IIOAB JOURNAL

ARTICLE

HUMAN-MACHINE INTERACTION IN FACIAL EXPRESSION RECOGNITION SYSTEM

K. Suresh[1*] and C. Chellappan[2]

[1]*Dep. of CSE, GKM College of Engineering and Technology, Anna University Recognized Research Centre, Chennai, INDIA*

[2] *GKM College of Engineering and Technology, Chennai, INDIA*

ABSTRACT

Image processing is an image-in and image-out work, in that artificial intelligence in affective computing is scientifically challenged Research area in the field of computer vision technology. Facial expressions are the fundamental way to express human emotions and also an effective method of non-verbal communication, development of artificial intelligence and pattern recognition, researchers paying more and more attention to the facial expression recognition system. The Human can easily recognize facial expression, but it is quite a challenging task for the machine to do this many application which use facial expression to evaluate the human nature, feelings, judgment and opinion. This paper presents a broad review of various modules in the facial expression recognition system and robust techniques used at each level. An ongoing challenge in this field is to design an intelligence system for effective human-machine interactions.

KEY WORDS
Emotion, expression, computer vision, recognition, affective computing

Published: 8 Nov 2016

***Corresponding Author**
Email:
sureshtrack@gmail.com

INTRODUCTION

Human invented computers are faster and more accurate than human beings in processing numerical data but human beings scoring over computers in recognition capability. A human can recognize an object in a few seconds, without any difficulty. The methods by which human gathers knowledge for recognition is unique because human beings use all the five sensory organs (visual, hear, taste, smell, and touch) to gather knowledge about the outside world. Among these perceptions of visual information plays a major role in understanding the environment. There are many ways a human can express emotions the most natural way of expressing emotions is using facial expressions because the face is the index of the human mind. Facial expression detection means finding the face in an image and classifying into one of the emotional states among six universal labels of angry, disgust, fear, happy, sad and surprise. If none of the above states presents it consider as neutral.

Facial expressions are playing an important role in human communication. A picture worth thousands of words, i.e. a face contains much information about a person id (identification) and also about mood and state of mind. Artificial intelligence in expression classifier system has lots of attention due to its potential application in many commercial fields like the advance driver assistance system, criminal investigation, gaming, human behavior analysis, human-machine intelligent interaction, health support appliances, intelligent robotic system, lie detection, smart rooms, surveillance, and depression detection. Facial expression recognition system [FERS] utilized in the automatic analysis of human emotion. Human interpersonal behaviors are affected by facial expression during communication. Although the technology for emotion recognition is an important one which demanded in various fields, it still remains the unsolved problem. Automatic facial expression recognition has undergone substantial advancement over the past two decades. To truly achieve effective human-computer intelligent interaction, there is a need for the computer to understand a facial expression scene to naturally interact with the user as similar to the way how human to human interactions are taking place.

Recognizing human facial expression is not just an easy and straightforward task due to several circumstances [Fig. 1] like illumination, facial occlusion, face shape, low-resolution images, head pose variations, dark-skin faces, the color of the image, the distance between facial landmarks variation from person to person, an absence of a neutral face for comparison, low intensity expression and so on. Facial expression is a specific and a hard case for the object recognition. Our faces can express the things which are difficult to put into words. In modern computer vision technology, face to face encounters are expected in next generation computer systems to convey information about the state of the mind of a person may be inferred. As per Albert Mehrabian's [1] in communication research, 7% of meaning in the words that are spoken, 38% of meaning is paralinguistic (the way that the words are said), 55% of communicating cues can be judged by facial expression, and hence recognition of facial expression becomes a major modality. In 1872- Darwin's Charles [2] demonstrated the universality of facial expression and their continuity in man and animals. The early 1970s, Paul Ekman has performed an extensive study of human facial expressions. 1971 American Psychologist Ekman and Friesen [3] defined six basic emotions: angry, disgust, fear, happiness, sadness, and surprise. In 1978 Facial Action Coding System was developed by Ekman [5] for facial expression description. The approach in FACS [4] which separates the expression into upper and lower face actionIn1978-suwea and et.al, [6] presented a preliminary investigation on automatic facial expression analysis from the sequence of images.

