Received July 4, 2019, accepted July 27, 2019, date of publication August 27, 2019, date of current version September 10, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2937810

Real-Time Surveillance Through Face Recognition Using HOG and Feedforward Neural Networks

MUHAMMAD AWAIS[1], MUHAMMAD JAVED IQBAL[1], IFTIKHAR AHMAD[2], MADINI O. ALASSAFI[2], RAYED ALGHAMDI[2], MOHAMMAD BASHERI[2], AND MUHAMMAD WAQAS[3]

[1]Department of Computer Science, University of Engineering and Technology, Taxila 47080, Pakistan
[2]Department of Information Technology, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah 80221, Saudi Arabia
[3]Usman Institute of Technology, Hamdard University, Karachi 75300, Pakistan

Corresponding author: Iftikhar Ahmad (iakhan@kau.edu.sa)

This article was funded by the Deanship of Scientific Research (DSR) at King Abdulaziz University, Jeddah, under grant No. DF-133-611-1441. The authors, therefore, acknowledge with thanks DSR for technical and financial support.

ABSTRACT Real-time security requirements continue to increase due to the occurrence of various suspicious activities in open and closed environments. Day-to-day security threats may seriously affect everyone lives. Many techniques have been introduced in this regard, but still some issues remain unaddressed. The work presented in this paper provides video surveillance with improved accuracy and less computational complexity. The most significant part of the system consists of face localization, detection and recognition. The system obtains underlined facial data through a video dataset or from the real-time environment. Subsequently, face/foreground and background keyframes are extracted from at hand captured video data. Finally, extracted facial image data is compared with the facial images in the database. In case no match is found with the existing data, a security alarm or signal is generated, alerting security personnel to take action. The proposed system is more accurate, has better performance, and is low cost compared with existing systems.

INDEX TERMS Face recognition, HOG features, feedforward backpropagation neural network, surveillance video, principal component analysis.

I. INTRODUCTION

Facial recognition has significant applications in the domain of biometrics and several systems related to security and surveillance [1], [2]. With the advancement in time, there is a drastic development in the field of technology, which is leading in the formation of a smart society. A smart society has a goal of minimizing the human-to-machine intervention. This idea has further opened a new paradigm in the field of security and surveillance. All public and private security organizations must overcome the main issue, which is how to observe a person that enters an area. Organizations face a number of challenges regarding this matter. Even though a person is already authenticated, the personnel must still be monitored, which affects the performance of the system. Real-time recognition and detection of any unauthorized activity in the monitored area is also a major challenge in this domain. Environmental changes such as weather and lightning also affect the system performance. False alarms are also one of the issues. A false alarm can deny access to those who are authorized [3]–[5]. One of the hot research areas in the field of security is authentication during surveillance using different parameters. These parameters include biometric verification and radio frequency identification tags. The biometric system includes facial detection and recognition and fingerprint and eye retina comparisons. Among the biometric verification parameters, facial recognition for authentication has a vital role.

A number of frameworks can be used to solve the particular issues of facial detection and recognition, such as geometric and template matching methods [6]. Principal component analysis (PCA) is features/variables reduction method, which converts a large number of features/variables into a smaller number that must possesses the maximum available information with reduced complexity [7], [8].

Linear combination method is also related to linear discriminant analysis (LDA) [9]–[11]. The LDA framework separately performs analysis on multiple types of

The associate editor coordinating the review of this article and approving it for publication was Kostas Psannis.

objects, which are then separated into different types or classes [5], [12].

The major objective of this work is to recommend a facial detection and recognition system that identifies a person using their face. This work uses selected facial features and a popular multilayer feedforward neural network for the task of classification. The extracted features are determined and presented as a pattern vector to the neural network. The learning algorithm recognizes people's faces by learning the approximation of facial features, regardless of different facial movements. The feature matrix changes depending on the face motion. The use of four different video sequences provides enough data to train the classifier to identify a person in a crowd.

This research paper is arranged as follows. Section 2 provides a brief description of existing work in the domain of facial image identification and recognition. Section 3 provides the illustration of proposed methodology of surveillance facial detection and recognition. Section 4 comprehensively discusses the experimental results of training and testing the learning algorithm using benchmark and real-time databases. Finally, Section 5 provides a summary of this work and make recommendations for further developments.

II. RELATED WORK

Video surveillance has become an active research area due to an increase in everyday security threats. The identification and authentication of a person's face is an effective method to mitigate security threats. The facial recognition system was first developed through a semi-automated algorithm in 1960. Over the years, Viola and Jones improved continuous real-time facial identification with the wavelet-based function [13], [14]. In their framework, the first stage was movement and location detection. The rest of the stages depends exceptionally on facial detection and recognition.

In videos, detection and localization of the moving human face is achieved through background subtraction models. Any object detection method is primarily built on this principle: at the start, a set of rules is applied to the frames obtained from the digital camera feed, utilizing the background subtraction models. Images obtained through this method do not contain any transferring item. More details can be found in a study on background modeling [15]. Background subtraction has various limitations, such as films with terrible signal to noise ratio caused by low-resolution cameras, blur caused by jittering movement in the camera optic, surrounding environment noise, compression artifacts, illumination adjustments, and shaking trees [16].

