The current challenges of automatic recognition of facial expressions: A systematic review

Audrey Masson [a,b,*], Guillaume Cazenave [b], Julien Trombini [b] and Martine Batt [a]

[a] *Interpsy – GRC, University of Lorraine, France*
E-mails: audrey.masson@two-i.fr, martine.batt@univ-lorraine.fr
[b] *Two-I, France*
E-mails: guillaume.cazenave@two-i.fr, julien.trombini@two-i.fr

Abstract. In recent years, due to its great economic and social potential, the recognition of facial expressions linked to emotions has become one of the most flourishing applications in the field of artificial intelligence, and has been the subject of many developments. However, despite significant progress, this field is still subject to many theoretical debates and technical challenges. It therefore seems important to make a general inventory of the different lines of research and to present a synthesis of recent results in this field. To this end, we have carried out a systematic review of the literature according to the guidelines of the PRISMA method. A search of 13 documentary databases identified a total of 220 references over the period 2014–2019. After a global presentation of the current systems and their performance, we grouped and analyzed the selected articles in the light of the main problems encountered in the field of automated facial expression recognition. The conclusion of this review highlights the strengths, limitations and main directions for future research in this field.

Keywords: Artificial intelligence, affective computing, facial expression, AFER, emotion theories

1. Introduction

For several decades now, emotional computing has been a flourishing multidisciplinary field of research with broad economic and social applications:

- mass emotional analysis (general mood of a population, level of well-being generated by a recreational or tourist structure, etc.);
- security (risks of aggression in a stadium or on public transport, detection of drowsy drivers, etc.);
- marketing and entertainment (emotional reactions to products, ads or films);
- personal assistants;
- health (pain detection, support for people with communication disorders, up to assistance with medical and psychopathological diagnosis).

Within this field, the automatic recognition of facial expressions (also called AFER) is the subject of an increasing number of researches.

Indeed, the idea is now widely accepted that facial expressions are a powerful non-verbal means of communication that provides a lot of information on the subjective experience of individuals (mental states, interests, opinions, physiological states, emotions) but also on their social motivations, their representations of the world and their intentions of action [163].

Therefore, they play a vital role in human interactions and are necessary for the progressive humanization of human-machine interactions.

Thanks to the evolution of techniques, in particular with the generalization of convolution networks and other deep learning approaches, progress in this field is constant but much remains to be done.

Indeed, the majority of the systems of AFER are based on a linear model of the emotions:

> Specific configurations of facial-muscle movements appear as if they summarily broadcast or display a person's emotions, which is why they are routinely referred to as emotional expressions and facial expressions (Barrett et al., 2019, p. 2 [15]).

---

*Corresponding author. E-mail: audrey.masson@two-i.fr.

Theories postulating a direct link between a finite number of emotions and facial expressions are currently being challenged. There is a consensus that human emotions are ephemeral, subjective and changing processes, the interpretation of which depends on many factors related to individuals as well as to the physical and social environment.

These characteristics make it an infinitely complex phenomenon to model and therefore to decode in all its nuances. Endowing a machine with the ability to recognize human emotional states is a scientific challenge around which researchers from many disciplinary fields (artificial intelligence, computer vision, robotics, psychology, cognitive sciences, neuroscience, sociology, etc.) gather and which still poses many theoretical and technical problems.

In this context, it appears necessary to make a systematic inventory of research in the field of the AFER, in order to update knowledge and critically analyse the theories and techniques that serve as a basis for current systems, as well as the challenges that remain to be taken up.

This article is organized as follows:

After describing our literature search methodology and the article selection process, we will quantitatively analyze the data from these articles. Finally, a thematic and qualitative synthesis will lead us to consider relevant perspectives in the field of automated facial expression recognition.

2. Methodology

2.1. Databases and search items

We searched 13 databases:

Pubmed; Biomed; ScienceDirect; Web of Science; Springerlink; Wiley; IEEExplore; ACM Digital Library; Cochrane; CAIRN; Scopus; Openedition; Open Grey.

The search was conducted using two sets of keywords in combination:

"automatic detection/analysis/automatic recognition/affective" "emotion/facial expression/emotion expression/facial display/action units/computing/ emotion context".

For example, the search terms were:

"automatic detection AND emotion", "automatic recognition AND facial expression".

These terms were used in the article metadata (title, keywords and abstract).

Table 1
Inclusion and exclusion criteria

Inclusion criteria	Exclusion criteria
Experimental studies, other study designs, reviews, meta-analysis	Emotion recognition by humans
Publications between 2014 and 2019	Face detection
Language: English	Pathological population
Automatic recognition of facial displays	
Typical adults	

2.2. Inclusion and exclusion criteria

Because automatic recognition of facial expressions is a rapidly evolving field, we have selected only studies published between 2014 and September 2019 (Table 1).

2.3. Selection process and flow diagram

As shown in Fig. 1, 890 unique citations were identified, of which 475 were excluded as out of scope, according to the inclusion and exclusion criteria.

The remaining 415 articles were evaluated on the basis of their full text, allowing authors to exclude 195 articles that were not relevant to the topic or contained incomplete data. Finally, 220 articles were included in the review.

2.4. Data collection and analysis

The data extracted from each study included the following:

- the full reference and year of publication
- the research question or main topic
- the technique used (classifier, training and validation datasets)
- the number and type of labels (basic, compound emotions, dimensions, Action Units)
- performance metrics (accuracy, F1 score)
- strengths and limitations of the study

The data were organized in tables to present the basic information for each study. An analysis of variance (ANOVA) was also conducted to compare the performance of the systems according to:

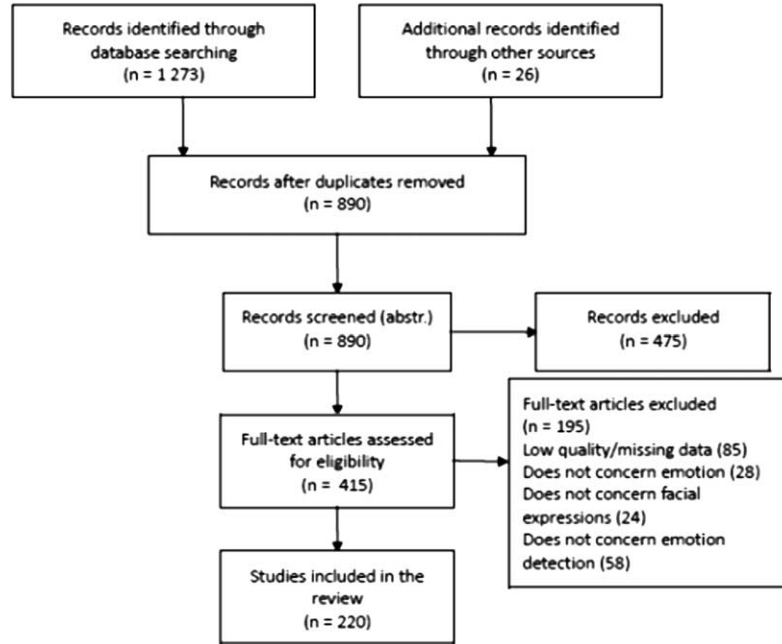- the general type of algorithm (machine learning/deep learning)

Fig. 1. Flow diagram of the documentary search process.

- the type of label (basic/complex/dimensional emotions)
- the type of data used (posed/spontaneous expressions)

Finally, a qualitative and critical analysis was carried out in order to identify the main issues addressed, the strengths and limitations of the studies.

3. Results and main challenges

3.1. Number of publications

First of all, the literature review made it possible to visualize the number of publications year by year (Fig. 2).
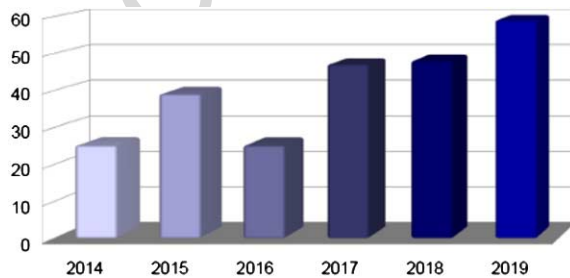

Fig. 2. Number of studies published year by year.

Despite a slight decrease in 2016, we can observe a general increase in the number of publications since 2014, showing a growing interest in this field of research.

3.2. The systems used

Machine learning techniques are mainly based on statistics in order to model functions and derive predictions that will allow classification of the input data.

There are a wide variety of machine learning algorithms (regressions, k-nearest neighbors, AdaBoost, Naïve Bayes, Decision Tree, Random Forest, SVM,...) which rely on different statistical methods but operate according to the same main steps: extraction of features that are handcrafted by data analysts, model estimation, then classification.

Deep Learning has allowed many technical advances, in particular because algorithms, unlike traditional machine learning ones, require very little or no human intervention before learning. However, they require a very large amount of training data to generalize satisfactorily.

Two main types of deep learning algorithms are currently used:

- Convolutional neural networks (CNN), particularly suitable for image processing, consist of a set of interconnected neuron layers which extract

the features of the input images (convolutional filtering), transform them and send them to the layers that will operate the classification

- Recurrent neural networks (RNN) contain connections by which information can circulate in a recurrent (circular) fashion, and thus make it possible to keep information in memory and therefore to process time sequences or elements of context.

The studies analysed reveal a wide variety of algorithms used, with a preponderance of SVM (Support Vector Machine) and convolution networks (CNN).

For convenience, traditional machine learning techniques (SVM, K-Nearest Neighbors, etc.), as opposed to deep learning algorithms, will be referred to as "machine learning" in the remainder of the text.

Deep learning algorithms (on the right in Fig. 3) account for only 42.7% of the total, but are increasingly used over time, becoming the majority from 2018 onwards (Fig. 4), reflecting a constant evolution of techniques.
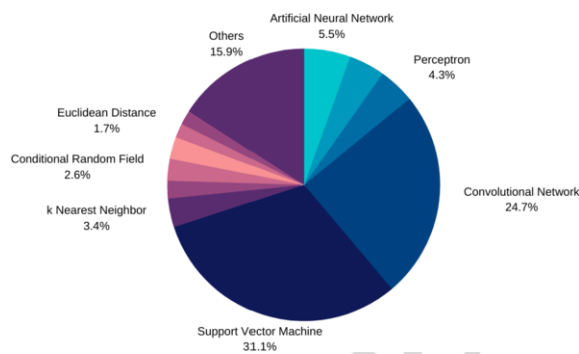

Fig. 3. Type of classifiers used in the studies.

Studies using machine learning algorithms report a wide variety of feature extraction techniques, with Gabor filters and local patterns (LBP, LDP) being the most widely used.
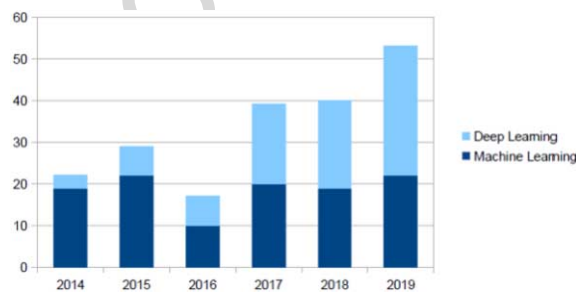

Fig. 4. Evolution over time of the type of algorithms.

Most systems operate in real time and deal with most of the problems such as occlusions, different head orientations or differences in brightness. However, some technical challenges remain.

The question of the temporal dynamics of facial expressions. The temporal dimension is an important element of facial expression that has not been fully explored by research, but is attracting the attention of many teams today. Indeed, as modalities of communication and caused by the interactions of the individual with his physical and social environment, emotions and their expressions fluctuate over time.

Since approaches based on convolutional neural networks cannot account for the temporal dynamics of facial expressions, several researchers [19,23,64,84,121, 140,181,183,190,201] have used other types of networks, including Recurrent Neural Networks (RNN) and Long-short term memory (LSTM), a special type of RNN capable of learning and linking temporal dependencies between several inputs.

Even if it poses some technical problems (especially the question of segmentation of continuous sequences), temporal dynamics can be rich in information and can greatly increase the performance of facial expression recognition systems [29,53,84].

For example, [1] compared a static approach (ORB and SVM) and a dynamic approach (measurement of facial landmarks and Conditional Random Field) and reported a significant increase in classification performance (average accuracy increases from 74.5% to 80.5%)

Temporal dynamics provides a continuous view of facial expressions, allowing them to be observed in terms of change and to align recognition systems with the natural process of emotional onset [29,84,88,118, 190].

According to [118] and [204], the temporal characteristics of facial expressions are also a key factor in distinguishing posed from spontaneous expressions.

Finally, they can also facilitate the detection of micro-expressions, as we will see below.

Recognizing subtle, low-intensity expressions. Micro-expressions are defined as quick, low amplitude facial expressions. They appear when people experience low intensity emotions, begin to feel an emotion (onset), or when emotions are concealed, whether deliberate or unconscious [59,93].

The detection and recognition of subtle expressions, and even more micro-expressions is difficult because of:

- their low amplitude, which requires precise movement or appearance descriptors;
- their very short duration (1/25 to 1/15 of a second for micro-expressions), which means that they can only be captured by very fast video acquisition systems;
- their appearance only in specific high-stakes situations, making them difficult to provoke;
- their potentially fragmented character (e.g. a slight tightening of the lips may be a sign of anger)

Because morphological changes in the face are by definition negligible, many researchers use a dynamic approach to facilitate the detection and classification of micro-expressions [66,88,99,116,118,127,190].

Some studies [202,205] use a motion magnification technique that improves the detection of micro-expressions by increasing the amplitude of facial movements. The most commonly used spotting technique is the threshold technique [101,107,134,202] (for a review, see [133]) and peak detection. [26] use a high-speed acquisition system (averaging 170 fps) as well as motion descriptors based on the absolute image difference between the current frame (potentially the apex of the expression) and a previous frame at half the duration (potential onset). In their review, [133] mention the problem of short non-emotional movements (e.g. head movements or eye blinks) which can be detected as peaks of micro-expressions. To deal with this problem, [134] have developed a threshold technique to eliminate head movements, blinks and changes in gaze.

Beyond micro-expressions, the estimation of the intensity of a facial expression provides valuable information, especially in the interpretation of complex or ambiguous behaviours (e.g. an expression that can have several meanings).

For example, the temporal evolution of the intensity of an expression can make it possible to distinguish between a posed smile and a spontaneous smile [118] or even to relate it to the intensity of a person's emotional experience, or even the level of pain. The structural and temporal dependencies between facial actions and their intensities also make it possible to increase detection performance by "constraining" the detection of one action in relation to another.

[109] and [200] have developed systems for estimating the intensity of Action Units (AU) taking into account the structural dependencies and time dynamics between AUs and their intensities. For example, AU 6 (cheek raiser) and 12 (lip corner puller) are correlated

in terms of presence but also in terms of intensity and dynamics of appearance: AU 12 appears and increases in intensity as the intensity of AU 6 increases.

In their study, [152] estimate the intensity using a Conditional Random Field that takes into account several elements that they call contextual: the subject observed (Who), the changes in appearance of facial expressions (How) and the temporal relationships between the different levels of AU intensity (When).

This consideration of spatio-temporal interactions allowed a significant increase in the performance of their systems.

