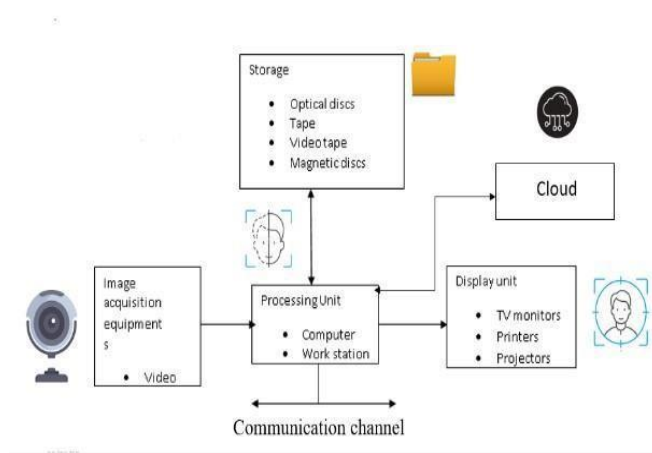

Fig 1 : System Architecture

The Proposed System of Face Detection

First stage was creating a face detection system using Haar-cascades. Although, training is required for creating new Haar-cascades, OpenCV has a robust set of Haar-cascades that was used for the project. Using face-cascades alone caused random objects to be identified and eye cascades were incorporated to obtain stable face detection. The flowchart of the detection system can be seen in figure .
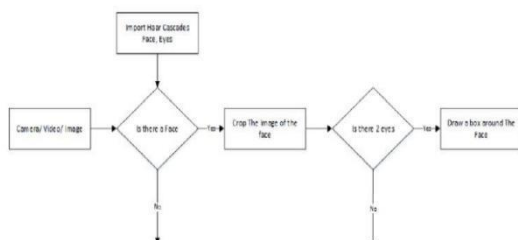

Fig 3 The Flow chart of the face detection application

classifier objects are created using classifier class in OpenCV through the cv2.CascadeClassifier() and loading the respective XML files. A camera object is created using the cv2.VideoCapture() to capture images By using the

CascadeClassifier.detectMultiScale() object of various sizes are matched and location is returned.

Face Recognition Process

For this project three algorithms are implemented independently. These are Eigenface, Fisherface and Linear binary pattern histograms respectively.

All three can be implemented using OpenCV libraries. There are three stages for the face recognition as follows:

1. Collecting images IDs

2. Extracting unique features, classifying them and storing in XML files

3. Matching features of an input image to the features in the saved XML files and predict identity.
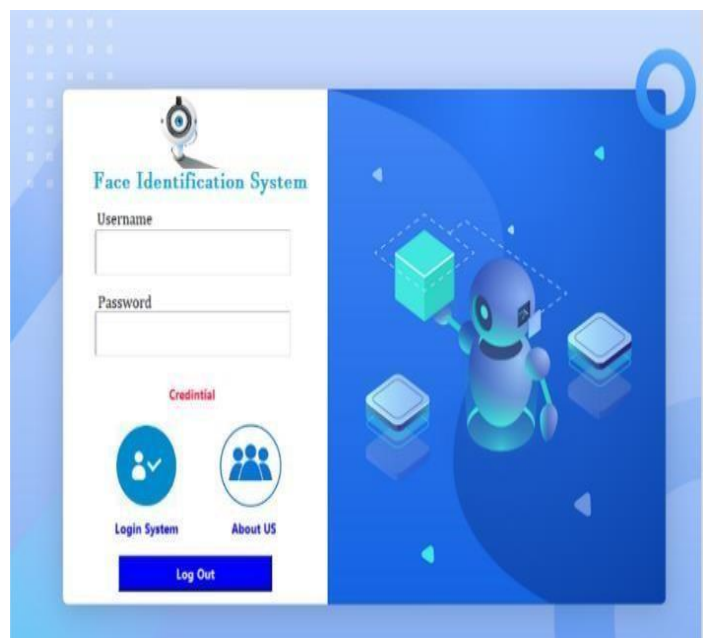

