Facial expression recognition based on a mlp neural network using constructive training algorithm

**Hayet Boughrara · Mohamed Chtourou ·
Chokri Ben Amar · Liming Chen**

Abstract This paper presents a constructive training algorithm for Multi Layer Perceptron (MLP) applied to facial expression recognition applications. The developed algorithm is composed by a single hidden-layer using a given number of neurons and a small number of training patterns. When the Mean Square Error MSE on the Training Data TD is not reduced to a predefined value, the number of hidden neurons grows during the neural network learning. Input patterns are trained incrementally until all patterns of TD are presented and learned. The proposed MLP constructive training algorithm seeks to find synthesis parameters as the number of patterns corresponding for subsets of each class to be presented initially in the training step, the initial number of hidden neurons, the number of iterations during the training step as well as the MSE predefined value. The suggested algorithm is developed in order to classify a facial expression. For the feature extraction stage, a biological vision-based facial description, namely Perceived Facial Images PFI has been applied to extract features from human face images. To evaluate, the proposed approach is tested on three databases which are the GEMEP FERA 2011, the Cohn-Kanade facial expression and the facial expression recognition FER-2013 databases. Compared to the fixed MLP architecture and the literature review, experimental results clearly demonstrate the efficiency of the proposed algorithm.

H. Boughrara (✉) · M. Chtourou
Control and Energy Management Laboratory, University of Sfax, ENIS, BP 1173, Sfax 3038, Tunisia
e-mail: hayet.boughrara@laposte.net

M. Chtourou
e-mail: mohamed.chtourou@enis.rnu.tn

C. Ben Amar
Research Groups on Intelligent Machines, University of Sfax, ENIS, BP 1173, Sfax 3038, Tunisia
e-mail: chokri.benamar@ieee.org

L. Chen
Laboratoire d'InfoRmatique en Image et Systémes d'information, University of Lyon,
69134 Ecully, France
e-mail: liming.chen@ec-lyon.fr

1 Introduction

Automatic facial expression analysis is becoming an increasingly important research field from automatic face recognition due to its multiple applications: human-computer intelligent interfaces, video games, human emotion analysis, talking heads or educational software, among others [24].

Basic facial expressions typically recognized in automatic affect-recognition tasks are happiness, sadness, fear, anger, disgust and surprise [14]. Several other emotions and many combinations of emotions have been studied but remain unconfirmed as universally distinguishable. Thus, most of the researches up to now have been oriented towards detecting these basic expressions. Approaches for facial expression recognition from both static images [62] and videos [17] have been proposed in the literature.

An automatic facial expression recognition system generally comprises three crucial steps [57]: face detection, facial feature extraction, and facial expression classification. Face detection is a preprocessing stage to detect or locate the face regions in the input images [61]. Facial feature extraction attempts to find the most appropriate representation of facial images for recognition. There are mainly two approaches: geometric features-based systems and appearance features-based systems. Geometric feature-based methods extract shapes and locations of facial components information including mouth, eyes, eyebrows, nose to form a feature vector. Nevertheless, the geometric features-based systems [45]-[43] require the accurate and reliable facial feature detection. So, it is difficult to realize in real-time applications where illumination changes with time and images are recorded in very low resolution.

Alternatively, the appearance features present the skin texture changes of the face. The appearance features can be extracted on either the whole face or specific regions in a face image. The most frequently used texture features are Gabor filter [70], pixel intensities [19], Discrete Cosine Transform (DCT) features [44], skin color information [29], Haar-like features [66], Local Binary Pattern (LBP) [27, 69], and Local Phase Quantization (LPQ) [71]. Accordingly, feature extraction methods based on Principal Component Analysis (PCA) [6], Linear Discriminant Analysis (LDA) [63], Regularized Discriminant Analysis (RDA) [31] and Independent Component Analysis (ICA) [39], have been used in order to enhance the performance of texture information.

