2020 6th International Conference on Advanced Computing & Communication Systems (ICACCS)

Face Emotion Classification using AMSER with Artificial Neural Networks

M.Kalpana Devi
Department of Information Technology
Sri Ramakrishna Institute of Technology
Coimbatore, India
lakross6.sidharaj@gmail.com

K.Prabhu
Department of Information Technology
Sri Ramakrishna Institute of Technology
Coimbatore, India
Prabhuit.23@gmail.com

Abstract— **The Facial expression recognition phases provides a leading challenges and issues which are volatile while representing the different facial expression. The test sample complexity is not measured during the training phase of classification model. In certain parts of the faces the emotions are detected by the physiological changes that are expressed externally on face. The human cognition laws like principle of simplicity are unreliable in finding out the suitable test samples. To overcome, a new algorithm is proposed Advanced Maximally Stable Extremal Regions (AMSER) method for extracting the features. With the corresponding dataset and Artificial Neural Networks (ANN), the classification process is extracted with better facial expressions. The experimental result shows the perfectness and correct accuracy in classifying the facial expressions.**

Keywords— Facial expression recognition, Advanced Maximally Stable Extremal Regions (AMSER), Artificial Neural Networks (ANN).

I. INTRODUCTION

In image processing, the analysis portion of facial recognition is most important in present era. This is huge trials which are confronted often during the face recognition. The well-known faces are recognized easily when compared to the unfamiliar faces which play a vital role. The images are sorted into preprocessing, feature extraction and classification. These are coupled with kernel PCA to enhance the recognition performance by using deep neural networks [1]. Now-a-days it becomes massive in detecting the different facial emotional expression. The most challenging tasks are acknowledged to enhance the better performance while detecting the facial emotions. The images are captured and retrieved in surveillance video systems. The expressions are strong and very powerful while the conversation takes place between each human being. The facial expression processed by machine plays a wide area of interest in world of research. The system analyzes the different facial expression features by positions and sizes are defined. These active features are monitored with the help of voting priority [2]. The different categories of problem areas in facial recognitions are: identifying faces, the feature extraction and classification. The scenario of this paper is to identify the presence posture of facial expressions. The facial expressions of seven different categories are shown in Fig.1.

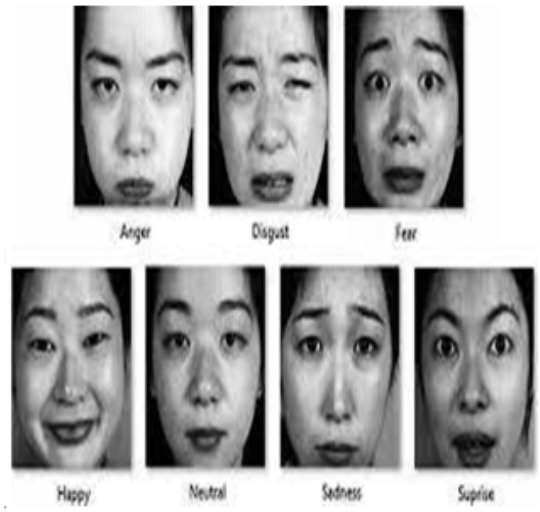

Fig.1. Facial Expression

The facial expression gathers the static images and features are extracted for edge region by using the directional information and ternary pattern in order to provide a smooth performance [3].

The emotions are enforced directly or indirectly based on decision making, machine learning, deep learning etc. The localized features are identified by using Partial F-test values by decreasing the low variance from the expression classes. Two levels are employed based on the Hidden Conditional Random Fields such as finding the category of the images and labeling the images from the categories [4]. The communication takes places with high ended emotional values.

