مجلة كلية التربية للعلوم الصرفة - جامعة ذي قار
المجلد ٨، العدد ٢، حزيران ٢٠١٨
Website: jceps.utq.edu.iq Email: jceps@eps.utq.edu.iq

Intelligent Monitoring to Detect and Recognized the Unauthorized Persons

Hassanein Samir Abdullah(1) , Sana Ali Jabber(2)

qhasanen@yahoo.com , sana.jebber@gmail.com

(1) Computer Science Department, University of Technology, Baghdad, Iraq

(2) University of Muthanna / Administration and Economics College

Abstract

Due of the urgent needs for Security and surveillance systems to account for the security conditions experienced by the country we contribution to the development and support of these traditional security systems with intelligent surveillance systems based on artificial intelligence algorithms to track and monitor any security breach. The proposed system will issue a warning whistle to alert the security men as well as build a video that summarizes all the unauthorized (strangers) movements which facilitates the process of retrieving important data quickly as well as reducing the storage space used to store videos. The proposed system is based on distinguishing faces and tracking them by using Viola and Jones approach to detected face, and then classifying the target face as an unauthorized by using Local Binary Patterns (LBP).

The experiments have shown that the proposed system has the efficient and robust detection; recognition and tracking face in surveillance multi-cameras are representing by 95% of the detection faces and more than 97% of the tracking of unauthorized (strangers) which represent a high proportion compared to similar works with high speed in the retrieval and analysis of video content.

1. Introduction

Recently, there is a growing need for video surveillance systems. Governments as well as companies have started to equip their property with Closed-Circuit Television (CCTV) cameras and other sensor systems for the sake of increasing the safety for their citizens or customers [1]. Video surveillance could be defined as an integrated system that has strong prevention abilities and is of common use in military, customs, police, fire fighting, airports, railways, urban transport and a great number of other public places. It is a significant part of security system due to the fact that it is of visualized, precise, fast and rich content of information. Video surveillance has become the basic tool because it is of rich, intuitive and precise information [2] .

A great deal of researchers has shown an enormous interest in the development of an intelligent surveillance system. This kind of systems deals with the real-time monitoring of persistent and transient objects within a particular environment. The basic objectives of those systems are providing an automatic scene interpretation and understanding and predicting the actions and interactions of the monitored items according to the information obtained by the sensors [3,4,5]. The basic phases of processing in an intelligent visual surveillance system are: detecting and recognizing moving object, in addition to their tracking, behavioral analysis and retrieval. Those phases are involved with the applications of machine vision, pattern analysis, artificial intelligence and data management. In these surveillance systems, a big number of sensors are used for collaboratively processing data [6] but the problems we faced when using regular CCTV cameras are that they work all the time, which makes screen monitoring a difficult task. Mayneeds large number of security guards to monitor the screens, because the normal man cannot monitor more than eight cameras at a time [7]. However, when there is a warning whistle that triggers the time of danger, it is easy to monitor and follow it. It attracts the attention of the observer at specific times. This paper, contributed to the development of conventional surveillance cameras using artificial intelligence algorithms by detecting and distinguishing strange faces.The system consists of the following stages:

- System training stage for authorized persons
- Monitoring and tracking stage
- Video reconfiguration phase for tracking unauthorized personnel

2. Literature Survey

The proliferation of surveillance devices to expand the search in this area to keep pace with the security developments witnessed by the whole world as a result of terrorist attacks has turned many researchers to developed and support security devices with intelligent techniques of these, and from these researches:

In [8] the authors have proposed an approach to synthesis face images-based on 3D face modeling and blurring. In the proposed algorithm, firstly a 2D frontal face with high-resolution was used to build a 3D face model, and then several virtual faces with different poses were synthesized from the 3D model. Experimental results indicated that the proposed method was effective and could be considered to be applied in intelligent video monitoring system. In A new gun-dome camera cooperative system that is capable of solving the automatically realize acquisition, refining and fast obtaining of the needed information in a surveillance video partly. In [2]the authors have presented framework, the approach of Deformable Part Model (DPM) achieves real-time detection of pedestrians. And concerning the facial detection, are utilized the Haar-Like feature and LQV classifier to execute the frontal face image capture. The results of the study demonstrated the effectiveness and sufficiency of the dual-camera system in close-up facial image capturing. In [9] researchers have suggested adaptive Appearance Model Tracking (AAMT) for gradually learning a track-face-model for every one of the individuals that appear in the scene. The Sequential Karhunen–Loeve approach is utilized for online learning of those track-face-models within a particle filter-based facial tracker. The most significant benefit of AAMT compared to the traditional still-to-video FR systems is that it’s of greater variation of face representation which could be acquired during procedures, which might result in a better discrimination for spatiotemporal identification.

