****

**K L E F**

**KONERU LAKSHMAIAH EDUCATION FOUNDATION**

**(Deemed to be university estd, u/s, 3 of th• UGC Act. 1956)**

**(NAAC Accredited "A" Grade UnJverslty)**

PROJECT BASED LAB REPORT

On

# PERSONALITY CLASSIFICATION

Submitted in partialfulfilment of the Requirements for the award of the Degree of Bachelor of Technology

In

Computer science and Engineering Under the esteemed guidance Nilu Singh By

|  |  |
| --- | --- |
| Student Name | ID number |
| VISHWAJEET SHARMA | 2000031822 |
|  |  |

(DST-FIST Sponsored Department) K L EDUCATION FOUNDATION

Green Fields, Vaddeswaram, G1mtur District-522 502 2022-2023

K L EDUCATION FOUNDAT ION

DEPARTMENT OF COMPUTER SCIENCE ANDENGINEERING

(DST-FIST Sponsored Depa rtment)

A picture containing text, gambling house, room

Description automatically generated

CERTIFICATE

This is to certify that this project-based lab report entitled "Personality Classification" is a bonafide work done by Vishwajeet Sharma(2000031822) in the course ?????????in partial fulfilment of the requirements for the award of Degree in Bachelor of Technology in COMPUTER SCIENCE AND ENGNEERING during the Even Semester of Academic year 2021-2022.

**Faculty in-Charge** Head of the Department

Dr .. Nilu Singh Prof. HARI KIRAN VEGE

K L EDUCATION FOUNDATION

DEPT OF COMPUTER SCINCE AND ENGINEERING

( DST-FIST Sponsored Department)

A close-up of a sign

Description automatically generated with low confidence

(DECLARATION )

We here by declare that this project-based lab report entitled "Personality Classification" has been prepared by us in the course Data Warehousing and Mining, 19CS3052P in partial fulfilment of the requirement for the award of degree bachelor of technology in COMPUTER SCIENCE AND ENGINEERING during the EVEN Semester of the academic year 2022-2023.We also declare that this project-based lab report is of our own effort and it has not been submitted to any other university for the award of any degree.

**DATE:-**

**PLACE:-**

Signature of Student

|  |  |
| --- | --- |
| Name | ID number |
| Vishwajeet Sharma | 2000031822 |
|  |  |

**ACKNOWLEDGEMENT**

Our sincere thanks to Dr Nilu Singh in the Lab for her outstanding support throughout the project for the successful completion of the work.

We express our gratitude to Dr. Nilu Singh the instructor and Dr. Pragnya ban Mishra Couxse Co-Ordinator for the course Data Warehousing and Mining, 19CS3052P in the Computer Science and Engineering Department for providing us with adequate planning and support and means by which we can complete this project-based Lab.

We express our gratitude to Prof. V.HARI KIRAN, Head of the Department for computer science and Engineering for providing us with adequate facilities, ways and means by which we can complete this project-based Lab.

We would like to place on record the deep sense of gratitude to the honorable Vice Chancellor,

KL University for providing the necessary facilities to carry the project-based Lab.

Last but not the least, we thank all Teaching and Non-Teaching Staff of our department and especially our classmates and our friends for their support in the completion of our project-based

Lab.

|  |  |
| --- | --- |
| Name | ID number |
|  |  |
|  |  |

**ABSTRACT**

In this paper, we propose an Automated Face Recognition Attendance System for Educational and Work Organizations for marking and maintaining attendance with the use of Face Recognition. Keeping up participation is imperative in all learning foundations for checking the execution of understudies. In most learning foundations, understudy attendances are physically taken by the utilization of participation sheets issued by the office heads as a major aspect of control. The understudies sign in these sheets which are then filled or physically signed in to a PC for future examination. This strategy is dull, tedious and wrong as a few understudies regularly sign for their missing partners. This technique additionally makes it hard to track the participation of individual understudies in a huge classroom condition In this undertaking, we propose the outline and utilization of a face recognition and acknowledgment framework to consequently distinguish understudies going to an address in a classroom and stamp their participation by perceiving their countenances