The biometric plays an important role in face emotion recognition which is superior now-a-days. The emotional intensity level is measured for facial action and

found there is a highest accuracy improvement by using Gabor filters, a Histogram of Oriented Gradients (HOG), and Local Binary Pattern (LBP) [5]. The speech of a person is observed and stored. The features are mined with the help of speech by using segment level. The Deep Neural Networks (DNN) and Support Vector Machine (SVM) are combined together to gather the perfect emotion from the speech recognition [6]. The facial features can be detected by the sequences of images which are captured. The captured images are segmented and normalized by setting the threshold values, to investigate into binary images. Based on this, the absolute facial feature points are constructed [7]. The different emotions are noticed from various sequences of methods with the occurrence happening in the physiological movements. This paper classifies the facial emotions by using the advanced maximally stable extremal regions (AMSER) method. This proposed paper concentrates on preprocessing, feature extraction, feature selection and classification techniques respectively. It suggests a classification results by using the AMSER and ANN method. This paper focuses in various aspects by showing Section 1 shows the introduction in recognizing the Face emotion approaches and its methods, section 2 describes the review of literature or related work, section 3 describes about existing work and problem statement, section 4 describe about proposed method, section 5 shows the simulation work of the proposed method, section 6 contains the conclusion of this paper.

II. LITERATURE REVIEW

Changxing Ding et al [8] suggested a new concept namely Multi-Directional Multi-Level Dual-Cross Patterns which performs the various aspects that are reflected in face. This method MDML-DCPs scheme decreases computation takes place. DCP is robust and provides an efficient computation result in form of pose and expression around the face.

The face recognition is standardized in two geometric representations namely similarity transformation and affine transformation by Multi-directional multi-level dual cross patterns (MDML-DCP). The similarity transformation is used to identify the originality of the face features and affine transformation reduces the intra-personal variations. It encodes the image of the face from multilevel to pattern level which represents the variation effectively. The results are compared for verification and identification of the face with FERET, CAS-PERL-R1, FRGC 2.0, and LFW.

Thierry Pun, et.al., [9] analyze the reading emotions from facial expression and speech is a milestone in Human-Computer Interaction. Recent sensing technologies, namely the Microsoft Kinect Sensor, provide basic input modalities data, such as RGB imaging, depth imaging and speech that can be used in Emotion Recognition. Moreover, Kinect can track a face in real time and present the face fiducial points, as well as 6 basic Action Units (AUs). In this work we explore this information by gathering a new and exclusive dataset. This is a new opportunity for the academic community as well to the progress of the emotion recognition problem. The database includes RGB, depth, audio, fiducial points and AUs for 18 volunteers for 7 emotions. The support vector machines, Neural Networks classifiers and k-Nearest Neighbor algorithm are implemented to compare the efficient result for emotional facial expressions. This technique concentrates on multimodal representations.

Vandana Patidar [10] analyzes the applications which could be easy to use and are user-friendly, that even people with specific disabilities use them easily. Some of the challenges that are faced in Facial Expression Recognition are pattern and features recognition which plays a most important part of the vital role. Many algorithms are designed to establish either static representation such as uniform background, identical poses, similar illuminations nor dynamic representation namely position variation, partial occlusion orientation, varying lighting. In general, it can be organized by three ways: firstly detecting the face secondly features Extraction and finally classification.

Ali Moeini , Hossein Moeini, et .al., [11] proposed a novel method with the representation of single image for face recognition concept A 3D Probabilistic Facial Expression Recognition Generic Elastic Model (3D PFER-GEM) and Feature Library Matrix (FLM) is proposed to redesign the framework of 3D model and to extract the features of the face in a rotated manner. The triplet angle is estimated by using the head pose and based on this sequence the FLM is nominated. The elastic model is constructed for designing the 3D model to the human face by considering or not considering the expression of 2D frontal image expressions. It is compared with the features which are extracted for classification using support vector machine. This proposed algorithm provides a high accuracy.

III. PROBLEM STATEMENT

The facial expression can be categorized into two groups Intrinsic factors and Extrinsic factors.

The intrinsic factors are related to observe the originality of the human face. These are broadly divided into intrapersonal and interpersonal factors. The intrapersonal factors are used to show the unlike features of the similar person. The interpersonal are used to show the unlike features of different persons. The Extrinsic factors is stated that the face appearances are adjusted according to the illumination, scaling, parameter etc.