Fig 2 : Administrator Login Page

Accessing your account is as easy as entering your username and password on the login screen, or clicking one of our Single login system options. Whenever you enter the project dashboard of a given project or any of its service you'll see a menu item called Admin appears in the project menu . Only project members can access the Project Administration page. All other users are denied access to this part of a project. In addition to the project main menu that was introduced in the description of the Project Dashboard, the Project Administration page has its own menu.
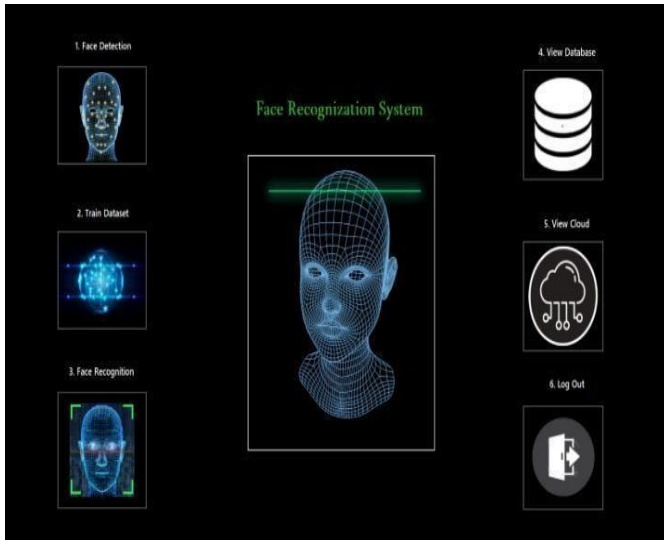
Fig 4 : Main interface in recognization system

of system and is designed such a way that it is expected to provide the user insight of the system. UI provides fundamental platform for human-computer interaction.
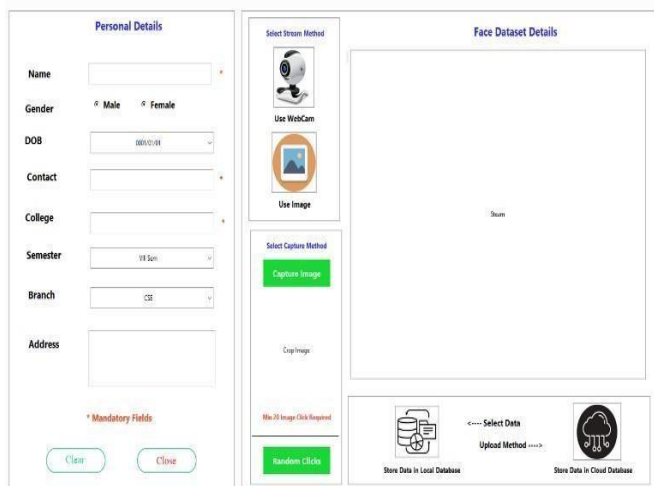

Fig 5 : Registration person Information for train data in data detection

A registration form is a list of fields that a user will input data into(text , image form) and submit to a system or individual. There are many reasons why you would want a person to fill out a registration form. system use registration forms and capture images to train data for recognize , identifies and person information

There are basically two different ways to store data in an system databases: Local database and cloud database. Cloud storage is a model of computer data storage in which the digital data is stored in logical pools. The physical storage spans multiple servers, and the physical environment is typically owned and managed by a hosting system and we have use as MangoDB to store system database. MongoDB is an open-source document database and leading NoSQL database. MongoDB is written in C++. This tutorial will give you great understanding