Background subtraction can also be accomplished by using Gaussian mixture version [1]. Machine learning and image processing [17] benefit from multimodal mean live video streaming background modeling techniques. To overcome changes caused by tiny moving actions, background multimodal techniques are used with Gaussian mixture model (GMM) chance distribution features [18]. Other segmentation techniques use pixel differences among two to three frames to extract successive shifting regions, as time-based changing models. Dynamic background modeling is a time-based changing model [8]. This operation is established with a neuro-fuzzy approach to reduce shadows.

Most feedforward techniques for face recognition depend on skin color data [19]. Skin color pixels have value in the ranges of $0.36 < r < 0.456$ and $0.28 < g < 0.363$ [20], [21]. The histogram and Gaussian models are utilized for skin color pixel recognition to defeat singular constraints and build precision up to 90.00%. Hue, saturation, and value (HSV) color model shows pixel classification, where H has values of 0 to 50 and S has values of 0.20 to 0.68 [22], [23]. Models HSV and $YCbCr$ are combined. Accordingly, the range of Cb is less than or equal to 125 or Cb is greater than or equal to 160; Cr is less than or equal to 100 or Cr is greater than or equal to 135. H has a range of 26 to 220. Skin-shading-based technique is sensitive to changes in brightness and cannot be used if the background contains a skin-shading-like object. Facial recognition develops facial edge map and skin-shading threshold value using the $YCbCr$ space for skin color pixel identification and Viola–Jones technique to verify recognition. Viola–Jones employed a constant Adaboost calculation for the arrangement of rectangular facial highlights [12], [13]. Progressive *HEER* features are utilized to make up for the present variety. Viola–Jones facial identifier is joined with skin color depth information to decrease the number of false positives. Furthermore, the next stage after facial identification is acknowledgment, where an individual's personality is confirmed via matching request with a facial coordinate in the dataset or database.

The significance of facial recognition techniques in different hot areas is becoming very critical and through it many highly reliable daily life applications are being developed. Recently, artificial neural networks has been used in many facial recognition tasks for the objective to obtain better accuracy results. It consists of different components that work in parallel with some objective function. Neural network has been also used for human face feeling and gender classification. The neural network works well on the photographs with varied lighting conditions and increases accuracy. The foremost drawback of the NN may be a great deal of your time needed for its training. ANN [24], [25] acknowledges the facial through learning former expertise. NN together with progressive intelligence was conjointly employed for the face recognition purpose [26], [27]. The Probabilistic Neural Network (PNN) [28] approach was designed by Vinitha and Santosh that detects and recognizes the faces from the grayscale pictures containing the frontal faces. The most advantage of PNN exploitation is that it needs short training time. The Network within the PNN is split into subnets because of this network isn't fully connected. SOM (Self-Organizing Map Neural Network) [29] have the property of topological preservation is a man-made neural network employed in face recognition. After the feature extraction of facial images, the RBFN or FFNN are alternative popular NN classifiers which accomplish the facial recognition task in easier and accurate way.

In [30], various issues found in the implementation of security and surveillance were highlighted. These include; various protection checks to validate the identification of an authenticated person, the changes in the atmosphere affects and an incorrect false alarm error. The work presented in [30] acceptable answer to the mentioned drawbacks with higher accuracy results.

The work presented in this paper intends to overcome or mitigate of effects of the issues mentioned in the existing literature.

III. PROPOSED MODEL

Several approaches have been described in the previous section. The potential problems with each method and alternative strategies for solving these problems are discussed subsequently. The major components of the utilized methodology are illustrated in Figure 1.

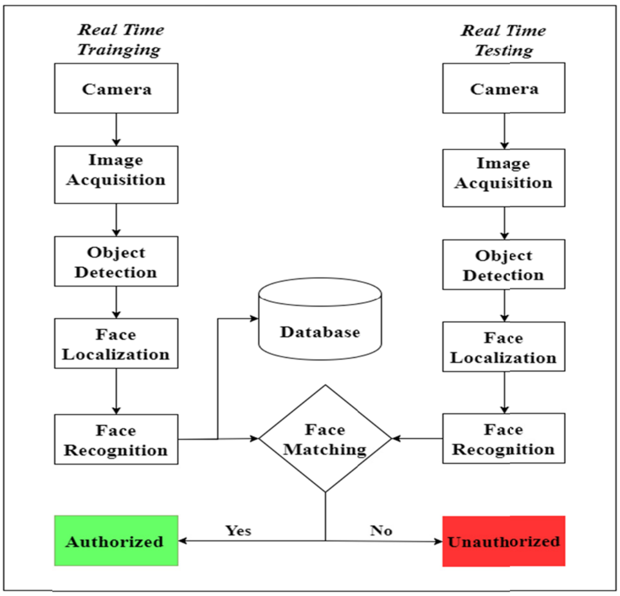


FIGURE 1. Proposed Model of surveillance via face recognition.

