ABSTRACT:

Human Face expression Recognition is one of the most powerful and challenging tasks in social communication. Generally, face expressions are natural and direct means for human beings to communicate their emotions and intentions. Face expressions are the key characteristics of [non-verbal communication](https://www.sciencedirect.com/topics/computer-science/nonverbal-communication). This paper describes the survey of Face Expression Recognition (FER) techniques which include the three major stages such as [preprocessing](https://www.sciencedirect.com/topics/computer-science/preprocessing), [feature extraction](https://www.sciencedirect.com/topics/computer-science/feature-extraction) and classification. This survey explains the various types of FER techniques with its major contributions. The performance of various FER techniques is compared based on the number of expressions recognized and complexity of algorithms. Databases like JAFFE, CK, and some other variety of [facial expression](https://www.sciencedirect.com/topics/computer-science/facial-expression) databases are discussed in this survey. The study on classifiers gather from recent papers reveals a more powerful and reliable understanding of the peculiar characteristics of classifiers for research fellows.

Classification

Face Expression Recognition (FER)

Feature extraction

Preprocessing

**1. Introduction**

Human [facial expressions](https://www.sciencedirect.com/topics/computer-science/facial-expression) are extremely essential in social communication. Normally communication involves both verbal and nonverbal. [Non-verbal communications](https://www.sciencedirect.com/topics/computer-science/nonverbal-communication) are expressed through facial expressions. Face expressions are the delicate signals of the larger communication. Non-verbal communication means communication between human and animals through eye contact, gesture, facial expressions, body language, and paralanguage.

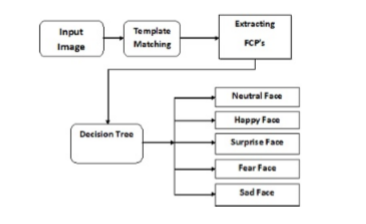
Eye contact is the important phase of communication which provides the mixture of ideas. Eye contact controls the contribution, discussions and creates a link with others. Face expressions include the smile, sad, anger, disgust, surprise, and fear. A smile on human face shows their happiness and it expresses eye with a curved shape. The sad expression is the feeling of looseness which is normally expressed as rising skewed eyebrows and frown. The anger on human face is related to unpleasant and irritating conditions. The expression of anger is expressed with squeezed eyebrows, slender and stretched eyelids. The disgust expressions are expressed with pull down eyebrows and creased nose. The surprise or shock expression is expressed when some unpredicted happens. This is expressed with eye-widening and mouth gaping and this expression is an easily identified one. The expression of fear is related with surprise expression which is expressed as growing skewed eyebrows.

FER has the important stage is [feature extraction](https://www.sciencedirect.com/topics/computer-science/feature-extraction) and classification. Feature extraction includes two types and they are geometric based and appearance based. The classification is also one of the important processes in which the above-mentioned expressions such as smile, sad, anger, disgust, surprise, and fear are categorized. The geometrically based feature extraction comprises eye, mouth, nose, eyebrow, other facial components and the appearance based feature extraction comprises the exact section of the face ([Zhao and Zhang, 2016](https://www.sciencedirect.com/science/article/pii/S1319157818303379" \l "b0270)).

Generally, the face offers three different types of signals such as static, slow and rapid signals. The static signals are skin color which includes the several lasting aspects of face skin pigmentation, greasy deposits, face shapes, the constitution of bones, cartilage and shape, location and size of [facial features](https://www.sciencedirect.com/topics/computer-science/facial-feature) such as brows, eyes, nose, mouth. The slow signals are permanent wrinkles which include the changes in [facial appearance](https://www.sciencedirect.com/topics/computer-science/facial-appearance) such as muscle tone and skin texture changes that happen slowly with time.

