A Review of Facial Expression Recognition

[1]Priyanka Nathawat, [2]Dr. Vivek Chaplot
College name: Bhupal Noble's University nathawat.pre@gmail.com

Abstract – Emotions are very present in our daily life. They are manifested by physiological reactions (palpitations, heat, and acceleration of the pulse), motor reactions (gestural expressions, facial expressions) and changes in the voice. The complexity of emotions was capable to seek attention of various researchers. Differing representations of emotions were then proposed to produce a set of theories. The automation of facial expression detection can then be approached in a variety of ways. We concentrated on the face expression detection. Concerning to this the recognition is based on different visual characteristics of the expression.

This paper is dedicated at first to the exhibition of the definition of emotions, followed by the great theories that have emerged in this area. In a second step, we present the methods of extraction of the facial characteristics and the methods of selection. Finally, some databases used for automatic recognition of emotions are presented.

Keywords – AAM, AU, CAU, FACS, FAPs, LDN, Neural Network, STAAM, SVM.

1. INTRODUCTION

Face recognition is a vast research area with numerous problems. This is due to its numerous applications in different sectors such as Mugshot album, surveillance camera, identification card, access control card etc. The face classification was firstly introduced in. In this proposed paper, the author started gathering the faces for the database and classified it according to its deviation from norm. This generated vectors which could be compared with other database vectors. This is called multimodal classification. Due to the vast advancement, Face Recognition systems are being used in real-time applications widely [1].
From last few decades, face expression recognition technology is grabbing more and more interest of the researchers due to its numerous applications. Face expression recognition is a part of face detection. Other face detection methods include facial alignment, head alignment, recognizing facial expression, face authentication, age and gender of a person etc. According to the human vision, it is very easy to recognize the face of a person in any circumstances but it is that much difficult for computer vision because of image variations (e.g. face orientation, brightness of image, pose, ageing, occlusion, face expressions, hairstyle, makeup etc.) [2].

Faces are the most widely researched visual precipitant. Most experts investigate face processing as a group mean approach and that neutralizes the behavioral responses and considers the variations between individuals as ‘noise’. Somehow, these variations are actually the differences between faces of individuals that facilitate valuable insights about face processing that complements the findings of group-mean studies. It is depicted through the studies that the question about the operation of face processing can be answered through the association and dissociation examining of different approaches. In short, these researches permit the association and dissociation of mechanisms employed for face processing, relate behavioral and neural face processing mechanisms, link face processing to broader capacities and its influences. The individual differences approach that is illustrated over here is more dominant than others and hence holds the scope for further exploration in the field of face processing as well as in other cognitive sciences.

One survey extensively scrutinizes the neural as well as cognitive bases of face processing. Most of the studies consider group-mean mechanism which concentrates on neural as well as cognitive response and treats variations between individuals as ‘noise’. The experts here tend to spotlight a different growing approach of literature which doesn't consider such variations as noise, moreover it considers it as valuable signal. These two theories complement each other but are powerful in their own arena. They are the means of examining neural bases, functional organization and developments in skilled face processing. The proposed research not only aims at the existence of individual differences in face processing. The theoretical understanding is enhanced by shifting the focus towards associations and dissociations across individuals. This theory associate and dissociates behavioral face processing such as holistic processing or familiar/unfamiliar face processing with neural processing and determines its relationships. [3]

Since early 90s, the automated facial expression detection has been a topic-of-interest among researchers. Previously there has been enough research done for face recognition, tracking, extracting features and classification techniques. The proposed work takes into account such surveys and proposes a time-line view for the advancements done in the field of face expression recognition. It states the ideal characteristics of the system and the databases that are being used by face expression recognizers. Here, facial-parameterization using FACS-AUs (Action Units) and MPEG-4 Facial Animation Parameters (FAPs) are also discussed as well as the advancement is made concerning to the standardization and a detailed state-of-the-art summary. The studies propose the face expression is categorized on the basis of features, emotions, expressions. According to recent studies, there are mainly six prototypic expressions. This paper was basically introduced with the objective of guiding student as well as researcher new to this field. It also discusses the challenges and future work. [4]

The key challenge in the facial emotion recognition field is the automatic classification of static images on the basis of seven basic emotions. Automatic analysis and detection of face expressions has been a subject undergoing intense study especially for discrete emotion detection as well as FACS Action Unit (AU) detection. Numerous database exist for facial-expression-recognition depending upon the compatibility and standardization. However, the comparison of systems turns to be difficult if the system is lacking in sufficient details to reproduce reported details or in common evaluation protocol. This affects the improvement of the face recognition field. This comparison can be executed in a fair manner through periodical challenge in analysis and recognition of facial expression. This allows us to explore new goals, targets and challenges. This paper presents the first challenge in the field of automated face recognition. Two challenges of sub-category are defined as:

- AU detection
- Discrete-Emotion-detection

The protocol of evaluation is outlined here and the outcomes of two sub-challenges are also discussed. [5]
For efficient face analysis and face expression recognition, a local feature descriptor and local directional number pattern (LDN) is proposed here as a new technique towards facial expression detection. The directional data about the texture of the face is encoded through local directional number pattern in a more compacted manner that yields more selective code in comparison with existing approaches. Further, the structure of every single micro-pattern is computed with the help of compass mask which is responsible for extracting directional information. Such data is encoded by using prominent directional numbers (directional indices) and signs that permit the discrimination of identical patterns possessing distinct intensity-transitions. The face is divided into various regions so that it is possible to extract the distributed local directional number features from them. The mentioned features are further concatenated as feature-vectors and are used as a face-descriptor. Multiple experiments are carried out so that the descriptor can perform well under any circumstances like variations in expression, time lapse and noise. Furthermore, the performance of these descriptors is analyzed with different masks under different face analysis conditions. [6]

2. FACE EXPRESSION CATEGORIZED ON THE BASIS OF Emotions

The recent advancement in the field of expression recognition emphasize on detecting more natural facial expressions. There are basically seven universal categories of emotions judged from face expressions. These are as follows [7]:

- Expression of Anger
- Expression of Disgust
- Expression of Fear
- Neutral expressions
- Happy emotions
- Sad emotions
- Surprise emotions

On the basis of these seven emotion categories, the datasets from AFEW and SFEW were considered to capture more natural expressions. The AFEW dataset contains static images whereas; SFEW-dataset is comprised of videos from movies. The reason behind considering movie scenes is that the emotions in movies are more spontaneous as compared to imitated expressions considered in previous survey of lab-controlled datasets. The grand challenge held in 2015 for detection of emotion in the Wild (EmotiW) was actually comprised of two sub-categories based on AFEW 5.0 and SFEW 2.0 datasets respectively. Both the datasets present numerous challenges as compared to conventional datasets because of their more natural features. The various hand-crafted features (including and Local Quantized Patterns (LPQ), Local-Binary-Patterns on Three-Orthogonal-Planes (LBP-TOP) and Pyramid Histogram of Oriented Gradients (PHOG)) which is carried out well for traditional datasets show considerably lower performance for the proposed two AFEW 5.0 and SFEW 2.0 datasets. [8]

In this article [9] a new expression recognition algorithm has been presented which is simpler and faster than other existing algorithms. The key goal of this algorithm is to make available emotion awareness in real-time basic applications running in secondary plane. The proposed algorithm is carried out in three stages – pre-processing, feature extraction and machine learning. This standard algorithm is developed on the basis of extraction of 19 characteristic face landmarks which are used to define basic facial movements and also recognize some of the AUs defined in the FACS. Moreover, a new concept of CAUs is introduced which stands for Combined Action Units. Combined Action Units means number of Action Units (AUs) grouped together as a single unit. Here, the classification of emotions is performed on the basis of logical rules. Hence no learning or training is required. The initial implementations are done on mobile platform. The detections of AU and CAU is performed on the basis of a specific pattern obtained from comparing the actual and neutral (obtained) image. This pattern is obtained from a group of facial landmarks. Aus and CAUs are defined on the basis of three factors – distances between the points, motion vectors and area enclosed in the triangles. These three factors are then compared with pre-defined threshold values that justify the activation of an AU or CAU [10].