Some of the challenges faced are:

1. When there is some motion reflection continuously and as well as crowd, it is very difficult to detect the face is shown in Fig.2.

Fig.2. Crowd image

2. The same facial expression with different lighting conditions makes dramatic changes in appearances is displayed in Fig.3.

Fig.3. Variation in illumination

3. If there is a change in position, viewing angle and rotation of the head might cause a challenge during the detection of face recognition is shown in Fig.4.

Fig 4. Pose variation

4. Based on the different emotion, the human facial expression varies.
5. When the ageing differs, the appearance of human faces gets reformed.
6. Occlusion is one of the widest challenges. The blockage of any features in the face image is called occlusion.
7. Low Resolution problem in the images.
8. Similarity of the same faces.
9. The system faults like camera distortion, improper techniques, background noise etc…

IV. PROPOSED METHODS

In recent years face recognition had an achieved most attention and interest from researchers due to its broadly used applications in computer vision. Unlike other biometric i.e palm, iris, fingerprint etc., face biometrics are captured without user's knowledge and later can be used for surveillance systems. Over the past decades of research, various methodologies were implemented under the concept of face recognition techniques to achieve good classification of the person feelings, but it fails in recognitions of the expression. To finding the accurate classification is our main aim. For the classification object here this paper uses the ANN method of classification. For doing this classification many phases were included which is shown in Fig.5.

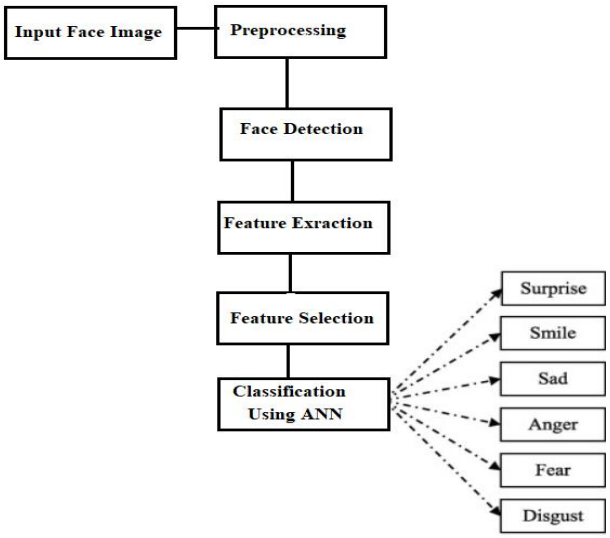

Fig 5. Proposed Architecture

A. Image processing

The image analysis is processed in order to provide a better solution and high accuracy in result. Due to the different factors and analysis, the result is compared to various existing methods. The few steps are followed for preprocessing the images are image resizing, SNR and image enhancement. To decrease the noise ratio image clipping is performed by cropping a few areas that are not needed for the result. To smoothen the image, low frequency components are used instead of high frequency [13]. Image smoothing is done by using the smoothing filter. Increasing the contrast is performed by Image enhancement.

B. Face Detection

Face detection are used for identifying the human faces. All the determined images are converted into bits and the images are authenticated from the database by match pattern verifications. The methods are complex based on the pose, expression, orientation, position, skin, facial hair etc. [14]. this is a part of object detection which concentrates mainly on security, authorization, entertainment, biometrics for face images. Some of the methods are followed for

detecting the face such as template matching, appearance based, feature based and knowledge based.

For detecting the facial images some of the sequences steps followed are intensity, size, shape and detecting feature points by using the rectangle box according to the author Ming-Hsuan yang et.al's [15].