**Keyword: -** *Face Recognition, LBP, SVM*

1. **INTRODUCTION**

Face is used as a one of the biometric identification parts of a human. Everyone has different Facial Features and stores them as a biometric print to uniquely identify the person. The traditional method of Attendance is very time taking and a burden on the faculties and in the current covid time, it is also not safe to touch a surface that is touched by many others. Other biometric methods have higher chances of spreading the coronavirus. Face Recognition Based Attendance system is totally contact-less method. Face Recognition is widely used in Phones and in other security systems to lock and unlock a particular system. This is implemented by many websites/apps like Whatsapp, Instagram and it helps in also Crime investigation. Face Recognition is also works from some short distance without any contact with the other person. So, there is no need to come near to the biometric system to activate and mark your attendance. In this type of system, there is no chance of false attendance or proxy. It works in two parts: Capture and verification. It compares the stored image with the image which is captured by the camera. If a person's identity totally matches then it will proceed further. Face Recognition works on the 1:1 Matching process. The research shows that there is so much improvement in the Face Biometric system in the last 10 years. There is a huge development in Face Recognition. Now it is working on the different programming languages and much easy to develop a Face Recognition system because of the lots of resources present. Attendance system based on Face Recognition. Here the face of a person works as an identity. It will be matched and verified by the system by the data of the person already stored in the system. If it will perfectly match it will mark attendance and store it in the database. It will take less time and data will be properly stored in a proper format. Stored data can be extracted into an excel sheet

Keeping up the participation is essential in every one of the establishments for checking the nearness of understudies. Each foundation has its own strategy in such manner. Some are taking participation physically utilizing the conventional pen and paper or document based approach and some have received strategies for programmed participation methods. There are numerous strategies existing for this reason they are:

* + Fingerprint Based System
  + Iris Recognition
  + RFID Based System
  + Face Recognition

The initial three techniques demonstrated wasteful in light of the fact that understudies need to make a line to touch their thumb on the checking gadget which devours the time. This framework utilizes the face acknowledgment approach for the programmed participation of understudies in the classroom condition without understudy's intervention. This participation is recorded by utilizing a camera connected in the classroom that is constantly catching pictures of understudies, identify the appearances in pictures and contrast the recognized countenances and the understudy database and check the attendance. The framework comprises of a camera that catches the pictures of the understudies sitting in the classroom and sends it to the picture upgrade module. In the picture improvement module, pictures are upgraded with the goal that coordinating can be performed effectively. After improvement, the picture comes in the face discovery and acknowledgment modules. At the season of enlistment, formats of face pictures of individual understudies are put away in the face database. The countenances are recognized from the caught by camera. In the acknowledgment module the identified appearances are continually looked at against put away database. On the off chance that any face is perceived the participation is refreshed and can be gotten to by anybody, the data will be sent to the truants' folks utilizing GSM innovation. This framework utilizes iris acknowledgment administration framework that does catching the picture of iris acknowledgment, extraction, putting away and coordinating. In any case, the difficulty strikes lay the trans-mission lines in the spots where the geology is terrible. The Authors have proposed a framework in view of constant face recognition which is reliable, secure and fast which needs improvement in different lighting conditions

1.2 Problem Statement

Traditional student attendance marking technique is often facing a lot of trouble. The face recognition student attendance system emphasizes its simplicity by eliminating classical student attendance marking technique such as calling student names or checking respective identification cards. There are not only disturbing the teaching process but also causes distraction for students during exam sessions. Apart from calling names, attendance sheet is passed around the classroom during the lecture sessions. The lecture class especially the class with a large number of students might find it difficult to have the attendance sheet being passed around the class. Thus, face recognition student attendance system is proposed in order to replace the manual signing of the presence of students which are burdensome and causes students get distracted in order to sign for their attendance. Furthermore, the face recognition based automated student attendance system able to overcome the problem of fraudulent approach and lecturers does not have to count the number of students several times to ensure the presence of the students. The paper proposed by Zhao, W et al. (2003) has listed the difficulties of facial identification. One of the difficulties of facial identification is the identification between known and unknown images. In addition, paper proposed by Pooja G.R et al. (2010) found out that the training process for face recognition student attendance system is slow and time-consuming. In addition, the paper proposed by Priyanka Wagh et al. (2015) mentioned that different lighting and head poses are often the problems that could degrade the performance of face recognition based student attendance system. Hence, there is a need to develop a real time operating student attendance system which means the identification process must be done within defined time constraints to prevent omission. The extracted features from facial images which represent the identity of the students have to be consistent towards a change in background, illumination, pose and expression. High accuracy and fast computation time will be the evaluation points of the performance.

