Polytechnic University of Turin

Master degree course of Engineering

and Management

Master Degree Thesis

Face Recognition and its Multiple Applications

Candidate:

Federica Michienzi

Thesis Supervisor:

Federica Marcolin

Thesis Advisor:

Emilio Paolucci

Year: 2020

I declare and guarantee that all data, knowledge and information in this document has been obtained and presented in accordance with academic rules and ethical conduct.

Based on these rules and conduct, I have fully cited and referenced all material and results presented in this work.

Summary

Abstract ***4***

Introduction ***1***

1. Facial expression Recognition ***2***

2. Facial Expression Recognition (FER) Techniques ***4***

2.1 Facial Expression Recognition (FER) Evolution **6**

2.2 Facial Expression Recognition (FER) Current Trends **9**

2.2.1 Estimating intensity of facial expressions 9

2.2.2 Microexpressions analysis 10

2.2.3 AFER for detecting non-primary affective states 11

2.3 Methods and Algorithms **12**

2.3.1 Pre-processing 13

2.3.2 Feature extraction 14

2.3.3 Classification 14

3. Fields of Applications of Facial Expression Recognition (FER) ***17***

3.1 Human-Computer Interaction, Human-Robotic Interaction and Healthcare System **18**

3.1.1 Human-Computer Interaction (HCI) 18

3.1.2 Human-Robotic Interaction (HRI) 23

3.1.3 Healthcare System 30

3.2 Computer Animations **34**

3.2.1 Creation of lexicon of virtual character's facial expressions 35

3.3 Police Interrogation **43**

3.3.1 Dialogue in Police Training with children Interviewing 44

3.4 Drivers' Safety **47**

3.5 Marketing **49**

4. Actual Commercial Applications of Facial Expression Recognition (FER) ***54***

4.1 Face Recognition and Big Data Analysis **55**

4.1.1 Biometric and Big Data Technology 57

4.1.2 Biometric, Big Data and Cloud Technologies 58

4.2 Face Recognition and Social Networks **59**

4.3 Face Recognition and Machine Learning **60**

4.4 Face Recognition and Cloud Computing **62**

4.5 Real-life Applications of FER Technology **63**

4.5.1 Device authentication (phones, PCs, connected cars) 64

4.5.2 Security sector (airports, hotels…) 70

4.5.3 Health 71

4.5.4 Retail, Marketing, and Advertising 71

4.5.5 Customer or visitor ID (bank branches, casinos, private companies) 73
4.6 Best Facial Recognition Software **73**
5. Facial Expression Recognition (FER) Market ***79***
5.1 Face Recognition Industry Overview **79**
5.2 Facial Recognition Market Segmentation **80**
5.3 Technology Readiness and Availability **85**
5.4 The Face Recognition Market Trends **87**
5.5 Controversies of Face Recognition **89**
5.5.1 Privacy Violation 89
5.5.2 Imperfect technology in law enforcement 91
6. Conclusions ***1***
References ***3***

Abstract

Face Recognition is presently an emerging technology with numerous uses.
Facial Expression Recognition (FER) can be widely applied to various research areas, such as mental diseases diagnosis and human social/physiological interaction detection. Thanks to the emerging advanced technologies in hardware and sensors, FER systems have been developed to support real-world application scenes.

The global market for facial recognition is a growing one, estimated to reach USD 11.30 Billion by 2026, due to the benefits the technology provides over conventional security methods.

This thesis provides an analysis of all the applications deriving from the facial expression recognition technology up to the subdivision by geographical area of the different areas of interest.

Introduction

Facial expressions (FE) forms, together with voice, language, hands and posture of the body, a fundamental communication system between humans in social contexts.

Facial recognition, by analysing and comparing patterns based on the person's facial contours uniquely identifies a person. Facial recognition is used in a wide variety of areas (Computer Science, Engineering, Neuroscience, Medicine…).

There are different facial recognition techniques in use. Most facial recognition systems are based on the different fiducial points on a human face. The values measured against the variable associated with points of a person’s face help in identifying/verifying the person. Facial recognition techniques are evolving with new approaches such as 3-D techniques, which helped to overcome some critical issues such as light or variation of expressions.

The aim of this thesis is to describe the history and the actual utilization of the different technology based on facial recognition. In the next chapters we will see how Face Recognition has evolved and has become one of the most promising markets currently existing.

The thesis is divided into chapters. The first chapter briefly explains what is meant with Facial Recognition. The second chapter describes the method used for face recognition. It is divided into three sections describing the current state of art of facial expressions recognition, its history, current trends, methods used. The third chapter describes the subject areas for the applications of the Face Recognition technologies, while the fourth chapter analyses their real-life application. In the fifth chapter the Face Recognition Market is described.

1. Facial expression Recognition

A human face reveals a great deal of information to a perceiver. It can tell about mood and intention and attentiveness, but it can also serve to identify a person. A person can be identified by other means than the face, for example voice, body shape, gait or even clothing may all establish identity in circumstances where facial detail may not be available. Nevertheless, a face is the most distinctive and widely used key to a person’s identity. Face detection is the first step of many machine vision applications such as security applications, face recognition, photography and recognition of facial expressions.

However, as H. Ellis (1975, 1981) pointed out, this considerable empirical activity was not initially accompanied by developments in theoretical understanding of the processes underlying face recognition. It is only comparatively recently that serious theoretical models have been put forward (Bruce, 1979, 1983; Baron, 1981; H. Ellis, 1981, 1983, in press a; Hay & Young, 1982; Rhodes, 1985; A. Ellis et al., in press).

Facial expressions are defined as the facial changes in response to a person’s internal emotional states, intentions, or social communications.

Generally, the face offers three different types of signals:

- Static signals;
- Slow signals;
- Rapid signals.

The static signals are face shapes, skin colour, constitution of bones, cartilage and shape, location and size of facial features such as brows, eyes, nose, mouth, greasy deposits.
The slow signals are permanent wrinkles which include the changes in facial appearance such as skin texture and muscle tone changes. These signals are called Slow Signals because they happen slowly with time. The rapid signals flashes on the face and remain for a few seconds. It is very hard to alter static and slow signals, but rapid signals can be altered with individual options (Ekman and Friesen, 2003).

Depending on context, FEs may have varied communicative functions such as:

- regulate conversations by signaling turn-taking;
- convey biometric information;
- express intensity of mental effort, and signal emotion.

By far, the latter has been the one most studied.

2. Facial Expression Recognition (FER) Techniques

Facial expression analysis has been a research topic since Darwin's work in 1872. After that, progress has been made to build computer systems to help in understanding and using this natural form of human communication and a number of facial expression recognition (FER) systems have been created.

The main stages in a FER system for detecting emotions are two: extracting features and classifying emotions. A FER system robustly performs when the extracted features can (1) lessen the within-class variations and (2) maximise between-class variations. Good feature representation can guarantee an accurate and efficient recognition process. There are two types of feature-based facial expression recognition techniques: geometric-based and appearance-based[1]. The second key stage in a FER system is the classification of the emotions. The feature vectors are employed to train the classifier used to assign one of the expression labels to the input face. High-dimensional feature vector challenges the face analysis, which impacts the performance and speed of FER systems. Therefore, various feature reduction techniques are used to reduce the dimension of features for improving computation efficiency, especially for designing a real-time FER system [1].

Most of the facial expression recognition (FER) systems attempt to recognize six basic emotional expressions:

- Fear;
- Disgust;
- Anger;
- Surprise;
- Happiness;
- Sadness.

[1] See 2.2.2 Feature extraction.

These six basic emotions, shown in Figure 1, were introduced by Ekman.

Some FER systems define more emotions such as envy, drowsiness, contempt, and pain.

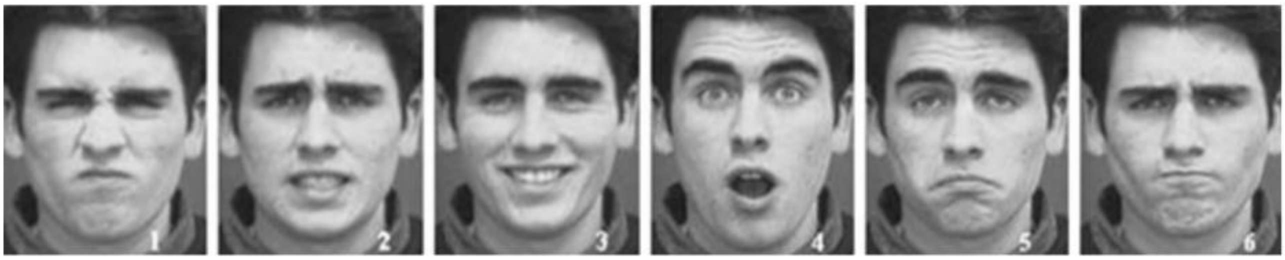

Figure 1. Primary emotions expressed on the face. From left to right: disgust, fear, joy, surprise, sadness, anger.

FER systems can be classified following different paths. Some systems are grouped into two types: spontaneous and pose-based. Spontaneous FER systems detect the facial expressions explicitly appearing on people's faces when they deal with daily situations, for example while watching movies. Pose-based expressions instead refer to artificial expressions which people mimic when they are asked to do in a sequence.

Other FER systems are categorized into micro-expression and macro-expression detectors. According to Ekman, micro-expressions indicate hidden expressions, which are quick (less than ½ second), brief and difficult to understand by humans, and individuals try to hide their pure emotions, especially in high-stake situation. Ekman has asserted that micro-expressions may be considered as the best cues for lie detection. In contrast, macro-expressions last ½ second to 4 seconds and occur in our daily interactions with others.

Although FER systems have broad applications in various areas (computer interactions, health-care systems and social marketing, and so on) facial expression analysis is incredibly challenging due to subtle movements of the foreground people and complex, noisy environment of the background in the real-world images/videos.

FER systems present in general three main challenges:

- illumination variation;
- subject-dependence;
- head pose-changing.

These challenges affect the performance of the systems.

In this chapter we will present how facial expression recognition evolved through time, the current trends, which methods are used and the main application fields.

2.1 Facial Expression Recognition (FER) Evolution

Facial expressions (FE) together with voice, language, hands and posture of the body, form a fundamental communication system between humans in social contexts. Automatic FE recognition (AFER) is an interdisciplinary domain standing at the crossing of behavioural science, neurology, and artificial intelligence [2].

Studies of the face were influenced in premodern times by popular theories of physiognomy and creationism. Physiognomy assumed that a person's character or personality could be judged by their outer appearance, especially the face [3]. Leonardo Da Vinci was one of the first to refute such claims stating they were without scientific support [4]. In the 17th century in England, John Buwler studied human communication with particular interest in the sign language of persons with hearing impairment. His book "Pathomyotomia or Dissection of the significant Muscles of the Affections of the Mind" was the first consistent work in the English language on the muscular mechanism of FE [5]. About two centuries later, influenced by creationism, Sir Charles Bell investigated FE as part of his research on sensory and motor control. He believed that FE was endowed by the Creator solely for human communication. Subsequently, Duchenne de Boulogne conducted systematic studies on how FEs are produced [6].
In the 19th century, Duchenne de Boulogne conducted experiments on how FEs are produced. Approximately in the same historical period, Charles Darwin firmly placed FE in an evolutionary context [7]. This marked the beginning of modern research of FEs.
More recently, important advancements were made through the works of researchers like Carroll Izard and Paul Ekman who inspired by Darwin performed seminal studies of FEs.

To the aim of this thesis we will focus on Automatic Face Expression Recognition (AFER).
The first work on AFER was published in 1978 [8]. It was tracking the motion of landmarks in an image sequence. Mostly because of poor face detection and face registration algorithms and limited computational power, the subject received little attention throughout the next decade.

The work of Mase and Pentland and Paul Ekman marked a revival of this research topic at the beginning of the nineties [9], [10].

In 2000, the *CK dataset* was published marking the beginning of modern AFER [11]. While a large number of approaches aimed at detecting primary FEs or a limited set of FACS action units (AUs) [12], [13], [14], [15], others focused on a larger set of AUs [16], [17], [18]. Some AUs are shown in Figure 2. Most of these early works used geometric representations, like vectors for describing the motion of the face [14], active contours for describing the shape of the mouth and eyebrows [15], or deformable 2D mesh models [13]. Others focused on appearance representations like Gabor filters [12], optical flow and LBPs [19] or combinations between the two [18].

Figure 2 Eight basic Action unit: AU 1, 4, 6, 9, 12, 15, 26, and 28.

The publication of the *BU-3DFE dataset* [20] was a starting point for consistently extending RGB FE recognition to 3D. While some of the methods require manual labelling of fiducial vertices during training and testing [21], [22], [23], others are fully automatic [24], [25], [26], [27]. Most use geometric representations of the 3D faces, like principal directions of surface curvatures to obtain robustness to head rotations [22], and normalized Euclidean distances between fiducial points in the 3D space [21]. Some encode global deformations of facial surface (depth differences between a basic facial shape component and an expressional shape component) [27] or local shape representations [28]. Most of them target primary expressions [22] but studies about AUs were published as well [28]. In the first part of the decade static representations were the primary choice in both RGB [12], [11], 3D [21], [29], [26], [22], [27] and thermal [30].

In later years various ways of dynamic representation were also explored like tracking geometrical deformations across frames in RGB [31], [13] and 3D [32], [33] or directly extracting features from RGB [17] and thermal frame sequences [34], [35].

Besides extended work on improving recognition of posed FEs and AUs, studies on expressions in ever more complex contexts were published. Works on spontaneous facial expression detection [36], [37], [38], [39], analysis of complex mental states [40], detection of fatigue [41], frustration [42], pain [43], [44], [45], severity of depression [46] and psychological distress [47], and including AFER capabilities in intelligent virtual agents [48] opened new territory in AFER research.

In summary, research in automatic AFER started at the end of the 1970's, but for more than a decade progress was slow mainly because of limitations of face detection and face registration algorithms and lack of sufficient computational power. From RGB static representations of posed FEs, approaches advanced towards dynamic representations and spontaneous expressions. In order to deal with challenges raised by large pose variations, diversity in illumination conditions and detection of subtle facial behaviour, alternative modalities like 3D and Thermal have been proposed. While most of the research focused on primary FEs and 11 AUs, analysis of pain, fatigue, frustration or cognitive states paved the way to new applications in AFER.

In Figure 3 we present a timeline of the historical evolution of AFER.

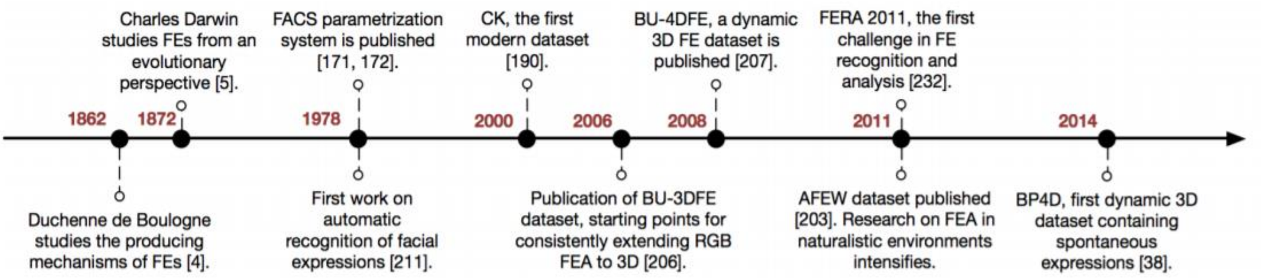

Figure 3 Historical evolution of AFER.

2.2 Facial Expression Recognition (FER) Current Trends

2.2.1 Estimating intensity of facial expressions

Detecting FACS AUs facilitates a comprehensive analysis of the face. However, being able to estimate the intensity of these expressions would have even greater informational value especially for the analysis of more complex facial behaviour. For example, differences in intensity and its timing can distinguish between posed and spontaneous smiles [49] and between smiles perceived as polite versus those perceived as embarrassed [50]. Moreover, intensity levels of a subset of AUs are important in determining the level of detected pain.

In recent years estimating intensity of facial expressions and especially of AUs has become an important trend in the community. As a consequence, the Facial Expression Recognition (FER) challenge added a special section for intensity estimation [51], [52]. This was recently facilitated by the publication of FE datasets that include intensity labels of spontaneous expression in RGB and 3D.

Even though attempts in estimating FE intensity have existed before, the first seminal work was published in 2006 [53]. It observed a correlation between a classifier's output margin, in this case the distance to the hyperplane of SVM classification, and the intensity of the facial expression. Unfortunately, this was only weakly observed for spontaneous FEs. A number of studies question the validity of estimating intensity from distance to the classification hyperplane. In two works published in 2011 and 2012 Savran et al. made a study of these techniques providing solutions to their main weak points [54], [55].

Other works consider the possible advantage of using 3D information for intensity detection because, even if 3D sensing noise can be excessive in the eye region and 3D misses the eye texture information, larger deformations on the lower face make 3D more advantageous.

A different line of research analyses the way geometrical and appearance representations could combine for optimizing AU intensity estimation [56], [57].

In summary, estimating facial AUs intensity followed a few distinct approaches. First, some researchers made a critical analysis about the limitations of estimating intensity from classification scores. As an alternative, direct estimation from features was analysed.

Further studies on optimal representations for intensity estimation of different AUs were published either from the points of view of geometrical vs appearance representations or the fusion between RGB and 3D. Finally, a third main research direction was focused on modelling the correlations between AUs appearance and intensity priors. Some works are treating a limited subset of AUs while others are more extensive. All the presented approaches use predesigned representations. While the vast majority of the works are performing a global feature extraction with or without selecting features there are cases of sparse representation.

2.2.2 Microexpressions analysis

Microexpressions are brief FEs that people in high stake situations make when trying to conceal their feelings. They were first reported by Haggard and Issacs in 1966. They are difficult to recognize. Even after extensive training, human accuracies remain low and an automatic system is highly useful. Microexpressions differ from other expressions also because of their subtleness and localization. These issues have been addressed by employing specific capturing and representation techniques.

As with spontaneous FEs (shorter and less intense than exaggerated posed expressions), methods for recognizing microexpressions take into account the expression dynamics. For this reason, a main trend in microexpression analysis is to use appearance representations captured locally in a dynamic way [58], [59], [60].

A problem in the evolution of microexpression analysis has been the lack of spontaneous expression datasets. Before the publication of the CASME and the SMIC dataset in 2013, methods were usually trained with posed non-public data. "*Recognising spontaneous facial micro-expressions*" [60] proposes the first microexpressions recognition system. LBP-TOP, an appearance descriptor is locally extracted from video cubes. Microexpressions detection and classification with high recognition rates are reported even at 25fps. Alternatively, existing datasets, such as BP4D, could be mined for microexpression analysis. One could identify the initial frames of discrete AUs, to mimic the duration and dynamic of microexpressions.

In conclusion, microexpressions are brief, low intensity FEs believed to reflect repressed feelings. Even highly trained human experts obtain low detection rates. An automatic microexpression recognition system would be highly valuable for spotting feelings humans are trying to hide. Due to their briefness, subtleness and localization most of methods in recent years have used local, dynamic, appearance representations extracted from high frequency video for detecting and classifying posed and more recently spontaneous microexpressions.

2.2.3 AFER for detecting non-primary affective states

Most of Automatic Facial Expression Recognition (AFER) was used for predicting both primary and non-primary affective states of basic emotions (such as, respectively anger/ happiness; complex mental states, fatigue, frustration, pain, depression, mood and personality traits).

Approaches related to mood prediction from facial cues have pursued both descriptive (e.g., FACS) and judgmental approaches to affect. In a paper from 2009, Cohn et al. studied the difference between directly predicting depression from video by using a global geometrical representation (AAM), indirectly predicting depression from video by analysing previously detected facial AUs and prediction depression from audio cues and they concluded that specific AUs have higher predictive power for depression than others suggesting the advantage of using indirect representations for depression prediction.

As humans rely heavily on facial cues to make judgments about others, it was assumed that personality could be inferred from FEs as well. Usually studies about personality are based on the Big-Five personality trait model which is organized along five factors: openness, conscientiousness, extraversion, agreeableness, and neuroticism. While there are works on detecting personality and mood from FEs only, the dominant approach is to use multimodality either by combining acoustic with visual cues or physiological with visual cues.

In conclusion, recent analysis of non-primary affective states mainly focused on predicting depression. For predicting levels of depression, local, dynamic representations of appearance were usually combined with acoustic representations. Studies of FEs for predicting personality traits had mixed conclusions until now.

2.3 Methods and Algorithms

In this thesis, facial expression analysis refers to computer systems that attempt to automatically analyse and recognize facial motions and facial feature changes from visual information. Sometimes the facial expression analysis is confused with emotion analysis in the computer vision domain. For emotion analysis, higher level knowledge is required: emotion's interpretation is aided by context, body gesture, voice, individual differences, and cultural factors as well as by facial configuration and timing [61]. Computer facial expression analysis systems need to analyse the facial actions regardless of context, culture, gender, and so on.

Facial expression analysis includes both measurement of facial motion and recognition of expression and the general approach to automatic facial expression analysis (AFEA) [62] consists of three steps:

- face acquisition/preprocessing: processing stage to automatically find the face region for the input images or sequences;
- facial data extraction and representation: after the face is located, facial changes caused by facial expressions are extracted and represented;
- facial expression recognition: the facial changes can be identified as facial action units or prototypic emotional expressions.

Face Expression Recognition (FER) techniques, as Figure 4 shows, include the three stages of pre-processing, feature extraction and classification.

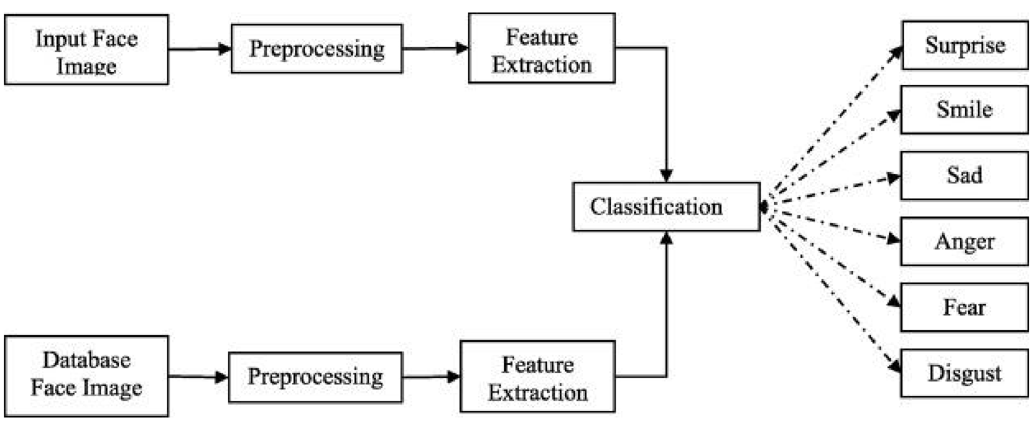

Figure 4 Architecture of face expression recognition system.

2.3.1 Pre-processing

Pre-processing 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 [63]. Image pre-processing includes different types of processes such as image clarity and scaling, contrast adjustment, and additional enhancement processes [64] to improve the expression frames [65] . Some processing techniques are:

- Bessel down sampling: used for face image size reduction but it protects the aspects and also the perceptual worth of the original image [66].
- The Gaussian filter: used for resizing the input images which provides the smoothness to the images [67].
- Normalization method: designed for reduction of illumination and variations of the face images [68] with the median filter and to achieve an improved face image. The normalization method also used for the extraction of eye positions which make more robust to personality differences for the FER system and it provides more clarity to the input images.
- Face alignment method: performed by using the SIFT (Scale Invariant Feature Transform) flow algorithm. For this, first calculate reference image for each face expressions. After that all the images are aligned through related reference images [69].
- ROI (Region of Interest) segmentation method: it is one of the important type of pre-processing method which includes three important functions such as regulating the face dimensions by dividing the colour components and of face image, eye or forehead and mouth regions segmentation [70]. In FER, ROI segmentation is most popular because for convenient segmentation of face organs from the face images.
- The histogram equalization method: used to conquer the illumination variations [71]. This method is mainly used for enhancing the contrast of the face images and for exact lighting also used to improve the distinction between the intensities.

In FER, more pre-processing methods are used but the ROI segmentation process is more suitable thanks to its ability to detect the face parts accurately.

2.3.2 Feature extraction

Feature extraction is finding and depicting of positive features of concern within an image for further processing. In image processing computer vision this stage spots the move from graphic to implicit data depiction and then these data depiction can be used as an input to the classification.

There are two types of feature-based facial expression recognition techniques:

- geometric-based;
- appearance-based.

The former refers to the FER systems, which extract local facial features including shape, positions, and angles between various facial elements, i.e., ear, eye, mouth and nose, and the feature vector is illustrated based on the geometrical relationship. The latter refers to the FER systems, which describe the appearance and employ texture information of face as a feature vector. The appearance-based approaches can obtain higher recognition rate and are more popular than geometric-based methods since it is a complicated task to find the efficient and proper geometric features in unconstrained environments and real-world applications. Wang et al. have comprehensively reviewed the facial feature extraction methods and pointed out the challenges brought by unconstrained or real-world environments [1].

2.3.3 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.

There is a huge number of methods used to classify expressions, some of them are:

- The directed Line segment Hausdorff Distance (dLHD) method: used for recognition of expressions [72];
- Euclidean distance metric: used for classification purpose which uses the normalized score and similarity score matrix for estimating Euclidean distance [73];
- Minimum Distance Classifier (MDC): one of the distance based classifier used for classification which estimates the distance between the feature vectors every sub image [74];

- The KNN (k – Nearest Neighbors) algorithm: classification method in which the relationship among the assessment models and the other models are estimated during the training stage [63];
- The Hidden Markov Model (HMM) classifier: statistical model which categorizes the expressions into different types [65];
- Hidden Conditional Random Fields (HCRF) representation: used for classification through full covariance Gaussian distribution for superior classification performance [75];
- Online Sequential Extreme Learning Machine (OSELM): method that uses RBF for classification. OSELM mainly contains two stages. They are initialization and sequential learning stages. Initialization stage includes the training samples [71]. Pair wise classifiers are also used for expression classification. It uses the one against one classification approach so exacting separation is utilized [76];
- ID3 Decision Tree (DT) classifier is a rule based classifier which extracts the predefined rules 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 [77];
- Learning Vector Quantization (LVQ): unsupervised clustering algorithm [64] which has two layers namely competitive and output layers. The competitive layer has the neurons that are known as 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 (MLP): used for classification and it contains three layers such as input layer, output layer and processing layer in which neurons are present [78];
- The Multilayer Feed Forward Neural Network (MFFNN) classifier: uses three layers such as input, hidden and output layers and back propagation algorithm for classification. In the training stage the weights are initialized and the activation units are estimated [66];
- Bayesian neural network 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 for its better accuracy [79];
- Convolution Neural Network (CNN) consists of two layers such as 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 with the two dimensional images where as in the subsampling layer, pooling

and redeployment are performed [80]. The CNN also contains two important perceptions likely shared weight and sparse connectivity [78]. 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 [81].

