2014 International Joint Conference on Neural Networks (IJCNN)
July 6-11, 2014, Beijing, China

Local Binary Pattern based Facial Expression Recognition using Self-organizing Map

Anima Majumder*, Laxmidhar Behera and Venkatesh K. Subramanian

***Abstract*—This paper presents an appearance feature based facial expression recognition system using Kohonen Self-Organizing Map (KSOM). Appearance features are extracted using uniform Local binary patterns (LBPs) from equally sub-divided blocks applied over face image. The dimensionality of the LBP feature vector is further reduced using principal component analysis (PCA) to remove the redundant data that leads to unnecessary computation cost. Using our proposed KSOM based classification approach, we train only 59 dimensional LBP features extracted from whole facial region. The classifier is designed to categorize six basic facial expressions (happiness, sadness, disgust, anger, surprise and fear).**

To validate the performance of the reduced 59 dimensional LBP feature vector, we also train the original data of dimension 944 using the KSOM. The results demonstrates, that with marginal degradation in overall recognition performance, the reduced 59 dimensional data obtains very good classification results. The paper also presents three more comparative studies based on widely used classifiers like; Support vector machine (SVM), Radial basis functions network (RBFN) and Multi-layer perceptron (MLP3). Our KSOM based approach outperforms all other classification methods with average recognition accuracy 69.18%. Whereas, the average recognition rated obtained by SVM, RBFN and MLP3 are 65.78%, 68.09% and 62.73% respectively.

***Index Terms*—Self-organizing map, Facial expression recognition, Local binary patterns, Principal component analysis, Support vector machine, Multi-layer perceptron, Radial basis function.**

I. Introduction

A recent challenge in designing computerized environments is to keep the human user at the core of the system. To have a truly affective human-computer intelligent interaction systems (HCIIs), the computer needs to be capable of interacting with the user in a natural way (the way in which two person interacts with each other). To recognize an emotional state, HCIIs should interpret non-verbal behavior like: voice, body gesture and facial expressions. Among three, facial expression is the most natural means of communicating human emotions, intentions and opinions to each other. [1] showed in their research work, that 55% of the emotional information is conveyed by facial expression alone. Remaining voice tone and spoken words conveys 38% and 7% of the information respectively. [2] did a psychological research on facial expression and they concluded that there are six basic facial expressions which are universal. The six basic expressions include happiness, sadness, disgust, anger, surprise and fear. Facial expression recognition system can broadly be categorized into three basic parts: Face detection, features extraction and classification.

The organization of the paper is as follows: section II presents related work done during last few years, section III demonstrates a brief overview of the proposed approach, section IV shows the method of extracting LBP features from an image, section V gives an explanation about how LBPs features are extracted using sub-blocks. Section VI explains the motivation behind using SOM based facial expression recognition system and the fundamental theory behind it, section VII depicts all the confusion matrix of recognition performances obtained using the proposed KSOM based classifiers and 3 other widely used classifiers, such as SVM, RBFN and MLP3. Finally in VIII conclusions are drawn.

II. Previous work

Researchers are rigorously working in this field for decades. Numerous work had been done in facial expression analysis and recognition [3], [4], [5], [6], [7], [8], [9], [10], [11], [12], [13], [14], [15].

The two most challenging aspects of facial expression recognition system are to get the relevant facial features and to apply the classifier which can best describe those features. Generally, there are two common types of facial features extraction techniques: geometric based methods and appearance based methods. Geometric based approach mainly depends on how we perceive an expression; i.e, based on shrinking/ widening of eyes, eyebrows, lips, nose etc. This approach mainly focuses on finding out the displacement/ angular changes of the feature points with respect to a neutral expression. But, the main drawback with this technique is it's sensitivity noise and accumulation of errors while tracking those features which makes the accurate detection of those feature points quite difficult. Whereas, appearance features are less sensitive to noise and can encode micro patterns present in the skin texture of face which are very important information for expression recognition.