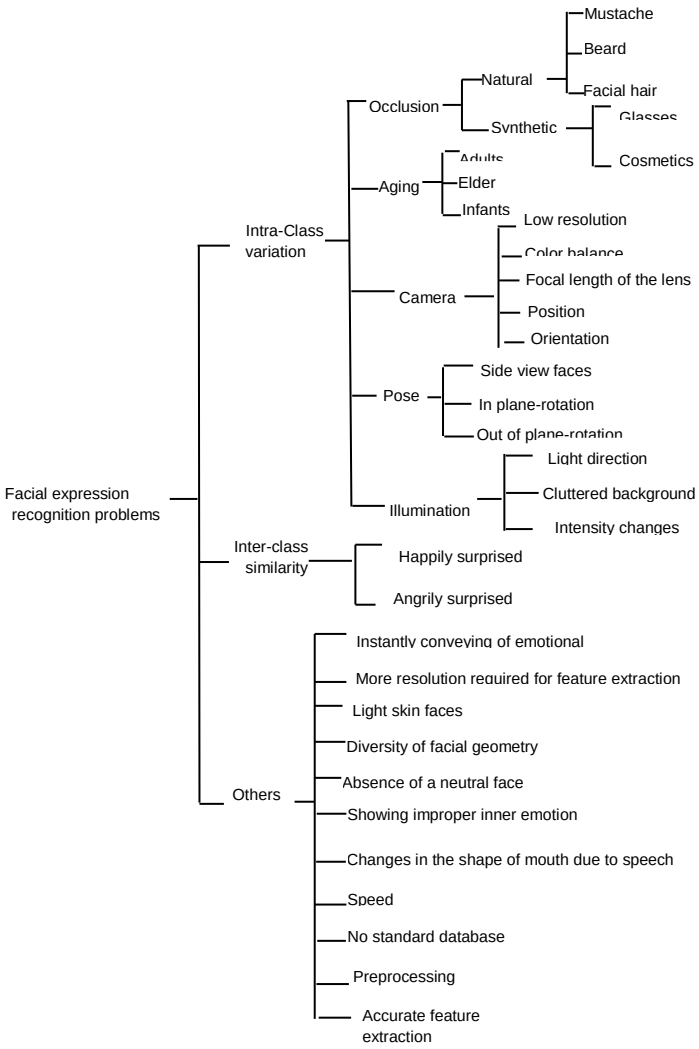

Fig. 1: Taxonomy of various issues in facial expression recognition system.

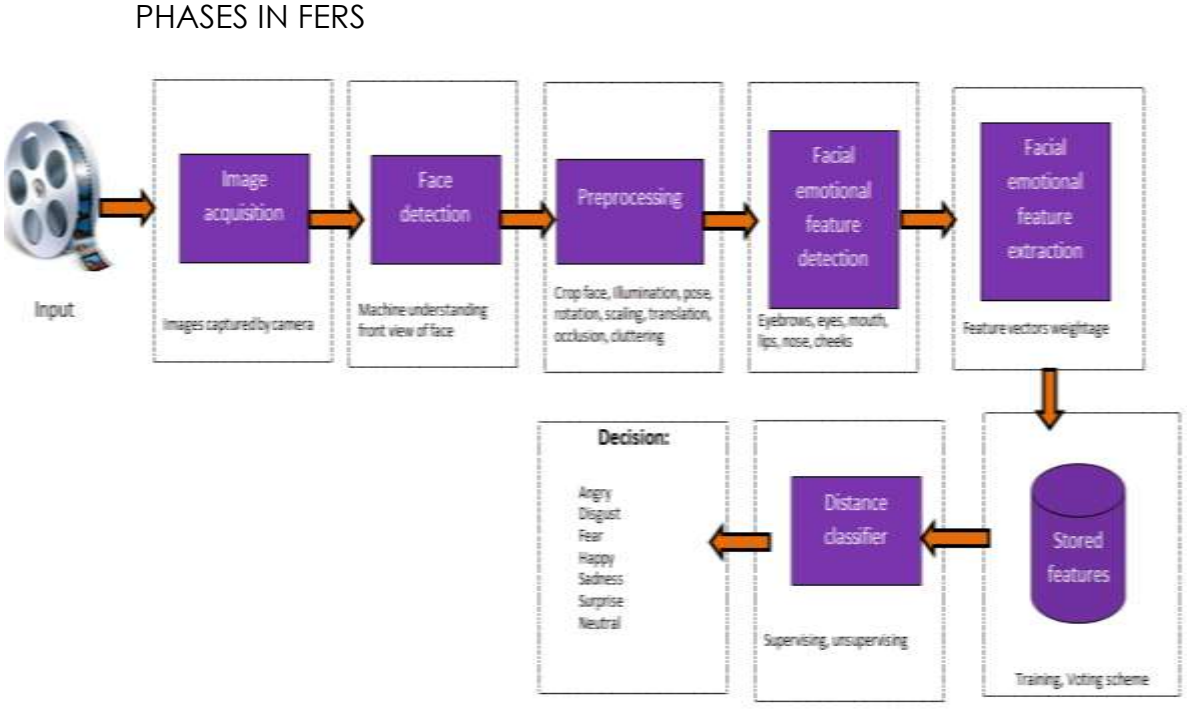

Fig. 2: Structure of facial expression recognition system.