3. Fields of Applications of Facial Expression Recognition (FER)

Facial Expression Recognition (FER) can be widely applied to various research areas, such as mental diseases diagnosis and human social/physiological interaction detection. Thanks to the emerging advanced technologies in hardware and sensors, FER systems have been developed to support real-world application scenes.

The ability to automatically recognize FEs has a wide range of applications:

- AFER, usually combined with speech, gaze and standard interactions like mouse movements and keystrokes can be used to build adaptive environments by detecting the user's affective states [82], [83].
- Similarly, one can build socially aware systems or robots with social skills like Sony's AIBO and ATR's Robovie [84].
- Detecting students' frustration can help improve e-learning experiences [85].
- Gaming experience can also be improved by adapting difficulty, music, characters or mission according to the player's emotional responses [86], [87], [88].
- Pain detection is used for monitoring patient progress in clinical settings [89], [90], [91].
- Detection of truthfulness or potential deception can be used during police interrogations or job interviews [92].
- Monitoring drowsiness or attentive and emotional status of the driver is critical for the safety and comfort of driving [93].
- Depression recognition from FEs is a very important application in analysis of psychological distress [94], [95], [96].

In this chapter we will analyse in details some of the main FER's fields of application.

3.1 Human-Computer Interaction, Human-Robotic Interaction and Healthcare System

Today, the majority of our time is spent on interacting with computers and mobile phones due to technology progression. However, adding emotional expression recognition to expect the users' feelings and emotional state can drastically improve human–computer interaction (HCI), human–robotic interactions (HRI) and healthcare system.

Humans usually employ different cues to express their emotions, meaning facial expressions, hand gestures and voice. Facial expressions represent up to 55% of human communications therefore, considering facial expressions in an HRI system enables simulation of natural interactions successfully.

Indeed, robots can easily interact with humans in a friendly way when they can analyse facial expressions and figure out their emotional states. In this way, they can be used in a healthcare system to detect humans' mental states through emotion analysis and improve the quality of life. The mental states are unfolded in daily situations where robots can inspect positive and negative emotions. Positive facial expressions, such as happiness and pleasure, demonstrate healthy emotion states while unhealthy emotion states are represented by fetching negative facial expressions (e.g., sadness and anger). An efficient facial expression system (FER) can significantly help people to improve their mental emotion state by exploring their behaviour patterns. McClure and Coleman have shown that some mental diseases such as anxiety or autism are diagnosed by investigating the emotional conflicts, which appear on the patients' expressions.

3.1.1 Human-Computer Interaction (HCI)

Every day, computers and computer-based applications become more and more sophisticated and increasingly involved in our everyday life, whether at a professional, a personal or a social level. For this reason, the ability to interact with them in a natural way, similarly to the way we interact with other human agents becomes ever more important.

The most crucial feature of human interaction that grants naturalism to the process is our ability to infer the emotional states of others based on covert and/or overt signals of those emotional states. This allows us to adjust our responses and behavioural patterns accordingly, thus ensuring convergence and optimisation of the interactive process.

For the efficient use of computer systems, most computer applications need more and more communication. For that motive, human-computer interaction (HCI) has been a dynamic field of study in the last decades.

Human-computer interaction (HCI) is a discipline concerned on the communication between users and computers. It has the objective of improving safety, effectiveness and usability of computer-based products. One important attribute of an interactive system is accessibility meaning that it can be used by people with disabilities [97]. Vision-based interaction, computer input techniques based on a camera and computer vision techniques, is a flexible approach with great potential as computer input method for people with physical disabilities that make standard input devices difficult or impossible to use [98].

Early facial expression recognition (FER) systems detected seven basic emotions and are based on the already mentioned AUs. However, it is still challenging to classify facial expressions in real-life conditions because of pose and lighting variations. FER systems may fail because expressions never solely represent one emotion. Researchers have yet to adopt a continuous emotion framework to break the facial expressions into two dimensions: arousal and valence as Figure 5 shows. With that, expressions are typically classified in a broader sense of emotions.

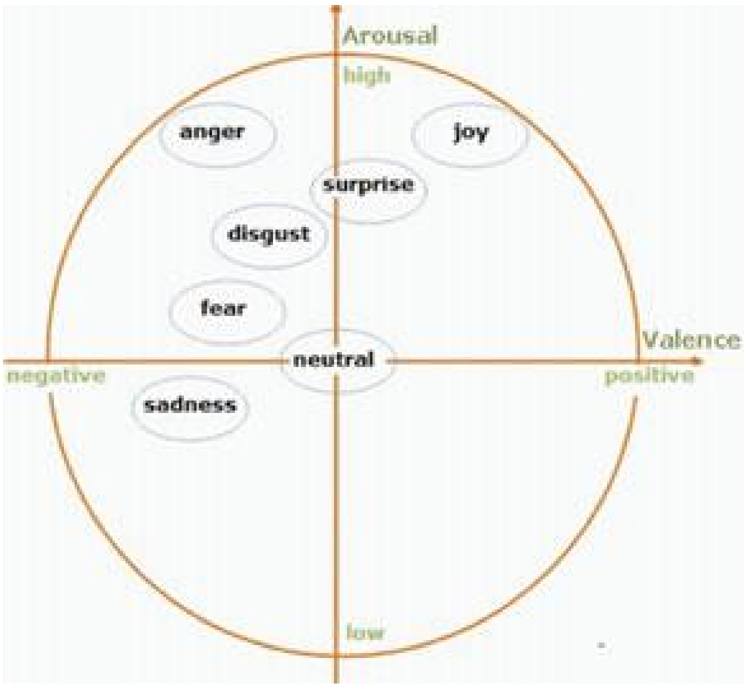

Figure 5 Two dimensions of emotion space and the distribution of seven basic emotions in arousal-valence (A-V) space [99].

To create a basic FER system in human-computer interaction interface, we need signal processing, both image processing and machine learning algorithms. It would start with a face detection algorithm followed by the facial landmark localization. A pre-processing algorithm can be included to normalize the images by removing any noises or imbalanced lighting. Then, the so-called features would need to be extracted from images to represent them in a way that computers can make use of them. All these steps (Figure 6) require accuracy, therefore advanced image processing techniques.

Recent developments in signal processing techniques and higher resolution cameras would allow more accurate face detection from images and facial landmark localization. The last step would be training a computer model that classifies/predicts the emotion label of any given face image and it can benefit from accessing thousands of images on internet in those days.

Excitingly, companies are investing in these types of technologies. MIT Labs created Affectiva, the world's largest facial expression database, by using deep learning and analysing almost four million face images. The facial expression analysis software Noldus classifies six basic emotions as well as gaze direction and head orientation. In the second half of 2016, Microsoft released the API for emotion recognition following. In January 2017 Apple bought Emotient, a start-up using AI to recognize facial expressions.

Graphical HCI systems are the mouse and the keyboards. Even if the innovation of the mouse and keyboard is a great development, there are still circumstances in which these devices are irreconcilable for HCI, for example in the case of communication with 3D objects. The two points of freedom of the mouse could not suitably emulate the 3 dimensions of space. The use of hand gestures offers a smart and natural optional to these burdensome interface tools for human computer communication. With use of hands gesture recognition, system can help people to interactive with computers in a more intuitive mode. Hand gesture recognition owns wide applications in sign language recognition [100], [101], computer games [102], virtual reality [103] and HCI systems [104].

Physical disability can, in certain cases, be a barrier for traditional human-computer interaction based on keyboard and mouse devices. Alternative ways of interaction based on computer vision may be successfully adapted in particular cases of disability [105]. To combat social exclusion of people with disability in interaction with technology, the idea to use the facial expression recognition as interaction type arises. The use of computer vision and digital image processing methods, namely, facial expression recognition appears as an alternative to human-computer interaction. The facial expressions are pre-programmed, being attributed a specific function in the computer for each expression made by user.

In the case of people with severe motor impairment, only some facial muscles can be activated, reducing the communication abilities to "yes" or "no" responses which could be recognized through two different facial expressions. Andreia Matos, Vítor Filipe and Pedro Couto exanimated the case of a child with a degenerative neuromuscular disease: the child is unable to move the members or the head and he can only use some facial muscles to make two expressions: “smile” and “kiss”. Their system uses a standard webcam and computer vision techniques to recognize the user’s facial expression and execute the corresponding computer command.

Facial expression is a demonstration of our affective state, cognitive activity and thought. It is understood that the mental state of a person can be recognizable by its face expression (in non-verbal communication) [106].

Usually, there are three stages in this type of analysis: face detection, facial feature extraction and expression recognition. In the first stage of the process, the system detects face image regions in the input image. Face detection can be a difficult task because there are factors such as hair, make up, shave, moustache, glasses or hats, that difficult the location of different facial characteristics. Moreover, scale and the orientation of human face on image, is another major difficulty, which leads to the non-use of fixed models to find the features. The presence of others objects and image noise are also problematic factors that add difficulties to these techniques. Also, in this particular case, as the subject is always with her tongue out of her mouth, another major difficulty, perhaps the biggest, arises. In a second stage, facial characteristics are extracted. These characteristics can be of two types: geometric or permanent characteristics and appearance or transient characteristics. In the last stage of the process, the facial expression is recognized.

The camera continuously acquires video frames, represented in RGB colour space, which are individual processed to recognize the facial expression. Firstly, a face detector should tell whether the image contains a human face and if so, where it is. The face image region is detected using the Viola and Jones algorithm [107]. This algorithm is one of the methods most used to detect human faces and has three major steps:

- Creation of integral image: a new image representation is created in a feature space, based on Haar basis functions;

- Classifier training: a variant of AdaBoost learning algorithm is used both to select the "best" features and to train the classifier;
- Construction of a cascade of classifiers: early stages of the cascade use simpler classifiers to reject many of the negative sub-windows while more complex classifiers are called to focus on promising regions.

After face detection, the mouth and eyes are extracted, also using Viola and Jones algorithm. The eye and mouth are the two facial characteristics used to recognize the expression. A scale normalization procedure is applied in order to obtain images always with the same dimension. The reference eye image dimension has been established in 92x112 pixels and the reference image mouth dimension has been established 145x92 pixels. These dimensions were defined after several tests with different values to obtain the best results. The mouth and eye appearance vary among images depending on the viewpoint, illumination conditions and the person itself.

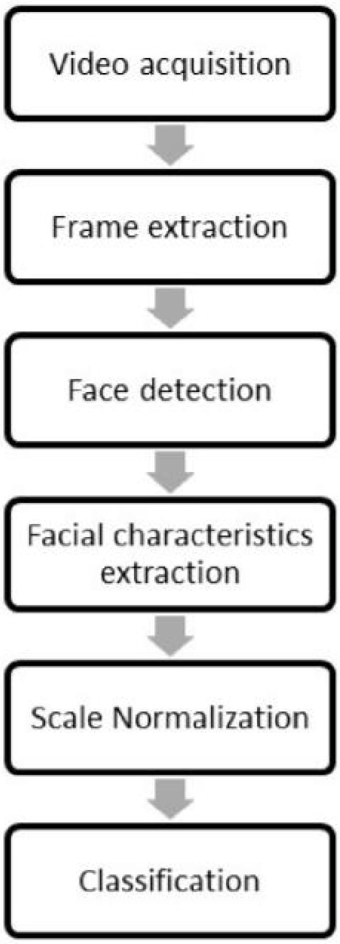

Figure 6 Diagram of facial expression recognition algorithm.

3.1.2 Human-Robotic Interaction (HRI)

"An intelligent user-interface not only should interpret the face movements but also should interpret the user's emotional state" (Breazeal, 2002).

To be effective in the human world, robots must respond to human emotional states. Knowing the emotional state of the user makes machines communicate and interact with humans in a natural way: intelligent entertaining systems for kids, interactive computers, intelligent sensors, social robots [108].

Knowing the human user's intentions and feelings enables a robot to respond more appropriately during tasks where humans and robots must work together [109] [110] [111], which they must do increasingly as the use of service robots continues to grow. In robot-assisted learning, a robot acts as the teacher by explaining the content of the lesson and questioning the user afterwards. Being aware of human emotion, the quality and success of these lessons will rise because the robot will be able to progress from lesson to lesson just when the human is ready [112]. Studies of human-robot interaction will be improved by automated emotion interpretation.

Natural human-robot interaction also requires detecting whether or not a person is telling the truth. Micro expressions within the face express these subtle differences. Specifically, trained computer vision applications would be able to make this distinction. Today's techniques for detecting human emotion approach this challenge by integrating dedicated hardware to make more direct measurements of the human [113] [114] [115].

There are several examples of artificial emotional systems that have been used to communicate emotions. Some of these have implemented animal-like designs to emphasize an emotional behaviour (e.g. Kismet [116] and Eddie [117]). Others have focused on emulating the elements of the human face, like the eyes, nose, nose, lips, eyelids, and eyebrows. These humanoid designs commonly rely on electromechanical [118] or digital [119] [120] robotic faces. There are two approaches for the appearance of the humanoid robotics face: the first mimics the look of the human face (e.g. android Actroid-SIT [121]) and the second uses a minimal number of facial elements (e.g. robot MiRae [122]). However, universal facial expressions must be legible, independently of the nature of the robotics face.

The first robots to explore facial expression legibility were WE-3RIV and WE-4RII that focused primarily on universal emotions [123] [124]. Other humanoid robots extend their capabilities to diverse facial expressions and social applications, such as the comedian robot Kobian-R14 based on the WE-3RIV and WE-4RII robots and Kaspar, a robot that focuses on the expression of emotions for medical purposes. The expression of emotion allows robots to communicate emotional states during the execution of a task like Lisa8 and Bender [125].

In social interaction, facial expressions help to communicate an emotional and mental state which is expected to be consistent. A common problem is how to access people's mental state [126] as this cannot be perceived directly and must be inferred by observers [127]. Thereby, facial expression of emotions can be a window to a mental state in the context of social interaction. Robotics and artificial systems can display an emotional state in order to influence and persuade their users [128].

In contrast to human–machine interactions with desktop computers, during interactions with robots the position and orientation of the human is less constrained, which makes facial expression recognition more difficult. The human may be further from the robot's camera, e.g he or she could be on the other side of a large room. The human may not be directly facing the robot. Furthermore, he or she may be moving, and looking in different directions, in order to take part in the robot's task. As a result, it is difficult to ensure a good, well illuminated view of the human's face for FER. Moreover, robots are mobile. Most image subtraction approaches cannot cope with this problem, because image subtraction is intended to separate facial activity from the entire surrounding that is expected to be static.

A technical challenge is to meet the real time computational constraints for human-robot interaction. The robot must respond to human emotions within a fraction of a second, to have the potential to improve the interaction. Many researchers have studied FER for robots. Some model–based methods cannot provide the real time performance needed for human–robot interaction, because model fitting is computationally expensive and because the required high-resolution images impose an additional computational burden [129]. Kim et al [129] use a set of rectangular features and train using AdaBoost. Yoshitomi et al fuse speech data, face images and thermal face images to help a robot recognize emotional states [130]. An HMM is used for speech recognition. Otsuka and Ohya [131] use HMMs to model facial expressions, for recognition by robots.

We are going to mention some innovative approaches to human-robotic interaction based on facial expression recognition:

<u>Matthias Wimmer, Bruce A. MacDonald, Dinuka Jayamuni and Arpit Yadav's Prototype:</u>

Matthias Wimmer, Bruce A. MacDonald, Dinuka Jayamuni and Arpit Yadav's approach [132] makes use of model-based techniques, which exploit a priori knowledge about objects, such as their shape or texture. Reducing the large amount of image data to a small set of model parameters facilitates and accelerates the subsequent facial expression interpretation, which mitigates the computational problems often associated with model–based approaches to FER.

According to the usual configuration [133], this model-based approach consists of seven components, which are illustrated in Figure 7. This approach fits well into the three-phase procedure of Pantic et al. [134], where skin colour extraction represents a pre-processing step, mentioned by Chibelushi et al. [135]. Phase 1 is contained by the core of model-based techniques: the model, localization, the fitting algorithm, and the objective function. Phase 2 consists of the facial feature extraction, and Phase 3 is the final step of facial expression classification.

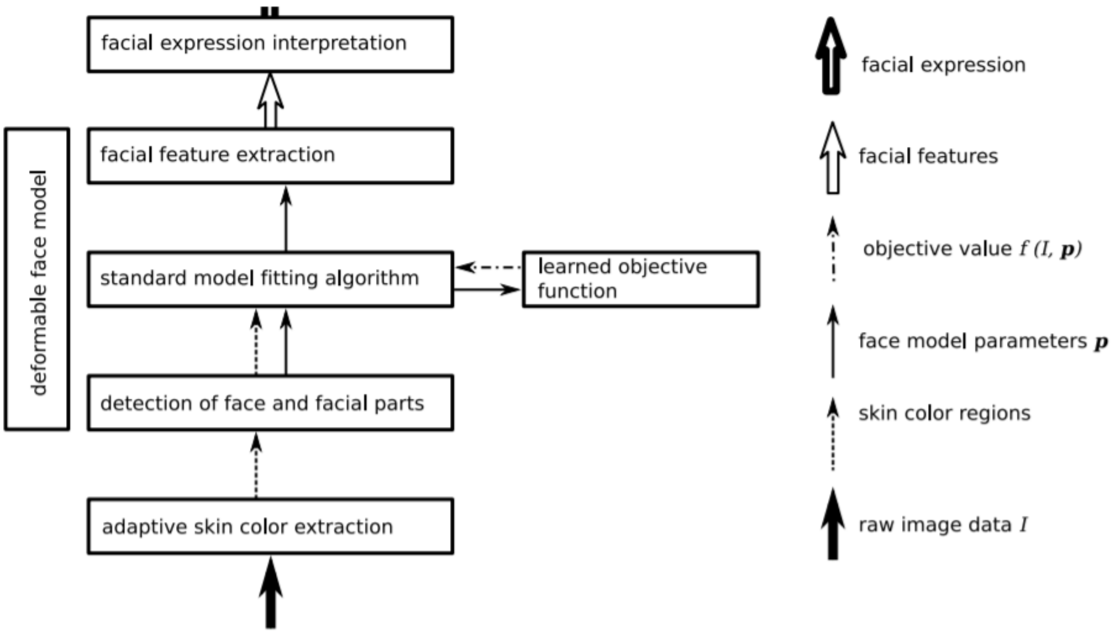

Figure 7 Model-based image interpretation splits the challenge of image interpretation into computationally independent modules.

The model contains a parameter vector p that represents its possible configurations, such as position, orientation, scaling, and deformation. Models are mapped onto the surface of an image via a set of feature points, a contour, a textured region, etc. The approach makes use of a statistics-based deformable model.

This approach imposes lower computational requirements by specializing model-based techniques to the face scenario. Adaptive skin colour extraction provides accurate low-level information. Both the face and the facial components are localized robustly. Specifically learned objective functions enable accurate fitting of a deformable face model. Experimental evaluation reports a recognition rate of 70% on the Cohn–Kanade facial expression database, and 67% in a robot scenario, which compare well to other FER systems. The technique provides a prototype suitable for FER by robots.

<u>The AIBO Robot:</u>

The AIBO robot: AIBO is a biologically-inspired robot able to show its emotions through an array of LEDs situated in the frontal part of the head as shown in Figure 8. The input to the system is a video stream capturing the user's face. In addition to the LEDs' configuration, the robot response contains some small head and body movements. From its concept design, AIBO's affective states are triggered by the Emotion Generator engine. This occurs as a response to its internal state representation, captured through multi-modal interaction (vision, audio and touch). For instance, it can display the 'happiness' feeling when it detects a face (through the vision system) or it hears a voice. But it does not possess a built-in system for vision-based automatic facial-expression recognition. For this reason, Fadi Dornaika and Bogdan Raducanu created an application for AIBO whose purpose is to enable it with this capability.

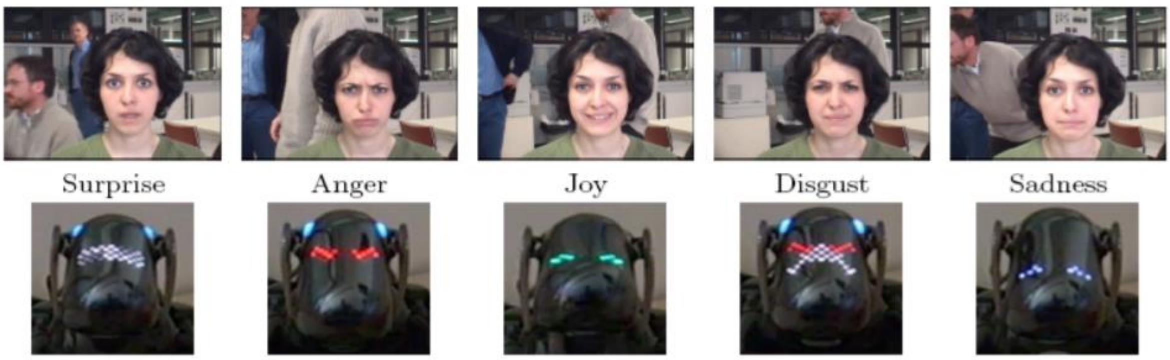

Figure 8 Five detected keyframes. These are shown in correspondence with the robot's response. The middle row shows the recognized expression. The bottom row shows a snapshot of the robot head when it interacts with the detected and recognized expression.

ROMAN and ROBIN Robots:

ROMAN robot has been developed by RRLAB as a test platform for human-robot interaction [136]. It consists of an upper body, two arms, neck, and an expressive head. ROMAN can generate facial expressions, gestures, and expressive body postures. It can generate nonverbal expressions using 47 DOF. Furthermore, it is also equipped with an expressive speech synthesizer. Figure 9 shows the ROMAN robot.

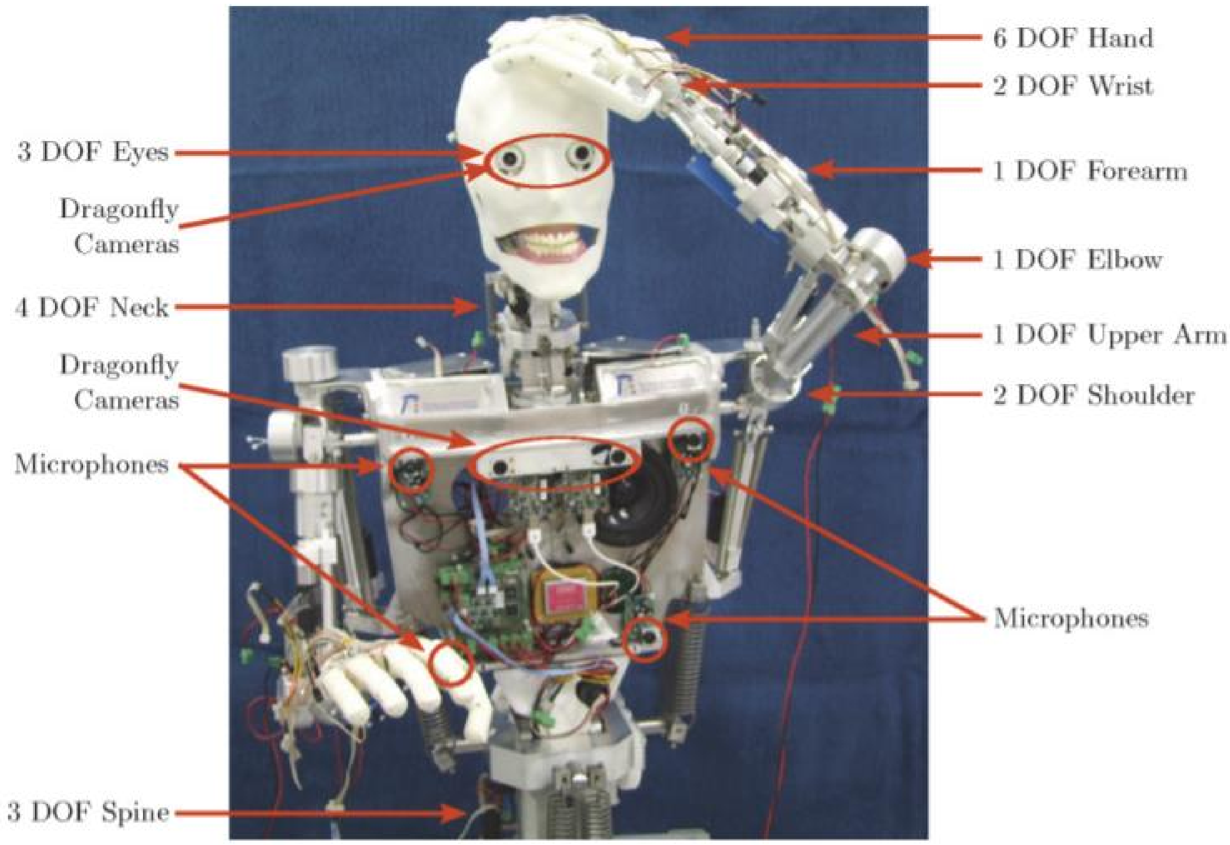

Figure 9 Interactive Humanoid Robot, ROMAN.

The control architecture of the robot, emotion-based architecture, is developed by Hirth [137]. The goal of the implementation is to realize the functions of emotions: regulative (emotion), selective (percepts), expressive (habits), motivational (motives), and rating (emotion), as well as the secondary functions.

Compared to previous publications where the rating and the regulative function were not included, all five functions of emotion mentioned in are realized. The perception system perceives and interprets information of the environment. Depending on this information, direct responses, performed by the habits, are activated and the motives calculate their satisfaction.

Beside the control architecture, ROMAN has been equipped with expressive function which enables it to produce different emotional states. Figure below depicts the facial expression of six basic emotions. A silicon skin has been glued to ROMAN's head as shown in Figure 10. By moving small metal plates, the different expressions have been generated [138].