Figure 1 shows that the model uses video data captured through a high-resolution CCTV camera. Next, image acquisition draws the keyframe of the image. Then, the foreground and background is extracted from the image, leading to face detection. Furthermore, the face from the detected image needs to be localized. After the facial localization, the facial recognition is matched with stored images in the database. The present study is helpful in such context. The proposed system is implemented to enhance the facial localization and recognition mechanism. Finally, the extracted image is matched with the previously stored images in the database. In case no match is found, a security alarm or signal is generated for further actions. The currently processed image parameters are then updated in the database. The saved facial data could be used in the future to maximize the identification of source location, minimize complexity and time consumption.

A. DATABASE DESIGN AND CREATION

The suggested system uses CCTV cameras as image acquisition tools from a real-time environment. The facial recognition process works on real-time video streaming images provided by the camera. The underlined method extracts a single image frame from the video stream. This acquisition image window size is set to 480 x 640 pixels. Human facial detection is the primary and important step for this facial recognition task. The primary aim of a human facial detection system is to detect any human face in an image. The system decides in the preprocessing phase if a face is located in the obtained image. The system repeatedly checks the human face from the input data. If a human face is not found in the keyframes of real-time video streaming, then the system will automatically repeat previous instructions until a face is detected, as shown in Figure 2.

FIGURE 2. Simple input image using the real-time video stream.

FIGURE 3. Cropped facial image from the database.

Figure 3 shows a cropped human face image using the keyframe. The cropped facial image is a small part of the original keyframe that encompasses a human face. These facial images are helpful in creating the required database. In case of offline work in the system, two different datasets are used. These systems include ChokePoint Faces [20] and AT&T datasets [21], [31].

The camera used in this face recognition tasks is directly connected to a computer or a network and supplied with a live video feed. Such cameras usually does not have their own storage memory and thus use system memory. For this study, we have used a typical camera (HP Wide Vision HD). The specifications of cameras used in the research implementation are presented in Table 1.

TABLE 1. Camera properties.

Name:	HP Wide Vision HD	Backlight Compensation:	0
Mega Pixel	0.9	Exposure	−6
Resolution	640 x 480	Sharpness	2
Available Resolutions	(1 x 6 cell)	Contrast	32

The camera was used to detect faces from the images. The system starts with image acquisition. Then, the object detector captures the image cascade via the Viola–Jones algorithm to detect various parts of the face, as nose, eyes, and mouth etc. Image labeling can also be applied in the devised system to train a neural network (NN) classifier. The detected facial feature is then saved in a database with an authorized perso's name and a facial image, as shown in Figure 4. Each facial image is detected in real-time system with a delay of 2 seconds.

FIGURE 4. Object detector.

B. FACIAL DETECTION AND LOCALIZATION

Facial detection and localization is the primary and important step to find if there is any human face in an image or not. Figure 4 represents the process of detection and localization. The obtained facial image is generated from a 3D RGB color scale to a monochromatic grayscale. True color image (color map) to a grayscale fixed intensity image is analyzed in this work.

Furthermore, histogram equalization is applied to automatically adjust intensity values, generating the grayscale image. This method involves converting intensity values of histogram. Moreover, the intensity values of the produced facial image are approximately equivalent to the specified histogram, as shown in Figure 5.

The histogram matrix applies the median filter of the facial image in two dimensions. Each output pixel of the image contains the approximate median value in a 3x3 neighbor around

FIGURE 5. RGB scale to grayscale image, histogram equalization of image, and median filter applied to the image (left to right).

the corresponding pixel within the input of the histogram image.

Each facial image is then labeled. The given images have many non-human-connected alternatives. The algorithmic labeling rule is used to name each human face within distinct intervals with a distinct number. This rule helps in managing every human face individually.

C. FACIAL RECOGNITION

Facial recognition is the new trend in security authentication systems. For real-time applications, security and surveillance must address the challenges of recognizing faces from live video streaming. In the proposed system (at the entrance of a building), the environment and lighting are constant, but variations in terms of lighting direction and shadow are expected. The system normalizes the image using a particular custom format of 150x150 pixels (Row, Column) for convenience [32].

The presented technique utilizes a histogram of oriented gradients (HOG) for fetching of facial features followed by the feedforward backpropagation neural network classifier. A sliding window approach is performed for the detection of an image. The sliding window of detection has a fixed size 150x150. The sliding window detects prominent image pixels from the whole face image. The analysis is divided into two phases. First, the descriptor value is calculated for each sliding window of detection via the HOG feature method [23], [33]. Subsequently, the descriptor values are categorized by applying an FFNN classifier.

The HOG feature extraction process works as follows:

I. The extracted face image is partitioned into small associated pixels values called cells. The HOG edge direction of the pixels in a cell is calculated, as shown in Figure 6.
II. Each cell is separated in an angular bin in accordance with the edge of the gradient.
III. The pixels of every cell contribute to the weighted gradient of the corresponding angular bin.
IV. Neighboring group cells considers spatial areas known as blocks. Grouping cells in a block is the basis of normalizing histograms.
V. The normalized histogram group represents the histogram of the block. The collection of these blocking histograms optimizes the HOG feature of an image.

