DOI: 10.1111/phor.12426

ORIGINAL ARTICLE

A deep learning approach for facial emotions recognition using principal component analysis and neural network techniques

Mudassir Khan[1] | S. Hariharasitaraman[2] | Shubham Joshi[3] | Vishal Jain[4] | M. Ramanan[5] | A. SampathKumar[6] | Ahmed A. Elngar[7]

[1]Department of Computer Science, College of Science and Arts Tanumah, King Khalid University, Abha, Saudi Arabia

[2]School of Computing Science and Engineering, VIT Bhopal University, Madhya Pradesh, Bhopal, India

[3]Department of Computer Engineering, SVKM'S NMIMS MPSTME, Shirpur, Maharashtra, India

[4]Department of Computer Science and Engineering, School of Engineering and Technology, Sharda University, Uttar Pradesh, Greater Noida, India

[5]Department of Physical Sciences and IT, Agricultural Engineering College and Research Institute, TNAU, Coimbatore, India

[6]Department of Applied Cybernetics, Faculty of Science, University of Hradec Kralove, Hradec Kralove, Czech Republic

[7]Faculty of Computers and Artificial Intelligence, Beni-Suef University, Beni-Suef, Egypt

Correspondence
A. SampathKumar, Department of Applied Cybernetics, Faculty of Science, University of Hradec Kralove, Hradec Kralove, Czech.
Email: sampathkumar.arumugam@uhk.cz

Abstract

In this work, advanced facial emotions are recognized using Neural network-based (NN) PCA methodology. The earlier models are cannot detect facial emotions with moving conditions but the CCTV and other advanced applications are mostly depending on moving object-based emotion recognition. The blurring, mask, and moving object-based facial image are applied to the training process, and at the testing condition, real-time facial images are applied. The PCA is extracting features and pre-processing the images with NN deep learning process. The proposed facial emotion recognition model is most useful for advanced applications. The design is finally verified through confusion matrix computations and gets measures like accuracy 98.34%, sensitivity 98.34% Recall 97.34%, and F score 98.45%. These output results compete with present models and outperformance the methodology.

KEYWORDS

deep learning, facial emotion recognition, neural network, principal component analysis

Résumé

Dans ce travail, les émotions faciales sont reconnues à l'aide d'une méthodologie ACP basée sur un réseau neuronal (NN). Les modèles précédents ne pouvaient pas détecter les émotions faciales dans des conditions de mouvement, alors même que la vidéosurveillance et autres applications avancées dépendent principalement de la reconnaissance d'émotions basée sur des objets en mouvement. Le flou, les occultations et les images d´objets en mouvement sont pris en compte dans le processus d´entraînement, et en conditions de test, les images faciales en temps réel sont utilisées. L'ACP extrait les caractéristiques et pré-traite les images avec un processus d'apprentissage profond utilisant un réseau neuronal. Le modèle proposé pour la reconnaissance des émotions faciales est très utile pour des applications avancées. La conception est finalement vérifiée par le calcul de matrices de confusion, qui donne une précision de 98,34%, une sensibilité de 98,34%, un rappel de 97,34% et un score F de 98,45%. Ces résultats concurrencent les modèles actuels et surpassent la méthodologie.

Zusammenfassung

In dieser Arbeit werden fortgeschrittene Gesichtsemotionen mithilfe der neuronalen netzwerkbasierten (NN) PCA-Methodik erkannt. Die früheren Modelle können Gesichtsemotionen bei Bewegung nicht erkennen, aber CCTV und andere fortschrittliche Anwendungen sind hauptsächlich auf die Erkennung von Emotionen auf der Grundlage von sich bewegenden Objekten angewiesen. Die Unschärfemaske und das auf einem sich bewegenden Objekt basierende Gesichtsbild werden auf den Trainingsprozess angewendet, und unter der Testbedingung werden Echtzeit-Gesichtsbilder angewendet. Die PCA extrahiert Merkmale und verarbeitet die Bilder mit dem NN-Deep-Learning-Prozess vor. Das vorgeschlagene Gesichtsemotionserkennungsmodell ist am nützlichsten für fortgeschrittene Anwendungen. Das Design wird schließlich durch Konfusionsmatrixberechnungen verifiziert und erhält Maße wie Genauigkeit 98,34 %, Empfindlichkeit 98,34 % Recall 97,34 % und F-Measure 98,45 %. Diese Ausgabeergebnisse konkurrieren mit aktuellen Modellen und übertreffen die Methodik.

Resumen

En este trabajo, las emociones faciales se reconocen utilizando la metodología PCA basada en redes neuronales (NN). Los modelos anteriores no pueden detectar emociones faciales en condiciones de movimiento, a pesar de que CCTV y otras aplicaciones avanzadas dependen principalmente del reconocimiento de emociones basado en objetos en movimiento. El desenfoque, las oclusiones y las imágenes de objetos en movimiento se tienen en cuenta en el proceso de entrenamiento y, en las condiciones de prueba, se usan imágenes faciales en tiempo real. El PCA extrae características y preprocesa las imágenes con un proceso de aprendizaje profundo usando una red neuronal. El modelo de reconocimiento de emociones faciales propuesto es muy útil para aplicaciones avanzadas. El diseño finalmente se verifica mediante cálculos de matriz de confusión, obteniendo una precisión del 98,34 %, una sensibilidad del 98,34 %, una recuperación del 97,34 % y una puntuación F del 98,45 %. Estos resultados compiten con los modelos actuales y superan el rendimiento de la metodología.

摘要

这项工作使用了基于神经网络的 PCA方法来识别面部情绪的细节。早期的模型无法检测到移动条件下的面部情绪，然而闭路电视和某些其他应用却依赖对移动中的对象进行情绪识别。在本模型训练过程中，采用了模糊、遮罩和基于移动对象的面部图像，在测试条件下使用了实时面部图像。PCA 提取特征，并通过神经网络深度学习过程对图像进行预处理。所提出的面部情绪识别模型对高级应用是最可用的。该模型最终通过混淆矩阵的计算得到了验证，准确率为 98.34%、敏感度 98.34%、召回率 97.34% 和F值 98.45%。 这些结果可以与目前的模型相媲美。

INTRODUCTION

Facial recognition, as an interdisciplinary research topic, has a long history of research by the scientific community since the 1980s (Zhao et al., 2003). The importance of facial recognition is due to the many applications that can be encountered in different areas, such as security, monitoring devices in public spaces, forensic applications, identity of people in image or video databases, human–machine interface, smart cards, and the biometric passport also known as the E-Passport (Bowyer, 2004).