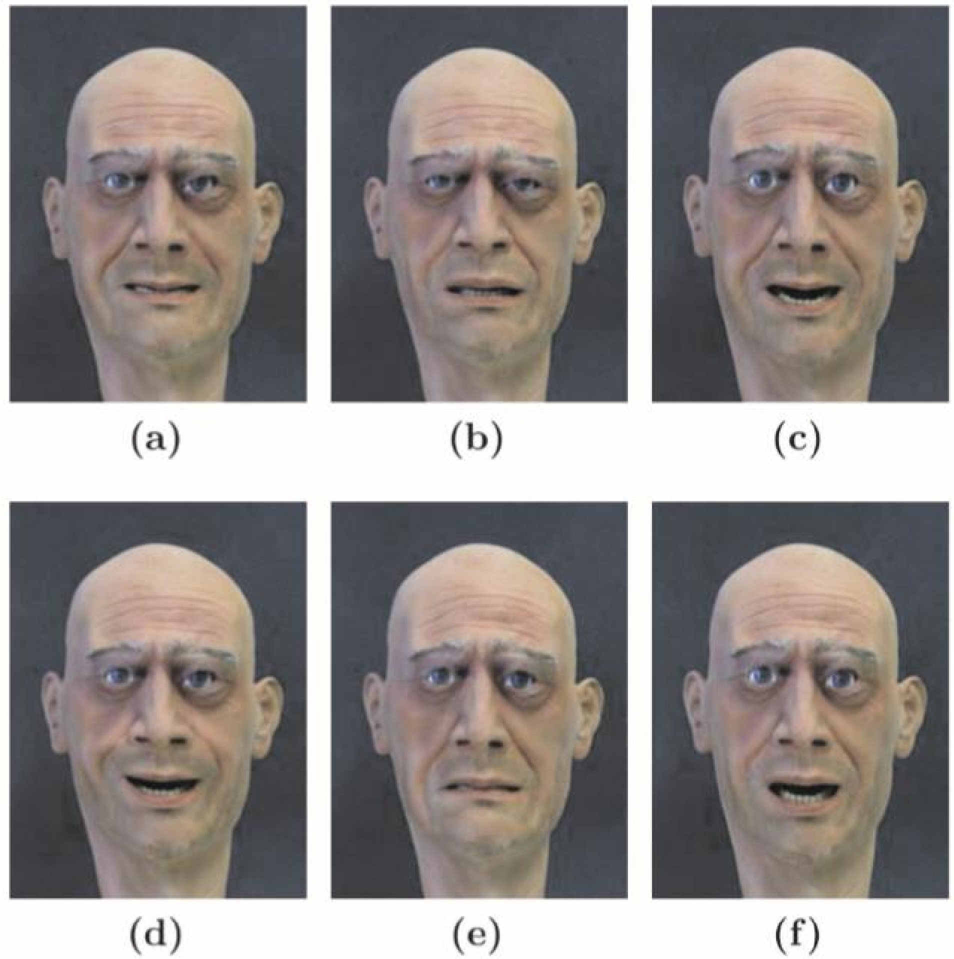

Figure 10 Facial expressions of ROMAN realized with a silicone skin (anger, disgust, fear, happiness, sadness and surprise).

However, ROMAN lacks in robustness when performing different gestures in real-time. Apart from emotion-based control architecture, important interactive modules like human feedback, context-aware perception and personality traits inference are missing in ROMAN. For that reason, Technical University of Kaiserslautern has developed a social humanoid robot, ROBot-human-INteraction (ROBIN). It is equipped with a backlit projected face, arms, hands and torso. It can speak via its built-in speech synthesis module in English and German language. The face makes use of projective technology to express almost any facial expression using action units. The head is able to move sideways for about ±45 degrees. RGB-D sensor is installed on the chest of robot. Additionally, a high definition camera is also installed on the head. The whole arm has 14 degrees of freedom, where hands are able to perform nearly all gestures (see Figure 11). ROBIN has its own processor that can handle all the movements of joints.

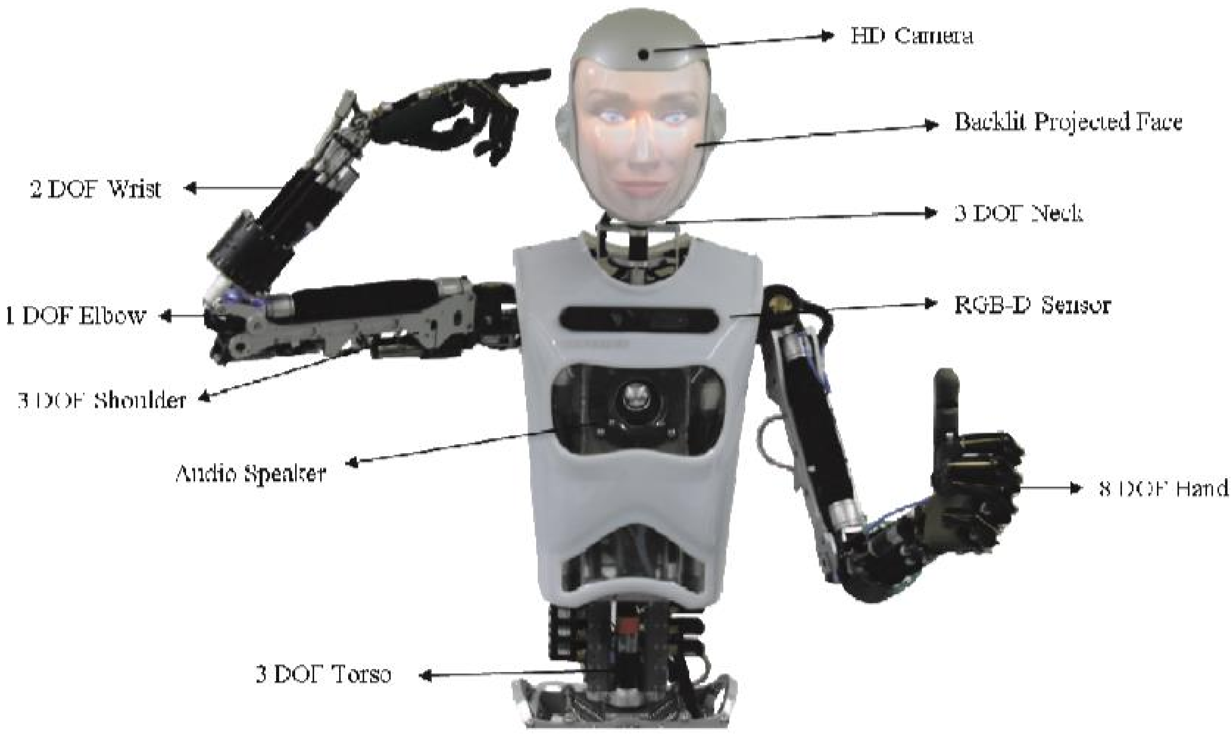

Figure 11 Interactive humanoid robot, ROBIN.

The human perception system is a part of the central nervous system, which is responsible for detecting and interpreting the surrounding environment. Analysing and interpreting the nonverbal cues is a challenging task even for humans. In order to achieve emotion-based human-robot interaction, there is a need of robust perception system that is able to perceive and understand some basic human nonverbal behaviour as well as some high-level perception tasks which include understanding human feedback communication channel, context-aware interaction and personality traits inference. ROBIN perception system can be divided into 2 subcategories: low-level perception and high-level perception system. Former deals with low-level features, e.g., facial expressions, posture, head pose, etc. while latter deals with complex human behaviours, e.g., feedback perception, context-aware perception and personality traits perception. The system is tested in real-time with active low-level and high-level perception features.

3.1.3 Healthcare System

With the rise of the robot and cloud computing technology, human-centric healthcare service attracts widely attention in order to meet the great challenges of traditional healthcare.

Along with the continuously growing aging trends and the dramatic increase in the elderly population, the healthcare sector is under a heavy burden due to the shortage of medical facilities and healthcare workers [139] [140]. In recent years, the expenses on healthcare services all over the world have been increased continuously, which has caused serious financial burden to the society and the government. In view of this situation, it is essential to develop economical, convenient and scalable home healthcare systems to effectively enhance the informatization degree of healthcare services and balance the uneven distribution of medical resources [141]. Moreover, with the explosion of mobile terminals [142], user's physiological signals can be easily collected which greatly encourages the development of intelligent healthcare.

In recent years, considerable research on sensors has contributed to the popularity of the home healthcare system, and various healthcare devices are developed for accurately measuring, analysing the user's physiological data and corresponding feedback [143]. However, the measurement range of physiological data, battery lifetime and the accuracy of data acquisition have been slightly restricted since the size and weight of devices cannot be too large considering comfortable and wearable [144] [145].

In order to overcome the defects and provide users with high Quality of Service (QoS) and Quality of Experience (QoE) healthcare services, the approach based on mobile robot is applied to collect environmental Cluster Compute (2015) and physiological data [140].

Thanks to the fact that robots can travel through freely, search the user and open the camera to transport video or photo to remote healthcare centre in health emergencies, physiological data acquisition modules can be integrated, and the measurement accuracy and the battery lifetime can be greatly improved. Moreover, due to robots humanized appearance (and the following functions: capture the user's voice, image and video; play audio and video), robots can be completed with the purpose of the human emotional communication [146]. For example, in the elderly healthcare system, the robot will collect the old man's voice and facial expressions, then such as data is uploaded to the remote healthcare centre. The healthcare system can estimate the old man's mental condition by multimodal emotional analysis. When the system determines that the human was lonely or depressed mood, the robot can initiatively play alleviate discomfort mood music or video for the old man, even more the robot can give the elderly psychological comfort by a robot dance.

Healthcare Systems in Smart Cities:

According to Ghulam Muhammad, Mansour Alsulaiman, Syed Umar Amin, Ahmed Ghoneim and Mohammed F. Alhamid, a facial-expression recognition system is proposed to improve the service of the healthcare in a smart city [147]. The proposed system applies a bandlet transform to a face image to extract sub-bands. Then, a weighted, center-symmetric local binary pattern is applied to each sub-band block by block. The CS-LBP histograms of the blocks are concatenated to produce a feature vector of the face image. An optional feature-selection technique selects the most dominant features, which are then fed into two classifiers: a Gaussian mixture model and a support vector machine. The scores of these classifiers are fused by weight to produce a confidence score, which is used to make decisions about the facial expression's type. Several experiments are performed using a large set of data to validate the proposed system. Experimental results show that the proposed system can recognize facial expressions with 99.95% accuracy.

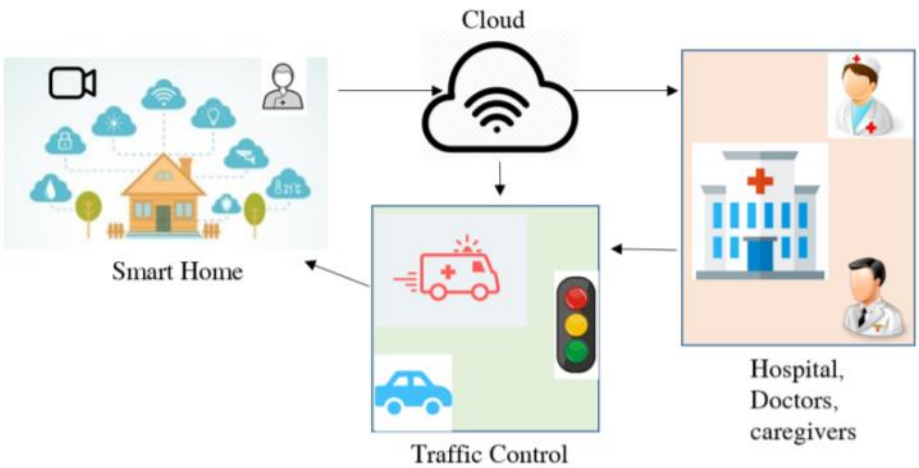

Figure 12 . A framework of a smart city for a smart healthcare.

Figure 12 shows a general healthcare framework in a Smart City. There are many smart homes in the city, each equipped with various required smart devices (smart cameras, smart appliances, smart video, smartphones, smart alarm systems, smart switches, smart locks, and so on). In the figure, we see the flow of data, decisions, and actions in healthcare in the Smart City. The sensors capture signals or data from a resident in the smart home. These signals are transferred to the cloud for processing. A cloud manager handles authentication and access issues, while a cloud server processes the signal and makes a decision. The decision is then passed to certain registered hospitals, doctors, and caregivers. The final decision comes from the doctor, who then alerts the caregivers, traffic managers, and hospitals to take appropriate actions.

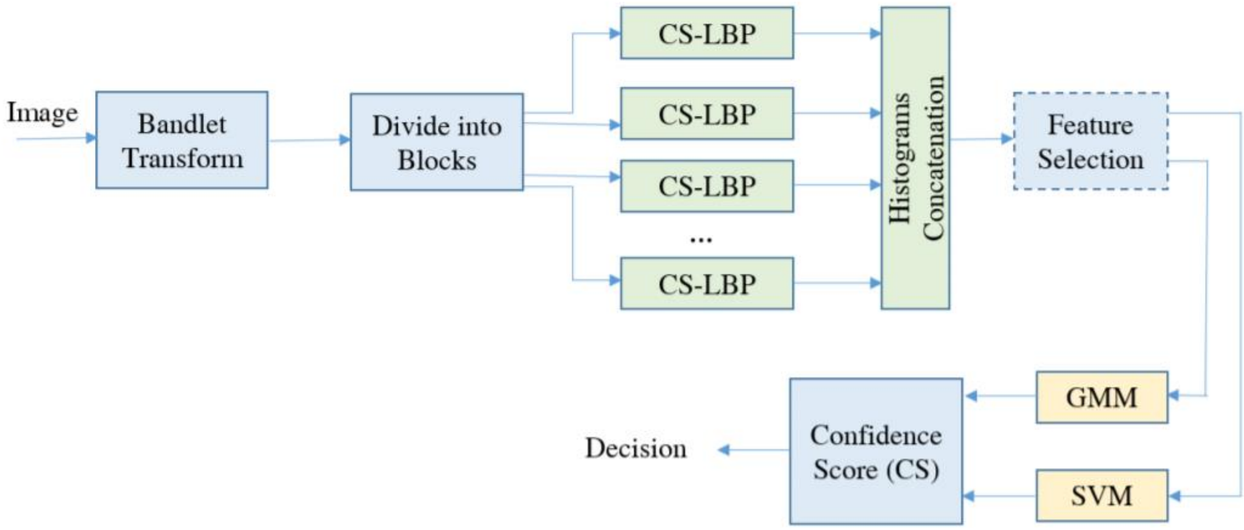

Figure 13 . Block diagram of the proposed facial-expression recognition system.

Figure 13 shows a block diagram of the proposed facial expression recognition system. A smart video or a smart camera constantly takes images of the patient in the smart home. The input to the system is the image taken by these sensors. Once an image is captured, a face detector locates the facial region in the image. Nowadays, almost all smart cameras have embedded face-detection modules. The detected face image is then transferred to the cloud. In the cloud server, the bandlet transform decomposes the image into several sub-bands at different scales: scale 0, scale 1, and scale 2. The blocks are of sizes 2×2, 4×4, 8×8, and 16×16. The bandlet transform is an improved model of traditional wavelet transforms [148]. In the bandlet transform, the geometrical structures of facial expressions are represented by some orthogonal bandlet bases. To accurately represent the geometric flow, the image is divided into small blocks, where a block can contain only one contour. Normally, smaller blocks can capture the geometrical flows more accurately than can larger blocks. The CS-LBP is applied to each block of the sub-band.

The calculation of the CS-LBP is expressed by the equation (see Figure 14):

$$CSLBP_{P,R} = \sum_{j=0}^{(P/2)-1} 2^j q\left(p_j - p_{j+(P/2)}\right)$$

$$q(x) = \begin{cases} 1, & x \geq T \\ 0, & \text{otherwise} \end{cases}$$

Figure 14 Equation for the calculation of the CS-LBP

Where:

- pj is the gray-scale intensity of the pixel (p);
- P and R are the numbers of pixels in a circular neighborhood, where the radius of the circle is R. In this work P = 8 and R = 1.

To preserve the spatial information in the CS-LBP histogram, it is calculated block-by-block. The histogram of each block is assigned a weight. The weight of each block is calculated by the information entropy of the block.The weighted CS-LBP histograms from the blocks are concatenated to produce a feature vector of the image. The number of features in the feature vector depends on the number of blocks and the number of bins in the CS-LBP histogram. Typically, the number of features is high, which may increase the time required to make a decision. Therefore, in the proposed system, a simple feature-selection technique is applied in the form of the Kruskal-Wallis (KS) test [149]. The KS test is a non-parametric, one-way analysis of variance that works on two or more classes. For a certain feature, it checks whether or not the medians of the classes are similar; based on this similarity, it returns a value p. If the value of p is close to 0, the feature is selected, because it is considered discriminative. In this work, 30 features were chosen according to their high p values. In the proposed system, two classifiers are used: the GMM and the SVM. The GMM is a stochastic method of modeling, frequently used in multiclass problems including speech/speaker recognition, emotion recognition, and environment recognition [150]. The SVM is a powerful binary classifier [151] that is also used in many image-processing applications. The scores of the two classifiers were combined using a weighted coefficient to make the recognition decision.

Experiments were performed using three datasets: one locally recorded and the two others public. Using the local dataset, the system achieved 99.95% accuracy, obtaining 99.9%

accuracy with the two public datasets. Experimentally, the following key points were determined: (1) Bandlet coefficients at scale 0 gave better accuracy than those at higher scales. (2) Block-based weighted CS-LBP on the bandlet subband achieved better accuracy than non-weighted CS-LBP. (3) As classifiers, both the GMM and the SVM had comparable performance. (4) Accuracy was improved by using both classifiers' scores.

The proposed facial-expression recognition system can be used in a smart healthcare framework. With this system, registered doctors and caregivers can constantly monitor patients' feelings remotely and take appropriate actions as required. The system can also give stakeholders automatic feedback from patients without needing to ask them for feedback.

3.2 Computer Animations

A long time ago most of the cartoons and animated movies were made with months and sometimes years of work to create realistic characters. Nowadays computer graphics became more realistic than ever. Computer vision field includes facial motion capture and facial animation, which makes these characters as vivid and live as real ones [152]. Computer facial animation uses data to create animated humans, animals or even imaginary creatures, which is applied not only in cinematographic industry, but in education, scientific simulation and communication as well. Technological advancements made improvements in facial animation.

An efficient FER plays a crucial role in computer graphics where human faces had to been modelled and parameterized as avatars and computer animations by accurately characterizing of face geometry and muscle motion. Many technologies such as virtual reality (VR) and augmented reality (AR) employ a robust FER to implement a natural, friendly communication with humans.

Generation of virtual human via computer animation model is in exponential progression. Currently, the natural interaction capability of an avatar is guided by the developments of artificial intelligence, diverse sensing technology, and advanced computer graphics [153]. The realistic facial animation applications offered an opportunity to bring facts and expressions of human to the social reality. It is realized that the virtual characters in computer games and simulations must relate to real situations with greater geometric details. The actions and approaches for automatic blending of the body together with facial motions are well recognized [154]. The task is becoming challenging because the animations at each event require the synchronization of facial components in speech which involve facial bones, muscles, and lips. Animation of highly accurate human face is complicated. Consequently, inclusion of several notable attributes in the facial expressions to precisely visualize the people feelings and mental states is essential. The facial action coding system (FACS) is used to process and describe the facial behaviours via every facial muscle [155]. For each facial muscle, it is a standard practice to change the appearances based on the anatomy analysis of human facial muscles conduct together with the tongue and jaw motion [156]. Facial anatomy deals with the changes in facial expression caused by their actions. Unlike muscle-based techniques, FACS works on facial actions and facial behaviours by studying the muscle actions rather than the muscle itself. Understanding the origin of facial expressions is complicated because they originate from cooperative effects of many muscles. Facial AUs are introduced to support the actions of facial muscles. Each muscle is divided into two or more AUs to completely explain autonomous actions of muscle parts. FACS classifies the human face into forty-six AUs because each unit embodies individual muscle action or groups of muscles describe a single facial state. Exact taxonomy of AUs on the face is to mimic entire facial muscle actions and the blending of different AUs yields diverse facial expression. For example, merging the AU4 (brow raiser), AU15 (lip corner depressor), AU1 (inner brow raiser), and AU23 (lip tightener) produces sad appearance. FACS explains true facial expressions built by all possible facial animation units prepared by the human face gauging head and eye point only.

3.2.1 Creation of lexicon of virtual character's facial expressions

To create a lexicon of virtual character's emotional facial expressions, two methods may be distinguished [157]:

1. A first method consists in exploiting the empirical and theoretical research in Human and Social Sciences on the characteristics of human's emotional faces.

2. A second method is based on the study of annotated corpus containing the expressions of emotions displayed by humans or virtual characters.

3.2.1.1 Theoretical based lexicon of facial expressions

To create a repertoire of a virtual character's facial expressions, the method that is commonly used consists in exploiting the empirical and theoretical studies in Psychology that have highlighted the morphological and dynamic characteristics of human's facial expressions. Different theories lead to different approaches.

<u>The categorical approach by Ekman</u> [61]: This approach is widely used in the domain of virtual characters to simulate emotional facial expressions. The Moving Pictures Experts Group MPEG-4 standards support facial animation by providing Facial Animation Parameters (FAPs) as well as a description of the expression of the six basic emotions [158].

<u>Dimensional approach</u>: This approach allows virtual characters to express a large number of emotional expressions. In dimensional models, a new expression is often created by applying some arithmetical operations, such as linear interpolation, on numerical definitions of discrete emotions placed in the multi-dimensional space. For instance, the model called Emotion Disc [159] uses a bi-linear interpolation between two basic expressions and the neutral one. In this approach, six expressions are spread evenly around the disc, while the neutral expression is set at its centre. The distance from the centre of the circle and an expression represents its intensity. The spatial relations in 2D are used to establish the expression corresponding to any point of the Emotion Disc.

Several dimensional approaches are based on 3D space models (PAD). These models define emotions in terms of pleasure (P), arousal (A) and dominance (D) [160]. The model proposed by Zhang and colleagues is based on PAD and a new parameterization of facial expressions: Partial Expression Parameters (PEPs). Each PEP defines a facial movement in a specific area of the face. Compared to other existing parameterizations (e.g., MPEG-4), PEPs ensure a similar amount of details, while using a minor number of parameters. The authors linked PEPs with values of P, A and D by conducting an experimental study.

The dimensional approach has the advantage allowing the generation of a large number of emotional facial expressions. However, the dynamic and the temporal characteristics of the expressions are generally not considered. Moreover, the large number of facial expressions poses the problem of the evaluation of all the generated emotional expressions.

Appraisal approach: Other models are based on an appraisal approach [161]. This cognitive psychological approach considers that facial expressions of emotions reflect how an individual appraises and deals with his environment. In this approach, values of appraisal variables (e.g., novelty, intrinsic pleasantness, conduciveness and coping potential) are associated to the activation of action units (smallest units of perceptible facial activity defined in FACS). The final animation that is generated on the virtual character's face is a sequence of several sub-expressions linked to the SECs cognitive evaluations.

However, the appraisal theories are still incomplete regarding the facial action predictions. They are also complex and difficult to implement since the modelling of cognitive capabilities to infer the values of appraisal variables is required.

3.2.1.2 Corpus-based-lexicon of facial expressions

To gather more subtle and natural expressions, other approaches are based on the analysis of annotated corpus of human or virtual faces.

Synthesis of emotional facial expressions from annotated human faces: To collect real data of persons expressing emotions, a first method consists in recording videos of actors having the instructions to express specific emotions. Another method consists in collecting spontaneous expressions by putting people in situations triggering various emotions. For instance, a common method to generate frustration is to simulate a bug in a computer-program-participants have to interact with. The second step is the annotation of the corpus to attribute labels to expressions, and to find out the morphological and dynamic characteristics of the emotion to create the lexicon of emotional facial expressions. Based on an annotated corpus of humans expressing emotions, two approaches to synthesize virtual emotional faces have been explored. The facial expressions can be synthesized at a very low-level by retargeting the points tracked on a human face to a virtual mesh or, at the higher level, using copy-synthesis approach. In the later the virtual character's expressions are synthesized from the manual annotation of the human facial behaviour. The synthesized facial expressions are labelled and stored in the lexicon using low-level animation format such as MPEG4 or FACS. The first approach consists of manually annotating the facial expressions using FACS coding. Then the FACS based manual annotation of each episode is converted into Behaviour Mark-up Language (BML). BML is an XML-like standard script language used to control the behaviour of a virtual character, including the face. The other method uses machine learning algorithm and motion capture data. The 3D points of 27 markers are captured for each frame of the expression and then are retargeted to the virtual mesh using Temporal Restricted Boltzmann Machines [162].

These two approaches offer different degrees of flexibility and control over the expression and different levels of realism and precision of the movements. Motion capture-based animation is usually richer in movements and consequently it may be perceived as more realistic. Also, the motion capture data permits maintaining the temporal and dynamic characteristics of the original expression. At the same time, optical motion capture system is invasive as markers need to be placed on the actors' face and may limit their spontaneous reactions. It is also resource and time consuming. On the other hand, describing animation by sequences of action units allows one to control precisely an animation and its meaning (e.g., by adding or removing AU6, a marker of the Duchenne smile) but has all the weaknesses of procedural approaches to facial animation. The animation is poor in details and the dynamics of the movements is not very realistic.

Facial motion capture records and converts person's movements into a digital database, using different scanners or cameras. From that database we can produce various applications of computer-animated characters. The database consists of various coordinates and relative positions of different reference points on the actor's face. For example, MPEG-4 standard includes standardized feature points which are used for various purposes like expression tracking. These feature animation points are a set of parameters that represent a complete set of facial actions, with motion, head, eye and mouth control as well, i.e. each facial animation point is an action that somehow *deforms* a face model, differing from its neutral state. This tracking can be two-dimensional, and even three-dimensional using multi-camera rigs or laser marker systems.

<u>Human facial expression and skin colour analysis:</u>

Generating 3D human faces is exceedingly complex in computer graphics design [163]. Current technologies of facial animation synthesis are partially able to generate realistic facial expressions containing basic emotions [164]. De Melo and Gratch simulated facial emotional expressions for blushing, sweating, tears of fear, anger, and wrinkles to express the physiological emotions in human [165]. They still relied on unrealistic texture-based method which is a static one. Facial colours of humans manifest the virtual effect, emotional estimate, facial image, remote health care, and individual identification. Indeed, the colours of human face are regarded as one of the most peculiar forms of expressions [166]. Former studies focusing mainly on the geometric features of these alterations considered the facial surface animation including skin stretching and wrinkled structures only. However, the changes in haemoglobin concentration are known to alter the skin colour. It may originate from the reaction of histamine or other skin conditions such as blushing and rashes. Generally, blushing is comprised of joy, shame, and anger. Regardless of its ability to convey emotion, the dynamic changes that occur in skin pigmentation are mostly ignored in the existing skin appearance models [167] [168].

A novel technique for the creation of realistic human facial expressions is introduced by combining the effects of sweat and tears with the skin colour. This approach is capable of creating realistic animation expression by producing sweat and tears. Pulse oximeter and the 3D skin analyser are used to determine the influences of oxygenation on altering skin colour appearance. The methods and expressions for creating animated and realistic fluid sweat and tears based on the particle system and SPH method are emphasized. The facial skin appearance is found to be influenced by the physical and physiological state of it. Extreme appearances of sweating with tears and blushing are generated to accomplish the 3D games and the actual facial animation simulation. Detailed animation and realistic facial animation expression by utilizing the same techniques are applied to create the facial expressions.

The approaches used to develop the computer graphics for facial animation to achieve the realism by generating actual sweating, tears, and blushing are highlighted in the Figure 15 [169].