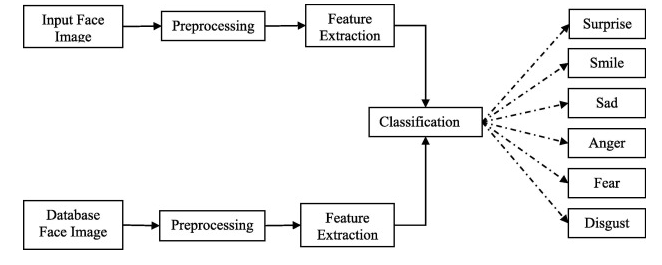
The rapid signals are raising the eyebrows which include the face muscles movement, impermanent face appearance changes, impermanent wrinkles and changes in the location and shape of facial features. These flashes on the face remain for a few seconds. These three signals are altered with individual option while it is very hard to alter static and slow signals. Also, the face is a multi-message system and it is not only a multi-signal system. Messages are transmitted through a face which includes emotion, feel position, age, quality, intelligence, [attractiveness](https://www.sciencedirect.com/topics/computer-science/attractiveness) and almost certainly other substances as well ([Ekman and Friesen, 2003](https://www.sciencedirect.com/science/article/pii/S1319157818303379" \l "b0060)).

This paper mainly focuses on various FER techniques with three major steps respectively [preprocessing](https://www.sciencedirect.com/topics/computer-science/preprocessing), feature extraction and classification. Also, this paper shows the advantages of different FER techniques and the performance analysis of different FER techniques. In this paper, only the image based FER techniques are chosen for the literature review and the video based FER techniques are not chosen. Mostly FER systems meet the problems of variation in illumination, pose variation, lighting variations, skin tone variations. Also this paper gives an essential research idea for future FER research.



Face expression recognition system

The overview of the FER system is illustrated .The FER system includes the major stages such as face image [preprocessing](https://www.sciencedirect.com/topics/computer-science/preprocessing), [feature extraction](https://www.sciencedirect.com/topics/computer-science/feature-extraction) and classification.

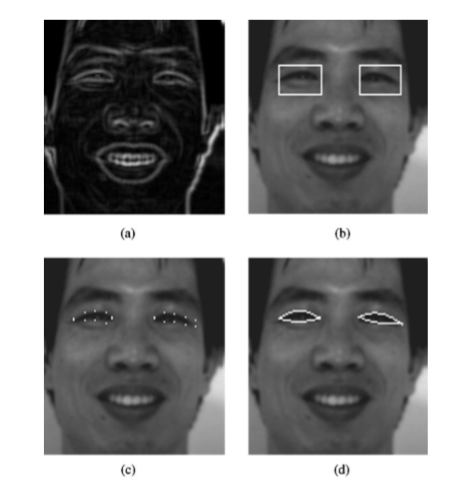


### 2.1. Preprocessing

Preprocessing is a process which can be used to improve the performance of the FER system and it can be carried out before feature extraction process ([Poursaberi et al., 2012](https://www.sciencedirect.com/science/article/pii/S1319157818303379" \l "b0175)). Image preprocessing includes different types of processes such as image clarity and scaling, contrast adjustment, and additional enhancement processes ([Bashyal et al., 2008](https://www.sciencedirect.com/science/article/pii/S1319157818303379" \l "b0015)) to improve the expression frames ([Taylor et al., 2014](https://www.sciencedirect.com/science/article/pii/S1319157818303379" \l "b0220)).

The cropping and scaling processes were performed on the face image in which the nose of the face is taken as midpoint and the other important facial components are included physically ([Zhang et al., 2011](https://www.sciencedirect.com/science/article/pii/S1319157818303379" \l "b0255)). Bessel down sampling is used for face image size reduction but it protects the aspects and also the perceptual worth of the original image ([Owusu et al., 2014](https://www.sciencedirect.com/science/article/pii/S1319157818303379" \l "b0170)). The [Gaussian](https://www.sciencedirect.com/topics/computer-science/gaussian) filter is used for resizing the input images which provides the smoothness to the images ([Biswas, 2015](https://www.sciencedirect.com/science/article/pii/S1319157818303379" \l "b0020)).