One of the vital issues in automatic facial features extraction method is to represent visual information that can reveal the

Anima Majumder is doctoral student, Department of Electrical Engineering, Indian Institute of Technology Kanpur, PIN 208016, Uttar Pradesh, India (email: animam@iitk.ac.in)

Laxmidhar Behera is professor, Department of Electrical Engineering, Indian Institute of Technology Kanpur, PIN 208016, Uttar Pradesh, India (email: lbehera@iitk.ac.in)

Venkatesh K. Subramanian is professor, Department of Electrical Engineering, Indian Institute of Technology Kanpur, PIN 208016, Uttar Pradesh, India (email: venkats@iitk.ac.in)

slightest facial movement usually occurs due to change in expression. Some of the attempts those were made during the successive years are: Optical flow analysis [16], Local binary patterns (LBPs)[17], Level Set [15], Active appearance model (AAM) [16], Geometric analysis of facial features [18].

Though, methods like AAMs and ASMs are recognized to be generating accurate facial features detection results, but they have one major drawback. Such model based methods need prior information about the shape features. Mostly the initial facial feature points' locations are marked manually. [17] proposed Local binary Patterns, which is a powerful means of texture description. Most important properties of LBPs are its robustness to change in illumination and computational simplicity. [12] found appearance based features by dividing the face image into sub-blocks.

Previous studies reports, that a facial expression recognition system needs a classifier which can perform well to classify six basic facial expressions (Happiness, Sadness, Disgust, Anger, Surprise and Fear). The powerful classifiers those are generally applied in facial expression recognition systems are: Multi Layer Perceptron (MLP3) [19], Radial basis function network (RBFN) [20], [21], Support Vector Machine (SVM) [22] and rule based classifiers [11] etc.

III. Overview of the proposed approach

We extract the face from the given image using Viola Jones' face detection algorithm [23] [24]. We further process the face image to crop out the unnecessary regions such as: hairy regions, regions near to ears, that generally leads to extra computation cost. The face image is resized into a standard image size 240×240 using linear interpolation. The re-sized image is sub-divided into 16 equally placed sub-blocks each of size 60×60. We extract uniform LBP features of dimension 59 from each of the sub-blocks and concatenate them in order, resulted into 944 dimensional feature vector that retains location ans structural information of facial features. We apply PCA based data dimensionality reduction approach over extracted 944 dimensional data and found that only 59 principal components are providing the significant contribution in classification process; rest are mostly redundant data. The most significant 59 dimensional eigen vectors are used to train our KSOM based classification approach. The KSOM [25] has an inherent property of clustering the data in an order that preserves the topology of input data. This characteristics of KSOM helps in arranging the data with small change in feature space to get clustered into closer zones. As there is a direct relationship between the small change in expressions and small change in feature zones, this property of KSOM in terms helps to have a better facial expression classification approach.

In this paper we compare our approach with SVM, RBFN and MLP3.

IV. Local Binary Patterns

The original Local Binary Pattern (LBP) was introduced by [17]. The basic assumption was that texture has locally two complementary aspects: pattern and strength. [26] [24] Original version works for 3×3 pixel operator size. The operator within the block of size 3×3 pixels thresholds the neighborhood pixels based on the value of center pixels.

For a gray scale image $I(x, y)$, assume that the gray level at an arbitrary location $x, y)$ be given as g_c. I. e., $g_c = I(x, y)$. For an evenly space circular neighborhood with P sampling points and R radius around the center pixel (x, y), the gray value of pixel at p^{th} sampling point is g_p, given by

$$g_p = I(x_p, y_p), \qquad p = 0, \dots, P-1 \tag{1}$$

and the coordinate values of the sampling point is defined as

$$x_p = x + R\,cos(2\pi p/P), \tag{2}$$
$$y_p = y - R\,sin(2\pi p/P). \tag{3}$$

The thresholding binary operator $S(g_p - g_c)$ can be given as

$$S(g_p - g_c) = \begin{cases} 1 & \text{if } g_p \geq g_c \\ 0 & \text{otherwise.} \end{cases}$$

In basic LBP, the LBP_{PR}, decimal equivalent of the P bit binary operator is computed by applying a binomial factor to each of the $S(g_p - g_c)$. It is calculated as

$$LBP_{P,R}(x_c, y_c) = \sum_{p=0}^{P-1} S(g_p - g_c)2^p. \tag{4}$$

An example of a basic LBP operator is demonstrated in Fig. 2. The neighborhood pixels are threshold based on the value of center pixel. The threshold pattern gives 8 bit binary value, which is further converted to decimal equivalent value. Thus, the $LBP_{8,1} = 145$.