3. The Proposed System Overview

The proposed system deals with multi-surveillance cameras distributed in the building.Figure (1) shows the overall framework of this work.Firstly, train the system on the group of people (security guards) who are allowed to enter the building and store them in a dataset using the AdaBoost learning algorithm, and then create a special database by using Oracle SQL Plus for each security guard to show all his data on the monitor screen if it appears in the building. Secondly, detect anyone who shows up on surveillance cameras using Viola-Jones face detection algorithm and then recognized him by LBPH by comparison face frame with the dataset that the system was trained it to detect infiltrator's (strangers) into the building and issuing an alarm whistle to alert the security guards and store each of the frames which appears the infiltrator's in a folder .Thirdly, create surveillance video to the infiltrator's who represents all their movements in the building.

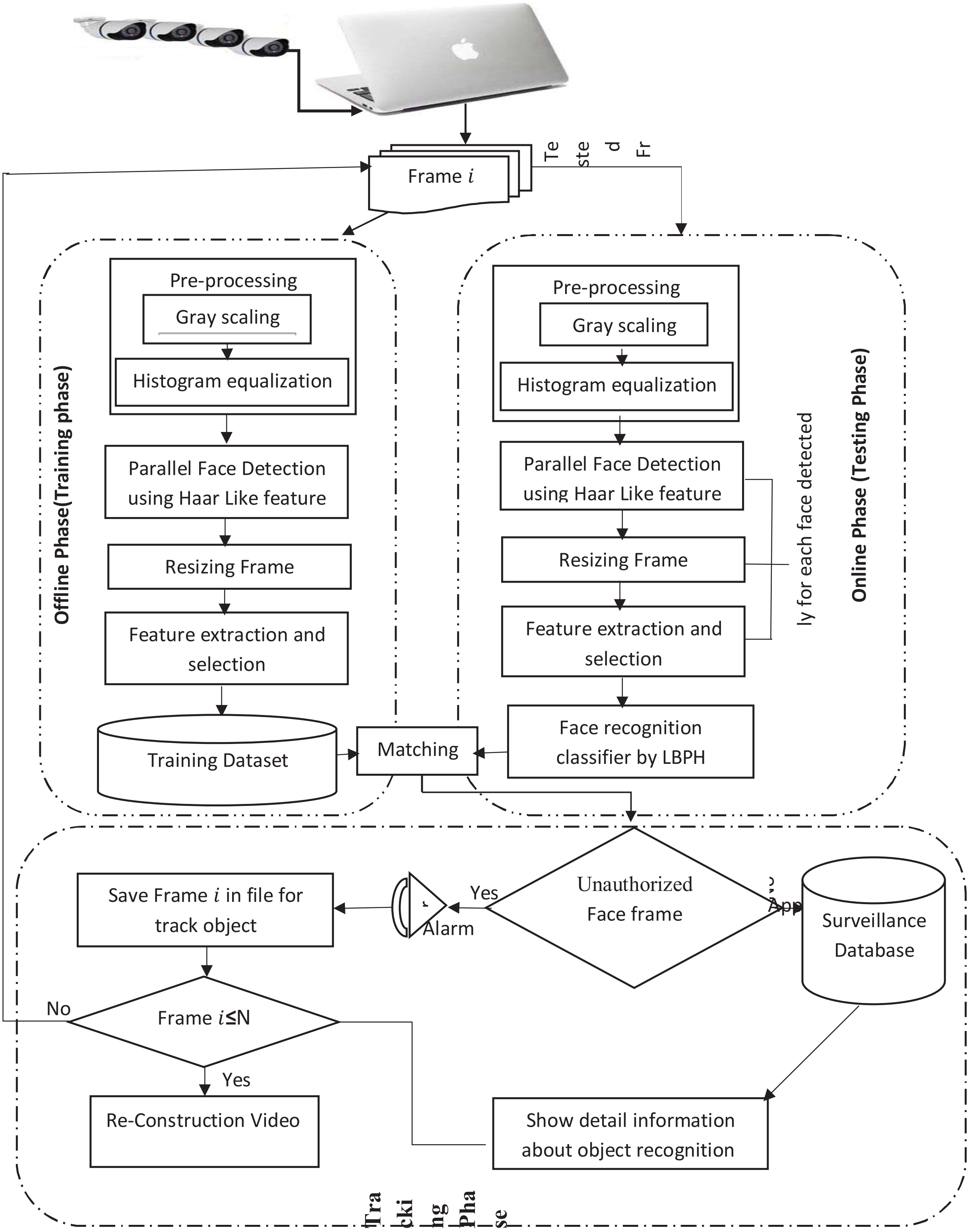

Figure 1.The Proposed System Architecture

4. System Training Stage for Authorizedpersons

This stage consists of the fallowing sections:

i. Sample Selection for Training a System

For training the proposed system on a set of persons authorized, The system was trained by taking a set of samples for each person, by recording video from surveillance cameraFor each of them . This method was used instead of capturing a set of static images to obtain high accuracy in recognation phase. Because distinguishing the faces in the video is affected by the distance between the camera and the person as well as the lighting plays a key role in the accuracy of detection and discrimination, so it was captured the frames from different distances from the camera, different illuminations and different face expressions. Figure (2) shows capture frames from different distances from the camera.Each color frame(RGB) is converted to gray frame and then smoothed it using Histogram Equalization [10] and pass it to Viola-Jones approach to face detection [11,12,13] post capturing frames process, the task was extracting only the individuals' faces from the captured images by The AdaBoost machine learning algorithm [14] andit saved the face image in file. This process is well illustrated in Figure (3).

