The AirNav Semarang Employee Presence System Using Face Recognition Based on Haar Cascade

Fidela Azzahra*, Christy Atika Sari, Eko Hari Rachmawanto

Faculty of Computer Science, Universitas Dian Nuswantoro, Jl. Imam Bonjol No.207, Pendrikan Kidul, Kec. Semarang Tengah, Kota Semarang, Jawa Tengah 50131, (024)3517261

*111202113417@mhs.dinus.ac.id

Abstract. The presence of employees is a key factor in supporting the needs of the workplace. At present, the employee presence system at PT. AirNav Indonesia Semarang Branch still uses fingerprint and RFID-based employee ID cards for authentication. This RFID-based system can increase employee fraud by allowing employees to misuse each other's ID cards. To avoid such fraud, a system needs to be built and it will be using face recognition technology as the primary authentication method, with the Haar Cascade Algorithm. This algorithm has the advantage of being computationally fast, as it only relies on the number of pixels within a rectangle, not every pixel of an image. In addition to fast computation, this algorithm also has the advantage of identifying objects that are relatively far away. With the implementation of the Haar Cascade algorithm, the results indicate the capability of face recognition in detecting the faces of registered employees within the system based on facial angles with an accuracy rate of 60%, expressions with an accuracy rate of 100%, as well as obstructive parameters such as glasses and masks with an accuracy rate of 33.33%. The ability to detect objects from various camera angles, recognize faces with different expressions, and identify objects obstructed by parameters can serve as reasons why this algorithm needs to be implemented.

Keywords: Attendance System, Authentication, Face Recognition, Haar Cascade Classifier

(Received 2024-06-06, Accepted 2024-07-02, Available Online by 2024-07-04)
fraud, such as entrusting ID cards to other employees to record attendance on their behalf.

However, as technology continues to develop, innovations allow the utilization of human biological data characteristics for various purposes. The uniqueness of this technology provides useful information in individual identification processes, such as fingerprints, retinas, voice patterns, and face patterns (face recognition). Face recognition, just like fingerprints [2] and the retina of the eye, is an image processing technique using the concept of computer vision [3] where the captured face image will be compared with the image data already available in the system [4]. This technology falls under the category of biometrics [5] and has been widely noticed and developed by experts as it uses algorithms to distinguish individuals based on the data available in the face database. Facial recognition itself is one of the distinctive features in security systems that identify faces as identity data, considering that the human face stores a lot of information has unique characteristics, and is widely utilized for individual identification purposes. Biometric technology provides a biologically-based authentication method, allowing the system to identify users more accurately. Biometric recognition systems are a subset of authentication systems that use biometric data. These systems automatically recognize a person's identity by comparing biometric traits [6] with the data stored in the database. As an authentication system, a biometric system can determine whether the recognition result is valid or not [7], whether to accept or reject, and whether the user is recognized or not recognized [8] because basically, a face recognition system works by comparing the input image with an image that has been stored in a database and finding the face match that best matches the existing input data [9].

Although many systems utilize face detection features such as security and privacy access systems, many systems utilize face detection [10]. It is important to note that the system also applies certain algorithms such as the Eigenface algorithm. This method is applied to calculate decoding using eigenvectors and represented in a large matrix. These eigenvectors describe the characteristics of variations in the facial image. Each face image contributes differently to each eigenvector, so the eigenvector displays the face image with varying clarity [11]. When referring to research conducted by Rifki Kosasih [12], the Eigenface algorithm has disadvantages such as digital face images taken must be under the same lighting conditions, normalized, and then processed with the same resolution as the $m \times n$ matrix, and components are derived based on pixel values. Because this method requires computationally intensive matrix operations, especially in applications that have a large amount of data, in some cases the computational complexity of the Eigenface algorithm itself can be an obstacle when wanting to detect objects with fast and responsive time.

Based on the weaknesses of the algorithm applied in the explanation above, the employee attendance application will adopt the Haar Cascade algorithm approach. According to research conducted by Munawir et al. [13], the principle of this algorithm is to recognize objects based on simple features rather than pixel values. This allows the computational process to run quickly while detecting objects in a short time, making it highly efficient, as demonstrated by the research conducted by Nikita Sharma and Indrajit Ghosal [14]. Haar Cascade algorithm uses a type of face detector known as a Cascade Classifier. When an image is captured from a video capture, the face detector examines each location in the image and classifies it as a face or non-face.