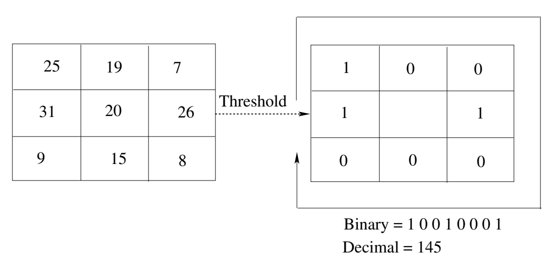

Fig. 2: An example of basic LBP operation showing the generation of decimal value after thresholding the neighborhood pixels.

The main motivation behind using LBP patterns is its robustness to change in illumination and computational simplicity. Moreover, the operator is capable of detecting texture primitives, such as spot, line end, edge and corner are detected by operators. An example of the texture premitives those can be detected by LBP operator is shown in Fig. 3.

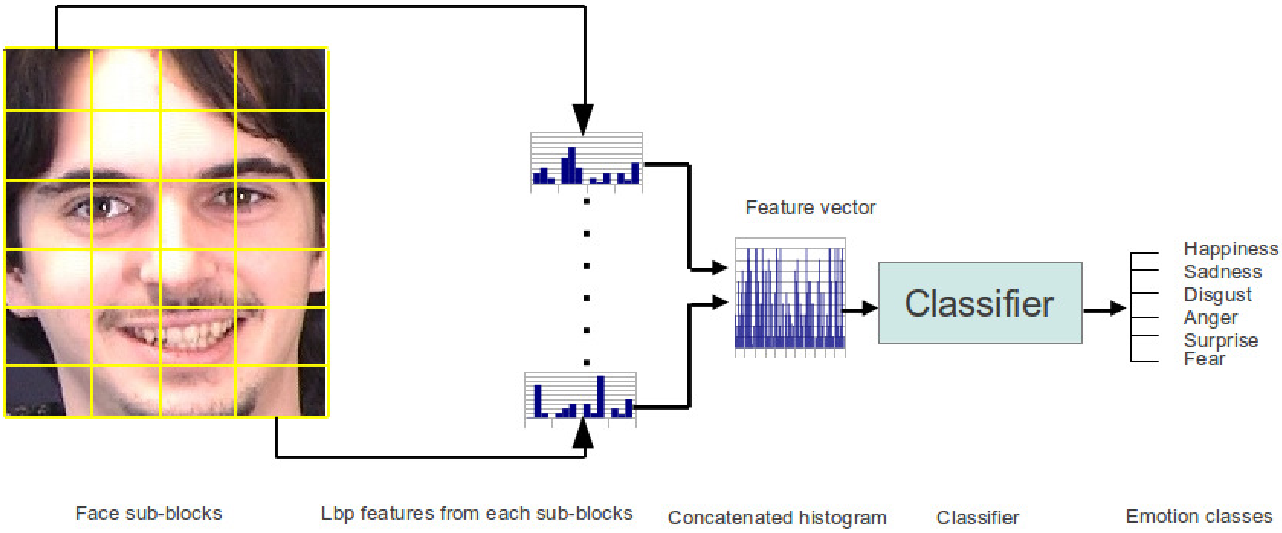

Fig. 1: An example of sub-block based classification technique using concatenated LBPs obtained from each sub-block.

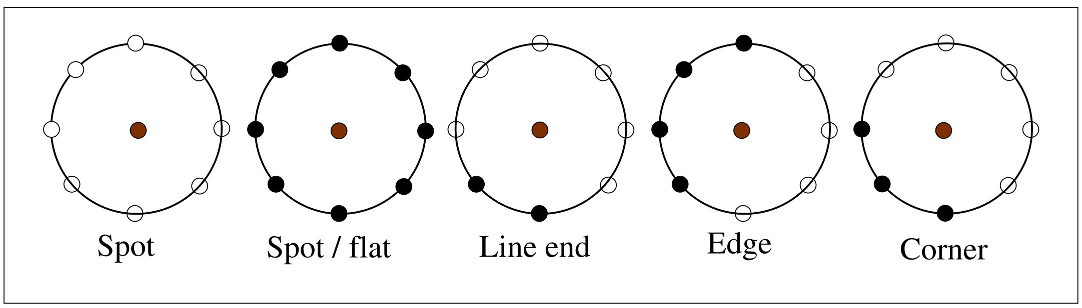

Fig. 3: Figure showing the texture primitives those are detectable by using LBP. circles with white color shows ones and the circles in black shows zeros.