In the last step of a facial expression recognition system, a classifier is employed to identify different expressions based on the extracted facial features. The most known classifiers used for facial expression recognition are the k-nearest neighbor classifier [15], template matching [64], Hidden Markov Models [7], Adaboost algorithms [46], Support Vectors Machines [65] and neural networks [11, 25].

This paper proposes to use a biological vision-based facial description, namely Perceived Facial Images PFI, in the facial expression recognition problem. The proposed PFI simulates the response of complex neurons to gradient information within a certain neighborhood and possess properties of being highly distinctive as well as robust to illumination and geometric transformations [22, 23]. The PFI is an intermediate facial representation to deal with the face image and the feature extraction vector is obtained using the PCA to reduce the dimension of the image. This vector is the input of the neural network classifier.

Due to their superior performance, neural networks have been widely used methods in facial expression recognition [30, 38]. Especially, Multi Layer Perceptron MLP neural network has shown good performance in emotion recognizing [2]. Many traditional algorithms fix the neural network structure before training. However, it is difficult to determine a proper structure in advance, which can both guarantee convergence and avoid over-fitting. Two approaches have been studied in order to determine the adequate neural network structure for a given problem. The first one, known as constructive approach, starts with a small structure and adds hidden neurons when it is necessary. Several constructive learning algorithms have been proposed in the literature [40, 48]. The second approach, known as pruning one, starts with a great structure and eliminates hidden neurons during the training procedure [21, 47].

Constructive algorithms have major advantages over the pruning ones [50]:

– Specifying the initial network is easier in constructive methods, whereas in pruning algorithms one usually has to decide a priori how large the initial network should be.
– Generally, constructive algorithms are more economical in terms of training time and network complexity and structure than pruning algorithms. In fact, small networks have been usually built using constructive algorithms due to their incremental learning nature. While in the construction of a neural network using the pruning algorithm, overly large efforts may be spent in pruning the redundant weights and hidden neurons.
– In constructive algorithms, a smaller number of parameters (weights) is to be updated in the initial stage of the training process thus requiring less training data for good generalization, while a sufficiently large training data is required in pruning algorithms.
– In pruning algorithms, several problems depending on parameters need to be properly specified or selected in order to obtain an acceptable network giving satisfactory performances. This requirement makes these algorithms more difficult to be used in real life applications.

Due to these reasons, the constructive training approach has been considered in this work.

The constructive training algorithm proposed in this paper is essentially based on the idea studied in the work by Liu et al. [35]. The main difference between the proposed algorithm and the one of Liu et al. [35] is the number of examples needed for a successful training step. Indeed, in the work by Liu et al. [35], patterns are trained incrementally by considering them one by one. But, in the proposed constructive training algorithm, patterns are trained incrementally by a subset of each class. In the proposed approach, the recruitment of hidden neurons is based on subsets of training data and not on each pattern as in Liu et al. [35]. A disadvantage of the algorithm of Liu et al. [35] is its high training time when a large training data are under consideration. In the proposed algorithm, huge data training are subdivided into subsets to reduce the learning time.

Authors of reference [41] have developed a modified version of the algorithm of Liu et al. [35] to a speech recognition problem. In this work, the problem of facial expression recognition based on a constructive training algorithm is investigated.

The proposed algorithm starts with a small number of training patterns and a single hidden-layer neural network using a certain number of neurons. During neural network training, the hidden neurons number is increased when the Mean Square Error MSE on the Training Data TD is not reduced to a predefined parameter called ϵ. Input patterns are trained incrementally subset by subset until all patterns of TD are learned.

The main contribution in this paper is to adapt the modified constructive training algorithm to expression recognition problem. This work is interested to the determination of

the MLP structure. Indeed, the proposed MLP constructive training algorithm allows the determination of the number of training patterns in the subsets of each class, the initial number of hidden neurons, the number of iterations during the training steps as well as the MSE threshold ϵ. This paper presents an exploration of the suitable values of these parameters which give the best performances of the neural classifier. Therefore, the process of determining the architecture of the network and the learning process happen simultaneously. This algorithm is used in the classification stage of facial expression recognition system. In addition, a method for applying Perceived Facial Images in the feature extraction stage is presented.