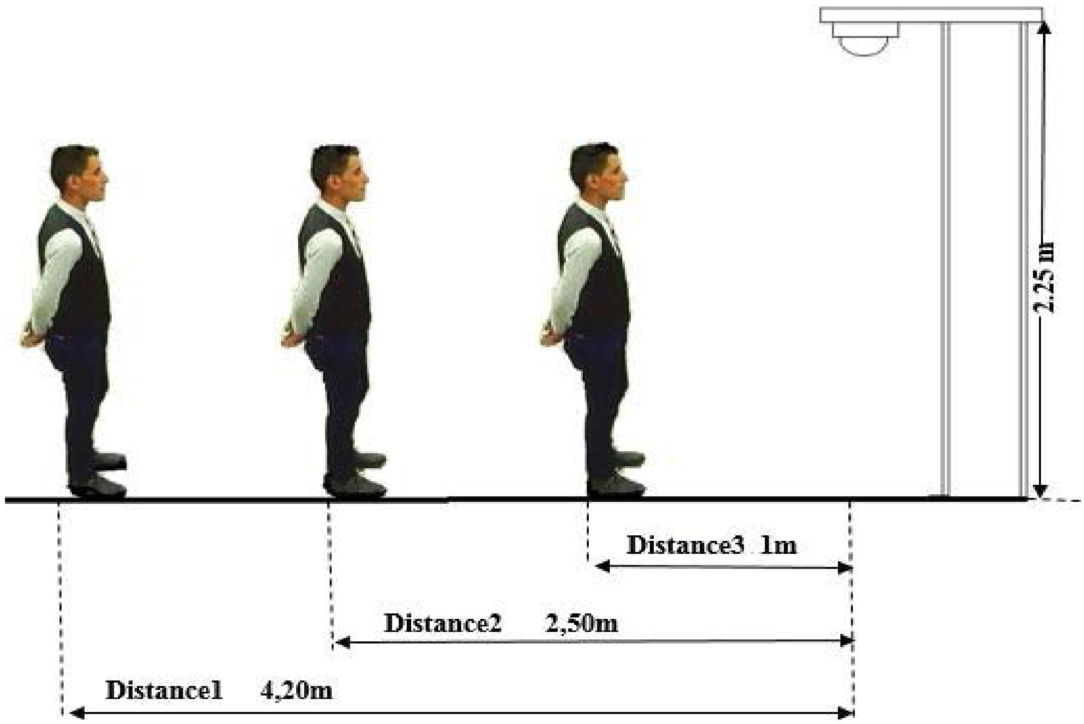

Figure 2.Capture image in different distance

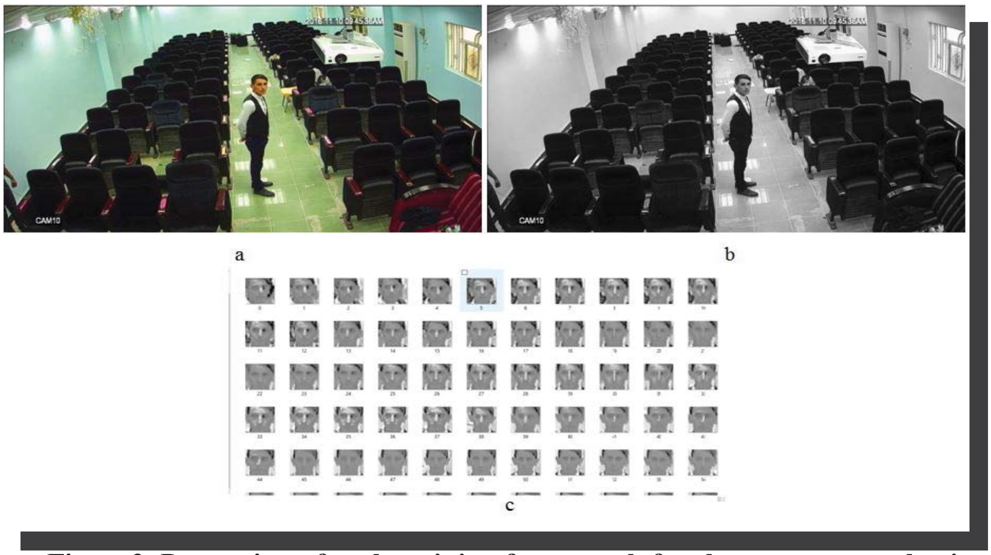

Figure3. Processing of each training framework for the system to authorized Persons. (a) Original frame (b) Enhancement grayscale frame (c) Sample of face authorized person after cut from frames.

ii. **Training the proposed System**

Training of the system on a set of persons authorized to enter the building during guard outside the official working to produce a strong classifier for dataset of static images of human faces. In addition to creating a database of these persons, both the dataset and database are use in the recognition phase. Figure (4) show the General schema of supervised training stage.