Normalization is the preprocessing method which can be designed for reduction of illumination and variations of the face images ([Ji and Idrissi, 2012](https://www.sciencedirect.com/science/article/pii/S1319157818303379" \l "b0120)) with the [median filter](https://www.sciencedirect.com/topics/computer-science/median-filter) and to achieve an improved face image. The [normalization method](https://www.sciencedirect.com/topics/computer-science/normalization-method) also used for the extraction of eye positions which make more robust to [personality differences](https://www.sciencedirect.com/topics/computer-science/personality-difference) for the FER system and it provides more clarity to the input images. Localization is a preprocessing method and it uses the Viola-Jones algorithm ([Noh et al., 2007](https://www.sciencedirect.com/science/article/pii/S1319157818303379" \l "b0165), [Demir, 2014](https://www.sciencedirect.com/science/article/pii/S1319157818303379" \l "b0055), [Zhang et al., 2014](https://www.sciencedirect.com/science/article/pii/S1319157818303379" \l "b0260), [Cossetin et al., 2016](https://www.sciencedirect.com/science/article/pii/S1319157818303379" \l "b0040), [Salmam et al., 2016](https://www.sciencedirect.com/science/article/pii/S1319157818303379" \l "b0195)) to detect the facial images from the input image. Detection of size and location of the face images using Adaboost [learning algorithm](https://www.sciencedirect.com/topics/computer-science/learning-algorithm) and haar like features ([Happy et al., 2015](https://www.sciencedirect.com/science/article/pii/S1319157818303379" \l "b0090), [Mahersia and Hamrouni, 2015](https://www.sciencedirect.com/science/article/pii/S1319157818303379" \l "b0145)). The localization is mainly used for spotting the size and locations of the face from the image.



### Feature extraction

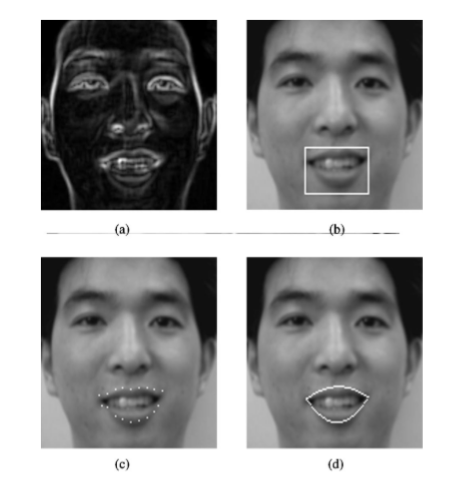
Feature extraction process is the next stage of FER system. Feature extraction is finding and depicting of positive features of concern within an image for further processing. In [image processing](https://www.sciencedirect.com/topics/computer-science/image-processing) [computer vision](https://www.sciencedirect.com/topics/computer-science/computer-vision) feature extraction is a significant stage, whereas it spots the move from graphic to implicit data depiction. Then these data depiction can be used as an input to the classification. The feature extraction methods are categorized into five types such as texture feature-based method, edge based method, global and local feature-based method, geometric feature-based method and patch-based method.

The descriptors which extract the features based on the texture feature-based methods are described as follows. Gabor filter is a [texture descriptor](https://www.sciencedirect.com/topics/computer-science/texture-descriptor) for feature extraction and it includes the magnitude and phase information. The Gabor filter with the magnitude feature confines the information about the organization of the face image. The phase feature precincts the information about the complete description of the magnitude features ([Bashyal and Venayagamoorthy, 2008](https://www.sciencedirect.com/science/article/pii/S1319157818303379" \l "b0015), [Owusu et al., 2014](https://www.sciencedirect.com/science/article/pii/S1319157818303379#b0170), [Zhang et al., 2014](https://www.sciencedirect.com/science/article/pii/S1319157818303379#b0260), [Hernandez-matamoros et al., 2015](https://www.sciencedirect.com/science/article/pii/S1319157818303379" \l "b0100), [Hegde et al., 2016](https://www.sciencedirect.com/science/article/pii/S1319157818303379" \l "b0095)). Local Binary Pattern (LBP) is also a texture descriptor and it can be used for feature extraction. Generally LBP features are produced with the binary code and it can be obtained by using thresholding between the center pixel and its locality pixels ([Happy et al., 2015](https://www.sciencedirect.com/science/article/pii/S1319157818303379#b0090), [Cossetin et al., 2016](https://www.sciencedirect.com/science/article/pii/S1319157818303379#b0040)).

