Advancing Attendance: A Facial Recognition System Empowered by Deep Learning Techniques

Huda Al-Nayyef*

Department of Computer Science, College of Science, Mustansiriyah University, 10052 Baghdad, Iraq.

ABSTRACT

Many studies during the last decade tried to provide an automatic technique for facial recognition and identification challenge especially in security systems. In this study, we suggested two methods for student attendance problem based on image processing and machine learning algorithms. The first method uses Haar cascade classifier with the Local Binary Patterns Histograms (LBPH) model and the second method composed from Histograms of Oriented Gradient (HoG) followed by the Convolutional Neural Network (CNN) model. Both methods take a collection of random student images taken from low quality sources as input. A set of image processing filters are first applied on images to enhance the method of extracting the face boundary. Then, each model will be trained using random images from student dataset. The trained model is tested using testing set. The results showed that the method that employ CNN model with HoG provides high accuracy value of 98.44%. While, the accuracy of LBPH model with Haar Cascade classifier is 95.63%.

Keywords: Attendance system, Face recognition, Deep learning, convolutional neural networks (CNN), Haar-Cascade classifier, Local Binary Pattern Histogram (LBPH), Histograms of Oriented Gradient (HOG).

1. Introduction

Image processing and facial recognition are two fascinating disciplines right now. Facial recognition is quickly taking the lead in biometrics and fingerprints [1]. In order to identify people, sophisticated software and technology can even evaluate blurry photographs. Face recognition (or facial recognition), is the process of examining features of a person’s picture taken by a camera. When the subject interacts with the camera, it measures their face structure and the separations between their features, storing the results in a database for comparison [2] [3].

The non-intrusive nature of facial recognition is one of its main benefits. It doesn’t need people to wait or do anything more than face the camera in order to confirm or identify them at a distance of two feet or more [4] [5]. Traditionally,
attendance is kept on paper by students using forms that instructors hand them during class. Face detection and face identification are the two primary processes in the face recognition process. Face detection and recognition techniques have been developed in a variety of ways [6]. Appearance-based techniques encompass the full range of facial features, whereas feature-based techniques concentrate on geometric features such as the eyes, nose, eyebrows, and cheeks [7] [8].

This study presents the use of facial recognition technology in an attendance system. By utilizing machine learning techniques, our solution seeks to address the shortcomings of current attendance monitoring systems and ensure the secure and effective recording of lecture attendance. Our main goal is to improve and simplify the procedure for monitoring and controlling student attendance. We also hope to provide students with a dependable attendance service, automate tasks to minimize manual errors, protect privacy and security by avoiding impersonation, accurately detect faces in a variety of environmental elements, and enable automated database updates devoid of human intervention.

The remainder of the paper is organized as follows. Section 2 depicts the state of the art. Section 3 discusses the methodologies used in the proposed attendance system. Section 4 presents the experimental results and analyses. Section 5 presents the conclusion.

2. Related Works

Face recognition technology has become increasingly popular in recent years, particularly in the field of student attendance. The state of the art in face recognition for student attendance includes the following:

- In the work of [9], They contributed a real-time surveillance system based on Raspberry Pi and CNN (Convolutional Neural Network) for facial identification. We presented the system with a labeled dataset. First, the system is trained on the labeled dataset to extract various facial features and recognize landmark faces, and then it compares the query image to the dataset based on features and landmark face detection.

- By using face recognition algorithms to track attendance in distance learning environments, [10] introduced a student attendance system using face recognition technology, which aims to improve success indirectly and increase student engagement. Experiments using specialized image processing filters were used to test the suggested model, and the results showed that when the sample size was increased, accuracy reached above 80%. The design model was implemented using the C# programming language in the Visual Studio.Net environment. The SQL Server was used to handle the database, and phases of image processing were made easier with the EmguCV library.

- In [11], the goal is to propose an Android-based course attendance system that uses facial recognition. To ensure that students attend the course, a QR code containing course information was created and placed at the front of the classroom. The student merely needed to use his or her smartphone to capture an image of his or her face and display a QR code. The image was then sent to the server for attendance purposes. The testing results demonstrate that the suggested attendance system achieved a face recognition accuracy of 97.29 by utilizing linear discriminant analysis and only took 0.000096 seconds to recognize a face image in the server.

