An efficient framework using visual recognition for IoT based smart city surveillance

Manish Kumar[1] · Kota Solomon Raju[2] · Dinesh Kumar[1] · Nitin Goyal[3] · Sahil Verma[4] ·
Aman Singh[5] (ID)

Abstract

Smart city surveillance systems are the battery operated light weight Internet of Things (IoT) devices. In such devices, automatic face recognition requires a low powered memory efficient visual computing system. For these real time applications in smart cities, efficient visual recognition systems are need of the hour. In this manuscript, efficient fast subspace decomposition over Chi Square transformation is proposed for IoT based on smart city surveillance systems. The proposed technique extracts the features for visual recognition using local binary pattern histogram. The redundant features are discarded by applying the fast subspace decomposition over the Gaussian distributed Local Binary Pattern (LBP) features. This redundancy is major contributor to memory and time consumption for battery based surveillance systems. The proposed technique is suitable for all visual recognition applications deployed in IoT based surveillance devices due to higher dimension reduction. The validation of proposed technique is proved on the basis of well-known databases. The technique shows significant results for all databases when implemented on Raspberry Pi. A comparison of the proposed technique with already existing/ reported techniques for the similar applications has been provided. Least error rate is achieved by the proposed technique with maximum feature reduction in minimum time for all the standard databases. Therefore, the proposed algorithm is useful for real time visual recognition for smart city surveillance.

Keywords Fast subspace decomposition · Local binary pattern · Feature reduction · Minimum error rate

✉ Aman Singh
  amansingh.x@gmail.com

Extended author information available on the last page of the article

1 Introduction

The real-time implementation of a computer vision system on IoT based surveillance system is the need of the hour for contemporary society. Pattern recognition is one of the ground breaking recognition techniques to serve major applications such as biometric security, forensic investigation, Quick Response (QR) code and smart door locking systems, etc. [19]. The major challenges for developing a feature recognition system based on IoT applications are computational efficiency, accuracy, power consumption, and portability. There are a lot of existing techniques for pattern recognition are local binary pattern (LBP) and its variants, principle component analysis (PCA), and linear discrimination analysis (LDA) developed by the various researcher. Among these techniques, the local binary pattern technique is the most popular, investigated, and scrutinized due to its quality features such as tolerance against illumination changes, ease of implementation, computational simplicity, and fast response [31]. LBP, along with its variants, are investigated by the authors for classification problems [2]. The face image is divided into 8×8 or 16×16 regions, and then LBP feature distributions are extracted. The histogram of such features is computed region-wise, and a global concatenated histogram is used as a face descriptor. The performance of the proposed method is evaluated under different challenges [14]. The main idea of the EVBP descriptor is based on Virtual Electric Field (VEF). Authors combined Local Binary Pattern (LBP) based on the VEF. The neighbourhood of each pixel is assumed as a grid of virtual electric charges that are electrostatically balanced. The LBP concept is applied to the neighbourhood to generate the EVBP based representation of the face. This representation is computed for all four directions using the corresponding four electrical interactions [9]. A novel face feature extraction approach based on LBP and Two Dimensional Locality Preserving Projections (2DLPP) is explored. This approach aims to enhance the texture features without disturbing the space structure properties of a face image. LBP nullifies the variation in illumination and noise due to which the detailed texture characteristics of face images are enhanced. 2DLPP is performed to keep prominent features and decrease the feature size. In the proposed mechanism, the Nearest Neighbourhood Classifier (NNC) is used to classify the faces [45]. A new approach named the Two Directional Multi-level Threshold-LBP Fusion (2D–MTLBP-F) is proposed to illuminate invariant face recognition. The Threshold Local Binary Pattern (TLBP), combined with the Discrete Cosine Transform (DCT), is investigated. The LBP with different thresholds and neighbourhoods can be used to generate information. This information can be used to enhance the recognition rate. In the proposed method face images are normalized using DCT normalization technique, the resultant images are transformed into 61 levels of TLBP with different thresholds, and then the normalized DCT image is fused into these TLBP layers and face recognition is performed using the sparse sensing classifier (SRC) [3]. A novel technique called Weber Local Binary Image Cosine Transform (WLBI-CT) merges the frequency components of images obtained through Weber local descriptor and local binary descriptor in frequency domain [15]. These frequency components are invariant to multi-scale and multi-orientation facial images for facial expressions. Selection of significant and prominent feature set is the key to highly accurate face recognition, texture classification [16, 18] and scene classification [35, 42]. Despite exotic properties and applications of LBP, its extracted features are very sensitive to the image noise. In any image small variations may drastically

modify the LBP features [22]. The number of LBP codes occurs very significantly thereby infrequent features are difficult to measure from the particular histogram bin and compact features of the image are difficult to calculate and become almost incomprehensible for the system.

Further, Uniform LBP is used for the dimensions reduction [12]. For the binary codes contain less than three transitions from 1 to 0 and vice versa are called uniform patterns. It has been observed that the uniform patterns are less than 90% of total patterns for (8, 1) neighbourhood and almost 70% for (16, 2) neighbourhood but still the further reduction of the dimensions of the image poses a serious challenge. To target the issue of significant dimensional reduction of LBP descriptor, many subspace approaches are reported in literature [25]. The Principal Component Analysis (PCA) approach is reported to remove co-occurrence features [11]. Still, PCA is hypersensitive to the noise, and its suitability is restricted to the small data analysis, and the recognition rate remains insufficient [4]. LDA is another useful method to reduce the feature dimensions but its computational complexity and rotation variant approach limits its uses [28]. In order to address computational complexity and for computational load reduction other prominent techniques reported are Power Method [33], QR factorization [21] and subspace iteration methods [8] but these approaches suffer from slow convergence under situations, such as low signal to noise ratio and unknown subspace dimensions.

The algorithms and their variants as shown in Table 1 achieve optimum accuracy. For IoT based surveillance systems these methods are too complex. The complexity is in terms of computation time and run time memory requirements. Multimodal biometric identification approach is proposed for human verification based on voice and face recognition fusion, for the surveillance system voice recognition module is difficult to implement [1]. Reducing the effect of noise due to the illumination on face database has been proposed for face recognition [6].Hybrid feature extraction (HFE) technique is proposed for overcome the anti ageing effect of face recognition. Results of the algorithm is proved on different database but the complexity in terms of training and testing is not suitable for IoT based fast recognition systems [30]. A multi-feature fusion framework is proposed in literature with Gabor and deep feature for small sample face recognition, accuracy and performance of algorithm is up to the mark still extraction of feature process is lengthy and time consuming [46]. Also, there are various deep learning algorithms involving neural networks like Convolution Neural Network (CNN) and its variants [13, 29] which were explored by various researchers in past few years. These algorithms are computationally expensive and require specialized hardware like GPU's for their development and deployment [32]. These algorithms are not suitable for IoT applications in real time where limited resources and power are available. Thus, there is a requirement of an approach which can address the above mentioned issues efficiently and is also suitable for feature dimensionality reduction in a short span of time.