Several studies have explored techniques to provide intensity estimation in discrete levels, mainly through Action Units whose 5-level intensity is provided in several databases (Bosphorus, DISFA, GFT, UNBC). [11] use a local-global ranking technique that compares the differences in Action Unit intensity between two images of the same person, in order to avoid the problem of morphological differences between individuals. [17] trained in a multi-task CNN that simultaneously detects the occurrence and intensity of all Action Units, taking into account different head orientations. [74] propose the detection of the presence and intensity of AUs with a feature extraction system separated into several ROIs, in order to reduce the problem of AU co-occurrences and to avoid errors due to occlusions of certain parts of the face to be reflected in the final output.

For their part, [219] have developed a regression system (SVR) in order to continuously estimate the intensity of the Action Units. They report that this regression model exceeds the performance of classifiers trained to estimate discrete levels.

3.3. What the algorithms are trying to recognize: The question of theoretical models

The term "emotion", frequently used in literature and in everyday life, is relatively difficult to define from a scientific point of view. Indeed, this multidimensional phenomenon is based on physical, physiological, cognitive and behavioral considerations.

From antiquity to the present day, psychologists, physiologists, anthropologists, philosophers, sociologists, ethologists and neuroscientists have carried out a great deal of work to define and classify human emotions.

Each of them has proposed a different definition.

Stemming from the classical debate between nature and culture, many theoretical controversies currently
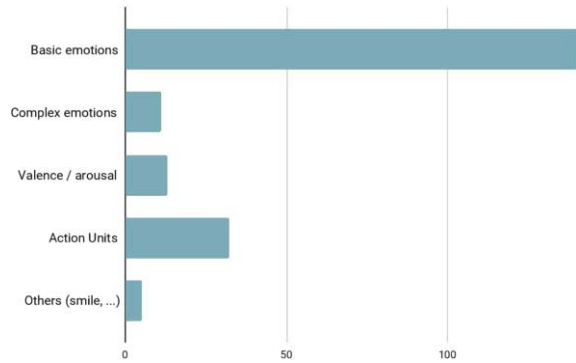
Fig. 5. Type of emotions considered in the studies.

Fig. 6. Type of datasets used in the studies.

remain concerning the nature of emotions and their relationship to facial expressions.

In the field of the AFER, these competing theoretical models of emotions lead to choices of measurement and thus of different data.

The large majority of studies postulate a direct link between emotions and facial expressions. The difference lies in the way these emotions are described, in terms of discrete categories or on a continuum of dimensions.

Figure 5 summarizes the type of emotions taken into account in the various studies.

Discrete emotions. Drawing on Darwin's original work, many authors, such as Ekman, Izard, Tomkins and Plutchik, argue that there is a finite number of basic emotions, which are linked to a biologically determined and pre-wired "affect program" to adapt to the demands of the environment, with a dual function of adaptation and social communication.

Each of these basic emotions would be expressed according to a prototypical and specific pattern of facial expression, which would vary little or not at all depending on the culture or context.

Proponents of this theory have proposed different lists of emotions that more or less overlap. Ekman, the most emblematic contemporary representative of this point of view, considers 7 basic emotions: joy, fear, surprise, sadness, anger, disgust and contempt. As can be seen in Fig. 6, a very large majority of studies (94.2%) focus on basic discrete emotions (see for example [1,6,9,10,80,149,167,168,217,226]).

These are inferred from images annotated in terms of emotion labels or from the detection of Action Units [11,73,75,94,129,144,180], see also [224] for a review.

Based on the assumption that Action Units are very rarely observed in isolation, [63,177] and [219] have taken into account co-occurrences and mutual exclu-
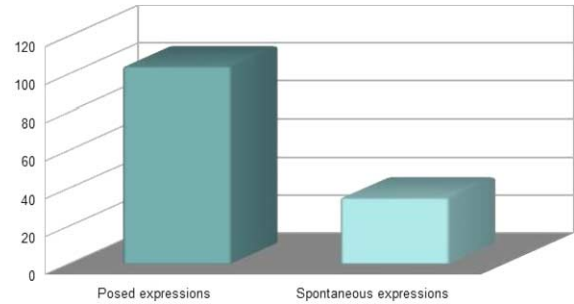
sions (anatomical impossibility, such as opening and closing the mouth) between AUs. They developed a correlation matrix to increase the accuracy of detection, and to model changes in the appearance of some AUs when they occur in combination with another (non-additive effect).

According to the discrete emotion theory, so-called "complex" or "compound" emotions would come from a combination of two or more basic emotions. They can appear either as a juxtaposition of basic emotions (e.g. "happily surprised", "sadly angry"), or they can be combined, like a mixture of colours, to form new emotions. For example, in the model of 141, the 8 primary emotions can combine to generate second-order emotions (e.g. disgust + anger = contempt).

In this model, there are also other emotions, sometimes called derived emotions, which are represented in the form of discrete labels belonging to the family of basic emotions and differing mainly in terms of intensity, such as rage, fury, irritation, etc. are related to anger. There are still very few studies on complex or compound emotions, mainly due to the lack of theoretical consensus on the existence of specific facial expressions for this type of emotion.

[180] used Plutchik's theoretical model to develop a system trained to detect Action Units (AU), then implemented AU combinations forming the expressions of the 6 basic emotions, which allows combinations corresponding to secondary emotions (notably awe, a combination of fear and surprise).

In the same vein, [111,112] revisited the databases annotated in basic emotions (notably JAFFE, CK+, DISFA and IMED) by applying a system called "fuzzy inference engine", giving a multilabel classification in order to represent compound emotions.

[76] compared the recognition performance of their system on 3 databases of spontaneous expressions depicting the 6 basic emotions as well as frustration and fun. With the aim of determining the characteristics al-

lowing the identification of 18 emotions (including the 6 basic emotions and 12 complex emotions, such as interest, pride, disappointment, shame, etc.), [2] created a system allowing the joint detection of facial expressions and head movements.

Dimensions. Sometimes called continuous, the dimensional model suggests that emotions are not discrete entities, but that they are located on a continuum organized according to 2 or 3 dimensions: valence/arousal, approach/avoidance and dominance.

The valence determines the positive or negative aspect of the emotional experience, the arousal describes the degree of excitement (more or less related to the intensity of the emotion).

The notion of dominance makes it possible to distinguish between emotions related to submission (lack of control) and dominance (high potential for control, such as anger). Each emotion can therefore be represented by a point on a two- or three-dimensional space, allowing more nuances than with the categorical model.

[4] have developed a mixed system allowing the detection of Action Units as well as a regression in terms of valence and arousal. The arousal is determined from the degree of intensity of each Action Unit. The authors do not provide information on the model that allowed them to determine the valence.

[85,126,207] also used regression and valence/arousal labelled databases (RECOLA, Amigos individual and groupDB, AffectNet) to represent 3D facial expressions in two-dimensional space. 162] used Russell's circumplex model to determine the links between discrete emotions classified from facial expressions (grouped in 4 classes: happy, surprise, neutral and negative) and the dimensional self-reporting of subjects in human-computer interaction situations (task or computer game).

In the same way, some studies [33,43,57,118] are dedicated to smile detection, without interpretation in terms of emotion categories.

The question of context. Psycho-constructionist models (for a review, see [15]), on the other hand, postulate that emotions are changing and dynamic phenomena, and that they depend on many individual factors but also on the environment with which people interact daily.

Indeed, facial expressions always manifest themselves to serve a particular purpose in a given social context, and it is this context that will determine the correct interpretation of an expression.

For example, depending on the situation, a smile may express joy, but also pride, embarrassment, or a polite greeting; a frown may indicate anger in some cases, and in others, strong concentration or confusion.

Anchored in the concrete, these theories help to highlight and explain the intra- and inter-individual variability in the expression of emotions.

Since 2014, only 8 studies have attempted to explore the context in which facial expressions appear.

Context is commonly defined as all the information resulting from the situation in which an event or phenomenon occurs, in this case an emotion.

It is therefore a broad and complex notion, which can be the subject of many definitions according to the authors: cultural origin [40], physical context and events [13,22], purpose of social interaction [157], or context including several parameters.

[89] approaches the notion of context, but as a source of variation to be eliminated. Rather than describing and formalizing context, it seeks to model the specificities of emotional expressions (and their temporal transitions) in order to isolate them from the multiple sources of variation that can be confusing.

[40] explored the effects of culture using 3 databases of subjects from different cultures (CK+, JAFFE and Bosphorus). They performed a cross-database validation to compare recognition performance, and concluded that performance was lower in cases where the training databases involved a different culture (subjects from different ethnies) from the validation database.

[13] have developed a system for interpreting the temporal history of facial expressions in the context of the events that caused the emotion expressed.

In other words, they consider that the temporal dynamic of a facial expression is determined by different types of socially significant events (type of emotion that may be provoked by the event, e.g. funny or sad) and different intensities.

[22] have also used an event-based method. They developed post-processing classifiers to identify the emotional context during a time window. These algorithms allowed them to predict sudden or progressive changes in emotional expressions through event detection.

[157] focused on the link between a virtual agent and a human, in particular on the purpose of the interaction: competition, information, education, collaboration, negotiation, guidance or purely social.

[132] were based on a broad definition of the context. They used formalisms to describe situations in terms of environment, social norms, salient objects and behaviours, and proposed a method for the detection of

Fig. 7. Sample images from Aff-Wild dataset (spontaneous expressions, dimensions).

8 types of complex social events (interview, wedding, sports event, party, dinner, birthday, class and night-club) with an accuracy between 37.9 and 51.8%.

On the basis of multimodal data from digital interviews, [182] developed a system for detecting emotions that takes into account the direction of the gaze, prosody, speech, facial expressions and some contextual information: the interaction phase (introduction, negative, positive), the sex of the subject and the content of the previous question.

More recently, [158] developed an emotion recognition system based on contextual information (presence and type of social partners, activities, temperature, physiological state, location and time of day) provided by 32 participants through a mobile application. This information led to individual, general and gender-based models that allowed to weight the prediction of emotions. The authors showed that contextual information is relevant, especially at the individual level, and increases detection performance.

3.4. The data used

There are currently very many databases of facial expressions used for the training and the validation of the algorithms of AFER.

They differ in their characteristics: number of images or sequences, static or dynamic data (videos), number of categories of emotions, posed or spontaneous expressions, real or controlled conditions, etc.

Except certain databases which are composed of images coming from the Internet (e.g. AffectNet, SFEW) or collected in natural situations (Aff-Wild, Fig. 7), the great majority of the databases are constituted in a controlled environment and under controlled conditions of luminosity and background. As can be seen in Fig. 6, more than 74% of the studies use databases with posed expressions, with a clear preference for CK+ (Fig. 8) and JAFFE, which concern 42% and 29% of the studies respectively.

Table 2 below provides a brief description of the most used datasets in terms of type of support (video or still image), numer of subjects and frames, type of expression (posed or spontaneous), and labels.

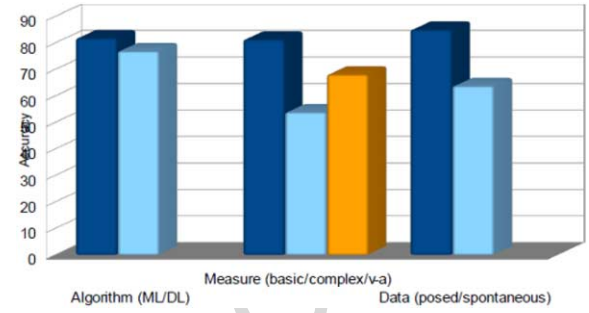Fig. 8. Sample images from CK+ dataset (posed basic emotion expressions.

Fig. 9. Average accuracy by type of algorithm, emotion measured and type of data.

3.5. Systems performance

Most of the studies analysed present their classification performance in terms of accuracy, which corresponds to the rate of correct predictions in relation to the total number of examples:

$$\text{Accuracy} = \frac{\text{true positives} + \text{true negatives}}{\text{total examples}}.$$

There has been a general increase in performance over time: the overall accuracy increased from 83.54% in 2014 to 85.38% in 2019.

An ANOVA was performed to determine the effect of system type (machine learning/deep learning), type of emotion (basic/complex/dimension) and type of expression (posed/spontaneous) on accuracy (Fig. 9). The results are shown in Fig. 5.

Contrary to what is sometimes argued, the ANOVA revealed no significant effect of the type of algorithm. Deep learning systems ($M = 76.92$, $SD = 20.73$, $n = 50$) do not outperform the conventional machine learning approaches ($M = 81.81$, $SD = 17.72$, $n = 80$, $F(1, 130) = 2.229$, $p = 0.13764$) in terms of average classification accuracy.

Systems trained to recognize basic emotions ($M = 81.2$, $SD = 18.43$, $n = 120$) obtain the best results, followed by the valence/arousal dimensions ($M = 70.1$, $SD = 15.69$, $n = 4$). Complex emotions are those that give the worst results and seem to be the most difficult to recognize ($M = 53.86$, $SD = 17.82$,

Table 2
Description of the most used datasets

Dataset	Type	Subjects	Frames or sequences	Posed/spontaneous	FACS coded	Emotion labels
AFEW	V	330	1747	P	N	7
AffectNet	I	450000	1450000	S	N	8 + V/A
Aff-Wild	V	200	248	S	N	V/A
Bosphorus	I (3D)	105	4666	P	Y+intensity	6
BP4D	V (3D)	41	368036	S	Y	–
BU-3DFE	I (3D)	100	5000	P	N	6+intensity
BU-4DFE	V (3D)	101	60600	P	N	6
CASME	V	35	195	S	Y	8 micro
CASME II	V	26	247	S	Y	5 micro
CFEE	I	230	6670	P	Y	22
CK+	V	123	593	S+P	Y	7
CMU Multi-PIE	I	337	750000	P	N	6
DISFA	V	27	130000	S	Y+intensity	–
Enterface	V	42	1166	S+P	N	6
FABO	V	23	*nc*	P	N	10+gestures
FER 2013	I	*nc*	35887	S+P	N	7
GEMEP-FERA	V	10	289	P	N	5
JAFFE	I	10	213	P	N	7
KDEF	I	70	4900	P	N	7
MMI	I+V	75	1500	P	Y	7
MUG	V	52	1462	P	N	6
OULU-CASIA	I+V	80	2880	P	N	6
RaFD	I	29672	29672	S	N	12
SAMM	V	32	159	S	Y	8 micro
SAVEE	V	4	480	P	N	7+audio
SFEW	I	68	663	S	N	7
SMIC	V	16	77	S	N	3 micro

$n = 6$). The type of emotion therefore has a significant effect on the accuracy: $F(2, 130) = 7.363$, $p = 0.00091$).

The nature of the data used also has a significant effect on recognition performance: systems validated on posed expression databases ($M = 85.04$, $SD = 14.42$, $n = 95$) obtain a better average accuracy than those that use spontaneous expressions ($M = 63.75$, $SD = 22.67$, $n = 35$, $F(1, 130) = 41.239$, $p = 0.00001$).

Indeed, the 3 datasets that generated the best performances are made up of posed expressions: FABO (average accuracy 95.51), BU-4DFE (90.94) and CK+ (88.06).

On the other hand, the most challenging datasets contain spontaneous expressions: Enterface (average accuracy 66.51), SFEW (67.51) and the two micro-expressions datasets CASME II (63.59) and SMIC (61.03).

