IJDACR
ISSN: 2319-4863

International Journal of Digital Application & Contemporary Research
Website: www.ijdacr.com (Volume 6, Issue 11, June 2018)

Facial Expression Recognition using LBP, DWT and SVM Classifier

Kavita Suthar
M. Tech. Scholar
Shrinathji Institute of Technology and Engineering
Nathdwara (India)
kavi06124@gmail.com

Komal Paliwal
Assistant Professor
Shrinathji Institute of Technology and Engineering
Nathdwara (India)
erkomalpaliwal.cs@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 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.**

This paper analyze the hybrid approach of Local Binary Pattern (LBP) and DWT (Discrete Wavelet Transform) features for facial expression recognition. A subspace is created by this algorithm for training of feature vectors and Support Vector Machine (SVM) Classifier calculates the similarity score for performance evaluation which will provide improved results in terms of recognition accuracy.

Keywords – **DWT, JAFFE, LBP, SVM, Viola Jones.**

I. Introduction

The biometric recognition of people consists of measuring, storing and comparing specific characteristics of individuals. Among these characteristics we can highlight: the fingerprints, the facial images, the geometry of the hands and fingers, the iris, the retina, the signature and the voice. For a biometric feature to be effective in the recognition of people, the following properties should be fulfilled to the greatest possible extent: universality (for all users), singularity (different for each user), invariant (with respect to capture conditions) and resistant (to fraud attempts). The biometric method selected will depend, to a large extent, on the type of application in which it will be used and the possible acceptance it generates among its users. Thus, for example, the signature of an individual is a very socially accepted method but the techniques based on the exploration of the retina by means of a low intensity infrared beam provoke a strong distrust on the part of the users [1].

One of the reasons for developing biometric systems has been to complement the use of information known to the user (for example, a secret number or a keyword) or possessed by it (for example, a magnetic card). These traditional methods are based on properties or elements that can be lost, stolen or forgotten. Such problems disappear with the use of biometric characteristics for personal identification, because they are their own and permanent for each individual. This advantage, together with the fact that it can be extracted quickly, makes biometric techniques valuable identification methods and suitable for use in automatic systems. However, these systems currently present problems of precision (in the rates of false acceptance and false rejection), of non-viability for certain disabled individuals who do not have a certain biometric characteristic required by the system, of vulnerability in certain cases and of acceptance by the parties of the users. Despite these drawbacks, biometric recognition constitutes a good additional authentication system. Even the simplest and cheapest biometric solutions can appreciably increase the overall security of a system if they complement traditional recognition methods and adapt flexibly to the characteristics of the particular application for which they are used [2].

The face recognition though being simple to human eyes is a tedious task for computational approaches. The face recognition scheme should possess sufficient parameters to recognize a face and also robust against noise. The easy scheme of face recognition is matching the pixels of test input with database image pixels in their corresponding

IJDACR
International Journal Of Digital Application & Contemporary Research

International Journal of Digital Application & Contemporary Research
Website: www.ijdacr.com (Volume 6, Issue 11, June 2018)

position. If the total number of pixels matched is larger than the defined threshold percentage of matching, the face is considered to be authentic. But on a practical note, the input images captured are not always in standard position for matching. For example, the images from security cameras placed at higher altitude than the height of a person captures the images that have different face angles. In these input images, the faces are tilt and could not be recognized with easy schemes.

Another problem that arises in face recognition are the quality of images that are given as input. The captured images have variation in actual size of face due to distance from where they are captured and also have varied skin colour tone due to sunlight [1]. Since the quality cannot be assured for input image, the algorithms are expected to perform with high rate of accuracy in given conditions. The solution for first consideration is a secondary step, primary step being the angle of face captured and the computational model used. Since at most of places only 2D face recognition models are installed, hence the accuracy of face recognition is subjected to angle of face or in high rotational cases is dependent of efficiency of observer. The noisy images in 2D computational model, the efficiency of mathematical algorithms are studied in terms of accuracy of detection [2]. Though many researches claim their methods to be robust and having high efficiency, the assumptions they make are not validated in practical world.