Therefore, a low power and less computational method need to be evolved for IoT based surveillance systems. Some authors proposed fast space decomposition [37, 38] to perform feature dimensional reduction. It is useful for optimum compact feature extraction. Face and texture datasets can be used for the validation of these methods. Recently, many authors are using Raspberry Pi board for IoT based surveillance systems as it is available at an affordable price for prototyping of the systems. The major contribution in the paper are:

Table 1 Comparison of different algorithms with respect to database and real time implementation

Technique Used	Accuracy	Databases Tested	Real Time	Complexity
Optimal [19] coefficients vector and sparse representation.	Recognition accuracy for face database ORL is 93%, for Facial Recognition Technology (FERET) is 79%, for Extended Yale face database is 83% achieved	ORL, FERET, Extended Yale B, Carnegie Mellon University (CMU), Pose, Illumination and Expression (PIE) and Aleix and Robert (AR) databases.	No	High
Local [31] ternary patterns (LTP)	achieving a face verification rate of 88.1% at 0.1% false accept rate	Extended Yale-B, Chinese Academy of Sciences-Pose, Expression, Accessories, and Lighting (CAS-PEAL-R_1) and Face Recognition Grand Challenge (FRGC)-204 data set.	Yes	Low
Local Binary [2] Pattern	Accuracy is more than 95%	FERET Database	No	Low
Noise-resistant [22] LBP (NRLBP), Extended noise-resistant LBP (ENRLBP)	Optimum accuracy is achieved.	AR Database extended Yale dataset, O2FN Dataset. With noise	No	Medium
Chi-squared [25] transformation (CST) with asymmetric principal component analysis (APCA) of LBP Histogram.	Error rate reduced to 0.38% for AR dataset. 1.5% for O2FN 9.5% for LFW 8.6% for 2D Hela	AR Database, LFW dataset, O2FN Dataset. 2D hela database also texture database	No	Medium
Karhunen–Loeve transform and SVM [11]	Accuracy achieved 97.5%, 97.5%, 91.5% and 90.6%	ORL database	No	High
Haar wavelet [4] transform (HWT) and Principal Component Analysis (PCA) using Levenberg-Marquardt backpropagation (LMBP) neural network	Accuracy achieved up to optimum level.	Faces 94 Directory dataset	No	High
Compound dictionary learning based classification (CDLBC) [17]	Accuracy achieved with number of the training samples of each Subject is 90.14%, 65.4%, 63.88% and 37.01% respectively	ORL, FERET, Georgia Tech (GT), LFW database,	No	High
Linear representation-based classification method using Fisher discriminant [28] features (LRFC) with linear regression (LR) Algorithm	Accuracy achieved for ORL database is 85%, for XM2VTS face database 95%	ORL, AR and XM2VTS (Multi Modal Verification for Teleservices and Security applications) face databases.	No	High

1. An efficient framework is proposed for surveillance systems for smart cities using IoT devices. The proposed framework works well for real time applications in industries for employee identification and surveillance systems in smart cities.
2. Efficient fast subspace decomposition over Chi Square transformation is proposed. This transformation has yield better recognition rates over various datasets.

3. Least error rate is achieved by the proposed technique for AR database, LFW, O2FN and Daynmic texture database. Further higher TAR is achieved using proposed method as compared to existing recognition techniques for varying values of FAR.

The paper is divided into four sections. The first section has already introduced the research problem and presented the literature survey. The second section of the paper explores and explains the algorithm of real time face recognition including face detection, face image enhancement, features extraction, dimensional reduction, face classification and other recognition applications. The penultimate section of the paper presents the experimental setup, results, and their discussion; and the last section concludes the paper.

2 Methodology

The proposed algorithm is elucidated in Fig. 1. Face detection; alignment and enhancement are achieved by using standard algorithms available in OpenCV library. The face image is normalized as preprocessing step to convert the complete dataset in the common range. Normalization process helps to handle various image datasets having different size and format of face images. Face identification is accomplished by extracting the required features using modified LBP. The LBP features are further reduced using proposed technique and the reduced features are stored along with Face ID. The stored features are uploaded on cloud or on IoT device for real time deployment. Once system is deployed and any face is detected
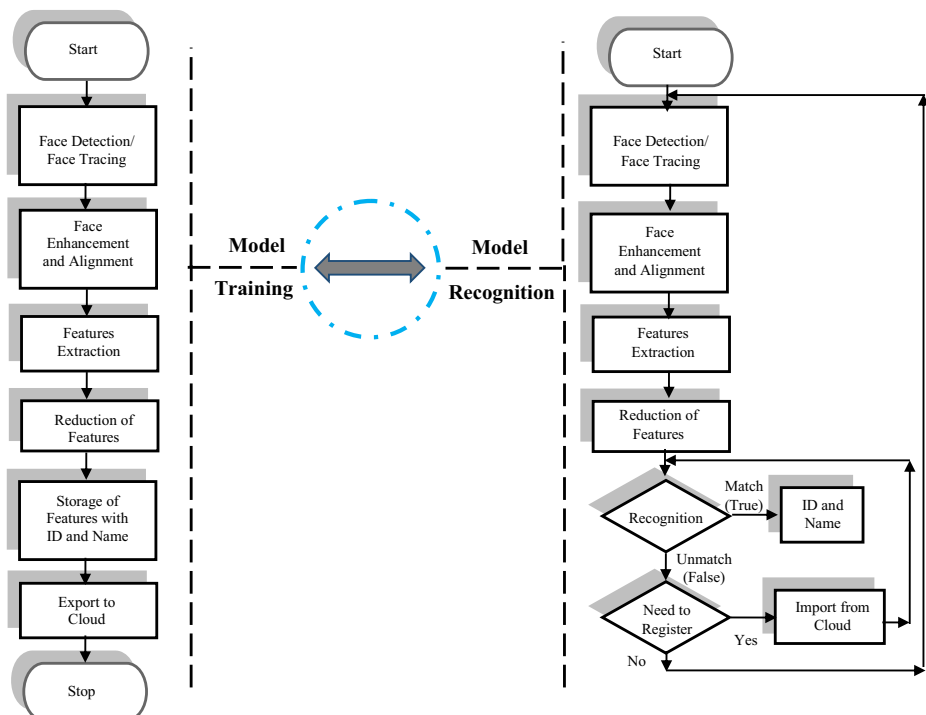

Fig. 1 Block of real time vision system on Raspberry Pi board

by the camera the features are extracted and reduced in real time for recognition. If the face is recognized from the database, the name and identity are displayed otherwise the face may be registered in the database for future recognition. The authors tested this system on raspberry Pi board but this can be further extended to other IoT devices.

Pseudo code of the proposed algorithm is as follows:

Training Phase:

Step 1: Normalize the dataset to convert the dataset by substracting the mean image and divide by variance. Also divide the image into 8×8 bins for computation of Local Binary Patterns

Step 2: Apply Local Binary Pattern of each image using uniform LBP.