[17] proved that only a subset of the 2^P patterns extracted using basic LBP operator is sufficient enough to describe most of the texture information within the image.

In this work we use uniform patterns in-stead of basic LBP patterns. A uniformity measure of a local binary pattern U is the number of bit-wise transitions from 0 to 1 for a circular bit pattern. The uniformity measure of U pattern is at most 2. As an example, for $LBP(8, 1)$ the patterns 00000000, 11111111 falls under category of zero transition. and patterns 00001111 falls under category of one transition and patterns 00111000 falls under two transition. [17] observed that for textured images, more than 90% of the information conserved in 8, 1) neighborhood. The total number of labels in uniform $LBP^{u2}_{8,2}$ is 59 (58 labels for uniform patterns with at most transition 2 and the patterns with transitions $U(x) \geq 2$, called non-uniform patterns are assigned to a single label, totally 59 labels in $LBP^{u2}_{8,1}$).

The histogram of the LBP image $LBP(x, y)$ can be computed as

$$H_k = \sum_{x,y} S(LBP(x, y) = k, \quad k = 0, \ldots, n-1 \tag{5}$$

where n is the number of different labels produced by LBP operator. In this case n is equal to 59 for uniform LBP.

V. Sub-block based LBP features extraction and dimensionality reduction

When LBP histogram of dimension 59 is extracted from the whole face image, the features encodes only the presence of micro patterns. The patterns contain no information about it's location. Face image of size 240×240 is equally sub-divided into blocks of size 60×60, totally 16 sub-blocks. Each sub-block $(R_1, \ldots R_{15})$ contributes uniform LBP feature vector of dimension 59. When all the feature vectors are concatenated in order, we obtain a new feature vector of dimension 944, that preserves location and shape information of the facial region.

The concatenated histogram can be defined as

$$H_{k,r} = \sum_{x,y} S\{LBP(x,y) = k\} \quad LBP\{(x,y) \in R_j\} \tag{6}$$

where $k = 0, \ldots, 58$, $j = 0, \ldots, 15$.

To check redundancy among data and remove them, we perform principal component analysis over the 944 dimensional LBP feature vector. It is observed that only first 59 principal components are contributing significant information and rest are found to be of almost zero variance. We keep only the first 59 eigen vectors and discard the rest.

VI. Expressions Recognition using KSOM

Kohonen Self Organizing Map (KSOM) [25] has an inherent property of clustering the data in an order that preserves the topology of the input vector; i.e, the data with small changes gets clustered in closer zones. This property in terms helps in better classification of facial expression data, as similar data with small change in expression gets clustered near to each other.

Figure 4 shows the flow diagram of the SOM based facial expression recognition system. Each frame from every video clip is processed separately. Face is detected initially followed by face normalization after cropping and removing some unnecessary facial regions. We resize the face image to 240×240 pixel resolution. The image is then sub-divided into 16 equal non-overlapping blocks with each block size 60×60 pixels. We apply LBP filter of size 3×3 on each block and obtain corresponding histogram images. The LBP features are concatenated in order to obtain 944 dimensional feature vector. The feature vector is further processed by PCA that keeps only 59 principal components which are found to be retaining significant information needed for expressions recognition. The feature vector $\mathbf{x} \in \mathbb{R}^{59}$ is thus given as input to the KSOM network of size 10×8. The Figure 5 demonstrates a pictorial overview of the parameters mapping from lattice space to the emotion space. A linear mapping is established between input and output space. Since we need six distinct classes as output classes (Happiness, Sadness, Disgust, Anger, Surprise and Fear), we set the desired output to be either 1 or -1 format. I.e, if the desired output is happiness, we set it as $\{1\ -1\ -1\ -1\ -1\ -1\}$. 1 represents true state of the class occurrence and -1 represents the false state.

A. SOM based model

The SOM network updates 3 parameters: a weight vector $\mathbf{W}_j = [w_{j,1}, w_{j,2}, \ldots, w_{j,59}] \in \mathbb{R}^{59}$, matrix $\mathbf{A}_j \in \mathbb{R}^{6\times 59}$ and $\mathbf{b}_j \in \mathbb{R}^6$ a bias vector parameter. Here j denotes the j^{th} node in the KSOM network. The work done by [27] gives a complete explanation about the KSOM based model and update rules used for all three parameters. For a KSOM network, the entire network participates in generating output for a given input vector $\mathbf{x} \in \mathbb{R}^{59}$.

