IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

AUTOMATIC FACE DETECTION ATTANDENCE SYSTEM USING EXCEL SHEET

AUTHORS' NAMES (BABU AMAN RAI SAXENA)
Department (Information Technology)
University (Jawaharlal Nehru Technological University)
Address (Ashok Nagar, Kukatpally, Hyderabad, Telangana 500085)
COUNTRY (India)

Abstract: - Recording attendance has been a part of our daily life, several methods are used for evaluating attendance, one of the basic methods is to use a book and call out names. New approaches are being used such as biometrics to assign attendance, although these techniques are dependable and efficient, they have disadvantages, in case of biometrics the major disadvantage is its cost of implementation which is higher. Here we are proposing the use of face detection for recording attendance. Although the cost would be high but, to reduce it, we are proposing a model where the memory cost can be reduced and this model can run on legacy systems thereby reducing the cost of implementation. This paper will show and test the use of face detection for marking attendance and the unique element about this model will be it's use of excel sheet instead of database. Although we cannot eliminate the database completely but, we can use normal data allocation method in any the folder of the computer instead of database. Although there are several methods to compare faces, we are using the eigen faces algorithm which uses the PCA (Principal Component Analysis), we have demonstrated this model using OpenCV library of python.

Key-Words: -Haar cascade classifier, OpenCV, Adaboost, Local Binary Histogram Pattern, FisherFaces Face Recognizer, EigenFaces Face Recognizer.

1 Introduction

1.1 Background Introduction

The present method that institutions works is that the faculty passes an attendance sheet or makes roll calls and mark the attendance of students, which occasionally disturbs the discipline of class and this sheet further goes to the admin department, which is then updated to automation on university. This process is hectic and time-consuming. Also, for professors or employees at institutes, the biometric system serves everyone one at a time. So, why not shift to an advance automated attendance system which functions on face recognition technique? Be it a classroom or entrance gates it will mark the attendance of the employees, professors, students etc.

1.2 Introduction to Face Recognition

Face recognition is technology which is used to identify a person from a video or photo source [1]. In the 1960's face recognition technique was introduced by Woodrow Wilson Bledsoe. Bledsoe developed a device that can classify photos of faces by hand using RAND tablet, a device that people could use to input horizontal and vertical coordinates on grid using a pen like stylus that emitted electromagnetic pulses. Ever since then the recognition system is being improved and optimized constantly, the technology becomes gradually mature and widely used in human daily life. It has been used increasingly for forensics by law enforcement and military professionals. This system contains three modules which are detection, training and recognition. Basically, the detection module detects the face, which gets in the field of visi1on of camera and saves the face in form of image in JPG format. Then the training modules trains the system using Haar-cascade algorithm [12].

1.3 Haar Cascade Classifier

It is a machine learning based method where a cascade function is trained from lot of positive (images with a face) and negative images (images without a face). The algorithm is proposed by scientists Paul Viola and Michael Jones.
The algorithm has four stages:

1.Haar Feature Selection: Haar features are calculated in the subsections of the input image. The difference between sum of pixel intensities of adjacent rectangular regions is calculated to differentiate the subsections of the image. A large number of haar are required for getting facial features[2].
2.Creating an Integral Image: Too much computation will be done when operations are performed on all pixels, so an integral image is used that reduce the computation to only four pixels. This makes the algorithm quite fast [6].
3.Adaboost: All the computed features are not relevant for the classification purpose. Adaboost is used to classify some of relevant features [3].
4.Cascading Classifiers: Now we can use the relevant features to classify a face from a non-face but algorithm provides another improvement using the concept of cascades of classifiers.
5 EigenFaces Face Recognizer : Every region of the image is not a facial region so it is not useful to apply all features on all the areas of the image. Instead of using all features at a time, group the features into different stages of the classifier. Apply each stage one by one in way to find the common facial region. If on any stage the classifier fails, then that region will be discarded from further iterations. Only the facial region will pass all stages of the classifier.