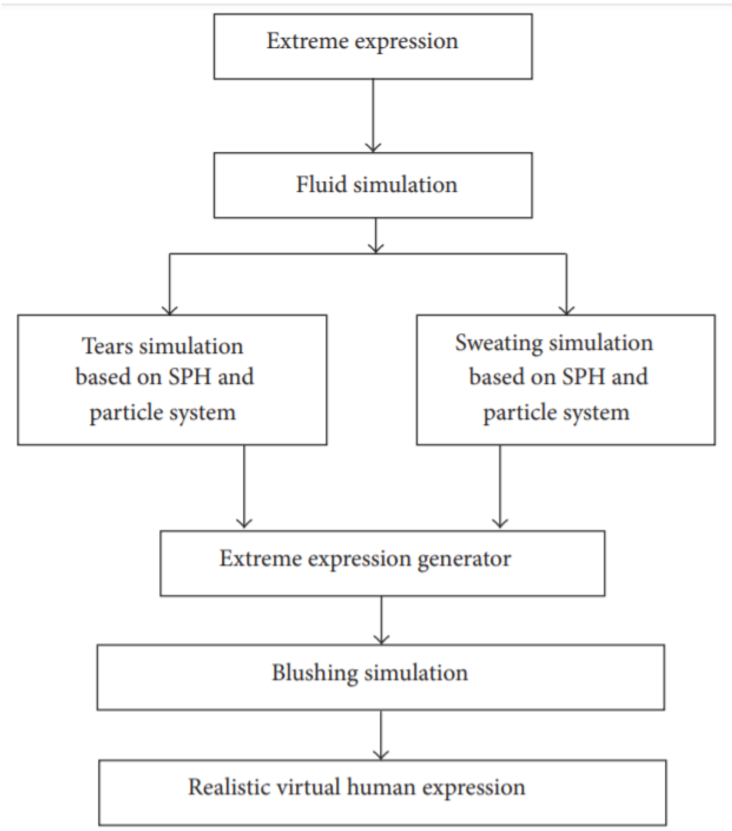

Figure 15 Research methodology.

Extreme Expression: Facial animation alone cannot simulate the extreme emotion of virtual human. Extreme expressions such as scared with sweating and happy until having tears are very strong emotional appearances. Therefore, supplementary elements including fluid mechanism (particle system) and SPH are required to provide these types of features. The facial animation is combined with these elements to perform the simulation.

Fluid Generator: This method creates a fluid simulation which is as realistic as touching tears. The art of creating small drops of rolling fluid requires a particle system which is certainly an efficient technique. The mixture of height area estimates per particle-based liquid generates another unit which is considered as a basic particle system reliable for producing splashes, spray, or water particles. Normally, this is not required to simulate particles system with one another but particles without intraparticle interaction. This simple particle can efficiently be simulated. It needs only a set of N particles $0 \leq i \leq N$ with masses mi, positions xi, velocities Vi, and accumulated external forces f. These particles are usually prepared and placed in substantial positions and velocities before being simulated. Normally, an emitter cuts down the particles that have fixed particle rates on particular velocity in the same direction. It is also ensured that the particles are not only created but also disappear within a specific time interval. They simply fade out once maximum lifetime is attained. In addition, the uses of the lives to paint the particles are properly accounted for. Only per-particle forces exist without any trace of particle-particle interactions and the prevailing equation is decoupled into regular differential equations [170].

SPH Interaction for Particle System: The creation of the particles is more realistic for the simulation of tears or sweating if properties of SPH method including pressure and viscosity are properly implemented. The SPH method being used for stars simulation can be smoothen via smoothing kernels (r). The kernel ($|x - xi|$) is a scalar weight function near the xi position of the particle I possessing exchange symmetry.

Extreme Expression Generator: Generating sweat and tears and their functional dependence on facial mesh. The contribution is based on the creation of extreme expression such as fear and happiness which are common in virtual human emotions. Consequently, the facial animation operation with one of the two techniques which is to simulate actual sweat or tears is highlighted.

Facial Skin Color (Blushing): As mentioned before, the facial skin colours are often used as factors to determine the age or ethnicity. The pulse oximeter system and the 3D skin analyser are used to measure the effects of oxygenation of the facial skin colour appearance.

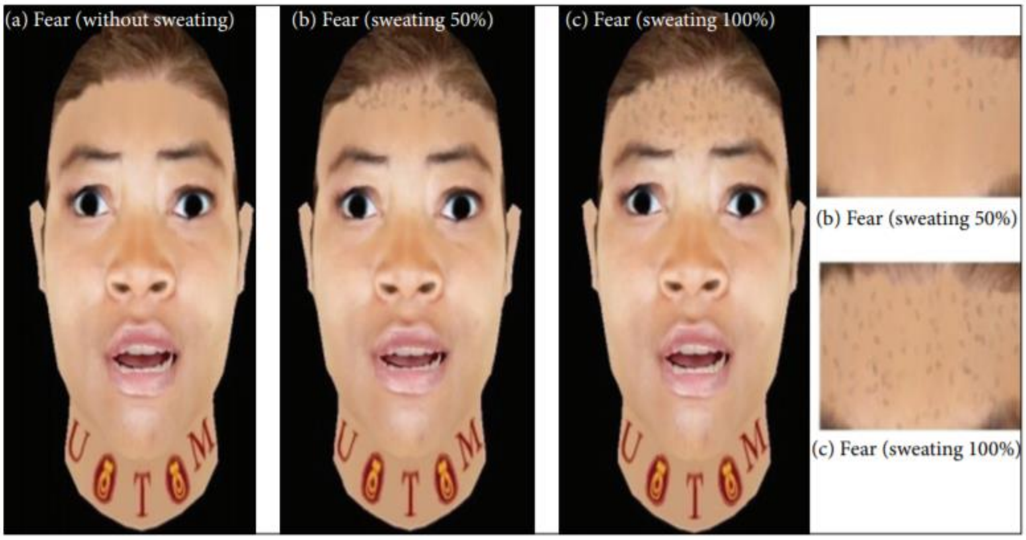

Figure 16 Simulation of sweating due to fear in real-time based on the particle system and SPH method.

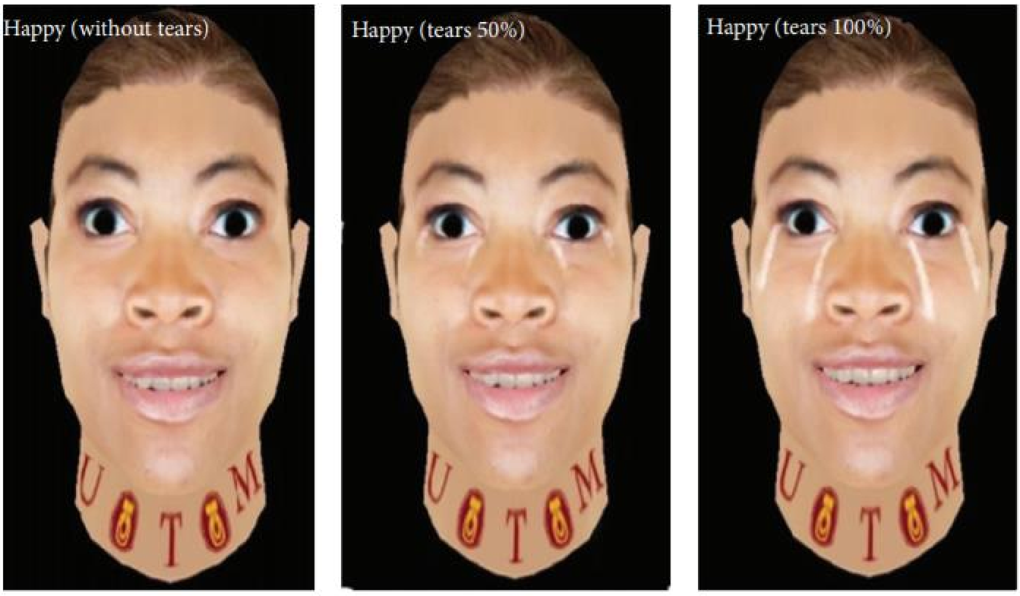

Figure 17 : Simulation of tears due to happiness in real-time based on the particle system and SPH method.

In summary, the emotional facial expressions of virtual characters are generally created with the assumption that virtual characters should display emotions as humans do, i.e. with the same morphological and dynamic characteristics of the face. Most of the models of emotional facial expressions are often based on either empirical or theoretical research in Human and Social Sciences. During human-machine interaction, the same emotional facial expression of a virtual character may have different effect, from positive to negative, on the user's perception depending on the situation in which the emotion is expressed. The display of emotions should be appropriate or plausible in the situation of the interaction. The emotional facial expressions are appropriate if they meet expectations of what one is supposed to feel in a given situation. However, an emotional expression may be inappropriate but plausible when the expression is displayed in a situation even if the expression is not the appropriate one. The recent work of De Melo et al. [165] shows that the user applies "reverse appraisal" to interpret the virtual character's emotional expression and then to deduce information from virtual character's facial expressions regarding for instance its goal conduciveness. An emotional expression may then be displayed depending on the values of the appraisal variables the virtual character wants to convey to the user.

3.3 Police Interrogation

Communication usually consists of a mix between verbal communication and several types of nonverbal behaviours (gestures, body language, tone of voice and facial expression). Sometimes the verbal and nonverbal language conflict with each other, which often results in the nonverbal cues being perceived as being more authentic than the verbal, as these are harder to control [171]. Nonverbal information has been termed as 'leaky' clues, as they can reveal the true feelings or intensions of the speaker [172]. Therefore, it is important for an interviewer to be able to pick up and recognize the nonverbal cues, as this provides insights in the emotional state of an interviewee, and can have big implications that can either have a positive or negative impact as to how well an interview may progress. By restricting these cues during police training we could increase the potential for more effective cognition and internalizing of these tacit interviewing behaviours.

Lie Detector:

Detecting lies is crucial in many areas. One technique to detect lies is through the identification of facial micro-expressions, which are brief, involuntary expressions shown on the face of humans when they are trying to conceal or repress emotions.
Manual measurement of micro-expressions is hard labour, time consuming, and inaccurate, so a Lie Detection System using Facial Micro-Expressions was designed and developed. It is an automated vision system designed and implemented using LabVIEW [173]. An Embedded Vision System (EVS) is used to capture the subject's interview. Then, a LabVIEW

program converts the video into series of frames and processes the frames, each at a time, in four consecutive stages.

- The first two stages deal with colour conversion and filtering.
- The third stage applies geometric-based dynamic templates on each frame to specify key features of the facial structure.
- The fourth stage extracts the needed measurements in order to detect facial micro-expressions to determine whether the subject is lying or not.

Testing results show that this system can be used for interpreting eight facial expressions: happiness, sadness, joy, anger, fear, surprise, disgust, and contempt, and detecting facial micro-expressions. It extracts accurate output that can be employed in other fields of studies such as psychological assessment. The results indicate high precision that allows future development of applications that respond to spontaneous facial expressions in real time.

3.3.1 Dialogue in Police Training with children Interviewing

Police officers when dealing with interviewing children have to cope with a complex set of emotions from a vulnerable witness. Triggers for recognising those emotions and how to build rapport are often the basis of learning exercises. Serious games' interface can provide valuable training because they can restrict emotional complexity and increase focus during the interactions on key factors for emotional recognition [174].

The review of literature reveals that emotion recognition, through facial expressions, can contribute significantly to the perceived quality of communication. For this study an 'emotions map' was created and tested by 41 participants to be used in the development of a targeted interface design to support the different levels of emotion recognition. The emotions identified were validated with a 70% agreement across experts and nonexperts highlighting the innate role of emotion recognition.

Emotions' recognition is a basic social skill and is identified as the ability to recognize the major group of human emotions. Emotions conveyed by facial expressions, gestures, words or situations contain critical information for the regulation of social interactions. Significant research has been conducted during the last three decades within the field of serious games, in the development of systems that attempt to mimic human cognitive processes by automatically analysing and interpreting facial expressions. In order to address the challenges related to this particular field of research, facial expression analysis has been distinguished between two main streams: facial affect analysis and facial muscle motion analysis [61] [175]. Most facial expression analysis systems focus on facial expressions to estimate emotion related activities. In addition, many studies have introduced the interpretation of real-life situations based on the correlation of multiple channels, such as both speech and facial expressions [176]. The majority of serious games developed around this scientific field mainly focus on health studies and affect recognition for individuals with Autism Spectrum Disorder (ASD). The use of virtual humans as a way to teach emotion recognition through games enables the contextualization of emotions, as they simulate real conversation without the actual social interaction that people with ASD find difficult [177].

The Child Interview Simulator (CIS) consists of a serious game developed to enable and complement current police training practices in the field of child interviewing, targeting mainly new recruits. The CIS simulates a real-life situation that allows the player to be a police officer whose goal is to obtain a first statement from a child that has witnessed an alleged criminal offence, and then conduct an interview.

A nine-year old boy is walking on his way back home from school, when he witnesses a man grabbing a lone female from the bushes and attempting to drag her off the common pathway. As the woman screams, the attacker notices the child watching him and in panic, runs away. The gameplay is based on Experimental Learning Theory [178], which uses the learner's experiences in order to facilitate learning. Following Kolb's four stages of learning theory, the player learns about the incident and collects information on the witness through various sources (State 1: Concrete Experience), then reflects on the obtained information (State 2: Reflective Observation), in order to develop his/her own understanding (Stage 3: Abstract Conceptualization), and finally act accordingly by making the correct choices in gameplay (Stage 4: Active Experimentation).

The CIS has two main parts as illustrated in Figure 18:

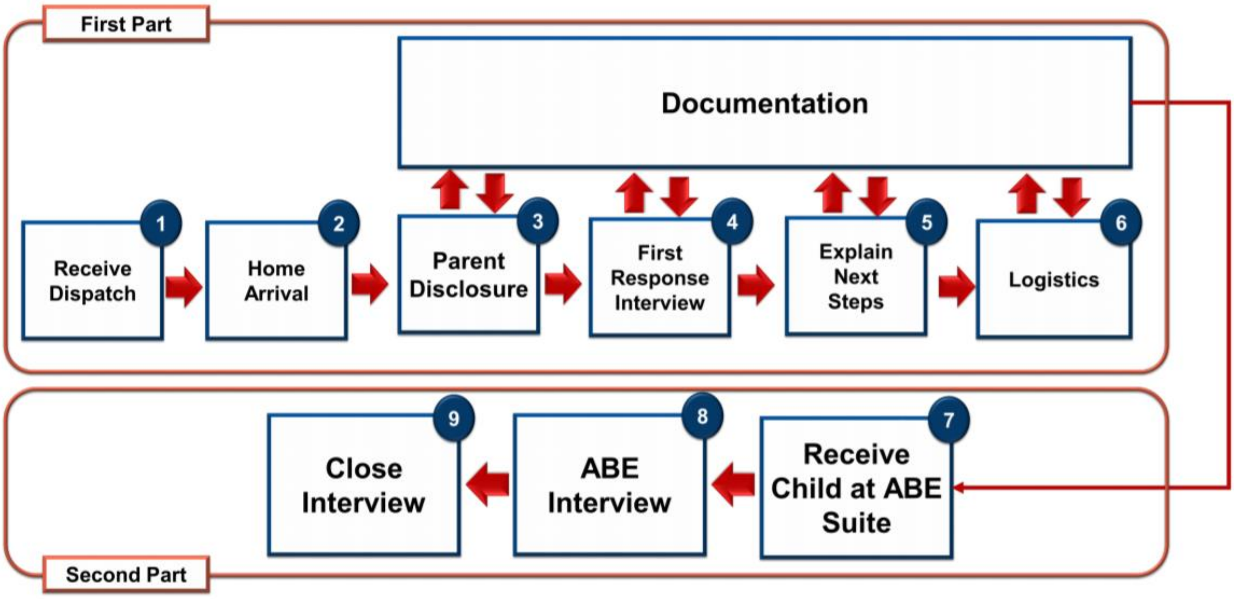

Figure 18 The Child Interview Simulator (CIS) showing the two parts broken down in episodes.

- the first begins when the police trainee receives a dispatch call about the incident and visits the child at his home in order to take a first response statement.
- the second part consists of the police trainee conducting an interview in what is known as Achieving Best Evidence (ABE) suite.

In the first part, the key episodes are the parent disclosure where the police trainee needs to obtain a statement of account from the parent, without the child present, and the first response interview where the police trainee obtains the first account statement of the alleged offence from the child. In the second part, the key episode is documenting the police interview with the child, which is captured on video, and used as evidence in court, thus not requiring the child to be present within the judicial court procedure. In all the episodes, it is necessary for the police trainee to engage with the child in verbal communication. However, the hard challenge to master is the recognition and interpretation of the non-verbal communication cues, which are complex and difficult to implement in the form of a serious game.

The CIS uses the child's facial expression to make explicit the emotional state of the child, and to convey their mood which provides an indication as to how the interview is proceeding.

To ensure that the emotions are recognizable, a set of facial expressions were generated for the child-witness character and these were tested with different stakeholders, ranging from experienced police officers to interviewing experts. The different emotions were organized into an Emotions Map consisting of eight different branches of emotions. The categorization was done based on the intensity of each emotional state. Consequently, the facial expressions marked in dark blue are the most easily recognizable by people (Level 1) and those marked in light blue are less easily recognizable (Level 2). Following this approach, an Emotions Map consisting of eight pairs of emotions (ranging from more to less intense facial expressions) and a neutral one was formed as Figure 19 shows.

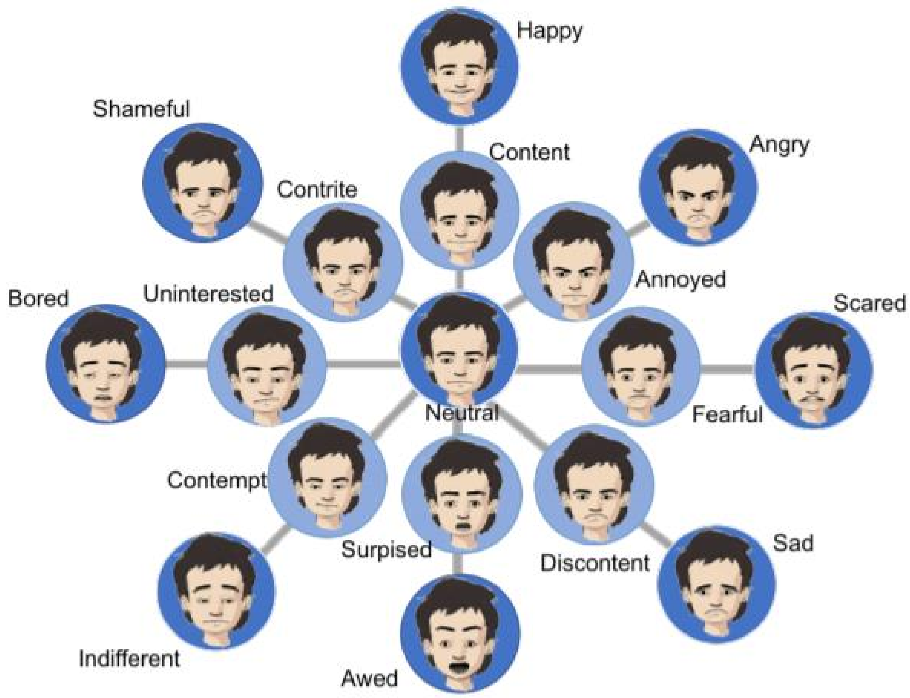

Figure 19. Emotions Map.

3.4 Drivers' Safety

FER is one of the most important factors of advanced driver assistant systems (ADASs), because it can be used to detect driver fatigue and, in conjunction with the rapidly developing intelligent vehicle technologies, assist safe driving.

Although FER has been studied for many years in the computer vision, it still presents many challenges related to the complexity of facial expression; changes in facial pose and illumination conditions; and occlusions and variations between individuals in terms of attributes such as age, gender, ethnic background and personality. To overcome these challenges the research on FER approaches proceeded in the following three research directions:

- The first FER approach consists of action unit (AU)-based methods.
- The second FER approach utilizes feature representation.
- The third approach is the deep neural network (DDNs).

In recent years, the third approach, deep neural networks (DNNs), has emerged as a general approach to machine learning, yielding state-of-the-art results in many computer vision studies that utilized the availability of big data [179]. In addition, improved results have been reported for DNN-based FER methods as compared to conventional FER methods because of their ability to construct discriminative features from learning tasks. In DNN-based FER methods, a variety of versions of DNN have been applied, such as convolutional neural networks (CNNs), long-short term memory (LSTM), generative adversarial networks (GANs) [180] [181] and inception and ResNet modules [182], according to the applications in which they are to be implemented. DNN-based methods recognize FEs by combining detected AUs, rather than using overall facial features for FER [183] [184].

Mira Jeong and Byoung Chul Ko System:

A fast FER algorithm for monitoring a driver's emotions is proposed. It is capable of operating in low specification devices installed in vehicles. For this purpose, a hierarchical weighted random forest (WRF) classifier that is trained based on the similarity of sample data, in order to improve its accuracy, is employed [185]. In the first step, facial landmarks are detected from input images and geometric features are extracted, considering the spatial position between landmarks. These feature vectors are then implemented in the proposed hierarchical WRF classifier to classify facial expressions. The method was evaluated experimentally using three databases, extended Cohn-Kanade database (CK+), MMI and the Keimyung University Facial Expression of Drivers (KMU-FED) database, and its performance was compared with that of state-of-the-art methods.

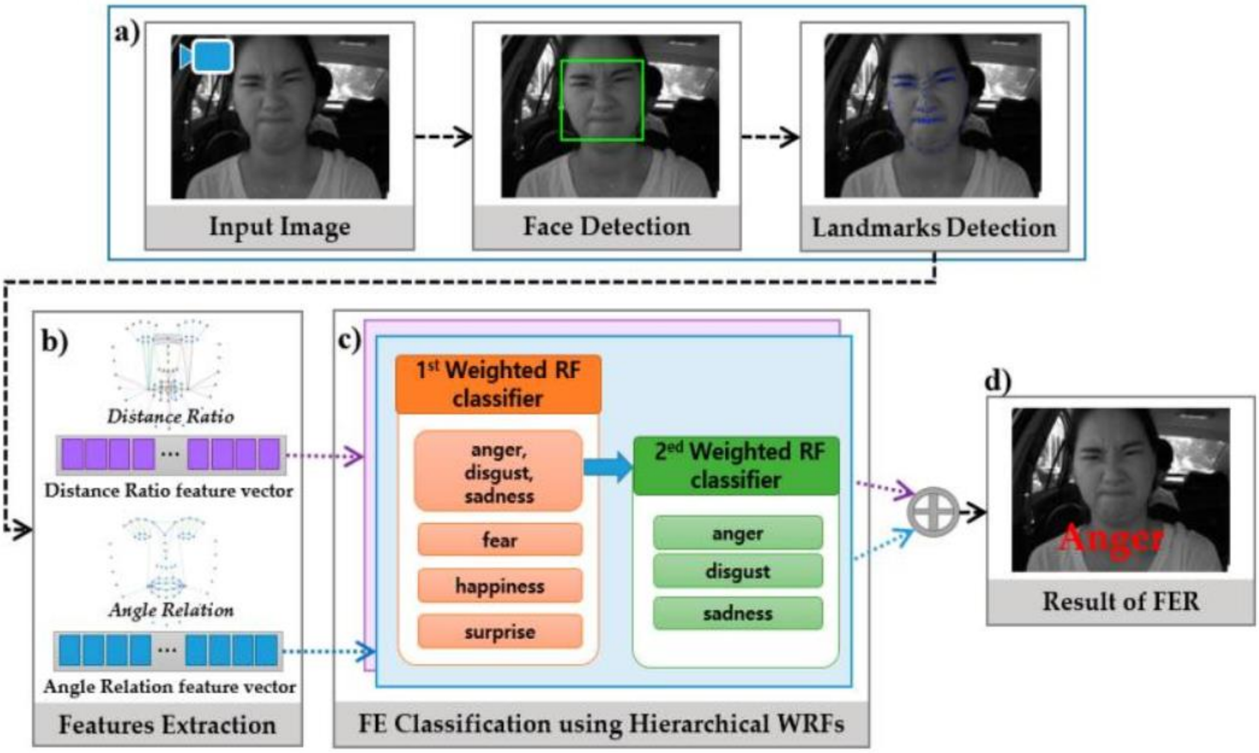

*Figure 20 Overview of the proposed method for facial expression recognition. (a) the face region and facial landmarks are extracted from the image; (b) two geometric features are extracted based on the distance ratio and angle relations; (c) the hierarchical weighted random forest classifies the facial expression (**d**).*

3.5 Marketing

Nowadays, in the marketing area, a new concept of marketing is emerging: the emotional marketing. The emotional marketing studies how to arouse emotions in people to induce them to buy a particular product/service. Recent studies have shown how purchasing choices and decisions are the result of a careful analysis of rational and emotional aspects. Psychological literature recognizes that the emotional conditions influence every stage of decision-making in purchasing processes. Emotions play a key role in any kind of social or business decision. The emotions are manifested in verbal, facial and textual expressions. In the paper techniques of emotional measurements are outlined. These measurements are very important for business goals [186].

Emotions represent another form of language universally spoken and understood. According to some researcher, the emotions are cognitive processes. Emotion is a process, in which the perception of a set of stimuli, allows a cognitive assessment that enables people to label and identify a particular emotional state. The emotional stimuli may be an event, a scene, a facial expression, a poster, an advertising campaign. These events, as a first reaction, put on alert the organism with somatic changes as heart rate, increase of sweat, acceleration of respiratory rhythm, rise of muscle tensions. Emotions give an immediate response that often doesn't use cognitive processes and conscious elaboration and sometimes they have an effect on cognitive aspects as concentration ability, confusion, loss of consciousness, alert and so on. This is what is asserted in evaluation theory, in which cognitive appraisal is the true cause of emotions [187]. Human emotions are deeply joined with the cognition. Emotions are important in social behaviour and to stimulate cognitive processes for strategies making.

The emotional reactions of our brain are measured through a series of techniques, biometric stimuli, which in combination with interpretations of psycholinguistics and cognitive psychology, explain the unconscious reactions of a person. Regarding facial emotions measurements, the feature points of a face are located at eyebrows, eyelids, cheeks, lips, chin and forehead. The first and the most important step in feature detection is to track the position of the eyes. Thereafter, the symmetry property of the face with respect to the eyes is used for tracking the rest of the features. We can also identify all differences and deformations from the "neutral" facial expression with measures on size ratio, distance ratio and orientation. Specifically, we locate and extract the corner points of specific regions of the face, such as the eyes, the mouth and the brows, and compute their variations in size from neutral expression. This information can be converted in datapixel of a higher-level representation of shape, motion, colour, texture and spatial configuration.

Other systems for treatment of facial expressions are based on computational images. The model can contain information on the geometry of the face and facial muscles or on movements of various portions of the face during a change of expression. In some sophisticated models, the patterns of expression are obtained combining together significant portions of the face such as mouth, eyes or eyebrows.

One of the models used for facial expressions and emotions recognition in the field of marketing adopts the Poem descriptor algorithm. Below a brief analysis of the model proposed by Moulay Smail Bouzakraoui, Abdelalim SADIQ and Nourddine Enneya.

POEM descriptor and SVM classifier:

Emotion has become a new trend in marketing, helping enterprise to understand the opinion being expressed on products. Currently the majority of enterprise use conventional marketing methods based on advertising, price, sale points, and satisfaction surveys etc.