Step 3: Create histogram of each bin and concatenate the histograms to get global histogram

Step 4: Apply Chi Square transformation of the resultant image to achieve Gaussian distribution

Step 5: Store the transformed data into the requisite format into the csv/xml file.

Testing/Deployment Phase:

Step 1: Read the stored csv/xml file.

Step 2: Input image from Camera/DataSet for recognition

Step 3: Apply LBP, Chi Square Transformation as per training on this single image.

Step 4:Compute Chi Square Distance of test image to all the images in the dataset and Classify the image to the minimum error class.

Testing/Deployment can be done on a local Machine or on IoT devices. In case of IoT device csv/xml file and trained model need to be deployed on IoT device.

2.1 Optimize computational efficiency

The proposed Chi-square transformed fast subspace LBP algorithm described as: Initially, uniform $LBP_{8,2}^{u2}$ is extracted the feature of a query image where subscript 8,2 represents eight neighbours at a distance of 2. Superscript u2 stands for using codes for uniform patterns and one code for all other patterns. The central pixel denoted as (x_c, y_c), P denotes the

$$LBP_{P,R}(x_c, y_c) = \sum_{P=0}^{P-1} S(i_c - i_p) 2^P \tag{1}$$

sampling points on a circle with radius R, i_c and i_P denotes, gray-scale values of the central pixel respectively [3]. Thresholding function $S(a)$ may be defined as

$$S(a) = \left\{ \begin{array}{l} 1, if\ a \geq 0 \\ 0\ otherwise \end{array} \right\} \tag{2}$$

$$h_{LBP} = \sum_a S(f_{LBP}(a)), P \tag{3}$$

The h_{LBP} is features histogram calculated by the standard LBP algorithm. Further, Chi-Square transformation is performed to make the distribution of the PDF of LBP as Gaussian thereby optimum usage of extracted LBP features is achieved.

This Chi square transformation is performed by taking two samples of LBP features denoted as 'a' and 'b'. These samples further introduce another feature vector $x = \{x_1, x_2, x_3 \ldots \ldots x_d\}$ where each element of x_i is represented as:

$$x_i = \frac{a_i - b_i}{\sqrt{a_i + b_i}} \tag{4}$$

To evaluate the Chi squared distance the normalization of 'x' is performed as:

$$x^T x = \sum_{i=0}^{d} \frac{(a_i - b_i)^2}{a_i - b_i} \tag{5}$$

Now, the fast sub space decomposition is applied on the input LBP feature 'x' for the dimensional reduction as:

$$x(t) = A(\theta) s(t) + n(t) \tag{6}$$

Where $s(t)$ is the LBP feature histogram, $A(\theta)$ is the subspace span with the dimension 'd', $n(t)$ represents additive noise and $x(t)$ is the array output observed at time $''t' = 1, \ldots, N$. In order to remove the unreliable features, the co-variance matrices of signal $x(t)$ is calculated as:

$$W_x = E\{x(t)x^H(t)\} = A(\theta) W_s A^H(\theta) \tag{7}$$

Where W_s is the co-variance matrix of the signal and the decomposition of W_x signal for the finite number of features 'N' (say) can be written as:

$$\widehat{W}_x = \frac{1}{N} \sum_{t=1}^{N} x(t) x^H(t) \tag{8}$$

$\widehat{W}_x$ is the signal subspace and its dimension is calculated from the 'd' eigenvectors $\{e_1, \ldots, e_d\}$ of $\widehat{W}_x$. Now the task is to calculate the optimal value of 'd' so that non-repeated feature of the LBP histogram can be extracted. The length of 'd' can be evaluated using the non-repeated Eigen values of the co-variance matrices [37]. This optimal length of 'd' is calculated by taking new statistics in consideration as reported in [37]. The extracted features of the trained data set are reduced by the signal subspace vector. Thereafter, the trained data set is stored in the system memory and the signal subspace vector extracts the optimal features for all the testing samples.

2.2 Recognition

The proposed approach reduces features of histogram of the trained data set stored in the system memory. For the recognition of the given query image, the reduced feature histogram can be computed. Thereafter minimum distance of the features is calculated by Chi square distance.

Various feature similarity approximation techniques between the test image features histogram and stored trained image feature histogram such as log-likelihood, Euclidean distance,
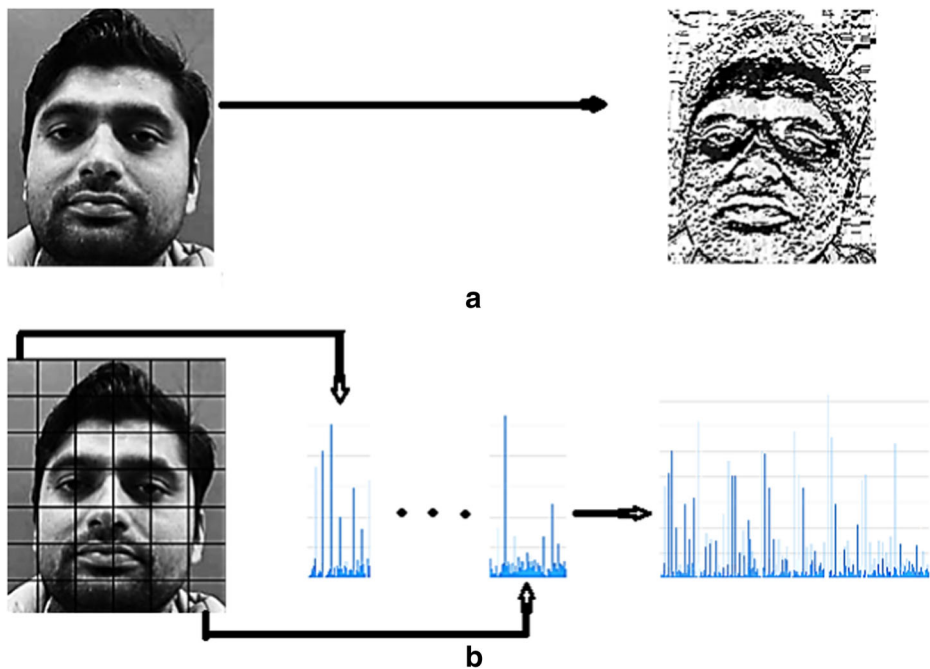
Fig. 2 **a** Extracted feature image, **b** Histogram vector

histogram intersection technique and Chi square distance are probed. In the proposed work, Chi square distance calculation is used for recognition. Further authors substantiate that after applying weights to the unique features on the image gives better results in terms of accuracy and time complexity. The extracted feature image and the histogram vector is shown in Fig. 2.

2.3 Proposed architecture for IoT applications

The proposed architecture is shown in Fig. 3. The High end server in the architecture is used to store dataset, module training. The trained model and the computed features are then stored in the common dataset. This common dataset is either on cloud or inside IoT device memory. The IoT Gadget is used to deploy the model in real time. The gadget is also connected to cctv/ web camera for real time input. The trained model inside the gadget will work as Identfication module for all applications like Employee Identification use Face, Security Surveillance in Industries and Security device for Vehicles. The computed decision can further be communicated to mobile device for further actions.