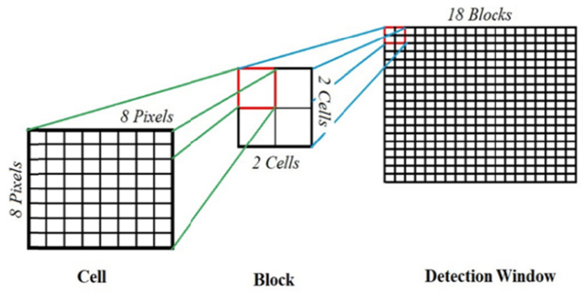


FIGURE 6. Detection sliding window divided into blocks and cells.

During HOG feature extraction, the face image is first split into several small cells. Using Equations 1 and 2, 2D gradient is determined for every pixel (x, y) in the given input image d.

Equation 3 is used to find gradient magnitude |GR (x, y)|, and Equation 4 to determine the gradient angle Ø (x, y) at position [23], [33].

$$\text{GRX} = \text{Imp(Xi + 1, Yj)} - \text{Img(Xi} - \text{1, Yj)} \quad (1)$$
$$\text{GRY} = \text{Img(Xi, Yj + 1)} - \text{Img(Xi, Yj} - \text{1)} \quad (2)$$
$$\text{GR(X, Y)|} = \sqrt{GRx^2 + GRy^2} \quad (3)$$
$$\tan(\varphi(X, Y)) = \frac{GR_X}{GR_Y} \quad (4)$$

The fixed size of 8x8 pixel detection sliding window is distributed into non-overlapping cells of the matrix. A block consists of four neighboring cells of the matrix. The blocks have vertical and horizontal correspondences of one cell. The obtained result is a block dimension of 18 x18, which is equal to a total of 324 cells in a detection sliding window.

For each cell, a histogram is produced using eight bins for the gradient angle. In each bin, the accumulated weighted gradient magnitudes are calculated. To connect the histograms of cells enclosed in a block, the block histogram is delivered with 4*8 = 32 elements. To increase the detection feature, the gradient magnitude at the outer pixels of the blocks has a weight less than the central ones. Then, the Gaussian matrix is sized to 16*16 multiplied by |G|Block, which is the gradient magnitude in a block.

$$|\text{GR}|\text{Gaussian} = \text{Fg} * |\text{GR}|\text{Block}, \quad (5)$$

where $\sigma = 8$ and * denotes the multiplication of the element .

Gradient magnitude added bins to create the histogram elements [23], [33]. For improved invariance to the illumination and contrast variations, the block of the histogram is labeled as L2 and the function of the normalized vector is good, as shown in Equation 6.

$$\text{HOG} = \sqrt{\frac{|GR|gaussian}{||GR|gaussian||1+\varepsilon^2}}. \quad (6)$$

Finally, a total of 324 blocks of histograms are concatenated to acquire the 324*32 = 10,368-dimensional HOG features of each image, as shown in Figure 6. After the extraction of features, a PCA is used to reduce facial feature dimensions. Dimension reduction is applied after extracting the HOG feature on each facial image.

A 2D facial image is characterized as a 1D vector by connecting each column (or row) into a large single vector. Consider a set of sampled images representing M vectors of size N (equal rows of image "x" column). Equation 7 describes that p' shows the pixel measurements. The mean centered facial image "I" is obtained by subtracting the image from each image vector using Equations 8 and 9 and "wi" as a mean-centered or scaled image. Figure 7 shows the resized image along with the HOG features.

$$x_i = [P_1 \ldots\ldots P_N]^T, \quad i = 1, \ldots, M \quad (7)$$
$$I = \frac{1}{M}\sum\nolimits_{i=1}^{M} x_i \quad (8)$$
$$w_i = x_i - I \quad (9)$$

To find a set of "M" orthonormal vectors "ei" for which the quantity,

$$\lambda_i = \frac{1}{M}\sum\nolimits_{n=1}^{M}\left(e_i^T w_n\right)^2 \quad (10)$$
$$e_i^T e_k = \delta_{lk} \quad (11)$$

is the orthonormality constraint, which is maximized in Equation 11.

Here, "ei's" and "λ_i's" are given by the covariance matrix using the eigenvectors and eigenvalues [34], [35].

The eigenvectors are sorted in a descending order in accordance with their corresponding eigenvalues. The eigenvector of the largest eigenvalue shows the highest variance in the data [32]. This decrease in exponential fashion (roughly 90% of the total variance) is controlled in the first 20% to 50% of the magnitudes/dimensions. A facial image is mapped into M' (<<M) dimensions.

In the experiments, PCA is utilized to reduce the size of facial image features from ten thousand to a few hundreds. To train a neural network classifier, FFNN gives a general structure to indicate a nonlinear operational mapping between an arrangement of the input and output sets. Activation operation is accomplished by showing the nonlinear operation of a single variable of the various variables in positions of arrangements [9], [34].