For each node j ranges from $1, \ldots, (10 \times 8)$, the the output vector $\mathbf{z}_j(n) \in \mathbb{R}^6$ at iteration n is given as

$$\mathbf{z}_j(n) = \mathbf{b}_j(n) + \mathbf{A}_j(n)[\mathbf{x} - \mathbf{w}_j(n)] \tag{7}$$

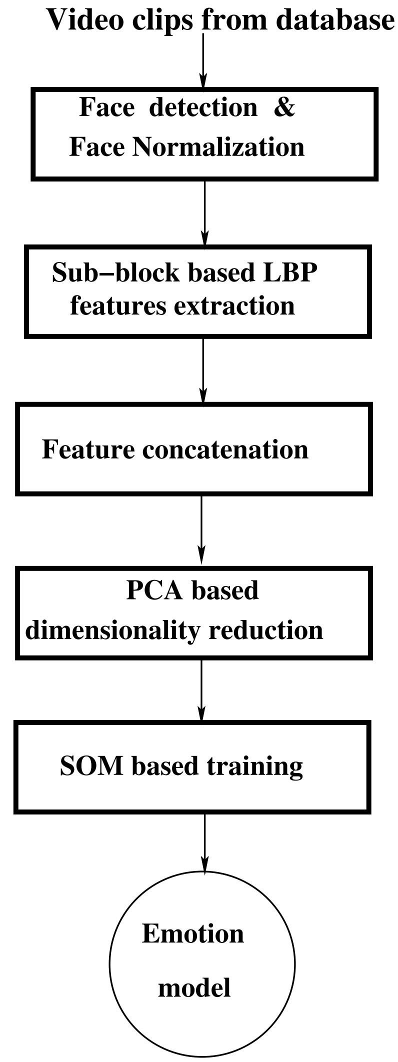

Fig. 4: System diagram of the proposed KSOM based training approach

and the output of entire network, which is an weighted average of the whole network, is given as

$$\mathbf{z}(n) = \frac{\sum_{j=1}^{M\times N} h_{j,i}(n)\mathbf{z}_j(n)}{\sum_{j=1}^{M\times N} h_{j,i}(n)} \tag{8}$$

where

$$h_{j,i}(n) = exp(-\frac{d_{j,i}^2}{2\sigma^2}) \tag{9}$$

is the neighborhood function that calculates the weight age of the neighboring node at each iteration using euclidean distance $d_{j,i}$ from the best matching unit (BMU).

The BMU is the node in the lattice structure whose euclidean distance from the input vector $\mathbf{x}$ is smallest.

Since the output is discrete in nature, we pass the output of the SOM network through a sigmoid function. Please refer [27] for the parameters' update rules.

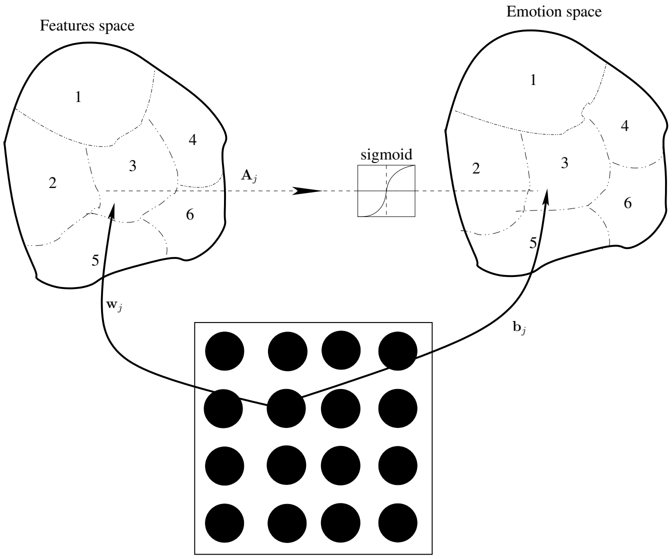

Fig. 5: Pictorial description of a 2 D KSOM network for input vector $\mathbf{x}$. This picture is taken from [27].