- In [12] presented a project of a Smart attendance system. An OpenCV-based facial recognition method is suggested by the project. This system uses a camera to take pictures for input, an algorithm to recognize, encode, and identify faces, a spreadsheet to track attendance, and a PDF file to store the attendance data. Thus, the database is created from the images taken from students. Those images are used to train the former. For training the model, labels from student images are stored in the database too. It is wealth to mention that this study embeds the LBPH model in learning the student images based on its characteristics.

- To improve the accuracy of recognition, the [13] suggested a new technique based mainly on combining the algorithm of Local Binary Pattern (LBP) with image processing filters. The idea behind this study is in the preprocessing by enhancing the students images using Contrast Adjustment, Bilateral Filter, Histogram Equalization, and Image Blending. This idea improves the overall accuracy of the LBP codes and result an accuracy improvement of the overall system.
• The [14] presented a novel system based mainly on tracking the student movement. Thus, the system senses the presents and identify students attendance. This method takes into consideration particular parts on recognizing specific features of the face, such as the nose and eyes. To evaluate the performance of various facial recognition algorithms, the system is tested in a variety of real-time settings. The report also suggests ways to improve the system’s dependability by addressing problems like spoofing and avoiding student proxies. This strategy saves teachers and administrators time by providing a more effective way to manage student attendance than previous techniques.

• The study of [15] presents a Raspberry-based face recognition system. The work focuses on a small-scale face detection and recognition problem for a Plymouth College visual observation and self-governance curriculum. It talks about the features offered by the Open Computer Vision (OpenCV) package and how to use Python to implement them. Haar-Cascades were used for face detection, and Eigenfaces, Fisher faces, and Local Binary Pattern Histograms were used for face identification. The procedure is comprehensive, with flowcharts showing each system step. Plots and images of the results are displayed, and a discussion of the difficulties faced comes next. The writers’ thoughts on the project and its possible uses are included in the paper’s conclusion.

• The study of [16] develops a facial recognition system using deep learning techniques. They computed and compared student face attributes for recognition using deep learning techniques, and they used the histogram of the oriented gradient method for face detection in photos. Their technology was able to recognize many faces in real time.

• In [17], the authors introduced an automated attendance management system. Automating face detection and recognition for educational institutions or organizations is the main goal of this technology. It makes it possible to track students’ and workers’ daily attendance in real time using facial recognition technology, which makes presence tracking more effective. This automatic system is user-friendly and runs smoothly in a variety of settings. Users can record video, and the system uses facial recognition to automatically record attendance. The system creates thorough attendance reports for simple administration, and this automation greatly improves accuracy.

3. Proposed Method

Fig. 1 presents the proposed pipeline for the student attendance system. This system can be divided into five main stages, from constructed our database for students images, until obtaining the attendance report, as described in the following:

![Students attendance proposed system pipeline](image)

3.1 Students database creation

A collected database of 800 images from 80 students is considered in this study. These images were taken using low-quality webcams that may located in cell phones, laptops, etc. Each student will be photographed more than once (five to ten times, depending on the gesture, lighting, perspective, and whether or not they are wearing glasses). The database is composed from a collection of 60 university students, and a collection from 20 young girls and boys taken from schools. Students’ information such as student name, birth, email, class id, and so on, are stored in SQLite database. The SQLite database is utilized by many of the most popular web browsers, operating systems, mobile devices, and other embedded systems, it is the most commonly used database engine.
and is used to create databases for storing photographs of pupils [18]. We utilize DB Browser for SQLite (DB4S) because it is an excellent, open-source, visual tool for designing, creating, and editing database files that work with SQLite [19].

3.2 Face detection

It is the second stage of the pipeline. It is a computer technique that determines where and how large a human face appears in a digital image. Face detection has become a standout within topics in computer vision literature. [20]. we preformed two methods for face detection as described below:

- **Haar Cascade**: This procedure involves using the HaarCascade Classifier and OpenCV package. The Haar Cascade method must be taught to recognize human faces before it can be utilized for face detection. This is termed feature extraction. The haar cascade training data were utilized. Here, we’re utilizing OpenCV’s detectMultiScale module. This is necessary to draw a rectangle around the faces in a picture (see Fig. 2). Haar features are used to extract features from a picture. Haar features are calculated by subtracting the sum of pixels in the white rectangle from the total of pixels in the black rectangle. Rectangle sizes and placements must be examined, resulting in a large number of characteristics for a picture.