Motivation

The present work is a part of efforts currently being made in the scientific community to obtain satisfactory solutions to the problems presented by automatic facial recognition by using two-dimensional information (Figure 1).

Objectives

The main goal of the first step of system is the detection and localization of the facial area in a given image. After the detection step, the biometric signatures of human face are extracted as a characteristic vector within the second step. The last stage is to identify the people from the representations of their faces. Each person is photographed many times, and data about them are retrieved and entered into the database (offline). When an input face picture arrives, face detection, feature extraction and finally comparison of these features are then performed to each face class stored in the database (online phase).

FIGURE 1 Facial emotions.

LITERATURE SURVEY

Approaches for face recognition

Generally, a facial recognition system is characterized by its classifier that can be designed according to two types of approaches.

Global approaches

The peculiarity of appearance-based algorithms is the direct use of values for the pixels of the entire image of the face on which the recognition decision will be based. The disadvantage of this approach is the size of the data to be processed. Indeed, in these methods an image of size $n = p \times q$ pixels is represented by a vector of the same size, in a large space. In order to reduce the size of the starting data, several methods have been to transform the vector of the original data into another space, without eliminating the discriminatory information that will be used during the classification stage.

Local methods

Local methods, based on models, use a priori knowledge that one possesses on the morphology of the face and generally rely on points its characteristics. These methods constitute another approach for non-linearity by constructing a space of local characteristics and using appropriate image filters, so that the face distributions are less affected by various changes. Bayesian approaches (such as the Bayesian information criterion [BIC] method), support vector machines (SVM) (Galbally et al., 2014), the active model of appearance (AAM) method or the local method binary pattern (LBP) were used for this purpose.

Hybrid methods

Hybrid methods consolidate the advantages of global and local characteristics of facial image. They increase stability recognition performance during installation changes, lighting of facial expressions. The analysis of local characteristics Logical Framework Analysis (LFA) (Gavrilova et al., 2017) and the wavelet extracted features of Gabor (such as elastic bunch graph matching—EBGM) are hybrid algorithms Typical.

Applications of automatic face recognition

The need to automate the task of facial recognition is used by the security forces (police albums, ID, passport, driving licences, etc.) and also by new commercial applications. In addition to those mentioned above, in which it is important to control fraud, verifying the identity of clients, for example, when accessing bank ATMs, to areas of restricted access, to multimedia telecommunications, in video surveillance systems, etc.

General architecture of the face recognition system

There is another type of facial recognition scenario when checking a watch list, where an individual is compared with a list restricted number of suspects (Wickens, 2014). The basic operating principle of a recognition system can be summarized in three steps when a standard dataset is acquired for the recognition model, and is shown in Figure 2.

The first step in the facial recognition detects (or tracks) a person's face in an image or video, aligns the face and allows normalization of the detected face image, so that it can be compared with other face images in the dataset that do not necessarily have the same size or lighting.

Independent component analysis (ICA)

ICA is a statistical feature projection technique. It is the further extension of principal component analysis (PCA). ICA projects the image data from high to low dimension. The fast-ICA method finds a maximum of non-Gaussianity of whitened data in order to compute independent components (ICs) (Cohen et al., 2014).

The ICs are actually the basic building blocks for representing the two-dimensional images. ICA is based on this basic idea, which is illustrated in Figures 3 and 4. This fundamental idea behind ICA makes it useful in facial recognition. Till now various ICA algorithms have been introduced (Lu et al., 2011). Some use an adaptive scheme based on stochastic gradient methods; others find the ICs through the minimization or maximization of high-order cumulates.

PROPOSED METHOD FOR AUTOMATIC FACIAL RECOGNITION

According to Figure 5, after face image database acquisition pre-processing performs and face detection is achieved with the Viola–Jones method. A further detected face region is sampled for feature extraction, once

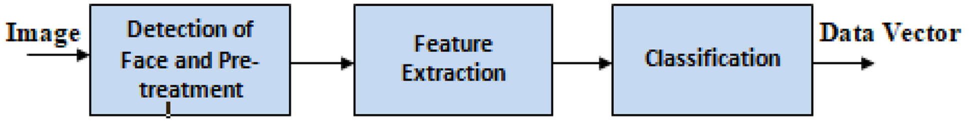


FIGURE 2 The process of recognizing a face.

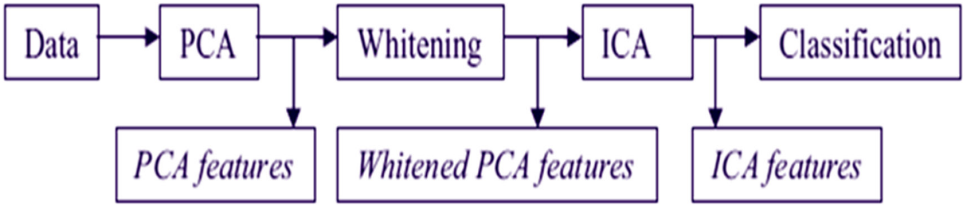


FIGURE 3 The ICA process.

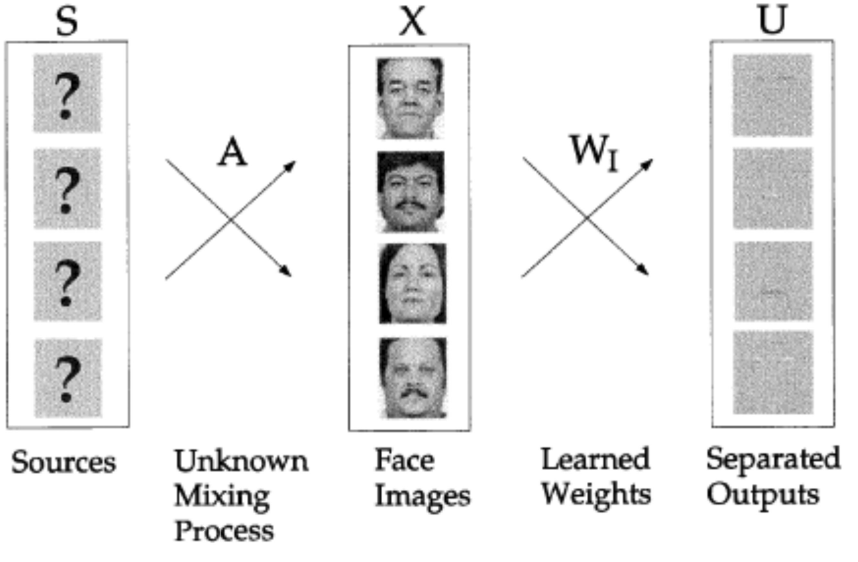


FIGURE 4 Fundamental idea of ICA for face recognition.

