Evaluation and Applying Feature Extraction Techniques for Face Detection and Recognition

R Arokia Paul Rajan[1], Angel Rose Mathew[2]
[1,2] Department of Computer Science, CHRIST (Deemed to be University), Bangalore, Karnataka, India

Article Info	ABSTRACT
	Detecting the image and identifying the face has become important in the field of computer vision in recognizing and analyzing, reconstructing into 3D, and labelling the image. Feature extraction is usually the first stage in detection and recognition of the image processing and computer vision. It supports the conversion of the image into a quantitative data. Later, this converted data can be used for labelling, classifying and recognizing a model. In this paper, performance of such feature extraction techniques viz. Local Binary Pattern (LBP), Histogram of Oriented Gradients (HOG) and Convolutional Neural Network (CNN) technique are applied to detect and recognize the face. The experiments conducted with a data set addressing the issues like pose variation, facial expression and intensity of light. The efficiency of the algorithms was evaluated based on the computational time and accuracy rate.

Corresponding Author:

R Arokia Paul Rajan,
Department of Computer Science,
CHRIST (Deemed to be University),
Bangalore, Karnataka, India,
Email: arokia.rajan@christuniversity.in

1. INTRODUCTION

The most important part to identify a human, is face which is unique physical characteristic except identical twins. Computers have just started to identify while for billion years, humans have been able to differentiate and recognize human faces. It has become a part of human-computer interaction by segmenting the background and modelling the face of the facial image [1]. Taking an account to remove the irregularities, including illumination, including shadowing and colour, environmental conditions, different poses of face and occlusion. Verification and identification are the two comparisons for face recognition. In authentication, in framework the face is compared with the given face in the database and decides true or false. In identification, the system compares the given face with other faces in the database and produces a match found. The following are the four steps in face recognition:

- Input: An image is captured and face is detected.
- Extract: A template is generated from the data extracted, including image scaling and rendering, segmentation etc.
- Compare: Template is compared with a new sample for identification.
- Result: System decides whether the sample matches with any extracted features.

Face recognition techniques are complex in nature because it analyses shape positions and patterns for an individual face. The main application of face recognition is in incorporating with biometric devices. Depending on the algorithm, a pattern is generated based on the characteristics of the human face. Usage of face recognition devices is increasing at a high rate and it is needed a high accuracy to make use of it as the user authentication and identification in banks, creating access codes, generating PIN and passwords for the security purposes . In such devices, face recognition techniques identifies a person by various methods like detecting the skin tone, face and eyes. These patterns are categorized into groups and are trained with labels.

So later, when an image is captured the classifier will compare with the other patterns stored in the database and check if any faces are present or not. For each comparison, lists of values are obtained [2].

Face recognition can replace the PIN and password methods as they can never be verified by another person and face cannot be misused except identical twins. Users can generate PIN and password that can be guessed easily by another person and which can be misused. It is one of the biggest challenging issues in security enforcement. Face recognition techniques are widely used in the many fields recently because,

- Users - physical interaction is not required.
- It gives accurate results for verification and authentication.
- Without any expertised interpretation, the results can be achieved.
- Helps to identify the suspects.
- Mainly for security reasons such as in banks, government offices etc.

In face recognition at first, a face recognition algorithm is applied to an image that has been selected and is called 'input image'. On the input image it searches for the faces. If matches are found, then its content and location of the face is returned. Once the face is detected, features are extracted which differentiates faces among other faces and then algorithms are applied to recognize these faces in the image [3]. Results are obtained accordingly if an image is identified as depicted in the Figure1.
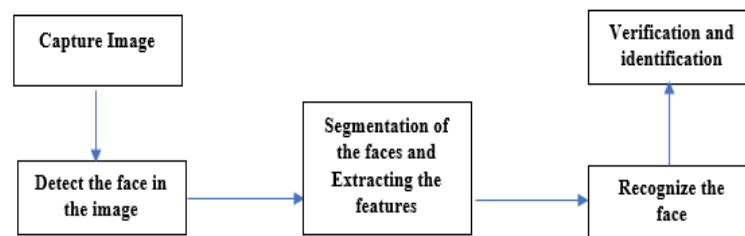