1.4 OpenCV Face Recognizers

OpenCV consists of three built face recognizers and you can use any of them by just changing a single line of code[3]. Below will be the names listed below of those face recognizers and their OpenCV calls.

1. EigenFaces Face Recognizer - cv2face.createEigenFaceRecognizer()
2. FisherFaces Face Recognizer - cv2.face.createFisherFaceRecognizer()
3. Local Binary Histograms- cv2.face.createLBPHFaceRecognizer()

1.5 EigenFaces Face Recognizer

This algorithm accepts the element that not all parts of face are equally important and equally useful. When you look at some one, you recognize him by his distinct features like eyes, nose, cheeks, forehead and how they differ with respect to each other. For example, from eyes to nose there is a significant change and same is case from nose to mouth[1,10]. When you look at multiple faces you compare them by looking at their parts of the faces because these parts are the most useful and important components of a face. In any case they are important because they catch the maximum change among faces, change the helps you differentiate one face from the other. This is exactly how EigenFaces works. These important components it extracts are called the principal components.

1.6 FisherFaces Face Recognizer

Eigenfaces face recognizer looks at all training faces of all persons at once and finds
the principal components from all of them combined. By capturing principal components from all of them combined you are not focusing on the features that discriminate one person from the other but the features that will represent all the persons in the training data as a whole FisherFace algorithm. Instead of extracting useful features that represent all the faces of all persons, it extracts useful features that discriminate one person from others. This way features of one person do not dominate over the others and you have the features that discriminate one person from others.

1.7 Local Binary Histogram Pattern

Local Binary Histogram Pattern is simple yet very effective texture operator which labels the pixels of an image by thresholding the neighbor area of each pixel and ponders the result as a binary number. It was first described in the year 1994. Local Binary Histogram Pattern has been found to be a powerful feature for all texture classification. It has further been evaluated that when Local Binary Histogram Pattern is combined with all histograms of oriented gradients signifier, it advances the detection performance significantly on some data sets. Using the Local Binary Histogram Pattern combined with histograms we can represent the face images with a simple data vector. As Local Binary Histogram Pattern is a visual descriptor it can also be used for face recognition, as can be seen in the following step-by-step explanation.

Face detection and recognition is technology which is being used to identify a faces or person from a video or photo source. In the 1960's face recognition was introduced by Woodrow Wilson Bledsoe. Bledsoe developed a device that could easily classify photos of faces by hand using what's known as a RAND tablet, a device that people could use to involvement of the horizontal and vertical coordinates on grid using a stylus that emitted electromagnetic pulse. From that point of time, the recognition system is being improved and optimized constantly. The technology becomes gradually mature and is more and more widely used in daily life. It has been used increasingly for forensics, law enforcement and military professionals. This system consists of three modules which are detection, training and recognition. Basically, the detection module detects the face which gets into the field of vision of camera

and saves the face in form of an image in JPG format. Then the training modules trains the system using the Haar cascade algorithm.

2. Problem Solution

2.1 Existing System

At present, attendance system, marking attendance involves manual attendance on the paper sheet by professors and teachers, but it is a very time-taking process and chances of fake attendance are also an issue that arises in such type of attendance marking techniques. Also, there is an attendance marking technique such as RFID, Biometrics etc. But these systems are currently not that popular in schools and classrooms for students and are very costly to implement.

2.2 Drawbacks in existing system

Manual arrangements put more pressure on people to be correct in all the details of their work at all the times, the problem is that people are not perfect.

- These attendance arrangements are manual.
- There is chance of forgery (one person can sign the presence of the other one) Since these systems are manually so there is a high risk of error.
- Extra manpower is required.
- Calculations linked to attendance are done manually (total classes attended in a month) which is usually open to errors.
- It is difficult to keep the database or register for the manual systems.
- It is difficult to search for a particular data from this system (especially if that data, we are asking for, is of very long ago).
- The ability to calculate the attendance ratio becomes big task as manual calculation produces errors, and also wastes a lot of time.
- This method could simply allow for impersonation and the attendance sheet can be stolen or lost.