The descriptors which extract the features based on patch-based methods are described as follows. Facial movement features are extracted as patches depending upon the distance characteristics. These are performed by using two processes such as extracting the patches and patch matching. The patch matching is performed by translating extracted patches into distance characteristics ([Zhang et al., 2011](https://www.sciencedirect.com/science/article/pii/S1319157818303379#b0255)).

The texture feature based descriptors are more useful feature extraction method than the others because it extracts the texture features like related to the appearance which provides the important feature vectors for FER. Also Local Directional Number (LDN) pattern ([Rahul and Cherian, 2016](https://www.sciencedirect.com/science/article/pii/S1319157818303379" \l "b0180)), Local Directional Ternary Pattern (LDTP) ([Ryu et al., 2017](https://www.sciencedirect.com/science/article/pii/S1319157818303379" \l "b0190)), KL-transform Extended LBP (K-ELBP) ([Guo et al., 2016](https://www.sciencedirect.com/science/article/pii/S1319157818303379" \l "b0080)) and Discrete [Wavelet Transform](https://www.sciencedirect.com/topics/computer-science/wavelet-transforms) (DWT) ([Nigam et al., 2018](https://www.sciencedirect.com/science/article/pii/S1319157818303379" \l "b0160)) texture feature based descriptors are used as feature descriptors in recent years FER.

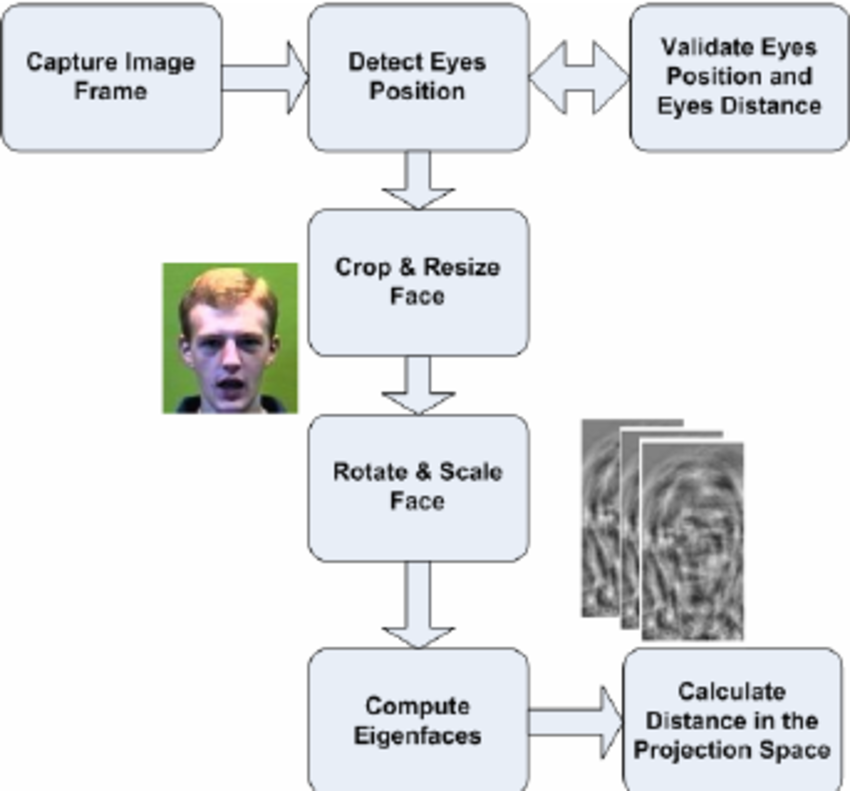
Several extracted features have high [dimensional vectors](https://www.sciencedirect.com/topics/computer-science/dimensional-vector). Generally these feature vectors are reduced by using various [dimensionality reduction](https://www.sciencedirect.com/topics/computer-science/dimensionality-reduction) algorithms such as PCA, Linear Discriminant Analysis, Whitened [Principle Component Analysis](https://www.sciencedirect.com/topics/computer-science/principle-component-analysis) and the important features are also selected with different algorithms such as Adaboost and [similarity scores](https://www.sciencedirect.com/topics/computer-science/similarity-score).