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

II. Proposed Method

Face recognition using extraction of LBP (Local Binary Pattern) and DWT (Discrete Wavelet Transform) features is proposed in this heading. In domain of inductive inference problem, source separation could be a challenging task. As to derive the solution one needs the sufficient information, the available information is exploited in maximum limits. The adaptive systems tend to inherit most of the available feature information to replicate the original set of input with elaborated clarity. The efficiency of an algorithm is subjected to its performance in case when evaluation parameters reflect sound values even in cases of noise, orientation and luminance conditions.

Figure 1 shows the flow diagram of proposed research work. The process of face recognition is a sequential task. The methods of face recognition are generally studied in three domains that are classified based on their approach. The template matching methods identifies the group of pixels in test image that resembles with the template image.

The geometrical local-feature-based methods concentrate on geometries of faces and select the relative sets of features for matching. Hybrid methods is the smart acquisition of template based approaches. Though the schemes are successful in their part of applications with supporting assumptions, they do possess certain advantages and disadvantages. One can say that the selection of algorithm should be subjected to the required task in a given application.

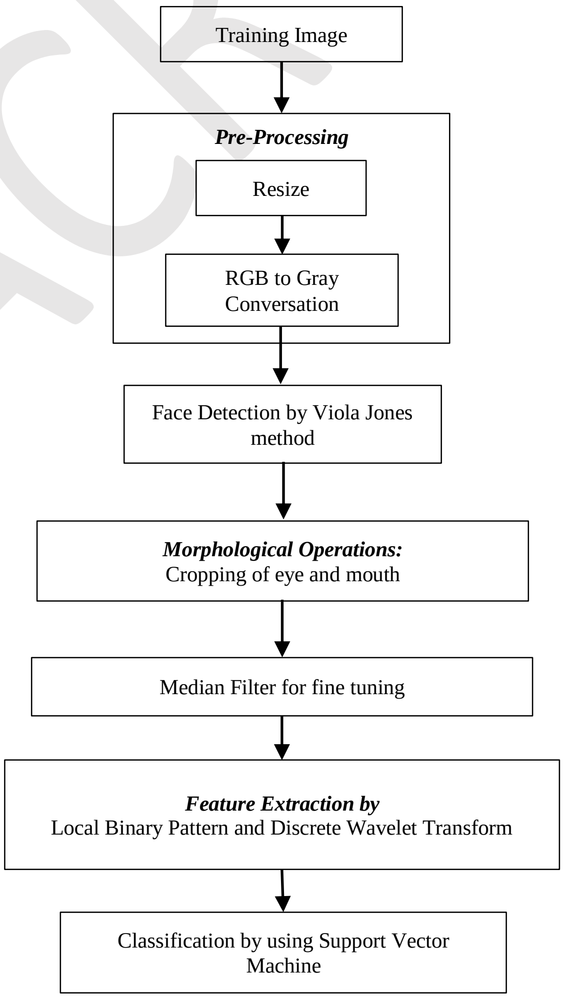


Figure 1: Flow diagram for proposed approach

Rest of methodology is explained as follows:

IJDACR
ISSN: 2319-4863

IJDACR
International Journal Of Digital Application & Contemporary Research

International Journal of Digital Application & Contemporary Research
Website: www.ijdacr.com (Volume 6, Issue 11, June 2018)

A. Pre-Processing of Face

Algorithm-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.
- After resizing the RGB image is converted to a gray scale image using rgb2gray function in case the input image is coloured.

B. The Viola Jones Method for Face Detection

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: 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:

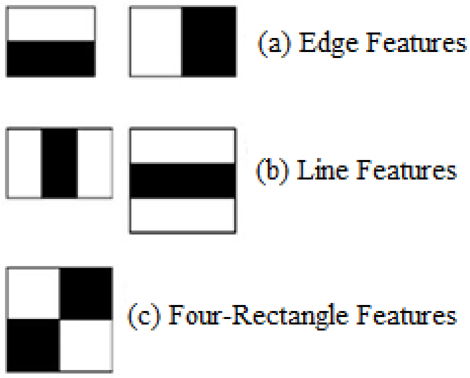


Figure 2: Haar features [4]

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 3.