a feature is extracted with Discrete Wavelet Transforms (DWT) + PCA, Harris corner and Speeded-up Robust Features (SURF), executed separately; as an outcome there will be three IDs achieved, respectively. All three IDs are further fused to obtain the final ID.

Table 1 represents the PCA and ICA component for the ORL datasets. The effectiveness of PCA and ICA as feature extraction techniques in appearance-based object identification systems has recently been examined. The majority of the vital information about the faces has been recorded in the rebuilt photographs, while the unimportant aspects have been omitted. Table 2 represents recognition rate in comparison with different datasets with the proposed model. The accuracy obtained in the proposed model outperforms other methods.

- Face detection.

Face detection can be performed, as explained below:

- Pre-processing:

1. Take images I_1 and I_2 as an input.
2. Resize both images for a 160 × 160 matrix.
3. Apply RGB to grey value conversion.
4. Face detection using the Viola–Jones method.

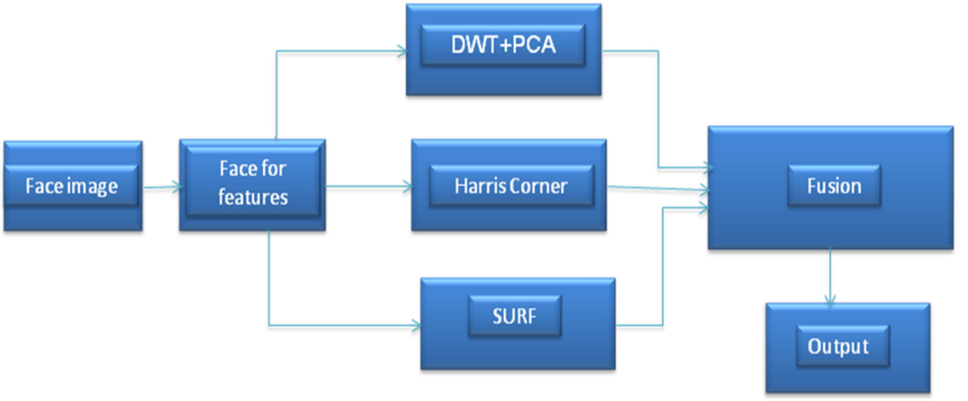
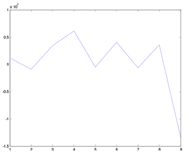
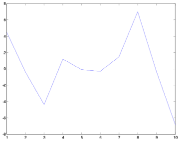
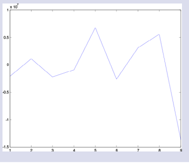
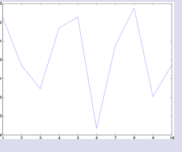
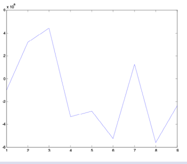
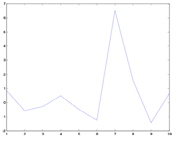
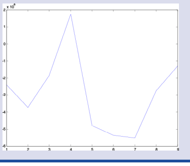
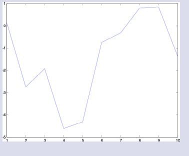


FIGURE 5 Block diagram of the face recognition system for Fusion.

TABLE 1 Principal component analysis (PCA) and independent component analysis (ICA) components for the ORL datasets.

Query image	Principal components	Independent components	Recognized image
			
			
			
			

Simulation analysis

While working on a specific approach, it is necessary for the researchers to work on a standard dataset in order to compare the results (Li et al., 2007). The following are some of the standard datasets used widely, and the results are simulated in MATLAB 7.8 for tabulation.

TABLE 2 Recognition rate comparison with different datasets with reference to the proposed method.

Methods	Dataset	Accuracy (%)
PCA + NN	ORL	88.86
ICA + NN	ORL	88.86
PCA + SVM	FERET	85.2
2DPCA + SVM	FERET	95.0
Common Eigen value	YALE B, FERET	97.5, 98.0
PCA	ORL	88.67
PCA	YALE	93.25
DWT	YALE	83.3
DWT	ORL	92.5
DWT + PCA	YALE	84.0
DWT-PCA-ICA-SVM	**ORL**	**88.7**
DWT-NN	ORL	91.6
PCA-NN	ORL	90.1
DWT-PCA-ICA-NN	**ORL**	**97.14**
DWT-PCA-ICA-SVM	**YALE**	**82.7**
DWT-NN	YALE	90.3
PCA-NN	YALE	92.1
DWT-PCA-ICA-NN	**YALE**	**96.23**

Bold represents the proposed method with different data sets.

Indian face dataset

The files in this database are in JPEG format. Each image is 640 × 480 pixels and a grey level of 256 per pixel (Tian et al., 2012). There are two main directories in which the image is saved accordingly, that is, males and females. Each directory contains details of the image, such as name, serial number, etc. Each directory possesses 11 different images per subject, with different looking orientations, that is, up, down, right, left; the stored emotions are smile, laughter, sad/disgust and neutral.

The training dataset contains front images of individuals from the Indian database and female and male datasets.

YALE face dataset

The files contain 165 greyscale images of 15 persons in GIF format. Each subject has 11 images, one for each facial expression: centre-light, with or without glasses, sad, sleepy, surprised, wink, etc. (Ho & Gopalan, 2014).

The training dataset contains front images of individuals from the YALE database. It contains a dataset of individuals with different expressions/emotions, such as sad, happy, smiling, wink, with or without glasses (Savvides et al., 2008).

ORL face database

The training dataset contains front images of individuals from the ORL database with different expressions/emotions such as sad, happy, smiling, wink, with or without glasses (Jin et al., 2001).