3 Results and discussions

The experiments are performed and validated on desktop and Raspberry Pi. The desktop machine is used with Octa core i5 processor of 2.7 GHz, 4 GB-DDR3 RAM, and Linux (Ubuntu 16.04) operating system using OpenCV (version 3.2.0). Raspberry Pi board having Quad core 1.2 GHz Broadcom BCM2837 64 bit CPU, 1 GB RAM with 8GB memory card. The proposed approach is validated on four different databases and the cross-validation
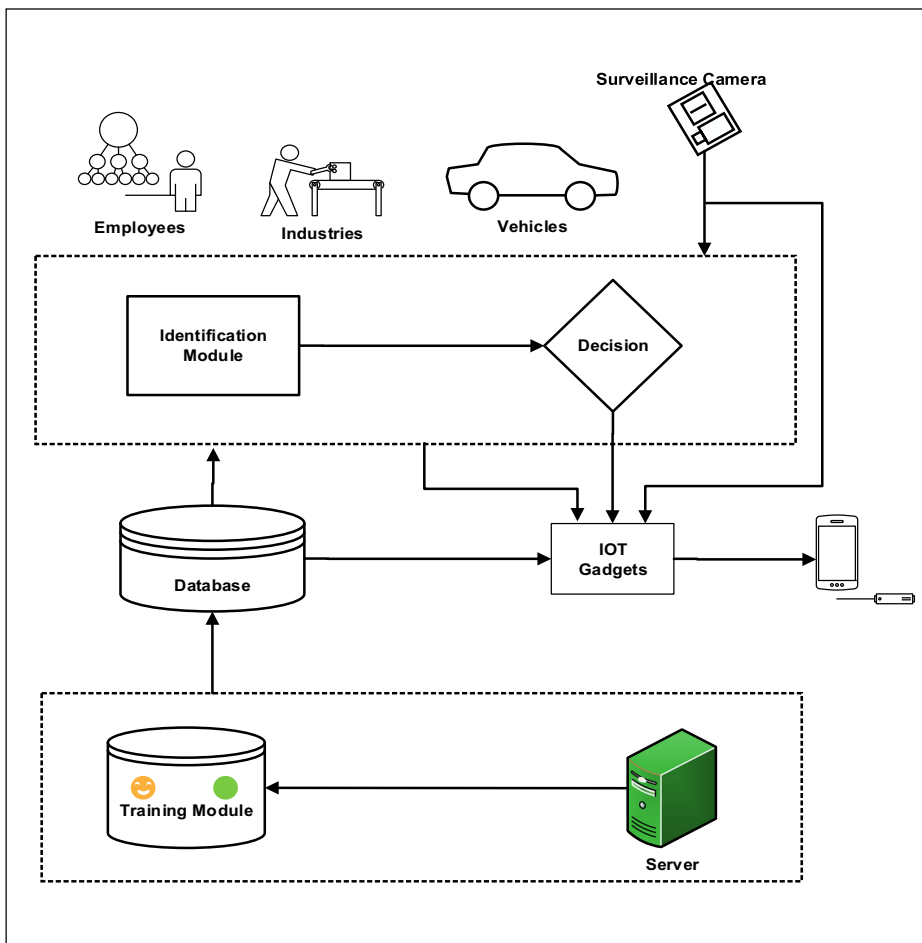
Fig. 3 Proposed architecture

technique gives the performance of the algorithm. The analysis and comparison of proposed algorithm is mentioned below for each dataset.

(a)   Analysis of AR dataset for face recognition

The AR database [5] has 4000 images of 126 different faces in which 56 female and 70 male faces are included and few samples are shown in Fig. 4. The images are normalized to 150×130 pixel and further divided into 8×8 blocks. The extracted features of LBP are $59 \times 8 \times 8 = 3776$ dimension. The proposed algorithm has been implemented on features for further reduction. Comparison of the proposed algorithm with existing approaches is shown in Table 2. It is observed that the performance of proposed algorithm is better than the existing approaches except PmSVM-Chi2 and PmSVM-HI as both of these are error free on the given dataset.

(b)   Analysis of O2FN Mobile dataset for face recognition

Fig. 4 AR dataset [5]

The O2FN mobile dataset [23] contains 2000 images of 50 different faces of 144 × 176 pixels and few samples of database are shown in Fig. 5. This database is chosen to corroborate the mobile face recognition. Comparison of proposed method with the extant approaches for the

Table 2 Comparison of the existing approaches on different database

S. No.	Algorithms	AR Dataset Error Rate (%)	O2FN Dataset Error Rate (%)	LFW Dataset Error Rate (%)	Dyn Tex++ Dataset Error rate (%)
1	Nearest-Neighbor Classifier with Euclidean distance (NNC-ED) [2]	3.43	33.20	28.70	10.43
2	Nearest-Neighbor Classifier with Chi Squared Distance (NNC-Chi2D) [22]	1.33	30.00	26.35	5.92
3	Principal Component Analysis with Mahalanobis distance (PCA-mMDC) [25]	2.48	5.00	21.05	5.48
4	Chi Squared Transformation-Asymmetric Principal Component Analysis (CST-APCA) [25]	0.38	1.50	9.50	3.02
5	Principal Component Analysis with Mahalanobis distance (SR-PCA-mMDC) [36]	1.33	2.80	15.88	4.60
6	Power mean support vector machine with Chi Squared Distance (PmSVM-Chi2) [34]	0	4.00	13.67	11.79
7	Power mean support vector machine with Histogram intersection kernel (PmSVM-HI) [34]	0	4.10	13.43	11.83
8	Joint Bayesian [7]	–	–	15.92	–
9	Distance Learning Pegasos (DL-PEGASOS) [10]	–	–	–	36.30
10	Dynamic fractal analysis (DFS) [39]	–	–	–	10.10
11	Dimensionality Reduced Local Directional Pattern (DR-LDP) [43]	–	–	26.87±0.7	–
12	Modified Dimensionality Reduced Local Directional Pattern (MDR-LDP) [44]	–	–	21.28±0.6	–
13	Compound Dictionary Learning Based Classification (CDLBC) [17]	2.10	4.20	20.8	–
14	Hybrid feature extraction (HFE) [30]	1.9	3.9	13.7	–
15	**Proposed**	**0.28**	**1.30**	**9.00**	**2.5**

Fig. 5 O2FN mobile dataset [23]

same dataset is shown in Table 2. It is observed that the performance of the proposed algorithm is better than the existing approaches.

    (c)   Analysis of LFW dataset for face recognition

The LFW dataset [41] contains 13,233 images of faces of 5749 different persons and eight diffenernt faces are shown in Fig. 6. In this dataset all the face images collected from internet showcase variation of expression, posture and illumination. The high dimension LBP feature gives robust performance than baseline LBP feature and baseline HOG feature [40]. Comparison of proposed algorithm with various existing approaches on the basis of percentage error rate is shown in Table 2. It is observed that the proposed approach performed better than extant subspace approaches although memory consumption and computation cost is quite high for this data set.