The descriptors which extract the features based on the edge based methods are described as follows. Line Edge Map (LEM) descriptor is a [facial expression](https://www.sciencedirect.com/topics/computer-science/facial-expression) descriptor which improves the geometrical structural features by using the dynamic two strip algorithm (Dyn2S) ([Gao et al., 2003](https://www.sciencedirect.com/science/article/pii/S1319157818303379" \l "b0075)). Based on the motion analysis two types of [facial features](https://www.sciencedirect.com/topics/computer-science/facial-feature) are extracted such as non discriminative and discriminative facial features ([Noh et al., 2007](https://www.sciencedirect.com/science/article/pii/S1319157818303379#b0165)). [Graphics-processing unit](https://www.sciencedirect.com/topics/computer-science/graphics-processing-unit) based Active Shape Model (GASM) is the feature extraction method which can be performed with edge detection, enhancement, tone mapping and local appearance model matching. After that the image ratio features are extracted from the expressed face images ([Song et al., 2010](https://www.sciencedirect.com/science/article/pii/S1319157818303379" \l "b0215)).

The descriptors which extract the features based on the global and local feature-based methods are described as follows. Principal [Component Analysis](https://www.sciencedirect.com/topics/computer-science/component-analysis) (PCA) method is used for feature extraction. It extracts the global and low dimensional features. Independent Component Analysis (ICA) is also a feature extraction method which extracts the local features using the multichannel observations ([Taylor et al., 2014](https://www.sciencedirect.com/science/article/pii/S1319157818303379#b0220)). Stepwise [Linear Discriminant Analysis](https://www.sciencedirect.com/topics/computer-science/linear-discriminant-analysis) (SWLDA) is the feature extraction technique which extracts the localized features with backward and forward [regression](https://www.sciencedirect.com/topics/computer-science/regression) models. Depends on the class labels the F-test values are estimated for both regression models ([Siddiqi et al., 2015](https://www.sciencedirect.com/science/article/pii/S1319157818303379" \l "b0210)).

The descriptors which extract the features based on the geometric feature-based methods are described as follows. Local Curvelet Transform (LCT) is a feature descriptor which extracts the geometric features which depends on wrapping mechanism. The extracted geometric features are mean, [entropy](https://www.sciencedirect.com/topics/computer-science/entropy) and standard deviation ([Demir, 2014](https://www.sciencedirect.com/science/article/pii/S1319157818303379#b0055)). Addition to these geometrical features energy, kurtosis are extracted by using three stage steerable pyramid representation ([Mahersia and Hamrouni, 2015](https://www.sciencedirect.com/science/article/pii/S1319157818303379" \l "b0145)).



### Classification

Classification is the final stage of FER system in which the classifier categorizes the expression such as smile, sad, surprise, anger, fear, disgust and neutral.

The [directed Line segment](https://www.sciencedirect.com/topics/computer-science/directed-line-segment) [Hausdorff Distance](https://www.sciencedirect.com/topics/computer-science/hausdorff-distance) (dLHD) method is used for recognition of expressions ([Gao et al., 2003](https://www.sciencedirect.com/science/article/pii/S1319157818303379#b0075)). [Euclidean distance](https://www.sciencedirect.com/topics/computer-science/euclidean-distance) metric is also used for classification purpose which uses the normalized score and similarity score matrix for estimating Euclidean distance ([Hegde et al., 2016](https://www.sciencedirect.com/science/article/pii/S1319157818303379#b0095)). Minimum Distance Classifier (MDC) is also one of the distance based classifier used for classification which estimates the distance between the feature vectors every sub image ([Islam et al., 2018](https://www.sciencedirect.com/science/article/pii/S1319157818303379" \l "b0110)). The KNN (k – Nearest Neighbors) algorithm is a classification method in which the relationship among the [assessment models](https://www.sciencedirect.com/topics/computer-science/assessment-model) and the other models are estimated during the training stage ([Poursaberi et al., 2012](https://www.sciencedirect.com/science/article/pii/S1319157818303379" \l "b0175)).

