IJDACR
ISSN: 2319-4863

IJDACR
International Journal Of Digital Application & Contemporary Research

International Journal of Digital Application & Contemporary Research
Website: www.ijdacr.com (Volume 5, Issue 9, April 2017)

Facial Expression Recognition using Hybrid Method of Local Binary Pattern and Gabor Filter Features with Random Forest Classifier

Priyanjali Kuruvila
M. Tech. Scholar
Department of Computer Science
Central India Institute of Technology, Indore
priyanjali.km@gmail.com

Rasna Sharma
Assistant Professor
Department of Computer Science
Central India Institute of Technology, Indore
rasna.sharma4@gmail.com

Abstract –**The face though seems an easy object to be recognized by retina but the artificial intelligence is not yet intelligent enough to do the task easily. As the source of a face is generally an image capturing object, there are lot of variations and complexions that persists with the image like (for example: noise, rotation etc.). There are many techniques that use some or other algorithm to find similarity in face model and the test image and most of them are successful on their part to attain better test similarities. However, considering the diverse scale of applications and mode of image sourcing, a single algorithm cannot get maximum efficiency everywhere. Even after using the best algorithm for a particular task, an application has to counter with challenges of face recognition. The main aim of this paper is to analyze the Hybrid method of Local Binary Pattern and Gabor Filter features and its performance when applied to facial expression recognition. This algorithm creates a subspace (face space) where the faces in a database are represented using a reduced number of features called feature vectors and Random Forest Classifier calculates the similarity score for performance evaluation which will provide improved results in terms of recognition accuracy.**

Keywords – **Gabor Filter, JAFFE, Local Binary Pattern, Random Forest Classifier.**

I. INTRODUCTION

Face Recognition has become an important issue in many applications such as credit card verification, safety systems, and proof of identity of criminal. For instance, the capability to model a specific face and separate it from a large number of stored face models would make it conceivable to tremendously enhance evidence of character of criminal. Identifying faces in photos for automating colour film improvement may be very useful, meanwhile the influence of many enrichment and noise reduction strategies relies on upon the content of image. The examination interest in Face Recognition has become altogether as an aftereffect of the accompanying certainties:

- The growth in prominence on commercial/civilian research projects,
- The re-emergence of neural network classifiers with emphasis on real time computation and adaptation [1].
- The accessibility of real-time hardware,
- The increasing need for surveillance related applications due to terrorist activities, drug smuggling, etc. [2].

Although it is noticeable that people are comfortable with Face Recognition, it is not at all obvious how the faces are decoded and encode by the human brain. So developing a computational model of Face Recognition is not easier, because faces are multi-dimensional and complex visual stimuli. Therefore, Face Recognition is a very higher level computer vision stimuli task, in which many more early vision techniques can be involved. Face Recognition system is a complex image-processing problem in real world applications with complex effects of occlusion, illumination and imaging condition on the live images. These images have some known properties like; same facial feature components, same resolution and analogous eye alignment. These images will be referred as “standard image” in the further sections. Recognition applications uses standard images and detection algorithms detect the faces and extract face images which include mouth, nose, eyes and eyebrows [1].

Biometrics recognizes the humans in its study through one or more traits of intrinsic or physical behaviours. The biometrics is gaining considerable interest and is accepted as technology that meets with growing concerns of security in sensitive areas for example: intelligence, forensic, government and commerce. Biometrics is a self-sufficient system with capability of an individual recognition via various biometric traits that could be physiological and behavioural. The fundamental use of a system could accumulate single or a combination of

IJDACR
International Journal Of Digital Application & Contemporary Research

International Journal of Digital Application & Contemporary Research
Website: www.ijdacr.com (Volume 5, Issue 9, April 2017)

physiological and behavioural traits like fingerprint, face, hand geometry, iris retina etc.

In previous research work [3] they use PCA and Euclidean distance for face recognition which shows lower accuracy (93.57%). Present research work uses Random Forest Classifier based approach for calculating the similarity score for performance evaluation which provide improved results in terms of recognition accuracy.