  ![Fig. 2. Haar cascade Classification.](image)

- **Histogram of Oriented Gradients (HOG)**: The HOG method detects objects by converting pixels to gradients. [14] [10]. Gradients are arrows going in a certain direction that depict how the brightness of pixels changes from light to dark. To determine the gradient for each pixel, we examine its neighboring pixels and note how they transition from light to dark. Gradients are extremely useful for identifying objects in photos when their brightness fluctuates. In face identification, for example, a person’s face may seem different in bright and dark illumination, yet the transition from light to dark pixels is quite consistent. Gradients allow us to find things even in low light. Calculating the gradient for each pixel would take forever, so we employ a technique. A sliding window used over the image; each window generates its own gradients, and we select the strongest one. We do this for each window in the image. When we want to locate something in a picture, we seek for the region of the test image that best fits the training image's HOG (Histogram of Oriented Gradients) representation. That's where we're likely to find what we're searching for. [21] [22].

3.3 Preprocessing

After the stage of image detection, the obtained image must enter in this stage, after convert it to grayscale, for solving the noise as following:

- **Median Filter**: is one popular technique for reducing noise in images. In 2D imaging, the median filter works by substituting the median value of each pixel inside a defined window (sometimes called the mask or kernel) with the value of its nearby pixels. The input image can be represented as \(I(x, y)\), where the spatial
coordinates are represented by \((x)\) and \((y)\). In order for the median filter to function, each pixel \((x, y)\) is centered around a \(((2k + 1) \times (2k + 1))\) window, with the size of the window determined by the parameter \((k)\) (e.g., in our case, \(k = 1\)). For a 2D images, Median filter is computed as follows:

\[
\text{Median}(I(x,y)) = \text{Median}(I(x+i,y+j) \mid i, j \in [-k,k])
\]  

(1)

The value of the pixel in the input image at coordinates \((x, y)\) is represented by \(I(x, y)\). The median function, which determines the median value of a set of numbers, is indicated by the symbol \(\text{Median}()\). The indices \((i)\) and \((j)\) indicate the displacement inside the window. Within the window, the range of indices is represented as \([-k, k]\).

- **Gaussian Filter**: it is used frequently with image processing applications [23], its main usage is to smooth or blur images using a Gaussian kernel. This kernel represents a form matrix of the Gaussian function. In image processing, this filter calculated by:

\[
G(x,y) = \frac{1}{2\pi \sigma^2} e^{-\frac{x^2+y^2}{2 \sigma^2}}
\]

(2)

It smooths a picture by averaging pixel values in a zone specified by the Gaussian kernel. It successfully eliminates noise and destroys small features, producing a blurred image. The standard deviation \(\sigma\) of the Gaussian distribution determines the degree of blurring. A higher \(\sigma\) value results in more blurring [24].

Following the filter-based denoising procedure, the images are normalized to remove any extremely different values. Furthermore, to improve the speed of face identification processing, images are downsized to 24x24 dimensions.

### 3.4 Face Identification and Recognition

This method may be broken into three phases. Prepare training data, train the face recognizer, and make predictions. Here, the images contained in the dataset will serve as training data. They will be allocated an integer label indicating which student they belong to. These photos are then utilized for facial recognition. This system employed two alternative approaches, which are as follows:

- **Local Binary Pattern Histogram (LBPH)**: Local Binary Pattern Histogram is the face recognition algorithm employed in this system. The list of all the face’s local binary patterns (LBP) is first acquired. After converting these LBPs to decimal numbers, all of those decimal values are created into histograms. Each image in the training data will ultimately result in one histogram. Later, the histogram of the face that needs to be recognized is computed during the recognition process, compared with the previously computed histograms, and the best matched label for the student that the face belongs to is returned [9].

- **Convolution Neural Network (CNN)**: CNN (Convolutional Neural Network)-based face recognition technology has emerged as the industry standard in the field of face identification thanks to the advancements in deep learning. CNNs were first proposed by Yann LeCun and Yoshua Bengio in 1995. As shown in Fig. 3, a convolutional neural network is a feed-forward network that can extract topological features from the input picture. The raw picture is processed to extract features, which are subsequently classified by a classifier. CNNs are resistant to distortions and basic geometric operations such as rotation, scaling, squeezing, and translation. To provide a certain level of shift, scale, and distortion invariance, Convolutional Neural Networks integrate three architectural concepts: shared weights, local receptive fields, and spatial or temporal sub-sampling. Typically, backpropagation is used to train the network similarly to a regular neural network [16] [8]. To avoid the overfitting issue, the data are split to 60% training, 30% testing, and 10% validation set with a dropout method. The activation functions used in CNN layers are RELU and SOFTMAX in last layer with 50 epochs of 30 iterations for each in the training phase.
3.5 Students Attendance

Throughout the face recognition procedure, the recognized faces will be marked as present in the excel sheet, while the others will be marked as absent, and the list of absentees will be emailed to the appropriate faculties. At the end of each month, faculty will receive an updated monthly attendance sheet.