However, given the methodological differences between studies, it is difficult to compare performance between systems in a precise manner, and these results, which allow general trends to be identified, should be treated with caution.

In addition, average accuracy is not a sufficient metric to assess an algorithm: it is also important to know the performance of a system according to the categories of emotions.

About half of the studies analyzed in this review have proposed confusion matrix category by category. We analyzed these matrices and noted the categories of emotions that were most often confused.

Overall, it turns out that negative valence emotions (especially fear and anger) are the most difficult to spot and are often confused with each other.

For example, the confusion between anger and sadness is the most frequent in 44% of studies, the one between fear and surprise is the most frequent in 32% of studies. While the confusion between fear and surprise is easily explained by the proximity of these two facial expressions, the one between anger and sadness

is more difficult to interpret. It is possible that it is the presence of a frown, present in both expressions, which is confusing for the algorithms.

More unexpectedly, a quarter of studies report frequent confusion between joy and fear. In most databases used with posed expressions, fear is expressed with the lips stretched, which could be mistaken for a smile.

3.6. Summary of the studies

Table 3 provides an overview of the studies included in the review, as well as a description of the classifier and databases used, the type of emotion, and the performance metrics.

4. Discussion and perspectives

A comprehensive review of recent literature was conducted, analyzing a total of 220 original studies.

The automatic recognition of emotional expressions is a rapidly evolving field, and we cannot guarantee that we have taken into account all the articles, especially those written in other languages or published during the writing of this review.

However, this work has made it possible to cover a large part of the research carried out in the field of the AFER. It also proposes a qualitative and thematic approach, which allows a better understanding of this complex field, but also to have an overall view of the issues addressed or to be developed in future research.

Our results suggest a large and steadily increasing number of publications between 2014 and 2019. Automated facial expression recognition is therefore currently attracting a great deal of interest.

The review has also highlighted a wide variety of techniques and models used: while deep learning is widely used, especially in recent studies, conventional machine learning techniques are still in the majority, and involve a wide range of algorithms (SVM, Random Forest, Naive Bayesian network, Decision Tree, etc.).

Despite the generally equivalent performance of conventional techniques and more recent approaches (deep learning), research has made it possible to develop increasingly sophisticated and high-performance models (majority use of deep learning from 2018 onwards, increase in average accuracy over time).

However, the vast majority of these performances are still tested in a very standardised framework, and

the data used for validation still concern almost exclusively posed expressions of a limited number of basic emotions, obtained under conditions of luminosity, pose and homogeneous background.

Performance drops significantly when systems are confronted with more natural data, i.e. data obtained in real-life conditions where emotions arise (everyday situations, natural conversations). This is due to the fact that emotion is a multifaceted phenomenon, which is not limited to a few prototypical facial expressions, and which is still subject to many theoretical controversies and therefore remains difficult to model in its complexity. This problem of definition of the concept of emotion leads to several questions, intimately related to each other, which will have to be tackled in the research in the field of the AFER.

Beyond the 6 basic emotions. Current automated systems, despite their increasing performance, use and require a cross-cultural and universal conception of facial expressions, i.e. the ability to translate them into easily interpretable data.

The classical theory of basic emotions, by its simplified and categorized description, makes it possible to give intelligibility to this very heterogeneous phenomenon, but it is now outdated:

> [. . .] the six categories of emotions have no use for the majority of everyday applications. This simplification of the task, while serving us well in the early days, needs to change significantly. (Gunes et al., 2016, p. 3 [61]).

Recent research has shown that there are many more categories of emotions, and that they are separated by more nuanced boundaries than conventional thinking would suggest, as their emotional characteristics (such as valence or arousal) may in some cases be similar.

It seems therefore necessary, in order to get as close as possible to reality, to develop theoretical models and to train our systems to recognize a wider range of expressions [21,36], relating to complex or compound emotions, as well as to cognitive states [72,117] or even non-emotional facial expressions, which appear more frequently than expressions related to emotions in the context of interactions between individuals.

Taking into account the variability of emotional states and facial expressions. Directly linked to the problem of categories of emotions, the question of the variability of expressions is essential to account for the richness of human emotional life.

In the model of basic emotions, the variability of facial expressions is often put down to methodologi-

Table 3
Brief description of the studies

Reference	System	Databases	Labels	Accuracy	F1 score
[1] Acevedo et al. (2016)	CRF-SVM	CK+	8 basic emotions	79.1–80.5	–
[3] Afdhal et al. (2014)	Wavelet Transform	CK	7 basic emotions	93.05	–
[6] Ali et al. (2016)	SVM	CK	7 basic emotions	93.41	–
[10] Bahreini et al. (2019)	Fuzzy logic rules	CK+	7 basic emotions	83.2	–
[12] Bandrabur et al. (2017)	MLP-SVM	CK+	7 basic emotions	94.28	–
[18] Battini-Sönmez & Cangelosi (2017)	CNN	CK+	7 basic emotions	99.38	–
[30] Canário & Oliveira (2015)	CNN	CK+	7 basic emotions	90	–
[39] Cui et al. (2016)	CNN	CK+	6 basic emotions	98.1	–
[79] Jain et al. (2016)	SVM	CK	6 basic emotions	97.2	–
[121] Mattela & Gupta (2018)	SVM	CK+	7 basic emotions	92.5	–
[143] Quan et al. (2014)	SVM	CK	6 basic emotions	88.32	–
[155] Said et al. (2015)	Wavelet Network	CK+	7 basic emotions	91.26	–
[178] Sönmez & Albayrak (2016)	Sparse representation	CK+	7 basic emotions	80	–
[188] Sun & Akansu (2014)	HMM	CK	6 basic emotions	86.67	–
[204] Wu et al. (2017)	CNN	CK+	Action Units	91.44	–
[215] Zeng et al. (2018)	ANN	CK+	6 basic emotions	95.79	–
[23] Bhandari et al. (2018)	CNN	JAFFE	6 basic emotions	69.35	–
[42] Das et al. (2014)	SVM	JAFFE	7 basic emotions	86.3	–
[48] Dou et al. (2016)	SVM	JAFFE	7 basic emotions	76.5	–
[80] Jameel et al. (2015)	Crisp logic ANN	JAFFE	7 basic emotions	90.47–94.04	–
[124] Mehta & Jadhav (2016)	Euclidean Distance	JAFFE	7 basic emotions	93.57	–
[128] Naik & Jagannath (2018)	ELM	JAFFE	6 basic emotions	90.44	–
[135] Patil et al. (2016)	Euclidean Distance	JAFFE	7 basic emotions	85	–
[137] Perikos et al. (2018)	ANFIS	JAFFE	7 basic emotions	90	0.89
[138] Perikos et al. (2014)	MLP	JAFFE	7 basic emotions	76.7	0.778
[140] Piparsaniyan & al (2014)	Naive Bayesian	JAFFE	7 basic emotions	96.73	–
[150] Richhariya & Gupta (2019)	SVM	JAFFE	6 basic emotions	77.13	–
[201] Wang et al. (2014)	SVM	JAFFE	7 basic emotions	84.9	–
[5] Alam et al. (2018)	S-DSRN	JAFFE-CK+	7 basic emotions	94.62	–
[14] Barman & Dutta (2017)	MLP NARX	JAFFE-CK+	5 basic emotions	92–97.6	–
[34] Chowdhury & Alam (2014)	Naive Bayesian	JAFFE-CK+	6 basic emotions	81.5–87.2	–
[65] Happy & Routray (2015)	SVM	JAFFE-CK+	6 basic emotions	94.14	–
[98] Lajevardi (2014)	GFSSIM	JAFFE-CK+	6 basic emotions	89.9	–
[113] Liu et al. (2017)	SVM	CK+-JAFFE	6 basic emotions	79	–
[119] Mansouri–Benssassi & Ye (2018)	Unsup. CNN	JAFFE-CK+	6 basic emotions	88.5	–
[122] Mayya et al. (2016)	CNN	JAFFE-CK+	6 basic emotions	97.07	0.9538
[125] Miao et al. (2019)	CNN	JAFFE-CK+	6 basic emotions	96	–
[139] Perikos et al. (2015)	SVM-MLP	JAFFE-CK+	7 basic emotions	93.2	–
[153] Saabni (2015)	ANN	JAFFE-CK+	7 basic emotions	88.84	–
[159] Salmam et al. (2016)	DT	JAFFE-CK+	7 basic emotions	89.9	–
[167] Shan et al. (2017)	CNN-kNN	JAFFE-CK+	6 basic emotions	71.19–78.52	–
[171] Shi et al. (2019)	SVM	CK+-JAFFE	6 basic emotions	96.31	–
[184] Suja et al. (2014)	ANN-kNN	JAFFE-CK	7 basic emotions	86.5	–

Table 3

(Continued)

Reference	System	Databases	Labels	Accuracy	F1 score
[193] Uçar et al. (2016)	ELM	JAFFE-CK+	6 basic emotions	94.91	–
[199] Vishnu et al. (2019)	SVM-MLP-EML	JAFFE-CK+	6 basic emotions	97.25	–
[206] Xiao et al. (2019)	DBN	JAFFE-CK+	6 basic emotions	86.37	–
[208] Xie & Hu (2019)	CNN	JAFFE-CK+	6 basic emotions	94.1	–
[4] Al-Darraji et al. (2017)	CNN	RaFD-CK+	AU + 7 basic emotions	90.85	–
[7] Ashir et al. (2019)	SVM	JAFFE-MMI-CK+	7 basic emotions	94.3	–
[37] Cornejo et al. (2015)	SVM-kNN	JAFFE-MUG-CK+	7 basic emotions	93.32–93.49	–
[40] da Silva & Pedrini (2015)	SVM	JAFFE-MUG-Bosphorus-CK+	6 basic emotions	70.6	–
[54] Farajzadeh et al. (2014)	MPC-SVM-ANN	JAFFE-CK+-TFEID	7 basic emotions	87.2–93.3	–
[63] Hao et al. (2018)	Naive Bayesian	CK+-Semaine-BP4D	Action Units	–	0.6578–0.833
[67] Hasani & Mahoor (2017)	CNN-CRF	CK+-MMI-FERA	6 basic emotions	53.33–93.04	–
[111] Liliana et al. (2017)	SVM-CRF	CK+-private db	12 emotions	86.9	–
[112] Liliana et al. (2019)	Fuzzy logic rules	CK+-JAFFE-DISFA-IMED	6 emotions multi-label (mixed)	90	–
[114] Liu et al. (2017)	SVM	CK+-SFEW-RAFB-RAF compound	6–11 emotions	45.2–96.59	–
[115] Lo Presti & La Cascia (2017)	K votes	CK+-Painful	7 basic emotions + pain	90.66	–
[154] Sadeghi et al. (2019)	SVM	CK+-SFEW-MMI-RaFD	7 basic emotions	72.41	–
[156] Sajjad et al. (2018)	SVM	JAFFE-MMI-CK+	7 basic emotions	97.86	–
[160] Salmam et al. (2019)	CNN	JAFFE-Oulu-Casia-CK+	7 basic emotions	76.78	–
[162] Samara et al. (2019)	SVM	CK+-KDEF-DEAP	8 emotions + v/a	80.75–96.94	–
[165] Sen et al. (2019)	SVM	CK+-MUG-JAFFE	6 basic emotions	91.11	–
[170] Sharma et al. (2019)	SVM	MMI-JAFFE-CK+	5 basic emotions	91	–
[211] Xue & Gertner (2014)	SVM	JAFFE-AT&T	7 basic emotions	71.5	–
[212] Yaddaden et al. (2018)	SVM	JAFFE-KDEF-RaFD	6 basic emotions	90.61	–
[218] Zhang et al. (2015)	SVM	CK+-MMI-FERA2013	7 basic emotions	93.6	–
[222] Zhao et al. (2015)	SVM	CK+-GFT-BP4D	Action Units	–	0.47
[226] Zia et al. (2015)	kNN	JAFFE-CK+-Feedtum	7 basic emotions	64.52	–
[100] Li et al. (2018)	SVM	CASME I&II	5 micro-expressions	76.1	0.454
[101] Li et al. (2018)	CNN-SVM	CASME I&II	5 micro-expressions	56.77	–
[202] Wang et al. (2017)	SVM	CASME II	5 micro-expressions	75.3	–
[225] Zhu & Chen (2019)	CNN	CASME II	3 micro-expressions	–	0.85
[26] Borza et al. (2017)	Center displacement	CASME II-SMIC	3 micro-expressions	79.5	0.79
[66] Happy & Routray (2018)	SVM	CASME-SMIC	5 micro-expressions	58.49	–
[99] Le Ngo et al. (2017)	SVM	CASME II-SMIC	5 micro-expressions	53	0.55
[102] Li et al. (2019)	CNN	SMIC-CASME I&II	5 micro-expressions	56.34	–
[103] Li et al. (2018)	CNN	CASME I&II-SMIC	5 micro-expressions	56.35	–
[106] Li et al. (2018)	SVM	SMIC-CASME II	5 micro-expressions	42.42	–
[116] Lu et al. (2015)	HMM-RF	SMIC-CASME II	4 micro-expressions	71.38	–
[127] Muna et al. (2017)	MLP-SVM	CASME II-SMIC	5 micro-expressions	85.07	–

Table 3

(Continued)