FFNN is less complex and one o the commonly used neural network model used for recognition tasks. In it, inpu image features are enclose within the input layer, whic is shown within the figure a a grou of circles labeled *input*1 to *input* i^{th}. Figure 7 shows that the hidden layer has neuron weights. The hidden neurons as circles labeled *H1* to *Hj* have sigmoi transfer operation. The output layer obtains the hidden layer neurons generated from the output, which is the set of neuron weights from the hidden layer. A linear transfer function is placed inside each neuron in the output layer. The issue is a regression type, and the sigmoidal function and linear transfer method are utilized on the hidden and output layers, respectively. Figure 8 shows the construction of the multilayer FFNN [5], [10], [12], [17].

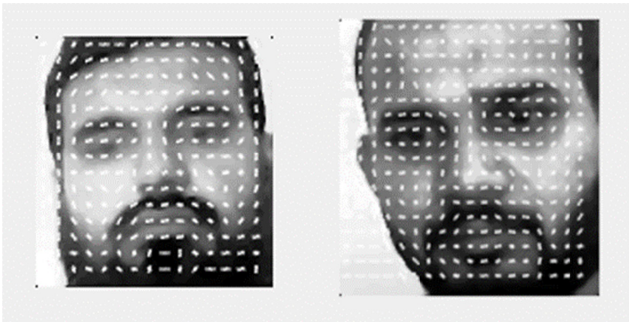


FIGURE 7. Resized image (150x150 pixel) and applied HOG feature function.

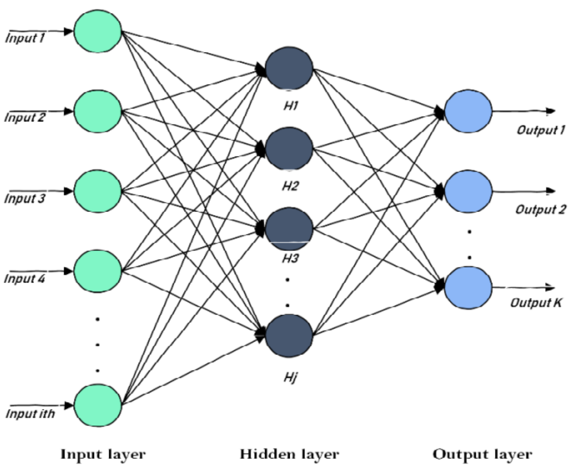


FIGURE 8. Architecture of the multi-layer FFNN classifier.

Figure 8 presents the architecture of the FFNN classification model. The preprocessed input image consists of 4,356 features. The hidden and output layers have 100 and 25 neurons, respectively. The best validation of the performance by cross-entropy of the system has been found in the 53 epochs.

IV. EXPERIMENTAL RESULTS

New faces images taken by an image acquisition device are compared with image information in the databases and then saved in the database. The proposed system uses the values of prediction and score comparison with a face classifier to test an image. If an authorized person is found from the database, the current image matches with the existing data. If a match is not found in the database, a security alarm will be generated. Moreover, the information of an unauthorized person model will be added in the database.

The output of the real-time face identification and authentication system is shown in Figure 9. The labeling of the live stream of a human facial image is marked with his or her name stored in the database. The result of the FFNN is a vector demonstrating a person's name from the register. Moreover, if the detected face is not found in the database, then the input image is labeled as "Unknown Person." The accuracy and error rate are described using Eq 12 and Eq 13.

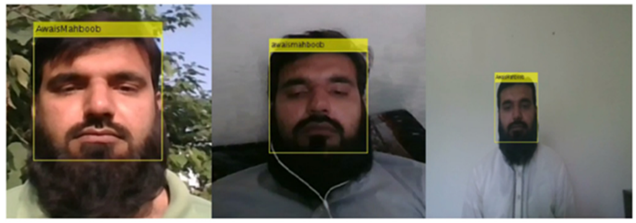


FIGURE 9. Match decision.

$$Accuracy = \frac{TP}{TP + FN} \tag{12}$$

$$Error\ Rate = 1 - Accuracy \tag{13}$$

where, TP is true positive and FN is false negative.

Real-time security and surveillance system through facial recognition recognizes the face of an individual. Furthermore, the particular system obtains images from a video camera through video streaming. The experimental results of this study are presented in Tables 2 and 3. The results are stable, with fair recognition rate in normal conditions such as frontal face view, normal light condition, and fixed face size (150x150). Results based on the variations in different environmental factors are shown in Table 2.

TABLE 2. Experimental results of real-time facial recognition based on environmental changes.

Facial Condition		Recognition Rate	Error
Normal	Frontal Face View, Normal Light	99.88%	0.12%
Light Changes Fix Face Size (150x150)	Sunny Day	97.50%	2.50%
	Fluorescent	96.44%	3.56%
	Cloudy Day	92.34%	7.66%
	Dark	90.83%	9.17%

TABLE 3. Experimental results of real-time facial recognition based on facial detection size.

Facial Condition		Recognition Rate	Error
Facial Detection Size	150x150	94.85%	5.15%
Normal light	200x200	98.66%	1.34%

Comparison between Real-time and Offline Datasets

The results show that the standard dataset, which is prepared for this type of experiments has excellent accuracy due to its highly restricted nature. The employed databases provides face images related to security and surveillance.

Proposed work is applicable for offline and real-time environments. However, the proposed dataset has several challenging execution complexities mentioned in Table 5, such

TABLE 4. Comparison between Standard and Proposed datasets.