3.6 System Evaluation

In order to evaluate the performance of the system, the accuracy metric is used to compute the training testing, and validation accuracies. Indeed, this metric needs to compute the True Positive (TP), True negative (TN), False positive (FP), and False Negative (FN) values. The following equation used to compute the Accuracy:

\[
\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}
\]

4. Obtained Results

4.1 System implementation

The proposed system has been implemented using python programming language with GUI on a DELL latitude E5530 using windows 10 64-bits. The system is tested on various operating systems including (Windows 32 and 64 bits, Ubuntu linux 20.04 LTS and 22.04 LTS). The system shows a good performance on previous platforms and succeed to provide the attendance report.

4.2 Attendance Modes

As seen in Fig. 4, the system provides two modes to register the attendance: the single mode and the multiple mode. In the single mode the instructors can communicate with the system using a GUI. In fact, the two primary alternatives that users will have access to are: mark attendance, administration registration, and student registration. In the student registration form, students are expected to fill in all necessary information. However, the multiple mode
registers the student's attendance using one image that contain the student class. It depends mainly on the single mode, but the system tries to apply object detection method to identify students' faces separately.

![Students Attendance System Implementation](image)

**Fig. 4. Students Attendance System Implementation**

### 4.3 Register a student

As shown in Fig. 5, the system provides various fields to enter necessary student's information. Following the Add button click, the webcam launches automatically by pressing the startCamera button, when the student is sitting in front of the camera well, it presses the capture button to allow the system to capture 5-10 images every 5 seconds. These images are stored in previous identified folder.

![Add a new student to our database](image)

**Fig. 5. Add a new student to our database with all information (student ID, Student name, Department, Stage, Study Type, Branch, Gender, E-mail, and also student picture)**

### 4.4 Managing Student Database

Following preprocessing, these photos will be placed in the training pictures folder. The faculty members are expected to fill out the faculty registration form with their email address and the appropriate course codes. This is significant since the corresponding faculties will eventually receive a list of those who are missing. As seen in Fig. 6 and Fig. 7, the database for students could only be updated or deleted by the administration members.
4.5 Performance and Results

To evaluate the performance of the models proposed in this study, we utilize the accuracy metric. Two learning models are examined: the first employs HaarCascade for identification, followed by the LBPH algorithm. The second model utilizes the HoG transform for face identification, followed by the CNN deep algorithm for predicting and identifying student attendance. Table 1 highlights the division ratio of the training, testing, and validation sets to train the two proposed models. In this table, each model was set to equal division ratios with equal number of epochs for comparison matters. The CNN model provides a good training accuracy of 98.5 with an error ratio equal to $\sim0.02$ rather than LBPH model that accumulate an accuracy of 92% and the error ration equal $\sim0.124$.

<table>
<thead>
<tr>
<th>Training Phase</th>
<th>Model</th>
<th>Training set</th>
<th>Testing set</th>
<th>Valid set</th>
<th>Target classes</th>
<th>Epochs</th>
<th>Training Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LBP</td>
<td>60% (480 images)</td>
<td>30% (240 images)</td>
<td>10% (80 images)</td>
<td>80</td>
<td>50</td>
<td>92%</td>
</tr>
<tr>
<td></td>
<td>CNN</td>
<td>60% (480 images)</td>
<td>30% (240 images)</td>
<td>10% (80 images)</td>
<td>80</td>
<td>50</td>
<td>98.5%</td>
</tr>
</tbody>
</table>

For the test and validation phases, the testing and validation sets are tested on both trained models separately. Table 2 illustrates the accuracy for each model for given testing and validation sets. To compute the overall system performance, the average accuracy is computed from summing the accuracies obtained from the testing and validation sets from a model divided by 2. It is clear from Table 2 that CNN provides good accuracy than LBPH.
think that the high accuracy of CNN is because the extraction of face features using the HoG transform, while in the case of LBPH, HaarCascade retrieve good features but the learning model is weak due to its learning mechanism.