2.3 Excel Sheets:

Excel sheets are widely being used for allocation of data and these do not require a person to know SQL command to retrieve data. The main purpose of introduction of Excel Sheets is to remove the database functionality thereby, reducing the memory costs and make easy for the normal computers to be able to access the given attendance, the major part of this project depends on excel sheet. Making the new excel sheets every class hour is beneficial as it not only gives the in-charge a relief but also gives the subject faculty more time to concentrate on teaching. As the recorded attendance will be directly accessible in the given laptops or desktops.

3 Proposed Solution:

To overcome the problems in the existing attendance system we shall develop a excel sheet-based face detection attendance system over simple attendance system. There are many solutions to automate the attendance management system like thumb-based system, simple computerized attendance system, Iris scanner, but all these systems have limitations overwork and security point of view. Our proposed system shall be a “Face Recognition Attendance System” which uses the basic idea of image processing which is used in many security applications like banks, airports, Intelligence agencies etc.

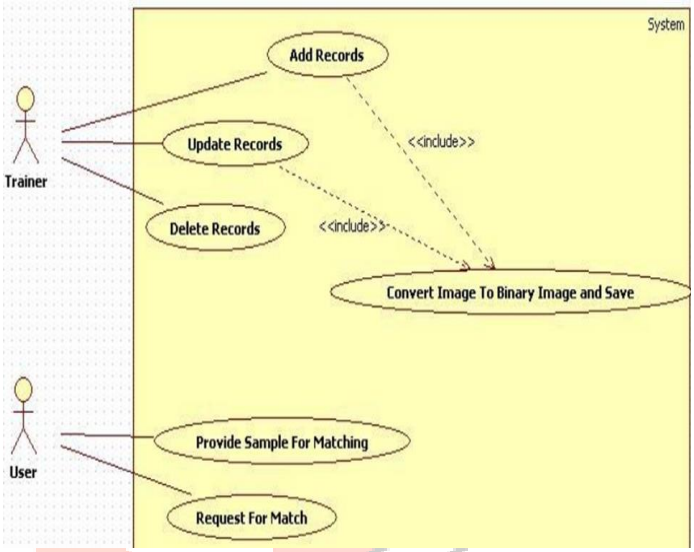

Figure 1: Use Case Diagram for proposed system

3.1 Proposed System Components:

These are the main components of the proposed attendance system

1. Student Registration
2. Face Detection
3. Face Recognition
   - Feature-Extraction
   - Feature-Classification
4. Attendance marking system
   Attendance management will handle:
   - Automated Attendance recording
   - Manual Attendance recording

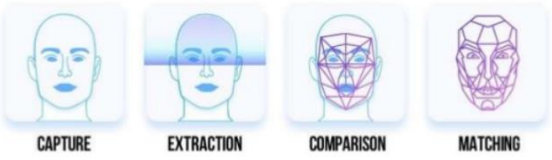

Fig 1: Step by Step process of detecting faces

3.2 DESIGN METHODOLOGY

3.2.1 SYSTEM ARCHITECTURE

As we can see in the below system architecture, firstly we are going to require a camera in the classroom to capture the images of the students. And that camera will be sending that image data to the required system or a basic computer or laptop based on the above requirements. On that particular laptop the required program of python face recognition will run and it will display the student present in an excel sheet.

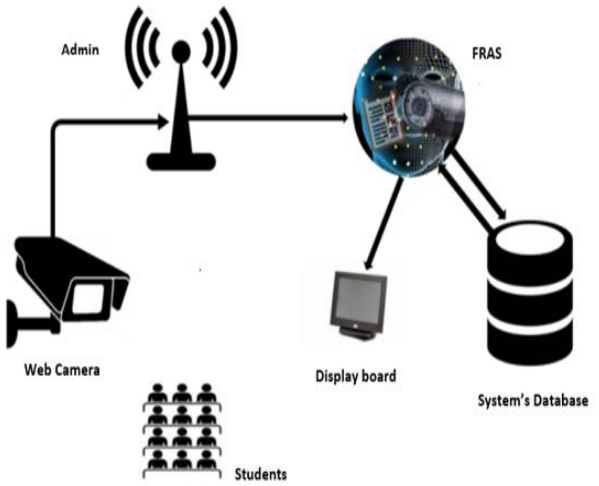

Fig 2 proposed hardware architecture of face detection

3.2.2 MODULES

There are three main modules in this system:

- Training Model on Images.
- Face Detection & Face Recognition

Training Model:

To train model on images first we should have a basic idea of an image, image is a standard Numpy array having pixels of data points. More the number of pixels in the image, better is its resolution quality. You can think of pixels to be tiny blocks of information arranged in form a 2 D grid and the depth of a pixel refers to the colour information present in it. In order to be processed by a processer, an image needs to convert into a binary system.

Eigenfaces:

The problem with the image depiction is that we are given a high dimensionality. Two-dimensional p×q grayscale images span a m=pq-dimensional vector, so an image with 100×100 pixels lies in a 10,000- dimensional image space. If there is any variance in data, so what we are looking for is the components that account for much of the information. The Principal Component Analysis (PCA) was autonomously proposed by Karl Pearson and Harold Hotelling. The idea was that a high-dimensional dataset that is often described by correlated variable quantity and therefore only few dimensions account for most of the information. The PCA method will find directions with the greatest variance in the data, called principal components.Let X={x1,x2,…,xn} be random vector with the observations as xi∈Rd.

1. Compute the mean μ

$$\mu = \frac{1}{n}\sum_{i=1}^{n} x_i$$

2. Compute the Covariance Matrix S

$$S = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)(x_i - \mu)^{T},$$

3. Calculate the eigenvalues λi and eigenvectors vi of S

$$Sv_i = \lambda_i v_i, i = 1, 2, \ldots, n$$

4. Order the eigenvectors as descending by their eigenvalue. The k principal mechanisms are the eigenvectors matching to the k largest eigenvalues.
The k principal elements of the observed vector x are then given by:

$$y = W^T(x - \mu)$$

5. where W=(v1,v2,…,vk).
The rebuilding from the PCA basis is given by:

$$x = Wy + \mu$$

where W=(v1,v2,...,vk).

Fisherfaces:

The Principal Component Analysis (PCA) is the main component of the Eigenface method in the face detection, it discovers a combination of features that will maximize the total variance in the data. While this is undoubtedly a good way to represent the data, it doesn't consider any classes and thus a lot of discriminative information can be lost when throwing components away. Imagine a condition where the variance in your data is been produced by an external source. Components that are recognized by the PCA do not always contain any discriminatory data at all, the expected samples are tarnished together and a classification becomes almost impossible.

Let X be the random vector which has samples drawn from c classes:

$$X = \{X_1, X_2, \ldots, X_c\}$$
$$X_i = \{x_1, x_2, \ldots, x_n\}$$

The matrices SB and S_{W} are calculated as:

$$S_B = \sum_{i=1}^{c} N_i(\mu_i - \mu)(\mu_i - \mu)^T$$
$$S_W = \sum_{i=1}^{c} \sum_{x_j \in X_i} (x_j - \mu_i)(x_j - \mu_i)^T$$

where μ is the total mean:

$$\mu = \frac{1}{N} \sum_{i=1}^{N} x_i$$

μi is mean of class i∈{1,…,c}:

$$\mu_i = \frac{1}{|X_i|} \sum_{x_j \in X_i} x_j$$

Fisher's classic algorithm will look for the projection W, that will maximize the class separability:

$$W_{opt} = \arg\max_W \frac{|W^T S_B W|}{|W^T S_W W|}$$

Following the solution for optimization problem is given by solving the common Eigenvalue Problem:

$$S_B v_i = \lambda_i S_w v_i$$
$$S_W^{-1} S_B v_i = \lambda_i v_i$$

The rank of SW is the most (N−c), with N samples and the c classes. In the pattern recognition problem the number of examples N is always smaller than the measurement of the input data (the number of pixels), so the scatter matrix SW becomes a singular. In this was solved by performing PCA on the data and projecting the samples into the (N−c)-dimensional space.