II. Proposed Method

Pre-Processing of Face

1. The input photograph to this module is resized using the inbuilt resize function available in MATLAB. We have resized the image to 250×250 pixels. This is saved in JPEG data folder.
2. After resizing the RGB image is converted to a gray scale image using *rgb2gray* function in case the input image is coloured.

Face Detection by Viola Jones Method

Viola Jones is a real time face detection algorithm, developed by Paul Jones in 2001 [4]. Primarily this algorithm was developed for face detection using Haar features in a classifier. The algorithm has 4 basic steps:

- Haar Features
- Integral Image
- AdaBoost
- Cascade Classifiers

Haar Features

These are the features that are used to compute a single value. All the pixels in the black rectangle are added and all the pixels in the white rectangle are added. Now the two values are subtracted. The different types of Haar features given below:

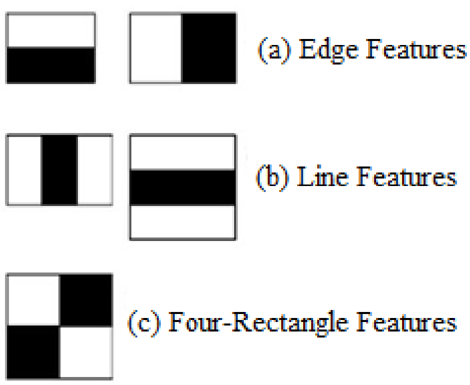


Figure 1: Haar features [4]

Integral Image

In an input image with a resolution of 224×224 pixels. This will amount to a lot of computation, thus instead of using an input image an integral image is used. An integral image is formed by adding all the pixels on the left and top of the pixel under test. This can be seen in the Figure 2.

1	1	1
1	1	1
1	1	1

Input Image

1	2	3
2	4	6
3	6	9

Integral Image

Figure 2: The Integral Image [4]

This integral image is used for calculating the sum of pixels inside any rectangle using four values. This value is calculated using the four corner values which is depicted in the Figure 3.

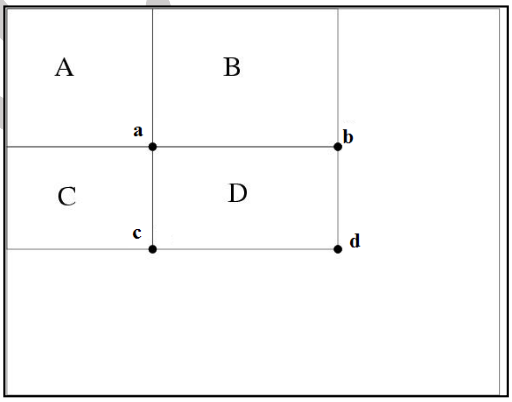


Figure 3: Four rectangles [4]

Integral Sum at $D = d - (b + c) + a$

AdaBoost

It is referred as adaptive boost. It is used to select only those features which are likely to give higher values when they match the features. This is done by linearly combining the weight of weak classifiers to form a strong classifier. The equation is given as:

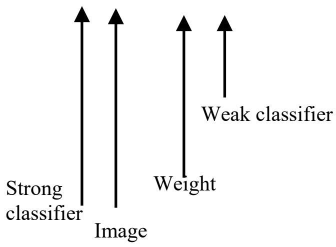
$$F(x) = \propto_1 f_1(x) + \propto_2 f_2(x) + \propto_3 f_3(x) + \cdots \quad (1)$$

A weak classifier is formed by determining the value of each feature on an image and then putting an appropriate weight to it. The mathematical description of a weak classifier is given below:

$$h(x, f, p, \theta) = \{1\ if\ pf(x) > p\theta \quad (2)$$

IJDACR
International Journal Of Digital Application & Contemporary Research

International Journal of Digital Application & Contemporary Research
Website: www.ijdacr.com (Volume 5, Issue 9, April 2017)



This explain that an image, v is classified as positive or negative by the factors as f, the feature applied, p the polarity and θ the threshold.

Cascading

The cascade classifier consists of stages, where each stage is an ensemble of weak learners. The weak learners are simple classifiers called decision stumps. Each stage is trained using a technique called boosting. Boosting provides the ability to train a highly accurate classifier by taking a weighted average of the decisions made by the weak learners.