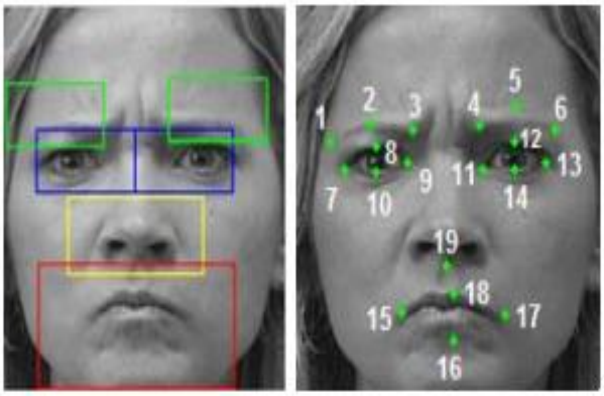


Figure 1: Detection of facial landmarks [10]

For accurate emotion detection, system is trained. This training is carried out in three stages –

1. Learning of Feature
2. Selection of Feature
3. Classifier modeling

The heuristic studies should be extensively performed for searching an optimal combination of feature-representation, feature-set, and classifier in order to attain better performance. A new method of Boosted Deep Belief Network (BDBN) has been proposed here for training the unified looped architecture in three stages. This proposed Boosted Deep Belief Network (BDBN) helps in learning specific set of features which efficiently classify expressions based on appearance or shape. Further, this training helps in the formation of strong boosted classifier in a statistical way. With continuous training and learning sessions, the classifier is strengthened as well as the capabilities to differentiate among various expressions also boosted relatively with the strong classifier through a joint fine-tuned process in the BDBN framework. The extensively performed heuristic studies state that the BDBN architecture produces considerable improvements in the field of face expression detection. [11]

A. Facial Expression Detection using Geometrical Feature Based Approach

The basic step in geometric-feature-based face expression recognition approach is the localization and tracking of densely situated facial points. For tracking dense facial points, almost all geometric feature-based approaches use AAM (active-appearance-model) and its variations. The extraction for the movement of facial features and shape of facial features is done with help of expression variation using locations of these features in different ways. In order to determine face emotions a recent AAM-based approach is proposed which compares various AAM algorithms and evaluates their performance. Certainly, chosen candid nodes are geometrically displaced and are given as the differences of node-coordinates among first and the greatest expression intensity frames. Further, it is utilized as an input for the novel multi-class SVM classifier. Detecting face in single image involves multiple geometrical features like:

- Facial features
- Skin colour
- Skin texture
- Multiple geometric features

These geometrical features are then used to detect emotions. This method employs manual detection of facial feature points and Piecewise-Bezier-volume-deformation tracking was utilized to trail those manual face points. Numerous machine learning techniques has been studied as well as experimented which states that the simple k-nearest neighbor technique provides the best results [12]. Another survey introduces STAAM (STereo-AAM) in order to enhance tracking and fitting of standard AAM with the help of multiple cameras which is used to model three-dimensional shape and rigid motion parameters. Furthermore, in order to combine three-dimensional shape and registered two-dimensional appearance, a generalized layer discriminant analysis classifier is used. A manifold was proposed for a geometric features-based approach tracking, modelling and recognition of emotions on a low-dimensional expression. For three-dimensional motion-based features, the extraction of offset and onset segment features is performed using feature selection methods. These features are then employed for training Gentle Boost classifiers, and a HMM. Another paper introduces a head-pose invariant face expression detection which uses a set of feature face landmarks extracted with the help of AAMs [13][14].

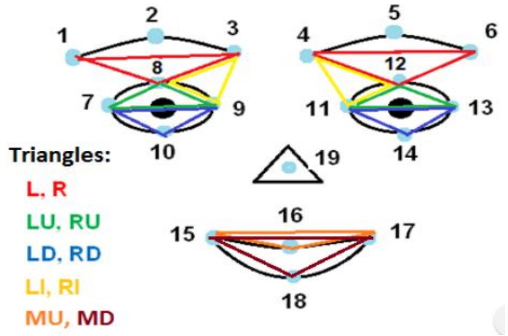
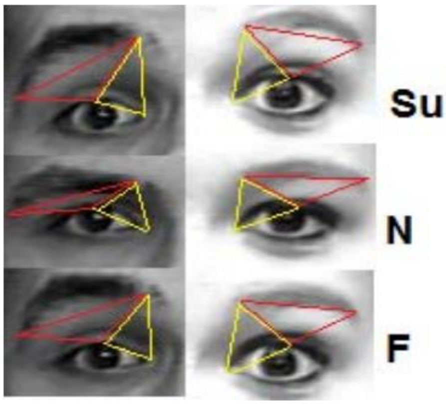
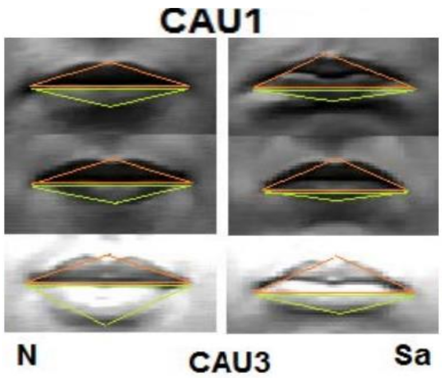


Figure 2: Triangles used to detect Aus and CAUs [13]



(a)



(b)

Figure 3: Triangles used for CAU detection under different emotional state – Surprise (Su), Fear (F), Sad (Sa) and Neutral (N) [13]

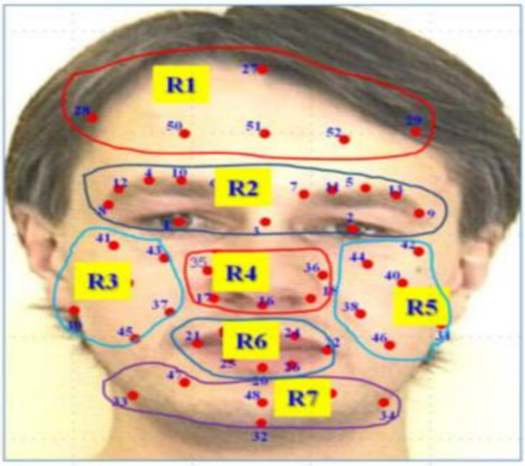


Figure 4: Grouping of facial points into various zones depending upon the facial geometry [14]

B. Recognition of Face Expressions using Appearance based Approach

The features based on the appearance patterns of an individual are used successfully for feeling acknowledgment, for example, neighborhood parallel example (LBP) administrator, nearby directional example (LDP), local Gabor binary patterns (LGBP), linear discriminant analysis (LDA), texture data achieved through Gabor filter, principle component analysis (PCA), histogram of direction inclinations (HOG) and non-negative lattice factorization (NMF) based surface element and so forth.

Some of the appearance-based features are given as follows which are based on:

- Distribution
- Eigenface

- Support Vector Machine (SVM)
- Neural Network
- Hidden Markov Model (HMM)
- Naive Bayes classifier
- Information-Theoretical approach-based features

The NMF theory from the appearance-based approaches has caused several promising works happen. In a face recognition method, a Gabor wavelets texture information extraction method is utilized to analyze the impact of partial occlusion on facial expression. For the same method, other algorithms used are - shape-based method and method of decomposing the image which is under supervision on the basis of discriminant NMF (DNMF). Another technique of face expression recognition is graph-preserving sparse NMF. Graph-preserving sparse NMF is a dimension reduction approach robust to occlusion effects. It transforms the face captured images with high dimensions into a locality-preserving subspace, by sparse denotations. Many experts also use local binary pattern (LBP) operator for the face expression analysis and recognition. One comprehensive study presents a theory about emotion recognition on the basis of LBP operator. The study uses SVM classifiers with boosted-LBP features with the goal of achieving best accuracy in face expression recognition. The LBP operator extensions i.e. – volume-LBP (VLBP) & LBP on three symmetrical planes (LBP-TOP) are utilized for outward appearance acknowledgment. Another example of a system using appearance-based features for facial expression detection is also presented in this study which is capable of automatically detecting facial expressions from videos. Various machine learning methods are investigated in this study and it is found that the method employing Gabor filters subset using AdaBoost and then trains SVM depending upon the outcome of the AdaBoost chosen filters which provide best results. [15] [16] [17]