4 1.3 Aims and Objectives

The objective of this project is to develop face recognition based automated student attendance system. Expected achievements in order to fulfill the objectives are: • To detect the face segment from the video frame. • To extract the useful features from the face detected. • To classify the features in order to recognize the face detected. • To record the attendance of the identified student. Figure 1.1 Block Diagram of the General Framework 1.4 Thesis Organization Chapter 2 includes a brief review of the approaches and studies that have been done previously by other researchers whereas Chapter 3 describe proposed methods and approaches used to obtain the desired output. The results of the proposed approach would be presented and discussed in Chapter 4. The conclusion, as well as some recommendations would be included in Chapter 5.

LITERATURE SURVEY

Traditionally attendance was taken manually which is very time consuming and often leads to human error. Additionally, there are many uncertainties towards the sources of the attendance records which in fact, most of the attendance records are not retrieved from the actual situation. The old method that uses paper sheets for taking student’s attendance can no longer be used. Based on the research, there are many solutions that are available to solve this issue. According to research journal “Attendance System Using NFC Technology with Embedded Camera on Mobile Device” (Bhise, Khichi, Korde,Lokare, 2015). The attendance system is improved by using Near Field. Communication (NFC) technology and mobile application. According to the research paper, each student is given a NFC tag that has a unique ID during their enrolment into the college. The face recognition problem is formulated as a problem in difference space, which models dissimilarities between two facial images. In different space they formulate face recognition as a two class problem. The cases are: (i) Dissimilarities between faces of the same person, and (ii) Dissimilarities between faces of different people. By modifying the interpretation of the decision surface generated a similarity metric between faces, that is learned from examples of differences between faces. The SVM-based algorithm is compared with a principal component analysis (PCA) based algorithm on a difficult set of images from the FERET database. Performance was measured for both verification and identification scenarios. The identification performance for SVM is 77-78% versus 54% for PCA. For verification, the equal error rate is 7% for SVM and 13% for PCA.

Face recognition is one of the few biometric methods that possess the merits of both high accuracy and low intrusiveness. It has the accuracy of a physiological approach without being intrusive. Over past 30 years, many researchers have proposed different face recognition techniques, motivated by the increased number of real world applications requiring the recognition of human faces. There are several problems that make automatic face recognition a very difficult task. However, the face image of a person inputs to the database that is usually acquired under different conditions. The important of automatic face recognition is much be cope with numerous variations of images of the same face due to changes in the following parameters such as 1. Pose 2. Illumination 3. Expression 4. Motion 5. Facial hair 6. Glasses 7. Background of image.

Face recognition technology is well advance that can applied for many commercial applications such as personal identification, security system, image- film processing, psychology, computer interaction, entertainment system, smart card, law enforcement, surveillance and so on. Face recognition can be done in both a still image and video sequence which has its origin in still-image face recognition. Different approaches of face recognition for still images can be categorized into three main groups such as 1. Holistic approach

Holistic approach :- In holistic approach or global feature, the whole face region is taken into account as input data into face detection system. Examples of holistic methods are eigenfaces (most widely used method for face recognition), probabilistic eigenfaces, fisherfaces, support vector machines, nearest feature lines (NFL) and independent-component analysis approaches.They are all based on principal component-analysis (PCA) techniques that can be used to simplify a dataset into lower dimension while retaining the characteristics of dataset

PROPOSED WORK

The proposed system face recognition-based attendance system can be divided into five main modules.The modules and their functions are defined as follows. Image Capture The camera is fixed at a distance from the entrance to capture the frontal image of the students.And remaining process goes for face detection. Face Detection A proper and efficient face detection algorithm always increases the performance of face recognition systems. Various algorithms are proposed for face detection such as face knowledge based methods, feature invariant methods, machine learning based methods. In this project, I implemented a system for locating faces in digital images. These are in JPEG format only. Before we continue, we must differentiate between face recognition and face detection. They are not the same, but one depends on the other. In this case face recognition needs face detection for making an identification to “recognize” a face. I will only cover face detection. Face detection uses classifiers, which are algorithms that detects what is either a face(1) or not a face(0) in.