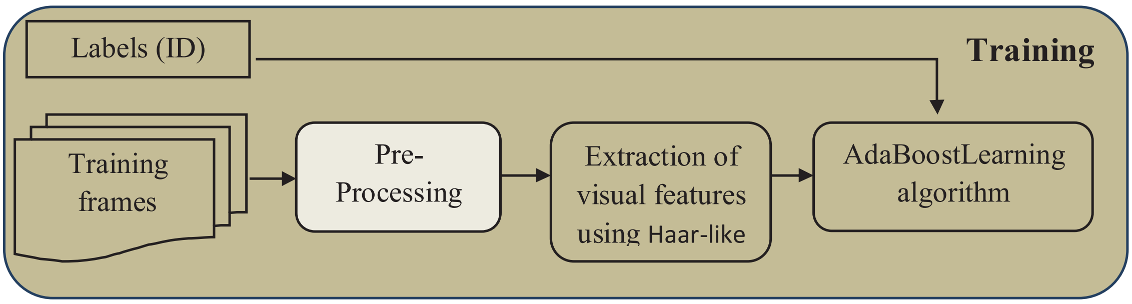

Figure 4. General Schema of Training system

Facial identification is valuable for the purposes of national security as it is for smaller scale surveillance systems. Nevertheless, for the sake of being capable of claiming that any facial identification system is sufficient, robust and reliable, it has to pass through thorough testing and verification, preferably on real-world data-sets. In this proposed system described a dataset of stationary images of human faces. Images were captured in controlled indoor environment with the use of four video surveillance cameras of the same quality. Dataset includes 200 stationary images of 4 subject’s rate 50 images per person. Images in these datasets are basically acquired under controlled illumination conditions, and are of high resolution (high quality capturing equipment has been utilized). Surveillance cameras have been installed in room at the height of 2.25 meters and placed as illustrated in figure (3). Each person has a special folder containing 50 samples taken from video cameras in the computers hard drive and was named in the following way:

Folder-name / number of image face.jpg; subject label (ID)

In this way, each one of the images in the dataset got an individual name that carries label (ID) Where each person has their own ID (this mean each 50 image have a same ID). The system will not save the images as files but interpret the features of each face and saves them into files.

Each ID in the image dataset correlates with ID in the Oracle database. The database contains a details of information about the trained individuals that will be used in the stage of excellence to display all information about the individuals are appearing in the control cameras. Figure (5) show the flowchart of training stage.

5. Monitoring and tracking stage

After training, the strong classifiers are produced to do face detection. The proposed system would be to detect faces using Viola-Jones in each frame. The objective of face detection would be to determine of whether there are any faces in the frame, then return the position and the bounding box of every one of the faces in the frame despite illuminations, occlusions, facial pose and expression.

For a given video, first of all we try to read a frame and see if we are successful doing so. If we are, then we do further processing. This condition is checked right after the input. If the frame is read then apply pre-processing step. Pre-processing step include to convert frame to a gray-scale image and normalize it by Histogram equalization. Later we call to Viola-Jones Haar features trained Cascade Classifier to detect faces in the frame,if faces are found, then recognized these faces using LBPH [15, 16] in the subsequent frames. Figure (6) shows the general schema of Monitoring and tracking stage. Figure (7) show the flowchart of face recognition. If no faces are found, then we just read next frame.

Tracking the infiltrator's through the recognizing of the face in each frame and compare it with a dataset which the system was trained on it, in order to distinguish the infiltrator's who penetrate the place and issuing the warning whistle, Then save the frame containing the stranger. The process of tracking faces in each frame results in them F_1,F_2... F_n frames are using to create an mp4 file from the tracked object (infiltrator's) in surveillance camera. Figure (8) shows samples of frames obtained in the tracking phase that are using in this process. Proposed system assumes a single people in each frame, or multiple people, the process of recognize not affected by the number of people appearingin each frame. Figure (9) shows an example of a person who has been authorized to enter with display him information from the database.