C. Steps Involved in Facial Expression Recognition

The detection method for facial expressions for digital images, associated system and its applications are exhibited. This method of digital image analysis aids in determining whether there is a smile or blink present on a person's face or not. The classification of these digital images in face recognition or pose-determination allows a use of relative classifier depending upon the specific applications. The basic steps of face recognition are presented below:

- Image Acquisition
- Pre-processing
- Feature extraction
- Classification
- Post-processing
-

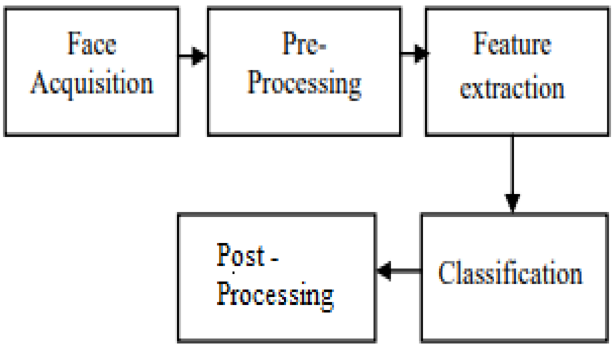


Figure 5: Steps involved in Facial Detection and Expression Classification

Image Acquisition: The sensor device such as camera is responsible for capturing the face of a subject and converting it into digital form with the help of analog to digital convertor. Face acquisition is the first step of face image processing. The key objective of this step is the automatic detection of facial portion from the provided images or videos. This image is passed through the detector which detects the face from each frame or video sequence. In order to identify the face region accurately from a scene or real-time detection, a new approach was proposed which characterizes a modified adaptation of the first 'Viola-Jones' face locator. This 'Viola-Jones' face identifier is contained prepared classifiers fell by AdaBoost. Further, integral image filters are utilized by each classifier which is similar to Haar Basis functions and can be calculated at any location and scale and that also at high speed. The Haar Basis function consists of rectangle feature, two-edge features and line features. For each cascade stage, an ‘AdaBoost’ based feature selection procedure is used for the selection of feature subset. This assists in detection of eyes, lips and nose which are customized in further research for including two more components which are cascaded objects like eye-brows and eye-balls. All the components which has been successfully detected from each frame are distinguished through bounding-box. Further the center-point of each bounding-box related to face components has been earmarked. [18]

Face Pre-Processing: The second step of face expression recognition is image pre-processing. In this stage, the gray scale conversion of captured image is carried out and even the size of the image is also reduced [19]. Compression of irrelevant data is carried out and this improves the important image-features. The motive of this step remains towards achieve unadulterated face pictures with standardized power, uniform size and shape. So as to achieve high-quality recognition performance in emotion recognition field, face pre-processing is a fundamental step which needs to be followed. It assists in removing irrelevant noise and converts all faces into a common domain. Since the experts decide to train the deep-network model on FER prior to face detection, SFEW dataset captured faces are resized to 48×48 and further transformed to gray scale in order to match it with FER data. Additionally, standard histogram equalizer remains used towards pre-process the face images obtained from FER and SFEW datasets followed by linear plane fitting which eliminates unbalanced illumination. In the final stage after plane fitting, the normalization of image pixels is performed towards a zero mean value and unit variance vector. [20]

Feature Extraction: Feature-extraction is the third stage of face expression recognition which is used to extract the required features of the image (face). Feature extraction is comprised of stages like – 1) Dimension reduction 2) feature extraction and 3) Feature selection. Out of these three, dimension reduction executes a pivotal role in recognizing the pattern of face/image. It uses SIFT to determine the key value points. Below various feature extraction techniques have been presented. [19] [20]

Discrete Cosine Transform (DCT) based Feature Extraction Technique: In reference paper [21], the input image is applied to Discrete Cosine Transform (DCT). The coefficients of DCT are arranged in zigzag manner and are implemented to execute the conversion of 2D-DCT image matrix to feature-vectors. The frequency domain components are placed at the initial points of vector. The image features of face such as eyes, lips and nose are extracted from the image fed as an input, the sub-image of which further undergoes DCT. Finally, the DCT coefficients for eyes, lips and nose are obtained and are combined with feature-vectors. ‘AdaBoost’ classifier is employed for further emotion recognition which generally achieves recognition rate of 75.94%.

Gabor Filter based Feature Extraction Technique: In reference paper [22], the faces in the images initially undergo pre-processing based on affine transform in order to normalize the images. The evaluation the pre-processed and normalized images is performed using different Gabor filters which are used to differentiate between various emotions. Dimension reduction techniques used for the recognition are Feasibility of Linear-Discriminant-Analysis (FLDA) and Principle-Component-Analysis (PCA) metaclassifier. The discriminate function is assumed as linear function of the feature data by the classifier. The obtained feature vector is the data in this scenario. As per the author, the proposed Gabor filter selection helps in reducing the computation complexity as well as dimension of feature-space. The database utilized here for evaluating the detection system is JAFFE database which achieves recognition rate of 90.22% and above.

Principle Component Analysis Based Feature Extraction Technique: In reference paper [23], WPCA (Weighted PCA) method is used for fusion of multiple features. These local self-related features having high dimensionality are extracted from the face images and further are divided into different zones. In the next step, Weighted PCA is applied for reducing the dimensionality of features. The weights are recognized on a faster basis with the help of facial action coding system in which RBF (Radial Basis Function) is also used. The facial emotions and face characteristics are further classified using Support Vector Machine algorithm. At last, in order to achieve the similarity among templates, the Euclidean distance is estimated and then facial expression recognition is performed using nearest algorithm. The presented technique attains a recognition rate of 88.25%.

Independent Component Analysis Based Feature Extraction Technique: In reference paper [19], it is shown that the ICA (Independent component Analysis) components are gotten from the rule of ideal data move through sinusoidal neurons. ICA remains applied to facial images achieved from FERET database depending upon the two architecture models. The first model assumes image as a random variable and pixel as a random trial. Hence ICA considers images to be statistically independent images. The space-resembling facial characteristics of these images are sparse and localized. The second model assumes pixel as a random variable and image as a random trial. The image coefficients are independent and hence yield factorial face code outcome.

Classification: This is the last or second-last step in the processing of face expression recognition, depending upon the type of application and requirement of the system. The classification is performed on the features obtained from feature-extraction outcome. Here, various classifiers have been proposed for accurate facial recognition depending upon optimized features, application and the system training. [19] [20] [24]
Hidden Markov Model (HMM) as classifier: In reference paper [25], HMM is proposed for the classification of the higher-level expressions which are critical to detect with normal classifier. These higher-level expressions include discouraging, encouraging, disagreeing, interested and unsure whereas, lower level expressions include joy, surprise, sad and neutral expressions. An emotional indexing model is employed to understand the emotional state hence it functions according to the database.

Neural Networks as classifier: In reference paper [26], there is a blend of two strategies Feature extraction and neural organization and two phases are included for face recognition and characterization. Pictures are pre-handled in order to diminish the time and to raise the nature of picture. Neural organization-based classifier is utilized to group the component vector.

Support Vector Machine as classifier: In this reference paper [27], first scanning of each video-sequence frame is performed in real-time for the recognition of frontal face then scaling of these faces is done to convert it into equally sized image-patches with a bank of Gabor-energy-filter [45]. The outcome of this filter is fed as an input to recognition classifier which performs coding of expressions into different dimensions. The features of the face are chosen from Gabor-filters with the help of AdaBoost which are further trained using SVM (Support vector machine). The author here has created end-to-end system which facilitates various codes for emotions at the rate of 24 frames/sec and animates the features generated by computer. Then, fully automated facial action coding is also applied. The proposed SVM classifier achieves a recognition rate of 93%.

AdaBoost as classifier: In this reference paper [28], the frontal facial images are classified depending upon the level of expressions which are critical to detect with normal classifier. The higher-level expressions include discouraging, encouraging, disagreeing, interested and unsure whereas, lower level expressions include joy, surprise, sad and neutral expressions. The recognition system doesn't contain any characteristic blocks. In this article, a method is used to recognize the colour of the face skin that is RGB, HSI and YCbCr which is applied to the facial images. Further, the region having skin-tone colours is separated from region having non-skin colour with the help of lower and upper bound threshold. Thus, the values are considered between 3 and 38 as the skin-tone colours range. Further, the features of face are detected using Color space transformation, Connected Component Labeling Technique, Pupils detection Face Region Verification like face segmentation, like referring height and width of the face region etc. This aids in detecting the facial landmarks like – eyebrows, eyes, lips and nose. Finally, the facial features of these are extracted and the features are separated by the displacement of face landmark features using proposed AdaBoost classification algorithm which provides accuracy of 90%.