Every FERS [Fig.2] must perform a few steps before to classifying the expression into a particular emotion. FERS categorized into following phases.

Image acquisition

Without image acquisition, the processing on the image is not possible so the first step of the FERS system is to capture an image. This phase would normally be done by using the standard video Camera. The Image acquisition phase itself includes several issues like image properties, a number of devices connected (camera, digitizer), the size of the face image, total image dimensions [Table 1], ambient lighting and so on. All of these factors may influence the facial expression analysis. Images acquired in low light or coarse resolution can provide less information about facial features. The methods that work well in studio lighting may perform poorly in more natural lighting when the angle of lighting changes across an image sequence. Most researchers using single-camera setups, usage of a single camera may difficult to standardize when the input is in out of a plane rotation in this circumstance multi camera setup is required, it supports the three-dimensional modeling and assesses the accuracy of the image alignment.

Table 1: Similar size face with different resolution

Same size face				
Various resolution	96 x128	69 x 93	48 x 64	24 x 32
Detect?	Yes	Yes	Yes	Yes
Pose?	Yes	Yes	Yes	Yes
Recognize?	Yes	Yes	Yes	May be
Features?	Yes	Yes	May be	No
Expressions?	Yes	Yes	May be	No

Pantic and Rothkrantz [7] were the first to use two cameras mounted on a headphone-like device; one camera is placed in front of the face and the other on the right side of the face. The cameras are moving together with the head to eliminate the scale and orientation variance of the acquired face images. Otsuka and Ohya [8] applied an adapted gradient-based optical flow algorithm to estimate the motion in the local facial [9] areas of the right eye and the mouth. The inputs of the facial images are acquired by a camera mounted on a helmet worn by the subject and sub sampled by eight in both directions. Professional grade PAL cameras provide very high-resolution images and by contrast, security cameras provide images that are seriously degraded.

Face localization

The face detection is a process of identifying the location of the face in the input image. Face detection method uses computer learning technique to detect the location [10] of the any faces in an image. For better performance in facial expression recognition, the importance of the facial landmark detection [11] is an undeniable fact. Face detection uses mathematical techniques on the pixel values or features in the facial area of an image to determine the human face. The human face is a dynamic object and has a high degree of variability in its appearance, which makes face detection [12] a difficult problem in computer vision. Face detection is the stepping stone to all the facial analysis algorithms, including face alignment, face relighting, face modeling, face recognition, head poses tracking, face verification/authentication, facial expression tracking/recognition and gender/age recognition. Only when computers can recognize faces well will they begin to truly understand people's thoughts and intentions.

Face detection appears as a trivial task for human beings, but an extremely tough task for computers. The difficulty associated with face detection [13] can be attributed to variations in scale, location, orientation (in-plane rotation), pose (out-of-plane rotation), facial expression, occlusions, illumination and so on. In the dynamic form of the input, face detection is performed to detect human faces in a plurality of frames. The location of the face in an image is detected using the well-known Viola–Jones method that's based on the Haar-like features and the AdaBoost learning algorithm [14]. The Viola–Jones method is familiar to an object detection algorithm providing competitive object detection rates in real time. The features used by Viola–Jones are derived from pixels selected from rectangular areas, which are imposed over the picture and show high sensitivity to the vertical and horizontal lines. [Fig. 3] show the different approaches used to detect the human face.

