Solving for Pose Variation in Group Face Recognition for Attendance Marking using Real Time Video

Lillmond Chandesh
StuID: 1903_17238

Daboo Mohesh
StuID: 1903_16386

Beeharry Tanweeruddin
StuID: 1903_16883

Bundhoo Vasish
StuID: 1903_16871

***Abstract*—One of the main problems in group face recognition is that some of the faces are too close to each other [11]. Thus, not all faces are identified. The aim is to recognize all faces in a group although there is pose variations. As methodology, OpenCV [6,9] is used as technique and Haar Cascade [12], HOG [11] and real time face recognition [9] as algorithm. We adapted a machine learning approach to help recognize human faces in real-time and were able to successfully detect multiple frontal faces from a live video. As result we were able to increase the accuracy level of group face recognition, up to 85.7% when faces are close to each other in real time. [11].**

Keywords— LBPH, OpenCV, Multiple Face Detection, Haar Cascade, HOG, Face Recognition, Pose variation, Real Time Video

I. INTRODUCTION

Group face recognition for attendance purposes has resulted in poor attendance marking due to misidentification [1]. This happens due to pose variations. Pose variations occurs in uncontrolled environments where there are variations in pose, light intensity and expression. It also occurs when people are close to each other [11]. Pose variation makes the system difficult to distinguish and detect faces.

However, many researches of group face recognition are done in controlled situations. Compared to the controlled environments, face recognition in uncontrolled environments comprise more variations [11] techniques. Different and algorithms have been used for group face recognition such as Viola-Jones Algorithm. The problem of using it is that it was unable to detect faces in other orientation or configurations (tilted, upside down, wearing a mask) [2]. Moreover, LBPH algorithm (Local Binary Patterns Histogram algorithm) and Haar Cascade classifier technique, but it is not effective as the system fails in real time [3]. Furthermore, MathLib technique was used and it resulted in 75% acceptance detection ratio. Therefore, it is not accurate in detecting faces [4]. 3D - based face recognition was implemented in order to solve pose variations, but It failed when it is used in uncontrolled or semi - controlled environments (illumination problems) [5].

Given that group face recognition for attendance marking still have misidentification problem, therefore, the aim of this research paper is to provide a solution by using more accurate algorithms and techniques to solve the problems.

II. LITERATURE REVIEW

A. Introduction

During attendance marking, group face recognition techniques are not able to identify students because their faces are close to each other [11]. Therefore, the aim of this paper is to recognize all faces clearly. This literature review describes the problem faced during real time face recognition as well as the different techniques and algorithms used to tackle the issues.

B. Problem

During lectures, when students are too close to each other, it is difficult to recognize their faces. This is because the techniques are unable to detect their faces in real time [11]. In addition, facial expression also leads to misidentification [1].

C. Technique

LBPH algorithm in its enhanced form with the use of OPENCV 2.4. OPENCV 2.4 has a new class called Face-Recognizer for the purpose of group face recognition. LBPH is a much enhanced and modified form of LBP algorithm which is notably known in providing good texture description. It has 83.7% of accuracy dealing with 0o face orientation and 0% of accuracy dealing with 90o face orientation. [6, 9]

For group face recognition, authors [7] used different algorithms. PCA and Eigen face algorithms were used to detect faces. Eigen with PCA has a success rate of 92.30% when identifying multiple faces but its recognition rate decreases for recognition under varying pose and illumination. [7].

Real time face recognition system is much accurate, precise, secure and reliable. However, it needs improvement in various lightning conditions [9].

OpenFace is a face recognition library. It is based on Google's FaceNet systems. It is very reliable when dealing with large number of subjects. OpenFace was trained with 500k images. It has an accuracy of 90% when dealing with 25 subjects at one time. Its performance drops when it is used in uncontrolled environments [10].

Histogram of Oriented Gradients (HOG) is another method for detecting objects which can also be used for detecting multiple faces. The strength of this method is that it

is not sensitive to a change in luminosity [11]. The disadvantage with HOG based face detection is that it doesn't work on faces at odd angles, it only works with straight and front faces and it is not a good fit for real-time video [11].

D. Implementation

This technique is implemented by using face extraction and real time video, a CCTV camera is used for capturing the faces of students

Dataset Creation: Dataset of students is created before the recognition process. Dataset is created only to train this system. We have created a dataset of 4 students which involves their name and images of student in different poses and variations. Whenever we register student's data and images in our system to create dataset, deep learning applies to each face to compute 128-d facial features and store in student face data file to recall that face in recognition process. This process is applying to each image taken during registration.