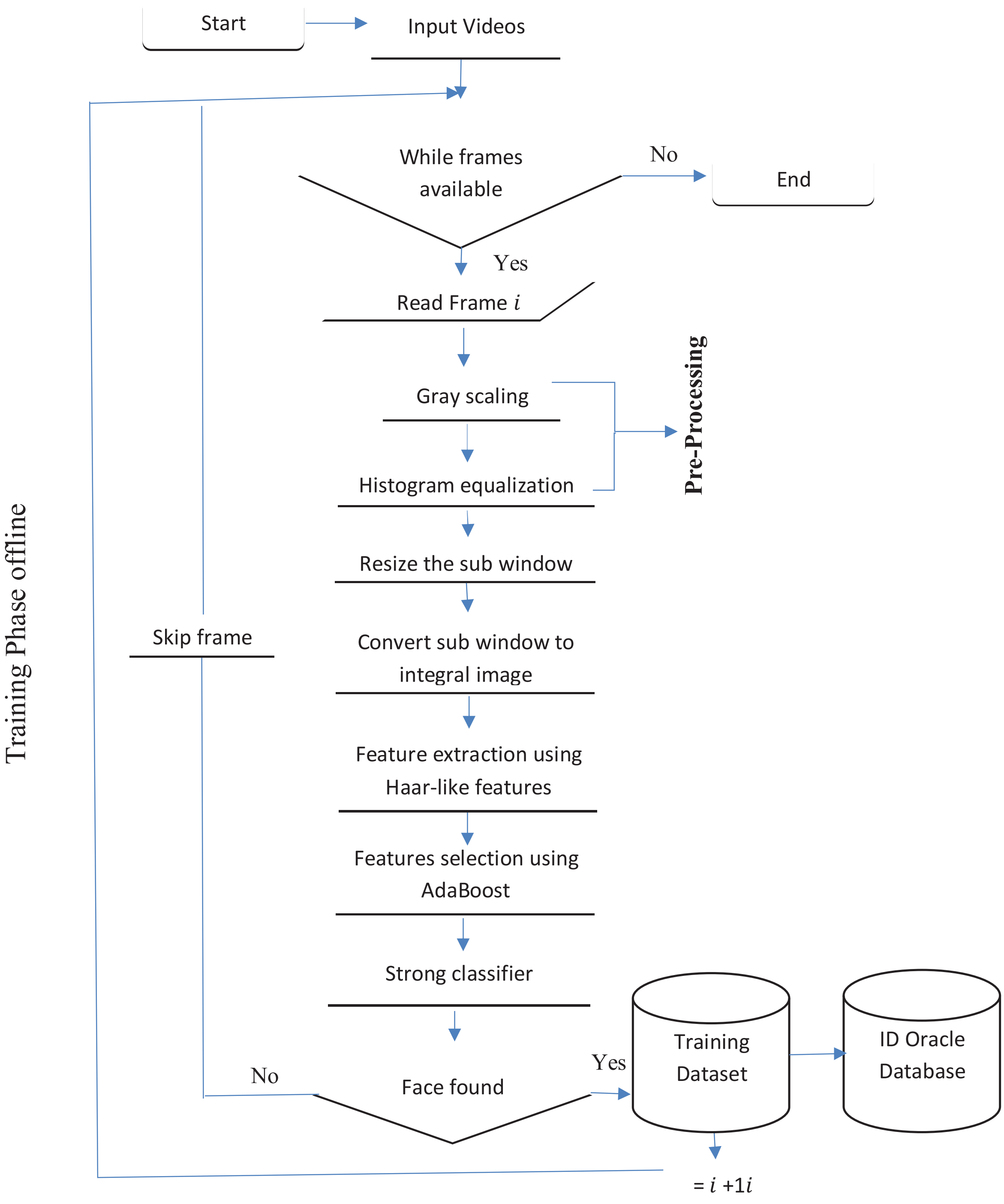

Figure 5.Block diagram of training stage

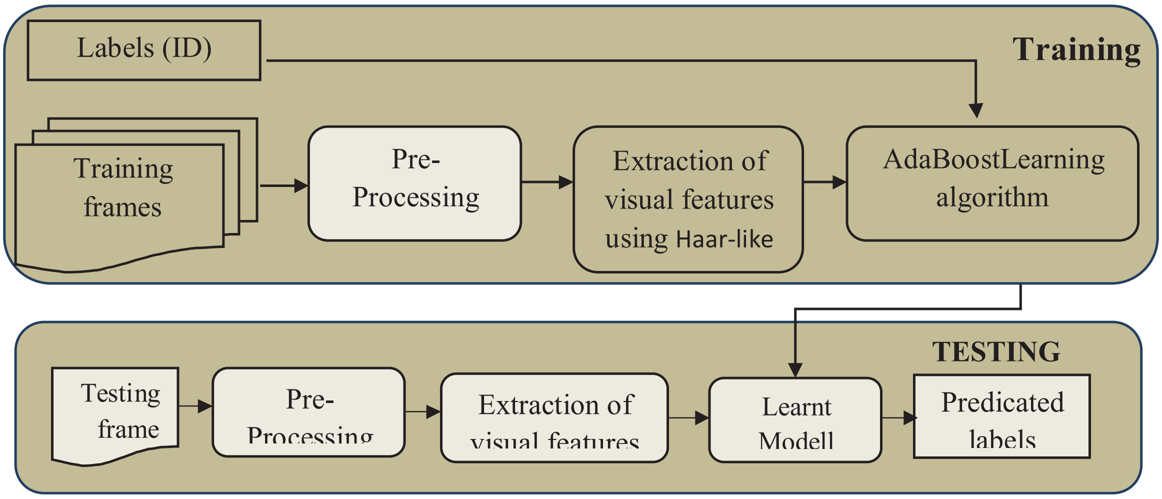

Figure (10) sho

Figure 6.General Schema of online phase (Testing phase) 's

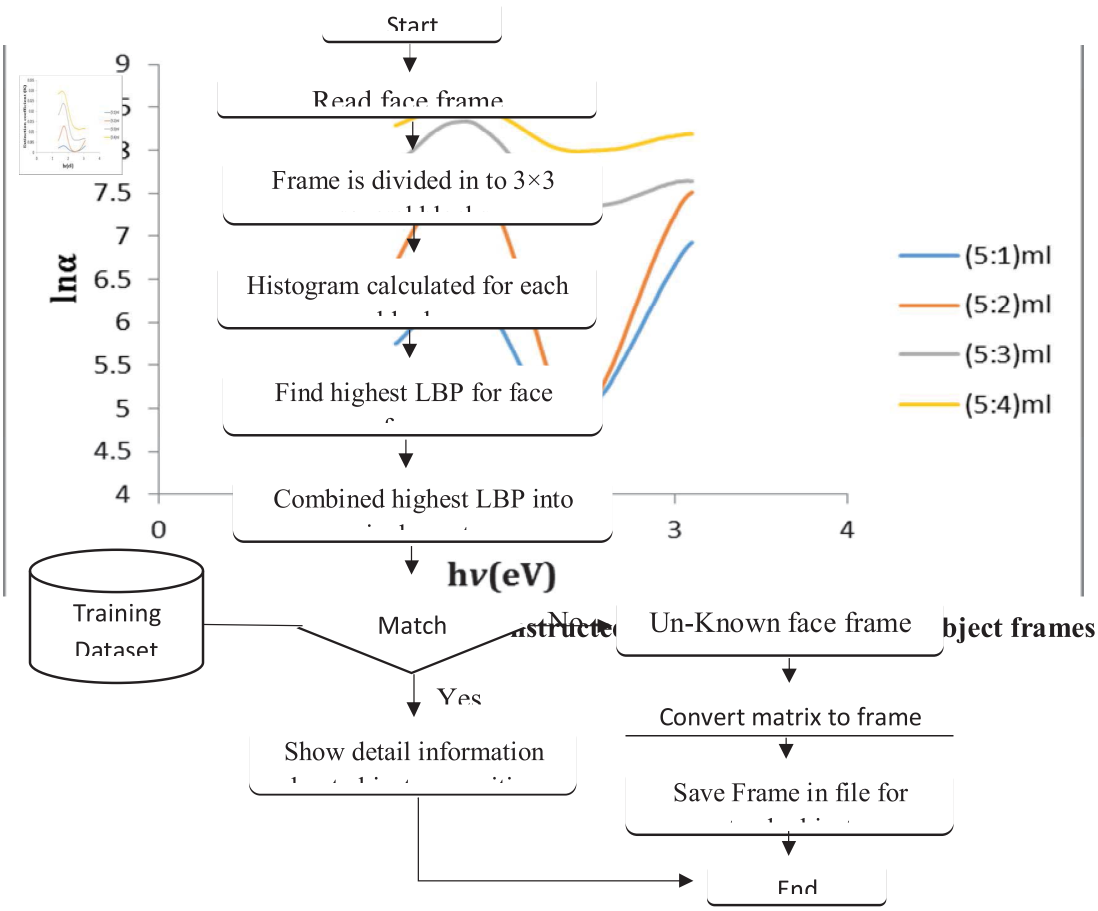

Figure 7. Block diagram of Face recognition steps

Figure 8. Sample of frames derived from trace object in four surveillance cameras

Figure 9.Example of inforamation authorized person that appear on monitor

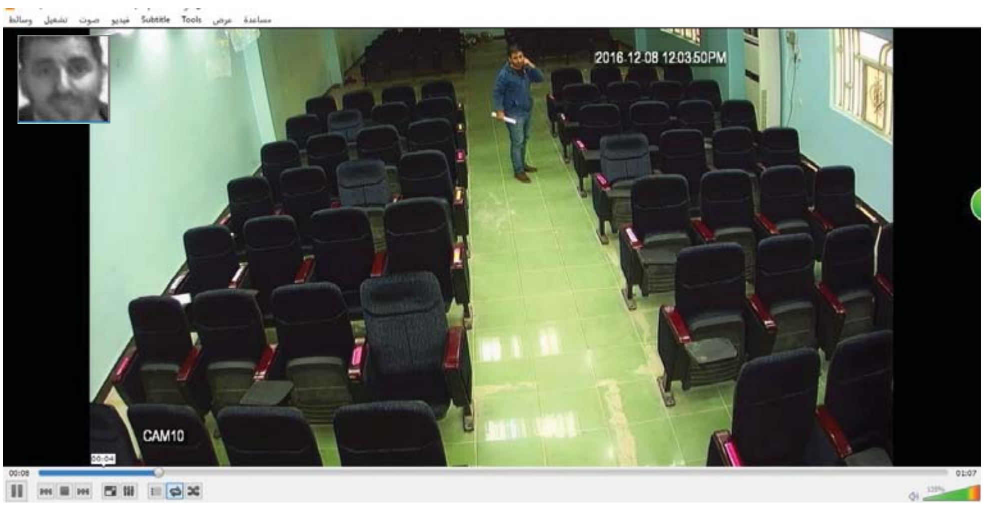

Figure 10.Reconstructed track video from track object frames

6. The Evaluation Proposed System

The proposed system detects the frontal face, the face detection process in the surveillance cameras is difficult since distance between the person and the camera and the difference of lighting affect the detection and discrimination. Therefore, the installation of surveillance cameras in places facing the entrance to the building and identify the specific areas (specific environment) are considered areas of detection so that anyone cannot enter the building without going through these areas.