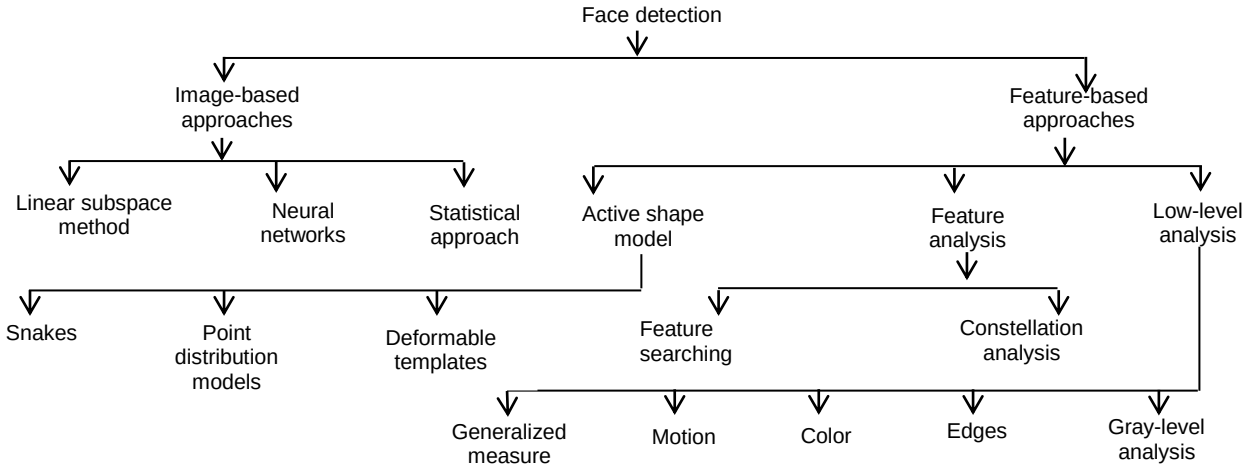

Fig. 3: Taxonomy of different face detection approaches.

Preprocessing

Image pre-processing techniques make the image easier to process the data and increase the chances of getting correct matches. Better chances of success with the change in illumination, pose, and picture quality can decrease the processing time and increase the feature detection performance, compared with a non-preprocessed image. Common pre-processing methods are resampling, edge detection and face alignment is an essential step and is usually carried out by detection and horizontal positioning of the eyes. A change in lighting conditions provides a considerable decrease in recognition performances. To solve this problem better preprocessing methods should be used before the feature extraction stage. The common preprocessing methods used to normalize the illumination [16] in facial images are gamma intensity correction, histogram equalization, logarithm transform and cosine transform. Gabor filter is the robust preprocessing technique in the FERS.

Feature extraction

Facial feature extraction is an important part in the FERS. The emotional model features are the features used to represent emotion in the face. It is the process of extracting relevant information from the image. The facial elements, especially key elements are constantly changing their positions when subjects are expressing emotions in other word feature extraction [Fig. 4] refers to the identification of facial feature points [17], preferably independent, which together represents a given emotional expression, which is responsible for particular facial behavior. The prominent emotional features of the face such as eyebrows, eyes, pupil diameter, nose, mouth, lips, and chin. The contours of the mouth, eyes, and eyebrows are playing an important role in the classification.

The accurate feature extraction method heavily depends upon the selection of the database and selection of the emotional features affects the classification accuracy. Feature extraction and representations are critical in FERS. The facial features can be extracted from either dynamic image sequences [18] or static images [19].The choices of features employed for emotion recognition are classified into two main categories, i.e., geometric features (facial points and shapes of facial components) and appearance features (descriptions of the texture of the face). Gabor wavelet-based technique& local binary pattern method is widely used for the facial emotional feature extraction process it can detect multi-scale, multi-direction texture changes [20], as well as by the effects of illumination is relatively small. The output of the feature extractor stage must contain separable and classifiable vectors.

10	60	71
150	65	53
140	200	18

3x3 neighborhood

0	0	1
1		0
1	1	0

Threshold

1	2	4
128		8
64	32	16

Binary weights

Pattern 11100100
LBP=4+32+64+128=228

Local binary pattern algorithm [LBP] is familiarly used in the facial emotion recognition system to extract the emotional features. LBP method is introduced by Ojala, it is a robust method for feature extraction. In a 3 x 3 image pixel are threshold by its center pixel value.

The neighborhood consists of 8 pixels; a total of 2^8 =256 different labels can be obtained depending on the relative gray values of the center and the pixels in the neighborhood. In a given image a pixel at (x_c ,y_c), the LBP result can be expressed in decimal format as

$$LBP(xc, yc) = \sum_{p=0}^{p-1} s(gp - gc)2^n$$

n-runs over the 8 neighbor of the central pixel, g_p and g_c are gray level values of the central pixel and the surrounding pixels, the function s(x) is defined as $s(x)=\begin{cases}1 \; if x >\_ \; 0 \\ 0 \; if x < \; 0\end{cases}$, limitation in the LBP operator is noise sensitivity and lack of rotational invariance.