1	1	1
1	1	1
1	1	1

Input Image

1	2	3
2	4	6
3	6	9

Integral Image

Figure 3: 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 4.

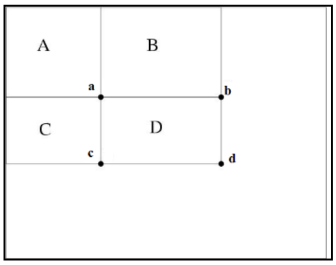
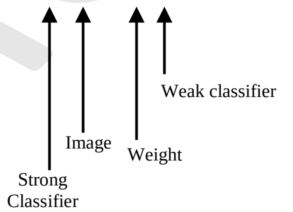


Figure 4: Four rectangles [4]

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:

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



(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:

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

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: Cascading is accomplished using cascade classifier which contains several phases, where each phase is a group of weak learners. This weak learners are trained by using boosting. Boosting provides the ability to train a highly accurate classifier by taking a weighted average of the decisions made by the weak learners.

C. Morphological Operations

Dilation: Dilation, also called expansion, filling, or growth, produces a thickening effect on the edges of the object. This algorithm is used to increase the contour of the objects and to join the discontinuous

IJDACR
ISSN: 2319-4863

International Journal of Digital Application & Contemporary Research
Website: www.ijdacr.com (Volume 6, Issue 11, June 2018)

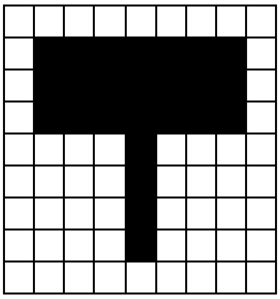
lines of these, produced by some filtration, mathematically the binary dilation is defined as:

$$A \oplus B = \{c \in E^{N|\,C=a+}b \text{ for all } a \in A \text{ and } b \in B\} \quad (3)$$

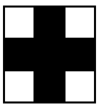
Erosion: Erosion is the dual function of expansion, but it is not the reverse, i.e. if erosion is done and then a dilation the resulting image will not be equal to the actual image, mathematically erosion is defined as:

$$A \ominus B = \{x \in E^N | x + b \in A \text{ for all } b \in B \quad (4)$$

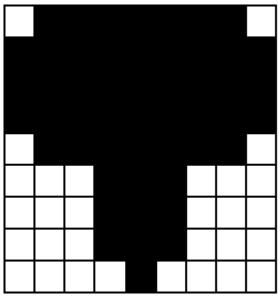
Figure 5 gives an example of dilation and erosion operations with the given structuring element.



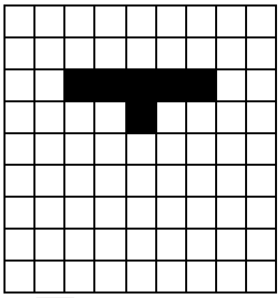
(a) Original image



(b) Structuring element called simple cross



(c) Dilation of (a) by (b)



(d) Erosion of (a) by (b)

Figure 5: Illustration of morphological operations [5]

D. Median Filter

Consider a discrete image F characterized by a gray level $f(x, y)$. Let $V(x_0, y_0)$ be the neighborhood associated with the coordinate point (x_0, y_0); it is assumed that this neighborhood has N coordinate pixels $(x_0 - u, y_0 - v)$ with odd N.
Let $\{f_1, f_2, \dots, f_i, \dots, f_{N-1}, f_N\}$ the gray levels associated with the N pixels of $V(x_0, y_0)$.