Morphological Operations

Since we are dealing with bimodal images, a pixel could be either 1 or 0. So, a bimodal image can be thought of set of 2-tupples, in which its elements are (row, column) coordinates of the non-zero elements of the image. Formally, let I and B are the sets corresponding to the image and structuring element, then $I = \{(x, y)|I[x, y] <> 0, \forall x, y \in I_R\}$ where I_R is the set of all possible (row, column) elements over and image I. B can be defined in a similar manner. There are two basic operations in morphology, which are called dilation and erosion. The dilation of I by B is denoted by $I \oplus B$ and it is defined as [5]:

$$I \oplus B = \{c | c = i + b, where\ i \in I, b \in B\} \quad (3)$$

In the mathematics literature, the dilation is also called as Minkowski addition to refer the inventor of the operator. To complete the dilation operation, B should be translated to the every image pixel and the union of the result should be taken as an overall result. Therefore, to be precise, translation of a set should be defined to shift the structuring element to a specific image point.

The translation of set B by t is defined as follows:

$$B_t = \{c | c = b + t, \forall b \in B\} \quad (4)$$

Dilation operation is defined as follows:

$$Dil(I, B) = \cup_{t \in I} I \oplus B_t \quad (5)$$

In the same sense, erosion and erosion operations are defined as in Equations 6, 7 respectively.

$$I \Theta B = \{c | c = i - b, i \in I, b \in B\} \quad (6)$$

$$Ero(I, B) = \cap_{t \epsilon I} I \Theta B_t \quad (7)$$

Implementation of the morphological operations on images is very similar to the filtering image via given kernel. Instead of taking weighted sum, the formulas given are used to evaluate the point at kernel center. Figure 4 gives an example of dilation and erosion operations with the given structuring element.

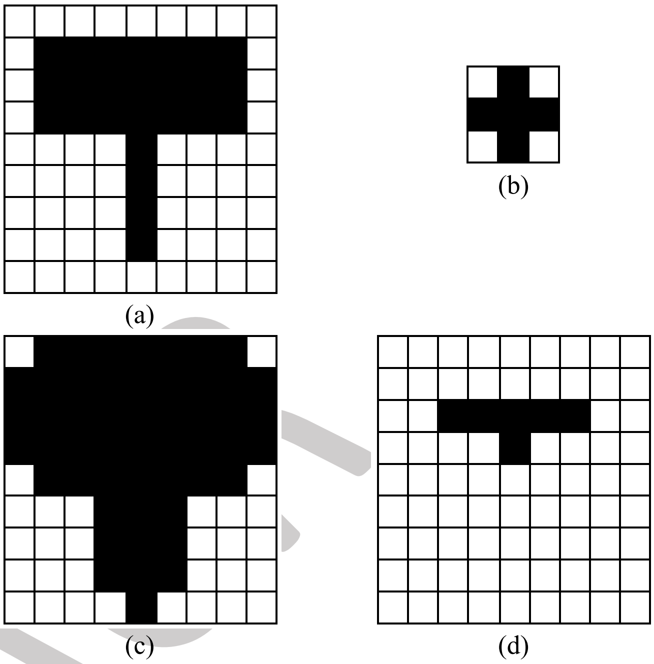


Figure 4: Example of basic morphological operations: (a) Original image (b) structuring element called simple cross, (c) dilation of (a) by (b), (d) erosion of (a) by (b) [5]

Median Filer

The best-known order-statistics filter is the median filter, which as its name implies, replaces the value of a pixel by the median of the gray levels in the neighbourhood of that pixel [6]:

$$\bar{f} = \underset{(s,t)\varepsilon s_{xy}}{median}\{g(s,t)\} \quad (8)$$

The original value of the pixel is included in the computation of the median. Median filters are quite popular because, for certain types of random noise, they provide excellent noise reduction capabilities, with considerably less blurring than linear smoothing filters of similar size. Median filters are particularly effective in the presence of both bipolar and uni-polar impulse noise [6].