C. Feature Extraction using AMSER

The MSER method is used for detecting the images. It mines the images into number of regions called as MSER. It converts the color images into gray scale images. It sets the threshold value and check whether the image pixel is above or below the threshold. If the pixel is below the threshold value, then images are converted into white else black. The AMSER feature is defined for preprocessing using morphological operations. To generate the features and pattern identification, feature extraction is enhanced and later the extracted feature is used for selection and classification to determine the appropriate result [16]. The input images are loaded and converted it into gray scale images. MSER region are detected from the gray scale images. The stable regions are selected by the intensity range based on the higher and lower intensity value. The morphological operations are done to improve the performance of the image clarity. The rectangle box is determined using AMSER for extracting the face [17, 18]. Some of these steps are followed to retrieve the components of images by AMSER intensity. In this paper, the faces are extracted using region extraction by setting the threshold value to convert into gray scale images for better results were shown in 5th section.

AMSER extraction implements the following steps:

- Load the input images from the databases and set the threshold values. Based on the threshold value convert the pixel images into gray scale are shown in Fig. 6.
- The connected gray scale components are extracted using AMSER Region extractor.
- The morphological operations are used to find the components for improving the clarity of the images.
- The rectangle boxes are enclosed to determine the extracted faces.

D. AMSER Algorithm

Step 1: Load the image Img with size A B C D M XN XP where $D_{i, j, k}$ refers the color intensity Img=$\sum\sum\sum$ D (i, j, k) where i → 1to A, j → 1to B, k → 1to C

Step 2: Image Img is converted into gray scale image Img1 where $D_{i, j}$ is the gray scale intensity Img1= $\sum\sum$ D(i, j)where i → 1to A , j → 1 to B.

Step 3: To smooth the image, find intensity value and apply the threshold value. Select the strong edge and reject the weak edges are shown in Fig.7.

Step 4: The morphological operations are done by dilation and reconstruction on the regions is shown in Fig. 8.

Step 5: The extracted faces are detected is shown in Fig.9.

E. Feature Selection using GA

The feature selection is done after the feature extraction by AMSER in order to improve the accuracy of the image and to reduce the time complexity. Some of the features like redundant, noisy images etc. are discarded to view the exact features by GA. Feature selection is analyzed using Genetic Algorithm (GA) to identify the pattern. The identified features are examined through information gain [19]. Information gains are defined to be as the variance between the entropy of the label available in the dataset and the entropy of the feature behavior. Next process is based on the ranking method for emotion among the selected features .The final step for selecting emotional images are done based on the highest features selections.

F. Classication using ANN

The artificial neural network (ANN) is vast area among the researchers. Rapidly growing areas is ANN which plays a vital role in many of the applications. The ANN can be used for classifying the emotional state in the face recognition. The parameter like weight, bias and learning rate are used for training the ANN. The adjustment of weight is made to improve the performance of the emotional images. The neurons are interconnected with each other and communication takes place based on the weight applied to it [20]. The adjustment of weight is varied in artificial neural network among the learning phase. Calculating the width of left eye, width of right eye, width of mouth, distance from left eyebrow to right eyebrow, height of left eye, height of right eye, height of mouth, distance from left eye to left eyebrow, distance from right eye to right eyebrow during the next step to produce the appropriate result.

V. EXPERIMENTAL RESULTS

The experimental result in the learning process is analyzed for the eligibility test. The following steps were established to find the result.

1. It is found and analyzed that the emotional expression is identified and classified using ANN perfectly.
2. The positive and negative emotional states with hidden reactions are well-known.

The seven emotional values like anger, neutral, disgust, fear, happy, sad and surprise are noticed. Japanese

Female Facial Expression (JAFFE) dataset are used for simulating the proposed work in MATLAB environment. The result of the emotional image is recorded and compared with the expected result. The following steps are involved for the proposed flow work which is shown in Fig.10.

1. The trained facial expressions are taken from the datasets are uploaded.
2. The preprocessing data are detected.
3. Apply the AMSER Algorithm for feature extraction.
4. The fitness values are processed based on the genetic algorithm to find the optimal solution.
5. Classification using ANN is performed based on the fitness value and trained.
6. If the images are matched, classify the result and then calculate the FAR, FRR and accuracy. If not matched, calculate the FAR, FRR and accuracy.