Face Detection and Extraction: Face detection is important as the image taken through the camera given to the system, face detection algorithm applies to identify the human faces in that image, the number of image processing algorithms are introducing to detect faces in an image and also the location of that detected faces. We have used HOG method to detect human faces in given image.

Face Encoding: Once the faces are detected in the given image, the next step is to extract the unique identifying facial feature for each image. Basically, whenever we get localization of face, the 128 key facial point are extracted for each image given input which are highly accurate and these 128-d facial points are stored in data file for face recognition.

Face Recognition: This is last step of face recognition process. We have used the one of the best learning techniques that is deep metric learning which is highly accurate and capable of outputting real value feature vector. Our system ratifies the faces, constructing the 128d embedding (ratification) for each. If the current image is matched with the 60% threshold with the existing dataset, it will move to attendance marking.

Attendance Marking: Once the face is identified with the image stored in the dataset, python generate roll numbers of present students and return that, when data is returned, the system generates attendance excel sheet.

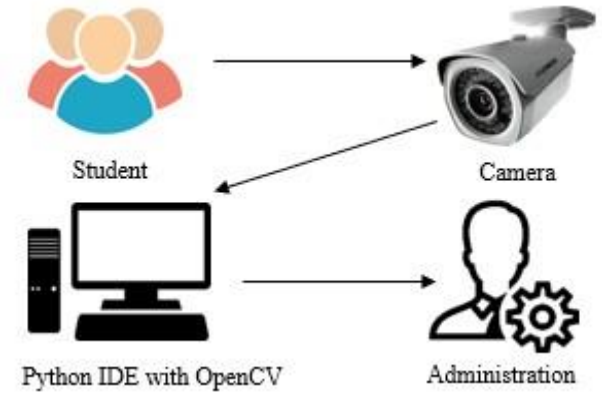
III. METHODOLOGY

A. Requirement Definition

Given that we are solving the problem of pose variation [1]. We will be using OpenCV [6,9] as technique and Haar Cascade [12], HOG [11] and real time face recognition [9] as algorithm. Whenever time for corresponding subject arrives the system automatically starts taking snaps and find whether human faces are appearing in the real time video. We have used Histogram of Oriented Gradient for face detection and deep learning techniques to calculate and compare 128 face features for face recognition. Once faces are detected and recognize with the existing dataset, the system calculate attendance for the recognize students with the respective subject id in real time. Then an excel sheet is generated.

B. Technique

The proposed model is based on face recognition mechanism. The basic system is presented in the diagram below. Whenever a student comes across the camera module of the system his image is captured by the system. If recognized then his attendance is automatically marked via post processing of the system else if not recognized, the student's attendance will not be marked.



The system will find the faces in the live streaming video. This is done using HOG (Histogram of Oriented Gradients) at the backend. Once we have the faces, they are warped to remove unwanted rotations. Then the image is feed to a pretrained neural network that out puts 128 measurements that are unique to that particular face. The parts that the model measures are not known as this is what the model learns by itself when it was trained. Once we have the face locations and the encodings, the system will proceed to Haar Cascade.

Haar cascade features shown in below image are used. Each feature is a single value obtained by subtracting sum of pixels under white rectangle from sum of pixels under black rectangle.

Now all possible sizes and locations of each kernel is used to calculate plenty of features. For each feature calculation, Haar features need to find sum of pixels under white and black rectangles. To solve this, they introduced the integral images.