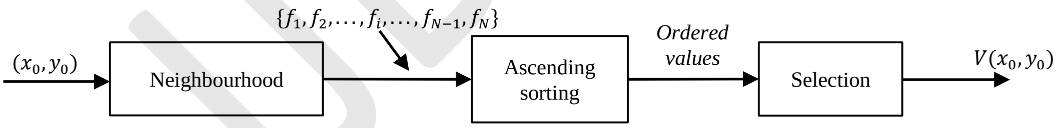


Figure 6: Median filter

The median filtering first proceeds by sorting the gray level values of the neighborhood followed by a selection of the middle element of the sorting.
Sorting is done in ascending order generally. It leads to the ordered set of gray values of the neighborhood of $f(x_0, y_0)$. Since the ordered elements are denoted by f_i, the ascending sort is characterized by:

$$f_q < f_2 < \cdots < f_{\frac{N+1}{2}} < \cdots < f_{N-1} < f_N \quad (5)$$

The middle element of the neighborhood is $f_{\frac{N+1}{2}}$. Its property is to be preceded by $\frac{N-1}{2}$ lower values and followed by as many higher values.
The filtering consists of replacing $f(x_0, y_0)$ by the median value of the neighborhood $f_{\frac{N+1}{2}}$ [6].

E. Feature Extraction

The purpose of feature extraction in the field of recognition is to express the feature in numerical or symbolic form called encoding. Depending on the case, the values of these features can be real, integer or binary. The vector composed of feature n represents a point in the new space of n dimensions. The steps involved in feature extraction are shown in the flow diagram in Figure 1. We use two feature extraction methods:
Local Binary Pattern (LBP): A face is separated into small region for computation of LBP for every region image pixel, further histogram of the LBP, is considered as feature vector of facial image. Let us N to form a large histogram representing the image

IJDACR
ISSN: 2319-4863

IJDACR
International Journal Of Digital Application & Contemporary Research

International Journal of Digital Application & Contemporary Research
Website: www.ijdacr.com (Volume 6, Issue 11, June 2018)

of facial features (see Figure 7) [7]. Efficacy of the LBP code as a facial index is explained by the fact that the LBP allows to characterize the details of a face. When only uniform LBPs are used, all non-uniform LBPs are labelled with a single label, while each of the uniform codes is grouped in a single histogram. For example, when P=8, we have 58 uniform codes but the histogram is of dimension 59. Similarly P=6 produces a histogram of dimension 33.

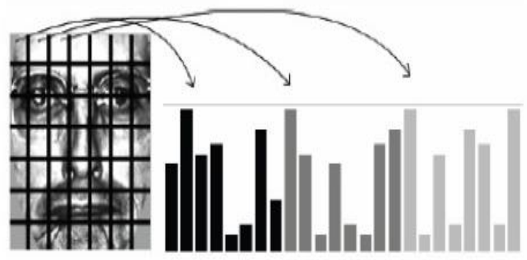


Figure 7: Representation of a face by the histograms of the LBP code [7]

Given two histograms of LBP H^1, H^2 of two faces, the subsequent phase is to use a metric to compute the similarity between these two histograms. In testing the three metrics χ^2, Histogram intersection and Log-likelihood statistic:

$$\chi^2(H^1, H^2) = \sum_i \frac{(H_i^1 - H_i^2)^2}{H_i^1 + H_i^2} \quad (6)$$

Discrete Wavelet Transform (DWT): DWT employees Fourier transform to convert time domain image into frequency domain. The mathematical expression of DWT is given by:

$$DWT_{x(n)} = \begin{cases} dd_{j,k} = \sum img(n) hh^*_s(n - 2^s r) \\ ap_{j,k} = \sum img(n) ll^*_s(n - 2^s r) \end{cases} \quad (7)$$

Where, $dd_{j,k}$ represents detailed coefficients.
$ap_{j,k}$ are the approximate coefficients of DWT transform.
$hh(n)$ are high pass filter functions.
$ll(n)$ low pass filter functions.
s is wavelet scale parameters.
r is translation factor.