2. Methods
2.1. Data Collection
The initial stage in system development is the data collection stage [15], including training data for face recognition and employee data compiled by the attendance system that will serve as a reference. In this context, the data collected is employee face data consisting of three employees. Two samples for permanent employees and one sample for interns/On Job Training (OJT) students are stored in the dataset. The dataset samples were directly obtained from each employee during an open internship conducted by the author. The specific date and time of data retrieval are documented in the Retrieval Time column of Table 1.
2.2. Adaboost Algorithm

The training process is performed using the Adaboost algorithm to select the most discriminative subset of Haar features for each classifier [20]. Each classifier is tested on the training dataset to select the features that are most effective in distinguishing between images containing objects of interest and those that are not. Error weights ($\epsilon$), class weights ($w$), and classifier weights ($\alpha$) are calculated in each iteration step.

$$\epsilon_t = \frac{\sum_{i=1}^{N} w_i \cdot e_i}{\sum_{i=1}^{N} w_i}$$  

(1)

Where in formula (3), $e_i$ is the error indicator function for the $i$ training sample and $w_i$, $t$ is the weight of the $i$ sample at the $t$ iteration step. Classifier Weight ($\alpha$) as in (2) and Class Weight ($w$) as in (3).

$$\alpha_t = \frac{1}{2} \ln \left( \frac{1-\epsilon_t}{\epsilon_t} \right)$$  

(2)

$$w_{t+1} = w_t \cdot \exp (-\alpha_t \cdot e_i)$$  

(3)

2.3. Haar-like Feature Classifier

Integral image is a simple calculated [18] representation of the intensity of pixels in an image. Integral images are very useful in determining the presence or absence of hundreds of Haar features in an image and at different scales efficiently. Integral image $I(x,y)$ of an image $G(x,y)$ can be defined by the formula (1). Where $I(x,y)$ is a variable that holds the total pixel value from the origin of the image to the selected point. This allows efficient calculation of the total pixel intensity within any rectangular area in the image. Haar features are a simple method to identify changes in light intensity in an image. For example, a basic Haar feature consists of two regions [19] with different total intensities. The evaluation of Haar features can be done with the formula found in the formula (2). Where $p(x,y)$ is a function that defines the intensity pattern for a particular Haar feature, and $w_i$ is the corresponding weight. N number of different Haar features are selected and scored at each classification step as in (4) and (5).

$$I(x,y) = \sum_{x',y'} x' \leq x, y' \leq y G(x',y')$$  

(4)

$$\int (x,y) = \sum_{i=1}^{N} w_i \cdot p_i(x,y)$$  

(5)

The first process performed in the Haar Cascade Classifier method to detect the presence of facial features in an image is to convert the image from an RGB image to a grayscale image as shown in Figure 3. The next process is to select the Haar features in the image which is done by boxing each region in the image from the top left to the bottom right. The goal is to find facial features such as eyes, nose, and mouth as shown in Figure 4. The next step is done by selecting the Haar features present in the grayscale image. Haar features are pixel-intensity patterns used to identify objects in an image. This process is done by dividing the image into small squares and looking for important patterns or features in them, such as edges and lines. These features are identified based on the difference in pixel values in certain areas within the grid. For example, eye features can be identified based on the difference in pixel values between the eye area and the surrounding area, as well as nose and mouth feature and once the important Haar features are identified, the next step in the Haar Cascade Classifier is to compare the pattern of the identified features with the pattern contained in the previously learned faces. This process allows the
classifier to distinguish the face from other objects in the image. The next step is to perform the optimal calculation using an integral image. Before determining the integral image, it is important to know the input value of the image. The input value of the image can change based on the quality, width, and height of the image. To get the input value of the image can be done using tools such as Matlab and Python.

After undergoing a series of calculations to obtain the input values, the values in the image can be ascertained. The values in Figure 5 are considered as the values of the input image, with the arrows pointing left indicating the vertical axis, and those pointing up indicating the horizontal axis.

The next step is done by selecting the Haar features present in the grayscale image. Haar features are pixel-intensity patterns used to identify objects in an image. This process is done by dividing the image into small squares and looking for important patterns or features in them, such as edges and lines. These features are identified based on the difference in pixel values in certain areas within the grid. For example, eye features can be identified based on the difference in pixel values between the eye area and the surrounding area, as well as nose and mouth feature and once the important Haar features are identified, the next step in the Haar Cascade Classifier is to compare the pattern of the identified features with the pattern contained in the previously learned faces. This process allows the classifier to distinguish the face from other objects in the image. The next step is to perform the optimal calculation using an integral image. Before determining the integral image, it is important to know the input value of the image. The input value of the image can change based on the quality, width, and height of the image. To get the input value of the image can be done using tools such as Matlab and Python. After undergoing a series of calculations to obtain the input values, the values in the image can be ascertained.