VII. Experimental Results

Publicly available MMI facial expression database [28] is used for our experiments. We used 81 different video clips taken from MMI database. The video clips contains 12 different characters where each character shows each of the six basic expressions separately. Among the data 60% of the data is taken for traning and remaining 40% is used for testing. We extract LBP features of dimensions 944. the features are passed through PCA to remove the unnecessary or redundant data. The reduced LBP features $\mathbf{x} \in \mathbb{R}^{59}$ and the original concatenated LBP features $\mathbf{x} \in \mathbb{R}^{944}$ are used for training KSOM network. Table I shows the confusion matrix for all the six basic emotions: (happiness (H), sadness (Sa), disgust (D), anger (A), surprise (Sur), fear (F)) where the feature vector $\mathbf{x} \in \mathbb{R}^{944}$ is trained as tested using KSOM. We observed that the average recognition accuracy obtained in this case is 71.21%. Whereas, when we train the KSOM network with $\mathbf{x} \in \mathbb{R}^{59}$ (after applying PCA), the average recognition performance is found to be 69.18%, which is just marginally less then that of the previous one. Table II shows the confusion matrix for KSOM based classification approach using feature vector $\mathbf{x} \in \mathbb{R}^{59}$.

The widely used classifiers: SVM, RBFN and MLP3 are used to compare with our recognition performance. feature vector $\mathbf{x} \in \mathbb{R}^{59}$ is used for each of the three case. We use multi-class libsvm library to train SVM classifier. Libsvm uses one against one classification technique. The kernel type used in this case of RBF as it is found to be giving better performance compared to the other kernels like; polynomial and lineal kernels. The radius of the kernel function is taken as 0.1 as it is found (by trial and error) to be giving best recognition accuracy. A parametric comparison among all the 4 classification techniques is shown in Table VI. Table III shows the confusion matrix generated after SVM based classification. An average recognition accuracy of 65.78% is calculated in this case. The recognition accuracy of other two neural network based classifiers: RBFN and MLP3 are shown in the Tables IV and V respectively. It is observed that RBFN outperforms among the three classifiers with average rate of recognition as 68.09%. Overall we see that, the proposed KSOM based method performs better than all other approaches with highest recognition accuracy 84% and average accuracy 69.18%.

VIII. Conclusions

The paper presents a new KSOM based facial expression recognition system using only 59 dimensional LBP features extracted from facial image. 16 non-overlapping sub-blocks, each of size 60×60 is placed over the facial region. 944 dimensional LBP feature vector is obtained after sequentially concatenating LBP vectors extracted from each of the sub-blocks. We applied PCA to compress data from 944 dimensions to only 59 dimensions. The results demonstrates that KSOM performs very well when feature vector $\mathbf{x} \in \mathbb{R}^{59}$ is applied. The average recognition rate obtained is just marginally less than KSOM based approach applied over $\mathbf{x} \in \mathbb{R}^{944}$. Use of PCA removes unnecessary redundant data and thus leads to less computation cost. To verify the outstanding performance of KSOM based approach we further train the feature vector $\mathbf{x} \in \mathbb{R}^{59}$ using 3 widely used classifiers such as; SVM, RBFN and MLP3. The average recognition rate of the classifiers are found to be 65.78%, 68.09% and 62.73% respectively. On the other hand, the KSOM based approach outperforms all other classifiers with average recognition rate 69.18%.

References

[1] A. Mehrabian, *Nonverbal communication*. Aldine, 2007.

[2] P. Ekman, W. V. Friesen, and J. C. Hager, *Facial action coding system*. A Human Face, Salt Lake City, 2002.

[3] J.-J. J. Lien, T. Kanade, J. Cohn, and C. Li, "Detection, tracking, and classification of action units in facial expression," *Journal of Robotics and Autonomous Systems*, July 1999.

[4] G. Donato, M. S. Bartlett, J. C. Hager, P. Ekman, and T. J. Sejnowski, "Classifying facial actions," *IEEE transactions on pattern analysis and machine intelligence*, vol. 21, no. 10, pp. 974–989, 1999.

[5] B. Fasel and J. Luettin, "Recognition of asymmetric facial action unit activities and intensities," in *15th International Conference on Pattern Recognition, 2000. Proceedings.*, vol. 1, 2000, pp. 1100 –1103 vol.1.

[6] I. Cohen, N. Sebe, A. Garg, M. Lew, and T. Huang, "Facial expression recognition from video sequences," in *IEEE International Conference on Multimedia and Expo, 2002. ICME '02. Proceedings. 2002*, vol. 2, 2002, pp. 121 – 124 vol.2.