Reference	Database	Techniques	Accuracy
[2]	AT&T faces	LBP, Multi KNN & BPNN	95.71%
[6]	LFW	image sensor (CIS), CNN	97.00%
[7]	PIE database, CMU, Yale face	VGG classifier, CNN	91.82% to 98.94%
[5]	-	Back Propagation Networks (BPC), CNN	96.66%
[12]	AT&T face	PCA, CASNN and FFNN	92.50% to 97.50%
[14]	Yale Face	PCA, FFNN	90.00%

TABLE 5. Accuracy comparison without PCA between standard and proposed datasets.

Dataset	Image Type	Image Back-ground	Face Size	Face Position
AT&T (ORL) Faces [31]	Png format	Uniform	Uniform	Uniform
Choke Point Faces [20]	Jpg format	Uniform	Uniform	Uniform
Proposed Real-time Dataset	Jpg format	No limitation	Different Scaling	Vary Significantly

as image dimensions, background, size, and position of the face.

The proposed real-time dataset has an improved system accuracy and this work also minimized the recognition error rate mentioned in Table 5.

TABLE 6. Accuracy comparison between standard and proposed datasets using PCA.

Dataset	Recognition Rate	Error
AT&T (ORL) Faces	92.00% to 96.00%	4.00%
ChokePoint Faces	70.00% to 88.00%	12.00%
Proposed Real-time Dataset	90.00% to 96.80%	3.20%

Table 6 presents the comparison of standard and proposed real-time datasets by using the principle component analysis.

When the limitations and restrictions are ignored, the developed system will be improved by the results from another dataset with less limitation for training and testing images.

Table 7 shows comparison of proposed results with the previous results. The results of our work are comparable with the previous approaches. Although deep learning based convolutional neural networks(CNN) have shown better results but those techniques are computational very inefficient. The ultimate objective of this work was low cost and simple solution.

TABLE 7. Comparison of results with existing work.

Dataset	Recognition Rate	Error
AT&T (ORL) Faces	97.50% to 99.00%	1.00%
ChokePoint Faces	91.00% to 99.91%	0.09%
Proposed Real-time Dataset	90.00% to 99.88%	0.12%

Proposed work implemented histogram of oriented gradients pattern of images to create the single facial feature vector. Two types of offline and real-time datasets were used in the experiments for validation of results. To reduce the dimensionality of feature vector, principal component analysis was used. The system was trained and tested using artificial neural network classifier. This work will be helpful in efficient detection of any authorized activities in the monitored area. Ultimately, the results of our system was evaluated and compared with the existing techniques mentioned in the literature. The proposed systems shows better or at least comparable performance in terms of accuracy and complexity.

V. CONCLUSION

Facial images captured in a video streaming surveillance environment typically suffer from very poor quality. In addition, due to the characteristics of the cameras, uncontrolled capturing conditions may lead to ambient variations, such as lighting changes, face pose, light shadowing, and body or face motion blur. Subsequently, the quality of facial captures using video surveillance cameras can affect the performance and effectiveness of video-based facial recognition systems.

Keeping in consideration today's security demands, we have proposed a video surveillance system that uses HOG features and FFNN classifier. However, the feature pattern can have variations due to the change of facial movement in a different video sequence. The successful implementation uses 25 facial images, each displaying four distinctive collections. The use of facial features through four movement sequence assigns the FFNN with an approximated understanding of identity. The proposed model was tested under extremely diverse conditions, and it performed efficiently and accurately. In the future, the system may be extended on some larger datasets using other deep learning methods, such as convolutional neural networks.

ACKNOWLEDGMENT

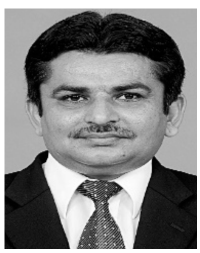
The authors would like to acknowledge with thanks DSR for technical and financial support.

REFERENCES

[1] Z. Shao, J. Cai, and Z. Wang, "Smart monitoring cameras driven intelligent processing to big surveillance video data," *IEEE Trans. Big Data*, vol. 4, no. 1, pp. 105–116, Mar. 2018.