The values in Figure 1 are considered as the values of the input image, with the arrows pointing left indicating the vertical axis, and those pointing up indicating the horizontal axis. The calculation of the integral image value (s) based on the input value referring to Figure 1 is performed using formula (7) to obtain the calculation result.

\[
s(x, y) = i(x, y) + s(x-1, y) + s(x, y-1) - s(x-1, y-1)
\]  

(7)

![Figure 1. Pixel Value of a Feature](image1.png) ![Figure 2. Integral Image Calculation Result](image2.png)

Table 2 is the calculation of the integral image that is done with the overall calculation based on the input image as shown in Figure 5. In doing the calculation to get the integral image value, we can use the formula as shown in formula (7) or by using a quick calculation technique. As an examples for the quick calculation technique, in the integral image calculation table for pixel value (2,2), the integral value is the pixel value plus the integral value of the pixel above it in the same column, plus the value of the previous pixel in the same row and column, which is \(1 + 3 + 0 + 3 = 7\). The integral image obtained from the calculations in Table 2 is shown in Figure 6. Based on the calculation results of the integral image in Figure 6, it will be grouped into 4 parts for regions A, B, C, and D. A = \{3, 17\}, \{7, 36\}, B = \{7, 29\}, \{17, 58\}, C = \{11, 19\}, \{29, 52\}, D = \{19, 34\}, \{51, 110\}.

For each region A,B,C, and D, the number of pixels will be calculated according to the same equation formula for each region for regions A,B,C, and D using formula (8).

\[
L_1 + L_4 - (L_2 + L_3)
\]  

(8)

Region A \(\Rightarrow\) \(L_1 = 3, L_2 = 17, L_3 = 7, L_4 = 36\) and \(A = L_1 + L_4 - (L_2 + L_3) = 3 + 36 - (17 + 7) = 15\)
Region B \(\Rightarrow\) \(L_1 = 7, L_2 = 29, L_3 = 17, L_4 = 58\) and \(B = L_1 + L_4 - (L_2 + L_3) = 7 + 58 - (29 + 17) = 19\)
Region C \(\Rightarrow\) \(L_1 = 11, L_2 = 19, L_3 = 29, L_4 = 52\) and \(C = L_1 + L_4 - (L_2 + L_3) = 11 + 52 - (19 + 29) = 15\)

02403011-04
Region D \rightarrow L_1 = 19, L_2 = 34, L_3 = 51, L_4 = 110 and D = L_1 + L_4 - (L_2 + L_3) = 19 + 110 - (34 + 51) = 44

After obtaining the integral image value of the input image and counting the number of pixels in a particular region, the next step is to compare the pixel values between the light and dark regions. If the difference in pixel values between light and dark regions exceeds the threshold value previously determined, then the region is considered to have significant features. Thus, the feature detection process in the image can be done by comparing the pixel value in each region against a predetermined threshold, and the regions that have a pixel value difference above the threshold can be identified as relevant features.