F. Similarity Measure using Support Vector Machine

SVM is a type of trainable classifier. The advantages of SVM when it is used in image classification problems are, first of all, its ability to work with data of very high dimension and secondly, its great power of generalization without the need to add a priori knowledge, even when counting with scarce training samples. Given an input pattern, an SVM classifier approximates the probability that this pattern belongs or not to a certain class, by means of functions of the input data.
The construction of an SVM classifier consists of finding an optimal hyperplane that separates the space into two parts, one for each of the two classes of a binary classification. This will serve to correctly classify input vectors (points in the input space) separating the points of two classes as much as possible. This type of classifiers can be extended to multiclass problems. Figure 8 shows an example in which given two classes, the objective is to identify the hyperplane that maximizes d (margin between classes) and at the same time correctly classify all the given examples. When the set of training examples is linearly separable, SVM provides the hyperplane that separates the two classes so that the distance between the hyperplane and the closest example of each class is maximum that is, the hyperplane is optimal. This distance is called margin and is interpreted as the maximum space between two classes, which does not contain any sample.

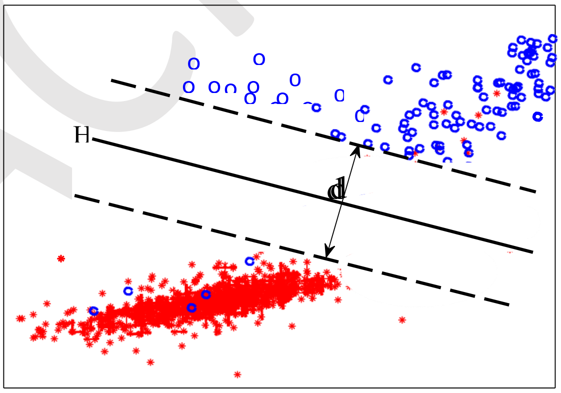


Figure 8: The optimal hyperplane H (in bold) with the maximum margin d [8]

Mathematical Foundation of SVMs

SVMs can be used to solve binary discrimination problems, that is, decide which class a sample belongs to. The solution to this problem is the construction of a function f which, at an input vector $x \in X$ matches an output $f(x)$: It is then decided that x is of class $+1$ if $f(x) > 0$ and of class -1 if $f(x) < 0$. It is a linear classifier. The decision boundary $f(x) = 0$ is a separating hyperplane.
Let H be a hyperplane, w its normal vector, and b its offset from the origin (see Figure 9). The hyperplane H is then given by:

$$f(x) = w^T x + b = 0 \quad (8)$$

The goal of the SVM learning algorithm is to find the parameters w and b of the best hyperplane through a learning set:

IJDACR
ISSN: 2319-4863

IJDACR
International Journal Of Digital Application & Contemporary Research

International Journal of Digital Application & Contemporary Research
Website: www.ijdacr.com (Volume 6, Issue 11, June 2018)

$$\mathcal{X} \times \mathcal{Y} = \{(x_1, y_1), \dots, (x_i, y_i)\} \in \mathbb{R}^N \times \{+1, -1\} \quad (9)$$

Where, y_i are the respective labels of x_i, N is the size of the learning set.

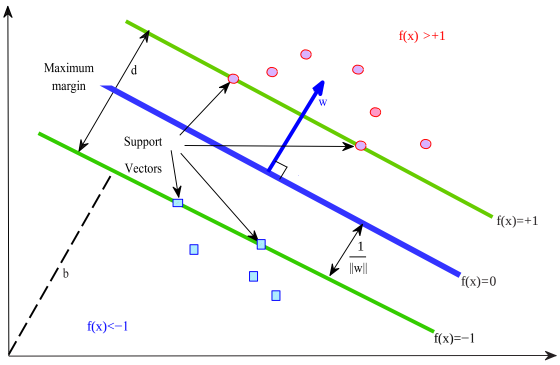


Figure 9: Margin illustration and support vectors [9]

III. Simulation And Results

The simulation is carried out by using image processing toolbox of MATLAB software.

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 [10].