Figure 1. Generic face recognition mechanism

1.1. Feature extraction techniques

There are many face recognition algorithms to recognize faces, includes Linear Discriminant Analysis (LDA), Genetic Algorithm, Support Vector Machine (SVM), Local Binary Pattern (LBP), Convolutional Neural Network (CNN, or ConvNet), Principal Component Analysis (PCA), Histogram of Oriented Gradients (HOG) and Back Propagation Neural Networks (BPNN). Feature extraction methods are used to extract features like eye, nose, mouth, and eyebrows for face recognition and reduce the dimensionality [4]. There are various techniques to extract features like Geometry, colour, template and appearance-based techniques. All these methods follow the different mathematical approach and functions depending on the database that is being experimented. This work aim to analyze the three feature extraction algorithms, namely LBP, HOG and CNN. The difference between these feature extraction methods is that both LBP and HOG take the information of gradients in a different way. LBP takes each pixel in eight directions, whereas HOG takes one direction. LBP is used for retrieving local patterns, whereas HOG takes corners and edges in an image. In case of CNN it is a deep learning hierarchical architecture where ConvNet operation trains the filters and repeatedly filters the information at every stage.

2. RELATED WORKS

The comparison of the reduced dimensionality algorithm such as Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) for a huge training dataset ORL database was carried out in [5]. The performance of the algorithm was improved by pre-processing the images. It also evaluated the algorithm by considering the face verification rate and recognition rate. It was concluded that for large dataset LDA performed better than PCA.

[6] reviewed different methods of face recognition like Independent Component Analysis (ICA), Linear Discriminant Analysis (LDA), Principal Component Analysis (PCA) and combined the hybrid techniques of soft computing tools in order to enhance the recognition of the face which includes Artificial Neural Networks (ANN), SOM and Support Vector Machine (SVM). It also analysed the challenges of facial recognition as the parameter like facial expression, illumination and pose variation along with other methods.

In [7] the performance of feature extraction methods, i.e. SpeeduUp Robust Features, Histograms of Oriented Gradients, Local Binary Patterns, Fully Affine SIFT, Scale Invariant Feature Transform and Gabor feature on three different datasets of face i.e. ORL, Yale and UMIST were evaluated. Through suitable experiments, it was calculated the matching time and recognition rate for different conditions. Found that SURF was better at matching time, but not much effective for application in real time, ASIFT at recognition rate, but time cost is limited and LBP is more adaptable to illumination.

The classification model viz. decision tree, artificial neural network and support vector machines along with extraction of features methods were compared in [8]. Local binary patterns, the histogram of gradients, and a pre-trained deep network on Histopathology Images gathered from KIMIA Path960 dataset. It was concluded that highest accuracy is obtained by SVM classification and LBP feature extraction methods respectively.

[9] conducted a detailed study of different techniques of face detection and feature extraction under various conditions. This work used Geometry based technique uses Gabor wavelet method which considers eyes, mouth and nose for feature extraction. The advantage of this work is that it could recognize for a small database. Also, a template based technique 'Deformable template' used which gives better accuracy for recognition but has a complexity description between the image and template. Colour-based technique uses the colour-based model, but performance varies with the external factors. It included an appearance-based approaches like PCA, ICA, LDA but it requires large database and quality images.

An experiment was conducted to find the best feature extraction by calculation the computational time of features of accelerated segment test (FAST), scale-invariant feature transform, binary robust invariant scalable key points (BRISK) and speeded-up robust features (SURF) [10]. It considered the shadow region for comparison and found that if combined all, it extracts the features in less time and gives higher accuracy. This work used satellite image dataset.

In [11], a comparison on the Histogram of Oriented Gradients and Haar-like cascade with Support Vector Machine which was applied for the detection of automatic runway in high-resolution satellite image results. It was shown that if LBP and Haar-like cascade combined it performs better than HOG+SVM. [12] presented the importance of detecting face is a crucial in this digital era and evaluated in terms of key parameters and came up with an approach for face detection. It is concluded that haar-like cascade is better for extracting features as it has less false alarm and low execution time.

[13] proposed CNN based face detector. This work experimented many parameters and trained a huge volume of data with the help of GPU technology. The experiment has been conducted on FDDB and MUST database which was better at detecting the occulated face images. Haar-like Cascade and HOG were used for detecting faces along with CNN. A model has been proposed by [14] for detecting expressions of the face using the feature extraction LBP method and CNN. The features of the face are extracted using the LBP method and it is converted into a LBP vector image. These images then trained using the CNN technique for face recognition.