### Table 2: Prediction results from learning models

<table>
<thead>
<tr>
<th>Model</th>
<th>Sets</th>
<th>No. of images</th>
<th>TP</th>
<th>TN</th>
<th>FP</th>
<th>FN</th>
<th>Accuracy</th>
<th>Avg. Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>HaarCascade+LBPH</td>
<td>Test (30%)</td>
<td>240</td>
<td>214</td>
<td>16</td>
<td>10</td>
<td>0</td>
<td>96%</td>
<td>95.63%</td>
</tr>
<tr>
<td></td>
<td>Validation (10%)</td>
<td>80</td>
<td>75</td>
<td>1</td>
<td>4</td>
<td>0</td>
<td>95%</td>
<td></td>
</tr>
<tr>
<td>HoG+CNN</td>
<td>Test (30%)</td>
<td>240</td>
<td>235</td>
<td>1</td>
<td>4</td>
<td>0</td>
<td>98.3%</td>
<td>98.44%</td>
</tr>
<tr>
<td></td>
<td>Validation (10%)</td>
<td>80</td>
<td>78</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>98.8%</td>
<td></td>
</tr>
</tbody>
</table>

### 4.6 Comparison with Related Work

In order to validate the proposed model, Table 3 compares the proposed models in this study with most recent studies in the same domain. In this table, each study collected its own dataset from various sources. As these studies did not provide the datasets online for testing our models, we tried to focus on providing student dataset of 800 images for 80 students from different environmental and physical conditions, in addition to different ages and genders. Besides, examine the dataset of two models with different features and leaning models.

### Table 3: Comparison of present study with related works

<table>
<thead>
<tr>
<th>Study No</th>
<th>Study name</th>
<th>Dataset name</th>
<th>No. of dataset images</th>
<th>Learning Model(s)</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>[9] VMU Face recognition 14 celebrity datasets Own dataset</td>
<td>156 2562 220 700</td>
<td>[Raspberry pi+CNN]</td>
<td>98% 98.24% 89.39% 95.71%</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>[10] (Dwi Sunaryono, Joko Siswantoro, and Radityo Anggoro, 2021), Collected face dataset</td>
<td>700 500</td>
<td>[HaarCascade+EmguCV]</td>
<td>linear discriminant analysis+QR code 80%</td>
<td>97.29%</td>
</tr>
<tr>
<td>3</td>
<td>Proposed study own collected student dataset</td>
<td>800</td>
<td>HaarCascade+LBPH HoG+CNN</td>
<td>95.63% 98.44%</td>
<td></td>
</tr>
</tbody>
</table>

### 5 Conclusion

In this study, we provide two techniques for precisely built solution for student attendance problem. This problem is highly connected to the exact facial identification and recognition problems using image processing and machine learning models. We illustrated the efficacy of using HaarCascade with the LBPH algorithm and the HoG transform with the CNN model using two different models. This technology went through thorough research and testing before being smoothly put into actual use. It was designed to properly identify students from varied datasets, including both university and secondary school contexts, using facial recognition technology to ensure careful attendance tracking. After successful identification, the system creates a thorough report, easily enclosed within text and excel files, that lists all attending pupils. Furthermore, the easy interaction with the department’s attendance and absence management system provides smooth data transfer. While both models perform well in terms of training and testing
scores, it is worth noting that the CNN model with HoG beats the LBPH model, with an impressive accuracy value of 98.44% vs 95.63% for the LBPH model.

The study efforts done as part of this project have provided several advantages, with major contributions including:

1. Enabling students to register their attendance more easily through the use of state-of-the-art Euclidean distance technology.
2. Using complex algorithms like the Haar cascade classifier for facial detection and subsequent student recognition and HOG (Histogram of Oriented Gradients) for extracting facial features.
3. The CNN model is highly recommended to apply for facial prediction than LBPH model.

This work not only tackles issues related to attendance management, but it also makes a substantial contribution to improving administrative effectiveness in higher education by bringing together cutting-edge technology and creative approaches.

Acknowledgements

The author is grateful to Mustansiriya University (http://uomustansiriyah.edu.iq) for their important support in completing this work.

Availability: the attendance system is available on github at https://github.com/bbgak2002/Advance_Attendance.

REFERENCES