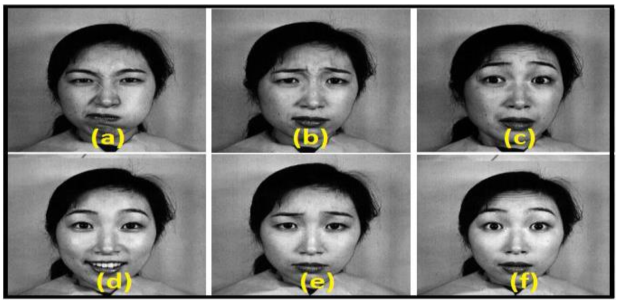


Figure 6: Image samples obtained from database with various types of emotions such as joy, surprise, sad and neutral expressions etc [20]

Post-processing: Post-processing is the last step of face expression recognition. The captured image is normally made to undergo post-processing stage after the successful classification of image. In this post-processing step, linear bank of filters are used for encoding the shift and illumination invariance. The key objective of post-processing is to enhance the recognition-accuracy by taking benefit of domain-knowledge to rectify the classification errors. The classification errors can also be corrected by coupling together several levels of a classification hierarchy. Using post-processing approach for images to gain invariance proves to be advantageous as it increases the chances of higher recognition accuracy. One of the surveys shows that the post-processing of tracked image is beneficial especially in reliable light conditions. In post-processing stage, the smoothness parameter is also set to default with no normalization. [29]

The rough outcome of the deep-network mainly represents the scores of confidences to determine whether the action unit is present or absent. Here, an extreme low value shows that the AU is absent whereas an extreme high value shows that the AU is present. Ideally, the decision threshold of the deep-network should depend upon the values located at middle point which represents whether an AU is present or absent as a ground-truth in the training-set. However, outcome of several conditions can skew the threshold to either of the extreme value. In order to mitigate this issue, the decision-threshold of AUs is optimized on a validation-set. This means the decision threshold is set at a worth that provides high F1 score when tested on validation-set. In the mentioned scenario, the dataset is separated into two categories – 1) Training set and 2) Validation set. In the next step, the training set is used to train the network and the decision threshold is tested on development validating set for all the best promising values of decision threshold. At last, the threshold having best performance is applied in decision making for the final network, which is further trained on the entire set of data. [30]

In this reference paper [31], a software tool i.e. Computer Expression Recognition Toolbox (CERT) has been proposed for automatically detecting the real-time face expressions. This software is also available for academic use for free of cost. CERT is capable of coding intensity of 19 distinct facial emotions obtained from FACS and 6 distinct prototypical face emotions. It also determines the 3-dimensional orientation of head (i.e. yaw, pitch and roll) and detects the location of 10 facial feature landmarks. When images containing posed facial expressions from the database are considered for experiment, it is found that CERT attains an average-recognition accuracy of 90.1% for facial expression analysis. But if a spontaneous dataset is considered for facial expression analysis, CERT attains a recognition accuracy of approximately 80%. CERT is capable of processing video images of size 320 x 240 in real-time at an average rate of around 10 frames/second in a standard dual-core laptop.

As we have just shown, all the methods relating to each approach have advantages as well as disadvantages. Our objective is to take advantage of the existing advantages and to propose new solutions so as to avoid the underlying disadvantages. Most of the techniques require manual intervention at the start of treatment.

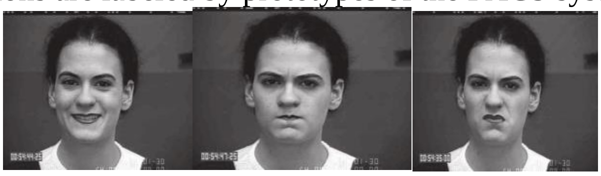
They propose a classification of an expression studied in a single universal category as postulated by Ekman: Joy, Surprise, Anger, Disgust, Sadness and Fear, except this is not real because the human being is not binary. To move from one expression to another the face passes through mixed transient expressions, thus the expression is classified into a simple category of un-realistic emotions and, hence it is desired that the classification system should be capable of identifying such an transitional mixed expressions. Finally most of these approaches are based mainly on the deformations of the permanent characteristics of the face (Eyes, Eyebrows and Mouth) because they consider that these traits carry enough useful information for the classification of facial expressions. Very few methods study the presence of transient features (wrinkles) on certain areas of the face in a post-treatment, in order to discriminate between two emotions.
Experimental setup and data sets description: [50]

3. FACIAL EXPRESSION DATASET

The evaluation of face expression detection and emotional recognition methods requires the use of one or more databases. In order to obtain the most objective evaluation possible, the databases must be general and independent of restrictions or assumptions related to a specific area. Thus, the performances of the methods can be referenced and compared more easily.
There are two basic types of knowledge: simulated emotions and spontaneous emotions. In this chapter we present the most used databases in the literature. Then, we study the constraints frequently encountered in the applications of recognition of emotions and evaluate their taking into account in the bases.

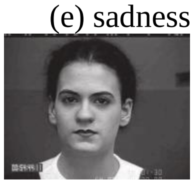
A. Cohn-Kanade(CK+) Facial Expression Dataset

The Cohn-Kanade Facial Expression Database is intended for automatic image analysis and synthesis, as well as perception studies such as facial expression recognition [33]. The first version of this database consists of expressions simulated by actors. Experts have described and visually shown the unitary action or the combination of unitary actions to achieve, the subjects reproduce them in front of the camera. Part of the database is open-access and contains 97 subjects, aged 18 to 30 years. These subjects are made up of 65% of women, 15% of African-Americans and 3% of Asians and Latinos [34]. The image sequences presented begin with the neutral expression and end with the maximum of the expression. In addition, the last image of the sequence is always coded by experts. Of these images, 17% are coordinated by two experts. The agreement of judgment of the two is calculated by the coefficient Kappa 2. This base includes expressions involving one or more muscular actions, among them the six universal emotions (joy, anger, fear, surprise, sadness, disgust). These emotions are labeled by prototypes of the FACS system [35].



(a) joy (b) anger (c) disgust

(d) fear (e) sadness (f) surprise



(g) neutral

Figure 7: Example extracted from the CK+ database

ISSN 2515-8260 Volume 7, Issue 4, 2020

B. MMIM&M Initiative (MMI) Dataset

The MMI base is a base of facial expressions, continually expanding. It consists of five parts: the first consists of 1767 image sequences of 20 participants from different origins (American, Asian, and European). The subjects perform 31 unit actions and aff states such as joy and boredom. These actions are repeated twice to increase the variability of the base. This first part has frontal views and profile views obtained by using a mirror positioned at a 45-degree angle to the front view of the subject and still respecting the position of the camera [36] [37].

Figure 8: Example image of the MMI base where the background is complex and natural light [37]

C. JAFFE

The Japanese Female Facial Expression Database (JAFFE) has 219 static images. These images concern ten Japanese women who simulate the six universal emotions and the neutral expression. The labeling of the base is made by 60 other Japanese women who quantify the emotion presentin each image on a 5 point scale for the six emotions [38]. The figure 9 presents an example taken from this database.

(a) joy (b) anger (c) disgust

(d) fear (e) sadness (f) surprise

(g) neutral

Figure 9: Example extracted from the JAFFE database [39]

D. MUG

MUG stands for "Multimedia Understanding Group" dataset image sequences achieved from MUG dataset [40]. It initiates and terminates in neutral state and tracks the apex, offset and onset temporal patterns. Further, few image sequences are recorded having different length for the six primary expressions. There are 50 to 160 images in each image sequence. A short training is provided to the subjects before the recording through a tutorial. An example of face expression sequence captured from MUG dataset is shown in Figure 10.