    (d)   Analysis of DynTex++ database for dynamic texture recognition.

The DynTex++ database [20] contains 36 classes and every class has 100 sequence of $50 \times 50 \times 50$ size. This dataset is widely using for dynamic texture recognition and it has large dimension as compared to face databases because of this it consumes relatively more memory and computation cost also soars. For the validation of approaches, the test bench is designed as five cross average validation; 80 sequences for training set, rest 20 sequences for testing. The same experiment is repeated for 5 time and average results are taken into consideration. Comparative analysis of proposed algorithm with existing approaches is shown in Table 2. It is clear that proposed approach performed better as compared to the existing approaches.

Fig. 6 LFW dataset [41]

In order to get better physical insight of the proposed technique a comparison analysis of error rate percentage with respect to reduced feature percentage and error rate percentage for standard algorithms have been performed. By using the proposed feature reduction technique, the percentage change in error rate with respect to the percentage change in reduction of features is studied for the standard databases as represented in Fig. 7. Less than 3% error rate is achieved with 27% reduction of features for the entire tested database except LFW in which it is less than 10%. In comparison to the existing recognition and detection algorithms the proposed technique exhibits least error rate with maximum feature reduction in minimum time for all the standard databases. Moreover, the proposed technique is dynamic in nature.

Further, the error rate performance of the existing algorithms for the standard databases has been compared to the proposed algorithm as represented in Fig. 8. It is observed that the percentage error rate is lowest for the proposed algorithm for all the standard databases as compared to existing algorithms.

Further, the proposed method is also contrasted for standard datasets with different algorithms. The parameters including precision, sensitivity and F-measure of the intended technique are compared with other existing algorithms for standard datasets as represented in Fig. 9a-d. It is observed that the precision of the proposed algorithm is comparable to the existing algorithm, but on the other hand, the sensitivity and F-measure is much larger, which proves the efficacy of the retrieved features through the use of proposed technique. Therefore, by applying the proposed technique the computation time for recognition as well as the memory usage has been reduced significantly. This makes the proposed algorithm suitable for real time applications and memory devices like Raspberry Pi etc.

False acceptance rates (FAR) and the true acceptance rates (TAR) are significant parameters for all surveillance related applications. As face recognition is now a days gaining popularity in surveillance environment so a comparison of FAR & TAR has been performed for all readily available datasets with the existing algorithms as shown in Fig. 10a-d. It is observed that TAR for varying for the proposed algorithm is higher than the existing recognition techniques for varying values of FAR which makes it highly efficacious for potential security and forensic investigation applications.
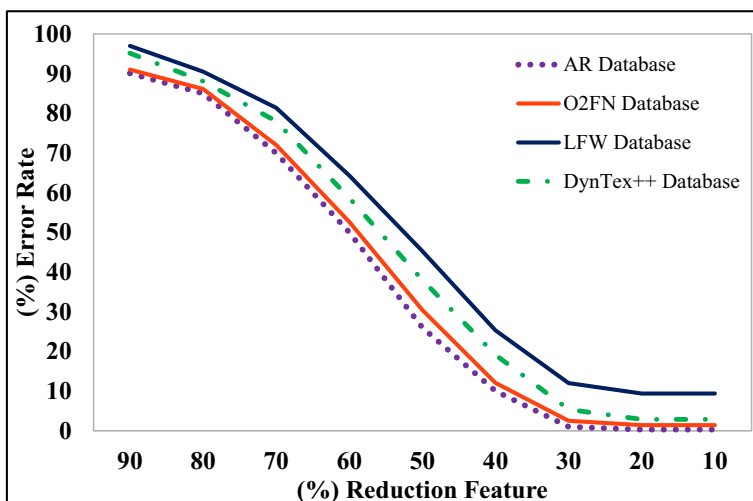

Fig. 7 Percentage reduced features with respect to percentage error rate
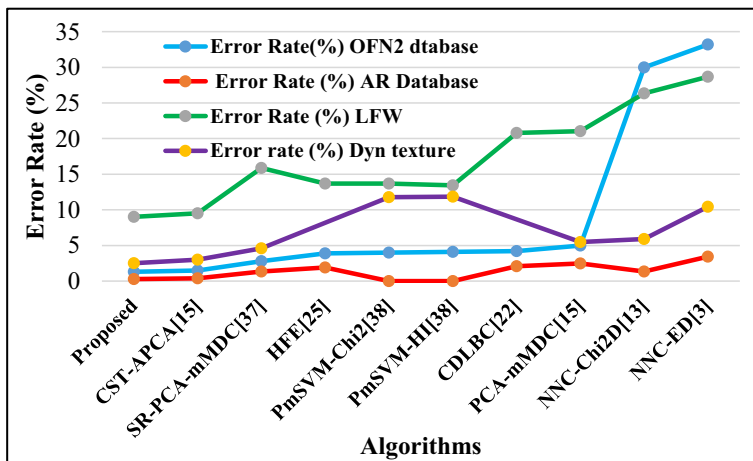
Fig. 8 Performance of different algorithms for standard databases

For verification of performance and accuracy of the proposed algorithm, it is compared with existing algorithms for standard databases as represented in Fig. 11. It is noticed that the proposed technique is as accurate as the other algorithms even after
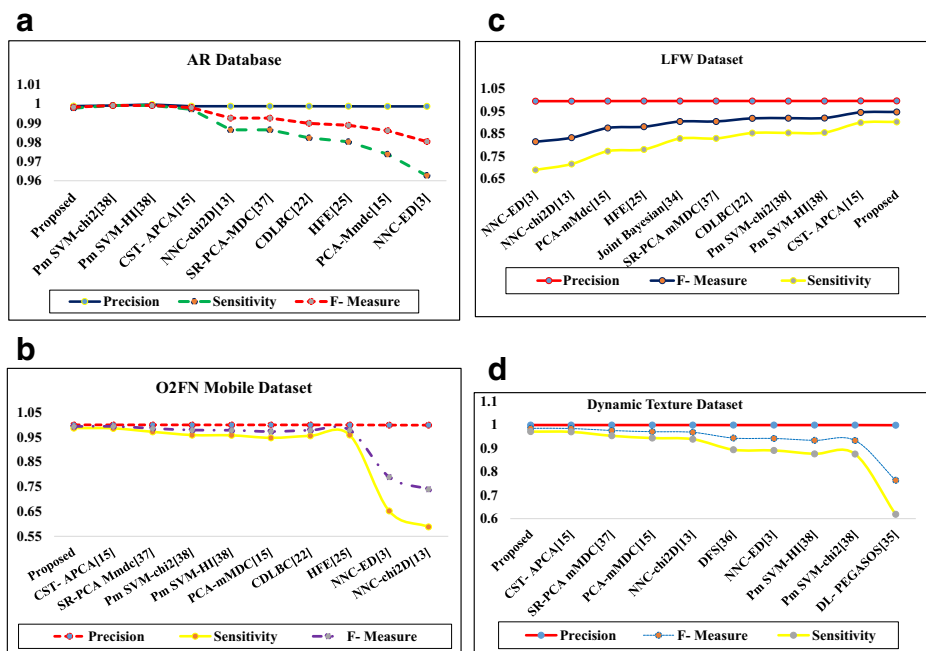