[Support Vector Machine](https://www.sciencedirect.com/topics/computer-science/support-vector-machine) (SVM) is one of the [classification techniques](https://www.sciencedirect.com/topics/computer-science/classification-technique) in which two types of approaches are involved. They are one against one and one against all approaches. One against all classification means it constructs one sample for each class ([Zhao and Pietikäinen, 2009](https://www.sciencedirect.com/science/article/pii/S1319157818303379" \l "b0265), [Zhang et al., 2011](https://www.sciencedirect.com/science/article/pii/S1319157818303379#b0255), [Zhang et al., 2014](https://www.sciencedirect.com/science/article/pii/S1319157818303379#b0260), [Biswas, 2015](https://www.sciencedirect.com/science/article/pii/S1319157818303379#b0020)). One against one classification means it constructs one class for each pair of classes ([Happy et al., 2015](https://www.sciencedirect.com/science/article/pii/S1319157818303379#b0090), [Kumar et al., 2016](https://www.sciencedirect.com/science/article/pii/S1319157818303379" \l "b0125), [Hegde et al., 2016](https://www.sciencedirect.com/science/article/pii/S1319157818303379#b0095)) and SVM is one of the strongest classification methods for advanced [dimensionality](https://www.sciencedirect.com/topics/computer-science/dimensionality) troubles ([Dahmane and Meunier, 2014](https://www.sciencedirect.com/science/article/pii/S1319157818303379" \l "b0050)). SVM is the supervised [machine learning technique](https://www.sciencedirect.com/topics/computer-science/machine-learning-technique) and it uses four types of kernels for its better performance ([Hernandez-matamoros et al., 2015](https://www.sciencedirect.com/science/article/pii/S1319157818303379#b0100)). They are [linear, polynomial](https://www.sciencedirect.com/topics/computer-science/linear-polynomial), [Radial Basis Function](https://www.sciencedirect.com/topics/computer-science/radial-basis-function) (RBF) and sigmoid. The linear kernel maps the [high dimensional data](https://www.sciencedirect.com/topics/computer-science/high-dimensional-data) and it is linearly separable ([Zhang et al., 2014](https://www.sciencedirect.com/science/article/pii/S1319157818303379#b0260), [Kumar et al., 2016](https://www.sciencedirect.com/science/article/pii/S1319157818303379#b0125)). The RBF kernel uses the function that maps the single feature into the high dimensional data ([Song et al., 2010](https://www.sciencedirect.com/science/article/pii/S1319157818303379#b0215), [Wang et al., 2010](https://www.sciencedirect.com/science/article/pii/S1319157818303379" \l "b0230), [Dahmane and Meunier, 2014](https://www.sciencedirect.com/science/article/pii/S1319157818303379#b0050), [Happy et al., 2015](https://www.sciencedirect.com/science/article/pii/S1319157818303379#b0090), [Hegde et al., 2016](https://www.sciencedirect.com/science/article/pii/S1319157818303379#b0095)). The polynomial kernel learns the [nonlinear models](https://www.sciencedirect.com/topics/computer-science/nonlinear-model) and also resolves their similarity ([Zhao and Pietikäinen, 2009](https://www.sciencedirect.com/science/article/pii/S1319157818303379#b0265), [Zhang et al., 2011](https://www.sciencedirect.com/science/article/pii/S1319157818303379#b0255), [Ji and Idrissi, 2012](https://www.sciencedirect.com/science/article/pii/S1319157818303379#b0120), [Biswas, 2015](https://www.sciencedirect.com/science/article/pii/S1319157818303379#b0020)).