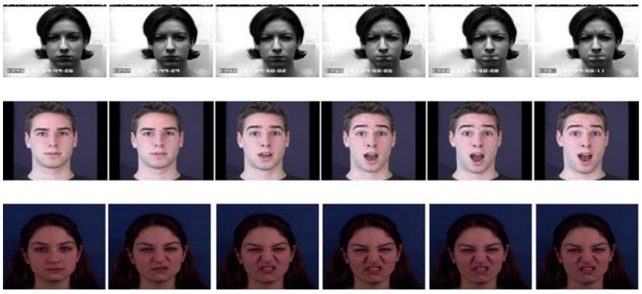


Figure 10: Example of face expression sequence obtained from various three datasets [50]

E. MHI Mimicry

It consists of 54 videos (sessions) of 40 subjects (28 men, 12 women) between 18 and 40 years old from Imperial College London. On each session, 2 participants interact and each session is divided into two parts. First a part debate in which the two participants will discuss politics. They will be either of the same opinion or of an opposite opinion.

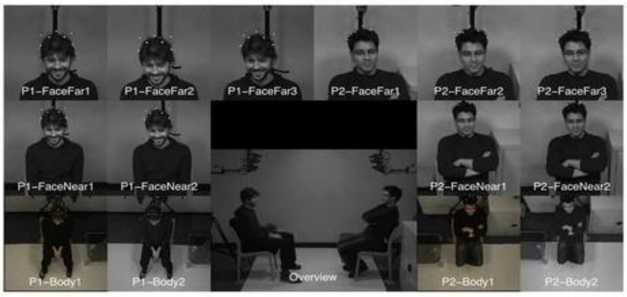


Figure 11: Example of MHI Mimicry Data [40]

The FEEDTUM database [41] is produced in association with European project FG-NET. She presents the six primary emotions (joy, anger, disgust, fear, sadness and surprise) and the neutral expression. The recording system consists of 18 computers with cameras (see Figure 12). It allows both a projection of videos on screens and a recording of stimulated emotions.

This base has 18 topics. Each subject visualizes three videos for each emotion. The reactions of the subjects are then recorded and labeled according to the stimulating video. For example, for a video intended to stimulate joy, the facial expressions of the subject are recorded and labeled as expressions of joy. These expressions are considered spontaneous.

(a) joy (b) anger (c) disgust

(d) fear (e) sadness (f) surprise

(g) neutral

Figure 12: Example from the FEEDTUM database

ISSN 2515-8260 Volume 7, Issue 4, 2020

Table 1: Comparison of 2D Databases (P / S: Pos / Spontaneous)

Nom	P/S	Size	contents	Label
CK[42]	P	97 subjects (65% women, African-Americans and 3% or South American) from 18	486 video sequences of the 6 Siques	FACS
CK+[43]	P/S	Number of subjects increased 27%	Number of sequences	FACS + emotion
MMI[44]	P/S	75 subjects (48% women, European, African and South from 19 to 62 years	2900 videos of the 6 basic of specific AUs	FACS
MHI Mimi-cry[40]	S	40 subjects (28 men, 12 women) between the ages of	54 videos (sessions)	Auto-evaluation
HCI Tagging [44]	S	27 subjects (11 men, 16 women) from 19 to 40 years	Part 1: 20 videos; Part 2: 28 and 14 videos	Auto-evaluation
SAL [46]	S	24 subjects	10 hours of video: the subjects an artificial intelligence and tions are changed according to different AI personalities	Feel trace
JAFFE[38]	P	10 Japanese women	213 images of the 7 basic emotions	Noted according to emotions by 60 Japanese

F. Other Bases of Facial Expressions used In Works of Literature

In Table 2 we present a set of bases used to evaluate emotion recognition systems. They are mostly simulated.

Table 2: Bases of face expressions used in the literature

s	Types of exprsions emotional	Base details
Hammal-Caplier[47]	simulated	• 21 subjects. • Joy, disgust, surprise. • Sequences initiates as well as terminates with the neutral state
Dailey-Cottrell[48]	simulated	• 16 subjects. • Joy, anger, fear, disgust, sadness and surprise. • For each emotion two images exist the expression and the neutral
		• 68 subjects. • Neutral expression, smile.
SALD[50]	spontaneous	• 20 subjects. •Labeling according to the dimensions Valencia, Arousal.
Yale[51]	simulated	• 15 subjects. • Several poses of head. • Neutral expression, sadness, surprise, drowsiness.
DaFEx	simulated	• 8 subjects. joy, surprise, sad, anger, fear, disgust and neutral
HUT[52]	simulated	• 2 subjects. • joy, surprise, sad, anger, fear, and disgust

G. Database Criteria

A generic database for the comparison of face and emotion recognition approaches needs to check certain criteria in order to get closer to the actual constraints of the applications. It must also provide reliable ground truths to enable the evaluation of the robustness of emotion recognition applications. We detail these criteria in the following points:

Choice of subjects

The choice of subjects in a database can have an important influence on the results of the recognition. Indeed, the expressions are different at the age of the participants. A baby does not express emotion in the same way as a grown person. It has a very smooth face, which does not contain texture, unlike the face of a senior with wrinkles. Thus, the level of texture present in a face varies according to age. In addition, face morphologies change from person to person, and origins also occur in the shape and appearance of faces. Asians, for example, have a specific eye shape. The experiments of P. Ekman [80] also prove that culture intervenes in the expression of emotions. Indeed, Ekman shows through an experiment conducted on Americans and Japanese that they are more reserved when expressing their emotions in front of foreign people. For these reasons, the databases must contain people of various ages and backgrounds. In all the bases that we have described previously there are no very young subjects (babies) and very old subjects (over 70 years old). However, most bases choose subjects of various origins.

Environment: Many of the constraints of the subject's scene affect the recognition of facial expressions. Here we mention the background, the ambient light and the cameras used. However, the approaches that make their learning on this type of image are not robust to sequences of images of real scenes where the backgrounds are non-uniform and dynamic. Despite the integration of complex backgrounds for some sequences of the MMI database (see Figure 13), the presence of dynamic backgrounds, due to camera movements or subject movements, remains an important criterion which is slightly included in the bases of recognition of emotions.

Ambient light variation is also a criterion that impacts the recognition results. In most bases, light is controlled inside laboratories. However, outdoor sequences with natural light are also essential for some applications. The variation of light in a real scene can project a shadow on the face, thus distorting the characteristics of it which will cause several errors.

The choice of cameras also has several impacts on the analysis of expressions. In fact, camera properties such as the number of frames per second and the resolution of images are very important in the construction of a base since they directly influence the learning outcomes of the methods. A database must include, in addition to high resolution images, images of low resolutions, the extraction of information achievable on high resolution images may become impossible for low resolution images. Thus, the methods making their learning on bases containing the two types of resolutions are more robust in real conditions. The majority of existing databases only include high resolution images. Among the few with low resolution images, there is PETS 2003 [81] with resolutions of approximately 69 × 93 pixels. Sampling videos canalso have an impact on performance. According to Kanade et al [82], the methods using the optical flows always assume that the movement from one image to another is very weak. Thus, the choice of the sampling of the video must be precise. The MMI base uses for example a sampling of 24 images per second.

Pose of the head: In real scenes, expressions are often stimulated by an event or by another person. Thus, in many cases emotions are expressed in an inclined view, see a profile view of the face expressions are often accompanied by a movement of the head. In these cases, the recognition of emotions must incorporate images with different inclinations of the head and different angles with respect to the front view of the head. A database must contain images of different views of the face. Some efforts have been made in this area, such as the MMI database, which has integrated images of frontal views of the face and side views. Databases incorporating this kind of information, however, are few in number and the majority of image databases only show the front or near frontal view of the face.