3. EVALUATION OF FEATURE EXTRACTION ALGORITHM

The objective of this research is to study and evaluation of the feature extraction methods namely, Local Binary pattern (LBP) and Histogram of Oriented Gradients (HOG) for face images and Detecting and recognizing the face using Convolutional Neural Networks (CNN) - deep neural network approach.

3.1. Local Binary Pattern (LBP)

Local Binary Pattern (LBP) has been playing a major role in image processing and computer vision for the last few years. It is a non-parametric descriptor which efficiently gives local patterns of an image by comparing each pixel value with neighbours as shown in the Figure2. It is widely used for analysing the texture. It is not much affected by the illumination compared to other methods and its simplicity in computation. It was introduced before HOG and SIFT. It is used in many applications like analysing face images and video modelling, analysing motion, biometrics [15], [16] etc.
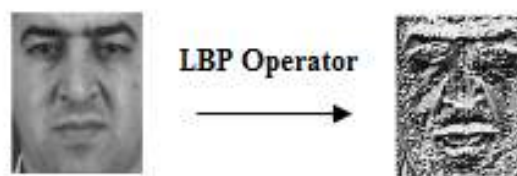

Figure 2. Generation of a grayscale image

The pixel of an image is labelled by the LBP operator called LBP codes, each pixel around the local pattern is encoded and a histogram is generated. The pixel values are in grayscale. The image is divided into 3x3 matrices where middle pixel p is a threshold with neighbour eight x pixel as shown in the Figure3. The binary threshold function s(p) is assigned to 0, when x < p and s(p) is assigned 1, when x >= p. Then it is multiplied by the power of two to convert the binary number obtained and is summed for obtaining a label as given in the equation (1).

$$LBP_{R,P} = \sum_{p=0}^{P-1} s(g_p - g_c).2^P \tag{1}$$

Where g_p is neighbourhood pixels, g_c is centre pixel thresholded value, p is sample points for example in 3x3 cell let p=0,1 up to 7 for P=8), R stands for radius, Coordinates of g_c = (0,0) and g_p = (x + Rcos(2π p/P), y - Rsin(2π p/P))
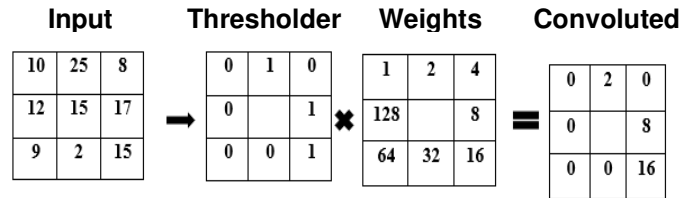

Figure 3. Basic LBP operator

LBP is calculated as follows for the above sample of data:

LBP = 2 + 8 + 16 = 26

C = (25 + 17 + 15) / 3 – (10 + 8 + 12 + 9 + 2) / 5 = -22

Limitation of LBP is it is not invariant to rotations, if there are a huge number of features, then the computation becomes complex and information is limited.

3.2. Histogram of Oriented Gradients (HOG)

In computer vision and image processing to detect the object, HOG feature descriptors are used. In an image the occurrences of gradient orientation are counted in the localized portions. The appearance shape of the local object can be described by the Descriptor of HOG which gives the edge direction. It is achieved by dividing the image into a cell which are small connected regions and foe each cell a histogram of gradient direction is compiled by edge directions [17]. In HOG, features are extracted by the following steps:

Step 1: Computation of Gradient by differential coefficients of the 1st order, Gx(i,j) and Gy(i,j) is calculated by this equation in (2) and (3) respectively.

$$G_x(i,j) = f(i+1,j) - f(i-1,j) \tag{2}$$

$$G_y(i,j) = f(i,j+1) - f(i,j-1) \tag{3}$$

where f(i,j) is lustrous at (i,j).