Database creation

The system model is designed such that it requires one image for training from each set and there are 40 sets of 10 images. Hence, the total number of training information is 40 (Table 3).

Let x_k be the set of images required for training, where x is the size of image ($m \times n$), m is row, n is column, and k is the number of sets depending on the selected dataset, that is, $k = 1, 2, 3, 4, \ldots, 40$.

The output received from the wavelet is the approximation coefficient of a Haar wavelet and will be the input of PCA. The extracted PCA will be the projected image feature of the image. If we say we are using 40 images for training, then we will be having 40 projected feature vectors for training. Now the rest of the images (i.e., 400 – 40 = 360) can be taken for testing as a test input (Li et al., 2015). The test input is processed through DWT/PCA and thus its projected vector is compared with the trained database.

Similarly, a measure is performed using Euclidean distance (D), which can be given by:

$$D = \sqrt{\sum_{i=0}^{n} (d_{1i} - d_{2i})^2}. \tag{1}$$

Based on this distance, accuracy is calculated. At the end the output of DWT + PCA is generated as the ID. In similar way each prototype (method) will have its unique ID as an outcome (Hafez et al., 2015). The resulted ID of each prototype may differ from each other. So it is difficult to choose which ID is accurate. Thus, we proposed a fusion-based approach for all IDs gathered by different prototypes. At last the fusion rule will decide the final outcome. For training purpose, only one single face is chosen. The test was performed using the same training set procedure.

HYBRID METHODS

Hybrid methods are approaches that use both global and local characteristics of a face image (Figure 6) (Lai et al., 2014). The performance of these methods includes the choice of the combination and how to combine them in such a way that their advantages are preserved and their disadvantages are avoided. These problems are similar to those multiple classifier systems or learning packages in the field of automatic learning. Local characteristics and overall characteristics will provide additional information relevant to the classification (Galbally et al., 2014).

Learning procedure for features

First, read the images from the learning database. These images are pre-processed, transformed into black and white and then decomposed by the Haar wavelet that provides four sub-images: an approximate image, a

TABLE 3 Database creation of face features.

1	2	3	–	–	–	–	–	–	10
2									
–									
–									
–									
–									
–									
40									

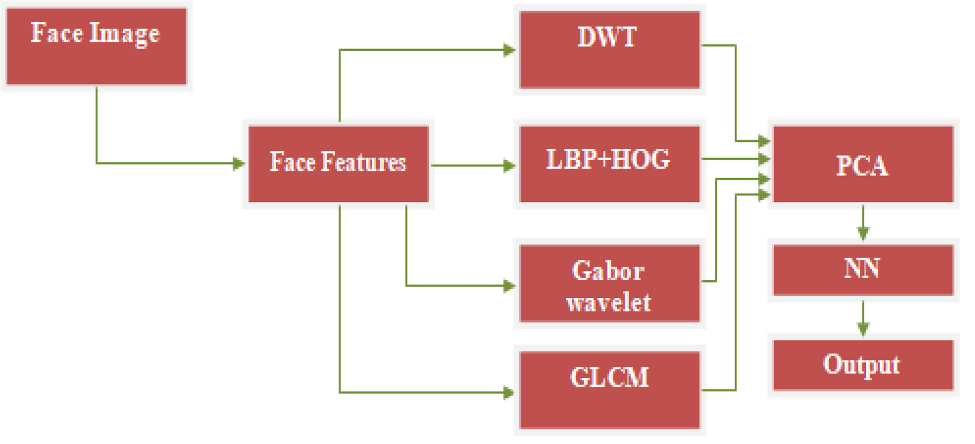


FIGURE 6 Block diagram of the Hybrid face recognition system.

horizontal detail, a vertical detail and a diagonal sized detail (33 × 33). The approximation of each image is taken for training (Gavrilova et al., 2017). Similarly, compute the LBP + Histogram of Oriented Gradient (HOG) of the image, a Gabor gradient of the image input image and a Gray Level Co-occurrence Matrix (GLCM) feature of the image and then covert it into a vector. Then put all the vectors of the images in a matrix called mat.

PCA is a mathematical method of data analysis that consists of searching the spatial directions, which represents the correlation between random variables. It is a linear orthogonal transformation that converts the information into a novel coordinate structure with the end goal. PCA is regularly used to minimize the size of the data by keeping only the main characteristics that contribute most to the overall variance and ignore the components of small variances. The retained components contain the most important information and the ignored components contain the least informative aspects of the data (Hron et al., 2016). Finally, a matrix of eigenvectors is obtained of size $C = 600$ and $L =$ contingent on the chosen size of the eigenvectors. The learning database presents 1000 images of 50 different people (20 different images for each person).

Projection on the PCA space

Project the images in a subspace of different dimensions from 1 to 1000. The image vector X_i is transformed into a vector Y_i by the relation:

$$Y_i = V^T \times X_i \tag{2}$$

$$V^T : (V_1, V_2, \ldots, V_i, V, \ldots, V_{1000}) \tag{3}$$

where V_i, $i = 1, \ldots, 1000$: are the eigenvectors.

Table 4 shows the results of the detection conducted on different subset of the YALE database which were categorized according to their illumination condition (Park et al., 2015). The average error rate was calculated as a performance evaluation parameter further, observation of average error rate according to Table 4 is as follows:

1. Generally, the use of subset 1 and 2 are the moderate lightening condition where images are clearer, and this leads to good performance for all the descriptors with the minimum average error rate. The subset

TABLE 4 Result of face detection conducted on different subset of the YALE database.

Method	Error rate versus illumination			Average error rate
	Subsets 1 and 2	Subset 3	Subset 4	
LBP	0.0	13.3	23.6	12.3
GLCM	1.2	34.3	45.3	26.93
Gabor features	0.0	0.0	12.1	4.03
DWT	0.0	13.0	34.5	15.83
Correlation (Ho & Gopalan, 2014)	0.0	23.3	73.6	32.3
Eigen face (Ho & Gopalan, 2014)	0.0	19.2	66.4	28.53
Liner subspace	0.0	0.0	15.0	5.0
Low-dimensional linear space	0.0	0.0	0.0	0.0
Proposed fused method	0.0	0.0	0.0	0.0

3 and subset 4 leads to produce poor performance for all the descriptors because of pose-variation and or illumination.