Spontaneous expression and posed: Placed and spontaneous expressions are generated by various areas of the brain. Thus the activation of the muscles and the dynamics of the expressions differ according to the type of emotions. Spontaneous expressions are more synchronized, more symmetrical and more coherent than forced expressions. According to [83], spontaneous expressions may correspond to innate expressions, whereas forced expressions correspond to socially learned behavior. A database must contain both types of expressions allowing a better recognition of emotions by automatic recognition approaches. However, there are several difficulties when it comes to collecting spontaneous emotions. Indeed, stimulating emotions such as anger, sadness and fear is very difficult in a lab environment. Bases integrating spontaneous emotions have adopted different protocols to stimulate them. Sebe et al [84] thus present videos to stimulate the emotions of subjects who remain alone in a room containing hidden cameras in order to give subjects the freedom to express themselves and preserve the authenticity of the expression. Sebe et al [84] only detected joy, disgust and surprise. Using a video-viewing technique similar to that used in [84], the six emotions are stimulated and recorded in the FEEDTUM database. Some emotions are barely visible and remain very close to the neutral expression. Another technique is used in the CK+ database to collect spontaneous emotions. The spontaneous smiles of participants are recorded before and after their participation in forced simulation experiments.

Labeling: The labeling of emotions is a task of major interest in the construction of databases. Two labeling approaches are possible [30]. The first uses a set of predefined emotion labels, usually universal emotions and judgment labeling, which is based on the opinion of several expert or non-expert participants. This technique is used in the FEEDTUM database. In the database presented by Sebe et al [84], subjects are asked about the emotions they experienced. The second approach is based on pre-defined prototypes for each emotion. These prototypes consist of a set of muscle movements. Several databases have used this approach, among which we mention the MMI database, which used the unit actions of the FACS system. The CK + database is also labeled by the FACS system. Two experts carry out the labeling independently. Then, the agreement rate between the two is calculated for a percentage of the data.

Occultation: Occultation is a very common constraint in emotion recognition applications. However, the bases presenting this kind of problem are very rare. Certain works dealing with the recognition of face expressions in the cases of occultations encounter difficulties in the evaluation of their methods. Some ideas have been put in place such as building bases containing such situations from existing databases. In Kotsia et al [85], the bases JAFFE and Cohn-Kanade were thus used to build bases taking into account this constraint.

4. EXPERIMENTAL RESULTS ACHIEVED IN DIFFERENT DATABASE

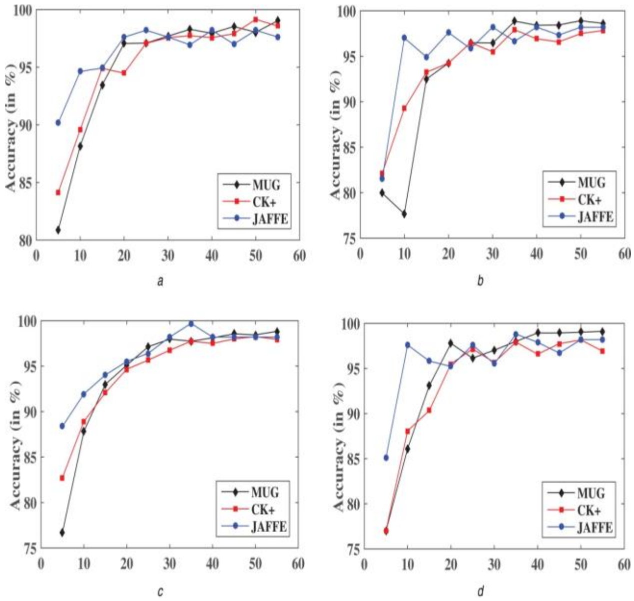


Figure 13: Suggested method and its performance for different image resolutions

Table 3: Comparison of FER performance using CK+ dataset using various methods

Reference	Method	Image data	Class	Accuracy (%)
[53]	HOG feature, ELM Ensemble	Frame	7	97.30
[54]	Texture and geometric features	Frame	6	92.20
[55]	LBP-TOP+VLBP	Sequence	6	96.26
[56]	LBP +KDIsomap	Frame	7	94.88
[57]	Gabor +Addaboost	Frame	7	94.50
[58]	HOG feature, ELM Ensemble	Frame	7	97.30
[59]	Stepwise LDA + hidden conditional random field	Frame	6	96.83
[60]	Geometric key displacement features	Sequence	6	99.70
[61]	Geometric features + dynamic Bayesian network	Frame	7	94.04
[62]	Local representation, LBP+NCM features	Frame	6	97.25

Table 4: Accuracy performance comparison in CK+ database with convolution neural network-based classifier in the literature

Method	Classifier	6-expression	7-expression
[63]	Cn classs	83.70 %	-
[64]	Cn class	-	88.70 %
[65]	Cb class	93.70%	-
[66]	Cn class	94.90 %	
[67]	Cn class	98.92%	98.80%
[68]	Cn class	-	95.8%

Table 5: Comparison the recognition rate of different methods on JAFFE

Feature extraction	**Dimensionality reduction**	**Classifier**	**Recognition rate (%)**
Gabor filter [69]	Adaboost	SVM	71.9
Gabor filter [69]	N/A	SVM	91.9
Gabor filter [70]	PCA	NN	90.6
LBP [71]	PCA	SVM	53.8
LBP[72]	LFDA	SVM	85.7
LBP [73]	GDA	SVM	83.3
LBP [71]	Adaboost	SVM	65.7
Boost LBP[71]	LPP	KNN	74.2
Boost LBP[71]	SONPP	KNN	66.8
ELBP [74]	KLT	SVM	93.3

Table 6: Performance of the state-of-the-art methods

Existing works	**Feature selection**	**Classifier**	**database**	**Accuracy**
[75]	local fisher discriminant	LFDA	MUG JAFFE	95.24% 94.32%
[76]	LBP LBP LBP-BOOST	SVM SVM SVM	CK+ JAFFE CK+	94.60 % 79.80% 95.00%
[77]	Gabor wavelet	KNN	JAFFE	92.37%
[78]	LBP	MTSL+SVM	CK+	97.70%
[79]	WPLBP	SVM	MUG JAFFE CK+	98.44% 98.51% 97.50%

5. CONCLUSION

We have presented the most famous theories and representations of emotions. Characteristic extraction approaches, classified into three categories (geometric, appearance and hybrid), as well as methods using descriptor selection were introduced. Finally, the databases frequently used have been described, as well as the constraints present in the applications of recognition of emotions. At the end of this study, we chose two bases for the rest of our work. Given the importance of emotions posed in social behavior and the importance of spontaneous emotions in everyday life, we have selected a basis for each type of emotions. For the emotions posed, the base CK +, which includes participants of different origins and different ages, was chosen. Concerning spontaneous emotions.

Automatic recognition of facial expressions based on transitory features is a new research focus. The results obtained proved that this approach can be an accepted basis for the recognition of facial expressions. Although doubt often exists between two expressions, we believe that it is better to preserve doubt than to take the risk of misclassification. It is for this reason that we have preferred to use the Theory of Evidence because the latter perfectly models the doubt and it is a very powerful fusion tool when it comes to data from several sources.

The comparison presented perfectly indicates that recognition based transient traits achieves results as well as recognition based permanent traits. The combination of the two approaches could yield optimal results in the future.