Fig. 9 **a** Analysis of precision, sensitivity and F-measure of AR dataset of the proposed approach with existing algorithms. **b** Analysis of precision, sensitivity and F-measure of O2FN mobile dataset of the proposed approach with existing algorithms. **c** Analysis of precision, sensitivity and F-measure of LFW dataset of the proposed approach with existing algorithms. **d** Analysis of precision, sensitivity and F-measure of Dynamic Texture dataset of the proposed approach with existing algorithms
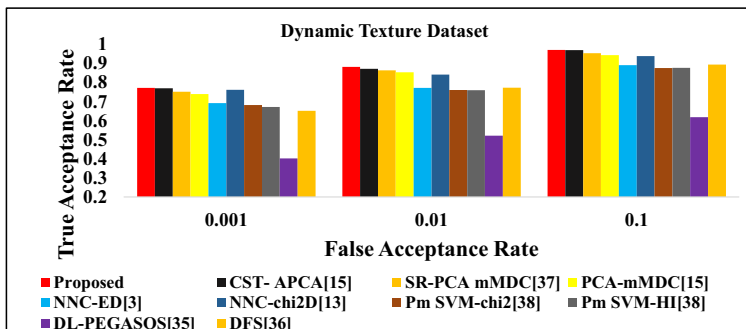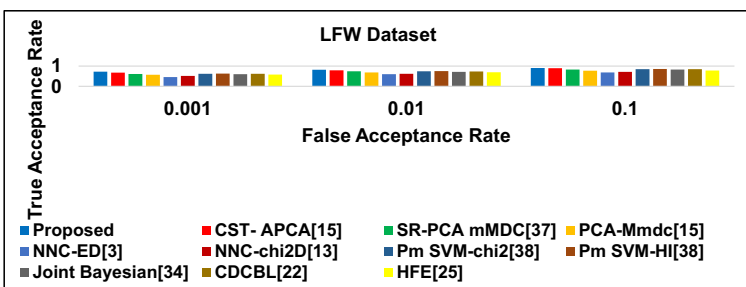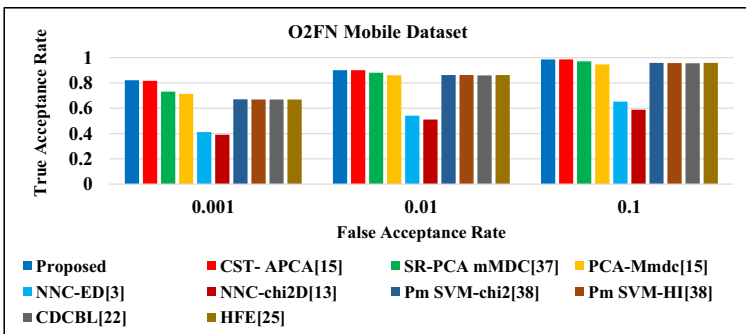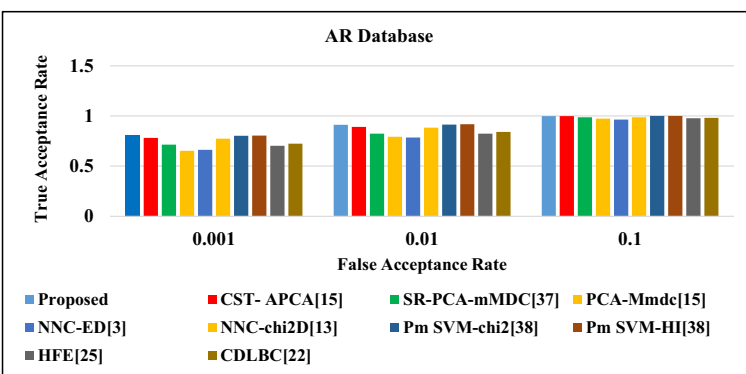
a

b

c

d

◀ **Fig. 10 a** TAR Vs FAR comparative analysis of Dynamic Texture database of existing and proposed algorithms. **b** TAR Vs FAR comparative analysis of LFW database of existing and proposed algorithms. **c** TAR Vs FAR comparative analysis of O2FN database database of existing and proposed algorithms. **d** TAR Vs FAR comparative analysis of AR database of existing and proposed algorithms

significant reduction in features. This shows that the features which are dropped or neglected were redundant and have no impact on accuracy. Therefore, the reduction in features while maintaining the accuracy of the technique saves time, memory and power consumption.

For further verification of the performance of the suggested method the feature dimensionality is compared with the existing techniques as illustrated in Table 3. It is noticed that the proposed algorithm exhibits maximum dimensionality reduction as compared to existing algorithm. Therefore, the proposed technique is capable of performing visual recognition efficiently with minimum feature size in minimum time span. This capability of the suggested approach makes it suitable for the real time implementation on the Raspberry Pi board for the potential uses in IoT applications such as forensic applications, identification in banking sector, in AADHAR database and texture recognition applications. The power consumption of the board is optimum due to better efficiency of the algorithm.

3.1 Real time implementation of vision recognition system

The proposed algorithm is validated through experimental results shown in results section. The features have been reduced effectively so that deployment of the algorithm on IoT devices is achieved in real time. For the real time application it can be implemented on the suitable IoT devices for prototype of the vision system. The proposed system is implemented using open source library OpenCV in C on Raspberry Pi running Ubuntu with USB camera.
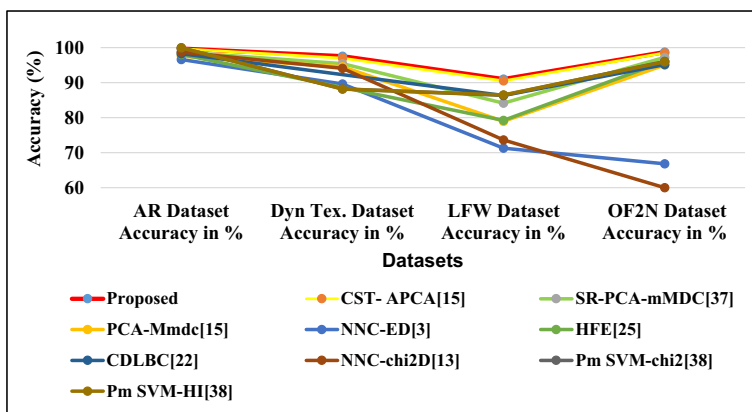

Fig. 11 Percentage accuracy comparisons

Table 3 Feature comparison of the proposed algorithm with the existing approaches in the literature