2. Proposed fused descriptor performs better than LBP by 12.3% and GLCM by 26.93% when taken average error rate tested on all subset. Somehow, the performance of GLCM, Correlation is poorer when compared with LBP, DWT, and Gabor descriptor.
3. Fused method shows the excellent robust behavior to illumination variations tested on sbuset 1, subset 2 and subset 3. This achieves a better-quality Average Error Rate by 0%.

The proposed fused descriptors are compared with other alternative descriptors running at comparatively lower computational cost, it is found that fused descriptors outperform the other descriptors except low dimensional linear subspace (Siddiqi et al., 2015). Though the low dimensional linear subspace descriptors and proposed fused method have exactly the same error rate but when considering pose variation, the proposed fused method is more efficient.

Face identification: ESSEX dataset

There are total 395 individuals in the database and for every individuals there are 20 images, hence total images are 7900 including male and female, the database are divided into four categories according to their variation of background and scale, versus extreme variation of expressions (Kirby & Sirovich, 1990).

Accuracy comparison with different learning rate

Table 5 shows the results of detection conducted on different subset of the Essex database which categorized according to their illumination condition (Yang et al., 2004). Average accuracy has been calculated as performance evaluation parameter further, observation of average accuracy according to Table 5 is as follows:

- Generally, the use of Faces 94 and 95 is moderate in background expression variation and where images are clearer, and this leads to good performance for all the descriptors with higher accuracy.

TABLE 5 Result of face detection conducted on different subset of the Essex database.

	Accuracy (%) at a 90% learning rate				
Method	Face 94	Face 95	Face 96	Grimace	Total
LBP	98.67	98.56	83.67	74.30	88.80
GLCM	93.56	92.45	63.45	62.50	77.99
Gabor features	98.56	99.01	85.34	77.98	90.225
DWT	97.56	97.67	85.67	74.56	88.91
Proposed fused method	99.56	99.80	98.60	97.40	98.84

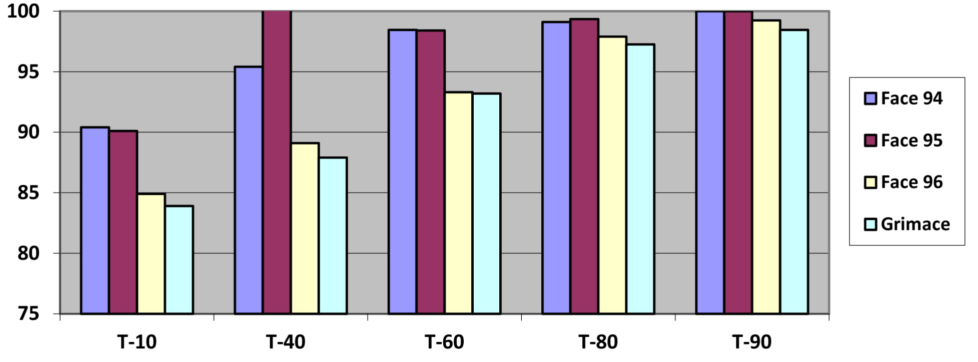


FIGURE 7 Accuracy in parentage according to different training rate for the Essex dataset.

- The proposed fused descriptor performs better than LBP by 10.0%, GLCM by 20.85%, Gabor features by 8.62% and DWT by 9.93% when taking the average accuracy tested on all types of dataset. Somehow, the performance of GLCM and LBP are poorer when compared with the DWT and Gabor descriptors.
- The fused method shows (Figure 6) excellent robust behaviour to illumination variations tested on Faces 94–96 and Grimace. This achieves a better quality average accuracy by 98.84% (Figure 7).

RESULTS

For training purposes, only a single face is chosen. The test is performed using the same training set procedure and evaluated based on the following:

$$\text{Accuracy} = \frac{\text{true positive} + \text{true negative}}{\text{true positive} + \text{true negative} + \text{false positive} + \text{false negative}} \tag{4}$$

$$\text{Precision} = \frac{\text{true positive}}{\text{true positive} + \text{false positive}} \tag{5}$$

$$\text{Sensitivity} = \frac{\text{true positive}}{\text{true positive} + \text{false negative}}. \tag{6}$$

The images used are of various combinations, that is, taken from different angles, varying the brightness, having different facial expressions for an individual, some images have occlusion, etc. The performance of the system using each algorithm is estimated individually based on sensitivity, precision and accuracy, and then the Hybrid algorithms are fused. The enhancement in system performance is represented in Figure 8. The recognition rate is calculated with the proposed hybrid method with a neural network (NN). This is represented in Table 6, which shows the greater recognition rate compared with the previous approach. The results are illustrated as follows.

Figure 9 represents the percentage accuracy against different nine orientations in the Essex database. The particular orientation single image was taken for training, and every orientation trained separately and tested for the entire dataset. Based on different orientation trained information, different accuracy is achieved. Figure 9 shows that there is less variation in accuracy according to the orientation. Hence, performance analysis claims better robustness against different orientations.

Figure 10 represents the percentage accuracy against five different illuminations (different light conditions) in the YALE B database. The particular illuminated single image was taken for training and tested for the entire dataset. Based on different illuminated trained information, different accuracy is achieved. Figure 10 claims that there is very little variation in accuracy according to different illumination. Hence, performance analysis claims better robustness against different lighting conditions.