6. REFERENCES

[1] Galbally, J., Marcel, S. and Fierrez, J., 2014. Image quality assessment for fake biometric detection: Application to iris, fingerprint, and face recognition. IEEE transactions on image processing, 23(2), pp.710-724.
[2] Mandal, B., Lim, R.Y., Dai, P., Sayed, M.R., Li, L. and Lim, J.H., 2016. Trends in machine and human face recognition. In Advances in Face Detection and Facial Image Analysis (pp. 145-187). Springer, Cham.
[3] Visconti, M.D.O.C., Halchenko, Y.O., Guntupalli, J.S., Gors, J.D. and Gobbini, M.I., 2017. The neural representation of personally familiar and unfamiliar faces in the distributed system for face perception. Scientific Reports, 7(1), pp.12237-12237.
[4] Bettadapura, Vinay. “Face expression recognition and analysis: the state of the art.” arXiv preprint arXiv:1203.6722(2012).
[5] Savran, A., Sankur, B. and Bilge, M.T., 2012. Regression-based intensity estimation of facial action units. Image and Vision Computing, 30(10), pp.774-784.
[6] Rivera, Adin Ramirez, Jorge Rojas Castillo, and OksamOksamChae. “Local directional number pattern for face analysis: Face and expression recognition.” IEEE transactions on image processing 22.5 (2013): 1740-1752.
[7] Mollahosseini, A., Hasani, B. and Mahoor, M.H., 2017. Affectnet: A database for facial expression, valence, and arousal computing in the wild. arXiv preprint arXiv:1708.03985.
[8] Dhall, Abhinav, et al. “Emotion recognition in the wild challenge 2014: Baseline, data and protocol.” Proceedings of the 16th International Conference on Multimodal Interaction. ACM, 2014.
[9] Li, Y., Wang, S., Zhao, Y. and Ji, Q., 2013. Simultaneous facial feature tracking and facial expression recognition. IEEE Transactions on Image Processing, 22(7), pp.2559-2573.
[10] Eckert, Martina, et al. “Fast facial expression recognition for emotion awareness disposal.” Consumer Electronics-Berlin (ICCE-Berlin), 2016 IEEE 6th International Conference on. IEEE, 2016.
[11] Liu, Ping, et al. “Facial expression recognition via a boosted deep belief network”, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2014.
[12] Sandbach, G.; Zafeiriou, S.; Pantic, M.; Rueckert, D. Recognition of 3D facial expressions dynamics. Image Vision Comput 2012, 762–773.
[13] Rudovic, O.; Pantic, M.; Patras, I. Coupled Gaussian processes for pose-invariant facial expression recognition. IEEE Trans. Pattern Anal. Mach. Intell.2012, 35, 1357–1369.
[14] Ghimire, Deepak, and Joonwhoan Lee. “Geometric feature-based facial expression recognition in image sequences using multi-class adaboost and support vector machines.” Sensors13.6 (2013): 7714-7734.
[15] Zhao, X.; Zhang, S. Facial expression recognition based on local binary patterns and kernel discriminant isomap. Sensors 2011, 11, 9573–9588.
[16] Moore, S.; Bowden, R. Local binary patterns for multi-view facial expression recognition. Comput. Vision Image Underst. 2011, 115, 541–558.
[17] Zhao, G.; Huang, X.; Taini, M.; Li, S.Z.; Pietikäinen, M. Facial expression recognition from near-infrared videos. Image Vision Comput. 2011
[18] Thorat, Bali, Ganesh Manza, and PravinYannawar. “Automatic Classification of Facial Expressions from Video Stream using Decision Tree.” International Journal of Computer Applications 121.22 (2015).
[19] A Study of Techniques for Facial Detection and Expression Classification
[20] A Hybrid Approach of Facial Emotion Detection using Genetic Algorithm along with Artificial Neural Network
[21] Owusu, E., Zhan, Y. and Mao, Q.R., 2014. A neural-AdaBoost based facial expression recognition

system. Expert Systems with Applications, 41(7), pp.3383-3390.

[22] Bashyal, S. and Venayagamoorthy, G.K., 2008. Recognition of facial expressions using Gabor wavelets and learning vector quantization. Engineering Applications of Artificial Intelligence, 21(7), pp.1056-1064.

[23] Hegde, G.P., Seetha, M. and Hegde, N., 2016. Kernel locality preserving symmetrical weighted Fisher discriminant analysis-based subspace approach for expression recognition. Engineering science and technology, an international journal, 19(3), pp.1321-1333.

[24] Hemalatha, G. and Sumathi, C.P., 2014. A study of techniques for facial detection and expression classification. International Journal of Computer Science and Engineering Survey, 5(2), p.27.

[25] Meng, H. and Bianchi-Berthouze, N., 2011. Naturalistic affective expression classification by a multi-stage approach based on hidden markov models. In Affective computing and intelligent interaction (pp. 378-387). Springer, Berlin, Heidelberg.

[26] Saudagare, P.V. and Chaudhari, D.S., 2012. Facial expression recognition using neural network–an overview. International Journal of Soft Computing and Engineering (IJSCE), 2(1), pp.224-227.

[27] Luo, Y., Wu, C.M. and Zhang, Y., 2013. Facial expression recognition based on fusion feature of PCA and LBP with SVM. Optik-International Journal for Light and Electron Optics, 124(17), pp.2767-2770.

[28] Owusu, E., Zhan, Y. and Mao, Q.R., 2014. A neural-AdaBoost based facial expression recognition system. Expert Systems with Applications, 41(7), pp.3383-3390.

[29] Beaudry, O., Roy-Charland, A., Perron, M., Cormier, I. and Tapp, R., 2014. Featural processing in recognition of emotional facial expressions. Cognition & emotion, 28(3), pp.416-432.

[30] Yao, A., Shao, J., Ma, N. and Chen, Y., 2015, November. Capturing au-aware facial features and their latent relations for emotion recognition in the wild. In Proceedings of the 2015 ACM on International Conference on Multimodal Interaction(pp. 451-458). ACM.

[31] Littlewort, G., Whitehill, J., Wu, T., Fasel, I., Frank, M., Movellan, J. and Bartlett, M., 2011, March. The computer expression recognition toolbox (CERT). In Automatic Face & Gesture Recognition and Workshops (FG 2011), 2011 IEEE International Conference on (pp. 298-305). IEEE.

[32] Happy, S.L. and Routray, A., 2015. Automatic facial expression recognition using features of salient facial patches. IEEE transactions on Affective Computing, 6(1), pp.1-12.

[33] Cohn-kanadeau-codedexpressiondatabasehttp://www.consortium.ri.cmu.edu/ckagree/

[34] Lucey, P., Cohn, J.F., Kanade, T., Saragih, J., Ambadar, Z. and Matthews, I., 2010, June. The extended cohn-kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression. In Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on (pp. 94-101). IEEE.

[35] Z.Ambadar,CohnJ.,andReedL.Allsmilesarenotcreatedequal:Mor-phologyandtiming ofsmilesperceivedasamused,polite,andembarras-sed/nervous.JournalofNonverbalBehavior,pages17–34,2009.

[36] Valstar, M. and Pantic, M., 2010, May. Induced disgust, happiness and surprise: an addition to the mmi facial expression database. In Proc. 3rd Intern. Workshop on EMOTION (satellite of LREC): Corpora for Research on Emotion and Affect (p. 65).

[37] Valstar, M.F., Pantic, M., Ambadar, Z. and Cohn, J.F., 2006, November. Spontaneous vs. posed facial behavior: automatic analysis of brow actions. In Proceedings of the 8th international conference on Multimodal interfaces (pp. 162-170). ACM.

[38] R.Gross.Face Databases, Handbookoffacerecognition.Springer-Verlag,2005.

[39] M.Lyons,M. Kamachi, and J.Gyoba. The Japanese female facial expression (jaffdatabase).

[40] Ghimire, D., Lee, J., Li, Z.N. and Jeong, S., 2017. Recognition of facial expressions based on salient geometric features and support vector machines. Multimedia Tools and Applications, 76(6), pp.7921-7946.LICHTENAUER, J., VALSTAR, M., Sun, X. and MAJA, A.N., 2012. Mahnob HMI IBug Mimicry Database (MHI-MIMICRY).

[41] Frank Wallhoff, “Facial expressions and emotion database, http://www.mmk.ei.tum.de/waf/fgnet/feedtum.html,2006.

[42] TakeoKanade,JeffreyF.Cohnet Yingli Tian. «Comprehensive Database for Facial Expression

Analysis». In: Proceeding sof the Fourth IEEE International Conference on Automatic Face and Gesture Recognition (2000).

[43] Lucey, Patrick, Jeffrey F. Cohn, Takeo Kanade, Jason Saragih, Zara Ambadar, and Iain Matthews. "The extended cohn-kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression." In 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition-Workshops, pp. 94-101. IEEE, 2010.

[44] Valstar, Michel, and Maja Pantic. "Induced disgust, happiness and surprise: an addition to the mmi facial expression database." In Proc. 3rd Intern. Workshop on EMOTION (satellite of LREC): Corpora for Research on Emotion and Affect, p. 65. 2010.

[45] Ruan, J., Yin, J., Chen, Q. and Chen, G., 2014. Facial expression recognition based on gabor wavelet transform and relevance vector machine. JOURNAL OF INFORMATION &COMPUTATIONAL SCIENCE, 11(1), pp.295-302.

[46] SiteofficieldeSAL.url:http://emotion research.net/toolbox/%20toolboxdatabase.2006-09-26.5667892524

[47] Hammalcaplier database, http ://facialexpressions.free.fr/,2003.