Feature Extraction

The feature extraction is a very important step for the design of a biometric system. The objective of this step is to extract variables that describe, unequivocally, the forms belonging to the same class while differentiating them from the other classes. The steps involved in feature extraction are shown in the flow diagram in Figure 5. We use two feature extraction methods:

IJDACR
ISSN: 2319-4863

IJDACR
International Journal Of Digital Application & Contemporary Research

International Journal of Digital Application & Contemporary Research
Website: www.ijdacr.com (Volume 5, Issue 9, April 2017)

- Local Binary Pattern (LBP)
- Log Gabor Filter

Local Binary Pattern

The Local Binary pattern operator is a simple yet powerful texture descriptor. Due to its discriminative power and computational simplicity, LBP texture operator has become a popular approach in various applications. It is found that certain fundamental patterns in the bit string account for most of the information in the texture. These patterns have the very strong capability of rotation invariance and gray invariance; these patterns are called uniform patterns that the binary pattern contains at most two bitwise transitions from 0 to 1 or vice versa when the bit pattern is traversed circularly. Given a central pixel in the image, a pattern code is computed by comparing it with its neighbours [7]:

$$LBP_{P,R} = \sum_{p=0}^{P-1} S(g_p - g_c)2^p \quad (9)$$

$$S(x) = \begin{cases} 1, x \geq 0 \\ 0, x < 0 \end{cases} \quad (10)$$

Where g_c is the gray value of the central pixel, g_p is the value of its neighbours, P is the total number of involved neighbours and R is the radius of the neighbourhood. Suppose the coordinate of g_c is $(0,0)$, then the coordinates of g_p are $(R * cos(2\pi p/P),\ R * sin(2\pi p/P))$. Figure 5 gives examples of circularly symmetric neighbour sets for different configurations of (P,R). The gray values of neighbours that are not in the center of grids can be estimated by interpolation.

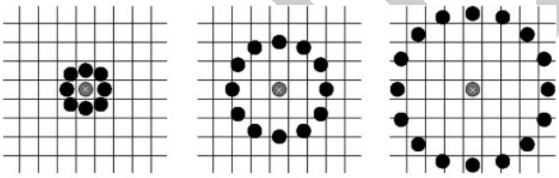


Figure 5: Circularly symmetric neighbour sets for different (P,R) [7]

Suppose the texture image is of size $N \times M$. After identifying the LBP pattern of each pixel (i,j), a histogram is built to represent the whole texture image:

$$H(k) = \sum_{i=1}^{N} \sum_{j=1}^{M} f\left(LBP_{P,R}(i,j),k\right), k\epsilon[0,K] \quad (11)$$

$$f(x,y) = \begin{cases} 1, & x = y \\ 0, & otherwise \end{cases} \quad (12)$$

Where K is the maximal LBP pattern value. The U value of an LBP pattern is defined as the number of spatial transitions (bitwise 0/1 changes) in that pattern:

$$U(LBP_{P,R}) = |s(g_{P-1} - g_c) - s(g_0 - g_c)| + \sum_{p=1}^{P-1} |s(g_P - g_c) - s(g_{p-1} - g_c)| \quad (13)$$

For example, the LBP pattern 00000000 has a U value of 0 and 01000000 has a U value of 2. The uniform LBP patterns refer to the patterns which have limited transition or discontinuities ($U \leq 2$) in the circular binary presentation. It was verified that only those "uniform" patterns are fundamental patterns of local image texture. In practice, the mapping from $LBP_{P,R}$ to $LBP_{P,R}^{u2}$ (superscript "$u2$" means that the uniform patterns have a U value of at most 2), which has $P * (P - 1) + 3$ distinct output values, is implemented with a lookup table of 2^P elements. The dissimilarity of sample and model histograms is a test of goodness-of-fit, which could be measured with a nonparametric statistic test.