The results are compared with the previous proposed methods (Table 7). Hybrid features give promising output, that is, 99.4% for face recognition. In fact, the Fusion-based accuracy, that is, 97.5%, provides results closer

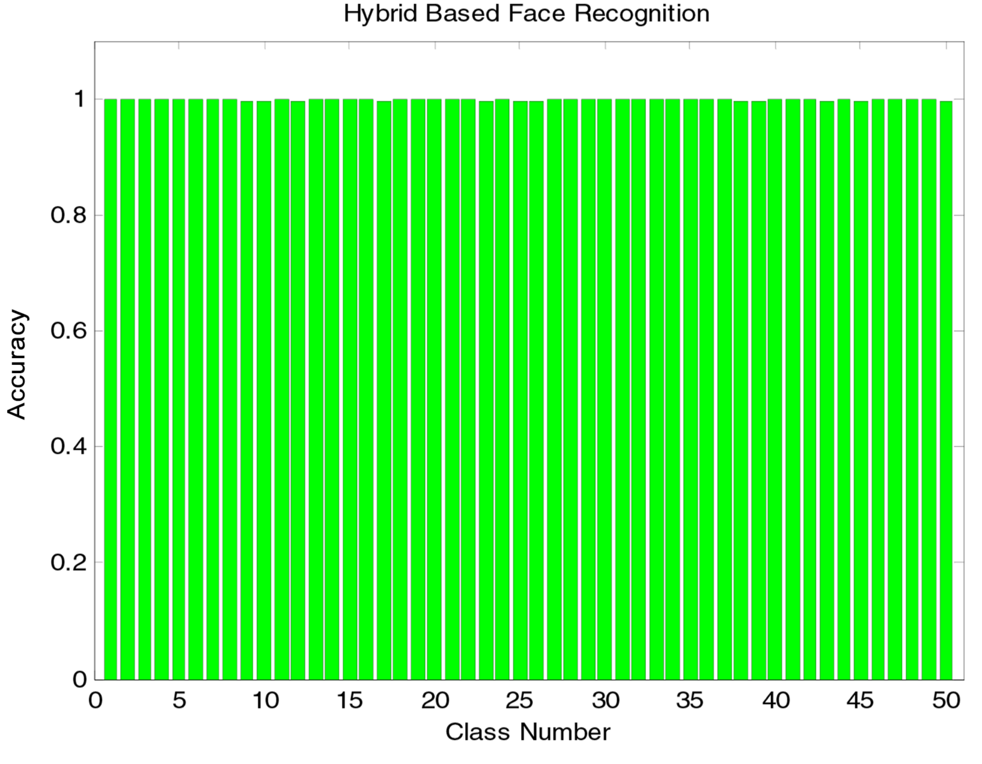


FIGURE 8 Performance of the Hybrid-based facial recognition for accuracy.

TABLE 6 Recognition rates of the proposed Hybrid method.

	Recognition rates for hybrid with PCA	Recognition rates for hybrid with NN
ORL database	97.50	99.45
FERET database	95.57	98.92
Essex database	96.99	99.40
YALE B database	95.23	99.28

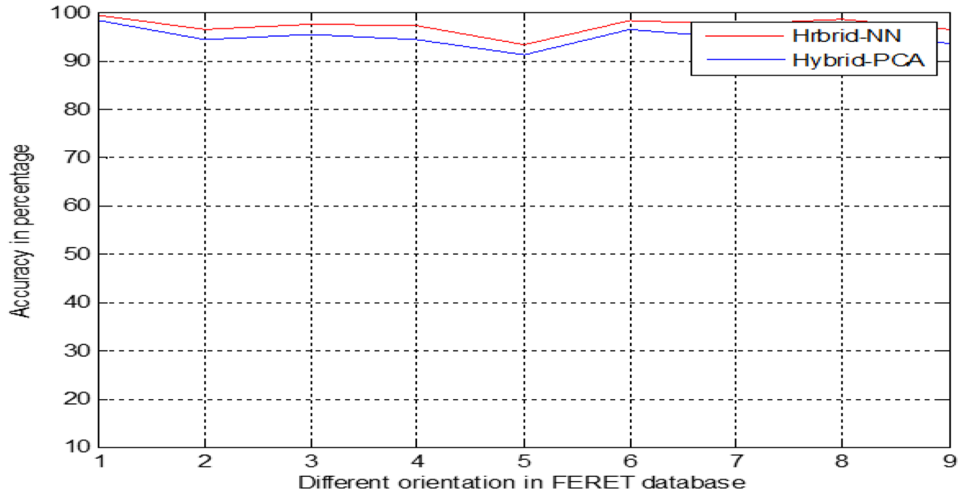


FIGURE 9 Accuracy performance of face recognition of the Essex dataset under different orientations.

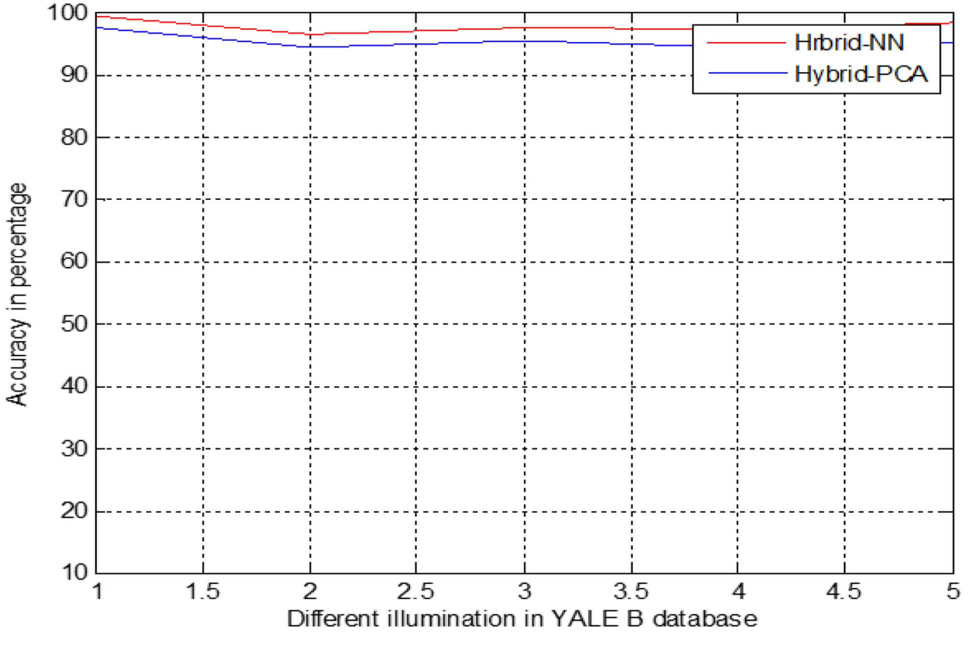


FIGURE 10 Accuracy performance of face recognition of the YALE B dataset under different orientations.

TABLE 7 Recognition rates for the ORL dataset.