The comparison chart of accuracy, FAR and FRR with existing and proposed work are shown in Fig.11, Fig.12 and Fig.13 respectively.

Fig 6. Input Face Image

Fig 7. Face Detected Image

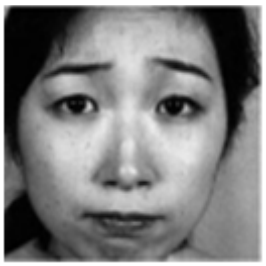

Fig 8. Region of Interest after Detecting Face

Fig 9. AMSER Feature Based Output Image

The proposed work results are calculated, verified and compared with the existing methods which are shown in the corresponding table.1, table.2 and table.3 respectively.

TABLE I. ACCURACY COMPARISON OF EXISTING AND PROPOSED METHOD

No of Samples	Accuracy		
	SIFT	***MSER***	***AMSER***
1	98.38	98.5	97.5
2	97.49	98.8	97.3
3	96.25	97.55	95.43
4	97.25	98.45	95.85
5	98.25	98.66	96.67
6	99.13	99.1	98.1
7	98.25	99.25	97.55
8	97.02	99.35	97.56
9	99.05	99.4	97.68
10	98.57	99.44	96.33

TABLE II. FAR COMPARISON FOR SIFT, MSER & AMSER METHOD

No of Samples	FAR		
	SIFT	***MSER***	***AMSER***
1	98.71	98.72	99.01
2	97.66	98.77	99.25
3	96.00	97.67	98.22
4	97.87	98.88	99.32
5	98.81	98.85	99.12
6	99.25	99.27	99.52
7	98.35	99.05	99.16
8	97.22	99.78	99.79
9	99.00	99.33	99.35
10	98.59	99.45	99.54

TABLE III. FRR COMPARISON FOR SIFT, MSER & AMSER METHOD

No of Samples	FRR		
	SIFT	*MSER*	*AMSER*
1	1.14	1.2	0.04
2	1.15	1.0	0.03
3	1.24	0.78	0.03
4	1.35	0.65	0.032
5	1.76	0.78	0.025
6	1.48	0.71	0.024
7	1.48	0.62	0.22
8	1.15	0.55	0.15
9	1.78	0.4	0.14
10	1.05	0.32	0.02