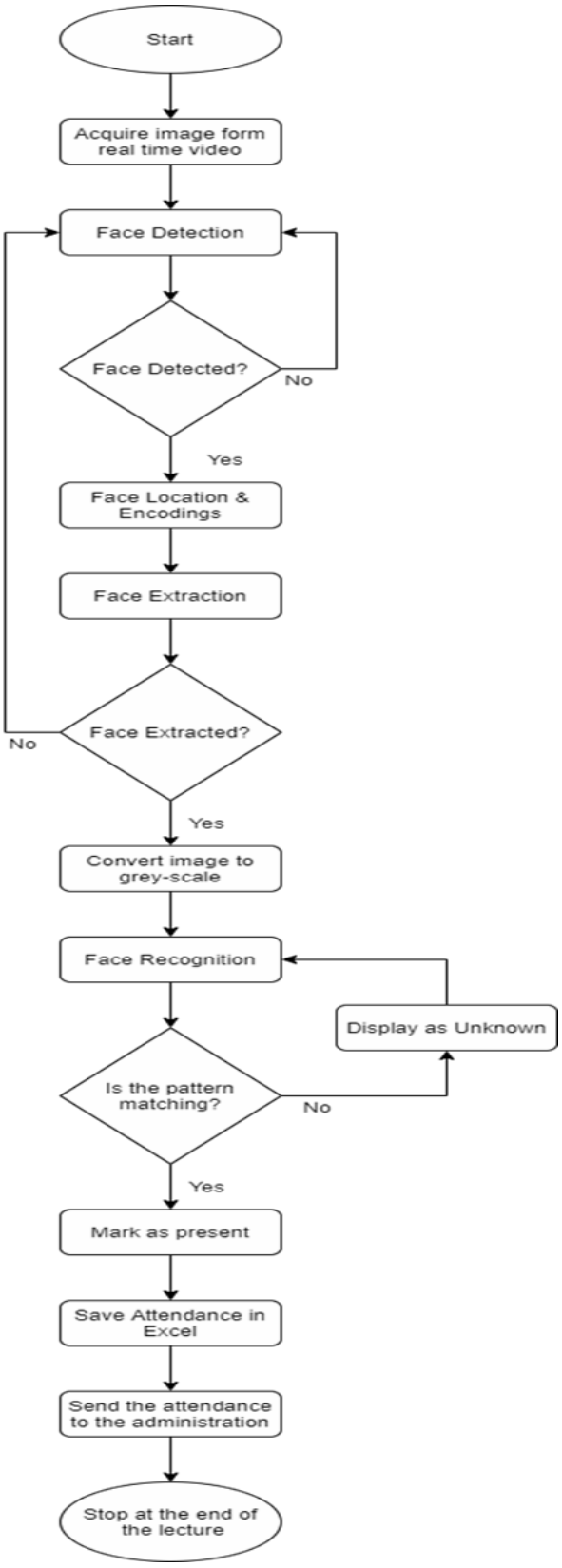
The second feature selected relies on the property that the eyes are darker than the bridge of the nose. But the same windows applying on cheeks or any other place is irrelevant

For this, we apply each and every feature on all the training images. For each feature, it finds the best threshold which will classify the faces to positive and negative. Haar cascade select the features with minimum error rate, which means they are the features that best classifies the face and non-face images. Each image is given an equal weight in the beginning. After each classification, weights of misclassified images are increased. Then again same process is done.

New error rates are calculated. Also new weights. The process is continued until required accuracy or error rate is achieved or required number of features are found.

The flowchart below shows an overview of the system and the steps to recognize group faces.

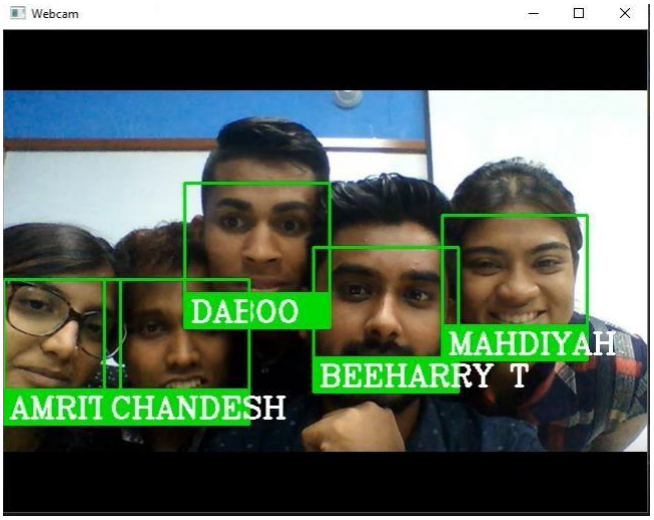
Flowchart



Once the program has launched, the camera will start detecting faces.
If the faces are detected, the features of faces are extracted and convert to gray-scale image.
Then, the patterns are compared to the ones in the dataset.
If the patterns match then the faces are recognized and are mark as present in attendance.
The attendance is saved on an excel sheet.