From the computed gradients m and Θ is calculated in the following equations (4) and (5) respectively:

$$m(i,j) = \sqrt{G_x(i,j)^2 + G_y(i,j)^2} \tag{4}$$

$$\theta(i,j) = \arctan\left(\frac{G_x(i,j)}{G_y(i,j)}\right) \tag{5}$$

Step 2: Generating the histogram from the values m and Θ by determining the class θ(i,j) belonging to and increasing the value, i.e. mn and mn+1 given in equation (6), (7) and (8).

$$n = \left[\frac{b\theta(i,j)}{\pi}\right] \tag{6}$$

$$m_n = (1 - \alpha)m(i,j) \tag{7}$$

$$m_{n+1} = \alpha m(i,j) \tag{8}$$

where b is the number of classes in total, proportional distribution of magnitude m(i,j) is α given in equation (9) which is the distance from θ(i,j) to class n and n+1.

$$\alpha = \frac{b}{\pi}\left(\theta(i,j) \bmod \frac{\pi}{b}\right) \tag{9}$$

Step 3: Normalizing histogram from the generated histogram and it eliminates the problem of illumination and contrast by L1-norm in the equation (10).

$$v = \frac{V_k}{\|V_k\| + \varepsilon}$$

<div align="right">(10)</div>

where V_k is the vector generated by combining the histogram, ε is constant and v is the final HOG feature which is normalized vector.

3.3. Convolutional Neural Networks (CNN)

It is a part of machine learning, modelled on the structure of the brain consisting of neurons i.e. network of training units. The neurons are trained to covert the captured image to labelled information for face recognition. Activation function is required to decide if an image has a face. The neurons are trained in such a way that when an image contains face, the 'face' label is activated. More the neurons are trained, it learns better to recognize the unlabeled images. After a long research in neural network the researchers, advanced into convolutional neural network. The visual field is entirely covered with layers of sub-regions over each other and process the images captured. Mainly used to identify the objects and traffic signs [18]. There are four steps to be followed:

Step 1: Convolution
Step 2: ReLU (non-linearity)
Step 3: Sub Sampling or Pooling
Step 4: Fully connected layer (Classification)

An image is always represented as a group of pixels. In an image any component is called channel. It mainly works on the grayscale image having 2D matrix representing an image. The range of pixel is between 0-255, where 0 is black and 255 is white in a matrix. The convolution operator extracts the features of the captured image and uses the square to learn and coordinate with the pixels. When a matrix slid on original image by one pixel it is called stride viewed as 3x3 matrix also known as filter or feature detector. And the generated matrix is called an activation map. The corresponding matrices are multiplied and the sum is calculated as shown in Figure4.
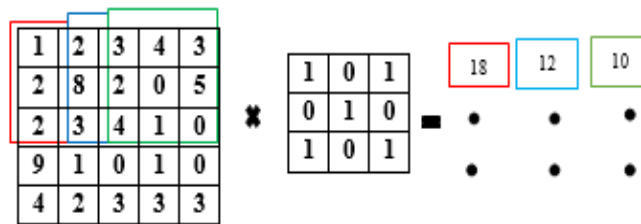

Figure. 4 Convolutional operator with stride one pixel and filter size 3x3

When input image size is n, stride size is t and the filter size is f then the output size is calculated as given in the equation (11).

$$output\ size = \frac{(n-f)}{t+1}$$

<div align="right">(11)</div>

In the process of training CNN learns from the values filtered. The model is efficient when number of filters are more which extracts a greater number of features. To perform convolution, we need to consider following parameters:

- *Stride:* When a matrix slide over the input image the quantity of pixel is decided. For stride 1, one pixel is moved by the filter and in stride 2, two pixels are moved. But less features will be mapped if stride is more.
- *Depth:* Number of filters required is defined for the convolutional operations as shown in the Figure5.
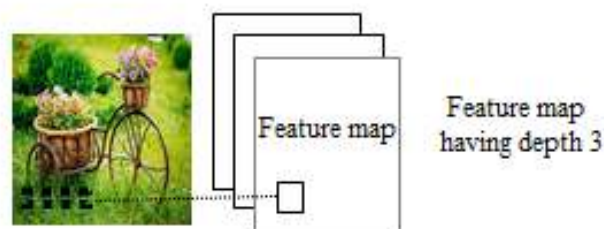

Figure 5. Feature map for depth three

- Zero-padding: It is also called wide convolution. It is a method to pad zeros on the borders of the matrix image. It is used to define the image properly.