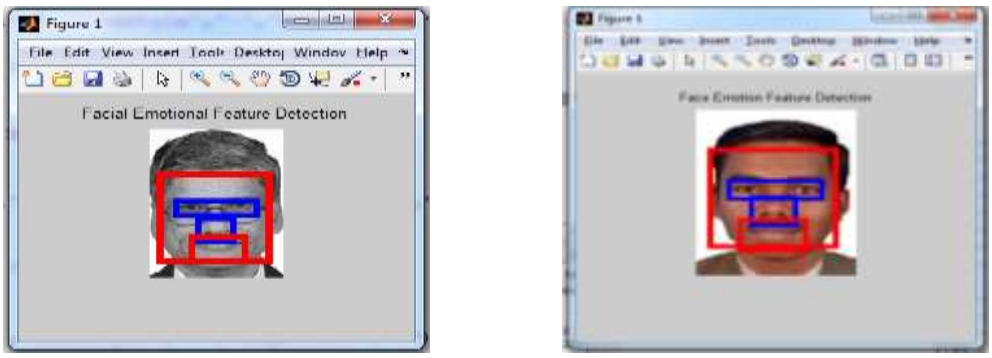

Fig. 4: Emotional features extraction from various faces

Database

To implement the new facial expression recognition system one of the most important aspects is the choice of the database [21] that will be used for testing the new system. The main drawback in the facial emotion recognition system is no standard database available it still an open challenge problem, the main reason is human face is defined with respect to age. If a common database is used by all the researchers, then testing the new system, comparing it with the other state of the art systems and benchmarking the performances are becoming a very easy and straightforward job. Table 2 describes the familiar databases available for facial emotion recognition system.

Table 2: Analysis-familiar facial emotion recognition database

S.no.	Name of database	Year	Image size	No. of picture per head	No. of unique people	Availability
1	FEI face database	2012	640x480	14	200	Yes
2	The MUCT face Database	2010	480x640	10-15	276	Yes
3	BU-3D FE database	2006	512x512	25	100	Yes
4	MMI database	2005	720x576	1280video	43	Yes
5	The yale face database	2003	320x243	11	15	Yes
6	PIE database	2002	640x486	9	68	Yes
7	CK database	2000	256x256	-	182	Yes
8	FERET database	2000	256x384	-	1,136	Yes
9	JAFFE database	1998	256x256	7	10	Yes
10	AR face database	1998	576x768	26	70male+56 female=126	Yes
11	Richard's MIT database	1996	480x640	6	82male+72female=154	Yes

MUCT-Milborrow University of Cape Town,BU-3DFE-Binghamton University 3D Facial Expression, CK-Cohn Kanad, FERET-Facial Recognition Technology, JAFFE-Japanese Female Facial Expression Database, AR-Aleix Martinez and Robert Benavente, MIT-Massachusetts Institute of Technology.

Expression classifier

After completion of the feature extraction phase, the final piece in the puzzle of facial emotion recognition system is a good classification [22] module. A classification aims at mapping emotional features into one of the six emotional labels. Information extracted from the face is the input for expression classifier that will classify the extracted features into particular expression. Image classification is a method of teaching the computer to make and improve predictions or behaviors based on some data. The robust algorithm

used to classify the facial expression is a Support Vector Machine (SVM) is a valid machine learning method [Fig. 5], it was originally proposed by C Cortes and V Vapnik [23] in the year of 1995.