Gabor Filter

An image represented in terms $f(x,y)$ where x and y signifies the coordinates of pixels having size M×N is convoluted in frequency domain. The Fourier transform is applied as the solution for this step [8]:

$$F\{f(x,y)\} = F(u,v) = \int_{-\infty}^{M-1} \int_{-\infty}^{N-1} f(x,y) e^{-j2\pi\left(\frac{ux}{M}+\frac{vy}{N}\right)} \, dx \, dy \quad (14)$$

In frequency domain the Gabor feature for an image $f(x,y)$ is the multiplication of convoluted image with Gabor filter bank $\Psi(x,y,\omega_m,\theta_n)$ given by [7]:

$$O_{m,n}(x,y) = F(u,v) * \Psi(u,v,\varpi_m,\theta_n) \quad (15)$$

Where, * is the convolution operator. The filter bank is created using m frequencies and n rotations $G(m \times n)$ that provides features points and is saved in form of vector.

Generate Feature Vector: Feature vectors are the matching templates that calculate the distance among the features of exerted information. For an M×N image with centre points (c_i, c_j), the spatial tessellation of region under interest is given by collection of sectors (S_i is the i[th] sector).

$$S_i = \{(x,y) | b(T_i + 1).r < b(T_i + 2), \theta_i \leq \theta \leq \theta_{i+1}, 1 \leq x \leq N, 1 \leq y \leq M\} \quad (16)$$

Where, (Let's say b=10 pixels, K=8 Sectors in each band) and,

$$T_i = i \ div \quad (17)$$

IJDACR
ISSN: 2319-4863

IJDACR
International Journal Of Digital Application & Contemporary Research

International Journal of Digital Application & Contemporary Research
Website: www.ijdacr.com (Volume 5, Issue 9, April 2017)

$$\theta_i = (i \bmod k)\left(\frac{2\pi}{k}\right) \quad (18)$$

$$r = \sqrt{(x - x_c)^2 + (y - y_c)^2} \quad (19)$$

$$\theta = \tan^{-1}\left(\frac{(y - y_c)}{(x - x_c)}\right) \quad (20)$$

On application of Gabor filter, the planes of Gabor response is generated is equal to the number of angles. The mean and standard deviation of Gabor response provides feature vector of single plane and thus in this manner total feature vector can be generated.

Similarity Measure using Random Forest Classifier

Random forests are recently proposed statistical inference tools, deriving their predictive accuracy from the nonlinear nature of their constituent decision tree members and the power of ensembles. Random forest committees provide more than just predictions; model information on data proximities can be exploited to provide random forest features. Variable importance measures show which variables are closely associated with a chosen response variable, while partial dependencies indicate the relation of important variables to said response variable.

The Generalization error (PE^*) of Random Forest is given as,

$$PE^* = P_{x,y}\big(mg(X,Y)\big) < 0 \quad (21)$$

Where, $mg(X,Y)$ is Margin function. The Margin function measures the extent to which the average number of votes at (X,Y) for the right class exceeds the average vote for any other class. Here X is the predictor vector and Y is the classification.

III. Simulation And Results

The performance of proposed algorithms has been studied by means of MATLAB simulation.

DATABASE: *The Japanese Female Facial Expression (JAFFE) Database:* The database contains 213 images of 7 facial expressions (6 basic facial expressions + 1 neutral) posed by 10 Japanese female models. Each image has been rated on 6 emotion adjectives by 60 Japanese subjects. The photos were taken at the Psychology Department in Kyushu University [9].

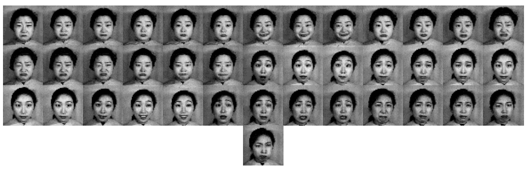


Figure 6: Test images for JAFFE (Japanese female facial expression) [9]