[48] DaileyM., CottrellG.W., andReillyJ. Californiafacialexpressions(cafe), unpublished digitalimages,UniversityofCalifornia, 2001.

[49] Piedatabase,http://www.ri.cmu.edu/research_project_detail.html.

[50] Douglas-Cowie, Ellen, Roddy Cowie, Ian Sneddon, Cate Cox, Orla Lowry, Margaret Mcrorie, Jean-Claude Martin et al. "The HUMAINE database: Addressing the collection and annotation of naturalistic and induced emotional data." In International conference on affective computing and intelligent interaction, pp. 488-500. Springer, Berlin, Heidelberg, 2007.

[51] YaleUniversity.Theyalefacedatabaseb, http://vision.ucsd.edu/isk-wak/extyaledatabase/extyaleb.html,2001.

[52] J.Katsyri,V. Klucharev, M. Frydrych, and M.Sams. Identification of synthetic and natural emotional facial expressions. InAudio-Visual Speech Pro-cessing Converence AVSP, 2003.

[53] Ghimire, D. and Lee, J., 2014. Extreme Learning Machine Ensemble Using Bagging for Facial Expression Recognition. JIPS, 10(3), pp.443-458.

[54] Poursaberi, A., Noubari, H.A., Gavrilova, M. and Yanushkevich, S.N., 2012. Gauss–Laguerre wavelet textural feature fusion with geometrical information for facial expression identification. EURASIP Journal on Image and Video Processing, 2012(1), p.17.

[55] Zhao, G. and Pietikainen, M., 2007. Dynamic texture recognition using local binary patterns with an application to facial expressions. IEEE transactions on pattern analysis and machine intelligence, 29(6), pp.915-928.

[56] Zhao, X. and Zhang, S., 2011. Facial expression recognition based on local binary patterns and kernel discriminant isomap. Sensors, 11(10), pp.9573-9588.

[57] Li, Y., Wang, S., Zhao, Y. and Ji, Q., 2013. Simultaneous facial feature tracking and facial expression recognition. IEEE Transactions on Image Processing, 22(7), pp.2559-2573.

[58] Ghimire, D. and Lee, J., 2014. Extreme Learning Machine Ensemble Using Bagging for Facial Expression Recognition. JIPS, 10(3), pp.443-458.

[59] Siddiqi, M.H., Ali, R., Khan, A.M., Park, Y.T. and Lee, S., 2015. Human facial expression recognition using stepwise linear discriminant analysis and hidden conditional random fields. IEEE Transactions on Image Processing, 24(4), pp.1386-1398.

[60] Saeed, A., Al-Hamadi, A., Niese, R. and Elzobi, M., 2014. Frame-based facial expression recognition using geometrical features. Advances in Human-Computer Interaction, 2014, p.4.

[61] Li, Y., Wang, S., Zhao, Y. and Ji, Q., 2013. Simultaneous facial feature tracking and facial expression recognition. IEEE Transactions on Image Processing, 22(7), pp.2559-2573.

[62] Ghimire, D., Jeong, S., Lee, J. and Park, S.H., 2017. Facial expression recognition based on local region specific features and support vector machines. Multimedia Tools and Applications, 76(6), pp.7803-7821.

[63] Fan, X. and Tjahjadi, T., 2015. A spatial-temporal framework based on histogram of gradients and optical flow for facial expression recognition in video sequences. Pattern Recognition, 48(11),

pp.3407-3416.
[64] Zavaschi, T.H., BrittoJr, A.S., Oliveira, L.E. and Koerich, A.L., 2013. Fusion of feature sets and classifiers for facial expression recognition. Expert Systems with Applications, 40(2), pp.646-655.
[65] Liu, M., Li, S., Shan, S. and Chen, X., 2015. Au-inspired deep networks for facial expression feature learning. Neurocomputing, 159, pp.126-136.
[66] Lee, S.H., Baddar, W.J. and Ro, Y.M., 2016. Collaborative expression representation using peak expression and intra class variation face images for practical subject-independent emotion recognition in videos. Pattern Recognition, 54, pp.52-67.
[67] Lopes, A.T., de Aguiar, E., De Souza, A.F. and Oliveira-Santos, T., 2017. Facial expression recognition with convolutional neural networks: coping with few data and the training sample order. Pattern Recognition, 61, pp.610-628.
[68] Agarwal, S., Santra, B. and Mukherjee, D.P., 2018. Anubhav: recognizing emotions through facial expression. The Visual Computer, 34(2), pp.177-191.
[69] Chang, C.C., 2011. " LIBSVM: a library for support vector machines," ACM Transactions on Intelligent Systems and Technology, 2: 27: 1--27: 27, 2011. http://www. csie. ntu. edu. tw/~ cjlin/libsvm, 2.
[70] Zhang, S., Zhao, X. and Lei, B., 2011, August. Facial expression recognition using local fisher discriminant analysis. In International Conference on Computer Science, Environment, Ecoinformatics, and Education (pp. 443-448). Springer, Berlin, Heidelberg.
[71] Yang, J. and Yang, J.Y., 2002. Generalized K–L transform based combined feature extraction. Pattern Recognition, 35(1), pp.295-297.
[72] Guo, G. and Dyer, C.R., 2003, June. Simultaneous feature selection and classifier training via linear programming: A case study for face expression recognition. In Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on (Vol. 1, pp. I-I). IEEE.
[73] Shan, C., Gong, S. and McOwan, P.W., 2006, June. A comprehensive empirical study on linear subspace methods for facial expression analysis. In Computer Vision and Pattern Recognition Workshop, 2006. CVPRW'06. Conference on (pp. 153-153). IEEE.
[74] Guo, M., Hou, X., Ma, Y. and Wu, X., 2017. Facial expression recognition using ELBP based on covariance matrix transform in KLT. Multimedia Tools and Applications, 76(2), pp.2995-3010.
[75] Rahulamathavan, Y., Phan, R.C.W., Chambers, J.A. and Parish, D.J., 2013. Facial expression recognition in the encrypted domain based on local fisher discriminant analysis. cognition, 4, p.6.
[76] Shan, C., Gong, S. and McOwan, P.W., 2009. Facial expression recognition based on local binary patterns: A comprehensive study. Image and vision Computing, 27(6), pp.803-816.
[77] Dongcheng, S., Fang, C. and Guangyi, D., 2013, July. Facial expression recognition based on Gabor wavelet phase features. In 2013 Seventh International Conference on Image and Graphics (pp. 520-523). IEEE.
[78] Liu, Ping, Joey Tianyi Zhou, Ivor Wai-Hung Tsang, Zibo Meng, Shizhong Han, and Yan Tong. "Feature disentangling machine-a novel approach of feature selection and disentangling in facial expression analysis." In European Conference on Computer Vision, pp. 151-166. Springer, Cham, 2014.
[79] Kumar, S., Bhuyan, M.K. and Chakraborty, B.K., 2016. Extraction of informative regions of a face for facial expression recognition. IET Computer Vision, 10(6), pp.567-576.
[80] Ekman, Paul. "Strong evidence for universals in facial expressions: a reply to Russell's mistaken critique." (1994): 268.
[81] Ferryman, J. "VS-PETS 2003." In IEEE Int'l Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance. 2003.
[82] Kanade, Takeo, Jeffrey F. Cohn, and Yingli Tian. "Comprehensive database for facial expression analysis." In Proceedings Fourth IEEE International Conference on Automatic Face and Gesture Recognition (Cat. No. PR00580), pp. 46-53. IEEE, 2000.
[83] Bartlett, Marian Stewart, Gwen Littlewort, Mark G. Frank, Claudia Lainscsek, Ian R. Fasel, and Javier R. Movellan. "Automatic recognition of facial actions in spontaneous expressions." Journal

of multimedia 1, no. 6 (2006): 22-35.

[84] Sebe, Nicu, Michael S. Lew, Yafei Sun, Ira Cohen, Theo Gevers, and Thomas S. Huang. "Authentic facial expression analysis." Image and Vision Computing 25, no. 12 (2007): 1856-1863.

[85] Kotsia, Irene, Ioan Buciu, and Ioannis Pitas. "An analysis of facial expression recognition under partial facial image occlusion." Image and Vision Computing 26, no. 7 (2008): 1052-1067.