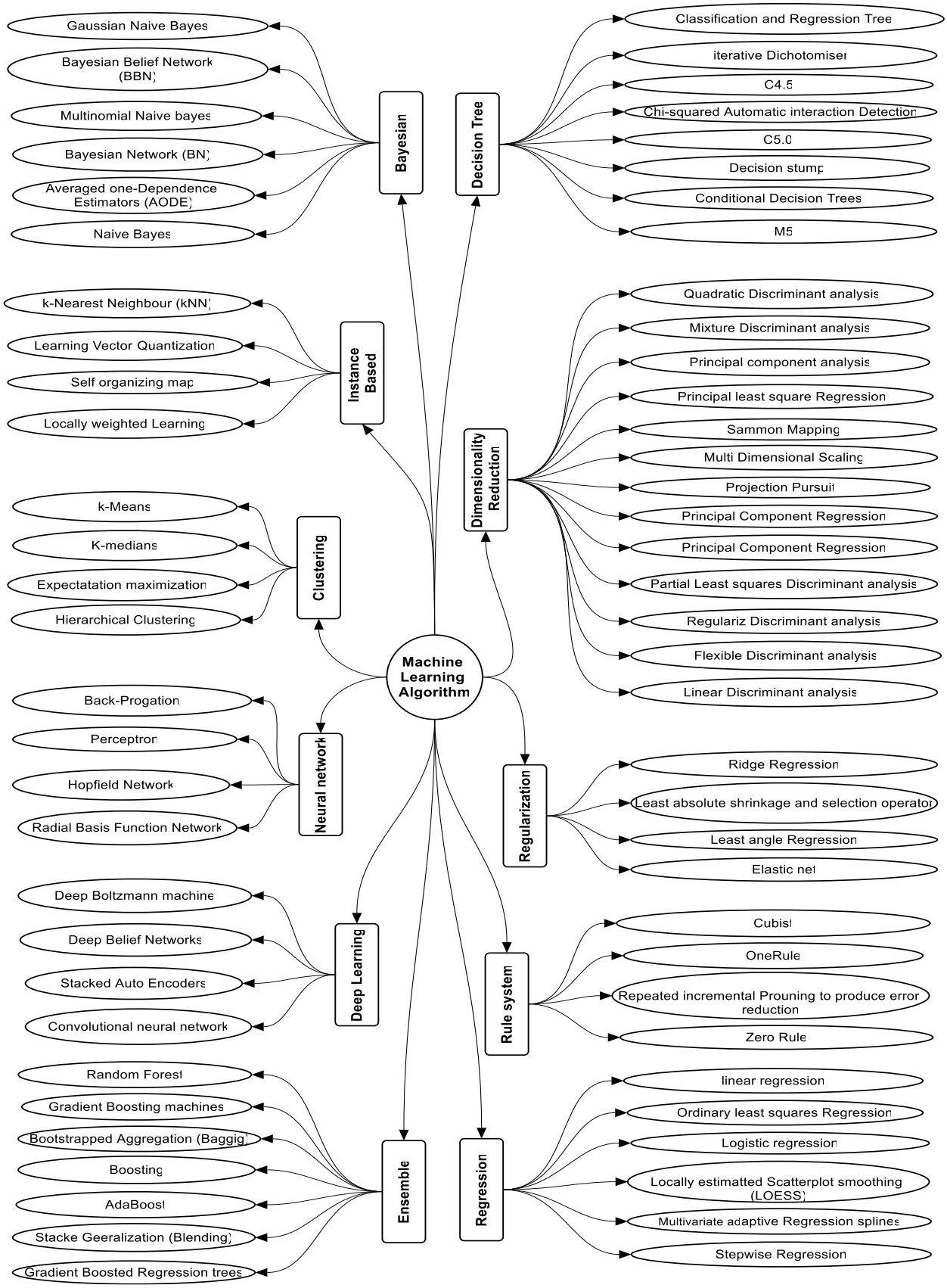

Fig. 5: Mind map-machine learning algorithm.

SVM is a supervised learning method used for classification and regression. A special property is that they

simultaneously minimize the empirical classification error and maximize the geometric margin, hence they are also known as maximum margin classifiers the dominating approach to doing so is to reduce the single multiclass problem into multiple binary classification problems. Each of the problems yields a binary classifier, which is assumed to produce an output function that gives relatively large values for examples from the positive class and relatively small values for examples belonging to the negative class. Kernels-support vector machines use a particular type of function class classifiers with large margins in a feature space induced by a kernel. The fuzzy rule-based system has been used for classification of six facial expressions of the face. Generally, each fuzzy image processing system has three main stages: image fuzzification, modification of membership values i.e., fuzzy Inference System and if necessary, image defuzzification. The emotion recognition systems of neural network algorithm [24] consist of three stages i.e. training of neural network, testing of neural network and performance evaluation of neural network. The accuracy of the emotion classification doesn't depend on only the selection of classifier algorithm; it is a framework of entire FERS.

EXPERIMENTAL RESULTS

Table 3 describes some recent works done on the facial emotion recognition system.