Reference	System	Databases	Labels	Accuracy	F1 score
[205] Xia et al. (2018)	CNN	CASME-SMIC	3 micro-expressions	62.03	–
[223] Zhao & Xu (2018)	SVM	CASME II-SMIC	5 micro-expressions	66.07	–
[88] Khor et al. (2018)	CNN-LSTM	CASME II-SAMM	7 micro-expressions	45.4	0.5
[35] Corneanu et al. (2018)	CNN	DISFA-BP4D	Action Units	57.65	–
[74] Hupont & Chetouani (2017)	SVM	DISFA-UNBC-BP4D	Action Units	–	0.55
[83] Jiang et al. (2014)	SVM	DISFA-FERA	Action Units	71.5	–
[109] Li et al. (2015)	SVM-DBN	DISFA	Action Units	81.06–86.6	–
[144] Racoviteanu et al. (2019)	CNN	DISFA-SPOS	Action Units	–	0.70
[192] Tran et al. (2017)	CNN	DISFA-FERA2015	Action Units	53.5	–
[200] Walecki et al. (2017)	CRF-CNN	DISFA-FERA2015	Action Units	–	–
[147] Rashid (2016)	DT-CNN-MLP	Bosphorus-JAFFE	7 basic emotions	84.44–92.24	–
[219] Zhang et al. (2015)	ANN-SVR	Bosphorus	AU + 6 basic emotions	92.2	–
[70] Hossain et al. (2019a)	CNN-ELM-SVM	Bigdata-Enterface	6 basic emotions	93.15	–
[71] Hossain et al. (2019b)	CNN-SVM	RML-Enterface	6 basic emotions	84.95	–
[149] Reddy et al. (2018)	kNN	Enterface	6 basic emotions	42.4	–
[164] Selvaraj et al. (2019)	SVM	Enterface-AFEW	6 basic emotions	94.62	–
[195] Vedantham et al. (2019)	ANN	Enterface-CK	6 basic emotions	84.16	–
[9] Azazi et al. (2014)	SVM/ANN/kNN	BU-3DFE	6 basic emotions	61.69–79.36	–
[60] Gui et al. (2017)	CNN	BU-3DFE-BU4DFE-CK+-MUG	7 basic emotions	83	–
[77] Huynh et al. (2016)	CNN	BU-3DFE	6 basic emotions	89.35	–
[82] Jan & Meng (2015)	SVM	BU-3DFE	7 basic emotions	90.04	–
[198] Vieriu et al. (2015)	RF	BU-3DFE-Bosphorus	7 basic emotions	69.8	–
[210] Xudong Yang et al. (2015)	SVM	BU-3DFE	6 basic emotions	83	–
[214] Yurtkan & Demirel (2014)	SVM	BU-3DFE	6 basic emotions	83.33	–
[19] Ben Amor et al. (2014)	Random Forest	BU-4DFE	6 basic emotions	93.21	–
[32] Chen et al. (2019)	CNN	BU-4DFE-CK+	6 basic emotions	77.4–95.4	–
[75] Hussain et al. (2017)	SVM-kNN	BU-4DFE	Action Units + happy/sad	91.53	–
[105] Li et al. (2018)	SVM-CNN	BU-4DFE	6 basic emotions	92.22	–
[168] Shao et al. (2015)	CRF	BU-4DFE-private db	6 basic emotions	87.63	–
[27] Boubenna & Lee (2018)	LDA-kNN	RaFD	8 basic emotions	90–98.67	–
[172] Shim et al. (2018)	CNN	RaFD	7 basic emotions	98.6	–
[28] Boughrara et al. (2016)	MLP	FERA 2013-CK+	5 basic emotions	89.47	–
[197] Verma et al. (2018)	SVM	FERA2013-ISED-Oulu-Casia-MMI	7 basic emotions	94.12	–
[217] Zhang et al. (2019)	CNN	FERA2013-RaFD	7 basic emotions	84.24	–
[130] Nguyen et al. (2018)	PathNet Transfer learning	SAVEE-Enterface	6 basic emotions	90.62	–
[145] Rahdari et al. (2019)	NB-RNN-BG-RF-SMO	SAVEE-RML-enterface	6 basic emotions	62.80–98.33	–
[221] Zhao et al. (2018)	CNN	SAVEE-AFEW-CK+	7 basic emotions	68.33	–
[11] Baltrusaitis et al. (2017)	SVM	BP4D-DISFA	Action Units	63	–
[17] Batista et al. (2017)	CNN	BP4D	Action Units	95.27	0.506
[57] Girard et al. (2015)	SVM	BP4D-Spectrum	AU + smile	–	–

Table 3

(Continued)

Reference	System	Databases	Labels	Accuracy	F1 score
[16]] Barros et al. (2015)	CNN	FABO	10 emotions	91.3	–
[187] Sun et al. (2018)	CNN-BLSTM-RNN	FABO	6 basic emotions + 4 cognitive states	99.71	–
[25] Borah & Konwar (2014)	ANN	FERET	7 basic emotions	85.5	–
[129] Natarajan & Muthuswamy (2015)	AdaBoost	FERET-CK+	AU + 5 basic emotions	94	–
[33] Chen et al. (2017)	SVM	UNBC	AU+pain	85.8	–
[148] Rathee & Ganotra (2018)	SVM	UNBC-DISFA	Action Units	55.81	–
[152] Rudovic et al. (2015)	CRF	UNBC-DISFA	Action Units + pain	–	0.35
[91] Kong (2019)	CNN	ORL-MultiPIE	6 basic emotions	91.28	–
[51] Eleftheriadis et al. (2014)	VC-GPM	MultiPIE-LFPW	6 basic emotions	87.92	–
[52] Fan et al. (2018)	CNN	AFEW-RaFD	7 basic emotions	57.75	–
[86] Kacem et al. (2018)	SVM	AFEW	7 basic emotions	38.38	–
[194] Van Huynh et al. (2019)	CNN	AFEW	7 basic emotions	51.96	–
[68] He et al. (2017)	LSTM-RNN	FERA 2017	Action Units	73.5	–
[108] Li et al. (2017)	SVM-RF-CNN	FERA 2017-BP4D	Action Units	69.4	0.498
[110] Lifkooee et al. (2018)	CNN	FERA 2017	Action Units	69	–
[179] Soysal et al. (2017)	CNN-ANN	FERA 2017	Action Units	69.7	–
[191] Tang et al. (2017)	ANN	FERA 2017	Action Units	77.8	0.574
[58] Gonzalez et al. (2015)	SVM	MMI	Action Units	64.7	–
[166] Senechal et al. (2014)	SVM	MMI-Bosphorus-CK+	Action Units + 6 basic emotions	91	–
[190] Talukder et al. (2016)	SVM	SMIC-MMI	3 micro-expressions	41.11	–
[169] Sharma et al. (2019)	MLP-ANN-SVM	MUG-JAFFE	7 basic emotions	95.5	–
[185] Sultan Zia et al. (2018)	DWMV	MUG-Feedtum-JAFFE-CK+	7 basic emotions	48.98	–
[59] Grobova et al. (2017)	SVM-RF	Private db	Action Units + micro-expression sad	73.86–95.72	–
[117] Ma et al. (2017)	SVM	Private db	AU + mental states	37.75	–
[186] Sumi & Ueda (2016)	DT	Private db	5 emotions +15 micro-expressions	75	–
[2] Adams et al. (2015)	Euclidean Distance	EESS	AU + 18 emotions	47	0.33
[8] Avots et al. (2019)	SVM-CNN	SAVEE-enterface-RML-AFEW	7 basic emotions	74.38	–
[20] Ben Tanfous et al. (2019)	LSTM-SVM	CK+-Oulu-Casia-CASME II-MSR Action 3D-Florence 3D-VTKinect	6 basic emotions	68.56–98.49	–
[24] Bishay & Patras (2017)	CNN-MLP-RNN	UNBC-DISFA-BP4D-Semaine	Action Units	76.32	0.627
[31] Chavan & Kulkarni (2019)	CNN	Kaggle	7 basic emotions	61	–
[46] Ding et al. (2017)	CNN	Oulu-Casia-TFD-SFEW-CK+	6 basic emotions	82.14	–
[47] Dino & Abdulrazzaq (2019)	CNN	FaceNetExpNet-CK+-Oulu-Casia-TFD-SFEW	8 basic emotions	96.8	–
[55] Fathallah et al. (2017)	CNN	MUG-RaFD-KDEF-CK+	6 basic emotions	93.42	–

Table 3

(Continued)

Reference	System	Databases	Labels	Accuracy	F1 score
[[64] Happy et al. (2019)	CNN	CK+-RaFD-FER2013-Lifespan	7 basic emotions	73.58–99.35	–
[72] Hu et al. (2018)	CNN	LSEMSW-Oulu-casia	13 emotions + cognitive states	36.72	–
[73] Hung et al. (2019)	CNN Transfer learning	JAFFE-KDEF-FER2013	AU + 6 basic emotions	84.95	–
[76] Hussein et al. (2017)	Sparse coder	VDMFP-MMI-BINED	6 basic emotions	–	–
[84] Jiang et al. (2014)	SVM	UNBC-SAL-FERA-MMI-CK+-Semaine	Action Units	67.7–86.5	–
[85] Jin et al. (2019)	SVR	AVEC 2015	valence-arousal	–	–
[92] Kukla & Nowak (2015)	ANN	KDEF-CK+	7 basic emotions	76	–
[93] Kulkarni et al. (2018)	CNN	SASE-FE-CK+-Oulu-Casia-BP4D	AU + 6 basic emotions	70.2	0.48
[94] Kulkarni & Bagal (2015)	SVM	FACES	7 basic emotions	88.8	–
[95] Kumar et al. (2019)	Semi-supervised SVM	CK+-MMI-Oulu-Casia-Manhob-realtime	6 basic emotions	90.05	0.68
[104] Li & Deng (2019)	CNN	RAF-ML-CK+-JAFFE-SFEW-MMI	6 emotions multi-label (mixed)	76.84	–
[118] Mandal & Ouarti (2017)	SVM	UVA Nemo Smile	Smile	78.1	–
[126] Mou et al. (2019)	SVM-LSTM	Amigos	Valence-arousal	–	0.70
[136] Peng et al. (2018)	CNN	Oulu-Casia-JAFFE-MUG-CK+-SAMM-CASME	5 basic emotions	56.34	–
[142] Prasada et al. (2019)	SVM	Indian Face – Berlin speech	6 basic emotions	77	–
[146] Rao et al. (2019)	CNN	FI-EmotionROI-IAP-Artphoto-Abstract	8 emotions	75.46	–
[173] Siddiqi (2018)	HMM	Youtube db	6 basic emotions	95	–
[175] Siritanawan et al. (2014)	Euclidean Distance	MyFace-CK+	7 basic emotions	35	–
[176] Slimani et al. (2019)	CNN	CFEE	22 compound emotions	47.73	–
[177] Song et al. (2015)	DBN	FERA2015-DISFA-CK+	Action Units	86.85	0.545
[180] Starostenko et al. (2019)	Naive Bayesian	Yale B-CAS PEAL	AU + 6 basic emotions + awe	96.75	–
[181] Stöckli et al. (2018)	FACET-AFFDEX	WSEFET-ADFES-RaFD	6 basic emotions	72–95	–
[189] Talele et al. (2016)	RNN	JFED-TFEID-CK	7 basic emotions	94.86	–
[196] Verburg & Menkovski (2019)	LSTM-RNN	SAMM	Micro-expression detection	–	0.0821
[207] Xiaohua et al. (2019)	RNN	AffectNet	8 emotions + v/a	–	–
[209] Xie et al. (2019)	CNN	JAFFE-TFEID-SFEW-FERA2013-BAUM2i-CK+	6 basic emotions	73.87	–
[213] Yuce et al. (2015)	SVM	Semaine-FERA2013-BP4D-CK+	Action Units	76.62	0.508
[216] Zhang et al. (2019)	CNN-DBN-SVM	BAUM1-RML-MMI	6 basic emotions	67	–
[220] Zhang et al. (2018)	CNN	ExpW-AFLW-CelebA-SFEW-CK+	7 basic emotions + 8 interpersonal traits	77.1	–

cal errors or secondary factors, such as "display rules", which lead individuals to inhibit pre-existing and invariable emotional reactions in all situations.

This classical view is now being replaced by theoretical models that seek to take into account the complexity of emotions emerging in the interaction between the individual and his environment (social and non-social) and facial expressions. Indeed, all observations in ecological environments show that facial expressions are highly variable and dependent on multiple factors.

These sources of variation come from the cognitive and psychological state of individuals as well as from other people and the environment. Variability can be observed at all levels of analysis, both within the same person over time and situations, and between individuals and cultures.

First, there may be several patterns of facial expressions or behaviours for the same emotion.

Indeed, depending on many factors such as their personality, their previous experience, their culture, the different social roles they are led to play but also depending on the situation, individuals do not experience or express their emotions in the same way or at the same intensity.

It is therefore necessary, even if exhaustiveness in this field is illusory, to collect a maximum of natural data in order to predict as many variants of facial expressions as possible. Automatic facial expression recognition systems must also, depending on the application, be adapted to the situation and even to each individual in certain cases.

Similarly, the same facial expression can have several meanings. For example, depending on the situation, a frown may mean anger, disgust, sadness, but also more complex mental states such as confusion or intense reflection.

The notion of context then plays an important role in order to disambiguate facial expressions and interpret them correctly.

Since 2015, this field has been the subject of increasing research interest, with numerous studies and workshops (see for example [62]) devoted to the formalization of context in the field of automatic facial expression recognition.

Although most of the studies we have reviewed are devoted to the analysis of elements of the physical context (location, ambient temperature, etc.), it is the notion of social context – in the sense of interactionality – which, for some authors, provides the most relevant information for the interpretation of facial expressions: the presence of other people, their behaviour, the type of relationship one has with them and the type of interaction that is taking place are determining factors in the experience and expression of emotions.

Observing emotions in all their modalities. In their daily interactions with the physical and social environment, humans use many modalities to express and decode their emotional states. For example, it is known that the arousal dimension is more easily perceived through non-visual modalities, such as vocal prosody and physiological signals [61].

It therefore seems natural and necessary for automated systems to be able to use several modalities, such as coupling audio and visual signals to differentiate emotional expressions from speech induced deformations [89], analysing vocal characteristics or emotional discourse, or even merging more than two modalities in order to increase their accuracy [56,78,131,161].

With the multiplication of portable technologies and sensors (touch, microphones, bio-signals, etc.), it is now easy to set up multimodal databases. The difficulty lies rather in the annotation of these data, which requires a deeper understanding of how humans fuse these modalities together to consistently produce or decipher emotional expressions. Moreover, it is often not feasible to apply the full set of sensors to an individual or individuals in the context of everyday life. It is therefore necessary to choose according to the objective.

Collecting and using spontaneous and natural data. As we have seen, most current systems are based on a paradox: while the theoretical model of basic emotions postulates the existence of a direct link between emotion and facial expression, the majority of systems are trained and validated on the basis of stated expressions, which by definition do not reflect an emotion actually felt.

The quality of the data being a paramount condition to train the ARM systems to recognize in a relevant way the facial expressions [50] [97], it is thus necessary to collect and to annotate a great number of spontaneous expressions, collected under natural conditions. However, obtaining the ground-truth is subject to several problems:

- data annotation is a time-consuming activity and is often subject to bias [81];
- as the research has mainly focused on annotation in terms of 6 basic categories of emotions, there is not yet a way to know what kind of annotation (Action Units, dimensions or discrete categories) is closest to the reality of the expressions;

- finally, as we will see below, facial expressions can vary significantly from one person or situation to another. It is thus necessary to constitute data bases presenting a maximum of these variants, in order to confront the AFER systems with a sufficiently broad sample of expressions.

In order to deal with these problems, some research teams have begun to develop the use of non-annotated data, in particular through transfer learning (transfer of learning from one algorithm to another, [73]), semi-supervised [64] or unsupervised learning [206].

5. Conclusion

The AFER is a field in constant evolution, and the researchers developed these last years more and more elaborate and powerful systems. However, a great majority of these systems have been tested in conditions far from the natural context, and one can advance that Artificial Intelligence is still far from the capacities of the human being as regards decoding emotional states.

One of the main problems is that the algorithms are designed to exploit data and derive stereotyped characteristics from it, which makes them unable to take into account special cases and novel configurations [45]. Emotion, according to the most recent theoretical conceptions, is a highly subtle and changing phenomenon, which varies according to many parameters that theory has not yet been able to formalize in their entirety. Its conceptual precision, reflecting its complexity, is still problematic and remains a "work in progress".

However, this lack of consensus in the definition and diversity of theoretical models apprehending the concept of emotion and facial expressions should not discourage researchers from studying these phenomena.

It is necessary to pursue research along the four promising avenues highlighted in this review (broadening the range of emotions, taking into account variability and context, tending towards multimodality and generating data from natural conditions), so that artificial intelligence cannot only describe but also apprehend our emotions and all the factors that accompany and determine them in a relevant manner.

All the disciplines concerned with this subject must work in synergy to achieve a more consensual model and to understand the subtle ways in which humans feel, express, decode and exchange their emotions on a daily basis.