Face recognition is achieved using machine learning and the basic pipeline used for it is as follows: 1. Finds face in an image. 2. Analyses facial features. 3. Compares against known faces and makes a prediction. Development of complete attendance system is achieved using UI and Android application. Here the application takes data like subject details, faculty details, date and time and provides a click to start the attendance. The images of students are clicked and sent to Linux server where python script runs to mark attendance and generate spreadsheet and PDF file which is then mailed.

|  |  |
| --- | --- |
| **Types of System** | **Drawbacks** |
| RFID based | Fraudulent usage |
| Fingerprint based | Time Consuming for students to wait and give their attendance |
| Iris-based | Evades the privacy of the user |
| Wireless-based | Poor performance if topography is bad Extraction Feature Image Capture Face Detection Pre−processing Image Capture Face Detection Pre−processing Classification |

Face Detection

Difference between face detection and face recognition are often misunderstood. Face detection is to determine only the face segment or face region from image, whereas face recognition is to identify the owner of the facial image. S.Aanjanadevi et al. (2017) and Wei-Lun Chao (2007) presented a few factors which cause face detection and face recognition to encounter difficulties. These factors consist of background, illumination, pose, expression, occlusion, rotation, scaling and translation

|  |  |  |
| --- | --- | --- |
| Face detection  method | Advantages | Disadvantages |
| Viola jones algorithm | 1. High detection speed 2. High accuracy. | 1. Long training time. 2. Limited head pose. 3. Not able to detect dark faces. |
| Local Binary pattern | 1. Simple computation. 2. High tolerance against the   monotonic  illumination changes. | 1. Only used for binary and grey images. 2. Overall performance is   inaccurate compared to Viola-Jones algorithm. |
| AdaBoost algorithm  (part of Viola  jones algorithm) | Need not to have any prior knowledge about face structure. | The result highly depends on the training data and affected by weak classifiers. |
| SMQT Features and SNOW Classifier Method | 1. Capable to deal with lighting problem in object detection. 2. Efficient in   computation. | The region contain very similar to grey value regions will be misidentified as face. |
| Neural-Network | High accuracy only if large size of image were trained. | 1. Detection process is slow and computation is complex. 2. Overall performance is   weaker than Viola-Jones algorithm. |

#### Pre-Processing

Subhi Singh et al. (2015) suggested cropping of detected face and colour image was converted to grayscale for pre-processing. They also proposed affine transform to be applied to align the facial image based on coordinates in middle of the eyes and scaling of image to be performed. Arun Katara et al (2017), Akshara Jadhav et.al (2017), Shireesha Chintalapati, and M.V. Raghunadh (2013), all of the 3 papers have proposed histogram equalization to be applied to facial image, and scaling of images was performed for pre-processing.

Pre-processing enhances the performance of the system. It plays an essential role to improve the accuracy of face recognition. Scaling is one of the important pre- processing steps to manipulate the size of the image. Scaling down of an image increases the processing speed by reducing the system computations since the number of pixels are reduced. The size and pixels of the image carry spatial information. Gonzalez, R. C. and Woods (2008) mentioned spatial information is a measure of the smallest discernible detail in an image. Hence, spatial information has to be manipulated carefully to avoid distortion of images to prevent checkerboard effect. The size should be same for all the images for normalization and standardization purposes. Subhi Singh et al (2015) proposed PCA (Principal Component Analysis) to extract features from facial images, same length and width of image is preferred, thus images were scaled to 120 × 120 pixels.