Table 3: Comparison of recent work results

S.no	Author	Feature extraction	Classifier	Database	Recognition rate	Emotion model
1.	S.l. happy, aurobind and routray (2015-IEEE) [25]	LBP	SVM	CK+	94.39%	6
2	Pojalachiranjeevi, et. al's (2015-IEEE) [26]	Key emotional) point	Personalized appearance model	CK+ ISL	98%	6
3	Ali moeini, et.al's (2015-IEEE) [27]	Feature library matrix	Support vector machine	FERET	99.09%	6+1 (neutral)
				CMU-PIE	98.24%	
4	Amin mohammadian, et.al's (2015) [28]	Feature vector	Style transfer mapping classifier	CK+	83.9%	6
5	Mostafa k et.al's (2014-IEEE) [29]	PCA	SVM	BU-3DFE	89.5%	6
6	Vassilis g. kaburlasos, et.al's (2013-IEEE)[30]	Orthogonal moment segmentation	Fuzzy ARTMAP neural classifier	JAFFE	93.96%	6
7	Jeemonikalita, et.al's (2013) [31]	Eigen vector	Euclidean distance	Self-database	95%	6
8	Yongqiang li, et.al's(2013-IEEE) [32]	Gabor transform	Dynamic bayesian network	CK+	87.43%	6
9	Michel f. valstar, et.al's (2012-IEEE) [33]	Gabor feature	Gentle boost+SVM	CK	95.3%	6
10	Govindukhandraokharat, et.al's(2008) [34]	Discrete cosine transform	Feed forward neural network	Self- database	100%	6

Table 4: System success rate of classification

Expression	Number of images experimented	Number of correct recognition	Success rate
Anger	25	22	88
Disgust	25	23	92
Fear	25	22	88
Happy	25	25	100
Sad	25	20	80
Surprise	25	24	96

Table 4 shows the successive of recognition rate 91% obtained for 150 samples with various expressions.

CONCLUSION

It is tricky to find the good coconut without opening similarly human emotion recognition system is mind intention related research work. In this paper briefly overviewed different phases in human-machine interaction in facial emotion recognition system. In the development of FERS fronting of various issues in each phase and robust practical algorithm available to solve various issues at each level are highlighted. This paper extremely guide & boost to the researcher who is new to in the area of human-machine interactions. Our future enhancement work will concentrate on the context aware online tutoring system i.e. our system automatically adjust the presentation style of an online tutor when negative attitude appeared on the candidate.

CONFLICT OF INTEREST
There is no conflict of interest.

ACKNOWLEDGEMENTS
The authors would like to thank the Reviewers for their detailed reviews and constructive comments which have helped to improve the quality of this paper and also Dr. Sujatha Balasubramaniam CEO GKM group of institution for providing facility.

FINANCIAL DISCLOSURE
None.