Rectified Linear Unit (ReLU) is a supplementary operation and is a non-linear which is performed in all convolution operators. It is used to replace all negative value with 0 for every pixel. We use ReLU because it performs consistently compared to tanh or sigmoid. For feature map the dimensionality is reduced by Spatial Pooling, which is also called Subsampling or downsampling. From a spatial neighbourhood, if the element with the large rectified map is found, then it called max pooling and if the average is taken, then its average pooling and when the sum is taken then it is called the sum pooling. The input structural size can be reduced and efficient computation is achieved by pooling. From the output layer, the softmax activation function is used by the Multilayer Perceptron. In fully connected layer all the neurons are connected to each and every neuron of the previous layer. Features are classified into various classes based on the trained database.

4. EXPERIMENTS, RESULTS AND DISCUSSION

The main aim of this WORK is to study and analysis the feature extraction method Local Binary pattern (LBP) and Histogram of Oriented Gradients (HOG) for face images and Detecting and recognizing the face using Convolutional Neural Networks (CNN) which is a deep neural network to find a method which can detect and recognize faces efficiently and accurately from the images. This is achieved by executing the algorithms in python on the created dataset [20]. The dataset consists of images which are occulated and clicked under different illuminations.

It is experimented with different face images that can detect and recognize easily than the complexity increases by giving more facial images. Complex dataset is taken so that we can identify the best method, that can be used for any real-time applications. The image to be recognized is trained properly with CNN so that when an image is given, it can give the desired output accurately. This is achieved by executing the methods one by one and time taken to recognize the image is observed. Table 1 gives the time taken to recognize the face in an image dataset with varying face numbers in an image for e.g. one image contains two faces, so the time taken to recognize both the faces is calculated and recorded.

Table 1. Comparison of computational time to recognize the face (in milliseconds)

Algorithm	Exp 1	Exp 2	Exp 3	Exp 4	Exp 5	Mean
HOG	01.786	05.143	01.076	02.199	01.446	02.330
CNN	04.263	30.926	10.867	09.685	12.976	13.743
LBP	00.394	02.724	00.817	00.373	01.018	01.065

The results are taken from the observation and plotted in the form of a graph as shown in the Figure6. Comparison was made based on the number of images vs computation time taken to recognize a face from the trained dataset. This work experimented all the three methods on the dataset of six different people consisting of 50 images, each including with different pose and illumination. Local Binary Pattern (LBP) showed better detection and recognition rate in a given period of time measured in milliseconds (ms). Also, as the complexity of images increased all showcased similar recognition rates.
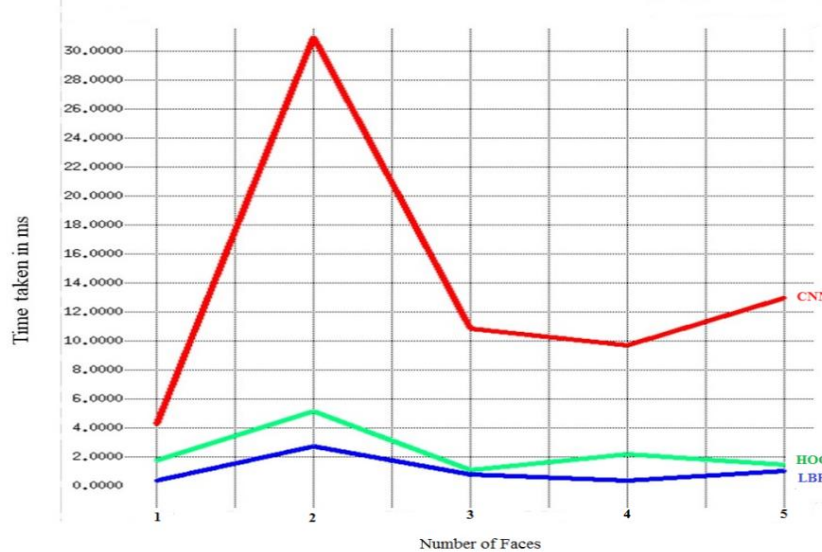