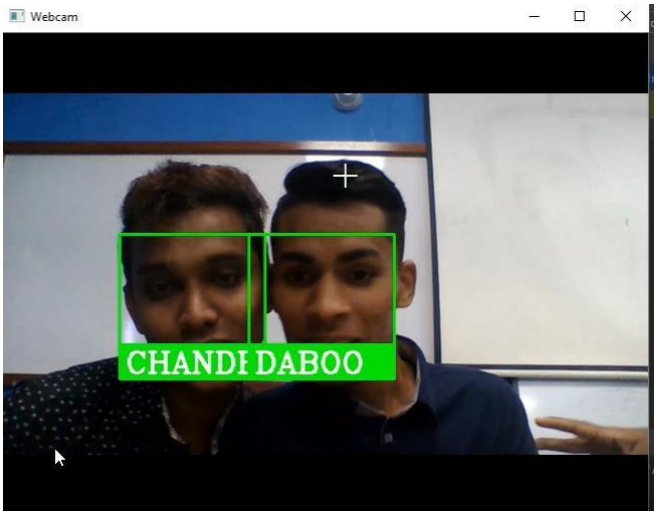
C. Testing

Attendance system proved to recognize faces when they are close to each other. The faces which are not in our training dataset are marked as unknown. The attendance of recognize images of students is marked in real time. And import to excel sheet and saved by the system automatically.



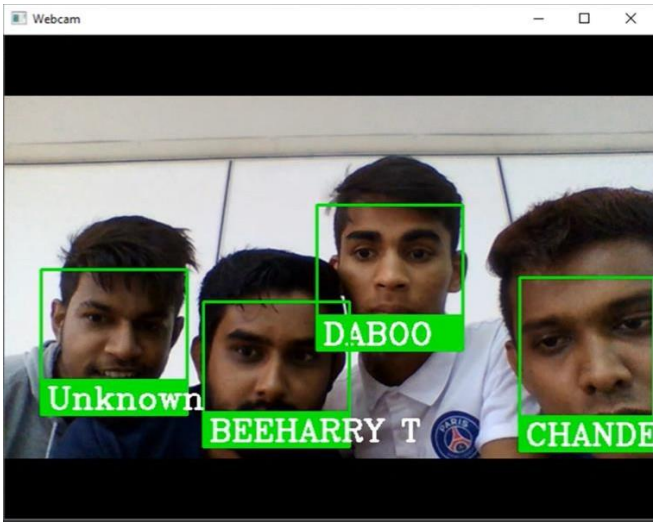
Distances between faces: Nearly 0cm

This picture above shows that the system was able to recognise five faces close to each other. The distance between the faces was nearly 0cm.



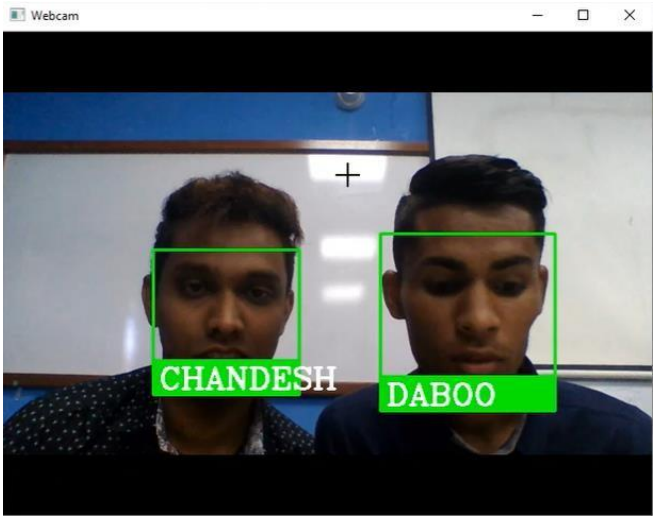
Distance between face: 1cm

This picture shows that the system was able to recognise two faces close to each other. The distance between the faces was 1cm.



Distance between faces: 3cm

This picture shows that the system was able to recognise four faces close to each other. The distance between the faces was 3cm.



Distance between faces: 5cm or more

This picture shows that the system was able to recognise five faces close to each other. The distance between the faces was 5cm or more.