Sr. No.	Technique(s) Applied	No. of Bins	Feature Reduction in Comparison to basic LBP
1.	Basic LBP [2]	$256 \times 8 \times 8 = 2^{14} = 16,384$	–
2.	LBP, CST and PCA [25] (*Time consuming due the use of multiple lengthy approaches*)	$59 \times 8 \times 8 = 3776$ reduced by APCA	Greater than $2^{13.6645} \sim 12,984$
3.	LBP-Three Orthogonal Plane [43]	$3 \times 2^P \times 8 \times 8 = 49,152$	Nil ($-32,738$)
4.	LBPHF [44]	$59 \times 8 \times 8 = 2^{11.885} = 3776$	$2^{13.622} \sim 12,608$
5.	Noise Resistance LBP [24]	$59 \times 8 \times 8 = 2^{11.885} = 3776$	$2^{13.622} \sim 12,608$
	Extended NRLBP [24]	$107 \times 8 \times 8 = 2^{12.6545} \sim 6448$	$2^{13.2785} \sim 9936$
6.	Concatenation of RLBP and DLBP [27]	$196 \times 8 \times 8 = 12,544$	**$2^{11.907} \sim 3840$**
7.	Maximum conditional mutual information LBP and PCA. [26] (*Time consuming due the use of lengthy PCA approach*)	$256 \times 8 \times 8 = 2^{14} = 16,384$	Up to some extent
8.	**Proposed**	$59 \times 8 \times 8 = 3776$ reduced by Fast Subspace Technique	Greater than $2^{13.718} \sim 13,476$

4 Conclusion

Effective dimensionality diminution by utilising fast subspace technique with Chi square transformation for Smart City Surveillance using IoT device. This technique is applied to the extracted feature histogram of local binary pattern for further reduction of redundant features. A reduction of 13,476 features is achieved in comparison to the basic LBP algorithm. The reduction of unreliable features improves the capacity of the system memory and reduces the response time of the system which is desired for IoT applications. The proposed algorithm is verified and validated on the sample face of author himself by using Raspberry Pi as the hardware development kit. The same steps can be implemented on other IoT devices like Arduino, RoboCV etc. The proposed algorithm exhibits minimum error rate with maximum feature reduction in minimum time for all the standard databases maintaining the accuracy as much as of the existing techniques. These characteristics of the proposed scheme prove it useful for real time implementation of face and other recognition for IoT based surveillance system.

In future, this method can be explored further in consideration with potential deep learning techniques for implementation of real time IoT applications. The same architecture and algorithm be deployed and tested for any visual recognition problem. The proposed architecture and algorithm is generic enough as shown in the results section that it works well on face as well as texture recognition. The real time speed to problems like highway surveillance may be a bottleneck and may need further investigation. The further improvement can be investigated in three areas. First to further reduce the computation complexity so that the frames per seconds of the system can be increased. Secondly, the power consumption factor needs to be investigated and reported for proposed architecture. In future, the proposed architecture can be extended to apply on datasets where human faces are having face masks in post COVID-19 era for person identification. Also this work can be utilized for automatic attendance during online sessions as in the pandemic. Furthermore, this scheme can be explored in doing fingerprint and iris recognition for complete biometric verification in banking or other high security services.

References

1. Abozaid A, Haggag A, Kasban H, Eltokhy M (2019) Multimodal biometric scheme for human authentication technique based on voice and face recognition fusion. Multimed Tools Appl 78(12):16345–16361. https://doi.org/10.1007/s11042-018-7012-3
2. Ahonen T, Hadid A, Pietikäinen M (2006) Face description with local binary patterns: application to face recognition. IEEE Trans Pattern Analys Machine Intell 28(12):2037–2041. https://doi.org/10.1109/TPAMI.2006.244
3. Alrjebi MM, Liu W, Li L (2018) Face recognition against illuminations using two directional multi-level threshold-LBP and DCT. Multimed Tools Appl 77(19):25659–25679. https://doi.org/10.1007/s11042-018-5812-0
4. Alwakeel M, Shaaban Z (2010) Face recognition based on haar wavelet transform and principal component analysis via levenberg-marquardt back-propagation neural network. Eur J Sci Res 42(1):25–31
5. Benavente R, Barcelona UAD (1999) The AR face database abstract vol 24
6. Chen GY (2019) An experimental study for the effects of noise on face recognition algorithms under varying illumination. Multimed Tools Appl 78(18):26615–26631. https://doi.org/10.1007/s11042-019-07810-y
7. Chen D, Cao X, Wen F, Sun J (2013) Blessing of dimensionality: high-dimensional feature and its efficient compression for face verification. Proc IEEE Comput Soc Confer Comput Vis Pattern Recognition:3025–3032. https://doi.org/10.1109/CVPR2013389
8. Comon P, Golub GH (1990) Tracking a few extreme singular values and vectors in signal processing. Proc IEEE 78(8):1327–1343
9. Dahmouni A, El Moutaouakil K, Satori K (2018) Face description using electric virtual binary pattern (EVBP): application to face recognition. Multimed Tools Appl 77(20):27471–27489. https://doi.org/10.1007/s11042-018-5932-6
10. Ghanem B, Ahuja N (2010) Maximum margin distance learning for dynamic texture recognition ECCV: 223–236
11. Gumus E, Kilic N, Sertbas A, Ucan ON (2010) Evaluation of face recognition techniques using PCA, wavelets and SVM. Expert Syst Appl 37(9):6404–6408. https://doi.org/10.1016/jeswa201002079
12. Guo Z, Zhang L, Zhang D (2010) A completed modeling of local binary pattern. IEEE Trans Image Process 19(6):1657–1663. https://doi.org/10.1109/TIP20102044957
13. Huan EY, Wen GH (2020) Transfer learning with deep convolutional neural network for constitution classification with face image. Multimed Tools Appl 79(17-18):11905–11919. https://doi.org/10.1007/s11042-019-08376-5
14. Kas M, El-merabet Y, Ruichek Y, Messoussi R (2020) A comprehensive comparative study of handcrafted methods forface recognition LBP-like and non LBP operators. Multimed Tools Appl 79(1-2):375–413. https://doi.org/10.1007/s11042-019-08049-3
15. Khan SA, Hussain A, Usman M (2018) Reliable facial expression recognition for multi-scale images using weber local binary image based cosine transform features. Multimed Tools Appl 77(1):1133–1165. https://doi.org/10.1007/s11042-016-4324-z
16. Li Z, Liu G, Yang Y, You J (2012) Scale and rotation-invariant local binary pattern using scale-adaptive text on and sub uniform-based circular shift. IEEE Trans Image Process 21(4):2130–2140. https://doi.org/10.1109/TIP20112173697
17. Li L, Peng Y, Liu S (2020) Compound dictionary learning based classification method with a novel virtual sample generation Technology for Face Recognition. Multimed Tools Appl 79:23325–23346. https://doi.org/10.1007/s11042-020-08965-9
18. Liao S, Law MWK, Chung ACS (2009) Dominant local binary patterns for texture classification. IEEE Trans Image Process 18(5):1107–1118. http://www.ncbinlmnihgov/pubmed/19342342. Accessed 21 Apr 2020
19. Liu S, Li L, Jin M, Hou S, Peng Y (2020) Optimized coefficient vector and sparse representation-based classification method for face recognition. IEEE ACCESS 8:8668–8674. https://doi.org/10.1109/ACCESS20192960928
20. Péteri R, Fazekas S, Huiskes MJ (2010) DynTex: a comprehensive database of dynamic textures pattern. Recognition Lett 31(12):1627–1632. https://doi.org/10.1016/jpatrec201005009
21. Reilly JP (1987) A real-time high-resolution technique for angle of-arrival estimation. Proc IEEE 75(12):1692–1694
22. Ren J, Jiang X, Yuan J (2013) Noise-resistant local binary pattern with an embedded error-correction mechanism. IEEE Trans Image Process 22(10):4049–4060. https://doi.org/10.1109/TIP20132268976
23. Ren J, Jiang X, Yuan J (2013) A complete and fully automated face verification system on mobile devices. Pattern Recogn 46(1):45–56