By analysing the non-verbal behaviour of customers, many cues can be extracted related to their perception of a product. Therefore, emotional factors are as important as classic functional aspects of products.

Moulay Smail Bouzakraoui, Abdelalim SADIQ and Nourddine Enneya propose a new system allowing to collect a set of video sequences on behaviours of an unlimited number of consumers, and during different periods, to a new product provided by enterpriseas shown in Figure 20 [188]. Then they extract from those videos sequences some representative's images named key frames. These images are those containing only faces. After that they extracted expression features to detect basic emotions (Joy, Sadness, Anger, fear, disgust, and surprise). Finally, they classify detected emotions as expressing a positive, negative or neutral opinion towards a product, and from these results, the enterprise constitutes an opinion on the impact of this product on the customer.

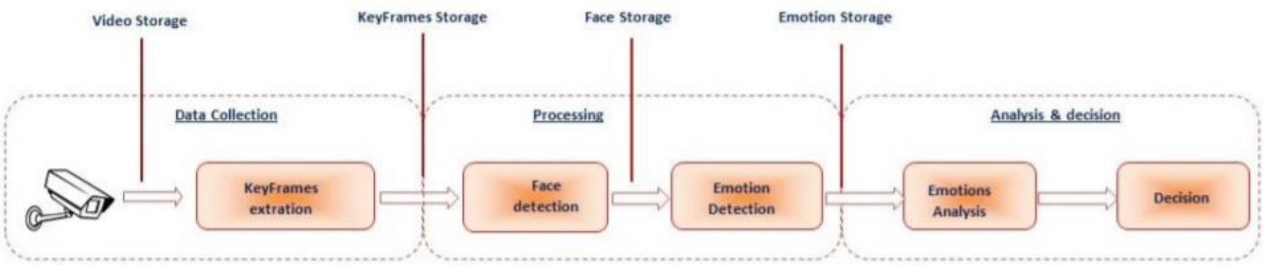

Figure 21 Customer emotions analysis process.

In this system, a facial expression classification algorithm is proposed which uses Haar classifier for face detection, Patterns of Oriented Edge Magnitudes (POEM), histogram of different block sizes of a face image as feature vector. In addition, classifies various facial expressions using Support Vector Machine (SVM). The steps of POEM are represented in Figure 22.

The first step of the model is the collection of data as a video sequence filmed by a camera set up in a well-defined place in a way to record the maximum possible behaviours of customers who frequent the testing product. For each video sequence, a set of representative images named key frame were extracted.

The second phase of the model consists in the processing phase, which is divided into two steps: Face detection and emotions detection. To extract features, the POEM algorithm is used. This descriptor is based on the application of LBP on oriented gradient. In order to calculate the POEM (Patterns of Oriented Edge Magnitudes) for a pixel, the values of the intensity in the calculation of traditional LBPs are replaced by the gradient values, calculated by accumulating a histogram local of gradient on all the pixels of a spatial patch (cellular). In addition, these calculations are carried out in different orientations. The first step is the gradient calculation of the image. At each pixel, the gradient is a 2D vector with its original size and discretized direction. The second step consists to incorporate gradient values of neighbouring pixels, and to calculate a local histogram of the gradient orientations on all pixels of the cell. Firstly, at the pixel position p, a POEM feature is calculated for each discretized direction θi. The final POEM descriptor for each pixel is the concatenation of these unidirectional POEMs at each m orientations.

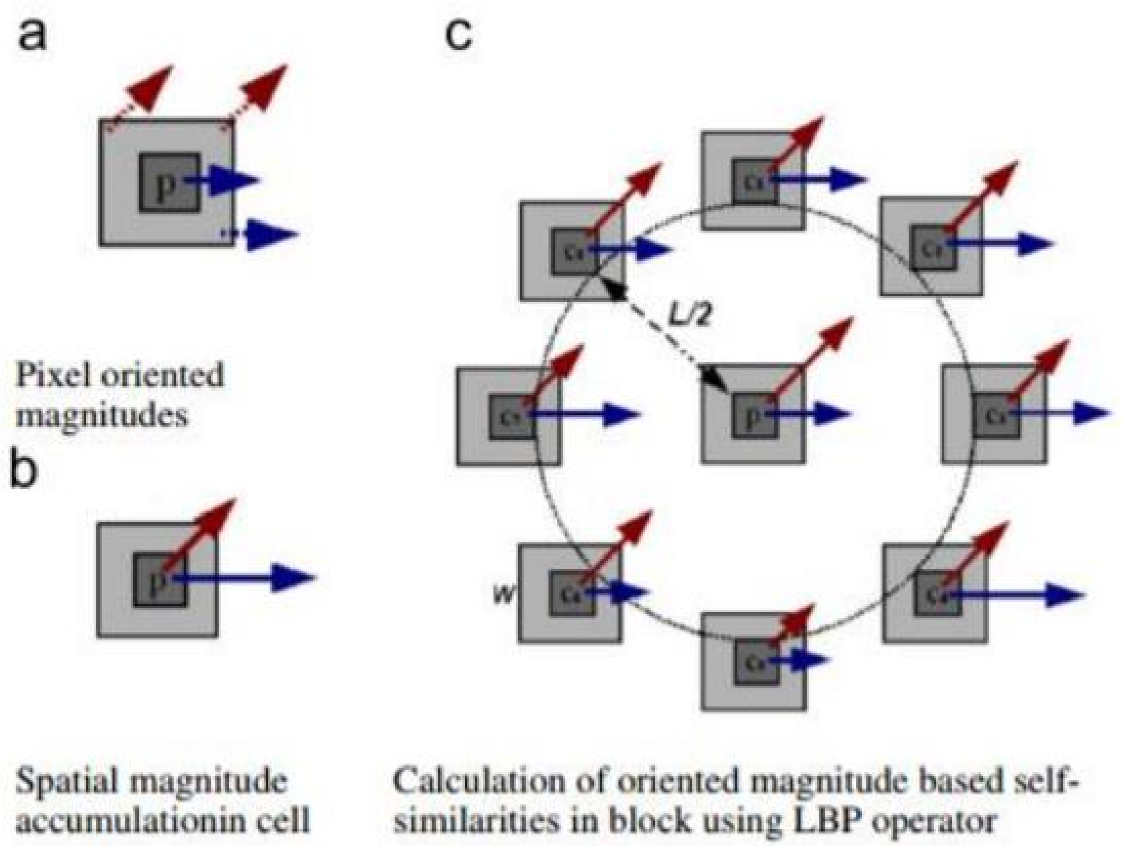

Figure 22 Main steps of POEM feature extraction); Pixel oriented magnitudes, magnitude based self-similarities in block using LBP operator.

Classifying data is one of the major parts in machine learning. The idea of Support Vector Machine (SVM) is creating a hyper plane in dimensional feature space and separate two classes of data with a maximum margin of hyper plane. The separating hyper plane maximizes the distance between two parallel hyper planes [189]. This optimum hyper plane is produced by maximizing minimum margin between two sets. Therefore, the resulting hyper plane will only be depended on border training patterns called support vectors.

The phase of analysis and decision allows having information about the preferences and behaviours of customers for the products. Development teams can quickly determine the needs of consumers with analyses of this information. Marketing teams can evaluate the emotions of consumers and create models to predict sales effectiveness, also optimize the effectiveness of advertising campaigns.

4. Actual Commercial Applications of Facial Expression Recognition (FER)

With the gradual maturity of the technology of face recognition, FER has been widely used in a huge variety of fields.

Developments in Big Data analysis, Cloud Computing, Social Networks and Machine learning have transformed the conventional view of how several problems in Computer Vision can be tackled [190].

FER has produced an attractive number of innovations, such as the advent of the Samsung Smart TVs, which incorporate FER technology by utilizing the built-in camera to provide unprecedented access control in terms of channel surfing and web browsing.

The last decade has been a good one for face recognition technology. Although it has existed since the 1960s, any large-scale attempts to successfully implement it failed due to the lack of precision and scalability. In 2011, both the Pinellas County Sheriff's Office and Panama's Tocumen airport implemented their own facial recognition systems with impressive results: drug smuggling and organized crime in Panama decreased significantly, and the time it took to process criminal investigation for Pinellas County Sheriff's Office went down drastically.

Mainstream applications followed next:

- Facebook has rolled out their face recognition feature that alerts you when untagged pictures and videos of you are uploaded in 2012.
- Google has leveraged the company's substantial experience in AI and machine learning to build Google Photos in 2015.
- Last year, Apple introduced Face ID on iPhone X.

Due to an overabundance of assorted FER technologies, the face has transformed to become an imperative smart object in the Internet of Things [191].

The increase in social networks and the invisibility of the Big Data that accompanies it underline the need to execute the FER to perform a multitude of associated activities such as Face Tagging, Criminal Apprehension, Missing Person Identification and many other disparate domains and disciplines due to the substantial number of facial images that are stored/retrieved by cloud services.

4.1 Face Recognition and Big Data Analysis

Big data is used for the processing and collection of large-scale and complex data. Big data, as a rule, is defined as a combination of large and complex data collections, and used when traditional applications fail to manage databases or to process the data. The main contents of Big Data are represented in Figure 23.

Figure 23 The Contents of Big Data

Big data analytics is in the scope of Data science. Many public and private organizations began to collect the large volume of various data that contain useful information about the problems. For example, use of big data technologies in national intelligence, cyber-security, marketing and medical informatics is of great importance now [192].

Recent advances in various related fields such as social networks and progress in nationalizing law enforcement databanks have led FER [193] and Big Data [194] [195] to converge and help effectively in the upcoming applications such as Face Tagging in social network hubs such as Facebook, as well as being a crucial component in critical applications such as criminal apprehension through the identification of mug shots, the search for missing persons and so on [194] [195].

Through the use of Big Data technologies, Japanese developers began to apply the software to recognize the persons in the black list of buyers of some stores, who were identified as thieves. The police in Tampa, Florida State use Big Data technology called Superbowl XXXV software to identify criminals basing on the scanned images of the face. Universities and airports in the United States also use Big Data in the field of biometric technologies for the identification of people [196].

The main advantage of Big Data Analysis (BDA) is the significant increase in accuracy and performance improvement it can offer. This accuracy increase is due to the fact that BDA relies on the dimension of sample size [194] [195] i.e. all the facial characteristics for comparison (as opposed to about five or six points in conventional mechanisms). The popular BDA technologies include: Apache Hive, Apache Giraph and the Horton works, Hadoop and so on.

Apache Hive big data software facilitates the management of the surveys and large-scale data in distributed memory [197].

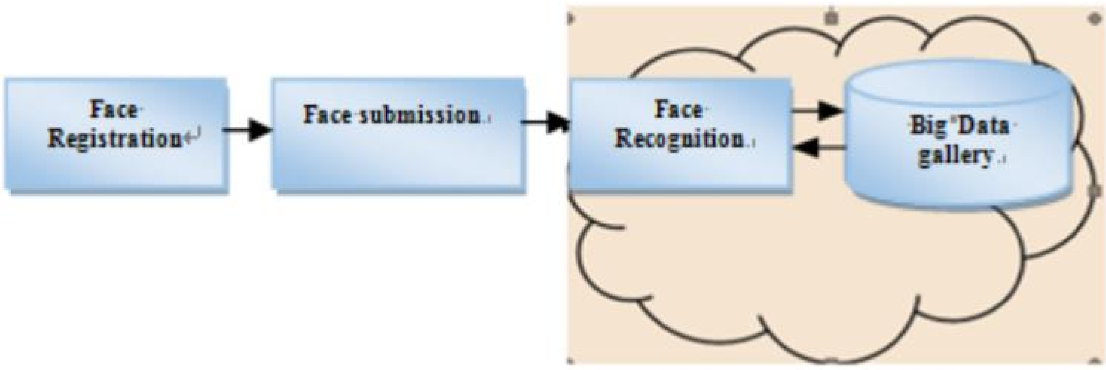

Figure 24 Recognition of Person Through the Use of Cloud and Big Data Technologies.

The figure 24 shows that, human face recognition process is implemented as follows:

- A new human face submitted to the system through the software is recorded, coded and transfer to cloud;
- It is processed by the face recognition algorithm, which is provided in advance; it consists of several stages (facial identification, etc.).
- In the next stage, the face is compared to the other images in the gallery and the recognition process is implemented.

Apache Giraph is an Apache project designed for graphics processing in large-scale data.

Yahoo developed Apache Giraph in 2010, on the basis of a Pregel paper published by Google [195]. Apache Giraph technology is based on the strategies of distributed computing and iterative graph processing system and is decisively reliant on the three critical components of parallel computation and processing: Concurrent computation, communication and barrier synchronization [194] [195].

Facebook also continues to widen its presence in the sphere of BDA by opting for the Apache Giraph methodology in order to proffer a novel social graph search that can effectively scale up to a trillion edges [195].

4.1.1 Biometric and Big Data Technology

Biometric data in the internal affairs agencies, banks and databases of criminal law are quite messy and mainly local. This causes certain difficulties for regional fight against crime. It also restricts the search opportunities in detecting suspects, that is, it ends before implementing the identification process. Wide use of biometric Big Data technologies is able to solve these problems [195] [198]. Biometric technologies, such as authentication, Biocryptography and Cloud-Based Architecture are ideal for security issues. SaaS, Tygart (Figure 25), MXSERVER and other scan be sited as an example of software [195] [199].

Figure 25 Recognition Technology Based Tygart.

Web applications are developed through Software as a service (SaaS) and managed independently [200].

MXSERVER is a system identifying a person based on a powerful server, and used for large-volume videos, photo collections and so on.

Tygart technology-based software is able to group appropriate sections of texts, video, and photos according to the interests of people and delete them [201].

Numerous seminars and conferences have been devoted to Big Data problems in Biometric technologies. Big Data Biometrics Symposium held in Washington in June 2014 was of great interest [195]. The symposium discussed some problems in the field of Biometric technologies, some of which are:

- Developing major biometric programs and improved standards for big data analysis of countries in the future;
- Compact management of biometric data and finding ways to handle them;
- Future contribution of large-scale analysts to big data and the role of cloud technologies in this;
- Biometric big data analysis and achieving the progress in identification and authentication solutions in a mobile environment to solve the tasks;
- Defining the identification prospects on request using Big Data;
- Studying solutions ways of the problems arising from the use of Big Data in biometric technology, and reducing errors, and so on [195] [202].

4.1.2 Biometric, Big Data and Cloud Technologies

Cloud is a new information technology providing single-point access to the distributed resources on request. The features of these technologies include the ability of the users to work independently, and the accessibility of the cloud anytime, through anywhere and any device (smart phone, laptop, tablet, notebook, and etc.) [195]. In addition, the resource sets are introduced to users as a service menu, providing the user to reduce or increase the volume of resources freely, which leads to the widespread use of this technology. The use of cloud in biometric technologies significantly reduces security and confidentiality problems and helps to prevent any unpleasant incidents as soon as possible [195].

MXSERVER ™ server system (Figure 26) processes large-scale video and photo collections using the files obtained through the computers, mobile phones, SIM cards and the video surveillance systems.

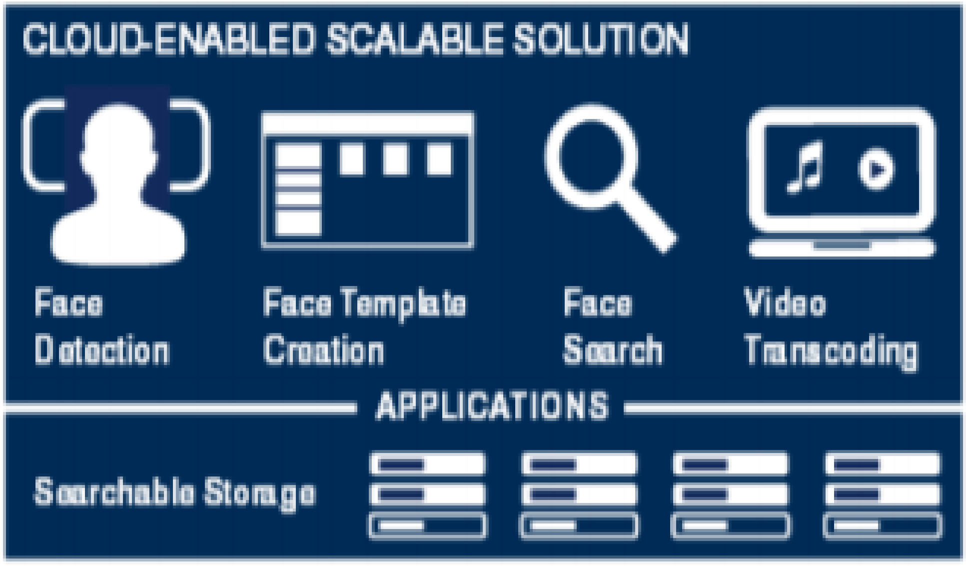

Figure 26 MXSERVER ™ Video and Photo Analysis System.

4.2 Face Recognition and Social Networks

Social networks (SN) have seen significant growth both in the personal and professional sectors [195]. Currently there are about 1.5 billion social users and the social value released according to the McKinsey Global Institute (MGI) is close to $ 1 trillion [195]. A recent MGI study on social technologies in terms of growth, power and intrinsic value has revealed crucial insights into the evolution of social media user behaviour and the various ways it expands in their personal and professional lives. SN has become intertwined in our daily life with elementary tasks such as sharing personal images with collective knowledge exploited through crowdfunding techniques [195].

The main SN hub is Facebook, which incorporates image processing techniques, particularly FER for a variety of features such as tag suggestions for further user interactions. The huge number of available facial images can be used in various ways to provide the user with advanced suggestions in order to improve the user's overall social media experience [195] [203].

4.3 Face Recognition and Machine Learning

Machine Learning (ML) [204] implies the understanding of complex patterns in a progressively abstract way by a system until it is able to cleverly learn new patterns or to recognize existing ones. One of the main ML techniques is the Artificial Neural Networks (ANNs), which is inspired by the mechanics of the human mind (as shown in Figure 27) and uses different parallel strategies such as the simulation of neurons and stratification to learn concepts from experience [205]. The RNA was instrumental in addressing a number of model recognition problems and, although in the past it has had limited success with real-time FR systems, this is changing rapidly with the advent of the Deep Learning methodology (DL) [206].

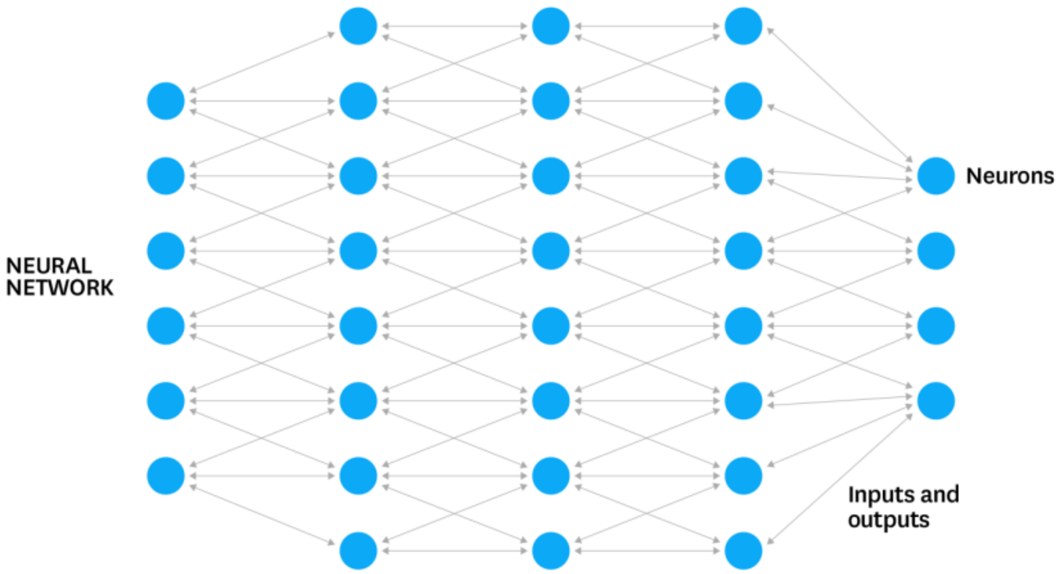

Figure 27 Computational models are inspired by the human brain, where neurons take input ans pass along outputs.

Deep Learning (DL) operates by modelling high-level abstractions based on how the human brain recognizes patterns. Deep neural networks (DNN) use a large number of layers in order to imitate human learning and consequently adjust the strength of the connections between the simulated neurons within the multiple layers in a similar way to the way in which the human brain strengthens understanding a concept [205]. Each layer is able to effectively model an increasingly abstract concept based on the most basic concepts learned at the previous levels. DL is particularly attractive due to its ability to improve efficiency by transmitting more data through its networks. Although these systems have existed for decades, the substantial volume and assortment of data needed to make its understanding to such an extent as to compete with the behaviour of human have only recently become available through the Internet of Things and Big Data.

The new FER development that incorporates Facebook's Deep Face is able to detect and distinguish two faces in a way that can compete with the capabilities of a human being. Deep Face employs Deep Neural Networks (DNN) in order to competently model complex high-level data and multiple functions [205]. Deep Face performs an effective recognition even in the presence of extreme variations in terms of lighting and camera angle and therefore has a potential applicability in new applications of photographic tagging and authentication technologies [205] [206]. Deep Face uses a 3D modelling technique to rotate a single flat image in 3 dimensions so as to allow the algorithms to powerfully visualize the various angles of the face. Using a DNN with over 100 million connections along with a substantial database with millions of facial images, Deep Face tracks the specific features used to recognize human faces and employs this knowledge in order to accurately discover very high-level similarities between facial images. Facebook reports an accuracy of around 97.35% on the Labeled Faces in the Wild (LFW) database and has managed to reduce the error rates of the current avant-garde techniques by more than 27% and is therefore close to human capacity. DNN has the additional advantage, compared to other methods like Support Vector Machines and Linear Discriminant Analysis (LDA), that it can be easily resized in substantially large data sets (which is a prerequisite for some FER problems). This involves a reduction in the costs of the hardware required to perform the processing and is therefore more feasible.

DNN is the ideal candidate for most of the current FER activity in which important characteristics must be detected among a substantial set of variables.

Other milestone systems are [207]:

- The DeepID, or “Deep hidden IDentity features”: a series of systems (e.g. DeepID, DeepID2, etc.), first described by Yi Sun, et al. in their 2014 paper titled “Deep Learning Face Representation from Predicting 10,000 Classes.” Their system was first described much like DeepFace, although was expanded in subsequent publications to support both identification and verification tasks by training via contrastive loss.

- The VGGFace: developed by Omkar Parkhi, et al. from the Visual Geometry Group (VGG) at Oxford and was described in their 2015 paper titled “Deep Face Recognition.”

In addition to a better-tuned model, the focus of their work was on how to collect a very large training dataset and use this to train a very deep CNN model for face recognition that allowed them to achieve then state-of-the-art results on standard datasets.

- FaceNet: a face recognition system developed in 2015 by researchers at Google that achieved then state-of-the-art results on a range of face recognition benchmark datasets. The FaceNet system can be used broadly thanks to multiple third-party open source implementations of the model and the availability of pre-trained models.

4.4 Face Recognition and Cloud Computing

Cloud Computing [208] [209] [210] is a model in which resources such as computational power, storage, network and so on are offered as services through a remote access mechanism. It has several desirable features such as quick elasticity, on-demand self-service, ubiquitous network access and resource pooling [208] [210].

In a FER system that incorporates cloud services, the FER engine is located in the cloud, not in the local processing unit (as in the case of traditional mechanisms) [211] [212] [208]. This attribute makes the system widely accessible, as well as providing the possibility of rapid and reliable integration with other applications [211] [208]. Furthermore, incorporating the cloud facilitates high scalability to ensure that the system can be adapted to a large user base [211] [208]. Furthermore, cloud-based RES systems have the advantage of parallel, real-time processing [208].

In the planning phase, the decision concerning the bifurcation of the components in the modules to be distributed on the cloud and those to be localized locally is crucial.

Several important commercial applications follow the client server model, in which the query face is acquired by the user and transmitted to the cloud server to conduct authentication with the faces of the FER database gallery located on the cloud [212] [208]. The aforementioned design choice requires the need to use security protocols that preserve privacy to protect network traffic between client and server [208] [213].

There are several reliable mechanisms [208] available for the cloud-based FER system such as Animetrics [211], BioID [214] and the solid Face.com technique [215] which acts as the main component in Facebook's FER implementation.

4.5 Real-life Applications of FER Technology

Today, facial recognition technology is used in many industries, from a way to pass the time to matters of national security, from unlocking phones, tagging people on social media to scanning crowds for security threats.

Facial recognition login software and biometric technology are making inroads into building robust security platforms - with a system that’s designed to prevent spoofing by masks or photos [216].

Apple deservedly gets much of the credit for making this technology available to common people in day-to-day life and for elevating the face as the most reliable, natural and secure biometric compared to its predecessors like fingerprint, IRIS sensors and voice recognition.

But, the use of facial recognition goes well beyond unlocking phones. In fact, it is mostly used for security purposes and is increasingly being adopted for a wide range of applications, including:

- Retail: Face recognition is being used to instantly identify when known shoplifters, organized retail criminals or people with a history of fraud enter retail establishments.
- Law Enforcement: Facial recognition is already helping police officers instantly identify individuals in the field from a safe distance. This intel can provide contextual data that tells them whether they need to proceed with caution.
- Advertising: Facial recognition is enabling more targeted advertising by making educated guesses at people’s age and gender. Companies like Tesco are already planning on installing screens at gas stations with facial recognition built in.
- School Safety: Facial recognition surveillance systems can instantly identify when dangerous parents, expelled students, drug dealers or other individuals that pose a threat to school safety enter school grounds.

- Missing Persons: Facial recognition can be used to find missing children and victims of human trafficking. As long as missing individuals are added to a database, law enforcement can become alerted as soon as they are recognized by facial recognition — be it an airport, retail store or other public space.
- Social Media: Facebook uses face recognition technology to automatically recognize when its members appear in photos. This makes it easier for people to find photos of themselves and can suggest when particular people should be tagged in photos.
- Casinos: Facial recognition helps casinos recognize the moment that a known cheater or member of voluntary exclusion lists enter a casino — these players can cost casinos hefty fines if they're caught gambling.
- Financial Services: A more concerning use of facial recognition (for reasons that will be enumerated below) has been in the financial services arena where facial recognition is being used for digital payments, account opening and online account access.

4.5.1 Device authentication (phones, PCs, connected cars)

Accessing information from mobile devices has become mainstream nowadays; besides the clear benefits that mobility provides as a mean to improve efficiency, productivity and user convenience, it in turn does require proper methods for secure access control.

Information access from smartphones and tablets has become mainstream both in business and personal environments over the last years. The use of these devices for accessing services like social networks, email or electronic commerce and banking has surpassed the access from traditional computers [217], turning mobile devices into essential tools in our everyday life.