[7] J. Cohn, L. Reed, Z. Ambadar, J. Xiao, and T. Moriyama, "Automatic analysis and recognition of brow actions and head motion in spontaneous facial behavior," in *IEEE International Conference on Systems, Man and Cybernetics, 2004*, vol. 1, oct. 2004, pp. 610 – 616 vol.1.

[8] Y.-I. Tian, T. Kanade, and J. Cohn, "Recognizing action units for facial expression analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence,*, vol. 23, no. 2, pp. 97 –115, feb 2001.

[9] Y. Zhang and Q. Ji, "Active and dynamic information fusion for facial expression understanding from image sequences," *IEEE Transactions on Pattern Analysis and Machine Intelligence,*, vol. 27, no. 5, pp. 699 –714, may 2005.

[10] M. Valstar, I. Patras, and M. Pantic, "Facial action unit detection using probabilistic actively learned support vector machines on tracked facial point data," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops, 2005. CVPR Workshops.*, june 2005, p. 76.

TABLE I: Confusion matrix of emotions detection for the 944 dimensional sub-block based LBP features data using KSOM. The emotion classified with maximum percentage is shown to be the detected emotion.

	H	Sa	D	A	Sur	F
H	**85.4**	1.6	4.83	0	6.45	1.6
Sa	8.3	**73.1**	1.85	3.7	8.3	4.6
D	27	9.6	**46.7**	6.7	5	5
A	13.8	18.7	0	**58.7**	3.75	5
Sur	1.9	3.85	0	0	**94.2**	0
F	5.1	17.9	0	7.7	0	**69.2**

TABLE II: Confusion matrix of emotions detection for the 59 dimensional sub-block based LBP features data using KSOM. The emotion classified with maximum percentage is shown to be the detected emotion.

	H	Sa	D	A	Sur	F
H	**84**	1.33	5.3	8	0	1.33
Sa	6.4	**53.2**	17.4	11.9	1.83	9.17
D	15	14.5	**50.75**	15.8	1.9	1.9
A	4	8	5.33	**76**	0	6.67
Sur	2.7	0	10.8	2.7	**83.8**	0
F	15.4	1.9	0	15.4	0	**67.31**

TABLE III: Confusion matrix of emotions detection for the 59 dimensional sub-block based LBP features data using SVM. The emotion classified with maximum percentage is shown to be the detected emotion.

	H	Sa	D	A	Sur	F
H	**81.0**	7.9	5.1	2.83	1.61	1.61
Sa	13.9	**59.5**	10.1	10.1	1.27	5.1
D	10.8	19.7	**53.2**	10.8	0	5.4
A	3.77	19.6	11.3	**63.3**	1.8	0
Sur	7.14	10.7	7.14	3.57	**71.4**	0
F	8.63	10.39	4.87	9.75	0	**66.3**

TABLE IV: Confusion matrix of emotions detection for the 59 dimensional sub-block based LBP features data using RBFN. The emotion classified with maximum percentage is shown to be the detected emotion.

	H	Sa	D	A	Sur	F
H	**75.6**	3.84	6.41	7.69	6.41	0
Sa	8.25	**62.4**	8.25	15.6	3.67	1.83
D	3.77	11.3	**56.04**	16.98	5.9	5.88
A	1.33	14.33	13.0	**70**	0	1.33
Sur	5.41	0	7.7	7.7	**79.2**	0
F	1.92	13.5	0	13.5	5.77	**65.38**

TABLE V: Confusion matrix of emotions detection for the 59 dimensional sub-block based LBP features data using MLP3. The emotion classified with maximum percentage is shown to be the detected emotion.

	H	Sa	D	A	Sur	F
H	**72**	2.67	9.33	6.67	4	5.33
Sa	8.25	**67.9**	11.9	7.34	0.91	3.67
D	9.43	11.3	**50.4**	14.3	8.89	5.66
A	1.33	20.33	13.7	**52.7**	5.33	6.67
Sur	8.1	0	10.41	7.7	**73.8**	0
F	1.92	15.38	5.77	17.3	0	**59.6**

[11] F. Tsalakanidou and S. Malassiotis, "Real-time 2d+3d facial action and expression recognition," *Pattern Recognition*, vol. 43, no. 5, pp. 1763 – 1775, 2010.