on MongoDB concepts needed to create and deploy a highly scalable and performance-oriented database. And next click train dataset.
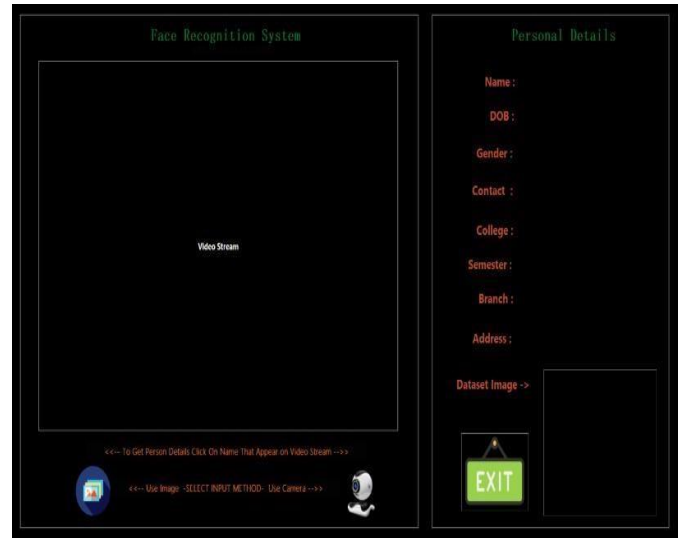

Fig 6 : Face recognize in system recognition

A face recognition system is expected to identify faces present in images and videos automatically. It can operate in either or both of two modes: i) Face verification (or authentication): involves a one-to-one match that compares a query face image against a template face image whose identity is being claimed.

ii)Face identification (or recognition): involves one-to-many matches that compares a query face image against all the template images in the database to determine the identity of the query face.

IV . METHODOLOGY

Face Detection using Haar-Cascades : A Haar wavelet is a mathematical fiction that produces square-shaped waves with a beginning and an end and used to create box shaped patterns to recognise signals with sudden transformations. An example is shown in figure 1. By combining several wavelets, a cascade can be created that can identify edges, lines and circles with different colour intensities. These sets are used in Viola Jones face detection technique in 2001 and since then more patterns are introduced [10] for object detection as shown in figure . To analyse an image using Haar cascades, a scale is selected smaller than the target image. Face detection on a human face is performed by matching a

combination of different Haar-like-features. For example, forehead, eye brows and eyes contrast as well as the nose with eyes as shown below in figure A single classifier is not accurate enough. Several classifiers are combined as to provide an accurate face detection system as shown in the block .
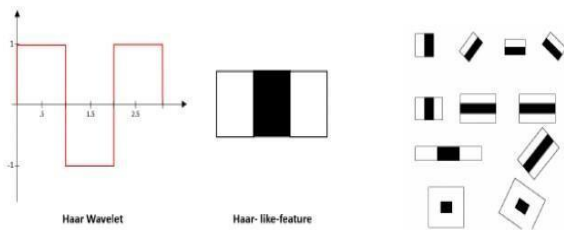
Fig 7 :- A Haar wavelet and resulting Haar-like features

In this project, a similar method is used effectively to by identifying faces and eyes in combination resulting better face detection. Similarly, in viola Jones method [7], several classifies were combined to create stronger classifiers. ADA boost is a machine learning algorithm that tests out several week classifiers on a selected location and choose the most suitable [7]. It can also reverse the direction of the classifier and get better results if necessary [7]. 2. Furthermore, Weightupdatesteps can be updated only on misses to get better performance. The cascade is scaled by 1.25 and re-iterated in order to find different sized faces. Running the cascade on an image using conventional loops takes a large amount of computing power and time. Viola Jones [7] used a summed area table (an integral image) to compute the matches fast. First developed in 1984 [11], it became popular after 2001 when Viola Jones implemented Haar-cascades for face detection. Using an integral image enables matching features with a single pass over the image.