Passwords have been the usual mechanism for user authentication for many years. However, there are many usability and security concerns that compromise their effectiveness: people use simple passwords, they reuse them on different accounts and services, passwords can be shared and cracked, etc. The amount of different accounts and passwords we deal with these days contributes in making harder the proper usage and maintenance. As a result, we often see news and reports that alert of stolen accounts and passwords [218]. This problem becomes critical in mobile devices, since they can be easily lost or stolen.

There are different biometric modalities that can be integrated in mobile devices: face, speaker, iris, fingerprint, etc. All of them have advantages and disadvantages, but one of the main benefits of face recognition (together with speaker recognition) is that, since smartphones already have integrated cameras, no additional hardware is required.

The most relevant, popular, and realistic use case is security. Buildings, border checkpoints, airports, and seaports all require authentication for access. ATM machines, banking applications, computer and network security, and email logins are some examples where authentication is also absolutely necessary for access. However, unlike the former examples, where there may be police or physical security present, the latter examples have to rely on other security measures. This is important because it makes it uncomplicated to fake one's identity and be easily verified as someone else. When people withdraw money from an ATM, they fear someone may see them entering their PIN number; this is something facial recognition can solve. If facial recognition was installed on ATMs and banking applications, having the credit card (account) number and pin, or login information, would not be enough to get access to the account. Facial recognition could also be beneficial even where physical security was in place. At border checkpoints, it would be more efficient, fast, and accurate for people to be verified through a machine rather than just by a person comparing pictures and names. This could also avoid biases and unlawful profiling. Facial recognition software is also applicable for access and entry into data centres or secured office buildings. This would limit the availability of sensitive data to only the necessary group of predetermined faces. Similar access to encrypted data could be enforced by using faces as private keys to decrypt data.

As phones and mobile devices have become omnipresent, the ability to access all of the owner's accounts and the vast amounts of information stored on the devices has become almost effortless. Android is one of the major operating systems that has been trying to perfect facial recognition technology to lock their devices. By using facial recognition to unlock a device or a specific application on the device, it not only ensures that the correct person is opening the device/application, but also allows for quick access without the need to remember anything, like a password and username [219] [220] [221] [222] [223].

Devices that offer facial recognition

Apple's iPhone X series and Samsung's Galaxy Note 8 and 9 are the most popular devices with facial recognition right now.

As other companies follow suit, it's estimated that over one billion smartphones will use digital 3D face scanning in the next two years.

In addition to Apple and Samsung, popular devices that offer this feature include Motorola Moto G6, OnePlus 6, Oppo Find X, Huawei Honor 7X, and LG G7. These devices scan faces using a 2D scanner, iris scanner, or a combination of the two. You can also download facial recognition applications like FaceLock, which lets you access or restrict other apps using facial recognition.

Samsung was the first to pack advanced facial recognition technologies into a top-tier flagship with iris scanning technology inside the ill-fated Galaxy Note 7. The technology stuck around inside the Galaxy S8 and new Note 8, which forms part of Samsung's security suite alongside a broader face recognition system and fingerprint options.

Samsung's iris scanning technology works by identifying the patterns in your irises. Just like fingerprints, these are unique to each person, making them very difficult to replicate. To do this, Samsung's latest flagships are equipped with an infrared diode that illuminates your eyes regardless of the surrounding lighting conditions. This light wavelength can't be detected by a regular front-facing camera, so a special infrared narrow focus camera then captures the detailed iris information. This image is then stored and processed locally on the device nothing is sent via the internet.

In more recent times, Samsung introduced its Intelligent Scan (see Figure 28) facial recognition system inside the Galaxy S9. Intelligent Scan is also included in the Galaxy Note 9. This promises to be harder to fool than the Galaxy S8's technology, which could be tricked with photographs and contact lenses.

The technology combines iris and facial scanning techniques together to improve accuracy and security. It even works better in low light. The facial scanning part of technology simply builds up a 2D image map of your face, which is common to all Android phones. The key is to combine the infrared iris scanning part with this 2D image to double up on the security layers.

Ultimately, Samsung's technology is only so secure. The company doesn't allow facial scanning biometrics to be used for highly sensitive security concerns like making payments. Instead, these can still only be made using the more secure fingerprint scanner.

Figure 28 Samsung Intelligent Scan

Apple unveiled its new Face ID technology as part of its iPhone X launch, the first 3D face scanning tech in a smartphone. Unlike the basic IR technology mentioned previously, 3D scanning is designed to map out a user's entire face in a highly secure manner. It doesn't just rely on the phone's familiar front-facing camera, there are actually lots of sensors crammed onto that strip at the top.

The iPhone X comes equipped with an array of sensors designed to capture details of your face. For starters, it uses an infrared flood light to illuminate your face, which will work regardless of your surrounding lighting conditions as it's outside of the visible spectrum. A secondary 30,000-point infrared laser matrix is then beamed out, which reflects off the flood light. Rather than snapping a picture of this infrared light, a special infrared camera detects subtle changes in the matrix point reflections as your face makes minute movements, which allows the camera to capture very accurate 3D depth data.

<u>FaceLock (Android)</u>

FaceLock is available in a free and paid version. The free version will lock settings, Google Play store, task manager, and one application of your choice. The reason Wise Orchard, or the application developer, chose to lock settings, Google Play, and task manager is to ensure that no one can go and uninstall the application to bypass the facial recognition authentication. The "pro" version, or the paid version, allows the user to lock any and all of the applications of their choosing, with the mandatory locking of settings, Google Play store, and task manager.

FaceLock Pro also allows the user to unlock the screen using their face. If the face is not recognized, a pin code or password is required to unlock the screen and all applications in both versions. However, it is not necessary to supply both the face and pin, just one or the other. FaceLock mandates that the user takes pictures of him or herself through the app and they suggest taking 10–15 different pictures in different settings and lighting. It is also suggested that one adds a training image whenever they are not recognized right away. This application uses eigenface technology to calculate and compare the training images (Figure 29) [216].

Figure 29 king a training image, FaceLock application download page, Google Play

AppLock (Android)

AppLock is similar to FaceLock, as it allows a user to lock down applications and uses facial recognition software to unlock them. However, unlike FaceLock, the owner of the Android device can choose the security level in which the application is locked down. There are two different levels: convenience mode and truly secure mode. These modes are based on enrollments saved to the application. After installing AppLock, it is necessary to configure the initial enrollment. In doing so, the application captures a picture of the face and the user picks one of three phrases, or a custom phrase of their own, which is recorded. For convenience mode either voice authentication or facial authentication is accepted, and for truly secure mode it is necessary to supply both the spoken phrase and face. Similar to

FaceLock, if facial recognition is not sufficient or the voice is not being recognized, a pre-setup pin can be used instead to get into the applications [216].

True Key (Android and IOS)

True Key, created by Intel Security, uses eigenface technology to authenticate a user. It is available on mobile devices, laptops, and desktops. This application stores passwords and usernames to popular websites and applications, like Facebook, Gmail, Netflix, Starbucks, and many more.

Android systems allow for the application to be present on all login pages of previously installed applications. If there is a username and password stored in True Key for the application, it will push it to the app. IOS systems allow for True Key to be easily accessed from web browsers to copy and paste the username and password into the correct fields [224].

The application can also create safe passwords, store passports, driver's licenses, credit cards, membership cards, and social security numbers. It encrypts this information and scrambles the provided passwords using the AES-256 algorithm. The data is decrypted once the user is authenticated by a recognition method using the face, fingerprint (depending on the device), or master password. The "basic security" level would be access into the application (and therefore to the applications that have stored usernames and passwords on True Key, and sensitive personal information) using only one of the recognition methods, whereas "advanced security" is a combination of two of the three methods. All information stored in the application is stored locally on the device. However, it is possible to set up more than one trust device, and the information is encrypted when being transferred during the sync process [216] [224].

BioID Facial Recognition (IOS)

BioID is available on all iOS versions for iPhones and iPads, but only some Android versions and devices. The application is similar to True Key, where it uses facial biometric authentication to log into other applications and services. BioID uses normal face metric technology, but needs a "liveness factor" to authenticate. The nodding or moving of the head is necessary to authenticate a person. They developers did this to prevent the entry of an adversary by displaying a picture or video [216].

4.5.2 Security sector (airports, hotels…)

The security sector is a veteran when it comes to using facial recognition, in fact face verification has been used at many country borders ever since 2006, when the digitized biometric passport was introduced.

- Police forces are users of face recognition technology and one of the most memorable accomplishments of this technology was in 2016 when the “man in the hat”, responsible for the Brussels terror attacks, was identified thanks to FBI facial recognition software [225].
- Many high-security facilities, such as government buildings and nuclear plants, implement facial verification technology to check the identities of employees.
- Further improvements will lead to facial identification even when the suspect is trying to conceal his or her features by wearing makeup, glasses, beards, or masks. This feature, called ‘hallucinating faces’, is currently being studied T Carnegie Mellon University [225].
- Using facial recognition at an airport check-in kiosk can potentially save time that would otherwise be spent entering a record ID or destination information. The Facial Recognition prototype is designed as a proof-of-concept for using a person’s face as a biometric identifier. Facial recognition can also be useful at other points in travel, such as when a passenger is boarding a plane or checking into a hotel [226].
  The tested use-cases are for kiosk check-in at an airport and for app login on a tablet. U.S.
  Customs and Border Protection (CBP) has been working with airlines to implement biometric face scanners in domestic airports to better streamline security. In fact, they are already in place in certain airports around the country. With the exception of Southwest, most major airlines in the U.S. are taking steps to include the CBP facial recognition technology as part of their security processes.
  In Figure 30 the Biometric scanning technology at Terminal F in Hartsfield-Jackson International Airport in Atlanta

Figure 30 Biometric scanning technology at Terminal F in Hartsfield-Jackson International Airport in Atlanta.

4.5.3 Health

In the last few years, some hospitals started to use facial recognition systems to make patient processing easier. The system matches the right patient to their records, thereby preventing record duplicates or record oversight.

Another area where facial recognition is used is emergency situations.

Either on site of an accident or during the transportation to a hospital, patients are often uncommunicative or unresponsive, which makes it hard to obtain medical information vital to the medical care. Facial recognition offers a fast way to access the medical information and, in many cases, speed up the process of providing necessary medical care.

Facial recognition can also help with screening for certain diseases. For example, researchers at Duke University developed an Autism & Beyond app that uses the iPhone's front camera and facial recognition algorithms to screen children for autism [225].

4.5.4 Retail, Marketing, and Advertising

Facial recognition is still at the beginning in retail and marketing, but there are some interesting trails done by big companies.

A way to utilize face recognition technology is by placing cameras in retail outlets, so that is possible to analyse and improve the customer purchasing process by accessing customer information from their social media profiles and offering customized offers and products [225]. The American department store Saks Fifth Avenue is already using such a system.

Amazon GO stores are reportedly using it as well. As Amazon has proved, when a retailer knows its customers, it can serve them more effectively. But how does that work in a physical, offline environment? One way is to integrate a facial recognition system with smart digital signage, making it possible to target different demographics (through facial analysis) and deliver specific offers that will appeal to them (via connected screens).

An example of this is the International Finance Center Mall in Seoul as shown in Figure 31 which has built facial recognition into its information kiosks. As a customer approaches, the cameras identify the person's age and gender in real-time, personalising interactive advertisements accordingly.

Figure 31 Facial recognition can identify key demographics, enabling connected digital advertising solutions to target different age groups.

Using facial recognition in stores, both online and brick-and-mortar, opens up significant opportunities. Clothing brands and cosmetics producers can use client-specific information to make highly targeted recommendations, sometimes helped by augmented reality. For example, clients could scan their faces to log into their store accounts and also try on clothes in a virtual mirror.

4.5.5 Customer or visitor ID (bank branches, casinos, private companies)

The banking industry is using facial recognition to both prevent fraud and making online banking safer.

HSBC launched a Face ID verification option for their corporate clients in more than 24 countries. The Face ID login is even faster than Touch ID [225].

Voting is another area where fraud constitutes a significant concern. Biometrics could help prevent multiple voting. Some steps have already been taken, allowing people to register and vote through the *Voatz* app that's based on blockchain technology.

4.6 Best Facial Recognition Software

The modern application of face detection technology led people and organizations in seeking out for new and innovative ways to use technology to improve everyday lives. Most common applications today are in retail, cities, university campuses, sports arenas, casinos, airports, among others.

In order to showcase the noteworthy face recognition software provided by renowned companies innovating the market, Analytics Insight has compiled the list of 'Top 10 Best Facial Recognition Software'.

Deep Vision AI

Deep Vision AI is a computer vision software company that excels at facial recognition. It applies proprietary advanced computer vision technology to understand images and video automatically, turning visual content into real-time analytics and valuable insights.

With more than 500M existing cameras worldwide today, Deep Vision AI provides with the ability to analyse camera streams through different AI-based software modules on a single offers plug-and-play platform, enabling users with real-time alerts and faster responsiveness. Its facial recognition software continuously monitors target zones to provide the identification of individuals over time matching against a watchlist of people of interest with highly accurate rates. The software is camera agnostic, connects into any existing camera infrastructure, and can be deployed from the cloud to the edge. Deep Vision AI offers the best performance solution in the market supporting real-time processing at +15 streams per GPU.

Different facial analysis capabilities are applied to business intelligence use cases to help recognize important customers in real-time, quantifies the frequency of visitors, or to improve the overall safety and security. Other highly granular attributes like the count, age and gender of people, is used to understand the changing demographics over time, customer behaviour, and measure marketing efforts. Customers typically combines facial recognition capabilities with advanced vehicle recognition and other AI-based features Deep Vision AI provides.

To comply with the EU General Data Protection Regulation (GDPR) and other international data protection laws, Deep Vision AI applies significant measures to assure a transparent and secure process of the data involved within the provision of its services to customers, partners, and its affiliates, having data privacy and AI ethics at the core of its value proposition.

Principal markets include cities and public venues, large retail stores, energy & infrastructure, public transportation, school campuses, gaming venues and event arenas. Deep Vision AI is a certified partner for NVIDIA’s Metropolis, Dell Digital Cities, Amazon AWS, Microsoft, Red Hat, and others.

SenseTime

As a leading technology platform developer, SenseTime is dedicated to creating industry solutions via innovations in AI and big data analysis. SenseTime’s multifunctional technology is rapidly expanding and already encompasses facial recognition, image recognition, intelligent video analytics, autonomous driving, and medical image recognition.

SenseTime has serviced over 400 well- known companies and government agencies including Honda, Qualcomm, China Mobile, UnionPay, Huawei, Xiaomi, OPPO, Vivo, and Weibo.

Among its platform software:

• *SensePortrait-S* is a Static Face Recognition Server, that provides functions of face detection from an image source, as well as feature extraction, attribute analysis, attribute comparison, and target retrieval from a vast facial image database

• *SensePortrait D* is Dynamic Face Recognition Server that provides functions of face detection in multiple surveillance video streams, as well as face tracking, feature extraction, and comparison

• *SenseFace* is a Face Recognition Surveillance Platform. Based on Face Recognition technology powered by a deep learning algorithm. It provides integrated solutions of intelligent video analysis, which functions in target surveillance, trajectory analysis, population management, and relevant data analysis, etc

Amazon Rekognition

Rekognition is a cloud-based Software as a service computer vision platform by Amazon, which makes it easy to add image and video analysis to applications using proven, highly scalable, deep learning technology that requires no machine learning expertise to use. With this platform, one can identify objects, people, text, scenes, and activities in images and videos, as well as detect any inappropriate content. It also provides highly accurate facial analysis and facial search capabilities that can be used to detect, analyse, and compare faces for a wide variety of user verification, people counting, and public safety use cases.

With Amazon Rekognition Custom Labels, organizations can identify the objects and scenes in images that are specific to their business needs. For example, they can build a model to classify specific machine parts on their assembly line or to detect unhealthy plants. Amazon's Custom Labels takes care of the heavy lifting of model development for organizations, so no machine learning experience is required. People simply need to supply images of objects or scenes they want to identify, and the service handles the rest.

FaceFirst

The FaceFirst software helps create safer communities and more secure transactions while providing great customer experiences. Its computer vision platform is used by companies for face recognition and automated video analytics to help retailers, event venues, transportation centres, and other organizations prevent crime and improve customer engagement. FaceFirst software is accurate, fast, scalable, secure and private and it also offers plug-and-play solutions for physical security, identity authentication, access control and visitor analytics, and robust SDK/APIs for easy integration into any system.

As a leading provider of effective facial recognition systems, it benefits to retail, transportation, event security, casinos, and other industry and public spaces. The FaceFirst platform is designed to be scalable, fast and accurate while maintaining the highest levels of security and privacy. FaceFirst leverages artificial intelligence to integrate with surveillance systems to prevent theft, fraud, and violence. FaceFirst is proudly designed, engineered and supported in the USA.

Trueface

TrueFace is an industry-leading computer vision model that helps people make sense of their camera data and turn it into actionable information. It offers only on-premise computer vision solutions which enhances data security and performance speeds for its partners. The platform-agnostic solutions are trained specifically for each deployment and work in a variety of ecosystems. It places the utmost priority on the diversity of training data ensuring equal performance for all ethnicities and genders.

Trueface has developed a suite of SDK's and a dockerized container solution that harnesses the powers of machine learning and artificial intelligence to transform the camera data into actionable intelligence. Facial recognition, weapon detection, and age verification technologies are all easily deployable on organizations' infrastructure, creating safer and smarter environments for their customers, employees, guests and more.

Face++

It is an AI Open Platform which offers computer vision technologies that enable people's applications to read and understand the world better. It allows people to easily add leading, deep learning-based image analysis recognition technologies into their applications, with simple and powerful APIs and SDKs.

Face++, a platform enabled by Chinese company Megvii, uses AI and machine vision in a variety of amazing ways to detect faces, analyse 106 data points on the face, and confirm a person's identity with a high degree of accuracy. Face++ also lets any developer create apps using its algorithms, which has helped make it the most extensive facial recognition platform in the world with 300,000 developers from 150 countries using it.

Significantly, Face++ is already integrated into Alibaba's City Brain platform that analyses the CCTV network in cities to optimize traffic flows and observe incidents that require police or medical attention.

Kairos

It provides state-of-the-art, ethical face recognition to developers and businesses across the globe. People can integrate Face Recognition via Kairos cloud API, or host Kairos on their servers for ultimate control of data, security, and privacy. They can create safer, more accessible customer experiences.

Kairos Face Recognition On-Premises control the data privacy and security 100 percent while keeping critical data in-house and away from third-parties/hackers. It also speeds up the face recognition enabled products that reduce latency risks associated with running on public cloud deployment.

Kairos is growing new markets with confidence along with its ultra-scalable architecture which means people can search 10 million faces at approximately the same time as 1 face.

Cognitec

Cognitec's FaceVACS Engine enables clients worldwide to develop new face recognition applications. It provides a clear and logical API for easy integration in other software programs. Cognitec provides the FaceVACS Engine through customized software development kits, with a set of functions and modules specific to each use case and computing platform, and based on tailored software licensing agreements. Such specific use cases include image quality check, verification for document issuance, and verification for access control.

Its features are:

- Powerful face localization and face tracking on images and video streams
- Industry-leading matching algorithms for enrolment, verification, and identification
- Accurate portrait characteristics check for gender, age, pose deviation, exposure, glasses, eyes closed, uniform lighting detection, unnatural colour, image and face geometry
- ISO 19794-5 full-frontal image type checks and formatting as required for ePassports
- Supports multiple algorithms to work with two-dimensional intensity data, or two-dimensional data and corresponding range data (3D data)

5. Facial Expression Recognition (FER) Market

5.1 Face Recognition Industry Overview

The global market for facial recognition is a growing market, estimated to reach USD 11.30 Billion by 2026, according to a new report by Reports and Data as shown in Figure 32 [227].

Due to the benefits it provides over conventional security methods, such as biometrics, facial recognition has gained prominence in recent times. Governments around the world have invested substantial resources in facial recognition technology, including leading adoptions from the United States and China. Technically advanced facial recognition technologies with mobile security apps and drones are likely to create more opportunities in the future.

The global face recognition market has been studied to provide some insights on demand forecasts, market trends and micro and macro indicators [228].

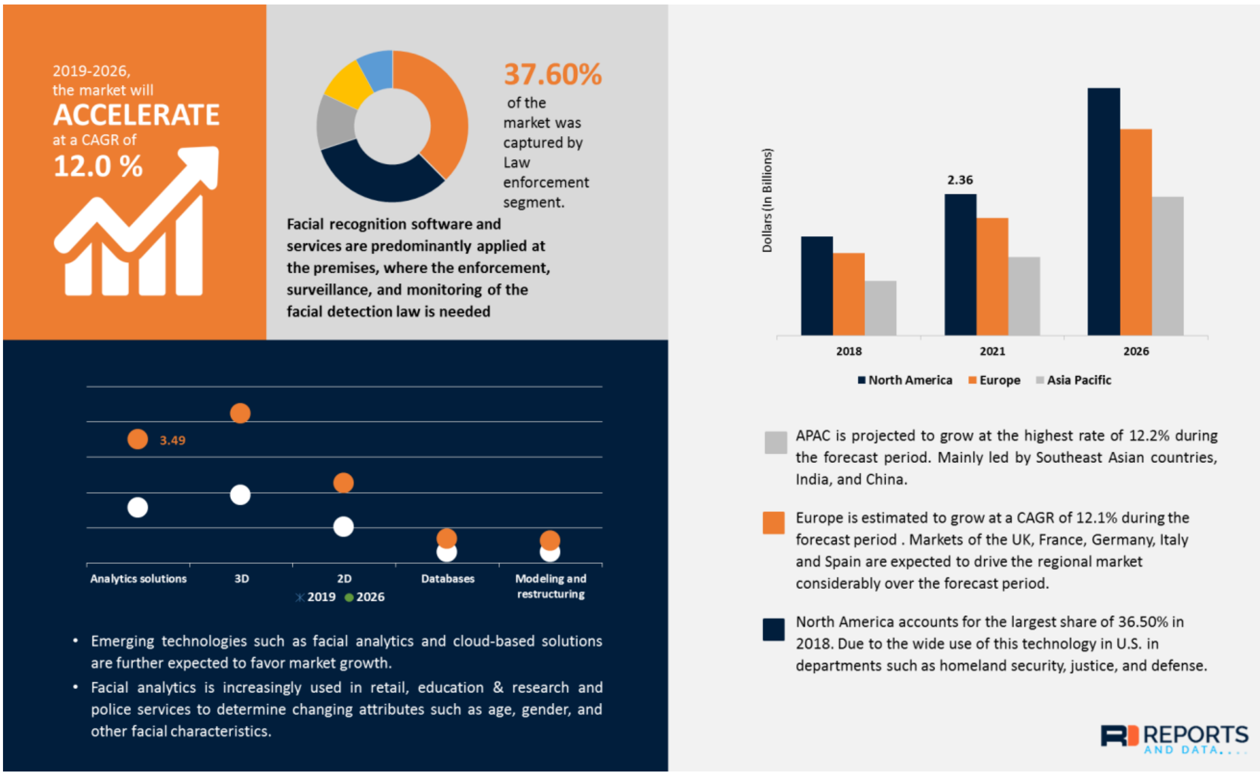

Figure 32 Attractive market opportunities in the Facial Recognition Market.

Due to the presence of key players, favourable reimbursement policies, and rising use of technology in the U.S along with several initiatives initiated by the government, North America accounts for the largest share of 36.5% of the market in 2018.

Technology improvements such as 3D recognition systems and cloud-based services are also anticipated to propel the overall market demand over the next couple of years.

5.2 Facial Recognition Market Segmentation

The Facial Recognition Market is divided on the basis of component, application and geography (see Figure 33).

The component/technology segment is divided into hardware (scanners, cameras, handheld devices, integrated devices), software (which includes 2D, 3D, and facial analytics) and services.

Stationary 3D facial recognition dominated the overall facial recognition market which occupied a 60% share in 2019 and is expected to maintain its position over the forecast period 2019-2025. 3D face recognition systems are capable of precisely recognizing humans faced with variant facial expressions and positions and under dim lights. Using an axis of measurement and depth that is not affected by illumination, 3D face recognition is used in darkness and possesses a capability to identify a subject at various view angles with the prospective to identify up to 90 degrees of a face in profile [229].

The application segment includes homeland security, criminal investigation, ID management, physical security, intelligent signage, web application, business intelligence, photo indexing & sorting, and others (VIP recognition, automotive and phone, PC & banking login).

SEGMENTATION	DETAILS
By Component	· Hardware · Software - 2D Facial Recognition - 3D Facial Recognition - Facial Analytics Recognition · Services - Training and Consulting Services - Cloud-based Facial Recognition Services
By Applications	· Civil Applications and Law enforcement · Security applications for electronic transactions and access control · Ambient Intelligence · Wearable systems · Automotive · Entertainment
By Vertical	· Banking, Financial Services, and Insurance · IT and Telecom · Government · Healthcare · Education · Travel and Hospitality · Defense · Retail and e-commerce · others
By Geography	· North America (the USA and Canada) · Europe (UK, Germany, France, Italy, Spain, and Rest of Europe) · Asia Pacific (Japan, China, India, Australia, Southeast Asia and Rest of Asia Pacific) · Latin America (Brazil, Mexico and Rest of Latin America) · Middle East & Africa (South Africa, GCC and Rest of the Middle East & Africa)

Figure 33 Face Recognition Market Segmentation

Based on geography (see Figure 34), North America provides profitable market growth possibilities, showing huge demand for facial recognition technology for human data security and criminal investigation. The FBI operates the region's largest facial recognition surveillance system. The ID scheme of the FBI holds a database of facial recognition with pictures of millions of Americans. On average, the FBI performs monthly searches to identify people with facial identification systems. Due to technological advances, North America is anticipated to have the largest market share in the face recognition industry and also houses most of the region's top fortune 500 businesses and tech giants. The US contributes a significant proportion of the national market share of North America.

The United States Department of State operates one of the largest facial recognition systems in the world using a database of 117 million American adults, with photos taken mainly from photos of a driver's license [229]. Although not yet completed, the database is used in some cities to provide clues as to who he was in the photo. The FBI uses photos as an investigative tool, not for positive identification [230]. As of 2016, facial recognition was used to identify people in photos taken by police in San Diego and Los Angeles [231] and use was expected in West Virginia and Dallas [232].

In recent years Maryland has used facial recognition by comparing people's faces with photos of driving licenses. The system sparked controversy when it was used in Baltimore to arrest rebel protesters after Freddie Gray died in police custody [233]. Many other states are using or developing a similar system; however, some states have laws prohibiting their use.

The FBI has also set up its new generation identification program to include facial recognition, as well as more traditional biometrics such as fingerprints and iris scans, which can be extracted from criminal and civilian databases [234]. The Federal Office of General Responsibility criticized the FBI for failing to address various privacy and accuracy concerns.

In 2019, researchers reported that Immigration and Customs Enforcement uses facial recognition software against state driver's license databases, even for some states that license undocumented immigrants.