Conflicts of interest

There are no known conflict of interest associated with this publication.

Author contributions

- Audrey Masson: Conceptualization, investigation, writing
- Guillaume Cazenave: Resources, review
- Julien Trombini: Resources, review
- Martine Batt: Supervision, review

References

[1] D. Acevedo, P. Negri, M.E. Buemi and M. Mejail, Facial expression recognition based on static and dynamic approaches, in: *2016 23rd International Conference on Pattern Recognition (ICPR)*, 2016, pp. 4124–4129. doi:10.1109/ICPR.2016.7900280.

[2] A. Adams, M. Mahmoud, T. Baltrusaitis and P. Robinson, Decoupling facial expressions and head motions in complex emotions, in: *2015 International Conference on Affective Computing and Intelligent Interaction (ACII)*, 2015, pp. 274–280. doi:10.1109/ACII.2015.7344583.

[3] R. Afdhal, R. Ejbali, M. Zaied and C. Ben Amar, Emotion recognition using features distances classified by wavelets network and trained by fast wavelets transform, in: *2014 14th International Conference on Hybrid Intelligent Systems*, 2014, pp. 238–241. doi:10.1109/HIS.2014.7086205.

[4] S. Al-Darraji, K. Berns and A. Rodić, Action unit based facial expression recognition using deep learning, in: *Advances in Robot Design and Intelligent Control*, A. Rodić and T. Borangiu, eds, Vol. 540, 2017, pp. 413–420. doi:10.1007/978-3-319-49058-8-45.

[5] M. Alam, L.S. Vidyaratne and K.M. Iftekharuddin, Sparse simultaneous recurrent deep learning for robust facial expression recognition, *IEEE Transactions on Neural Networks and Learning Systems* **29**(10) (2018), 4905–4916. doi:10.1109/TNNLS.2017.2776248.

[6] H. Ali, M. Hariharan, S. Yaacob, A.H. Adom, S.K. Za'ba and M. Elshaikh, Facial emotion recognition under partial occlusion using empirical mode decomposition, in: *2016 2nd IEEE International Symposium on Robotics and Manufacturing Automation (ROMA)*, 2016, pp. 1–6. doi:10.1109/ROMA.2016. 7847818.

[7] A.M. Ashir, A. Eleyan and B. Akdemir, Facial expression recognition with dynamic cascaded classifier, *Neural Computing and Applications.* (2019). doi:10.1007/s00521-019-04138-4.

[8] E. Avots, T. Sapiński, M. Bachmann and D. Kamińska, Audiovisual emotion recognition in wild, *Machine Vision and Applications* **30**(5) (2019), 975–985. doi:10.1007/s00138-018-0960-9.

[9] A. Azazi, S.L. Lutfi and I. Venkat, Analysis and evaluation of SURF descriptors for automatic 3D facial expression recognition using different classifiers, in: *2014 4th World Congress on Information and Communication Technologies (WICT 2014)*, 2014, pp. 23–28. doi:10.1109/WICT.2014.7077296.

[10] K. Bahreini, W. van der Vegt and W. Westera, A fuzzy logic approach to reliable real-time recognition of facial emotions, *Multimedia Tools and Applications.* (2019). doi:10.1007/s11042-019-7250-z.

[11] T. Baltrusaitis, L. Li and L.-P. Morency, Local-global ranking for facial expression intensity estimation, in: *2017 Seventh International Conference on Affective Computing and Intelligent Interaction (ACII)*, 2017, pp. 111–118. doi:10.1109/ACII.2017.8273587.

[12] A. Bandrabur, L. Florea, C. Florea and M. Mancas, Late fusion of facial dynamics for automatic expression recognition, *Turkish Journal of Electrical Engineering & Computer Sciences* **25** (2017), 2696–2707. doi:10.3906/elk-1607-113.

[13] E.I. Barakova, R. Gorbunov and M. Rauterberg, Automatic interpretation of affective facial expressions in the context of interpersonal interaction, *IEEE Transactions on Human-Machine Systems* **45**(4) (2015), 409–418. doi:10.1109/THMS.2015.2419259.

[14] A. Barman and P. Dutta, Facial expression recognition using shape signature feature, in: *2017 Third International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN)*, 2017, pp. 174–179. doi:10.1109/ICRCICN.2017.8234502.

[15] L. Barrett, R. Adolphs, S. Marsella, A.-M. Martinez and S.-D. Pollak, Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements, *Psychological Science in the Public Interest* **20**(1) (2019), 1–68. doi:10.1177/1529100619832930.

[16] P. Barros, D. Jirak, C. Weber and S. Wermter, Multimodal emotional state recognition using sequence-dependent deep hierarchical features, *Neural Networks* **72** (2015), 140–151. doi:10.1016/j.neunet.2015.09.009.

[17] J.C. Batista, V. Albiero, O.R.P. Bellon and L. Silva, AUMPNet: Simultaneous action units detection and intensity estimation on multipose facial images using a single convolutional neural network, in: *2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017)*, 2017, pp. 866–871. doi:10.1109/FG.2017.111.

[18] E. Battini Sönmez and A. Cangelosi, Convolutional neural networks with balanced batches for facial expressions recognition, in: *Proceedings Volume 10341, Ninth International Conference on Machine Vision (ICMV 2016)*, A. Verikas,

P. Radeva, D.P. Nikolaev, W. Zhang and J. Zhou, eds, 2017. doi:10.1117/12.2268412.

[19] B. Ben Amor, H. Drira, S. Berretti, M. Daoudi and A. Srivastava, 4-D facial expression recognition by learning geometric deformations, *IEEE Transactions on Cybernetics* **44**(12) (2014), 2443–2457. doi:10.1109/TCYB.2014.2308091.

[20] A. Ben Tanfous, H. Drira and B. Ben Amor, Sparse coding of shape trajectories for facial expression and action recognition, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **1**(1) (2019). doi:10.1109/TPAMI.2019.2932979.

[21] K.-I. Benta and M.-F. Vaida, Towards real-life facial expression recognition systems, *Advances in Electrical and Computer Engineering* **15**(2) (2015), 93–102. doi:10.4316/AECE.2015.02012.

[22] A. Bernin, L. Müller, S. Ghose, C. Grecos, Q. Wang, R. Jettke and F. Vogt, Automatic classification and shift detection of facial expressions in event-aware smart environments, in: *Proceedings of the 11th Pervasive Technologies Related to Assistive Environments Conference On–PETRA'18*, 2018, pp. 194–201. doi:10.1145/3197768.3201527.

[23] P. Bhandari, R.K. Bijarniya, S. Chatterjee and M. Kolekar, Analysis for self-taught and transfer learning based approaches for emotion recognition, in: *2018 5th International Conference on Signal Processing and Integrated Networks (SPIN)*, 2018, pp. 509–512. doi:10.1109/SPIN.2018.8474199.

[24] M. Bishay and I. Patras, Fusing multilabel deep networks for facial action unit detection, in: *2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017)*, 2017, pp. 681–688. doi:10.1109/FG.2017.86.

[25] S. Borah and S. Konwar, ANN based human facial expression recognition in color images, in: *2014 International Conference on High Performance Computing and Applications (ICHPCA)*, 2014, pp. 1–6. doi:10.1109/ICHPCA.2014.7045337.

[26] D. Borza, R. Danescu, R. Itu and A. Darabant, High-speed video system for micro-expression detection and recognition, *Sensors* **17**(12) (2017), 2913. doi:10.3390/s17122913.

[27] H. Boubenna and D. Lee, Image-based emotion recognition using evolutionary algorithms, *Biologically Inspired Cognitive Architectures* **24** (2018), 70–76. doi:10.1016/j.bica.2018.04.008.

[28] H. Boughrara, M. Chtourou, C. Ben Amar and L. Chen, Facial expression recognition based on a mlp neural network using constructive training algorithm, *Multimedia Tools and Applications* **75**(2) (2016), 709–731. doi:10.1007/s11042-014-2322-6.

[29] L.F. Bringmann, E. Ferrer, E.L. Hamaker, D. Borsboom and F. Tuerlinckx, Modeling nonstationary emotion dynamics in dyads using a time-varying vector-autoregressive model, *Multivariate Behavioral Research* **53**(3) (2018), 293–314. doi:10.1080/00273171.2018.1439722.

[30] J.P. Canario and L. Oliveira, Recognition of facial expressions based on deep conspicuous net, in: *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications*, A. Pardo and J. Kittler, eds, Vol. 9423, 2015, pp. 255–262. doi:10.1007/978-3-319-25751-8-31.

[31] U. Chavan and D. Kulkarni, Optimizing deep convolutional neural network for facial expression recognitions, in: *Data Management, Analytics and Innovation*, V.E. Balas,

N. Sharma and A. Chakrabarti, eds, Vol. 808, 2019, pp. 185–196. doi:10.1007/978-981-13-1402-5-14.

[32] J. Chen, Y. Lv, R. Xu and C. Xu, Automatic social signal analysis: Facial expression recognition using difference convolution neural network, *Journal of Parallel and Distributed Computing* **131** (2019), 97–102. doi:10.1016/j.jpdc.2019.04.017.

[33] J. Chen, Q. Ou, Z. Chi and H. Fu, Smile detection in the wild with deep convolutional neural networks, *Machine Vision and Applications* **28**(1–2) (2017), 173–183. doi:10.1007/s00138-016-0817-z.

[34] M.I.H. Chowdhury and F.I. Alam, A probabilistic approach to support Self-Organizing Map (SOM) driven facial expression recognition, in: *2014 17th International Conference on Computer and Information Technology (ICCIT)*, 2014, pp. 210–216. doi:10.1109/ICCITechn.2014.7073131.

[35] C. Corneanu, M. Madadi and S. Escalera, *Deep Structure Inference Network for Facial Action Unit Recognition*, V. Ferrari, M. Hebert, C. Sminchisescu and Y. Weiss, eds, Vol. 11216, Computer Vision – ECCV, 2018, pp. 309–324. doi:10.1007/978-3-030-01258-8-19.

[36] C.A. Corneanu, M.O. Simon, J.F. Cohn and S.E. Guerrero, Survey on RGB, 3D, thermal, and multimodal approaches for facial expression recognition: History, trends, and afect-related applications, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **38**(8) (2016), 1548–1568. doi:10.1109/TPAMI.2016.2515606.

[37] J.Y.R. Cornejo, H. Pedrini and F. Flórez-Revuelta, Facial expression recognition with occlusions based on geometric representation, in: *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications*, A. Pardo and J. Kittler, eds, Vol. 9423, 2015, pp. 263–270. doi:10.1007/978-3-319-25751-8-32.

[38] M.D. Costa-Abreu and G.S. Bezerra, FAMOS: A framework for investigating the use of face features to identify spontaneous emotions, *Patern Analusis and Applications.* (2017). doi:10.1007/sl0044-017-0675-y.

[39] R. Cui, M. Liu and M. Liu, Facial expression recognition based on ensemble of mulitple CNNs, in: *Biometric Recognition*, Z. You, J. Zhou, Y. Wang, Z. Sun, S. Shan, W. Zheng and Q. Zhao, eds, Vol. 9967, 2016, pp. 511–518. doi:10.1007/978-3-319-46654-5-56.

[40] F.A.M. da Silva and H. Pedrini, Effects of cultural characteristics on building an emotion classifier through facial expression analysis, *Journal of Electronic Imaging* **24**(2) (2015), 023015. doi:10.1117/1.JEI.24.2.023015.

[41] C. Darwin, *The Expresston of the Emotions in Man and Animals*, John Murray, London, 1872/1965.

[42] D. Das, M.M. Hoque, J.F. Ara, M.A.F.M.R. Hasan, Y. Kobayashi and Y. Kuno, Automatic face parts extraction and facial expression recognition, in: *2014 9th International Forum on Strategic Technology (IFOST)*, 2014, 128131. doi:10.1109/IFOST.2014.6991087.

[43] M. Del Líbano, M.G. Calvo, A. Fernández-Martín and G. Recio, Discrimination between smiling faces: Human observers vs. automated face analysis, *Acta Psychologica* **187** (2018), 1929. doi:10.1016/j.actpsy.2018.04.019.

[44] W. Deng, J. Hu, S. Zhang and J. Guo, DeepEmo: Real-world facial expression analysis via deep learning, *2015 Visual Communications and Image Processing (VCIP)* (2015), 14. doi:10.1109/VCIP.2015.7457876.

[45] J.-L. Dessalles, *Des Intelligences TRES Artificielles*, Odile Jacob, 2019.

[46] H. Ding, S.K. Zhou and R. Chellappa, FaceNet2ExpNet: Regularizing a deep face REcognition net for expression recognition, in: *2017 12th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2017)*, 2017, pp. 118–126. doi:10.1109/FG.2017.23.

[47] H.I. Dino and M.B. Abdulrazzaq, Facial expression classification based on SVM, KNN and MLP classifiers, in: *2019 International Conference on Advanced Science and Engineering (ICOASE)*, 2019, pp. 70–75. doi:10.1109/ICOASE.2019.8723728.

[48] Y. Dou, M. Zhou, J. Wang and J. Qiang, Facial expression recognition based-on saliency guided support vector machine, in: *2016 9th International Symposium on Computational Intelligence and Design (ISCID)*, 2016, pp. 389–393. doi:10.1109/ISCID.2016.2098.

[49] P. Ekman, An argument for basic emotions, *Cognition and emotion* **6**(3–4) (1992), 169–200. doi:10.1080/02699939208411068.

[50] O. Ekundayo and S. Viriri, Facial expression recognition: A review of methods, performances and limitations, in: *2019 Conference on Information Communications Technology and Society (ICTAS)*, 2019, pp. 1–6. doi:10.1109/ICTAS.2019.8703619.

[51] S. Eleftheriadis, O. Rudovic and M. Pantic, View-constrained latent variable model for multi-view facial expression classification, in: *Advances in Visual Computing*, G. Bebis, R. Boyle, B. Parvin, D. Koracin, R. McMahan, J. Jerald and M. Carlson, eds, Vol. 8888, 2014, pp. 292–303. doi:10.1007/978-3-319-14364-4-28.

[52] Y. Fan, J.C.K. Lam and V.O.K. Li, Multi-region ensemble convolutional neural network for facial expression recognition, in: *Artificial Neural Networks and Machine Learning – ICANN 2018*, V. Kůrková, Y. Manolopoulos, B. Hammer, L. Iliadis and I. Maglogiannis, eds, Vol. 11139, 2018, pp. 84–94. doi:10.1007/978-3-030-01418-6-9.

[53] H. Fang, N. Mac Parthalain, A.J. Aubrey, G.K.L. Tam, R. Borgo, P.L. Rosin and M. Chen, Facial expression recognition in dynamic sequences: An integrated approach, *Pattern Recognition* **47**(3) (2014), 1271–1281. doi:10.1016/j.patcog.2013.09.023.

[54] N. Farajzadeh, G. Pan and Z. Wu, Facial expression recognition based on meta probability codes, *Pattern Analysis and Applications* **17**(4) (2014), 763–781. doi:10.1007/s10044-012-0315-5.

[55] A. Fathallah, L. Abdi and A. Douik, Facial expression recognition via deep learning, in: *2017 IEEE/ACS 14th International Conference on Computer Systems and Applications (AICCSA)*, 2017, pp. 745–750. doi:10.1109/AICCSA.2017.124.