The [Hidden Markov Model](https://www.sciencedirect.com/topics/computer-science/hidden-markov-models) (HMM) classifier is the statistical model which categorizes the expressions into different types ([Taylor et al., 2014](https://www.sciencedirect.com/science/article/pii/S1319157818303379#b0220)). Hidden [Conditional Random Fields](https://www.sciencedirect.com/topics/computer-science/conditional-random-field) (HCRF) representation is used for classification. It uses the full covariance Gaussian distribution for superior classification performance ([Siddiqi et al., 2015](https://www.sciencedirect.com/science/article/pii/S1319157818303379#b0210)).

Online Sequential Extreme Learning Machine (OSELM) is a method that uses RBF for classification. OSELM mainly contains two stages. They are [initialization](https://www.sciencedirect.com/topics/computer-science/initialization) and sequential learning stages. [Initialization stage](https://www.sciencedirect.com/topics/computer-science/initialization-stage) includes the training samples ([Demir, 2014](https://www.sciencedirect.com/science/article/pii/S1319157818303379#b0055)). Pair wise classifiers are also used for expression classification. It uses the one against one classification approach so exacting separation is utilized ([Cossetin et al., 2016](https://www.sciencedirect.com/science/article/pii/S1319157818303379" \l "b0040)).

ID3 [Decision Tree](https://www.sciencedirect.com/topics/computer-science/decision-trees) (DT) classifier is a rule based classifier which extracts the [predefined rules](https://www.sciencedirect.com/topics/computer-science/predefined-rule) to produce competent rules. The predefined rules are generated from the decision tree and it was constructed by information gain metrics. The classification is performed using the least Boolean evaluation ([Noh et al., 2007](https://www.sciencedirect.com/science/article/pii/S1319157818303379#b0165), [Rashid, 2016](https://www.sciencedirect.com/science/article/pii/S1319157818303379" \l "b0185)). Classification and Regression Tree (CART) is a [machine learning algorithm](https://www.sciencedirect.com/topics/computer-science/machine-learning-algorithm) for classification. The metric likely Decision tree and Gini impurity are estimated. CART classifiers are signified by using the [distance vectors](https://www.sciencedirect.com/topics/computer-science/distance-vector) ([Salmam et al., 2016](https://www.sciencedirect.com/science/article/pii/S1319157818303379" \l "b0195)).

Learning [Vector Quantization](https://www.sciencedirect.com/topics/computer-science/vector-quantization) (LVQ) is the unsupervised [clustering algorithm](https://www.sciencedirect.com/topics/computer-science/clustering-algorithm) ([Bashyal et al., 2008](https://www.sciencedirect.com/science/article/pii/S1319157818303379" \l "b0015)) which has two layers namely competitive and output layers. The competitive layer has the neurons that are known as [subclasses](https://www.sciencedirect.com/topics/computer-science/subclasses). The neuron which is the greatest match in competitive layer then put high for the class of exacting neuron in the output layer. [Multi Layer Perceptron](https://www.sciencedirect.com/topics/computer-science/multilayer-perceptron) (MLP) is also used for classification and it contains three layers such as input layer, output layer and processing layer in which neurons are present ([Rashid, 2016](https://www.sciencedirect.com/science/article/pii/S1319157818303379#b0185)).