Europe has appeared as an intelligent solution for addressing current identification requirements along with identity claims verification. The German Ministry of the Interior Security intends to install video surveillance systems of facial recognition throughout the airports of country as well as train stations to identify the individuals that may pose safety hazards- Few key countries like UK, Germany and France are contributing to the technological development and application of facial recognition market in Europe.

Police forces in the United Kingdom have been trialling live facial recognition technology at public events since 2015. However, a recent report and investigation by Big Brother Watch found that these systems were up to 98% inaccurate [235].

In May 2017, a man was arrested using an automatic facial recognition (AFR) system mounted on a van operated by the South Wales Police. Ars Technica reported that "this appears to be the first time [AFR] has led to an arrest" [236].

The Australian Border Force and New Zealand Customs Service have set up an automated border processing system called SmartGate that uses face recognition, which compares the face of the traveller with the data in the e-passport microchip [237] [238]. All Canadian international airports use facial recognition as part of the Primary Inspection Kiosk program that compares a traveler face to their photo stored on the ePassport. This program first came to Vancouver International Airport in early 2017 and was rolled up to all remaining international airports in 2018-2019 [239]. The Tocumen International Airport in Panama operates an airport-wide surveillance system using hundreds of live face recognition cameras to identify wanted individuals passing through the airport [240].

During the forecast period, Asia Pacific is anticipated to grow at a remarkable CAGR. Huge investments of the government sector in safety and surveillance infrastructure enhanced public consciousness, and the advent of complex analytics-backed techniques are some if the variables driving this region’s market growth.

As of late 2017, China has deployed facial recognition and artificial intelligence technology in Xinjiang. Reporters visiting the region found surveillance cameras installed every hundred meters or so in several cities, as well as facial recognition checkpoints at areas like gas stations, shopping centers, and mosque entrances [241] [242].

Like China, but a year earlier, The Netherlands has deployed facial recognition and artificial intelligence technology since 2016 [243]. The database of the Dutch police currently contains over 2.2 million pictures of 1.3 million Dutch citizens. This accounts for about 8%

of the population. Hundreds of cameras have been deployed in the city of Amsterdam alone [244].

The Middle East and Africa are anticipated that the market of biometrics-as-a-service will increase in the Middle East and Africa over the forecast period with a significant amount of investment by the key players [245].

Figure 34 Facial Recognition Market Expected to Grow Rate by Region (2019-2024)

The global facial recognition market has strong competition among the well-established and new emerging players. These market players target to gain a competitive advantage over the other players by participating in partnerships, mergers, and acquisitions and expanding their businesses.

The main players in the sector are strongly focusing on innovation in production technologies to improve efficiency and durability. The best long-term growth opportunities for this sector can be seized by ensuring continuous process improvements and financial flexibility to invest in optimal strategies.

Major vendors in the global market include [246]:

- NEC (Japan);
- Aware (US);
- Gemalto (Netherlands);
- Ayonix Face Technologies (Japan);
- Cognitec Systems GmbH (Germany);
- NVISO SA (Switzerland);
- Daon (US);
- StereoVision Imaging (US);
- Techno Brain (Kenya);
- Neurotechnology (Lithuania);
- Innovatrics (Slovakia);
- id3 Technologies (France);
- DEMIA (France);
- Animetrics (US);
- MEGVII (China).

5.3 Technology Readiness and Availability

According to the Horizon 2020 EU Research and Innovation Framework Programme the readiness and availability of a given technology is assessed using nine different levels (Technology Readiness Levels, TRL):

- TRL 1: basic principles observed,

- TRL 2: technology concept formulated,

- TRL 3: experimental proof of concept,

- TRL 4: technology validated in laboratory,

- TRL 5: technology validated in relevant environment (industrially relevant environment in the case of key enabling technologies),

- TRL 6: technology demonstrated in relevant environment (industrially relevant environment in the case of key enabling technologies),

- TRL 7: system prototype demonstration in operational environment,

- TRL 8: system complete and qualified,

- TRL 9: actual system proven in operational environment.

Face technology has reached TRL 9, with multiple large-scale systems already deployed and working worldwide. Despites that, each operational scenario has its own specificities: the successful application of a certain technology to a given specific use-case and environment, does not necessarily guarantee the same level of success when those operational conditions are changed. In particular, for facial recognition technology to achieve the expected level of performance, there are parameters that have to be considered, such as (probably, the most important) the expected accuracy, which depends on the data (i.e. facial images and photographs) a system will have to deal with and with the quality of that data.

5.4 The Face Recognition Market Trends

The facial recognition market holds a substantial scope for growth on the global horizon. The market, which is in its growth stage however, is expected to contribute significantly to the global market within the next 10 years. Global facial recognition market share is shown in figure 35.

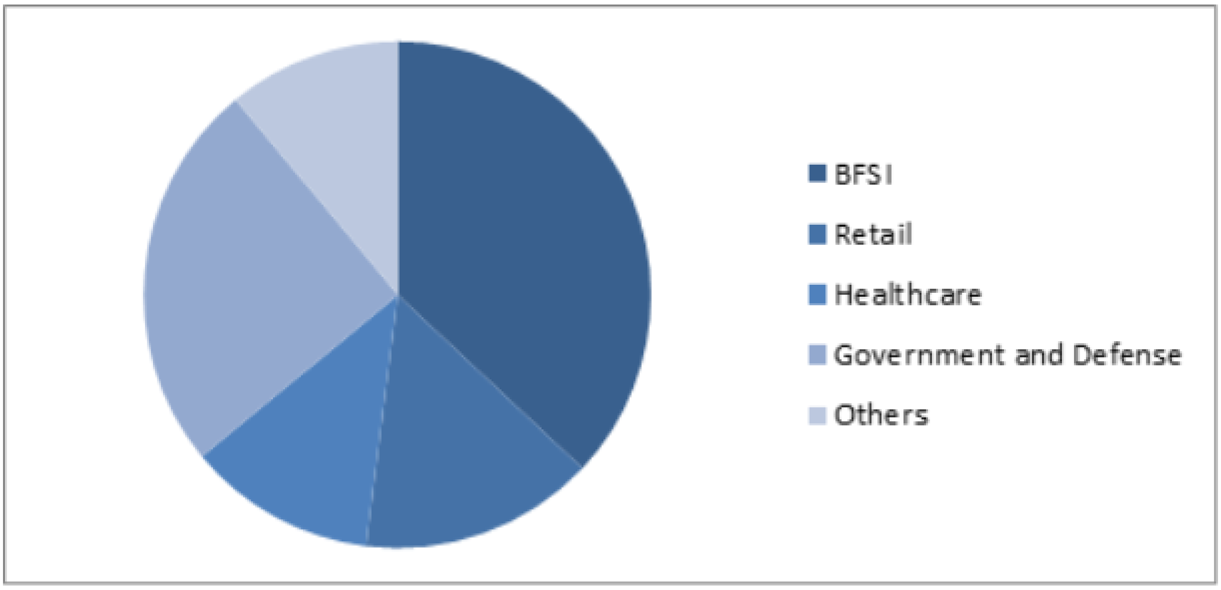

Figure 35 Global Facial Recognition Market Share, 2018 (%)

Based on the end-users the BFSI (Banking, financial services and insurance) segment is projected to hold a significant share in the global facial recognition market. Facial recognition is one of the numerous ways with which the banks can increase efficiency and accessibility. Many tech giants such as Apple and Microsoft are offering their services to the banking and financial industry. For instance, US-based banks such as Chase, HSBC, and the United Services Automobile Association (USAA) are using FaceID offered by Apple that allows customers to securely log into their mobile banking apps. In addition, UK-based Lloyds Bank is testing similar features using the biometric authentication technology of Microsoft.

The growing data breaches in the financial sector are anticipated to increase the use of facial recognition solutions in the financial sector.

Facial recognition ensures high-level security for the public and private sectors. The increase in terrorist attacks on government organizations in recent years creates the need for companies and governments to implement biometric facial recognition. In addition, the growing application of physical security and the intelligent progress of signage and technology, such as cloud-based services and 3D-based recognition systems, have increased the use of facial recognition. However, the increased sensitivity to shared personal data, the lack of accuracy and the high implementation costs do not favour the expansion of the market in some regions. Technically advanced face recognition systems with mobile security application and drones could create more opportunities in the future. Growing incidences of terrorist attacks and security concerns are expected to boost the market growth [247].

The main market factor that favoured the growth of face recognition technologies was the growing demand for advanced security and surveillance systems in civil and government agencies to improve security. The number of data breach cases has increased exponentially from 400 to 600 breaches in the period 2011-2015, due to the growing generation of digital content and the lack of data security laws [247]. Data breach cases and brute force attacks have increased the demand for advanced surveillance systems, which in turn has increased the demand for facial recognition solutions.

Currently, the negative aspects of the developed technologies are the high implementation costs and the low precision. Costs include maintenance and middleware costs which also contribute to implementation costs. However, few manufacturers, such as FaceFirst, Inc., have started using efficient algorithms, such as PCA, FFT, to improve the accuracy and reduce the costs of facial recognition technology. Therefore, with the development of the technology, the lack of accuracy and the high implementation costs of the facial recognition technology are likely to reduce their impact during the forecast period.

The growing demand for facial recognition in smart devices, such as smartphones, laptops, tablets and personal digital assistants, used for both personal and commercial purposes, presents various growth opportunities for the facial recognition market. In addition, the growing number of drones in various commercial sectors, such as media and entertainment, inspection and detection, contributes to the demand for facial recognition [247].

Technological advances will lead to lower prices for facial recognition systems in the future. The software development kit (SDK) has improved accuracy in terms of recognizing facial features. Therefore, better product quality increases awareness among users, which in turn will increase the adoption of biometrics for facial recognition in the future.

5.5 Controversies of Face Recognition

Facial recognition software has become increasingly popular in recent years. While there are some potential benefits to using technology to prevent and resolve crime, there are many concerns regarding privacy, security and legislation relating to the use of technology.

Facial recognition technology uses biometrics to map facial features and help verify identity through key facial features and uses a database of photos, such as mug shots and driving license photos to identify people in the photos and security videos.

The facial recognition market is growing exponentially and the most significant uses for technology are surveillance and marketing. This, however, raises concerns for many people [248]. The main reason for citizens' concerns is the lack of federal regulations regarding the use of facial recognition technology. Many are concerned about the accuracy of the technology and whether there are prejudices and misinformation in these technologies. One problem, for example, is that technology has shown in numerous studies that it is inaccurate in identifying black people, especially black women [248].

Another major concern is the use of facial recognition for law enforcement purposes. Today, many police departments in the United States, including New York, Chicago, Detroit and Orlando, have started using technology. According to a May 2018 report, the FBI has access to 412 million facial images for research. Currently, the Chinese government is already using facial recognition to arrest jaywalker and other petty crimes that provoke debate between what are considered fundamental civil rights and privacy issues with respect to protecting the public. Accuracy and accountability are needed when it comes to the use of technology, particularly as regards the judicial system [248].

5.5.1 Privacy Violation

The problem of compromising privacy due to surveillance technologies is voiced by civil rights organizations and privacy activists such as Electronic Frontier Foundation [249], Big Brother Watch and ACLU [250].

Some fear that this could lead to a "total surveillance company", with the government and other authorities having the ability to know where they are and the activities of all citizens around the clock.

Facial recognition can be used not only to identify an individual, but also to discover other personal data associated with an individual - such as other photos showing the individual, blog posts, social network profiles, Internet behaviour, travel patterns, etc. - through only facial features [251]. Concerns were raised about who would have access to knowledge of their position and people with them at any time [252]. In addition, individuals have a limited ability to avoid or counteract the detection of facial recognition unless they hide their faces. The dynamic of daily privacy changes allowing any marketer, government agency or random stranger to secretly collect the identities and associated personal information of any individual captured by the facial recognition system [251].. Consumers may not understand or be aware of what their data is used for, which denies them the ability to allow the way their personal information is shared [252].

Facial recognition has been used in Russia to harass women allegedly involved in online pornography [253]. In Russia there is a 'FindFace' app capable of identifying faces with an accuracy of about 70% using the social media app called VK. This app wouldn't be possible in other countries that don't use VK as their social media platform photos aren't stored the same way as VK [254].

A hearing was held in July 2012 before the Subcommittee on Privacy, Technology and Law of the Judicial Commission, the United States Senate, to address issues related to the meaning of facial recognition technology for privacy and civil liberties.

In 2014, the National Telecommunications and Information Association (NTIA) launched a multi-stakeholder process to engage privacy defenders and industry representatives in order to establish guidelines relating to the use of facial recognition technology by private companies [255]. In June 2015, privacy advocates left the bargaining table on what they believed was a dead end based on the fact that industry representatives were unwilling to accept the consent requirements for facial recognition data collection [256]. The NTIA and industry representatives continued without the privacy representatives and draft rules are expected to be presented in spring 2016 [257].

In July 2015, the U.S. government accountability office conducted a report to the ranking member, privacy, technology and law subcommittee, the judicial commission, the United

States Senate. The report discussed the commercial uses of facial recognition technology, privacy concerns and applicable federal law. He claims that issues related to facial recognition technology have been discussed previously and represent the need for updated federal privacy laws that continually correspond to the degree and impact of advanced technologies. In addition, some industrial, governmental and private organizations are in the process of developing or have developed "voluntary privacy guidelines". These guidelines vary between groups, but the overall goal is to obtain consent and inform citizens about the intended use of facial recognition technology. This helps to counteract the privacy issues that arise when citizens are unaware of the place where their personal data and privacy are used since the report indicates it as a prevalent issue.

The main concern for the development of biometric technology, and more specifically facial recognition, concerns privacy. The increase in facial recognition technologies has led people to worry that large companies, such as Google or Apple, or even government agencies will use it for mass surveillance of the public [258].

5.5.2 Imperfect technology in law enforcement

Law enforcement agencies around the world have started using facial recognition software to identify criminals. It is still disputed whether or not facial recognition technology works less accurately on black people [259]. A study by Joy Buolamwini (MIT Media Lab) and Timnit Gebru (Microsoft Research) found that the error rate for gender recognition for black women within three commercial facial recognition systems varied from 23.8% to 36%, while for lighter-skinned men it was between 0.0 and 1.6%. Overall accuracy rates for identifying men (91.9%) were higher than women (79.4%) and none of the systems was able to understand a non-binary understanding of gender [260]. However, another study showed that several commercial facial recognition software sold to law enforcement offices across the country had a lower false mismatch rate for blacks than whites [261].

Experts fear that the new technology may actually harm some communities [262]. With so much room for error in this technology, it is believed that both legal defenders and facial recognition software companies claim that the technology should provide only part of the case, no evidence that can lead to the arrest of an individual [263].

The lack of regulations that bind facial recognition technology companies to racially distorted test requirements can be a significant flaw in adopting law enforcement use. CyberExtruder, a law enforcement company, said it did not perform any tests or research on distortion in their software. CyberExtruder has noted that some skin colors are more difficult to recognize for software with the current limitations of technology. "Just as individuals with very dark skin are difficult to identify with high meaning through facial recognition, individuals with very pale skin are the same," said Blake Senftner, a senior computer engineer at CyberExtruder [263].

In 2018, the Scottish government created a code of conduct that dealt with privacy issues and garnered praise from the Open Rights Group [258].

The facial recognition technology market has reached a staggering $ 4.6 billion in 2019 - and will grow another 25% in the next 9 years [264].

In May 2019, the San Francisco Council of Supervisors voted to ban police and other government agencies from using facial recognition technology, making San Francisco the first city in the United States to ban this practice [265].

The Swedish Data Protection Authority (DPA) issued its first ever financial penalty for a violation of the EU’s General Data Protection Regulation (GDPR) against a school that was using the technology to replace time-consuming roll calls during class. The DPA found that the school illegally obtained the biometric data of its students without completing an impact assessment. In addition, the school did not make the DPA aware of the pilot scheme. A 200,000 SEK fine (€19,000/$21,000) was issued.

6. Conclusions

The main goal of thesis is to analyse and provide an overall view of the facial recognition market and technologies, since the beginning of its study to actual applications and limits, to future trends and opportunities.

Face recognition systems and facial image processing applications significance as a research area are increasing. Implementations of system are crime prevention, video surveillance, person verification, and similar security activities.

Due to the benefits it provides over conventional security methods, facial recognition has gained prominence in recent times and governments around the world have invested substantial resources in facial recognition technology, including leading adoptions from the United States and China.

Technically advanced facial recognition technologies with mobile security apps and drones are likely to create more opportunities in the future and despites the high costs of implementation, investments on technologies are made and improvements are not long in coming.

Facial recognition technology is already mainstream and will continue to grow, but there are still substantial obstacles.

The biggest drawback for facial recognition technology in most people's opinions is the threat to an individual's privacy. In fact, several cities have considered or will ban real-time facial recognition surveillance use by law enforcement.

Moreover, the technology is not as effective at identifying people of colour and women as it is white males. Exponential progress has already been made, plunging to 0.2 percent errors in 2018 compared to 4 percent in 2013.

In addition, there are issues that need to be resolved that can throw off the technology when a person changes appearance or the camera angle isn't quite right.

Another potential downside is the storage of sensitive personal data and the challenges that come with it.

In order to benefit from the positive aspects of facial recognition, our society is going to have to work through some significant challenges to privacy and civil liberties. Will individuals accept the invasion of their privacy as a proper cost to being more secure and for the conveniences facial recognition provides?

References

[1] N. Samadiani, G. Huang, B. Cai, W. Luo, C.-H. Chi, Y. Xiang and J. He, *A Review on Automatic Facial Expression Recognition Systems Assisted by Multimodal Sensor Data,* 2019.

[2] C. A. Corneanu, J. F. C. Marc Oliu and S. Escalera, *Survey on RGB, 3D, Thermal, and Multimodal Approaches for Facial Expression Recognition: History, Trends, and Affect-related Applications,* 2016.

[3] R. W. R. Highfield and R. Jenkins, *How your looks betray your personality,* 2009.

[4] A. Chastel, *Leonardo on Art and the Artist. Courier Corporation,* 2002.

[5] S. G. e. al., *Toward a universal language of motion: reflections on a seventeenth century muscle man,* 1994.

[6] G.-B. D. d. Boulogne and R. A. Cuthbertson, “The Mechanism of Human Facial Expression,” *Cambridge University Press,* 1990.

[7] C. Darwin, “The expression of emotion in man and animals,” *Oxford University Press,* 1872.

[8] M. Suwa, N. Sugie and K. Fujimora, *A preliminary note on pattern recognition of human emotional expression,* 1978.

[9] M. K. and P. A., *Automatic lipreading by optical-flow analysis,* 1991.

[10] P. Ekman, T. S. Huang, T. J. Sejnowski and J. C. Hager, *Final report to NSF of the planning workshop on facial expression understanding,* 1993.

[11] Y.-L. Tian, T. Kanade and J. F. Cohn, *Recognizing action units for facial expression analysis,* 2001.

[12] G. Littlewort, M. S. Bartlett, I. Fasel, J. Susskind and J. Movellan, *Dynamics of facial expression extracted automatically from,* 2004.

[13] I. Kotsia and I. Pitas, *Facial expression recognition in image sequences using geometric deformation features and support vector machines,* 2007.

[14] I. Cohen, N. Sebe, L. Chen, A. Garg and T. S. Huang, *Facial expression recognition from video sequences: Temporal and static modelling,* 2003.

[15] P. S. Aleksic and A. K. Katsaggelos, *Automatic facial expression recognition using facial animation parameters and multistream hmms,* 2006.

[16] M. Pantic and I. Patras, *Dynamics of facial expression: recognition of facial actions and their temporal segments from face profile image sequences,* 2006.

[17] S. Koelstra, M. Pantic and I. Patras, *A dynamic texture-based approach to recognition of facial actions and their temporal models,* 2010.

[18] Y.-L. Tian, T. Kanade and J. F. Cohn, *Recognizing action units for facial expression analysis,* 2001.

[19] C. Shan, S. Gong and P. W. McOwan, *Facial expression recognition based on local binary patterns: A comprehensive study,* 2009.

[20]]. L. Yin, X. Wei, Y. Sun, J. Wang and M. J. Rosato, *A 3D facial expression database for facial behavior research,* 2006.

[21] H. Tang and T. Huang, *3D facial expression recognition based on automatically selected features,* 2008.

[22] J. Wang, L. Yin, X. Wei and Y. Sun, *3D facial expression recognition based on primitive surface feature distribution,* 2006.

[23] H. Soyel and H. Demirel, *Facial expression recognition using 3D facial feature distances,* 2007.

[24] N. Vretos, N. Nikolaidis and I. Pitas, *3D facial expression recognition using Zernike moments on depth images,* 2011.

[25] W. Zeng, H. Li, L. Chen, J. M. Morvan and X. D. Gu, *An automatic 3D expression recognition framework based on sparse representation of conformal images,* 2013.

[26] P. Lemaire, M. Ardabilian, L. Chen and M. Daoudi, *Fully automatic 3D facial expression recognition using differential mean curvature maps and histograms of oriented gradients,* 2013.

[27] B. Gong, Y. Wang, J. Liu and X. Tang, *Automatic facial expression recognition on a single 3D face by exploring shape deformation,* 2009.

[28] G. Sandbach, S. Zafeiriou and M. Pantic, *Local normal binary patterns for 3D facial action unit detection,* 2012.

[29] S. Berretti, B. B. Amor, M. Daoudi and A. D. Bimbo, *3D facial expression recognition using sift descriptors of automatically detected keypoints,* 2011.

[30] B. Hernandez, G. Olague, R. Hammoud, L. Trujillo and E. Romero, *Visual learning of texture descriptors for facial expression recognition in thermal imagery,* 2007.

[31] P. Liu and L. Yin, *Spontaneous facial expression analysis based on temperature changes and head motions,* 2015.

[32] Y. Chang, M. Vieira, M. Turk and L. Velho, *Automatic 3D facial expression analysis in videos,* 2005.

[33] V. Le, H. Tang and T. S. Huang, *Expression recognition from 3D dynamic faces using robust spatio-temporal shape features,* 2011.

[34] H. Nguyen, K. Kotani, F. Chen and B. Le, *A thermal facial emotion database and its analysis,* 2014.

[35] S. Wang, Z. Liu, S. Lv, Y. Lv, G. Wu, P. Peng, F. Chen and X. Wang, *A natural visible and infrared facial expression database for expression recognition and emotion inference,* 2010.

[36] N. Sebe, M. S. Lew, Y. Sun, I. Cohen, T. Gevers and T. S. Huang, *Authentic facial expression analysis,* 2007.

[37] S. Lucey, A. B. Ashraf and J. F. Cohn, *Investigating spontaneous facial action recognition through aam representations of the face,* 2007.

[38] M. F. Valstar, M. Pantic, Z. Ambadar and J. F. Cohn, *Spontaneous vs. posed facial behavior: automatic analysis of brow actions,* 2006.

[39] Z. Zeng, Y. Fu, G. I. Roisman, Z. Wen, Y. Hu and T. S. Huang, *Spontaneous emotional facial expression detection,* 2006.

[40] R. E. Kaliouby and P. Robinson, *Real-time inference of complex mental states from facial expressions and head gestures,* 2005.

[41] Q. Ji, P. Lan and C. Looney, *A probabilistic framework for modeling and real-time monitoring human fatigue,* 2006.

[42] A. Kapoor, W. Burleson and R. W. Picard, *Automatic prediction of frustration,* 2007.

[43] A. B. Ashraf, S. Lucey, J. F. Cohn, T. Chen, Z. Ambadar, K. M. Prkachin and P. E. Solomon, *The painful face–pain expression recognition using active appearance models,* 2009.

[44] G. C. Littlewort, M. S. Bartlett and K. Lee, *Automatic coding of facial expressions displayed during posed and genuine pain,* 2009.

[45] G. C. Littlewort, M. S. Bartlett and K. Lee, *automated measurement of spontaneousallfacial expressions of genuine and posed pain,* 2007.

[46] J. M. Girard, J. F. Cohn, M. H. Mahoor, S. M. Mavadati, Z. H. mal and D. P. Rosenwald, *Nonverbal social withdrawal in depression: Evidence from manual and automatic analyses,* 2014.

[47] G. M. Lucas, J. Gratch, S. Scherer, J. Boberg and G. Stratou, *Towards an affective interface for assessment of psychological distress.*

[48] D. DeVault, R. Artstein, G. Benn, T. Dey, E. Fast, A. Gainer, K. Georgila, J. Gratch, A. Hartholt and M. L. e. al., *Simsensei kiosk: A virtual human interviewer for healthcare decision support,* 2014.

[49] J. F. Cohn and K. L. Schmidt, *The timing of facial motion in posed and spontaneous smiles,* 2004.

[50] Z. Ambadar, J. F. Cohn and L. I. Reed, *All smiles are not created equal: Morphology and timing of smiles perceived as amused, polite, and embarrassed/nervous,* 2009.

[51] M. F. Valstar, B. Jiang, M. Mehu, M. Pantic and K. Scherer, *The first facial expression recognition and analysis challenge,* 2011.

[52] M. Valstar, J. Girard, T. Almaev, G. McKeown, M. Mehu, L. Yin, M. Pantic and J. Cohn, *Fera 2015-second facial expression recognition and analysis challenge,* 2015.

[53] M. S. Littlewort, B. G. C., M. G. Frank, C. R. Lainscsek, I. Fasel and M. J. R., *Automatic recognition of facial actions in spontaneous expressions,* 2011.

[54] A. Savran, B. Sankur and M. T. Bilge, *Estimation of facial action intensities on 2d and 3D data,* 2011.

[55] A. Savran, B. Sankur and M. T. Bilge, *Regression-based intensity estimation of facial action units,* 2014.

[56] S. Kaltwang, O. Rudovic and M. Pantic, *Continuous pain in tensity estimation from facial expressions,* 2012.

[57] N. Zaker, M. H. Mahoor, W. Mattson, D. S. Messinger and J. F. C. e. a. m. o. s. f. actions:., *Cohn et al., "Intensity measurement of spontaneous facial actions: Evaluation of different image representations,* 2012.

[58] M. Shreve, S. Godavarthy, V. Manohar, D. Goldgof and S. Sarkar, *Towards macro-and micro-expression spotting in video using strain patterns,* 2009.

[59] M. Shreve, S. Godavarthy, D. Goldgof and S. Sarkar, *Macro-and micro-expression spotting in long videos using spatio-temporal strain,* 2011.

[60] T. Pfister, X. Li, G. Zhao and M. Pietikainen, *Recognising spontaneous facial micro-expressions,* 2011.

[61] P. Ekman and W. V. Friesen, Unmasking the Face: A Guide to Recognizing Emotions from Facial Clues, ISHK, 2003.

[62] Y. Tian, T. Kanade and J. F. Cohn, Handbook of face recognition, 2nd revised and expanded ed. ed., 2011.

[63] A. Poursaberi, H. Noubari, M. Gavrilova and Yanushkevich, "Gauss - Laguerre wavelet textural feature fusion with geometrical information for facial expression identification," 2012, pp. pp. 1-13.

[64] Bashyal and Venayagamoorthy, "Recognition of facial expressions using Gabor wavelets and learning vector quantization," Elsevier, 2008.