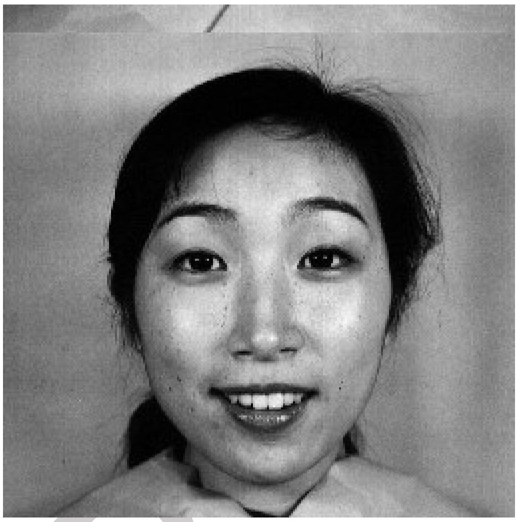


Figure 7: Input image

Figure 8: Resized to 224×224

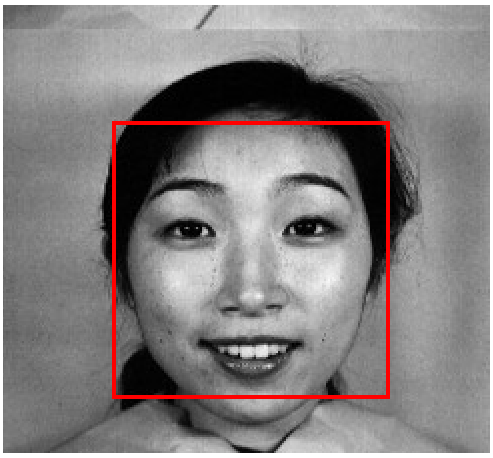


Figure 9: Face detection using Viola Jones method

IJDACR
International Journal Of Digital Application & Contemporary Research

International Journal of Digital Application & Contemporary Research
Website: www.ijdacr.com (Volume 5, Issue 9, April 2017)

Figure 10: Cropped image

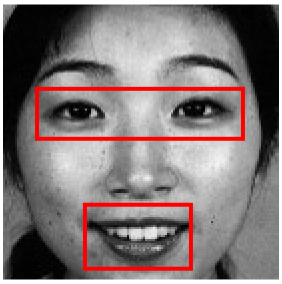


Figure 11: Eye and mouth detection

Figure 12: Cropped image

Figure 13: Resized to 110×110

Figure 14: Local Binary Pattern

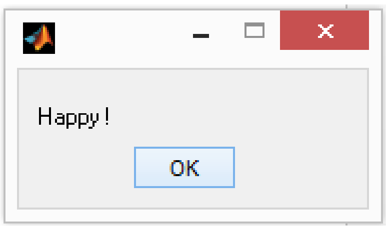


Figure 15: Result

In proposed Random Forest classifier based approach there is not any threshold value for face recognition. Random Forest Classifier itself does the similarity measure and recognizes test image. Finally, confusion matrix plot show the performance of LBP and Gabor filter based method.

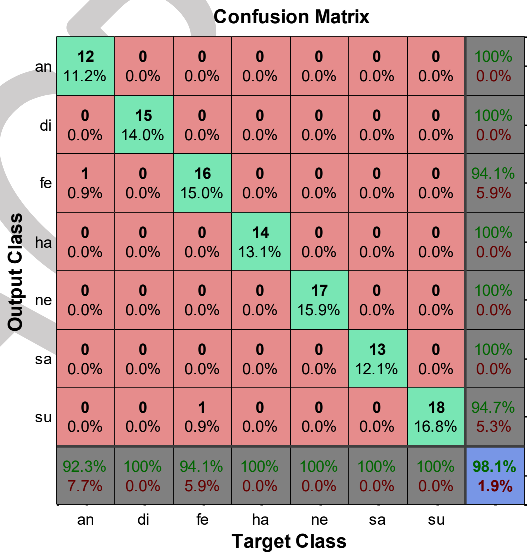


Figure 16: Confusion Matrix plot for proposed approach

The row and column are the classes of facial expression database. There are 7 sets of classes and each class having different set of expressions. The confusion matrix plot indicates the accuracy i.e. 98.1% for proposed algorithm.