Besides scaling of images, colour image is usually converted to grayscale image for pre-processing. Grayscale images are believed to be less sensitive to illumination condition and take less computational time. Grayscale image is 8 bit image which the pixel range from 0 to 255 whereas colour image is 24 bit image which pixel can have 16 77 7216 values. Hence, colour image requires more storage space and more computational power compared to grayscale images. (Kanan and Cottrell, 2012). If colour image is not necessary in computation, then it is considered as noise. In addition, pre-processing is important to enhance the contrast of images. In the paper of Pratiksha M. Patel (2016), he mentioned that Histogram equalization is one of the methods of pre-processing in order to improve the contrast of the image. It provides uniform distribution of intensities over the intensity level axis, which is able to reduce uneven illumination effect at the same time

#### Feature Extraction

The feature is a set of data that represents the information in an image. Extraction of facial feature is most essential for face recognition. However, selection of features could be an arduous task. Feature extraction algorithm has to be consistent and stable over a variety of changes in order to give high accuracy result

PCA is famous with its robust and high speed computation. Basically, PCA retains data variation and remove unnecessary existing correlations among the original features. PCA is basically a dimension reduction algorithm. It compresses each facial image which is represented by the matrix into single column vector. Furthermore, PCA removes average value from image to centralize the image data. The Principle Component of distribution of facial images is known as Eigen faces. Every single facial image from training set contributes to Eigen faces. As a result, Eigen face encodes best variation among known facial images. Training images and test images are then projected onto Eigen face space to obtain projected training images and projected test image respectively. Euclidean distance is computed by comparing the distance between projected training images and projected test image to perform the recognition. PCA feature extraction process includes all trained facial images. Hence, the extracted feature contains correlation between facial images in the training set and the result of recognition of PCA highly depends on training set image.

Diagram

Description automatically generated

LDA (Linear discriminant analysis) also known as Fisher face is another popular algorithm for face recognition. LDA extract features by grouping images of the same class and separate images of different classes. LDA is able to perform well even with different facial expressions, illumination and pose due to its class separation characteristic. Same class is defined by facial images of the same individual, but with different facial expressions, varying lighting or pose, whereas facial images of person with different identity are categorized as different classes. Same class images yield within-class scatter matrix meanwhile different class images yield between-class scatter matrix. LDA manage to maximize the ratio of the determinant of the between-class scatter matrix over the determinant of the within class scatter matrix. LDA is believed to have lower error rates compared to PCA only if more samples per class are trained and small size of different class. LBP has a few advantages which make it popular to be implemented. It has high tolerance against the monotonic illumination changes and it is able to deal with variety of facial expressions, image rotation and aging of persons

Feature Classification And Face Recognition

Classification involves the process of identification of face. Distance classifier finds the distance between the test image and train image based on the extracted features. The smaller the distance between the input feature points and the trained feature points, the higher the similarity of the test image and training image. In other words, the facial images with the smallest/minimum distance will be classified as the same person. Deepesh Raj (2011) mentioned several types of distance classifiers such as Euclidean Distance, City Block Distance and Mahalanobis distance for face recognition. Md. Abdur Rahim et al. (2013) implemented Chi-Square statistic as distance classifier for LBP operator. The equation of each classification method is defined below. Chi square distance is defined as (2.1) χ 2 = ∑ oberved frequency−expected frequency expected frequency . Chi-square statistic is usually used to compare between two bins of histogram. The City Block Distance or Manhattan Distance is known as L1-norm which is defined in (2.2) 𝑑(𝑥, 𝑦) = |𝑥 − 𝑦| The Euclidean distance is known as L2-norm which is defined in (2.3) 𝑑(𝑥, 𝑦) = |𝑥 − 𝑦| 2 where, X is the input feature points and Y is the trained featured points. 22 The Mahalanobis distance is defined in (2.4) 𝑑(𝑥, 𝑦) = (𝑦 − 𝑚𝑥) 𝑆𝑥 (𝑦 − 𝑚𝑥) 𝑇 where 𝑚𝑥 is mean of x and 𝑆𝑥 is covariance matrix of x. Md. Abdur Rahim et.al (2013), after performing the LBP feature extraction, Chi-Square statistic is suggested to be used as dissimilarity measures for histograms to compute the distance between two images. Abhishek Singh and Saurabh Kumar (2012) proposed Euclidean distance to compute the distance between two images after PCA feature extraction was performed. Threshold can be set for the distance calculated from the classifier. A face is classified as belonging to a class only if its distance is below the chosen threshold, otherwise the face is classified as unknown.