The Multilayer [Feed Forward Neural Network](https://www.sciencedirect.com/topics/computer-science/feedforward-neural-network) (MFFNN) classifier uses three layers such as input, hidden and output layers and [back propagation algorithm](https://www.sciencedirect.com/topics/computer-science/backpropagation-algorithm) for classification. In the training stage the weights are initialized and the activation units are estimated ([Owusu et al., 2014](https://www.sciencedirect.com/science/article/pii/S1319157818303379#b0170)). Bayesian [neural network](https://www.sciencedirect.com/topics/computer-science/neural-networks) classifier is the classification method which also includes three layers such as input, hidden and output layers. The classical back propagation algorithm is used with [Bayesian classifier](https://www.sciencedirect.com/topics/computer-science/bayesian-classifier) for its better accuracy ([Mahersia and Hamrouni, 2015](https://www.sciencedirect.com/science/article/pii/S1319157818303379" \l "b0145)). Convolution Neural Network (CNN) consists of two layers such as [convolutional layer](https://www.sciencedirect.com/topics/computer-science/convolutional-layer) and subsampling layer in which the two dimensional images are taken as input. In convolutional layer the feature maps are produced by intricate the [convolution kernels](https://www.sciencedirect.com/topics/computer-science/convolution-kernel) with the two dimensional images where as in the subsampling layer, pooling and redeployment are performed ([Shan et al., 2017](https://www.sciencedirect.com/science/article/pii/S1319157818303379" \l "b0205)). The CNN also contains two important perceptions likely shared weight and sparse [connectivity](https://www.sciencedirect.com/topics/computer-science/connectivity) ([Rashid, 2016](https://www.sciencedirect.com/science/article/pii/S1319157818303379#b0185)). In FER, the CNN classifier used as multiple classifiers for the different face regions. If CNN is framed for entire face image then first frame the CNN for mouth area and next for eye area likely for each other area CNNs are framed ([Cui et al., 2016](https://www.sciencedirect.com/science/article/pii/S1319157818303379" \l "b0045)).

[Deep Neural Network](https://www.sciencedirect.com/topics/computer-science/deep-neural-network) (DNN) contains various hidden layers and the more difficult functions are trained efficiently comparing with other neural networks ([Li and Lam, 2015](https://www.sciencedirect.com/science/article/pii/S1319157818303379" \l "b0130)). The Deep Belief Network (DBN) contains the hidden variable resides of the various number of Restricted [Boltzmann Machine](https://www.sciencedirect.com/topics/computer-science/boltzmann-machine) (RBM) which are the undirected generative pattern ([Lv, 2015](https://www.sciencedirect.com/science/article/pii/S1319157818303379" \l "b0135)). DBN contains the [Back Propagation](https://www.sciencedirect.com/topics/computer-science/backpropagation) (BP) layer classifies the high-level features using classification ([Yang et al., 2016](https://www.sciencedirect.com/science/article/pii/S1319157818303379" \l "b0240)). DBN generally includes two phases such as pre-learning and fine-tuning ([Wu and Qiu, 2017](https://www.sciencedirect.com/science/article/pii/S1319157818303379" \l "b0235)) in which RBM are developed separately in the first step whereas the BP are learning the input and output data in the last phase.

According to several classifiers SVM classifier gives better recognition accuracy and it provides better classification. The neural network based classifier CNN gives better accuracy than the other neural network based classifiers. In FER, SVM classifier is more exploitable comparing with other classifiers for recognition of expressions.

The various FER techniques with their algorithm is analyzed in [Table1](https://www.sciencedirect.com/science/article/pii/S1319157818303379" \l "t0005) which includes the algorithms that are used for three important requirements such as preprocessing, feature extraction and classification. The various preprocessing methods used in this table are, [face detection](https://www.sciencedirect.com/topics/computer-science/face-detection), [image enhancement](https://www.sciencedirect.com/topics/computer-science/image-enhancement), normalization, Gabor filter, localization, face acquisition, down sampling, histogram equalization, face region detection, face alignment, ROI segmentation and resizing. The different feature extraction methods used in this table are LEM, Action based model, Gabor filter, LBP-TOP, GASM, Patch based, GL wavelet, LBP, VTB, Moments, PCA, ICA, LCT, HOG, Steerable pyramid, DCT, SWLDA, WLD, SDM, WPLBP, haar like features, LDN, LDTP, DWT, K-ELBP, 2DPCA and eigenfaces. Classifiers used in this table are ID3 decision tree, LVQ, SVM, KNN, HMM, MFFNN, OSLEM, Bayesian neural network, HCRF, pair wise, CART, Euclidean distance, CNN, MDC, Chi square test and fisher discrimination dictionary.