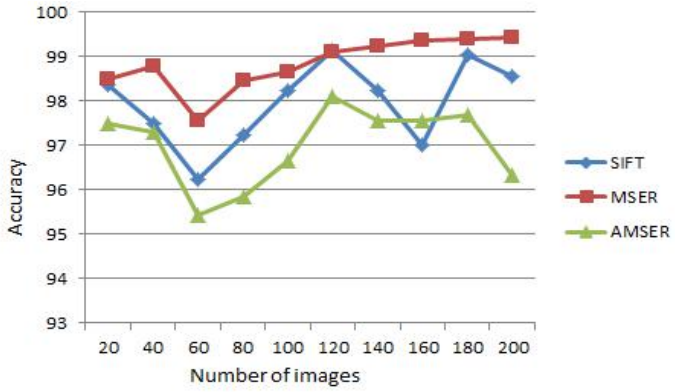

Fig 11. Accuracy Comparison of Existing and Proposed Method

Fig 12. Comparison of FAR with Existing and Proposed Method

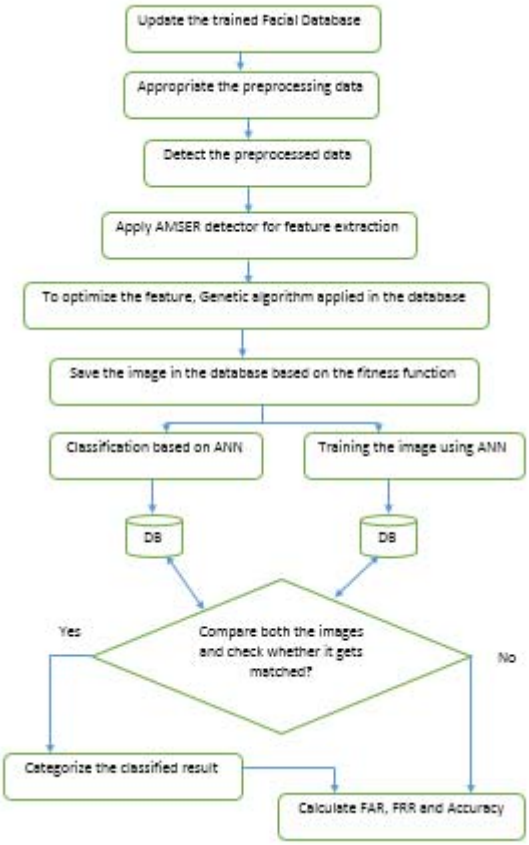

Fig 10. Proposed Algorithm Flowchart

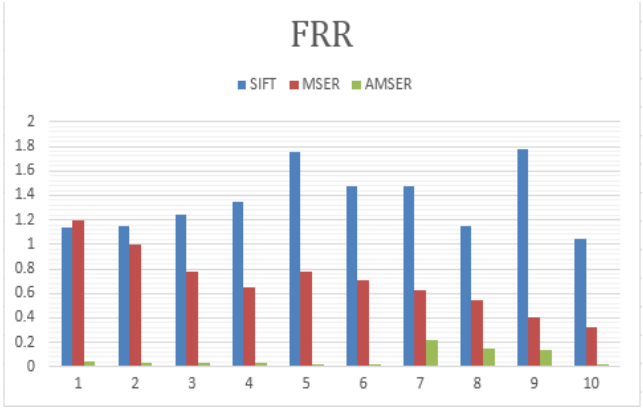

Fig 13. Comparison of FRR with Existing and Proposed Method

Individual Accuracy rates for each emotion

- Accuracy of Angry 9.666667e+01
- Accuracy of Neutral 100
- Accuracy of Disgusted 9.655172e+01

- Accuracy of Fear 100
- Accuracy of Happy 100
- Accuracy of Sad 9.666667e+01
- Accuracy of Surprise 100

VI. CONCLUSIONS

In this research work, a facial emotion classification using AMSER algorithm along with ANN has been proposed. In this system various emotions like fear, disgust, anger, happy, neutral, sadness and surprise are processed from facial expression. Emotions are recognized using preprocessing, feature extraction, selection and classification which plays a big challenging role in the field of research and achieves a better accuracy in minimum execution time. To extract the features for various facial expression, AMSER is used. The feature extraction from AMSER is initialized as random population, a genetic algorithm has been proposed to identify the fitness function. Artificial neural network are used to classify the emotional images from the different facial expressions. This proposed algorithm uses Japanese Female Facial Expression (JAFFE) dataset and the simulation work is done through MATLAB environment. Some of the controls are required to manage the images through facial expression for real world applications. The efficiency of the proposed system results with 99 % better accuracy compared to the existing system. In future, for extracting the feature a new trends with machine learning algorithm and artificial intelligence can be proposed.

Acknowledgment

We would like to thank the management of SNR trust for supporting the faculty in doing Research work. We also thank our beloved Principal of Sri Ramakrishna Institute of Technology for encouraging the staff members to publish the papers in various Scopus indexed and SCI journals. Lastly, we thank our HOD of Information Technology who motivates us and provide his valuable suggestion for publishing the paper.

References

[1] J. Li, E.Y. Lam , "Facial expression recognition using deep neural networks", IEEE International conference on Imaging Systems and Techniques , IEEE, 2015.

[2] S.L. Happy, et.al.,"Automatic facial expression recognition using features of salient facial patches", IEEE Transaction on Affective Computing, vol:6 Issue :1 ,2015,pp:1-12.

[3] Byungyong Ryu, et.al., "Local directional ternary pattern for facial expression recognition", IEEE Transaction on Image Processing, vol:26 Issue: 12, 2017,pp: 6006-6018.

