Human Detection and Identification for Home Monitoring System

D Arisandi¹, M Elveny¹, R Rahayu¹

¹Department of Information Technology, Faculty of Computer Science and Information Technology, Universitas Sumatera Utara, Medan, Indonesia
dedyarisdani@usu.ac.id, marischaelveny@usu.ac.id, rusnairahayu75@gmail.com

Abstract. Currently, a monitoring system is needed to improve security and productivity in various sectors, for example in public services such as education centers, offices, banking, stations, roads, and the industrial sector. With a monitoring system, public service activities can be well monitored. Apart from public services, a home surveillance system can be implemented, making it easier for household members to monitor home security. The monitoring system camera will record the object, the human object detection process is carried out with the Histogram of Oriented Gradient, after the object is detected, the facial recognition process of the object is carried out whether it is a member of the house or not by using Learning Vector Quantization. If the object is not a member of the house, the system will send a notification in the form of a photo of the object's face to the user's email. For object detection with the Histogram of Oriented Gradient, an average accuracy of 90% is obtained and Learning Vector Quantization for facial recognition is 85%.

1. Introduction

With the development of scientific and technological advances in various sectors, the need for a monitoring system is increasing where a monitoring system is implemented to increase security and productivity, for example public services such as shopping and education centers, offices, banking, stations, airports, and roads using a monitoring system for security, and the productivity aspect is applied to the industrial sector where management can monitor the production activities of workers, control process instrumentation, machine installation, and others. Apart from public services or public places, one place that uses a monitoring system for security aspects is the house. Home members will find it easier to monitor home security with a monitoring system. Another research utilize automatic monitoring system to control electrical appliance in hospital room especially patient room by using voice recognition [1].

This study uses an external camera and there are weaknesses that the installed camera always records even though there is no activity so that memory usage is too large and becomes useless. One of alternative to solve this problem is a system or software that will record when human movement is detected so that it can increase the efficiency of the camera. And this system can provide direct notification to users if something happens around the house. Whereas, researcher using voice to identify some command from user especially in patient room. By using this system, patients will not depend to ask nurse to adjust their room, so they are able to adjust appliance themselves such as mattress adjustment according to position (sleeping, rest, sitting position). The voice is received by
using microphone, motor servo utilize mini bed is simulated patients’ electrical bed, the lighting room controlled used small lamp with relay and need buzzer (help button) to accommodate to call nurse [1].

2. Identification of Problems
The camera of monitoring system always records even though there is no activity, as a result there is useless memory usage and there is no direct notification to users or household members. So an approach is needed to be able to increase the efficiency of camera, so that the camera only detects and records when there is movement of human objects and can provide direct notification to members of the house.

3. Related Research
Related research to the detection and recognition of human objects was carried out by Navneet Dalal in 2005 using the Histogram of Oriented Gradient method for detection of human objects with data from 1800 pedestrian databases [2]. The research by Thamrin Imanuel Panggabean in 2018 using a motion detector, where the camera will start recording only when there is movement of a human object and is able to recognize the detected human object, using the Support Vector Machine and Deep Convolutional Neural Network [3]. Suryadi & Sikumbang in 2015 in their research used the OpenCV-based Histogram of Oriented Gradients (HOG) to detect the human body by video input and determine the detection results whether human or not and marked with a square bounding box [4].

The research by Arisandi D et al in 2017 to identify the identity of students in a university through face so that students do not have to browse the student directory on the college website. Face recognition uses Fisherface Methods [5]. Research by Muhammad Iqshand Johanda in 2014 on human face pattern recognition. Training is carried out on each input pattern, the results of the training will be compared with the target and using Robert's edge detection method to obtain the outline of the face image [6]. Rudy Chandra in 2017 using Learning Vector Quantization to identify faces and make these faces the main key for the attendance system [7]. The difference between this study and previous research is the combination of algorithms used, notification provided, and testing have been carried out in real time, whereas in previous research it has not been done in real time.

4. Research Methods
In this research, there are several steps starting from the video input process or url stream, preprocessing, the detection process, face recognition or identification, and notification. For details, it can be seen in Figure 1.
4.1 Input
In this process, the camera that has been installed will start recording or stream video through a certain url, then the video changes into image frames.