REFERENCES

[1] Mehrabian. [1968] A Communication without words. Psychology Today, 9(2):52-55.
[2] Darwin C. [1872] The Expression of the Emotions in Man and Animals. John Murray.
[3] Ekman P, Friesen W V. [1978] Facial Action Coding System: A Technique for the Measurement of Facial Movement. Palo Alto: Consulting Psychologists Press.
[4] Ying-li Tian, et.al. [2001] Recognition Action Units for Facial Expression Analysis. IEEE Transactions on pattern analysis and machine intelligence, 23(2):1-19.
[5] P. Ekman and W.V. Friesen. [1978] The Facial Action Coding System: A Technique for the Measurement of Facial Movement. Consulting Psychologist Press, San Francisco.
[6] M. Suwa, N. Sugie, and K. Fujinmora. [1978] A Preliminary Note on Pattern Recognition of Human Emotional Expression. In proceedings of 4th International Conference on Pattern Recognition, 408-410.
[7] Pantic and Rothkrantz. [2000] Automatic Analysis of Facial Expressions: The state of the art. Transactions on pattern analysis and machine intelligence, 22(12):1424-1445.
[8] Takahiro Otsuka and J. Ohya. [1996] Recognition of Facial Expressions Using HMM with Continuous Output Probabilities. Proceeding International Workshop Robot and Human Communication, 323-328.
[9] Sakmongkon Chumkamon and Eiji Hayashi. [2013] Facial expression recognition using constrained local models and Hidden Markov models with consciousness-based architecture. SICE International Symposium on System Integration.
[10] Stefanos Zafeiriou, Cha Zhang and Zhengyou Zhang. [2015] A Survey on Face Detection in the wild: past, present and future. ELSEVIER, Science Direct, 1-33.
[11] GN Matre, SK Shah. [2013] Facial Expression Detection. IEEE International conference on computational intelligence and computing research.
[12] Kah Phooi Seng. [2004] Improved automatic face detection technique in color images. TENCON IEEE Region 10 Conference, 1:459-462.
[13] Erik Hjelmas, Boon Kee Low. [2001] Face Detection: A Survey. Computer Vision and Image Understanding, 236-274.
[14] Ebenezer Owusu et.al. [2014] A neural-AdaBoost based facial expression recognition system. ELSEIVER, 41(7):3383-3390.
[15] Bobick AF. [2002] The Recognition of Human Movement using Temporal Templates. Pattern Analysis and Machine Intelligence, 23(3):257-267.
[16] Haifeng Hu. [2015] Illumination invariant face recognition based on dual-tree complex wavelet transform. IET compute vision, 9(2):163-173.
[17] Robert Brunelli, Tomaso Poggio. [1993] Face Recognition: Features versus Templates. IEEE Transactions on Pattern Analysis and Machine Intelligence, 15(10):1042-1052.
[18] Yongqiang Li, Shangfei Wang. [2013] Simultaneous Facial Feature Tracking and Facial Expression Recognition. IEEE Transactions on image processing, 22(7):2559-2573.
[19] Ira Cohen. [2003] Facial expression recognition from video sequences: temporal and static modeling, Elseiver, 91:160-187.
[20] Guoying Zhao and Matti Pietik¨ainen. [2007] Dynamic Texture Recognition Using Local Binary Patterns with an Application to Facial Expressions. IEEE Transactions on pattern analysis and machine intelligence, 29(6):915-928.
[21] F Abdat, cMaaoui and APruski. [2011] Human computer Interaction using emotion Recognition from facial expression. UKsim 5th European symposium on computer modeling and simulation, 196-201.
[22] Ashim Saha, Anurag De, M .C. Pal and Nirmalya Kar. [2014] Different Techniques of Automatic Facial Expression Recognition: A Survey. Second international conferences on advances in computing, communication and information technology, 43-49.
[23] C Cortes, V Vapnik. [1995] Support vector networks. Machine Learning, 20(3):273-293.
[24] Anchal Garg and Rohitbajaji. [2015] Facial Expression Recognition and classification using hybridization of ICA, GA and Neural Network for Human Computer interaction. Journal of network communication and engineering technologies, 2(1):50-57.
[25] S.L.Happy, Aurobind and Routray. [2015] Automatic Facial Expression Recognition Using Features of Salient Facial Patches. IEEE Transactions on affective computing, 6(1):1-12.
[26] Pojala Chiranjeevi, Viswanath Gopala krishnan, and Pratibha Moogi. [2015] Neutral Face Classification using personalized appearance models for fast and robust emotion detection. IEEE Transactions on image processing, 24(9):2701-2711.
[27] Ali Moeini and Hossein Moeini. [2015] Real-World and Rapid Face Recognition toward Pose and Expression Variations via Feature Library Matrix. IEEE Transactions on Information Forensics and security, 10(5):969-984.
[28] Amin Mohammadian, Hassan Aghaeinia, Farzad Towhidkhah. [2015] Video-based facial expression recognition by removing the style variations", IET image processing, 9(7):596-603.
[29] Mostafa K. Abd El Meguid and Martin D. Levine. [2014] Fully Automated Recognition of Spontaneous Facial Expressions in Videos Using Random Forest Classifiers. IEEE Transactions on affective computing, 5(2):141-154.
[30] Vassilis GKaburlasos, Stelios E Papadakis, and George A Papakostas. [2013] Lattice Computing Extension of the FAM Neural Classifier for Human Facial Expression Recognition. IEEE Transactions on neural networks and learning systems, 24(10):1526-1538.
[31] Jeemoni Kalita and Karen Das. [2013] Recognition of facial expression using eigenvector based distributed features and Euclidean distance based decision making technique. International Journal of advanced Computer science and applications, 4(2):196-202
[32] Yongqiang Li, Shangfei Wang, Yongping Zhao, and QiangJi. [2013] Simultaneous Facial Feature Tracking and Facial Expression Recognition. IEEE Transactions on image processing, 22 (7):28-43
[33] Michel F. Valstar, Maja Pantic. [2012] Fully Automatic Recognition of the Temporal Phases of Facial Actions. IEEE Transactions on systems, man and cybernetics, 42(1):2559-2573.
[34] Govind Ukhandrao Kharat and Sanjay V. Dudul. [2008] Neural Network Classifier for Human Emotion Recognition from Facial expressions using Discrete Cosine Transform. ICETET First international conference, 653-658.