24. Ren J, Jiang X, Yuan J (2013a) Dynamic texture recognition using enhanced LBP features ICASSP, IEEE international conference on acoustics, speech and signal processing - proceedings :2400–2404. https://doi.org/10.1109/ICASSP20136638085
25. Ren J, Member S, Jiang X, Member S (2015) A chi-squared-transformed subspace of LBP histogram for visual recognition. IEEE Trans Image Process 24(6):1893–1904
26. Ren J, Jiang X, Yuan J (2015) Learning LBP structure by maximizing the conditional mutual information. Pattern Recognit:1–11
27. Satpathy A, Jiang X, Eng HL (2014) LBP-based edge-texture features for object recognition. IEEE Trans Image Process 23(5):1953–1964. https://doi.org/10.1109/TIP20142310123
28. Shao C, Gao S, Song X, Yang X, Xu G (2018) Linear representation of intra-class discriminant features for small-sample face recognition. J Eng 2018(16):1668–1673. https://doi.org/10.1049/joe20188306
29. Sharma S, Kumar V (2020) Voxel-based 3D face reconstruction and its application to face recognition using sequential deep learning. Multimed Tools Appl 79:17303–17330. https://doi.org/10.1007/s11042-020-08688-x
30. Shoba VBT, Sam IS (2020) A hybrid features extraction on face for efficient face recognition. Multimed Tools Appl 79:22595–22616. https://doi.org/10.1007/s11042-020-08997-1
31. Tan X, Triggs B (2010) Enhanced Local Texture Feature Sets for Face Recognition Under Difficult Lighting Conditions. IEEE Trans Image Process 19(6):1635–1650. https://doi.org/10.1109/TIP.2010.2042645
32. Tiong LCO, Kim ST, Ro YM (2019) Implementation of multimodal biometric recognition via multi-feature deep learning networks and feature fusion. Multimed Tools Appl 78(16):22743–22772. https://doi.org/10.1007/s11042-019-7618-0
33. Tufts DW, Melissinos CD (1985) Simple, effective computation of principal eigenvectors and their eigenvalues and application to high resolution estimation of frequencies. IEEE Recent Publ:320–323
34. Wu J (2012) Power mean SVM for large scale visual classification. Proc IEEE Conf Comput Vis Pattern Recognit Providence 2344–2351. https://doi.org/10.1109/CVPR.2012.6247946
35. Wu J, Rehg JM (2011) CENTRIST: a visual descriptor for scene categorization. IEEE Trans Pattern Analys Machine Intell 33(8):1489–1501. https://doi.org/10.1109/TPAMI2010224
36. Wu J, Zhang Y, Lin W (2014) Towards good practices for action video encoding proceedings of the IEEE computer society. Confer Comput Vision Pattern Recog:2577–2584. https://doi.org/10.1109/CVPR2014330
37. Xu G, Kailath T (1991) Fast subspace decomposition of data matrices. Conference record of the twenty-fifth asilomar conference on signals, systems & computers, Pacific Grove, CA, USA 908–912. https://doi.org/10.1109/ACSSC.1991.186578
38. Xu G, Kailath T (1994) Fast subspace decomposition. IEEE Trans Signal Process 42(3):539–551
39. Xu Y, Quan Y, Ling H, Ji H (2011) Dynamic texture classification using dynamic fractal analysis. Proc IEEE Int Confer Comput Vision:1219–1226. https://doi.org/10.1109/ICCV20116126372
40. Zeng C, Ma H (2010) Robust head-shoulder detection by PCA-based multilevel HOG-LBP detector for people counting. 2010 20th Int Confer Pattern Recognition, Istanbul 2069–2072. https://doi.org/10.1109/ICPR.2010.509
41. Zhang N, Deng W (2016) Fine-grained LFW database. 2016 International Conference on Biometrics ICB, Halmstad, 1–6 https://doi.org/10.1109/ICB.2016.7550057
42. Zhang S, Zhang L (2018) Combining weighted adaptive CS-LBP and local linear discriminant projection for gait recognition. Multimed Tools Appl 77(10):12331–12347. https://doi.org/10.1007/s11042-017-4884-6
43. Zhao G, Pietika M (2007) Dynamic texture recognition using local binary patterns with an application to facial expressions. IEEE Trans Pattern Anal Mach Intell 29(6):915–928
44. Zhao G, Ahonen T, Matas J, Pietikäinen M (2012) Rotation-invariant image and video description with local binary pattern features. IEEE Trans Image Process 21(4):1465–1477. https://doi.org/10.1109/TIP20112175739
45. Zhou L, Wang H, Liu W, Lu ZM (2019) Face feature extraction and recognition via local binary pattern and two-dimensional locality preserving projection. Multimed Tools Appl 78(11):14971–14987. https://doi.org/10.1007/s11042-018-6868-6
46. Zou G, Fu G, Gao M, Pan J, Liu Z (2020) A new approach for small sample face recognition with pose variation by fusing Gabor encoding features and deep features. Multimed Tools Appl 79:23571–23598. https://doi.org/10.1007/s11042-020-09076-1

Affiliations

Manish Kumar[1] · Kota Solomon Raju[2] · Dinesh Kumar[1] · Nitin Goyal[3] · Sahil Verma[4] · Aman Singh[5]

Manish Kumar
ermanish205@gmail.com

Kota Solomon Raju
kotasolomonceeri@gmail.com

Dinesh Kumar
dineshkumaresdkuk@gmail.com

Nitin Goyal
nitin.goyal@chitkara.edu.in

Sahil Verma
sahilverma@ieee.org

[1]   Electronic Science Department, Kurukshetra University, Kurukshetra, Haryana, India

[2]   Central Electronics Engineering Research Institute, CSIR, Pilani, India

[3]   Chitkara University Institute of Engineering and Technology, Chitkara University, Rajpura, Punjab, India

[4]   Department of Computer Science and Engineering, Chandigarh University, Mohali, Punjab 140413, India

[5]   Department of Computer Science and Engineering, Lovely Professional University, Phagwara, Punjab, India