Methodology Flow

The approach performs face recognition based student attendance system. The methodology flow begins with the capture of image by using simple and handy interface, followed by pre-processing of the captured facial images, then feature extraction from the facial images, subjective selection and lastly classification of the facial images to be recognized. Both LBP and PCA feature extraction methods are studied in detail and computed in this proposed approach in order to make comparisons. LBP is enhanced in this approach to reduce the illumination effect. An algorithm to combine enhanced LBP and PCA is also designed for subjective selection in order to increase the accuracy. The details of each stage will be discussed in the following sections. The flow chart for the proposed system is categorized into two parts, first training of images followed by testing images (recognize the unknown input image) shown in Figure 3.1 and Figure 3.2 respectively. 25 Start Conversion of colour image to grayscale image Median filtering Enhanced LBP and PCA feature extraction Contra

## Training database

Start

Read face images from file

Preprocessing



Size of image 250 x250?

No

Yes

Colour Image or Grayscale Image

Colou

Median filtering on 3 channel (R,G,B)

r

Scale to size 250 x250

Grayscale

Enhanced LBP and PCA feature

extraction

Contrast Limited Adaptive Histogram

Equalization(CLAHE)

End

#### Conclusion

1. In this approach, a face recognition based automated student attendance system is thoroughly described. The proposed approach provides a method to identify the individuals by comparing their input image obtained from recording video frame with respect to train image. This proposed approach able to detect and localize face from an input facial image, which is obtained from the recording video frame. Besides, it provides a method in pre-processing stage to enhance the image contrast and reduce the illumination effect. Extraction of features from the facial image is performed by applying both LBP and PCA. The algorithm designed to combine LBP and PCA able to stabilize the system by giving consistent results. The accuracy of this proposed approach is 100 % for high-quality images, 92.31 % for low-quality images and 95.76 % of Yale face database when two images per person are trained.
2. As a conclusion for analysis, the extraction of facial feature could be challenging especially in different lighting. In pre-processing stage, Contrast Limited Adaptive Histogram Equalization (CLAHE) able to reduce the illumination effect. CLAHE perform better compared to histogram equalization in terms of contrast improvement. Enhanced LBP with larger radius size specifically, radius size two, perform better compared to original LBP operator, with less affected by illumination and more consistent compared to other radius sizes.

#### Recommendation

In this proposed approach,there are a few limitations. First,the input image has to be frontal and a upright single facial image. Second,the accuracy might drop under extreme illumination problem. Third,false recognition might occur if the captured image is blurred. Besides, LBP is textural based descriptor which extracts local features. Hence, test image and train image have to be the same quality which is captured by using the same device in order to have high accuracy. Lastly, if an individual wears make up in the image for face recognition, the important features will be covered.

In fact, a better camera with a better lighting source able to reduce the illumination problem and also able to avoid the captured of blurred images. In this proposed approach, laptop built in camera is a default device. However the lighting source of the laptop camera is very dim, this cause the system to be unstable. For future work, a better camera and a better lighting source can be used in order to obtain better result. This can reduce the dependency on the brightness of environment, especially the places to capture test and train images. Furthermore, a face recognition system which has more faces other than a single facial image can be designed. This can increase the efficiency of the system. The test image and train image in this approach is highly related to each other and highly dependent on the image captured device. The capture device has to be the same for this approach to perform better. Thus, other algorithms can be used instead of LBP, for example A.I (artificial intelligence) algorithm which can be implemented to perform the face recognition. CNN (Convolution Neural Network) which is a hot topic recently, is a machine deep learning algorithm which is able to perform recognition with less dependency on a particular train image given a large database. However, CNN requires an extremely large database to increase its accuracy or having relatively small class size to have high performance.

In pre-processing stage, an algorithm, for instance affine transform can be applied to align the facial image based on coordinates in the middle of the eyes. This might help, especially in PCA algorithm, which it maps test image to train image to perform face recognition.