Figure 6. Detection and Recognition performance graph

Complex dataset of face images of different people is taken so that we can identify the best method, that can be used for any real-time applications as shown in the Figure7.
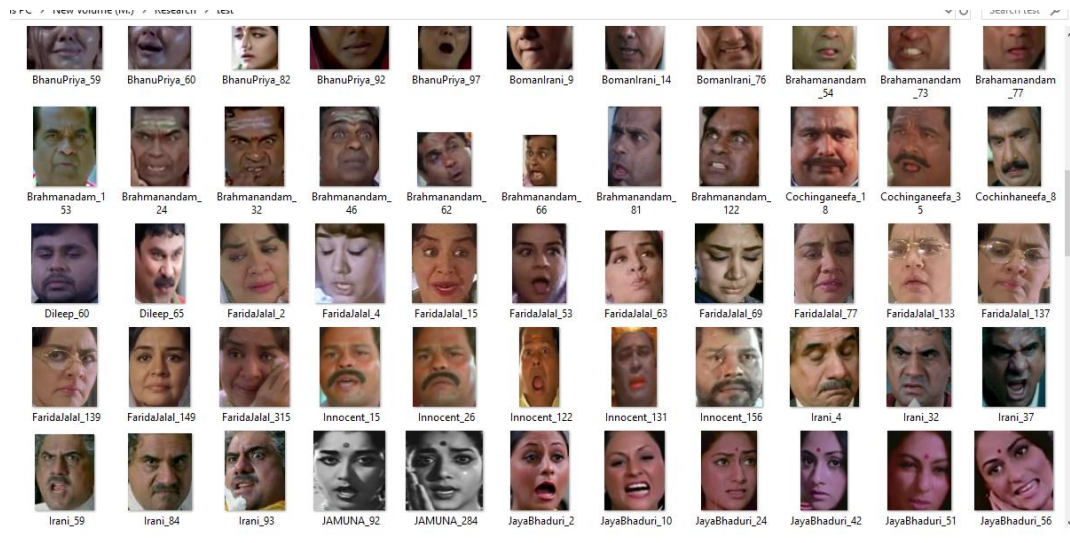

Figure 7. Sample Dataset of face images

When an image is given, it detects the face location and compares with the trained dataset, once it recognizes the face among a different set of faces the time is calculated to recognize the person as shown in the Figure8.
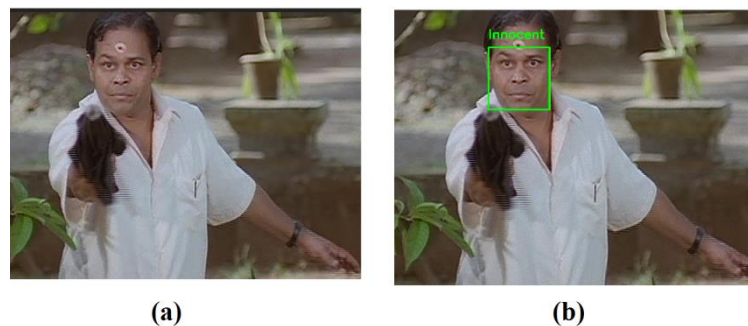

(a)   **(b)**

Figure 8. (a) Input image to be detected and recognized   (b) Successfully identied face of Actor Mr.Innocent

5.   CONCLUSION

This paper presented a few face detection and recognition techniques with feature extraction methods. The efficiency of LBP, HOG and CNN techniques has been evaluated and analysed based on the performance metrics. Based on this evaluation, it is concluded that LBP is better at extracting features fast and accurate no matter how large is the dataset. Also, this research experimented face detection and recognition using CNN technique. The accuracy of the CNN method is measured by the computational time and the recognition rate of a face. Since LBP and HOG produce a different set of data about the features extracted from an image, it is identified that the combination of these two methods along with CNN will produce more accurate results when applied to real time applications.