[65] P. Taylor, M. H. Siddiqi, R. Ali, A. Sattar, A. M. Khan and S. Lee, "Depth camera-based facial expression recognition system using multilayer scheme," *IETE Technical Review,* no. July 2014, 2014.

[66] E. Owusu, Y. Zhan and Q. Mao, "A neural-AdaBoost based facial expression recognition system," *Expert systems with applications,* vol. Volume 41, no. Issue 7, 1 June 2014.

[67] S. Biswas, "An Efficient Expression Recognition Method using Contourlet Transform," *Int. Conf. Percept. Mach. Intell.,* 2015.

[68] Y. Ji and K. Idrissi, "Automatic facial expression recognition based on spatiotemporal descriptors," *Pattern Recognition Letters,* vol. Volume 33, no. Issue 10, 15 July 2012.

[69] M. Dahmane and J. Meunier, "Prototype-Based Modeling for Facial Expression Analysis," *ISEE Transactions on Multimedia,* vol. Volume 16, no. issue 6, 2014.

[70] A. Hernandez-Matamorosa, A. Bonarini, E. Escamilla-Hernandeza, M. Nakano-Miyatakea and H. Perez-Meana, "Facial expression recognition with automatic segmentation of face regions using a fuzzy based classification approach," *Knowledge-based Systems,* vol. Volume 110, 15 October 2016.

[71] Y. Demir, “A new facial expression recognition based on curvelet transform and online sequential extreme learning machine initialized with spherical clustering,” *Neural Computing and Applications,* vol. Issue 1, January 2016.

[72] Y. Gao, M. Leung, S. C. Hui and M. Tananda, “Facial expression recognition from line-based caricatures,” *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans,* vol. Volume 33, no. Issue 3, 2003.

[73] G. Hegde, M. Seetha and N. Hegde, “Kernel Locality Preserving Symmetrical Weighted Fisher Discriminant Analysis based subspace approach for expression recognition,” *Engineering Science and Technology, an International Journal,* vol. Volume 19, no. Issue 3, September 206.

[74] D. I. Islam, S. R. N. Anal and A. Datta, Facial Expression Recognition Using 2DPCA on Segmented Images, 2018.

[75] M. H. Siddiqi, R. Ali, A. M. Khan, Y.-T. Park and S. Lee, “Human Facial Expression Recognition Using Stepwise Linear Discriminant Analysis and Hidden Conditional Random Fields,” *IEEE Transactions on Image Processing,* vol. Volume 24, no. Issue 4, 2015.

[76] M. J. Cossetin, J. C. Nievola and A. L. Koerich, *Facial expression recognition using a pairwise feature selection and classification approach,* 2016.

[77] S. Noh, H. Park, Y. Jin and J.-I. Park, “Feature-Adaptive Motion Energy Analysis for Facial Expression Recognition,” in *Advances in Visual Computing*, 2007, pp. pp 452-463.

[78] T. A. Rashid, “Convolutional Neural Networks based Method for Improving Facial Expression Recognition,” in *Intelligent Systems Technologies and Applications*, 2016, pp. pp 73-84.

[79] H. Mahersia and K. Hamrouni, “Using multiple steerable filters and Bayesian regularization for facial expression recognition,” in *Engineering Applications of Artificial Intelligence*, Elsevier, February 2015, pp. Pages 190-202.

[80] K. Shan, J. Guo, W. You, D. Lu and R. Bie, *Automatic facial expression recognition based on a deep convolutional-neural-network structure,* 2017.

[81] R. Cui, M. Liu and M. Liu, *Facial Expression Recognition Based on Ensemble of Mulitple CNNs,* 2016, pp. pp 511-518.

[82] Z. Duric, W. D. Gray, R. Heishman, F. Li, A. Rosenfeld, M. J. Schoelles, C. Schunn and H. Wechsler, *Integrating perceptual and cognitive modeling for adaptive and intelligent human computer interaction,* 2002.

[83]]. L. Maat and M. Pantic, *Gaze-x: adaptive, affective, multimodal interface for single-user office scenarios,* 2007.

[84] H. Ishiguro, T. Ono, M. Imai, T. Maeda, T. Kanda and R. Nakatsu, *Robovie: an interactive humanoid robot,* 2001.

[85] A. Kapoor, W. Burleson and R. W. Picard, *Automatic prediction of frustration,* 2012.

[86] S. Bakkes, C. T. Tan and Y. Pisan, *Personalised gaming,* 2012.

[87] C. T. Tan, D. Rosser, S. Bakkes and Y. Pisan, *A feasibility study in using facial expressions analysis to evaluate player experiences,* 2012.

[88] P. M. Blom, S. Bakkes, C. T. Tan, S. Whiteson, D. Roijers, R. Valenti and T. Gevers, *Towards personalised gaming via facial expression recognition,* 2014.

[89] P. Lucey, J. F. Cohn, I. Matthews, S. Lucey, S. Sridharan, J. Howlett and K. M. Prkachin, *Automatically detecting pain in video through facial action units,* 2011.

[90] S. Kaltwang, O. Rudovic and M. Pantic, *Continuous pain in tensity estimation from facial expressions,* 2012.

[91] R. Irani, K. Nasrollahi, C. M. O. Simon, A. Corneanu, S. Escalera, C. Bahnsen, D. H. Lundtoft, T. B. Moeslund, T. L. Pedersen and M.-L. e. al., *Spatiotemporal analysis of rgb-dt facial images for multimodal pain level recognition,* 2015.

[92] A. Ryan, J. F. Cohn, S. Lucey, J. Saragih, P. Lucey, F. D. I. Torre and A. Ross, *Automated facial expression recognition system,* 2009.

[93] E. Vural, M. Cetin, A. Ercil, G. Littlewort, M. Bartlett and J. Movellan, *Drowsy driver detection through facial movement analysis,* 2007.

[94] J. M. Girard, J. F. Cohn, M. H. Mahoor, S. M. Mavadati, Z. H. mal and D. P. Rosenwald, *Nonverbal social withdrawal in depression: Evidence from manual and automatic analyses,* 2014.

[95] S. Scherer, G. Stratou, M. Mahmoud, J. Boberg, J. Gratch, A. Rizzo and L.-P. Morency, *Automatic behavior descriptors for psychological disorder analysis,* 2013.

[96] J. Joshi, A. Dhall, R. Goecke, M. Breakspear and G. Parker, *Neural-net classification for spatio-temporal descriptor based depression analysis,* 2012.

[97] C. Mauri, T. Granollers, J. Lorés and M. García, “Computer vision interaction for people with severe movement restrictions,” *Interdisciplinary Journal on Humans in ICT Environments,* vol. Vol 2, p. pp. 38–54, 2006..

[98] R. Kjeldsen, *Improvements in vision-based pointer control,* 2006.

[99] G. Hatice, *Automatic, dimensional and continuous emotion recognition.,* 2010.

[100] S. Siddharth and A. Rautaray, “Vision based hand gesture recognition for human computer interaction: a survey,” *Springer Journal Artificial Intelligence Review,* p. pp. 1–54, 2012.

[101] R. Azad, B. Azad and I. T. Kazerooni, “ReaReal-Time and Robust Method for Hand Gesture Recognition System Based on Cross-Correlation Coefficient,” *Advances in Computer Science: an International Journal ,* Vols. Vol. 2, Issue 5, no. No.6, pp. pp. 121-125., 2013.

[102] J. P. Wachs, M. K. olsch, H. Stern and Y. Edan, “Vision based hand-gesture applications,” *Communications of the ACM,* pp. pp. 54-60., 2011.

[103] O. A. v. Nierop, A. v. d. Helm, K. J. Overbeeke and T. J. Djajadiningrat, “A natural human hand model,” *Springer journal The Visual Computer,* Vols. pp 31-44., pp. pp 31-44., Volume 24, Issue 12008.

[104] Y. Tu, C. Kao and H. Lin, *Human Computer Interaction Using Face and Gesture Recognition,* 2013.

[105] A. Matos, V. Filipe and P. Couto, *Human-computer interaction based on facial expression recognition: a case study in Degenerative Neuromuscolar Disease,* 2016.

[106] J. Kumari, R. Rajesh and K. Pooja, *Facial expression recognition: a survey,* 2015.

[107] P. Viola and M. Jones, “Robust Real-time Face Detection,” *International Journal of Computer Vision,* p. pp.137–154, 2004.

[108] F. Dornaika and B. Raducanu, *Facial Expression Recognition for HCI Applications,* 2009.

[109] C. Breazeal, *Function meets style: Insights from emotion theory applied to HRI.,* 2004.

[110] P. Rani and N. Sarkar, *Emotion-sensitive robots – a new paradigm for human-robot interaction,* 2004.

[111] M. Scheutz, P. Schermerhorn and J. Kramer, *The utility of affect expression in natural language interactions in joint human-robot tasks.,* 2006.

[112] R. Picard, *Toward agents that recognize emotion. In: Imagina98.,* 1998.

[113] C. Ikehara, D. Chin and M. Crosby, *A model for integrating an adaptive in formation filter utilizing biosensor data to assess cognitive load.,* 2003.

[114] R. Vick and C. Ikehara, *Methodological issues of real time data acquisition from multiple sources of physiological data.,* 2003.

[115] E. Sheldon, *Virtual agent interactions.,* 2001.

[116] C. Breazeal, *Emotion and sociable humanoid robots.,* 2003.

[117] S. S, B. A and e. a. Kuhnlenz K, *Design and evaluation of emotion-display EDDIE.,* 2006.

[118] F. Hegel, F. Eyssel and B. Wrede, *The social robot Flobi: Key concepts of industrial design.,* 2010.

[119] V. Seib, J. Giesen and D. G. e. al., *Enhancing human-robot interaction by a robot face with facial expressions and synchronized lip movements.,* 2013.

[120] S. Moubayed, G. Skantze and J. Beskow, *The Furhat back projected humanoid head–lip reading, gaze and multi-party interaction.,* 2013.

[121] Y. Kondo, K. Takemura and J. T. e. al., *A gesture-centric Android system for multi-party human-robot interaction.,* 2013.

[122] C. Bennett and S. Sˇabanovic´, *Deriving minimal features for human-like facial expressions in robotic faces.,* 2014.

[123] H. Miwa, T. Umetsu and A. T. e. al., *Robot personality based on the equations of emotion defined in the 3D mental space, Proceedings 2001 ICRA.,* 2001.

[124] K. Itoh, H. Miwa and M. M. e. al., *Various emotional expressions with emotion expression humanoid robot WE4RII, IEEE Conference on Robotics and Automation, 2004.,* 2004.

[125] J. R.-d. Solar, M. Correa and R. Verschae, *Bender: a general-purpose social robot with human-robot interaction capabilities.,* 2012.

[126] S. Gallagher and S. Varga, *Social constraints on the direcperception of emotions and intentions.,* 2014.

[127] S. Johnson, *The recognition of mentalistic agents in infancy.,* 2000.

[128] M. Roubroeks, J. Ham and C. Midden, *When artificial social agents try to persuade people: the role of social agency on the occurrence of psychological reactance.,* 2011.

[129] D. Kim, S. Jung, K. An, H. Lee and M. Chung, *Development of a facial expression imitation system.,* 2006.

[130] Y. Yoshitomi, S. Kim, T. Kawano and T. Kilazoe, *Effect of sensor fusion for recognition of emotional states using voice, face image and thermal image of face.,* 2000.

[131] T. Otsuka and J. Ohya, *Recognition of facial expressions using hmm with continuous output probabilities.,* 1996.

[132] M. Wimmer, B. A. MacDonald, D. Jayamuni and A. Yadav, *Facial Expression Recognition for Human-robot interaction - A Prototype,* 2008.

[133] M. Wimmer, *Model-based Image Interpretation with Application to Facial Expression Recognition.,* 2006.

[134] J. Pantic and L. Rothkrantz, *Automatic analysis of facial expressions: The state of the art.,* 2000.

[135] C. Chibelushi and F. Bourel, *Facial expression recognition: A brief tutorial overview.,* 2003.

[136] J. Hirth and K. Berns, *Humanoid robots,* 2009.

[137] J. Hirth, N. Schmitz and K. Berns, *International Journal of Social Robotics.,* 2011.

[138] K. Berns and J. Hirth, *Proc. Of IROS,* 2006.

[139] Chen, Gonzalez, Leung, Zhang and Li, *A 2G-RFID based e-healthcare system.,* 2010.

[140] González-Valenzuela, Chen and Leung, *Mobility support for health monitoring at home using wearable sensors.,* 2011.

[141] Chen, Ma, Ullah, Cai, Song and Rochas, *Rochas: robotics and cloud-assisted healthcare system for empty nester.,* 2013.

[142] Zhang, Xiong, Hsu and Vasilakos, *BASA: building mobile Ad-Hoc social networks on top of android.,* 2014.

[143] L. R., L. Y, L. N.D. and Z. L, *Moodscope: build ing a mood sensor from smartphone usage patterns.,* 2013.

[144] C. M., *NDNC-BAN: supporting rich media healthcare service via named data networking in cloud-assisted wireless body area networks.,* 2014.

[145] Wan, Ullah, Lai, Zhou and Wang, *Cloud enabled wireless body area networks for pervasive healthcare.,* 2013.

[146] Chen, Zhang, Li, Mao and Leung, *EMC: emotion-aware mobile cloud computing.,* 2015.

[147] G. Muhammad, M. Alsulaiman, S. U. Amin, A. Ghoneim and M. F. Alhamid, *A Facial-Expression Monitoring System for Improved Healthcare in Smart Cities,* 2017.

[148] S. Mallat and G. Peyré, *A review of bandlet methods for geometrical image representation,* 2007.

[149] S. Yvan, I. Inaki and L. Pedro, *A review of feature selection techniques in bioinformatics,* 2007.

[150] L. Rabiner and B. H. Juang, *Fundamentals of Speech Recognition.,* 1993.

[151] S. Abe, *Support Vector Machines for Pattern Classification,* 2005.

[152] "Visage Technologies," 2019-10-02. [Online].

[153] S. Lee, G. Carlson and S. J. e. al., *Designing an expressive avatar of a real person,* USA, 2010.

[154] W. v. Tol and A. Egges, *Real-time crying simulation,* 2009.

[155] P. Ekman and W. V. Friesen, *Detecting deception from the body or face,* 1974.

[156] F. Parke, K. Waters and T. R. Alley, *Computer Facial Animation,* 2008.

[157] M. Ochs, R. Niewiadomski and C. Pelachaud, *Facial Expressions of Emotions for Virtual Characters,* 2015.

[158] J. Ostermann, *Face animation in mpeg-4.,* 2002.

[159] Z. Ruttkay, H. Noot and P. T. Hagen, *Emotion disc and emotion squares: tools to explore the facial expressions face,* 2003.

[160] A. Mehrabian., *Basic Dimensions for a General Psychological Theory: Implications for Personality, Social, Environmental, and Developmental Studies.,* 1980.

[161] K. Scherer, A. Schorr and T. Johnstone, *Appraisal Processes in Emotion: Theory, Methods, Research.,* 2001.

[162] M. Zeiler, G. Taylor, L. Sigal, I. Matthews and R. Fergus, *Facial expression transfer with input-output temporal restricted boltzmann machines.,* 2011.

[163] M. D. Ilie, C. Negrescu and D. Stanomir, *Circular interpolation for morphing 3D facial animations,* 2011.

[164] Y. Zhang, Q. Ji, Z. Zhu and B. Yi, *Dynamic facial expression analysis and synthesis with MPEG-4 facial animation parameters,* 2011.

[165] C. M. d. Melo and J. Gratch, *Expression of emotions using wrinkles, blushing, sweating and tears,* 2009.

[166] C. Darwin, *The Expression of the Emotions in Man and Animals,* 1872.

[167] T. S. J. Jimenez and N. B. e. al., *A practical appearance model for dynamic facial color,* 2010.

[168] N. Ersotelos and F. Dong, *Building highly realistic facial modeling and animation: a survey,* 2008.

[169] M. H. Alkawaz, A. H. Basori, D. Mohamad and F. Mohamed, *Realistic Facial Expression of Virtual Human Based on Color, Sweat, and Tears Effects,* 2014.

[170] P. Moireau, N. Xiao and M. A. e. al., *External tissue support and fluid-structure simulation in blood flows,* 2012.

[171] M. Nešic and V. Nešic, *Neuroscience of Nonverbal Communication, Chapter in The Social Psychology of Non-verbal Communication.,* 2014.

[172] P. Ekman and V. Wallace, *Nonverbal leakage and clues to deception. Psychiatry .,* 1969.

[173] M. Owayjan, A. Kashour, N. A. Haddad, M. Fadel and G. A. Souki, *The design and development of a Lie Detection System using facial micro-expressions.,* 2012.

[174] M. Margoudi, J. Hart, A. Adams and M. Oliveira, *Exploring Emotion Representation to Support Dialogue in Police Training on Child Interviewing,* 2016.

[175] M. Moussa and M. Magnenat-Thalmann, *Applying affect recognition in serious games: The playmancer project.,* 2009.

[176] S. Ioannou, A. Raouzaiou, V. Tzouvaras, T. Mailis, K. Karpouzis and S. Kollias, *Emotion recognition through facial expression analysis based on a neurofuzzy network.,* 2005.

[177] S.L.Finkelstein, A. Nickel, L. Harrison, E. Suma. and T. Barnes, *cMotion: A new game design to teach emotion recognition and programming logic to children using virtual humans.,* 2009.

[178] M. Kebritchi, *Examining the pedagogical foundations of modern educational computer games.,* 2008.

[179] K. B., *A Brief Review of Facial Emotion Recognition Based on Visual Information. Sensors.,* 2018.

[180] H. Yang, U. Ciftci and L. Yin, *Facial Expression Recognition by De-expression residue Learning.,* 2018.

[181] F. Zhang, T. Zhang, Q. Mao and C. Xu, *Joint pose and expression modeling for facial expression recognition.,* 2018.

[182] B. Hasani and M. Mahoor, *Facial Expression Recognition Using Enhanced Deep 3D Convolutional Neural Networks,* 2017.

[183] K. Zhao, W. Chu and H. Zhang, *Deep Region and Multi-label Learning for Facial Action Unit Detection.,* 2016.

[184] M. Liu, S. Li, S. Shan and X. Chen, *Au-inspired deep networks for facial expression feature learning.,* 2015.

[185] M. Jeong and B. C. Ko, *Driver's Facial Expression Recognition in Real-Time for Safe Driving,* 2018.

[186] D. Consoli, *A New Concept of Marketing: The Emotional Marketing,* 2010.

[187] W. Jarrold, *Towards a theory of affective mind: computationally modeling the generativity of goal appraisal.,* 2004.

[188] M. S. Bouzakraoui, A. SADIQ and N. Enneya, *A Customer Emotion Recognition through Facial Expression using POEM descriptor and SVM classifier,* 2017.

[189] K. Sangwook and e. al., *Deep learning of support vector machines with class probability output networks.,* 2014.

[190] A. Vinay, S. Shekhara, J. Rituparna, A. Tushar, K. N. B. Murthya and S. Natarajan, *Cloud Based Big Data Analytics Framework for Face Recognition in Social Networks using Machine Learning,* 2015.

[191] "Face Time," *Metalife.*

[192] Najafabadi, Villanustre, Khoshgoftaar, Seliya, Wald and Muharemagic, *Deep learning applications and challenges in big data analytics. - Journal of Big Data,* 2015.

[193] R. Narayanan, R. Chellappa and A. K. Roy-Chowdhury, *Facial similarity across age, disguise, illumination and pose.,* 2004.

[194] C. Min, S. Mao, Y. Zhang and V. C. Leung, *Big Data Analysis.,* 2014.

[195] "How Big Data drives Facial Recognition," *Keywebmetrcs ,* 2015.

[196] S. Mahmudova, *Big Data Challenges in Biometric Technology.,* 2016.

[197] [Online]. Available: https://en.wikipedia.org/wiki/Apache_Giraph.

[198] M. Sh., *Mahmudova Sh. The Role of Biometric Networks in Recognition of Person. Communications.,* 2015.

[199] [Online]. Available: http://findbiometrics.com/face-recognition-and-big-data-analysis-bringing-efficiency-to-lawenforcement/.

[200] [Online]. Available: https://www.pac-online.com/saas-big-data-dynamic-duo-we-have-all-been-waitin.

[201] [Online]. Available: https://www.asdevents.com/event.asp?id=3225&desc=Biometrics+Big+Data+Symposium.

[202] [Online]. Available: https://hive.apache.org/.

[203] "Is Facebook Taking Big Data Analytics Too Far?," *Smartdatacollective.com,* 2015.

[204] C. M. Bishop, *Pattern recognition and machine learning,* 2006.

[205] “New Beginnings in Facial Recognition,” *Thoughtworks.com ,* 2014.

[206] S. Yi, Y. Chen, X. Wang and X. Tang, *Deep learning face representation by joint identification-verification.,* 2014.

[207] [Online]. Available: https://towardsdatascience.com/an-intro-to-deep-learning-for-face-recognition-aa8dfbbc51fb.

[208] P. Peter and J. J. VitomirŠtruc, *Building cloud-based biometric services.,* 2013.

[209] C. Hehua, L. Chi, M. Fang and J. Wu, *Facial Expression Recognition Based On Cloud Model.*

[210] M. A. A. Dakhane D. M., *Data Security in Cloud Computing for Biometric Application,* 2012.

[211] “Cloud Face Recognition Services,” *Animetrics,* 2015.

[212] “Image processing and cloud computing Architecture overview,” *Computervisiontalks.com,* 2015.

[213] C. C. T. Xiang, *Privacy-Preserving Face Recognition with Outsourced Computation..*

[214] “BioID Cloud-based Biometric Recognition System,” *BioID.com Homepage,* 2015.

[215] “Face.com Cloud-based Biometric Recognition System,” *Face.com Homepage,* 2015.

[216] M. G. Galterio, S. A. Shavit and T. Hayajneh, *A Review of Facial Biometrics Security for Smart Devices,* 2018.

[217] P. S., B. R.M. and J. A. Biometrics, *The Future of Identification. Computer,* 2000.

[218] J. J. Woodward, C. Horn, J. Gatune and A. Thomas, *Biometrics: A Look at Facial Recognition,* 2003.

[219] S. R., U. Uludag, A. Mink, M. Indovina and J. A., *Large-scale evaluation of multimodal biometric authentication using state-of-the-art systems.,* 2005.

[220] M. Ohana, O. Dunkelman, S. Gibson and M. Osadchy, *HoneyFaces: Increasing the Security and Privacy of Authentication Using Synthetic Facial Images.,* 2018.

[221] U. Mardikar, *Systems and Methods for Authenticating Facial Biometric Data against Secondary Sources.,* 2017.

[222] I. Traoré, Y. Nakkabi, S. Saad, B. Sayed and J. Ardigo, *P.M: Ensuring Online Exam Integrity through Continuous Biometric Authentication.,* 2017.

[223] M. Chowdhury, J. Gao and R. Islam, *Fuzzy rule based approach for face and facial feature extraction in biometric authentication.,* 2017.

[224] *Intel Security. True Key.,* 2018.

[225] V. Shchutskaya, *Real-life Application of Face Recognition Technology,* 2018.

[226] “Sabre.com,” [Online].

[227] “reportsanddata.com,” May 2019. [Online].

[228] “PR News Wire,” 2019. [Online]. Available: https://www.prnewswire.co.uk/.

[229] “Here's How Many Adult Faces Are Scanned From Facial Recognition Databases.”.

[230] “The trouble with facial recognition technology (in the real world).”.

[231] “Real-Time Facial Recognition Is Available, But Will U.S. Police Buy It?”.

[232] “Police Facial Recognition Databases Log About Half Of Americans”.

[233] Knezevic, K. R. and Alison., *Maryland's use of facial recognition software questioned by researchers, civil liberties advocates..*

[234] *"Next Generation Identification",* 2016.

[235] *"Face Off: The lawless growth of facial recognition in UK policing" (PDF). Big Brother Watch..*

[236] A. Sebastian, *UK police arrest man via automatic face-recognition tech.,* 2017.

[237] *"Smartgates". Australian Border Force.,* 11 March 2019.

[238] *"Our history". New Zealand Customs Service.,* 11 March 2019..

[239] *Facial recognition technology is coming to Canadian airports this spring". CBC News.,* 2017.

[240] V. Ben., *Panama puts names to more faces.,* 2014.

[241] *China's massive investment in artificial intelligence has an insidious downside.,* 2018.

[242] *China bets on facial recognition in big drive for total surveillance.,* 2018.

[243] T. J. Schellevis, *Politie gaat verdachten opsporen met gezichtsherkenning.,* 2019.

[244] L. Boon, *Meekijken met de 226 gemeentecamera's.,* 2019.

[245] “Fortune Business Insights,” 2019. [Online]. Available: https://www.fortunebusinessinsights.com/industry-reports/facial-recognition-market-101061.

[246] “marketsandmarkets.com,” 2019. [Online].

[247] “Allied Market Research,” 2016. [Online]. Available: www.alliedmarketresearch.com.

[248] N. Martin, “The Major Concerns Around Facial Recognition Technology,” 2019.

[249] *"EFF Sues FBI For Access to Facial-Recognition Records". Electronic Frontier Foundation.,* 2013.

[250] *"Q&A On Face-Recognition". American Civil Liberties Union..*

[251] H. Geiger, *"Facial Recognition and Privacy". Center for Democracy & Technology.,* 2018.

[252] C. A. Puente, *"FACIAL RECOGNITION TECHNOLOGY Commercial Uses, Privacy Issues, and Applicable Federal Law" (PDF).,* 2015.

[253] *Facial Recognition is getting really accurate, and we have not prepared.,* 2016.

[254] *This creepy facial recognition app is taking Russia by storm.,* 2016.

[255] *"Privacy Multistakeholder Process: Facial Recognition Technology". National Telecommunications and Information Association.,* 2016.

[256] M. David, *Facial recognition talks break down as privacy advocates withdraw.,* 2016.

[257] W. Dustin, *Business eyes facial recognition guidelines.,* 2016.

[258] [Online]. Available: https://en.wikipedia.org/wiki/Facial_recognition_system#Mobile_platforms.

[259] *"Photo Algorithms ID White Men Fine—Black Women, Not So Much". WIRED.,* 2018.

[260] J. Buolamwini and T. Gebru, *"Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification". Proceedings of Machine Learning Research, vol. 81: 77–91.,* 2018.

[261] G. Patrick, Q. George and P. Jonathon, *"Report on the Evaluation of 2D Still-Image Face Recognition Algorithms" (PDF). National Institute of Standards and Technology.,* 2011.

[262] B. Stephen, *"Rise of the racist robots – how AI is learning all our worst impulses". the Guardian.,* 2018.

[263] B. Ali, *"How white engineers built racist code – and why it's dangerous for black people". the Guardian.,* 2018.

[264] *"Facial Recognition Technology Market Report 2019-2029". Visiongain.,* 2019.

[265] P. Kari, *"San Francisco is first US city to ban police use of facial recognition tech". The Guardian.,* 2019.