Equation:

Accuracy = (SUM SUCCESS / SUM RECOGNIZABLE) * 100

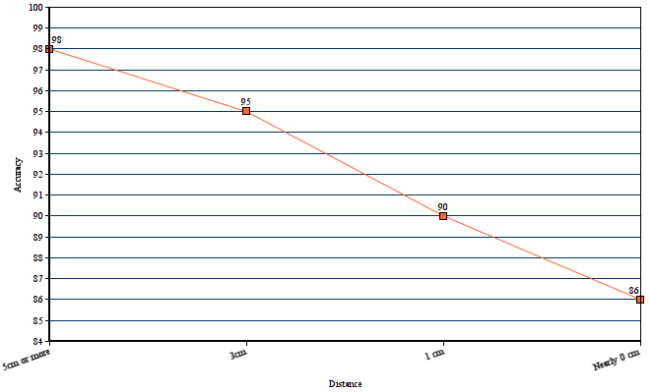
Accuracy = (12 / 14) * 100 = 85.7%

Test	Accuracy Level		
	Distance between faces	***Face Detection***	***Face Recognition***
1	5cm or more	99%	98%
2	3cm	96%	95%
3	1cm	93%	90%
4	Nearly 0cm	87%	85.7%

We performed a set of experiments to demonstrate the efficiency of the proposed system. We tested the system with a group of faces at different distances to each other.

D. Result

This graph below shows the accuracy level of different distance between the faces.



From the graph above it is been observed that with the decreasing distance between faces, face detection and recognition rate is become decreases. There are also variations in pose, light intensity and expression that can affect the system accuracy.

The test carried out shows that Haar Cascade is robust to recognize faces close to each other. From the conducted experiments it has been observed that other algorithms are more robust to illumination factor in general while LDA has been proven more robust to facial expression factor. [11]

S.No	Algorithm	Mean accuracy±Std
1	LBPH	84.242±0.0562
2	LDA	83.636±0.099
3	PCA	75.151±0.083

Accuracy Level with different algorithms [13]

From our survey, 96% of the users said that the system was able to detect and recognize their faces accurately in an uncontrolled environment. The most frequent problem the users faced was, the camera takes too much time to lauch.

IV. Conclusion

A. Achievement

Our algorithms can detect and recognize faces when they are close to each other with an accuracy of 85.7% in real-time. We have used face recognition concept to mark the attendance of student and make the system better.

B. Weakness

The camera takes time to launch as it trains the images when the system starts. Due to the low density of pixel of a webcam, the system takes time to detect multiple faces. Poor lighting conditions may affect image quality which indirectly degrades system performance.

C. Futurework

The future work is to improve the recognition rate of algorithms when there are unintentional changes in a person like tonsuring head, using scarf, beard. It also needs improve where the system sometimes fails to recognize students from some distance, also we have some processing limitation, working with a system of high processing may result even better performance of this system.

References

[1] Muhammad Sharif, Farah Naz, Mussarat Yasmin, Muhammad Alyas Shahid and Amjad Rehman, "Face Recognition: A Survey," March 2017

[2] Yashwant, "Various Techniques used for Face Recognition", 2020

[3] Shashwat Dobhal, Avinash Kumar Singh, Ajay Kumar, "Automated Attendance System using Multiple Face Detection and Recognition," April 2020.

[4] Abdelfatah Aref, Omaima N. A. AL-Allaf, Mohammad A. Alia, "Real-Time Group Face-Detection for an Intelligent Class-Attendance System," May 2015.

[5] Waldemar Wójcik, Konrad Gromaszek and Muhtar Junisbekov, "Face Recognition: Issues, Methods and Alternative Applications," July 2016.

[6] Yasaman Heydarzadeh, Abol Fazl ToroghiH Aghighat, "An Efficient Face Detection Method using AdaBoost and Facial Parts," August 2011.

[7] A. M. Patil, Dr. Satish R. Kolhe, Dr. Pradeep M. Pati, "Face Recognition by PCA Technique," Second International Conference on Emerging Trends in Engineering &Technology (ICETET) Dec 2019.

[8] T. A. Roshan Tharanga and S. M. Samarakoon, 'Smart attendance using real time face recognition,' 2013

[9] Nirmalya Kar, Dr. Mrinal Kanti Deb Barma, Ashim Saha, "Study of Implementing Automated Attendance System Using Face Recognition Technique," Jan 2012

[10] Nicolas Delbiaggio, "A comparison of facial recognition's algorithms," May 2017

[11] Alexander A. S. Gunawan and Reza A. Prasetyo, "Face Recognition Performance in Facing Pose Variation," Aug 2017

[12] Dr.Vipulsangram.K.Kadam , Deepali G. Ganakwar, "Face Detection: A Literature Review," Sept 2020

[13] Ali Rehman Shinwari; Asadullah Jalali Balooch; Ala Abdulhakim Alariki, Sami Abduljalil Abdulhak," A Comparative Study of Face Recognition algorithms under Facial Expression and Illumination," May 2019