Table 1: Notations for expression

S. No.	Abbreviation	Meaning
1	An	Angry
2	Di	Disgust
3	Fe	Fear
4	Ha	Happy

IJDACR
ISSN: 2319-4863

IJDACR
International Journal Of Digital Application & Contemporary Research

International Journal of Digital Application & Contemporary Research
Website: www.ijdacr.com (Volume 5, Issue 9, April 2017)

5	Ne	Neutral
6	Sa	Sad
7	Su	Surprise

Table 2: Comparison of result with previous research works

Experiments		**Recognition Rate (Accuracy In %)**						
		An	**Di**	**Fe**	**Ha**	**Ne**	**Sa**	**Su**
Previous Research	[8]	23.1	66.7	58.3	63.5	--	36.7	66.7
	[10]	96.7	82.8	84.4	83.9	93.3	83.9	80
	[11]	88	73	73	78	--	81	85
JAFFE Database Stage-I - 7 emotions	Method 1 [3]	100	90	100	95	100	75	95
	Method 2 [3]	100	85	95	90	85	80	90
	Method 3 [3]	100	80	90	95	85	70	85
JAFFE Database Stage-II - 5 emotions	Method 1 [3]	90.47	61.9	----	61.9	80.9	---	57.1
Proposed Approach: 7 emotions	**Using Random Forest Classifier**	**92.3**	**100**	**94.1**	**100**	**100**	**100**	**100**

IV. CONCLUSION

The foremost use of facial expression recognition is in security purposes, yet the technical advancements integrated this technology in present technical applications such as smile detection in camera etc. Face recognition is useful in cases when a person adopts disguise looks and makes hard for human eyes to recognize. The applications of face recognition are crucial in security aspects hence the need of this research is justified. However, the face recognition is not easy in artificial intelligence and suffers numerous challenges and the process has a specific model to follow.

This paper presents facial expression recognition system using extraction of LBP and Gabor Filter features classified by Random Forest Classifier. Confusion matrix demonstrates that the proposed random forest based approach gives more accuracy than the previous research works.

REFERENCE

[1] M.R.M. Rizk and A. Taha, "Analysis of Neural Networks for Face Recognition Systems with Feature Extraction to Develop an Eye Localization Based Method", IEEE, 2002.

[2] C. Liu and H. Wechsler: "Independent Component Analysis of Gabor Features for Face Recognition", IEEE Trans. Neural Networks, vol. 14, no. 4, pp. 919-928, 2003.

[3] Mehta, Neelum, and Sangeeta Jadhav. "Facial Emotion recognition using Log Gabor filter and PCA." In Computing Communication Control and automation (ICCUBEA), 2016 International Conference on, pp. 1-5. IEEE, 2016.

[4] Viola, Paul, and Michael Jones. "Rapid object detection using a boosted cascade of simple features." In Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on, vol. 1, pp. I-I. IEEE, 2001

[5] Shih, F. "Image processing and mathematical morphology". Boca Raton: CRC press, 2009.

[6] Sezan, M. Ibrahim, and Reginald L. Lagendijk, eds. Motion analysis and image sequence processing. Vol. 220. Springer Science & Business Media, 2012.

[7] Shan, Caifeng, Shaogang Gong, and Peter W. McOwan. "Facial expression recognition based on local binary patterns: A comprehensive study." Image and Vision Computing 27, no. 6 (2009): 803-816.

[8] Lajevardi, Seyed Mehdi, and Margaret Lech. "Facial expression recognition using neural networks and log-gabor filters." In Computing: Techniques and Applications, 2008. DICTA'08. Digital Image, pp. 77-83. IEEE, 2008.

[9] Lyons, Michael J., Shigeru Akamatsu, Miyuki Kamachi, Jiro Gyoba, and Julien Budynek. "The Japanese female facial expression (JAFFE) database." In Proceedings of third international conference on automatic face and gesture recognition, pp. 14-16. 1998.

[10] Nectarios Rose," Facial Expression Classification using Gabor and Log-Gabor Filters", IEEE 7th International Conference on Automatic Face and Gesture Recognition, 2006.

[11] Sanchez-Mendoza, David, David Masip, and Agata Lapedriza. "Emotion recognition from mid-level features." Pattern Recognition Letters 67, pp. 66-74, 2015.