[56] J.M. Garcia-Garcia, V.M.R. Penichet and M.D. Lozano, Emotion detection: A technology review, in: *Proceedings of the XVIII International Conference on Human Computer Interaction*, 2017, pp. 1–8. doi:10.1145/3123818.3123852.

[57] J.M. Girard, J.F. Cohn and F. De la Torre, Estimating smile intensity: A better way, *Pattern Recognition Letters* **66** (2015), 13–21. doi:10.1016/j.patrec.2014.10.004.

[58] I. Gonzalez, F. Cartella, V. Enescu and H. Sahli, Recognition of facial actions and their temporal segments based on

duration models, *Multimedia Tools and Applications* **74**(22) (2015), 10001–10024. doi:10.1007/s11042-014-2320-8.

[59] J. Grobova, M. Colovic, M. Marjanovic, A. Njegus, H. Demire and G. Anbarjafari, Automatic hidden sadness detection using micro-expressions, in: *2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017)*, 2017, pp. 828–832. doi:10.1109/FG.2017.105.

[60] L. Gui, T. Baltrusaitis and L.-P. Morency, Curriculum learning for facial expression recognition, in: *2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017)*, 2017, pp. 505–511. doi:10.1109/FG.2017.68.

[61] H. Gunes and H. Hung, Is automatic facial expression recognition of emotions coming to a dead end? The rise of the new kids on the block, *Image and Vision Computing* **55** (2016), 6–8. doi:10.1016/j.imavis.2016.03.013.

[62] Z. Hammal and M.T. Suarez, Towards context based affective computing introduction to the third international, in: *CBAR 2015 Workshop. 2015 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)*, 2015, pp. 1–2. doi:10.1109/FG.2015.7284841.

[63] L. Hao, S. Wang, G. Peng and Q. Ji, Facial action unit recognition augmented by their dependencies, in: *2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*, 2018, pp. 187–194. doi:10.1109/FG.2018.00036.

[64] S.L. Happy, A. Dantcheva and F. Bremond, A weakly supervised learning technique for classifying facial expressions, *Pattern Recognition Letters* **128** (2019), 162–168. doi:10.1016/j.patrec.2019.08.025.

[65] S.L. Happy and A. Routray, Automatic facial expression recognition using features of salient facial patches, *IEEE Transactions on Affective Computing* **6**(1) (2015), 1–12. doi:10.1109/TAFFC.2014.2386334.

[66] S.L. Happy and A. Routray, Recognizing subtle micro-facial expressions using fuzzy histogram of optical flow orientations and feature selection methods, in: *Computational Intelligence for Pattern Recognition*, W. Pedrycz and S.-M. Chen, eds, Vol. 777, 2018, pp. 341–368. doi:10.1007/978-3-319-89629-8-13.

[67] B. Hasani and M.H. Mahoor, Spatio-temporal facial expression recognition using convolutional neural networks and conditional random fields, in: *2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017)*, 2017, pp. 790–795. doi:10.1109/FG.2017.99.

[68] J. He, D. Li, B. Yang, S. Cao, B. Sun and L. Yu, Multi view facial action unit detection based on CNN and BLSTM-RNN, in: *2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017)*, 2017, pp. 848–853. doi:10.1109/FG.2017.108.

[69] A. Hernandez-Garcia, Perceived emotion from images through deep neural networks, in: *2017 Seventh International Conference on Affective Computing and Intelligent Interaction (ACII)*, 2017, pp. 566–570. doi:10.1109/ACII.2017.8273656.

[70] M.S. Hossain and G. Muhammad, Emotion recognition using deep learning approach from audio–visual emotional big data, *Information Fusion* **49** (2019), 69–78. doi:10.1016/j.inffus.2018.09.008.

[71] M.S. Hossain and G. Muhammad, Emotion recognition using secure edge and cloud computing, *Information Sciences* **504** (2019), 589–601. doi:10.1016/j.ins.2019.07.040.

[72] G. Hu, L. Liu, Y. Yuan, Z. Yu, Y. Hua, Z. Zhang and Y. Yang, *Deep Multi-Task Learning to Recognise Subtle Facial Expressions of Mental States*, V. Ferrari, M. Hebert, C. Sminchisescu and Y. Weiss, eds, Vol. 11216, Computer Vision – ECCV, 2018, pp. 106–123. doi:10.1007/978-3-030-01258-8-7.

[73] J.C. Hung, K.-C. Lin and N.-X. Lai, Recognizing learning emotion based on convolutional neural networks and transfer learning, *Applied Soft Computing* **84** (2019), 105724. doi:10.1016/j.asoc.2019.105724.

[74] I. Hupont and M. Chetouani, Region-based facial representation for real-time action units intensity detection across datasets, *Pattern Analysis and Applications*. (2017). doi:10.1007/s10044-017-0645-4.

[75] N. Hussain, H. Ujir, I. Hipiny and J.-L. Minoi, 3D Facial Action Units Recognition for Emotional Expression, 2017, arXiv:1712.00195.

[76] H. Hussein, M. Naqvi and J. Chambers, Study of image-based expression recognition techniques on three recent spontaneous databases, in: *2017 22nd International Conference on Digital Signal Processing (DSP)*, 2017, pp. 1–5. doi:10.1109/ICDSP.2017.8096062.

[77] X.-P. Huynh, T.-D. Tran and Y.-G. Kim, Convolutional neural network models for facial expression recognition using BU-3DFE database, in: *Information Science and Applications (ICISA) 2016*, K.J. Kim and N. Joukov, eds, Vol. 376, 2016, pp. 441–450. doi:10.1007/978-981-10-0557-2-44.

[78] M. Imani and G.A. Montazer, A survey of emotion recognition methods with emphasis on E-Learning environments, *Journal of Network and Computer Applications* **147** (2019), 102423. doi:10.1016/j.jnca.2019.102423.

[79] S. Jain, M. Durgesh and T. Ramesh, Facial expression recognition using variants of LBP and classifier fusion, in: *Proceedings of International Conference on ICT for Sustainable Development*, S.C. Satapathy, A. Joshi, N. Modi and N. Pathak, eds, Vol. 408, 2016, pp. 725–732. doi:10.1007/978-981-10-0129-1-75.

[80] R. Jameel, A. Singhal and A. Bansal, A comparison of performance of crisp logic and probabilistic neural network for facial expression recognition, in: *2015 1st International Conference on Next Generation Computing Technologies (NGCT)*, 2015, pp. 841–846. doi:10.1109/NGCT.2015.7375238.

[81] R. Jameel, A. Singhal and A. Bansal, A comprehensive study on facial expressions recognition techniques, in: *2016 6th International Conference – Cloud System and Big Data Engineering (Confluence)*, 2016, pp. 478–483. doi:10.1109/CONFLUENCE.2016.7508167.

[82] A. Jan and H. Meng, Automatic 3D facial expression recognition using geometric and textured feature fusion, in: *2015 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)*, 2015, pp. 1–6. doi:10.1109/FG.2015.7284860.

[83] B. Jiang, B. Martinez, M.F. Valstar and M. Pantic, Decision level fusion of domain specific regions for facial action recognition, in: *2014 22nd International Conference on Pattern Recognition*, 2014, pp. 1776–1781. doi:10.1109/ICPR.2014.312.

[84] B. Jiang, M. Valstar, B. Martinez and M. Pantic, A dynamic appearance descriptor approach to facial actions temporal modeling, *IEEE Transactions on Cybernetics* **44**(2) (2014), 161–174. doi:10.1109/TCYB.2013.2249063.

[85] H. Jin, X. Wang, Y. Lian and J. Hua, Emotion information visualization through learning of 3D morphable face model, *The Visual Computer* **35**(4) (2019), 535–548. doi:10.1007/s00371-018-1482-1.

[86] A. Kacem, M. Daoudi and J.-C. Alvarez-Paiva, Barycentric representation and metric learning for facial expression recognition, in: *2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*, 2018, pp. 443–447. doi:10.1109/FG.2018.00071.

[87] G. Khan, S. Samyan, M.U.G. Khan, M. Shahid and S.Q. Wahla, A survey on analysis of human faces and facial expressions datasets, *International Journal of Machine Learning and Cybernetics.* (2019). doi:10.1007/s13042-019-00995-6.

[88] H.-Q. Khor, J. See, R.C.W. Phan and W. Lin, Enriched long-term recurrent convolutional network for facial microexpression recognition, in: *2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*, 2018, pp. 667–674. doi:10.1109/FG.2018.00105.

[89] Y. Kim, Exploring sources of variation in human behavioral data: Towards automatic audio-visual emotion recognition, in: *2015 International Conference on Affective Computing and Intelligent Interaction (ACII)*, 2015, pp. 748–753. doi:10.1109/ACII.2015.7344653.

[90] B. Ko, A brief review of facial emotion recognition based on visual information, *Sensors* **18**(2) (2018), 401. doi:10.3390/s18020401.

[91] F. Kong, Facial expression recognition method based on deep convolutional neural network combined with improved LBP features, *Personal and Ubiquitous Computing* **23**(3–4) (2019), 531–539. doi:10.1007/s00779-019-01238-9.

[92] E. Kukla and P. Nowak, Facial emotion recognition based on cascade of neural networks, in: *New Research in Multimedia and Internet Systems*, A. Zgrzywa, K. Choroś and A. Siemiński, eds, Vol. 314, 2015, pp. 67–78. doi:10.1007/978-3-319-10383-9-7.

[93] K. Kulkarni, C. Corneanu, I. Ofodile, S. Escalera, X. Baro, S. Hyniewska and G. Anbarjafari, Automatic recognition of facial displays of unfelt emotions, *IEEE Transactions on Affective Computing* **1**(1) (2018). doi:10.1109/TAFFC.2018.2874996.

[94] K.R. Kulkarni and S.B. Bagal, Facial expression recognition, in: *2015 Annual IEEE India Conference (INDICON)*, 2015, pp. 1–5. doi:10.1109/INDICON.2015.7443572.

[95] M.P. Kumar and M.K. Rajagopal, Detecting facial emotions using normalized minimal feature vectors and semisupervised twin support vector machines classifier, *Applied Intelligence.* (2019). doi:10.1007/s10489-019-01500-w.

[96] Y. Kumar and S. Sharma, A systematic survey of facial expression recognition techniques, in: *2017 International Conference on Computing Methodologies and Communication (ICCMC)*, 2017, pp. 1074–1079. doi:10.1109/ICCMC.2017.8282636.

[97] T. Kundu and C. Saravanan, Advancements and recent trends in emotion recognition using facial, in: *Image Analysis and Machine Learning Models. 2017 International Conference on Electrical, Electronics, Communication, Computer, and Op-timization Techniques (ICEECCOT)*, 2017, pp. 1–6. doi:10.1109/ICEECCOT.2017.8284512.

[98] S.M. Lajevardi, Structural similarity classifier for facial expression recognition, *Signal, Image and Video Processing* **8**(6) (2014), 1103–1110. doi:10.1007/s11760-014-0639-2.

[99] A.C. Le Ngo, J. See and R.C.-W. Phan, Sparsity in dynamics of spontaneous subtle emotions: Analysis and application, *IEEE Transactions on Affective Computing* **8**(3) (2017), 396–411. doi:10.1109/TAFFC.2016.2523996.

[100] J. Li, C. Soladie and R. Seguier, LTP-ML: Micro-expression detection by recognition of local temporal pattern of facial movements, in: *2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*, 2018, pp. 634–641. doi:10.1109/FG.2018.00100.

[101] J. Li, C. Soladie, R. Seguier, S.-J. Wang and M.H. Yap, Spotting micro-expressions on long videos sequences, in: *2019 14th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2019)*, 2019, pp. 1–5. doi:10.1109/FG.2019.8756626.

[102] J. Li, Y. Wang, J. See and W. Liu, Micro-expression recognition based on 3D flow convolutional neural network, *Pattern Analysis and Applications* **22**(4) (2019), 1331–1339. doi:10.1007/s10044-018-0757-5.

[103] Q. Li, J. Yu, T. Kurihara and S. Zhan, Micro-expression analysis by fusing deep convolutional neural network and optical flow, in: *2018 5th International Conference on Control, Decision and Information Technologies (CoDIT)*, 2018, pp. 265–270. doi:10.1109/CoDIT.2018.8394868.

[104] Q. Li, S. Zhan, L. Xu and C. Wu, Facial micro-expression recognition based on the fusion of deep learning and enhanced optical flow, *Multimedia Tools and Applications.* (2018). doi:10.1007/s11042-018-6857-9.

[105] S. Li and W. Deng, Blended emotion in-the-wild: Multilabel facial expression recognition using crowdsourced annotations and deep locality feature learning, *International Journal of Computer Vision* **127**(6–7) (2019), 884–906. doi:10.1007/s11263-018-1131-1.

[106] W. Li, D. Huang, H. Li and Y. Wang, Automatic 4D facial expression recognition using dynamic geometrical image network, in: *2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*, 2018, pp. 24–30. doi:10.1109/FG.2018.00014.

[107] X. Li, S. Chen and Q. Jin, Facial action units detection with multi-features and -AUs fusion, in: *2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017)*, 2017, pp. 860–865. doi:10.1109/FG.2017.110.

[108] X. Li, X. Hong, A. Moilanen, X. Huang, T. Pfister, G. Zhao and M. Pietikainen, Towards reading hidden emotions: A comparative study of spontaneous micro-expression spotting and recognition methods, *IEEE Transactions on Affective Computing* **9**(4) (2018), 563–577. doi:10.1109/TAFFC.2017.2667642.

[109] Y. Li, S.M. Mavadati, M.H. Mahoor, Y. Zhao and Q. Ji, Measuring the intensity of spontaneous facial action units with dynamic Bayesian network, *Pattern Recognition* **48**(11) (2015), 3417–3427. doi:10.1016/j.patcog.2015.04.022.

[110] M.Z. Lifkooee, Ö.M. Soysal and K. Sekeroglu, Video mining for facial action unit classification using statistical spatial–temporal feature image and LoG deep convolutional neural network, *Machine Vision and Applications* **30**(1) (2018), 41–57. doi:10.1007/s00138-018-0967-2.

[111] D.Y. Liliana, C. Basaruddin and M.R. Widyanto, Mix emotion recognition from facial expression using SVM-CRF sequence classifier, in: *Proceedings of the International Conference on Algorithms, Computing and Systems – ICACS'17*, 2017, pp. 27–31. doi:10.1145/3127942.3127958.

[112] D.Y. Liliana, T. Basaruddin, M.R. Widyanto and I.I.D. Oriza, Fuzzy emotion: A natural approach to automatic facial expression recognition from psychological perspective using fuzzy system, *Cognitive Processing.* (2019). doi:10.1007/s10339-019-00923-0.

[113] Y. Liu, Y. Li, X. Ma and R. Song, Facial expression recognition with fusion features extracted from salient facial areas, *Sensors* **17**(4) (2017), 712. doi:10.3390/s17040712.

[114] Z. Liu, S. Li and W. Deng, Boosting-POOF: Boosting part based one vs one feature for facial expression recognition in the wild, in: *2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017)*, 2017, pp. 967–972. doi:10.1109/FG.2017.120.