Local Binary Pattern Histogram : Local binary patterns were proposed as classifiers in computer vision and in 1990 By Li Wang [4]. The combination of LBP with histogram oriented gradients was introduced in 2009 that increased its performance in certain datasets [5]. For feature encoding, the image is divided into cells (4 x 4 pixels).Using a clockwise or counter-clockwise direction surrounding pixel values are compared with the central as. The value of intensity or luminosity of each neighbour is compared with the centre pixel. Depending if the difference is higher or lower than 0, a 1 or a 0 is assigned to the location. The result provides an 8-bit value to the cell. The advantage of this technique is even if the luminosity of the image.
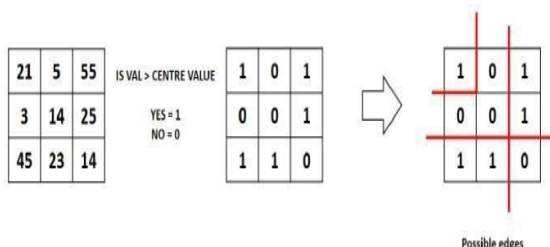

Fig 8 :- Local binary pattern histogram generating 8-bit number

is changed as in figure, the result is the same as before. Histograms are used in larger cells to find the frequency of occurrences of values making process faster. By analysing the results in the cell, edges can be detected as the values change.

By computing the values of all cells and concatenating the histograms, feature vectors can be obtained. Images can be classified by processing with an ID attached. Input images are classified using the same process and compared with the dataset and distance is obtained. By setting up a threshold, it can be identified if it is a known or unknown face. Eigenface and Fisherface compute the dominant features of the whole training set while LBPH analyse them individually.

Fig 9 :- The results are same even if brightness is changed

V . CONCLUSION

This paper describes the mega-project for visual perception and autonomy module. Next, it explains the technologies used in the project and the methodology used. Finally, it shows the results, discuss the challenges and how they were resolved followed by a discussion. Using Haar-cascades for face detection worked extremely well even when subjects wore spectacles. Real time video speed was satisfactory as well devoid of noticeable frame lag. Considering all factors, LBPH combined with Haar-cascades can be implemented as a cost effective face recognition platform. An example is a system to identify known troublemakers in a mall or a supermarket to provide the owner a warning to keep him alert or for automatic attendance taking in a class.

In this project face detection is carried out by using ViolaJones face detection algorithm, feature extraction by using Local Binary Patterns Histograms and classification by using Euclidean distance classifier. The proposed system is implemented using Opencv and anaconda. From the above graphs I can conclude that LBPH and Euclidean distance has better recognition rate. The results achieving for this methods is of accuracy 85%-95%.

VI . REFERENCES

1. International Journal of Computer Science and Information Technologies, Vol. 5 (5), 2014,pp 64916496.

2. Manisha V. Borkar, Bhakti Kurhade, A Research – Face Recognition by Using Near Set Theory, International Journal

of Advanced Research in Computer Science and Software Engineering, Volume 5, Issue 4, 2015 ISSN: 2277 128X.

3. Asit Kumar Datta, Madhura Datta, Pradipta Kumar Banerjee, Face Detection and Recognition: Theory and Practice, CRC Press, 28-Oct-2015.

4. White paper by Animetrics, Inc, Facial Recognition & Identity Resolution, 2012.

5. Paul Viola, Michael J. Jones, Robust Real-Time Face Detection, International Journal of Computer Vision 57(2), 2004.

6. Brunelli, R. and Poggio, T. (1993), Face Recognition: Features versus Templates. IEEE Transactions on Pattern Analysis and Machine Intelligence, 15(10):1042-1052

7. Craw, I., Ellis, H., and Lishman, J.R. (1987). Automatic extraction of face features. Pattern Recognition Letters, 5:183187, February.

8. Adelson, E. H., and Bergen, J. R. (1986) The Extraction of Spatio-Temporal Energy.

9. Analysis (pp. 151-155) Charleston, SC; May 7-9 Learning for Face Verification." IEEE Transactions on Pattern Analysis and Machine Intelligence 38.10 (2016): 19972009.

10. Goodfellow, Ian, Yoshua Bengio, and Aaron Courville. Deep learning. MIT Press, 2016