The optimization solution can be rewritten as:

$$W_{pca} = \arg\max_W |W^T S_T W|$$
$$W_{fld} = \arg\max_W \frac{|W^T W_{pca}^T S_B W_{pca} W|}{|W^T W_{pca}^T S_W W_{pca} W|}$$

The transformation matrix W, that projects a sample into the (c−1)-dimensional space is given by:

$$W = W_{fld}^T W_{pca}^T$$

FACE DETECTION:

This is the concept of recognizing faces in Python using OpenCV library and you can utilize it to perform tasks like Facial detection. Here we are going to see the algorithm usage for the face detection.

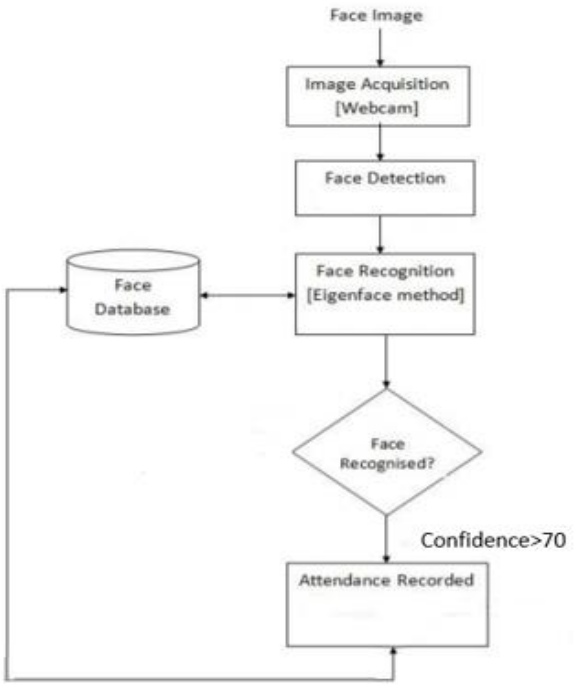

Figure 5: Block Diagram of Face Detection.

4 Proposed System Outcome

•It will mark the attendance of students via a face Id.
•It will detect faces via wireless camera (IP camera)/webcam and then recognizes the faces.
•After recognition, it will mark the attendance of the recognized student and then it updates the attendance record.
•The admin will be able to print these record details afterwards

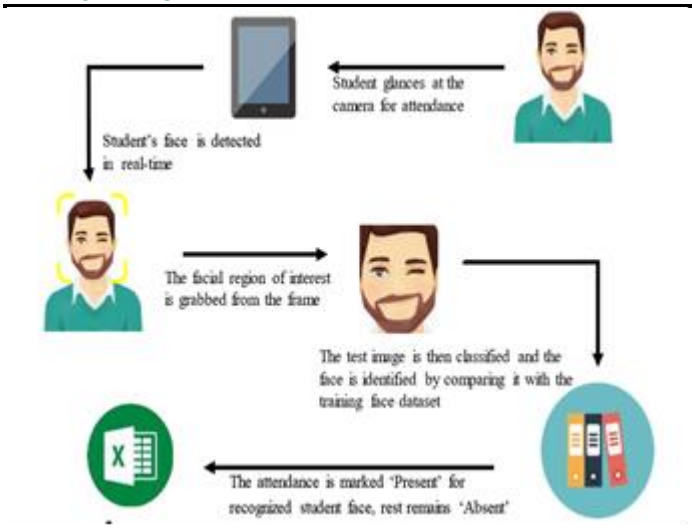

Figure 3: Generating Attendance Flow

4.1 Future Scope

- Provides a facility for the automated attendance of students.
- Uses the live face recognition to recognize each individual and mark their attendance automatically.
- Utilizes video and image processing technique to provide inputs to the system.
- Facility of marking a manual attendance.