The remainder of the paper is organized as follows. Section 2 introduces the Perceived Facial Images PFI for facial expression recognition application. Section 3 describes the facial expression recognition system developed in this study. Section 4 presents the proposed constructive training algorithm. The experimental results obtained on the GEMEP FERA 2011 database, the Cohn-Kanade database and the FER-2013 database are presented and analyzed in Section 5. To examine the efficiency of the proposed method, comparisons with fixed MLP architecture and some other methods in literature are conducted. Finally, conclusions are drawn in Section 6.

2 Perceived facial images

The feature extraction is an important step of the classification process. In fact, extracting an efficient representation of the face from images contributes to the success of recognition procedure. The present paper suggests making use of a biological vision-based facial description namely Perceived Facial Images PFI, which was initially applied to 3D face recognition [22, 23] and 2D face recognition [3]. As well, the proposed PFI was applied on facial expression recognition using SIFT matching method [5]. The obtained results in the study of [5] were competitive with respect to several methods of the state-of-the art. In order to improve performances, the present paper proposes to use jointly the PFI and the constructive training algorithm in facial expression recognition application.

PFI aims at giving a visual representation simulating the human visual perception. The PFI was inspired from the study of Edelman et al. [13], who proposed a representation of complex neurons in primary visual cortex. These complex neurons respond to a gradient at a particular orientation and spatial frequency, but the location of the gradient is allowed to shift over a small receptive field rather than being precisely localized.

The proposed representation PFI aims at simulating the response of complex neurons, based on a convolution of gradients in specific directions in a given circular neighborhood. The precise radius value of the circular area needs to be experimentally fixed. Specifically, given an input image I, a given number of gradient maps L_1, L_2, ..., L_o, one for each quantized direction o, are first computed. They are defined as:

$$L_o = \left(\frac{\partial I}{\partial o} \right)^+ \tag{1}$$

The $+$ sign means that only positive values are kept to preserve the polarity of the intensity changes.

Each gradient map describes gradient norms of the input image in a direction o at every pixel location. Then, the response of complex neurons is simulated by convolving its

gradient maps with a Gaussian kernel G. The standard deviation of G is proportional to the radius of the given neighborhood area, R, as in (2).

$$\rho_o^R = G_R * L_o, \tag{2}$$

where $*$ denotes the convolution operator.

The purpose of the convolution with Gaussian kernels is to allow the gradients to shift within a neighborhood without abrupt changes. At a given pixel location (x, y), we collect all the values of the convolved gradient maps at that location and form the vector $\rho^R(x, y)$ thus having a response value of complex neurons for each orientation o.

$$\rho^R = \left[\rho_1^R(x, y), \rho_2^R(x, y), ..., \rho_o^R(x, y) \right] \tag{3}$$

This vector, $\rho^R(x, y)$, is further normalized to unit norm vector, which is called the subsequent response vector and denoted by $\underline{\rho}^R$.

Facial image can be represented by its perceived values of complex neurons according to the response vectors. Specifically, given a facial image I, Perceived Facial Images J_o are generated using complex neurons for each orientation o defined as in (4).