[2] M. A. Abuzneid and A. Mahmood, "Enhanced human face recognition using LBPH Descriptor, multi-KNN, and back-propagation neural network," *IEEE Access*, vol. 6, pp. 20641–20651, 2018.
[3] D. Meena and R. Sharan, "An approach to face detection and recognition," in *Proc. Int. Conf. Recent Adv. Innov. Eng. (ICRAIE)*, Dec. 2016, pp. 1–6.
[4] B. S. Satari, N. A. A. Rahman, and Z. M. Z. Abidin, "Face recognition for security efficiency in managing and monitoring visitors of an organization," in *Proc. Int. Symp. Biometrics Secur. Technol. (ISBAST)*, Aug. 2014, pp. 95–101.
[5] N. Jamil, S. Lqbal, and N. Iqbal, "Face recognition using neural networks," in *Proc. IEEE 21st Century Int. Multi Topic Conf. (INMIC)*, Dec. 2001, pp. 277–281.
[6] K. Bong, S. Choi, C. Kim, D. Han, and H.-J. Yoo, "A low-power convolutional neural network face recognition processor and a CIS integrated with always-on face detector," *IEEE J. Solid-State Circuits*, vol. 53, no. 1, pp. 115–123, Jan. 2018.
[7] Y.-H. Kim, H. Kim, S.-W. Kim, H.-Y. Kim, and S.-J. Ko, "Illumination normalisation using convolutional neural network with application to face recognition," *Electron. Lett.*, vol. 53, no. 6, pp. 399–401, 2017.
[8] S. Z. Li and A. Jain, *Handbook of Face Recognition*. New York, NY, USA: Springer Verlag, 2011.
[9] D. Malik and S. Bansal, "Face recognition based on principal component analysis and linear discriminant analysis," *Int. J. Res. Electron. Commun. Technol.*, vol. 3, pp. 7–10, Jun. 2016.
[10] P. Padilla, J. F. Valenzuela-Valdés, J. L. Padilla, and F. Luna-Valero, "Electromagnetic near-field inhomogeneity reduction for image acquisition optimization in high-resolution multi-channel magnetic resonance imaging (MRI) systems," *IEEE Access*, vol. 5, pp. 5149–5157, 2017.
[11] F. Z. Chelali, A. Djeradi, and R. Djeradi, "Linear discriminant analysis for face recognition," in *Proc. Int. Conf. Multimedia Comput. Syst.*, Apr. 2009, pp. 1–10.
[12] A. J. Dhanaseely, S. Himavathi, and E. Srinivasan, "Performance comparison of cascade and feed forward neural network for face recognition system," in *Proc. Int. Conf. Softw. Eng. Mobile Appl. Modeling Develop. (ICSEMA)*, 2012, pp. 1–6.
[13] K. Vikram and S. Padmavathi, "Facial parts detection using Viola Jones algorithm," in *Proc. 4th Int. Conf. Adv. Comput. Commun. Syst. (ICACCS)*, Jan. 2017, pp. 1–4.
[14] T. H. Le and L. Bui, "Face recognition based on SVM and 2DPCA," 2011, *arXiv:1110.5404*. [Online]. Available: https://arxiv.org/abs/1110.5404
[15] R. Arroyo, J. J. Yebes, L. M. Bergasa, I. G. Daza, and J. Almazán, "Expert video-surveillance system for real-time detection of suspicious behaviors in shopping malls," *Expert Syst. Appl.*, vol. 42, no. 21, pp. 7991–8005, Nov. 2015.
[16] S. C. Loke, "Astronomical image acquisition using an improved track and accumulate method," *IEEE Access*, vol. 5, pp. 9691–9698, 2017.
[17] R. K. M. P, K. S. R, and K. M. Aishwarya, "Artificial neural networks for face recognition using PCA and BPNN," in *Proc. IEEE Region Conf. (TENCON)*, Nov. 2015, pp. 1–6.
[18] M. Slavković and D. Jevtić, "Face recognition using eigenface approach," *Serbian J. Elect. Eng.*, vol. 9, no. 1, pp. 121–130, Feb. 2012.
[19] H. M. Desai and V. Gandhi, "A survey: Background subtraction techniques," *Int. J. Sci. Eng. Res.*, vol. 5, no. 12, pp. 1365–1367, 2014.
[20] Y. Wong, S. Chen, S. Mau, C. Sanderson, and B. C. Lovell, "Patch-based probabilistic image quality assessment for face selection and improved video-based face recognition," in *Proc. IEEE Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2011, pp. 74–81.
[21] *Face Recognition Homepage*. Accessed: Oct. 15, 2017. [Online]. Available: http://www.face-rec.org/databases/
[22] M. Hahnle, F. Saxen, M. Hisung, U. Brunsmann, and K. Doll, "FPGA-based real-time pedestrian detection on high-resolution images," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops*, Jun. 2013, pp. 629–635.
[23] K. Mizuno, Y. Terachi, K. Takagi, S. Izumi, H. Kawaguchi, and M. Yoshimoto, "Architectural study of HOG feature extraction processor for real-time object detection," in *Proc. IEEE Workshop Signal Process. Syst.*, Oct. 2012, pp. 197–202.
[24] T. Sakamoto, T. Sato, P. Aubry, and A. Yarovoy, "Fast imaging method for security systems using ultrawideband radar," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 52, no. 2, pp. 658–670, Apr. 2016.
[25] J. Xu, S. Denman, S. Sridharan, and C. Fookes, "An efficient and robust system for multiperson event detection in real-world indoor surveillance scenes," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 25, no. 6, pp. 1063–1076, Jun. 2015.
[26] B. Hou, R. Zheng, and G. Yang, "Quick search algorithms based on ethnic facial image database," in *Proc. IEEE 5th Int. Conf. Softw. Eng. Service Sci.*, Jun. 2014, pp. 573–576.
[27] R. Sarkar, S. Bakshi, and P. K. Sa, "A real-time model for multiple human face tracking from low-resolution surveillance videos," *Procedia Technol.*, vol. 6, pp. 1004–1010, Jan. 2012.
[28] V. S. Rasmi and K. R. Vinothini, "Real time unusual event detection using video surveillance system for enhancing security," in *Proc. Online Int. Conf. Green Eng. Technol. (IC-GET)*, Nov. 2015, pp. 1–4.
[29] S. N. Jyothi and K. V. Vardhan, "Design and implementation of real time security surveillance system using IoT," in *Proc. Int. Conf. Commun. Electron. Syst. (ICCES)*, Oct. 2016, pp. 1–5.
[30] A. Rajan and V. P. Binu, "Enhancement and security in surveillance video system," in *Proc. Int. Conf. Next Gener. Intell. Syst. (ICNGIS)*, Sep. 2016, pp. 1–5.
[31] F. S. Samaria and A. C. Harter, "Parameterisation of a stochastic model for human face identification," in *Proc. IEEE Workshop Appl. Comput. Vis.*, Dec. 1994, pp. 138–142.
[32] R. Khokher, R. C. Singh, and R. Kumar, "Footprint recognition with principal component analysis and independent component analysis," *Macromol. Symposia*, vol. 347, no. 1, pp. 16–26, 2015.
[33] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, vol. 1, Jun. 2005, pp. 886–893.
[34] D. Meva and C. K. Kumbharana, "Identification of best suitable samples for training database for face recognition using principal component analysis with eigenface method," *Int. J. Comput. Appl.*, vol. 115, no. 12, pp. 24–26, 2015.
[35] X.-Y. Li and Z.-X. Lin, "Face recognition based on HOG and fast PCA algorithm," in *Proceedings of the Fourth Euro-China Conference on Intelligent Data Analysis and Applications*. Málaga, Spain: Springer, Oct. 2017.