4.2 Human detection
4.2.1 Resizing. The first process is resizing image which generated from the previously stage. The resizing process uses a library of the HOG algorithm to change image in a large size to a smaller one for the needs of the detection process.

4.2.2 Feature extraction. After the resizing process is carried out, we conduct the feature extraction process using the Histogram of Oriented Gradient method to determine whether the frame contains humans or not. Histogram of Oriented Gradient comes from an assumption which states that an object can be represented properly based on shape. To obtain differentiating information, the image will be divided into cells and each cell will be calculated as a Histogram of Oriented Gradient. The Histogram of Oriented Gradient process starts with normalizing gamma and color (converting a color image to a gray image), the gradient compute process or calculating the gradient value of each pixel, calculating the histogram of the gradient orientation of each cell, normalizing the block, and the SVM Classifier.

4.2.3 Bounding box. If there is a human object, it will be given a rectangle bounding box with the information "Number of people detected". Bounding box will provide a box for the detected object and count the number of detected objects based on the detection process carried out by the hog algorithm through body features.

4.3 Face recognition
4.3.1 Preprocessing. After a human object is detected, it is determined whether any faces are detected. If there is a face, then the preprocessing stage is carried out first, namely resizing, grayscaling, and thresholding, with the results of resizing only on the face area.
4.3.2 Face identification. Then, the identification process or face recognition using Learning Vector Quantization, whether the face is face of a member of the house or not. The process is conducted by comparing data testing and data training. Initial process is input of input pattern and output pattern, then consider for the nearest Euclidian distance, new weight calculations, and done as much as Epoch. The final result will be calculated use the closest distance, if the smallest score, then it belongs to the closest target category. Data is used in this research is 50, 35 for training and 15 for testing.

4.4 Output
The output of the system is in the form of images and videos.
- Images
  The result image is an image of a detected human object, and if there is a face, a marker will be assigned to the image with the name of the member of the house and if it is not recognized then an "Unknown" marker will be given.
- Video
  The result video are videos that only contain the activities or movements of human objects.

4.5 Notification
Notification is sent in the form of an image of an unrecognized human object. Notifications sent to the user's email or home members.

5. Result and Discussion
5.1 Human detection
In this process, the detection of human objects is carried out with differences in light intensity and distance of objects. The test was carried out at a light intensity of 5 lux, 150 lux, 750 lux, and 1000 lux. And testing is carried out on objects with a distance of 1 meter, 2 meters and 5 meters from the camera. Tests were also carried out on non-human objects, namely animals and objects that resemble humans and the system succeeded in not detecting that it was a human being in front of the camera. Examples of testing for non-human objects can be seen in Figure 2 and Figure 3.

![Figure 2. Testing with animal](image1)

![Figure 3. Testing with object](image2)

For tests with different light intensities and distances, it can be seen in Figure 4, Figure 5, Figure 6, and Figure 7.
Figure 4. Testing with intensity of 5 lux and distance of 1 meter

Figure 5. Testing with intensity of 150 lux and distance of 2 meter

Figure 6. Testing with intensity of 750 lux and distance of 5 meter

Figure 7. Testing with intensity of 1000 lux and distance of 5 meter
Then we perform an evaluation of the system in detection of human. The equation used is the calculation of precision, recall, and F-Score. Table 1 show the result of testing in this research.

**Table 1. Testing of Human Detection**

| Identification Result | Actual Condition | Human Object | Not a Human Object |
|------------------------|------------------|--------------|--------------------|
| Human Object           | True Positive (TP) | False Positive (FP) |
| Not a Human Object     | False Negative (FN) | True Negative (TN) |

Formulas:

\[
Recall = \frac{TP}{TP + FN} \tag{1}
\]

\[
Precision = \frac{TP}{TP + FP} \tag{2}
\]

\[
F-Score = \frac{2 \times (Precision \times Recall)}{Precision + Recall} \tag{3}
\]

Explanation:
- True Positive (TP) is a condition in which the actual condition of a human object detected has been identified by the system as a human object.
- False Positive (FP) is a condition where the actual condition of a non-human object is detected and failed to be identified by the system as a non-human object.
- True Negative (TN) is a condition where the actual condition of a non-human object is detected and identified by the system as a non-human object.
- False Negative (FN) is a condition where the actual condition of detected human objects fails to be identified by the system as human objects.