[12] S. Moore and R. Bowden, "Local binary patterns for multi-view facial expression recognition," *Computer Vision and Image Understanding*, vol. 115, no. 4, pp. 541–558, 2011.

[13] W. Gu, Y. Venkatesh, and C. Xiang, "A novel application of self-organizing network for facial expression recognition from radial encoded contours," *Soft Computing-A Fusion of Foundations, Methodologies and Applications*, vol. 14, no. 2, pp. 113–122, 2010.

[14] T.-H. Wang and J.-J. J. Lien, "Facial expression recognition system based on rigid and non-rigid motion separation and 3d pose estimation," *Pattern Recognition*, vol. 42, no. 5, pp. 962 – 977, 2009.

[15] A. S. M. Sohail and P. Bhattacharya, "Classifying facial expressions using level set method based lip contour detection and multi-class support vector machines," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 25, no. 06, pp. 835–862, 2011.

[16] R. Luo, C. Huang, and P. Lin, "Alignment and tracking of facial features with component-based active appearance models and optical flow," in *Advanced Intelligent Mechatronics (AIM), 2011 IEEE/ASME International Conference on*, july 2011, pp. 1058 –1063.

[17] T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on featured distributions," *Pattern recognition*, vol. 29, no. 1, pp. 51–59, 1996.

[18] Z. Zhang, M. Lyons, M. Schuster, and S. Akamatsu, "Comparison between geometry-based and gabor-wavelets-based facial expression recognition using multi-layer perceptron," in *Third IEEE International Conference on Automatic Face and Gesture Recognition, 1998. Proceedings.* IEEE, 1998, pp. 454–459.

[19] Z. Zhang and Z. Zhang, "Feature-based facial expression recognition:

TABLE VI: Parameters used in different classifiers.

KSOM	SVM	RBFN	MLP3
$2D$ lattice network of size 10×8.	15 SVMs for 6 class classification	50 radial centers	10 neurons.
Logistic sigmoid at output node	Radial basis function (RBF) as kernel	Gaussian radial function at centers	Sigmoid function
Neighborhood radius σ initial = 3.5, final = 0.001	penalty weight $C = 1$ RBF radius $\gamma = 0.10$	Sigma updated at each iteration	
10 generations	50 generations	10 generations	10 generations

Sensitivity analysis and experiments with a multilayer perceptron," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 13, pp. 893–911, 1999.

[20] M. Rosenblum, Y. Yacoob, and L. Davis, "Human expression recognition from motion using a radial basis function network architecture," *IEEE Transactions on Neural Networks,*, vol. 7, no. 5, pp. 1121–1138, 1996.

[21] D. Lin, "Facial expression classification using pca and hierarchical radial basis function network," *Journal of information science and engineering*, vol. 22, no. 5, pp. 1033–1046, 2006.

[22] M. Bartlett, G. Littlewort, M. Frank, C. Lainscsek, I. Fasel, and J. Movellan, "Recognizing facial expression: machine learning and application to spontaneous behavior," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005. CVPR 2005.*, vol. 2, june 2005, pp. 568 – 573 vol. 2.

[23] P. Viola and M. Jones, "Robust real-time object detection," *International Journal of Computer Vision*, vol. 57, no. 2, pp. 137–154, 2002.

[24] A. Majumder, L. Behera, and V. K. Subramanian, "Facial expression recognition with regional features using local binary patterns," in *CAIP (1)*, 2013, pp. 556–563.

[25] T. Kohonen, "The self-organizing map," *Proceedings of the IEEE*, vol. 78, no. 9, pp. 1464–1480, 1990.

[26] M. Pietikinen, A. Hadid, G. Zhao, and T. Ahonen, "Local binary patterns for still images," in *Computer Vision Using Local Binary Patterns*. Springer London, 2011, vol. 40, pp. 13–47.

[27] A. Majumder, L. Behera, and V. K. Subramanian, "Emotion recognition from geometric facial features using self-organizing map," *Pattern Recognition*, vol. 47, no. 3, pp. 1282 – 1293, 2014.

[28] M. Pantic, M. F. Valstar, R. Rademaker, and L. Maat, "Web-based database for facial expression analysis," in *Proceedings of IEEE Int'l Conf. Multimedia and Expo (ICME'05)*, Amsterdam, The Netherlands, July 2005, pp. 317–321.