MUHAMMAD AWAIS received the B.S. degree in computer science from National Textile University, Faisalabad, Pakistan, and the M.S. degree in computer science from the Computer Science Department, University of Engineering and Technology, Taxila, Pakistan. His research interests include image processing, real-time security, and surveillance system.



MUHAMMAD JAVED IQBAL received the M.Sc. degree in computer science from the University of Agriculture, Faisalabad, Pakistan, in 2001, the M.S./M.Phil. degrees in computer science from International Islamic University Islamabad, Pakistan, in 2008, and the Ph.D. degree in computer science/information technology from Universiti Teknologi PETRONAS, Malaysia, in February 2015. He is currently an HEC approved Ph.D. Supervisor and an Assistant Professor with the Computer Science Department, University of Engineering and Technology, Taxila, Pakistan. After completion of his doctoral studies, he has been actively involved in research. He has more than 15 international publications which includes four ISI indexed impact factor journals and one book chapter Springer LNEE. His research interests include machine learning, data science, pattern recognition, computational intelligence algorithms for biological data classification, bioinformatics, and big data mining.

IFTIKHAR AHMAD received the B.Sc. degree from Islamia University, Bahawalpur, Pakistan, in 1999, the M.Sc. degree in computer science from the University of Agriculture, Faisalabad, Pakistan, in 2001, the M.S./M.Phil. degrees in computer science from the COMSATS Institute of Information Technology, Abbottabad, Pakistan, in 2007, and the Ph.D. degree in information technology from Universiti Teknologi PETRONAS, Malaysia, in 2011. He has been a Faculty Member and a Research Supervisor with various universities, since 2001. He is currently a Faculty Member with the Information Technology Department, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah, Saudi Arabia. He has been involved in several funded projects as PI and Co-PI. He has published several articles in reputed journals and conferences. He is also a member of several scientific and professional bodies.

MADINI O. ALASSAFI received the B.S. degree in computer science from King Abdulaziz University, Saudi Arabia, in 2006, the M.S. degree in computer science from California Lutheran University, USA, in 2013, and the Ph.D. degree in security cloud computing from the University of Southampton, Southampton, U.K., in 2018. He is currently an Assistant Professor and the Chairman of the Information Technology Department, Faculty of Computing and Information Technology, King Abdulaziz University Jeddah, Saudi Arabia. He has published numerous conference papers, journal articles and book chapters. His research interests include cloud computing and security, distributed systems, the Internet of Things (IoT) security issues, cloud security adoption, risks, cloud migration project management, cloud of things and security threats.

RAYED ALGHAMDI received the master's and Ph.D. degrees in information and communication technology from Griffith University, Australia, in 2008 and 2014, respectively. He is currently an Assistance Professor with the Information Technology Department, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah, Saudi Arabia. He is also coordinates the research unit with the Faculty of Computing and Information Technology.

MOHAMMAD BASHERI received the Ph.D. degree from Durham University, U.K., in October 2013. He has received several trainings and certifications from well-reputed institutes and organizations. He is currently a Faculty Member and the Vice Dean of Development, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah, Saudi Arabia. He has published several articles in conferences and journals of international repute. He has served as a consultant for some institutes and organizations.

MUHAMMAD WAQAS received the bachelor's degree in electronics engineering from the Usman Institute of Technology, Hamdard University, Pakistan, in 2018. He was a recipient of fully funded scholarships in multiple semesters. His research interests include cyber security, mobile and wireless networks, and embedded systems.