[4] M.H. Siddiqi, et.al., "Human facial expression recognition using stepwise linear discriminant analysis and hidden conditional random fields", IEEE Transaction on Image Processing, vol:24, issue: 4 ,2015, pp: 1386-1398.

[5] Mehta, D.; Siddiqui, M.F.H.; Javaid, A.Y. "Recognition of emotion intensities using machine learning algorithms: A comparative study", Sensors (Basel) , vol: 19(8), 2019.

[6] Chen, R.; Zhou, Y.; Qian, Y. " Emotion Recognition Using Support Vector Machine and Deep Neural Network", National Conference on Man-Machine Speech Communication , Springer, 2018; pp. 122–131.

[7] Rajesh A Patil, Vineet Sabula, A.S.MandaI "Automatic Detection of Facial Feature Points in Image sequences", International Conference on Image Information Processing, 2011.

[8] Changxing Ding et al, "Multi-Directional Multi-Level Dual-Cross Patterns for Robust Face Recognition", IEEE Transactions on Pattern Analysis and Machine Intelligence , vol: 38 , Issue: 3 ,2016, pp:518-531.

[9] Thierry Pun, M. Soleymani, J. Lichtenauer, and M. Pantic, "A Multimodal Database for Affect Recognition and Implicit Tagging," IEEE Transactions on Affective Computing , vol: 3 , Issue: 1, 2012 , pp. 42–55.

[10] Vandana Patidar, "An Analysis of Facial Expression Recognition Techniques" International Journal on Future Revolution in Computer Science & Communication Engineering, vol: 3 Issue: 11, 2017, pp:510 – 519.

[11] Ali Moeini , Hossein Moeini et al, "Real-World and Rapid Face Recognition Toward Pose and Expression Variations via Feature Library Matrix", IEEE Transactions on Information Forensics and Security, vol: 10 , Issue: 5 , 2015, pp: 969 – 984.

[12] Amrendra Sharan, "A Hybrid Approach of Facial Emotion Detection using Genetic Algorithm along with Artificial Neural Network", International Journal of Computer Applications, vol: 175 – No.4, 2017.

[13] Haiyang Zhanget al, "Image Pre-processing Methods in Face Recognition", Symposium on Photonics and Optoelectronics (SOPO) , IEEE, 2011.

[14] Wei zha, Jeong·Sun Park, AND Sang Woong L EE," Fully Automatic Face Detection and Facial Feature Point Extraction Using Local Gabor Filter Bank and PCA", International Conference on Machine Learning and Cybernetics (ICMLC) , 2011.

[15] Ming-Husan Yang,David J.Kriegman,Narendra Ahuja ,"Detecting Faces in Images: A survey" IEEE Transaction on Pattern Analysis and Machine Intelligence",Vol.24,No.1,2002, pp:34-58.

[16] Changyuan Liu ; Yan Li ; Xiaojun Bi , et.al., "Face recognition based on relevance vector machine" Proceedings of 2011 6th International Forum on Strategic Technology, 2011,pp:1202- 1206.

[17] M. Donoser and H. Bischof. "Efficient maximally stable extremal region (mser) tracking" , IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2006, pp: 553–560.

[18] Ehab Salahat, "Recent Advances in Features Extraction and Description Algorithms: A Comprehensive Survey", Annual IEEE Industrial Electronics Societys 18th International Conf. on Industrial Technology, 2017.

[19] Boughrara.H, Liming,chen, Ben Amar.C,Chtourou.M, " Face Recognition under varying Facial Expression based on Perceived Facial images and Local Feature matching "2012 International Conference on Information Technology and e Services, 2012, pp:24-26.

[20] Anissa Bouzalmat, Naouar Beghini, et.al., "Face detection and Recognition using base propagation Neural Network and Fourier Gabor filters" , Signal & Image Processing : An International Journal (SIPIJ), vol 2, No.3 , 2011.