Several different measures are currently used to evaluate the performance of the facial recognition system. This section shows some of them. The normative approach to dealing with assessment of the face recognition system revolves around the concept of ground truth for positive and negative detection.Table 1 shows the confusion matrix. Positive and negative terms reveal the asymmetry of disclosure tasks where one category is the relevant class category,whereas, the other is unrelated.The system must differentiate between the criteria of the face and the non-face. A true positive means that the part of facial images to be detected by the system, while a false positive means the part of the non-facial images to reveal as face.The detection rate and recall have the same meaning as the term true positive hereFalse positives mean wrong matching individuals with images in the database, not catching people even when they are photographed in the database.These case, mean that they are false negatives.

Table 1.Confusion matrix.

Ground truth\detection	**Detected (Positive)**	**Rejected (Negative)**
Relevant	True positive (TP)	False negative (FN)
Non-relevant	False positive (FP)	True negative (TN)

1. The Evaluation of TrackingAuthorized Persons

The results of applying the system in a specific environment are displayed in In Table (2) and Figure (11). 300 frames wascaptured from videos recorded from two-hour surveillance cameras , 60 frame without person and 240 with persons,for detection face in frame to three authorized persons. In this test, detection and recognation rates were calculated separately for each one. Used the videoes without person to detect the false positives and true negatives cases. The videoes with person, also obtained the false negatives and true positives cases for each sample.

Samples	TN	TP	FP	FN	Precision	Recall	Accuracy	Error rate	TNR	FPR
Person1	100	70	2	4	0.972	0.946	0.966	0.034	0.980	0.020
Person2	98	70	2	2	0.972	0.972	0.977	0.023	0.980	0.020
Person3	100	55	15	0	0.786	1.000	0.912	0.088	0.870	0.130
Average					**0.910**	**0.973**	**0.952**	**0.048**	**0.943**	**0.057**

Table 2.Result of evaluatedproposedsystem

Where the :

$Precision = \frac{TP}{TP+FP}$ (1)

$Recall = \frac{TP}{TP+FN}$ (2)

$Accuarcy = \frac{TP+TN}{TP+TN+FP+FN}$ (3)

$Error\ rate = \frac{\boldsymbol{FP+FN}}{\boldsymbol{TP+TN+FP+FN}}$ (4)

Specificity $(TNR) = \frac{TN}{TN+FP}$ (5)

False positive rate $(FPR) = \frac{FP}{TN+FP}$ (6)

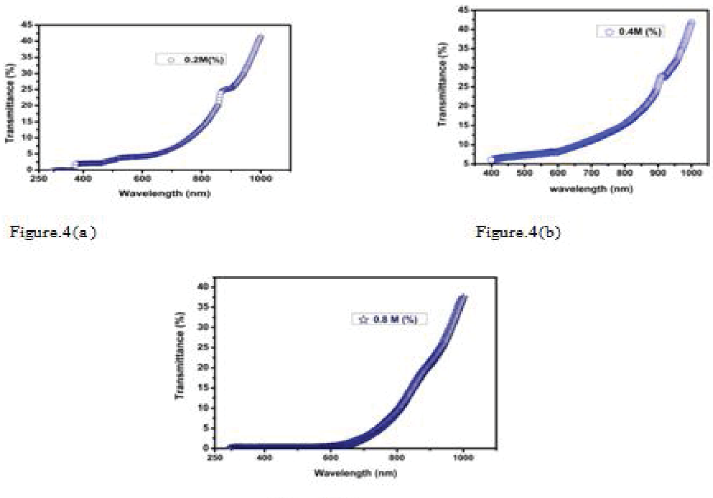

Figure.4(a) Figure.4(b) Figure.4(c)

Figure 11. Evaluation proposed system in specificenvairoment

2. **The Evaluation ofTracking untheorizedPerson**

The movement of untheorized (strangers) in the building was monitored through surveillance cameras and took more than one sample to measure the effectiveness of the proposed system. Table (3) and Figure (12) illustrates the system's achievement of six samples with a difference in the number of strangers between one sample and another and taking the average for each sample.

Samples	TN	TP	FP	FN	Precision	Recall	Accuracy	Error rate	TNR	FPR
1	500	44	5	2	0.898	0.957	0.987	0.013	0.990	0.010
2	660	26	2	3	0.929	0.897	0.993	0.007	0.997	0.003
3	395	10	8	3	0.556	0.769	0.974	0.026	0.980	0.020
4	500	40	3	0	0.930	1.000	0.994	0.006	0.994	0.006
5	600	20	2	0	0.909	1.000	0.997	0.003	0.997	0.003
6	420	29	10	4	0.744	0.879	0.970	0.030	0.977	0.023
Average					**0.827**	**0.917**	**0.986**	**0.014**	**0.989**	**0.011**