[115] L. Lo Presti and M. La Cascia, Boosting Hankel matrices for face emotion recognition and pain detection, *Computer Vision and Image Understanding* **156** (2017), 19–33. doi:10.1016/j.cviu.2016.10.007.

[116] Z. Lu, Z. Luo, H. Zheng, J. Chen and W. Li, A Delaunay-based temporal coding model for micro-expression recognition, in: *Computer Vision – ACCV 2014 Workshops*, C.V. Jawahar and S. Shan, eds, Vol. 9009, 2015, pp. 698–711. doi:10.1007/978-3-319-16631-5-51.

[117] Z. Ma, M. Mahmoud, P. Robinson, E. Dias and L. Skrypchuk, Automatic detection of a driver's complex mental states, in: *Computational Science and Its Applications – ICCSA 2017*, O. Gervasi, B. Murgante, S. Misra, G. Borruso, C.M. Torre, A.M.A.C. Rocha and A. Cuzzocrea, eds, Vol. 10406, 2017, pp. 678–691. doi:10.1007/978-3-319-62398-6-48.

[118] B. Mandal, D. Lee and N. Ouarti, Distinguishing posed and spontaneous smiles by facial dynamics, in: *Computer Vision – ACCV 2016 Workshops*, C.-S. Chen, J. Lu and K.-K. Ma, eds, Vol. 10116, 2017, pp. 552–566. doi:10.1007/978-3-319-54407-6-37.

[119] E. Mansouri-Benssassi and J. Ye, Bio-inspired spiking neural networks for facial expression recognition: Generalisation investigation, in: *Theory and Practice of Natural Computing*, D. Fagan, C. Martín-Vide, M. O'Neill and M.A. Vega-Rodríguez, eds, Vol. 11324, 2018, pp. 426–437. doi:10.1007/978-3-030-04070-3-33.

[120] P. Marrero-Fernández, A. Montoya-Padrón, A. Jaume-i-Capó and J.M. Buades Rubio, Evaluating the research in automatic emotion recognition, *IETE Technical Review* **31**(3) (2014), 220–232. doi:10.1080/02564602.2014.906863.

[121] G. Mattela and S.K. Gupta, Facial expression recognition using Gabor-mean-DWT feature extraction technique, in: *2018 5th International Conference on Signal Processing and Integrated Networks (SPIN)*, 2018, pp. 575–580. doi:10.1109/SPIN.2018.8474206.

[122] V. Mayya, R.M. Pai and M.M. Manohara Pai, Automatic facial expression recognition using DCNN, *Procedia Computer Science* **93** (2016), 453–461. doi:10.1016/j.procs.2016.07.233.

[123] D. Mehta, M. Siddiqi and A. Javaid, Facial emotion recognition: A survey and real-world user experiences in mixed reality, *Sensors* **18**(2) (2018), 416. doi:10.3390/s18020416.

[124] N. Mehta and S. Jadhav, Facial emotion recognition using Log Gabor filter and PCA, in: *2016 International Conference on Computing Communication Control and Automation (ICCUBEA)*, 2016, pp. 1–5. doi:10.1109/ICCUBEA.2016.7860054.

[125] Y. Miao, H. Dong, J.M.A. Jaam and A.E. Saddik, A deep learning system for recognizing facial expression in real-time, *ACM Transactions on Multimedia Computing, Communications, and Applications* **15**(2) (2019), 1–20. doi:10.1145/3311747.

[126] W. Mou, H. Gunes and I. Patras, Alone versus in-a-group: A multi-modal framework for automatic affect recognition, *ACM Transactions on Multimedia Computing, Communications, and Applications* **15**(2) (2019), 1–23. doi:10.1145/3321509.

[127] N. Muna, U.D. Rosiani, E.M. Yuniamo and M.H. Pumomo, Subpixel subtle motion estimation of micro-expressions multiclass classification, in: *2017 IEEE 2nd International Conference on Signal and Image Processing (ICSIP)*, 2017, pp. 325–330. doi:10.1109/SIPROCESS.2017.8124558.

[128] S. Naik and R.P.K. Jagannath, GCV-based regularized extreme learning machine for facial expression recognition, in: *Advances in Machine Learning and Data Science*, D. Reddy Edla, P. Lingras and K. Venkatanareshbabu, eds, Vol. 705, 2018, pp. 129–138. doi:10.1007/978-981-10-8569-7-14.

[129] P. Natarajan and V. Muthuswamy, Multi-view face expression recognition – a hybrid method, in: *Artificial Intelligence and Evolutionary Algorithms in Engineering Systems*, L.P. Suresh, S.S. Dash and B.K. Panigrahi, eds, Vol. 325, 2015, pp. 799–808. doi:10.1007/978-81-322-2135-7-84.

[130] D. Nguyen, K. Nguyen, S. Sridharan, I. Abbasnejad, D. Dean and C. Fookes, Meta transfer learning for facial emotion recognition, in: *2018 24th International Conference on Pattern Recognition (ICPR)*, 2018, pp. 3543–3548. doi:10.1109/ICPR.2018.8545411.

[131] T.-L. Nguyen, S. Kavuri and M. Lee, A multimodal convolutional neuro-fuzzy network for emotion understanding of movie clips, *Neural Networks* **118** (2019), 208–219. doi:10.1016/j.neunet.2019.06.010.

[132] M.F. O'Connor and L.D. Riek, Detecting social context: A method for social event classification using naturalistic multimodal data, in: *2015 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)*, 2015, pp. 1–7. doi:10.1109/FG.2015.7284843.

[133] Y.-H. Oh, J. See, A.C. Le Ngo, R.C.-W. Phan and V.M. Baskaran, A survey of automatic facial micro-expression analysis: Databases, *Methods, and Challenges. Frontiers in Psychology* **9** (2018), 1128. doi:10.3389/fpsyg.2018.01128.

[134] D. Patel, G. Zhao and M. Pietikäinen, Spatiotemporal integration of optical flow vectors for micro-expression detection, in: *Advanced Concepts for Intelligent Vision Systems*, S. Battiato, J. Blanc-Talon, G. Gallo, W. Philips, D. Popescu and P. Scheunders, eds, Vol. 9386, 2015, pp. 369–380. doi:10.1007/978-3-319-25903-1-32.

[135] M.N. Patil, B. Iyer and R. Arya, Performance evaluation of PCA and ICA algorithm for facial expression recognition application, in: *Proceedings of Fifth International Conference on Soft Computing for Problem Solving*, M. Pant, K. Deep, J.C. Bansal, A. Nagar and K.N. Das, eds, Vol. 436, 2016, pp. 965–976. doi:10.1007/978-981-10-0448-3-81.

[136] M. Peng, Z. Wu, Z. Zhang and T. Chen, From macro to micro expression recognition: Deep learning on small datasets using transfer learning, in: *2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*, 2018, pp. 657–661. doi:10.1109/FG.2018.00103.

[137] I. Perikos, M. Paraskevas and I. Hatzilygeroudis, Facial expression recognition using adaptive neuro-fuzzy inference systems, in: *2018 IEEE/ACIS 17th International Conference on Computer and Information Science (ICIS)*, 2018, pp. 1–6. doi:10.1109/ICIS.2018.8466438.

[138] I. Perikos, E. Ziakopoulos and I. Hatzilygeroudis, Recognizing emotions from facial expressions using neural network, in: *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications*, E. Bayro-Corrochano and E. Hancock, eds, Vol. 8827, 2014, pp. 236–245. doi:10.1007/978-3-662-44654-6-23.

[139] I. Perikos, E. Ziakopoulos and I. Hatzilygeroudis, Recognize emotions from facial expressions using a SVM and neural network schema, in: *Engineering Applications of Neural Networks*, L. Iliadis and C. Jayne, eds, Vol. 517, 2015, pp. 265–274. doi:10.1007/978-3-319-23983-5-25.

[140] Y. Piparsaniyan, V.K. Sharma and K.K. Mahapatra, Robust facial expression recognition using Gabor feature and Bayesian discriminating classifier, in: *2014 International Conference on Communication and Signal Processing*, 2014, pp. 538–541. doi:10.1109/ICCSP.2014.6949900.

[141] R. Plutchik, The nature of emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice, *American Scientist* **89**(4) (2001), 344–350. doi:10.1511/2001.4.344.

[142] R.K. Prasada, S.R. Chandra and H.N. Chowdary, An integrated approach to emotion recognition and gender classification, *Journal of Visual Communication and Image Representation* **60** (2019), 339–345. doi:10.1016/j.jvcir.2019.03.002.

[143] C. Quan, Y. Qian and F. Ren, Dynamic facial expression recognition based on K-order emotional intensity model, in: *2014 IEEE International Conference on Robotics and Biomimetics (ROBIO 2014)*, 2014, pp. 1164–1168. doi:10.1109/ROBIO.2014.7090490.

[144] A. Racoviteanu, C. Florea, M. Badea and C. Vertan, Spontaneous emotion detection by combined learned and fixed descriptors, in: *2019 International Symposium on Signals, Circuits and Systems (ISSCS)*, 2019, pp. 1–4. doi:10.1109/ISSCS.2019.8801755.

[145] F. Rahdari, E. Rashedi and M. Eftekhari, A multimodal emotion recognition system using facial landmark analysis, *Iranian Journal of Science and Technology, Transactions of Electrical Engineering* **43**(S1) (2019), 171–189. doi:10.1007/s40998-018-0142-9.

[146] T. Rao, X. Li, H. Zhang and M. Xu, Multi-level region-based convolutional neural network for image emotion classification, *Neurocomputing* **333** (2019), 429–439. doi:10.1016/j.neucom.2018.12.053.

[147] T.A. Rashid, Convolutional neural networks based method for improving facial expression recognition, in: *Intelligent Systems Technologies and Applications*, J.M. Corchado Rodriguez, S. Mitra, S.M. Thampi and E.-S. El-Alfy, eds, Vol. 530, 2016, pp. 73–84. doi:10.1007/978-3-319-47952-1-6.

[148] N. Rathee and D. Ganotra, An efficient approach for facial action unit intensity detection using distance metric learning based on cosine similarity, *Signal, Image and Video Processing* **12**(6) (2018), 1141–1148. doi:10.1007/s11760-018-1255-3.

[149] R.P. Reddy, P.M. Krishna, V. Narayanan and S. Lalitha, Affective state recognition using image cues, in: *2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, 2018, pp. 928–933. doi:10.1109/ICACCI.2018.8554441.

[150] B. Richhariya and D. Gupta, Facial expression recognition using iterative universum twin support vector machine, *Applied Soft Computing* **76** (2019), 53–67. doi:10.1016/j.asoc.2018.11.046.

[151] P.V. Rouast, M. Adam and R. Chiong, Deep learning for human affect recognition: Insights and new developments, *IEEE Transactions on Affective Computing* **1**(1) (2019). doi:10.1109/TAFFC.2018.2890471.

[152] O. Rudovic, V. Pavlovic and M. Pantic, Context-sensitive dynamic ordinal regression for intensity estimation of facial action units, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **37**(5) (2015), 944–958. doi:10.1109/TPAMI.2014.2356192.

[153] R. Saabni, Facial expression recognition using multi radial bases function networks and 2-D Gabor filters, in: *2015 Fifth International Conference on Digital Information Processing and Communications (ICDIPC)*, 2015, pp. 225–230. doi:10.1109/ICDIPC.2015.7323033.

[154] H. Sadeghi and A.-A. Raie, Human vision inspired feature extraction for facial expression recognition, *Multimedia Tools and Applications*. (2019). doi:10.1007/s11042-019-07863-z.

[155] S. Said, O. Jemai, M. Zaied and C. Ben Amar, Wavelet networks for facial emotion recognition, in: *2015 15th International Conference on Intelligent Systems Design and Applications (ISDA)*, 2015, pp. 295–300. doi:10.1109/ISDA.2015.7489242.

[156] M. Sajjad, A. Shah, Z. Jan, S.I. Shah, S.W. Baik and I. Mehmood, Facial appearance and texture feature-based robust facial expression recognition framework for sentiment knowledge discovery, *Cluster Computing* **21**(1) (2018), 549–567. doi:10.1007/s10586-017-0935-z.

[157] H. Salam and M. Chetouani, A multi-level context-based modeling of engagement in human-robot interaction, in: *2015 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)*, 2015, pp. 1–6. doi:10.1109/FG.2015.7284845.

[158] M.G. Salido Ortega, L.-F. Rodríguez and J.O. Gutierrez-Garcia, Towards emotion recognition from contextual information using machine learning, *Journal of Ambient Intelligence and Humanized Computing*. (2019). doi:10.1007/s12652-019-01485-x.

[159] F.Z. Salmam, A. Madani and M. Kissi, Facial expression recognition using decision trees, in: *2016 13th International Conference on Computer Graphics, Imaging and Visualization (CGiV)*, 2016, pp. 125–130. doi:10.1109/CGiV.2016.33.

[160] F.Z. Salmam, A. Madani and M. Kissi, Fusing multi-stream deep neural networks for facial expression recognition, *Signal, Image and Video Processing* **13**(3) (2019), 609–616. doi:10.1007/s11760-018-1388-4.

[161] 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, *Sensors* **19**(8) (2019), 1863. doi:10.3390/s19081863.

[162] A. Samara, L. Galway, R. Bond and H. Wang, Affective state detection via facial expression analysis within a human–computer interaction context, *Journal of Ambient Intelligence and Humanized Computing* **10**(6) (2019), 2175–2184. doi:10.1007/s12652-017-0636-8.

[163] A. Scarantino, How to do things with emotional expressions: The theory of affective pragmatics, *Psychological Inquiry* **28**(2–3) (2017), 165–185. doi:10.1080/1047840X.2017.1328951.

[164] A. Selvaraj and N.S. Russel, Bimodal recognition of affective states with the features inspired from human visual and auditory perception system, *International Journal of Imaging Systems and Technology*. (2019). doi:10.1002/ima.22338.

[165] D. Sen, S. Datta and R. Balasubramanian, Facial emotion classification using concatenated geometric and textural features, *Multimedia Tools and Applications* **78**(8) (2019), 10287–10323. doi:10.1007/s11042-018-6537-9.

[166] T. Senechal, K. Bailly and L. Prevost, Impact of action unit detection in automatic emotion recognition, *Pattern Analysis and Applications* **17**(1) (2014), 51–67. doi:10.1007/s10044-012-0279-5.

[167] K. Shan, J. Guo, W. You, D. Lu and R. Bie, Automatic facial expression recognition based on a deep convolutional-neural-network structure, in: *2017 IEEE 15th International Conference on Software Engineering Research, Management and Applications (SERA)*, 2017, pp. 123–128. doi:10.1109/SERA.2017.7965717.

[168] J. Shao, I. Gori, S. Wan and J.K. Aggarwal, 3D dynamic facial expression recognition using low-resolution videos, *Pattern Recognition Letters* **65** (2015), 157–162. doi:10.1016/j.patrec.2015.07.039.

[169] G. Sharma, L. Singh and S. Gautam, Automatic facial expression recognition using combined geometric features, *3D Research* **10**(2) (2019). doi:10.1007/s13319-019-0224-0.

[170] M. Sharma, A.S. Jalal and A. Khan, Emotion recognition using facial expression by fusing key points descriptor and texture features, *Multimedia Tools and Applications* **78**(12) (2019), 16195–16219. doi:10.1007/s11042-018-7030-1.