REFERENCES (
[1]   Rein-Lien Hsu. 2002. Face Detection and Modeling for Recognition. Dissertation, Department of Computer Science & Engineering, Michigan State University,.
[2]   Shang-Hung Lin. An Introduction to Face Recognition Technology. *Informing Science Special Issue on Multimedia Informing Technologies - Part 2,* IC Media Corporation, 2000; Vol 3, No. 1.
[3]   Tjokorda Agung Budi Wirayuda and Iping Supriana Suwardi. Opportunity and challenge on face recognition: generic-to-specific feature representation and recognition strategy. *International Journal of Latest Trends in Engineering and Technology,* 2017; Volume 9, Issue 1.
[4]   Rabia Jafri and Hamid R. Arabnia. A Survey of Face Recognition Techniques. *Journal of Information Processing Systems,* 2009; Vol.5, No.2, pp. 41-68,

[5] S. Suganya and D. Menaka. Performance Evaluation of Face Recognition Algorithms. *International Journal on Recent and Innovation Trends in Computing and Communication,* 2009; Volume 2, Issue 1, pp: 135 – 140,

[6] Sujata G. Bhele and V. H. Mankar. A Review Paper on Face Recognition Techniques. *International Journal of Advanced Research in Computer Engineering & Technology.* 2012; Volume 1, Issue 8, pp. 923-925.

[7] Yin Liu, *et al. Evaluation of feature extraction methods for face recognition*, 2013 IEEE Sixth International Symposium on Computational Intelligence and Design, pp. 313-316.

[8] Taha J. Alhindi, *et al. Comparing LBP, HOG and Deep Features for Classification of Histopathology Images.* IEEE World Congress on Computational Intelligence, July 2018.

[9] Urvashi Bakshi and Rohit Singhal. A Survey On Face Detection Methods And Feature Extraction Techniques of Face Recognition. *International Journal of Emerging Trends & Technology in Computer Science,* 2014; Volume 3, Issue 3, pp. 233- 238.

[10] Shahid Karim, *et al.* Comparative analysis of feature extraction methods in satellite imagery. *Journal of Applied and Remote Sensors,* 2017; Vol.11(4).

[11] Cruz, Shiguemori and Guimara˜es. A comparison of Haar-like, LBP and HOG approaches to concrete and asphalt runway detection in high resolution imagery. 2015; *J. Comp. Int. Sci.* Volume 6(3), pp.121-136.

[12] Mayank Chauhan and Mukesh Sakle. Study & Analysis of Different Face Detection Techniques. *International Journal of Computer Science Information Technologies,2014;* Volume 5 (2).

[13] Emine Cengil, Ahmet Cinars. Comparison of Hog (Histogram of oriented gradients) and Haar Cascade Algorithms with a Convolutional Neural Network Based Face Detection Approaches. *International Journal of Advance Research, Ideas and Innovations in Technology,* 2017; Volume 3, Issue 5..

[14] Sonali Sawardekar, Sowmiya Raksha Naik. Facial Expression Recognition using Efficient LBP and CNN. *International Research Journal of Engineering and Technology*, 2018; Volume 05, Issue 06.

[15] Pietikäinen *et al*, Computer Vision Using Local Binary Patterns, *Computational Imaging and Vision 40,* Springer-Verlag, 2011; pp. 13-47.

[16] T. Ahonen, A. Hadid, M. Pietikäinen. Face recognition with local binary patterns. *in Proc. Euro. Conf. Computer Vision (ECCV)*, pp. 469–481, 2004.

[17] Navneet DALAL. *Histogram of Oriented Gradients (HOG) for Object Detection*. 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), Vol. 1, pp. 886-893, June, 2005.

[18] John Murphy. An Overview of Convolutional Neural Network Architectures for Deep Learning. 2016; Microway Inc, pp. 1-22.

[19] Ludovic Trottier, *et al*, "Parametric Exponential Linear Unit for Deep Convolutional Neural Networks," 2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA), pp. 207–214, 2017.

[20] M. I. Ouloul, *et al,* An Efficient Face Recognition Using SIFT Descriptor in RGB-D Images, *International Journal of Electrical and Computer Engineering (IJECE),* 2015; Volume 5, No 6.

BIOGRAPHIES OF AUTHORS

Dr R Arokia Paul Rajan is currently working as Associate Professor, Department of Computer Science, CHRIST (Deemed to be University), Bengaluru, Karnatka, India. He holds Ph.D. in Computer Science & Engineering from Pondicherry University, India. His research area is data management in Cloud architectures. He published 17 research papers in various international journals and conferences.

Angel Rose Mathew pursued MCA from the Department of Computer Science, CHRIST (Deemed to be University), Bengaluru, India. Her area of research interest includes Computer Vision and Cloud Computing.