Table 2. Integral Image Calculation

<table>
<thead>
<tr>
<th>Pixel Value (i)</th>
<th>s(x-1,y)</th>
<th>s(x,y-1)</th>
<th>s(x-1,y-1)</th>
<th>s(x,y)</th>
<th>Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>Integral for (1,1)</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>6</td>
<td>Integral for (1,2)</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>6</td>
<td>0</td>
<td>11</td>
<td>Integral for (1,3)</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>11</td>
<td>0</td>
<td>15</td>
<td>Integral for (1,4)</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>15</td>
<td>0</td>
<td>15</td>
<td>Integral for (1,5)</td>
</tr>
<tr>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>Integral for (2,1)</td>
</tr>
<tr>
<td>1</td>
<td>6</td>
<td>3</td>
<td>3</td>
<td>7</td>
<td>Integral for (2,2)</td>
</tr>
<tr>
<td>5</td>
<td>11</td>
<td>7</td>
<td>6</td>
<td>17</td>
<td>Integral for (2,3)</td>
</tr>
<tr>
<td>8</td>
<td>15</td>
<td>17</td>
<td>11</td>
<td>29</td>
<td>Integral for (2,4)</td>
</tr>
<tr>
<td>8</td>
<td>15</td>
<td>29</td>
<td>15</td>
<td>37</td>
<td>Integral for (2,5)</td>
</tr>
<tr>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>Integral for (3,1)</td>
</tr>
<tr>
<td>0</td>
<td>7</td>
<td>3</td>
<td>3</td>
<td>7</td>
<td>Integral for (3,2)</td>
</tr>
<tr>
<td>0</td>
<td>17</td>
<td>7</td>
<td>7</td>
<td>17</td>
<td>Integral for (3,3)</td>
</tr>
<tr>
<td>7</td>
<td>29</td>
<td>17</td>
<td>17</td>
<td>36</td>
<td>Integral for (3,4)</td>
</tr>
<tr>
<td>14</td>
<td>37</td>
<td>36</td>
<td>29</td>
<td>58</td>
<td>Integral for (3,5)</td>
</tr>
<tr>
<td>8</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>11</td>
<td>Integral for (4,1)</td>
</tr>
<tr>
<td>4</td>
<td>7</td>
<td>11</td>
<td>3</td>
<td>19</td>
<td>Integral for (4,2)</td>
</tr>
<tr>
<td>0</td>
<td>17</td>
<td>19</td>
<td>7</td>
<td>29</td>
<td>Integral for (4,3)</td>
</tr>
<tr>
<td>4</td>
<td>36</td>
<td>29</td>
<td>17</td>
<td>52</td>
<td>Integral for (4,4)</td>
</tr>
<tr>
<td>11</td>
<td>58</td>
<td>52</td>
<td>36</td>
<td>85</td>
<td>Integral for (4,5)</td>
</tr>
<tr>
<td>6</td>
<td>11</td>
<td>0</td>
<td>0</td>
<td>17</td>
<td>Integral for (5,1)</td>
</tr>
<tr>
<td>9</td>
<td>19</td>
<td>17</td>
<td>11</td>
<td>34</td>
<td>Integral for (5,2)</td>
</tr>
<tr>
<td>7</td>
<td>29</td>
<td>34</td>
<td>19</td>
<td>51</td>
<td>Integral for (5,3)</td>
</tr>
<tr>
<td>0</td>
<td>52</td>
<td>51</td>
<td>29</td>
<td>74</td>
<td>Integral for (5,4)</td>
</tr>
<tr>
<td>3</td>
<td>85</td>
<td>74</td>
<td>52</td>
<td>110</td>
<td>Integral for (5,5)</td>
</tr>
</tbody>
</table>

3. Results and Discussion

The results of this study will describe the appearance of the employee attendance application. Figure 3 is a view of the home screen menu. This home screen is the display when the application is run for the first time. On the navbar menu, there are “Home” options and the “Employee Attendance”. The option that has bolded text is the page that is currently being visited, and if you want to do attendance, just press the Employee Attendance option on the navbar. Figure 4 displays the “Employee Attendance” menu, which is designed to facilitate employees in recording attendance. In the image, Yordan Dewo Ajie, one of the employees, is seen testing the application that has been created. The employee's face is surrounded by a green frame, indicating that their attendance was successfully detected by the system. Complete information about the employee's name and time of attendance is available in the table located on the right side of the image. After recording attendance, employee attendance data will be saved automatically into the system and can be exported in PDF or CSV format by pressing the “Download PDF” or “Download CSV” button. A file showing the successful export of data in PDF form can be seen as shown in Figure 5.
Three types of tests are carried out randomly. The first test is a test based on the angle at which the face image is taken, such as upright, looking down, lifting the head, and rotating the head left or right. The second test is a test based on facial expressions. Such as smiling facial expressions, frowning faces that show teeth, and gaping faces. The third test is parameter-based. For testing based on this parameter if there is a barrier parameter such as glasses or a mask covering the face. The first test conducted in Table 3 is a test conducted based on the angle at which the face image is taken. The results obtained show that 6 faces are recognized, 1 face is not recognized, and 3 faces are not detected. The second test conducted in Table 4 is a test conducted based on facial expressions. The results obtained show that 4 faces are recognized. The third test conducted in Table 5 is a test conducted based on barrier parameters for faces such as glasses and masks. The results obtained show that 1 face is recognized and 2 faces are not detected. There have been 10 trials for the first test, 4 trials for the second test, and 3 trials for the third test. Based on the experiments that have been carried out, the accuracy value in the test experiment can be obtained as described in the calculation (9), while the error value can be obtained using the formula in the
In the first test of accuracy based on the angle of taking facial images, 6 trials were successfully identified from 10 types of trials conducted. Thus, the employee attendance application achieved a success rate of 60% with an error rate of 40%.