5 Results:

Using the above process we successfully utilized the face detection algorithm for creating advance attendance system and we replaced the notion of having a database for working with face detection, in this process we also use excel sheets for showing attendance.

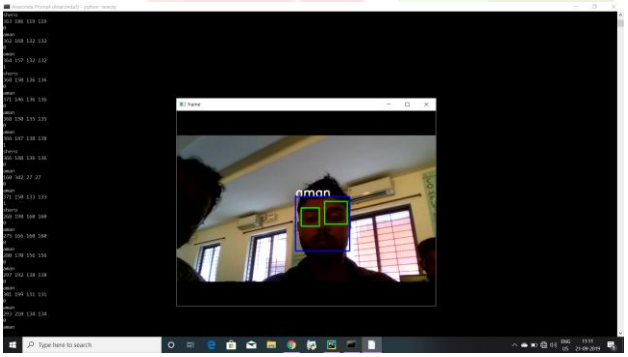

Fig 6: Real time face detection using eigen faces

The below figure shows the use of excel sheets for marking attandace this is done by using the python language which contains inbuilt classes for assigning attandance, here we can see in the below figure that the name and ID of the student are kept side by side for the display making the use of excel sheet to provide clarity and caplability to share the information with respective class in charges.

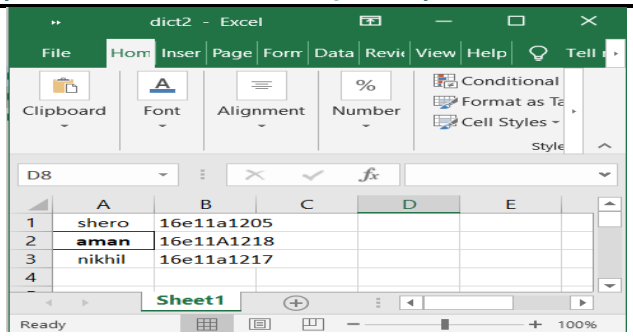

Fig 7: Automated attendance using excel sheet

6 Conclusion

Although face detection technique is useful for the attendance allocation but, there are ways through which we can get vulnerability in the system but reducing this is the primary aspect on any model, for the use of face detection algorithm we have used OpenCV library of python which has Haar cascade algorithm, the performance of this system is done through a model which created this working of this paper. We successfully created a system which can record attendance on excel sheets thereby, making the working of this paper accessible on conventional systems. In the future we would work on the face detection where we can add this technique on the systems with lower specifications.

References:

[1] O. M. Parkhi, A. Vedaldi, A. Zisserman Deep Face Recognition British Machine Vision Conference, 2015

[2] H.-W. Ng, S. Winkler. A data-driven approach to cleaning large face datasets. Proc. IEEE International Conference on Image Processing (ICIP), Paris, France, Oct. 27-30, 2014

[3] Beyeler, M. (2017). Machine Learning for OpenCV: Intelligent image processing with Python. London: Packt Publishing Ltd.

[4] Bruce, P., & Bruce, A. (2017). Practical Statistics for Data Scientists. California: O'Reilly Media.

[5] Chollet, F. (2018). Deep Learning with Python. New York: Manning Publications Co.

[6] Géron, A. (2017). Hands-On Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems. Brooklyn: O'Reilly.

[7] Grus, J. (2015). Data Science from Scratch: First Principles with Python. California: O'Reilly Media.

[8] Lutz, M. (2013). Learning Python. California: O'Reilly Media.

[9] Matthes, E. (2015). A Hands-On, Project-Based Introduction to Programming. San Francisco: William Pollock.

[10] [10] McKinney, W. (2017). Python for Data Analysis: Data Wrangling with

Pandas, NumPy, and IPython. California: O'Reilly Media

[11] [Phillips, D. (2014). Creating Apps in Kivy: Mobile with Python. California: O'Reilly.

[12] Solem, J. E. (2012). Programming Computer Vision with Python: Tools and algorithms for analyzing images. United States: O'Reilly Media.