Table 3.Evaluation of Tracking Infiltrator's Personin surveillance cameras

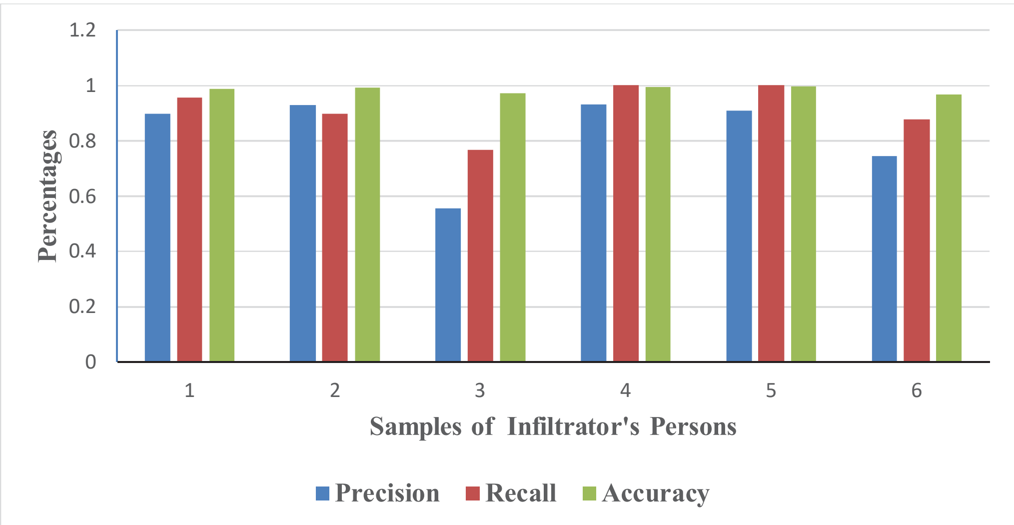

Figure 12.Evaluation of Tracking Infiltrator's Personin surveillance cameras

7. Conclusions

- Increasing the number of surveillance cameras does not affect the accuracy of the identification and tracking of people, and this increase affects only the speed of the system, so such systems require computers with a special specification (like server, super computer, mainframe etc.).
- There is a difference between the accuracy results of the open environment and the specific environment because the distance between the person and the camera has an important role in the process of detection and recognition.
- There is a need for continuous updating of training dataset to obtain an effective surveillance system.
- The system achieved efficient results and high accuracy for such systems that depend on the detection, recognition and tracking face in surveillance multi-cameras are represented by 95% of the detection faces,and 97% of the tracking of strangers which represent a high proportion compared to similar works.

Reference

[١] Lulu Chenet al. (2013), A survey of human motion analysis using depth imagery, Pattern Recognition Letters, Volume 34, Issue 15.

[٢] Yan, Zhiguo et al. (2017). The Big Data Analysis on the Camera-based Face Image in Surveillance Cameras. Intelligent Automation & Soft Computing: 1-9.

[٣] Choudharyei al. (2016).A Study of Digital Video Compression Techniques. PARIPEX-Indian Journal of Research 5.4.

[٤] Oh, J. M et al. (2016). Trajectory based database management for intelligent surveillance system with heterogeneous sensors. Multimedia Tools and Applications, 75(23), 15429-15444.

[٥] Hossen, M. K., and Tuli, S. H.(2016).A surveillance system based on motion detection and motion estimation using optical flow. In Informatics, Electronics and Vision (ICIEV), 2016 5th International Conference on (pp. 646-651). IEEE.

[٦] Shahabi C et al. (2014).multi source event detection and collection system for effective surveillance of criminal activity. J Inf Process Syst 10(1):1–22

[٧] Ball, Mark G. et al.(2016). A Review of the Use of Computational Intelligence in the Design of Military Surveillance Networks. In Recent Advances in Computational Intelligence in Defense and Security, pp. 663-693. Springer International Publishing.

[٨] Xiao et al.(2017). Surveillance Video Face Recognition with Single Sample per Person Based on 3D Modeling and Blurring.Neurocomputing.

[٩] Rasti et al.(2016).Convolutional Neural Network Super Resolution for Face Recognition in Surveillance Monitoring. In International Conference on Articulated Motion and Deformable Objects, pp. 175-184. Springer International Publishing.

[١٠] Ting et al.(2015). Visual Contrast Enhancement Algorithm Based on Histogram Equalization. Sensors.

[١١] P. Viola and M. J. Jones (2004).Robust real-time face detection, in: International journal of computer vision 57 (2), 137–154.

[١٢] P.ViolaandM.Jones(2001). Rabid Object Detection using a Boosted Cascade of Simple Features, in: Computer Vision and Pattern Recognition.

[١٣] Tikoo et al. (2017). Detection of Face using Viola Jones and Recognition using Back Propagation Neural Network. arXiv preprint arXiv:1701.08257.

[١٤] C. Niyomugaboet al.(2016).A Modified Adaboost Algorithm to Reduce False Positives in Face Detection, Mathematical Problems in Engineering.

[١٥] Md. Abdur Rahim et al.(2013). Face Recognition using Local Binary Patterns (LBP),in: Global Journal of Computer Science and Technology Graphics & Vision ,Volume 13 Issue 4 Version 1.0.

[١٦] Gao Xianget al.(2016).Face recognition based on LBPH and regression of Local Binary features”, in IEEE.