\[
\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \times 100\% \quad (9)
\]

\[
\text{Error Value} = 100\% - \text{Total Accuracy} \quad (10)
\]

- **Accuracy 1 (%)** = \( \frac{6 + 0}{6 + 4 + 0 + 0} \times 100\% = 60\% 

- **Error Value 1 (%)** = \( 100\% - 60\% = 40\% 

- **Accuracy 2 (%)** = \( \frac{4 + 0}{4 + 0 + 0 + 0} \times 100\% = 100\% 

- **Error Value 2 (%)** = \( 100\% - 0\% = 0\% 

- **Accuracy 3 (%)** = \( \frac{1 + 2 + 0 + 0}{1 + 2 + 0 + 0} \times 100\% = 33.33\% 

- **Error Value 3 (%)** = \( 100\% - 33.33\% = 66.67\% 

**Table 4. Testing Based on Facial Expressions**

<table>
<thead>
<tr>
<th>Name</th>
<th>Bhakti Yudha Setiyawan</th>
</tr>
</thead>
<tbody>
<tr>
<td>Testing</td>
<td></td>
</tr>
<tr>
<td>Information</td>
<td>Smiling Face, Frowning Face, Visible Teeth, Gawking Face</td>
</tr>
<tr>
<td>Result</td>
<td>Recognized, Recognized, Recognized, Recognized</td>
</tr>
</tbody>
</table>

**Table 5. Testing Based on Barrier Parameters**

<table>
<thead>
<tr>
<th>Name</th>
<th>Bhakti Yudha Setiyawan</th>
</tr>
</thead>
<tbody>
<tr>
<td>Testing</td>
<td></td>
</tr>
<tr>
<td>Information</td>
<td>With glasses, With Mask, With Goggles and Mask</td>
</tr>
<tr>
<td>Result</td>
<td>Recognized, Not Detected, Not Detected</td>
</tr>
</tbody>
</table>

In the second test of accuracy based on facial expressions, 4 trials were successfully identified from 4 types of trials conducted. Thus, the employee attendance application achieved a success rate of 100% with an error rate of 0%. In the third test of accuracy based on barrier parameters, there was 1 trial that was successfully identified from 3 types of trials conducted. Thus, the employee attendance application achieved a success rate of 33.33% with an error rate of 66.67%.

**4. Conclusion**

The level of accuracy produced in the Haar Cascade Classifier algorithm on the design of the AirNav Semarang employee attendance system in the first test resulted in an accuracy rate of 60% with an error rate of 40%, in the second test of 100% with an error rate of 0%, and in the third test of 33.33% with an error rate of 66.67%. The varying levels of accuracy in the results of several experiments stem from the unequal distribution of tests across different categories. For instance, within the three tests conducted on barrier parameters, only one can be deemed successful. Conversely, in the second test focusing on facial expressions, all trials involving the same face but with different expressions were successfully completed without any exceptions. This underscores the premise that the accuracy and error rates observed in these tests are contingent upon the quantity of tests conducted. According to the accuracy rate and error rate in the paragraph above, the system can be said to be imperfect if it has an accuracy rate below 100% and an error rate above 0%. As a solution to overcoming the low level of accuracy and the high level of error, several series of stages can be carried out. The first stage can be by testing again
using the same sample data but when the test is carried out it must pay attention to several important factors, such as adequate lighting levels, appropriate face angles to improve the readability process, distances that do not exceed the ideal distance, and parameters that can hinder the face reading process. The second stage can be by adding new sample data with a testing process that allows faces in webcam captures to be recognized. It is anticipated that this solution will contribute to similar research in the next phases. For the future research in the field facial attendance system could focus on enhancing the robustness and accuracy of the Haar Cascade Classifier algorithm under varying environmental conditions. One avenue of exploration could involve refining the algorithm's ability to handle obstructive parameters, such as masks, more effectively. This could entail developing advanced preprocessing techniques or incorporating additional feature extraction methods specifically tailored to address occlusions caused by facial accessories. In addition, systems that are directly connected to the database are considered to provide noticeably better performance than those that are limited to local computer use.

References