DWT-PCA-ICA-NN (first model)	DWT-PCA-ICA-SVM (first model)	SURF (second model)	Harris Corner (second model)	DWT + PCA (second model)	Fusion (SURF, DWT-PCA, Harris corner) (second model)	Hybrid features + PCA (third model)	Hybrid features + PCA + NN (third model)
96.0%	90.0%	97.3%	98.2%	95.4%	98.9%	97.5%	99.4%

to the hybrid, but it is based on hit and trail. Hence, the overall performance of the proposed system outperforms the existing results.

The results are compared with the previous proposed methods (Table 8). Hybrid features give a promising output, that is, 99.45% for face recognition. In fact, the Fusion-based accuracy, that is, 99.3%, provides results closer to the hybrid, but it is based on hit and trail. Hence, the overall performance of the proposed system outperforms the existing results.

The results are compared with the previous proposed methods (Table 9). Hybrid features give a promising output, that is, 99.45% for face recognition. In fact, the Fusion-based accuracy, that is, 98.93%, provides results closer to the hybrid, but it is based on hit and trail. Hence, the overall performance of the proposed system outperforms the existing results.

The results are compared with the previous proposed methods (Table 10). Hybrid features give a promising output, that is, 98.95% for face recognition. In fact, the Fusion-based accuracy, that is, 97.13%, provides results closer to the hybrid, but it is based on hit and trail. Hence, the overall performance of the proposed system outperforms the existing results.

The results are compared with the previous proposed methods (Table 11). Hybrid features give promising output, that is, 99.28% for face recognition. In fact, the Fusion-based accuracy that is, 95.23%, provides results closer to the hybrid, but it is based on hit and trail. Hence, the overall performance of the proposed system outperforms the existing results.

TABLE 8 Recognition rates for the Indian dataset.

DWT-PCA-ICA-NN (first model)	DWT-PCA-ICA-SVM (first model)	SURF (second model)	Harris Corner (second model)	DWT + PCA (second model)	Fusion (SURF, DWT-PCA, Harris corner) (second model)	Hybrid features + PCA (third model)	Hybrid features + PCA + NN (third model)
98.78%	92.12%	99.10%	99.20%	98.99%	98.90%	99.30%	99.45%

TABLE 9 Recognition rates for the Essex dataset.

DWT-PCA-ICA-NN (first model)	DWT-PCA-ICA-SVM (first model)	SURF (second model)	Harris Corner (second model)	DWT + PCA (second model)	Fusion (SURF, DWT-PCA, Harris corner) (second model)	Hybrid features + PCA (third model)	Hybrid features + PCA + NN (third model)
98.78%	92.12%	99.10%	98.30%	97.40%	99.40%	98.93%	99.45%

TABLE 10 Recognition rates for the FERET dataset.

DWT-PCA-ICA-NN (first model)	DWT-PCA-ICA-SVM (first model)	SURF (second model)	Harris Corner (second model)	DWT + PCA (second model)	Fusion (SURF, DWT-PCA, Harris corner) (second model)	Hybrid features + PCA (third model)	Hybrid features + PCA + NN (third model)
94.78%	91.12%	95.10%	93.40%	94.40%	96.40%	97.13%	98.95%

TABLE 11 Recognition rates for the YALE dataset.

DWT-PCA-ICA-NN (first model)	DWT-PCA-ICA-SVM (first model)	SURF (second model)	Harris Corner (second model)	DWT + PCA (second model)	Fusion (SURF, DWT-PCA, Harris corner) (second model)	Hybrid features + PCA (third model)	Hybrid features + PCA + NN (third model)
95.34%	94.10%	96.10%	92.40%	93.60%	95.90%	95.23%	99.28%

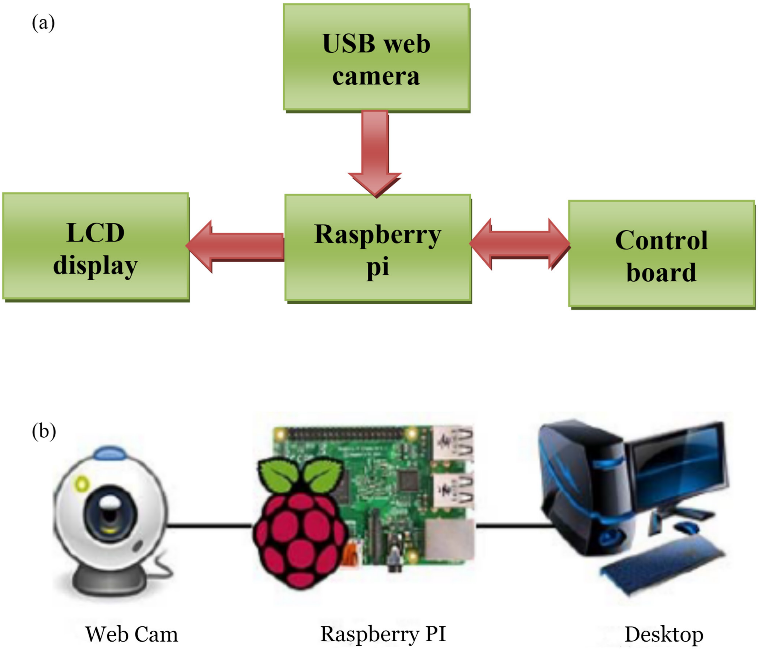


FIGURE 11 (a, b) The block diagram's proposed face recognition system.

Real-time face recognition system implementation

All algorithms are implemented using the OpenCv library, which is developed to provide assistance in building real-time systems for image processing applications. The library file has many built-in packages that provide assistance for the implementation of a real-time face recognition system and it performs face recognition operations separately with less processing time and high efficiency. The library file requires only a small amount of processing speed when incorporated with a Raspberry PI processor.

From Figure 11, it is understood that the output from the USB web camera is directly processed through Raspberry PI, which detects and recognizes the image of the face, and the resulting face is generated, which is displayed onscreen with that person's name.

CONCLUSIONS

The method of real-time face recognition system uses PCA and LBP algorithms, implemented on a Raspberry PI processor. The proposed work is more useful for society in various applications when face recognition plays a major role. In future work, different algorithms for feature extraction can be implemented to improve the accuracy of recognition.

CONFLICT OF INTEREST

The authors declare they have no conflict of interest.

DATA AVAILABILITY STATEMENT

All data are available in the paper.

CODE AVAILABILITY

All data available in the paper are custom mode.