[171] Y. Shi, Z. Lv, N. Bi and C. Zhang, An improved SIFT algorithm for robust emotion recognition under various face poses and illuminations, *Neural Computing and Applications*. (2019). doi:10.1007/s00521-019-04437-w.

[172] H. Shim, K.-H. Cho, K.-E. Ko, I.-H. Jang and K.-B. Sim, Multi-tasking deep convolutional network architecture design for extracting nonverbal communicative information from a face, *Cognitive Systems Research* **52** (2018), 658–667. doi:10.1016/j.cogsys.2018.08.006.

[173] M.H. Siddiqi, Accurate and robust facial expression recognition system using real-time YouTube-based datasets, *Applied Intelligence* **48**(9) (2018), 2912–2929. doi:10.1007/s10489-017-1121-y.

[174] M.H. Siddiqi, M. Ali, M.E. Abdelrahman Eldib, A. Khan, O. Banos, A.M. Khan and H. Choo, Evaluating real-life performance of the state-of-the-art in facial expression recognition using a novel YouTube-based datasets, *Multimedia Tools and Applications* **77**(1) (2018), 917–937. doi:10.1007/s11042-016-4321-2.

[175] P. Siritanawan, K. Kotani and F. Chen, Independent subspace of dynamic Gabor features for facial expression classification, in: *2014 IEEE International Symposium on Multimedia*, 2014, pp. 47–54. doi:10.1109/ISM.2014.48.

[176] K. Slimani, K. Lekdioui, R. Messoussi and R. Touahni, Compound facial expression recognition based on highway CNN, in: *Proceedings of the New Challenges in Data Sciences: Acts of the Second Conference of the Moroccan Classification Society on ZZZ–SMC'19*, 2019, pp. 1–7. doi:10.1145/3314074.3314075.

[177] Y. Song, D. McDuff, D. Vasisht and A. Kapoor, Exploiting sparsity and co-occurrence structure for action unit recognition, in: *2015 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)*, 2015, pp. 1–8. doi:10.1109/FG.2015.7163081.

[178] E. Sönmez and S. Albayrak, A facial component-based system for emotion classification, *Turkish Journal of Electrical Engineering & Computer Sciences* **24** (2016), 1663–1673. doi:10.3906/elk-1401-18.

[179] O.M. Soysal, S. Shirzad and K. Sekeroglu, Facial action unit recognition using data mining integrated deep learning, in: *2017 International Conference on Computational Science and Computational Intelligence (CSCI)*, 2017, pp. 437–443. doi:10.1109/CSCI.2017.74.

[180] O. Starostenko, C. Cruz-Perez, V. Alarcon-Aquino and R. Rosas-Romero, Real-time facial expression recognition using local appearance-based descriptors, *Journal of Intelligent & Fuzzy Systems* **36**(5) (2019), 5037–5049. doi:10.3233/JIFS-179049.

[181] S. Stöckli, M. Schulte-Mecklenbeck, S. Borer and A.C. Samson, Facial expression analysis with AFFDEX and FACET: A validation study, *Behavior Research Methods* **50**(4) (2018), 1446–1460. doi:10.3758/s13428-017-0996-1.

[182] G. Stratou and L.-P. Morency, MultiSense — context-aware nonverbal behavior analysis framework: A psychological distress use case, *IEEE Transactions on Affective Computing* **8**(2) (2017), 190–203. doi:10.1109/TAFFC.2016.2614300.

[183] G. Stratou, J. Van Der Schalk, R. Hoegen and J. Gratch, Refactoring facial expressions: An automatic analysis of natural occurring facial expressions in iterative social dilemma, in: *2017 Seventh International Conference on Affective Computing and Intelligent Interaction (ACII)*, 2017, pp. 427–433. doi:10.1109/ACII.2017.8273635.

[184] P. Suja, S. Tripathi and J. Deepthy, Emotion recognition from facial expressions using frequency domain techniques, in: *Advances in Signal Processing and Intelligent Recognition Systems*, S.M. Thampi, A. Gelbukh and J. Mukhopadhyay, eds, Vol. 264, 2014, pp. 299–310. doi:10.1007/978-3-319-04960-1-27.

[185] M. Sultan Zia, M. Hussain and M. Arfan Jaffar, A novel spontaneous facial expression recognition using dynamically weighted majority voting based ensemble classifier, *Multimedia Tools and Applications* **77**(19) (2018), 25537–25567. doi:10.1007/s11042-018-5806-y.

[186] K. Sumi and T. Ueda, Micro-expression recognition for detecting human emotional changes, in: *Novel User Experiences*, M. Kurosu and H.-C. Interaction, eds, Vol. 9733, 2016, pp. 60–70. doi:10.1007/978-3-319-39513-5-6.

[187] B. Sun, S. Cao, J. He and L. Yu, Affect recognition from facial movements and body gestures by hierarchical deep spatio-

temporal features and fusion strategy, *Neural Networks* **105** (2018), 36–51. doi:10.1016/j.neunet.2017.11.021.

[188] Y. Sun and A.N. Akansu, Automatic inference of mental states from spontaneous facial expressions, in: *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2014, pp. 719–723. doi:10.1109/ICASSP.2014.6853690.

[189] K. Talele, A. Shirsat, T. Uplenchwar and K. Tuckley, Facial expression recognition using general regression neural network, in: *2016 IEEE Bombay Section Symposium (IBSS)*, 2016, pp. 1–6. doi:10.1109/IBSS.2016.7940203.

[190] B.M.S.B. Talukder, B. Chowdhury, T. Howlader and S.M.M. Rahman, Intelligent recognition of spontaneous expression using motion magnification of spatio-temporal data, in: *Intelligence and Security Informatics*, M. Chau, G.A. Wang and H. Chen, eds, Vol. 9650, 2016, pp. 114–128. doi:10.1007/978-3-319-31863-9-9.

[191] C. Tang, W. Zheng, J. Yan, Q. Li, Y. Li, T. Zhang and Z. Cui, View-independent facial action unit detection, in: *2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017)*, 2017, pp. 878–882. doi:10.1109/FG.2017.113.

[192] D.L. Tran, R. Walecki, O. Rudovic, S. Eleftheriadis, B. Schuller and M. Pantic, DeepCoder: Semi-parametric variational autoencoders for automatic facial action coding, in: *2017 IEEE International Conference on Computer Vision (ICCV)*, 2017, pp. 3209–3218. doi:10.1109/ICCV.2017.346.

[193] A. Uçar, Y. Demir and C. Güzeliş, A new facial expression recognition based on curvelet transform and online sequential extreme learning machine initialized with spherical clustering, *Neural Computing and Applications* **27**(1) (2016), 131–142. doi:10.1007/s00521-014-1569-1.

[194] T. Van Huynh, H.-J. Yang, G.-S. Lee, S.-H. Kim and I.-S. Na, Emotion recognition by integrating eye movement analysis and facial expression model, in: *Proceedings of the 3rd International Conference on Machine Learning and Soft Computing–ICMLSC 2019*, 2019, pp. 166–169. doi:10.1145/3310986.3311001.

[195] R. Vedantham, L. Settipalli and E.S. Reddy, Modified back propagation neural network for facial expression classification using principal component analysis and ridgelet transform, in: *Computational Intelligence: Theories, Applications and Future Directions – Volume II*, N.K. Verma and A.K. Ghosh, eds, Vol. 799, 2019, pp. 175–187. doi:10.1007/978-981-13-1135-2-14.

[196] M. Verburg and V. Menkovski, Micro-expression detection in long videos using optical flow and recurrent neural networks, in: *2019 14th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2019)*, 2019, pp. 1–6. doi:10.1109/FG.2019.8756588.

[197] M. Verma, P. Sexena, S. Vipparthi and G. Singh, QUEST: Quadriletral senary bit pattern for facial expression recognition, in: *2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2018, pp. 1498–1503. doi:10.1109/SMC.2018.00260.

[198] R.-L. Vieriu, S. Tulyakov, S. Semeniuta, E. Sangineto and N. Sebe, Facial expression recognition under a wide range of head poses, in: *2015 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)*, 2015, pp. 1–7. doi:10.1109/FG.2015.7163098.

[199] R. Vishnu Priya, V. Vijayakumar and J.M.R.S. Tavares, MQS-MER: A mixed quadratic shape model with optimal fuzzy membership functions for emotion recognition, *Neural Computing and Applications*. (2019). doi:10.1007/s00521-018-3940-0.

[200] R. Walecki, O. Rudovic, V. Pavlovic, B. Schuller and M. Pantic, Deep Structured Learning for Facial Action Unit Intensity Estimation, 2017, arXiv:1704.04481.

[201] X. Wang, X. Liu, L. Lu and Z. Shen, A new facial expression recognition method based on geometric alignment and LBP features, in: *2014 IEEE 17th International Conference on Computational Science and Engineering*, 2014, pp. 1734–1737. doi:10.1109/CSE.2014.318.

[202] Y. Wang, J. See, Y.-H. Oh, R.C.-W. Phan, Y. Rahulamathavan, H.-C. Ling and X. Li, Effective recognition of facial micro-expressions with video motion magnification, *Multimedia Tools and Applications* **76**(20) (2017), 21665–21690. doi:10.1007/s11042-016-4079-6.

[203] C.-H. Wu, J.-C. Lin and W.-L. Wei, Action unit reconstruction of occluded facial expression, in: *2014 International Conference on Orange Technologies*, 2014, pp. 177–180. doi:10.1109/ICOT.2014.6956628.

[204] Y. Wu, G. An and Y. Zhang, Lite AU convolution network driven by a small amount of samples, *Electronics Letters* **53**(14) (2017), 920–922. doi:10.1049/el.2017.0272.

[205] Z. Xia, X. Feng, X. Hong and G. Zhao, Spontaneous facial micro-expression recognition via deep convolutional network, in: *2018 Eighth International Conference on Image Processing Theory, Tools and Applications (IPTA)*, 2018, pp. 1–6. doi:10.1109/IPTA.2018.8608119.

[206] Y. Xiao, D. Wang and L. Hou, Unsupervised emotion recognition algorithm based on improved deep belief model in combination with probabilistic linear discriminant analysis, *Personal and Ubiquitous Computing* **23**(3–4) (2019), 553–562. doi:10.1007/s00779-019-01235-y.

[207] W. Xiaohua, P. Muzi, P. Lijuan, H. Min, J. Chunhua and R. Fuji, Two-level attention with two-stage multi-task learning for facial emotion recognition, *Journal of Visual Communication and Image Representation* **62** (2019), 217–225. doi:10.1016/j.jvcir.2019.05.009.

[208] S. Xie and H. Hu, Facial expression recognition using hierarchical features with deep comprehensive multipatches aggregation convolutional neural networks, *IEEE Transactions on Multimedia* **21**(1) (2019), 211–220. doi:10.1109/TMM.2018.2844085.

[209] S. Xie, H. Hu and Y. Wu, Deep multi-path convolutional neural network joint with salient region attention for facial expression recognition, *Pattern Recognition* **92** (2019), 177–191. doi:10.1016/j.patcog.2019.03.019.

[210] Y. Xudong, D. Huang, Y. Wang and L. Chen, Automatic 3D facial expression recognition using geometric scattering representation, in: *2015 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)*, 2015, pp. 1–6. doi:10.1109/FG.2015.7163090.

[211] H. Xue and I. Gertner, Automatic recognition of emotions from facial expressions, in: *Proceedings Volume 9090, Automatic Target Recognition XXIV*, F.A. Sadjadi and A. Mahalanobis, eds, 2014. doi:10.1117/12.2057796.

[212] Y. Yaddaden, M. Adda, A. Bouzouane, S. Gaboury and B. Bouchard, One-class and bi-class SVM classifier comparison for automatic facial expression recognition, in: *2018 In-*

ternational Conference on Applied Smart Systems (ICASS)*, 2018, pp. 1–6. doi:10.1109/ICASS.2018.8651969.

[213] A. Yuce, H. Gao and J.-P. Thiran, Discriminant multi-label manifold embedding for facial action unit detection, in: *2015 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)*, 2015, pp. 1–6. doi:10.1109/FG.2015.7284871.

[214] K. Yurtkan and H. Demirel, Entropy-based feature selection for improved 3D facial expression recognition, *Signal, Image and Video Processing* **8**(2) (2014), 267–277. doi:10.1007/s11760-013-0543-1.

[215] N. Zeng, H. Zhang, B. Song, W. Liu, Y. Li and A.M. Dobaie, Facial expression recognition via learning deep sparse autoencoders, *Neurocomputing* **273** (2018), 643–649. doi:10.1016/j.neucom.2017.08.043.

[216] S. Zhang, X. Pan, Y. Cui, X. Zhao and L. Liu, Learning affective video features for facial expression recognition via hybrid deep learning, *IEEE Access* **1**(1) (2019). doi:10.1109/ACCESS.2019.2901521.

[217] X. Zhang and Y. Ma, Learning of complicate facial expression categories, in: *Proceedings of the 2019 International Conference on Image, Video and Signal Processing–IVSP 2019*, 2019, pp. 73–80. doi:10.1145/3317640.3317659.

[218] X. Zhang, M.H. Mahoor and S.M. Mavadati, Facial expression recognition using lp-norm MKL multiclass-SVM, *Machine Vision and Applications* **26**(4) (2015), 467–483. doi:10.1007/s00138-015-0677-y.

[219] Y. Zhang, L. Zhang and M.A. Hossain, Adaptive 3D facial action intensity estimation and emotion recognition, *Expert Systems with Applications* **42**(3) (2015), 1446–1464. doi:10.1016/j.eswa.2014.08.042.

[220] Z. Zhang, P. Luo, C.C. Loy and X. Tang, From facial expression recognition to interpersonal relation prediction, *International Journal of Computer Vision* **126**(5) (2018), 550–569. doi:10.1007/s11263-017-1055-1.

[221] J. Zhao, X. Mao and J. Zhang, Learning deep facial expression features from image and optical flow sequences using 3D CNN, *The Visual Computer* **34**(10) (2018), 1461–1475. doi:10.1007/s00371-018-1477-y.

[222] K. Zhao, C. Wen-Sheng, F. De la Torre, J.F. Cohn and H. Zhang, Joint patch and multi-label learning for facial action unit detection, in: *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 2207–2216. doi:10.1109/CVPR.2015.7298833.

[223] Y. Zhao and J. Xu, Necessary morphological patches extraction for automatic micro-expression recognition, *Applied Sciences* **8**(10) (2018), 1811. doi:10.3390/app8101811.

[224] R. Zhi, M. Liu and D. Zhang, A comprehensive survey on automatic facial action unit analysis, *The Visual Computer*. (2019). doi:10.1007/s00371-019-01707-5.

[225] X. Zhu and Z. Chen, Dual-modality spatiotemporal feature learning for spontaneous facial expression recognition in e-learning using hybrid deep neural network, *The Visual Computer*. (2019). doi:10.1007/s00371-019-01660-3.

[226] M.S. Zia and M.A. Jaffar, An adaptive training based on classification system for patterns in facial expressions using SURF descriptor templates, *Multimedia Tools and Applications* **74**(11) (2015), 3881–3899. doi:10.1007/s11042-013-1803-3.