The results of the accuracy of detecting human objects can be seen in Table 2 and Table 3.

**Table 2. Calculation results of human detection F-Score with intensity of 750 lux**

| Distance | 1 Meter | 2 Meter | 5 Meter |
|----------|---------|---------|---------|
| TP       | 6       | 6       | 5       |
| FP       | -       | -       | -       |
| TN       | 4       | 4       | 4       |
| FN       | -       | -       | 1       |
| Recall   | 100     | 100     | 83.3    |
| Precision| 100     | 100     | 100     |
| F-score  | 100     | 100     | 90      |
Table 3. Calculation results of human detection F-Score with intensity of 1000 lux

| Distance | 1 Meter | 2 Meter | 5 Meter |
|----------|---------|---------|---------|
| TP       | 6       | 6       | 5       |
| FP       | -       | -       | -       |
| TN       | 4       | 4       | 4       |
| FN       | -       | -       | 1       |
| Recall   | 100     | 100     | 83.3    |
| Precision| 100     | 100     | 100     |
| F-score  | 100     | 100     | 90      |

5.2 Face recognition

After the human object detection process is carried out, we conducted face recognition by comparing data from members of the house. The face recognition results can be seen in Table 4.

Table 4. Face recognition testing results

| Testing  | Data   | Actual Output | System Output |
|----------|--------|---------------|---------------|
| 1000 lux | Known | Known         |               |
| 1000 lux | Unknown | Unknown       |               |
| 750 lux  | Known | Known         |               |
| 150 lux  | Known | Known         |               |
| 5 lux    | Known | Known         |               |
| 5 lux    | Known | Unknown       |               |
| …        | …      | …             | …             |
| 5 lux    | Unknown | Unknown       |               |
To calculate face recognition accuracy can be seen in the equation (1).

\[
\text{Percentage of accuracy} = \frac{\text{Number of correct tests}}{\text{Number of total tests}} \times 100\% 
\]

\[
= \frac{17}{20} \times 100\% = 85\% 
\]

6. Conclusion
After testing, detection of human objects and face recognition, conclusions can be drawn, are:
1. The implementation of the Histogram of Oriented Gradient method is able to detect human objects in the monitoring system properly.
2. In the process of human detection as part of improving the performance of the camera, the monitoring system is influenced by light intensity and the distance of the object from the camera.
3. In light conditions of 750-1000 lux and with a distance of 1-2 meters, the system can detect human objects with an accuracy rate of 100% and a distance of 5 meters at 90% and face recognition results obtained an accuracy of 85%.

References
[1] Agustin E I, Yunardi R T and Firdaus A A 2019 Voice recognition system for controlling electrical appliances in smart hospital room *TELKOMNIKA* 17 2 pp 965-972
[2] Dalal N and Triggs B 2005 Histograms of Oriented Gradients for Human Detection. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition* 1 p 886-893.
[3] Panggabean, T I 2018 Pendeteksian dan Pengenalan Wajah Manusia untuk Peningkatan Kinerja Kamera Pengaman. Skripsi. Universitas Sumatera Utara
[4] Suryadi K and Sikumbang S 2015 Human Detection menggunakan Metode Histogram of Oriented Gradients (HOG) berbasis OpenCV. *Jurnal Pendidikan Teknik Elektro* 4 2 p.639-645
[5] Arisandi D, Syahputra M F, Putri I L, Purnamawati S, Rahmat R F and Sari P P 2017 A Real Time Mobile-Based Face Recognition with Fisherface Methods. *J. of Physics: Conference Series 978*
[6] Johandra M I 2014 Implementasi Learning Vector Quantization dan Deteksi Tepi Robert untuk Pengenalan Citra Wajah. Skripsi. Universitas Sumatera Utara.
[7] Chandra R 2017 Sistem Absensi menggunakan Wajah pada Jaringan Syaraf Tiruan dengan Algoritma *Learning Vector Quantization* (LVQ). Skripsi. Universitas Sumatera Utara.