INFORMED CONSENT

Informed consent was obtained from all individual participants included in the study.

ORCID

Mudassir Khan https://orcid.org/0000-0002-1117-7819
Shubham Joshi https://orcid.org/0000-0003-3815-1261
A. SampathKumar https://orcid.org/0000-0001-5318-5676
Ahmed A. Elngar https://orcid.org/0000-0001-6124-7152

REFERENCES

Bowyer, K.W. (2004) Face recognition technology: security versus privacy. *IEEE Technology and Society Magazine*, 23(1), 9–19. Available from: https://doi.org/10.1109/MTAS.2004.1273467

Cohen, P., West, S.G. & Aiken, L.S. (2014) *Applied multiple regression/correlation analysis for the behavioral sciences*. New York: Psychology Press. Available from: https://doi.org/10.4324/9781410606266

Galbally, J., Marcel, S. & Fierrez, J. (2014) Image quality assessment for fake biometric detection: Application to iris, fingerprint, and face recognition. *IEEE Transactions on Image Processing*, 23(2), 710–724. Available from: https://doi.org/10.1109/TIP.2013.2292332

Gavrilova, M.L., Ahmed, F., Azam, S., Paul, P.P., Rahman, W., Sultana, M. et al. (2017) Emerging trends in security system design using the concept of social behavioural biometrics. In: Alsmadi, I., Karabatis, G. & Aleroud, A. (Eds.) *Information fusion for cyber-security analytics*. Studies in Computational Intelligence, vol. 691. Cham: Springer, pp. 229–251. Available from: https://doi.org/10.1007/978-3-319-44257-0_10

Hafez, S.F., Selim, M.M. & Zayed, H.H. (2015) 2D face recognition system based on selected Gabor filters and linear discriminant analysis lda. *arXiv preprint*. arXiv:1503.03741. Available from: https://doi.org/10.48550/arXiv.1503.03741

Ho, H.T. & Gopalan, R. (2014) Model-driven domain adaptation on product manifolds for unconstrained face recognition. *International Journal of Computer Vision*, 109(1–2), 110–125. Available from: https://doi.org/10.1007/s11263-014-0720-x

Hron, K., Menafoglio, A., Templ, M., Hrůzová, K. & Filzmoser, P. (2016) Simplicial principal component analysis for density functions in Bayes spaces. *Computational Statistics & Data Analysis*, 94, 330–350. Available from: https://doi.org/10.1016/j.csda.2015.07.007

Jin, Z., Yang, J.Y., Hu, Z.S. & Lou, Z. (2001) Face recognition based on the uncorrelated discriminant transformation. *Pattern Recognition*, 34(7), 1405–1416. Available from: https://doi.org/10.1016/S0031-3203(00)00084-4

Kirby, M. & Sirovich, L. (1990) Application of the Karhunen–Loeve procedure for the characterization of human faces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(1), 103–108. Available from: https://doi.org/10.1109/34.41390

Lai, Z., Xu, Y., Chen, Q., Yang, J. & Zhang, D. (2014) Multilinear sparse principal component analysis. *IEEE Transactions on Neural Networks and Learning Systems*, 25(10), 1942–1950. Available from: https://doi.org/10.1109/TNNLS.2013.2297381

Li, H., Huang, D., Morvan, J.M., Wang, Y. & Chen, L. (2015) Towards 3D face recognition in the real: a registration-free approach using fine-grained matching of 3D keypoint descriptors. *International Journal of Computer Vision*, 113(2), 128–142. Available from: https://doi.org/10.1007/s11263-014-0785-6

Li, X., Hu, W., Zhang, Z., Zhang, X. & Luo, G. (2007, October) Robust visual tracking based on incremental tensor subspace learning. In: *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference*. Rio de Janeiro: IEEE, pp. 1–8. Available from: https://doi.org/10.1109/ICCV.2007.4408950

Lu, H., Plataniotis, K.N. & Venetsanopoulos, A.N. (2011) A survey of multilinear subspace learning for tensor data. *Pattern Recognition*, 44(7), 1540–1551. Available from: https://doi.org/10.1016/j.patcog.2011.01.004

Park, M., Jitkrittum, W., Qamar, A., Szabó, Z., Buesing, L. & Sahani, M. (2015) Bayesian manifold learning: the locally linear latent variable model (LL-LVM). *Advances in Neural Information Processing Systems*, 1, 154–162. Available from: https://doi.org/10.5555/2969239.2969257

Savvides, M., Heo, J. & Park, S.W. (2008) Face recognition. In: Jain, A.K., Flynn, P. & Ross, A.A. (Eds.) *Handbook of biometrics*. Boston, MA: Springer, pp. 43–70. Available from: https://doi.org/10.1007/978-0-85729-932-1

Siddiqi, M.H., Ali, R., Khan, A.M., Park, Y.T. & Lee, S. (2015) Human facial expression recognition using stepwise linear discriminant analysis and hidden conditional random fields. *IEEE Transactions on Image Processing*, 24(4), 1386–1398. Available from: https://doi.org/10.1109/TIP.2015.2405346

Tian, C., Fan, G., Gao, X. & Tian, Q. (2012) Multiview face recognition: from tensorface to v-tensorface and k-tensorface. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 42(2), 320–333. Available from: https://doi.org/10.1109/TSMCB.2011.2169452

Wickens, T.D. (2014) *The geometry of multivariate statistics*. New York: Psychology Press. Available from: https://doi.org/10.4324/9781315806334

Yang, J., Zhang, D., Frangi, A.F. & Yang, J.Y. (2004) Two-dimensional PCA: a new approach to appearance-based face representation and recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(1), 131–137. Available from: https://doi.org/10.1109/TPAMI.2004.1261097

Zhao, W., Chellappa, R., Phillips, P.J. & Rosenfeld, A. (2003) Face recognition: A literature survey. *ACM Computing Surveys (CSUR)*, 35(4), 399–458. Available from: https://doi.org/10.1145/954339.954342

How to cite this article: Khan, M., Hariharasitaraman, S., Joshi, S., Jain, V., Ramanan, M., SampathKumar, A. et al. (2022) A deep learning approach for facial emotions recognition using principal component analysis and neural network techniques. *The Photogrammetric Record*, 00, 1–18. Available from: https://doi.org/10.1111/phor.12426