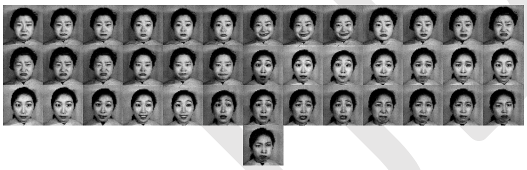


Figure 10: Test images for JAFFE (Japanese female facial expression) [10]

Figure 11: Input image

Figure 12: Resized to 224×224

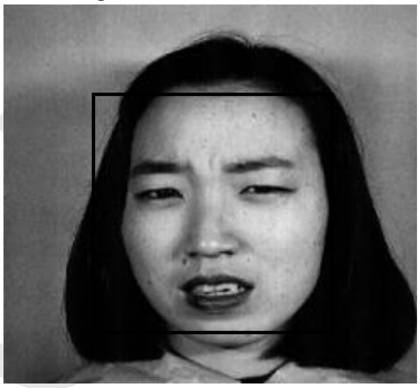


Figure 13: Face detection using Viola Jones method

Figure 14: Cropped image

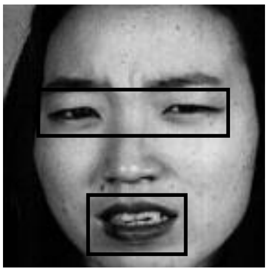


Figure 15: Eye and mouth detection

Figure 16: Cropped image

IJDACR
ISSN: 2319-4863

IJDACR
International Journal Of Digital Application & Contemporary Research

International Journal of Digital Application & Contemporary Research
Website: www.ijdacr.com (Volume 6, Issue 11, June 2018)

Figure 17: Resized to 110×110

In proposed support vector machine classifier based approach there is not any threshold value for face recognition. SVM itself does the similarity measure and recognizes test image. Finally, confusion matrix plot show the performance of LBP and DWT based method.

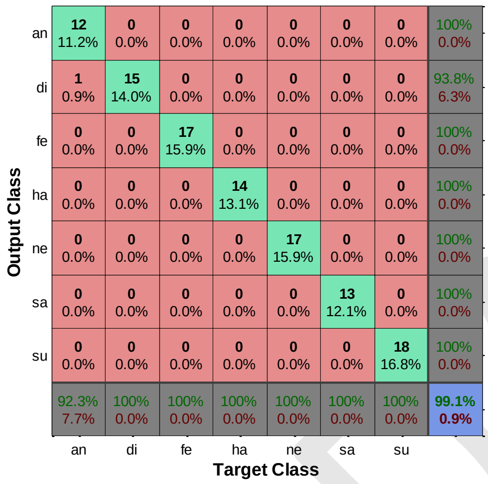


Figure 18: Confusion Matrix plot for proposed approach

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. 99.1% for proposed algorithm.

Table 1: Notations for expression

S. No.	Abbreviation	Meaning
1	An	Angry
2	Di	Disgust
3	Fe	Fear
4	Ha	Happy
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	[11]	23.1	66.7	58.3	63.5	--	36.7	66.7
	[12]	96.7	82.8	84.4	83.9	93.3	83.9	80
	[13]	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 Support Vector Machine Classifier**	**92.3**	**100**	**100**	**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 DWT features classified by Support Vector Machine Classifier. Confusion matrix demonstrates that the proposed SVM based approach gives more accuracy than the previous research works.

REFERENCE

[1] Meher SS, Maben P. “Face recognition and facial expression identification using PCA”, IEEE International Conference In Advance Computing (IACC), pp. 1093-1098, 2014.

[2] G. Wang and Z. Ou, “Face Recognition Based on Image Enhancement and Gabor Features”, Proceedings 6th World Congress on Intelligent Control and Automation, (2006), pp. 9761-9764.

[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] Li, Li. "Support vector machines." In Selected applications of convex optimization, pp. 17-52. Springer, Berlin, Heidelberg, 2015.

IJDACR
International Journal Of Digital Application & Contemporary Research

International Journal of Digital Application & Contemporary Research
Website: www.ijdacr.com (Volume 6, Issue 11, June 2018)

[9] Guenther, Nick, and Matthias Schonlau. "Support vector machines." Stata J 16, no. 4 (2016): 917-937.

[10] 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.

[11] 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.

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

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