$$J_o(x, y) = \underline{\rho}_o^R(x, y) \tag{4}$$

Figure 1 illustrates the process applied to a facial image. In this work, eight PFIs corresponding to eight directions are computed. Therefore, for each image from the database, eight images correspond to eight PFIs are obtained. It has been demonstrated that we can not decide which direction is to be used [3, 5, 22]. The proposed approach which will be described in the next section, is based on the use of the eight PFI directions independently.
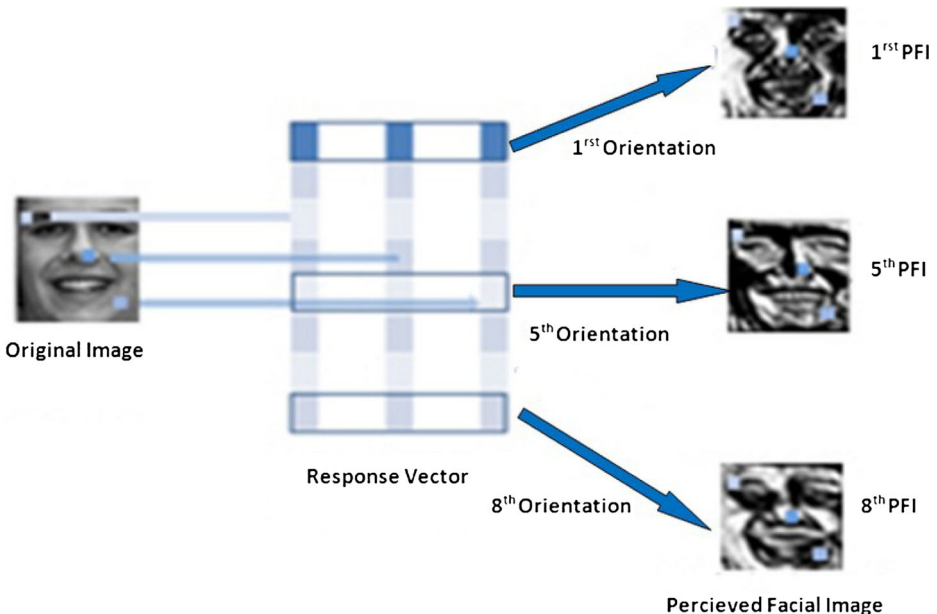

Fig. 1 An illustration of the PFI in eight orientations

3 Facial expression recognition system

A facial expression recognition system generally consists of three stages: face detection, feature extraction and feature classification. In this paper, the face detection is based on the OpenCV face detector. In the second stage, the PFI is used to extract the feature vector of the faces. In fact, eight PFIs are generated for each image in the database. After that, the PCA is applied to reduce the dimension of the image. For each direction, a classifier is developed and the final decision is obtained from the fusion of the ones corresponding to the eight directions.

Having the extracted feature vector for all the samples in the training and testing sets, the next step would be to design the classifier. In this study, Multi Layer Perceptron MLP architecture has been used as a classifier of facial expressions.

Three layer MLP have been used in this study. The number of input neurons is equal to the size of related feature vector. Similarly, the number of output neurons is equal to the number of facial expressions to be recognized. In the learning phase, the desired output neuron has 1 for the correct input pattern and 0 for all others output neurons. The hidden layer is constructed using the proposed constructive training algorithm which will be presented in the next section.

There are two steps on the realization of the facial expression recognition system using the MLP architecture: the training step and the testing step. The learning algorithm used this study is the standard back-propagation [36]. The back-propagation network undergoes a supervised learning process. The training algorithm is accomplished based on the following expressions:

$$\triangle \omega_{ji}(t) = -\eta \frac{\partial E_p(t)}{\partial \omega_{ji}(t)}, \tag{5}$$

where η is the learning rate and E_p designed the error of the network for the pth pattern (MSE) and defined as in (6).

$$E_p = \frac{1}{2} \sum (d_{pk} - s_{pk})^2, \tag{6}$$

where k is the number of output neurons in the MLP. d_p and s_p are the desired and the neural computed outputs for p^{th} training vector.

For the testing step, the class of each presented pattern is assigned to the maximal value of the neuron outputs. The performance of the recognition system has been measured in terms of recognition rate on testing set.

Three databases have been used to evaluate the proposed approach: FERA 2011 database, the Cohn-Kanade database and the FER-2013 database. The fera 2011 database is composed by video of facial expressions. The emotion detection concerns five discrete emotion classes. Each video has a single emotion label $e \in E$, where $E = \{$Anger, Fear, Joy, Relief, Sadness$\}$. Since the videos do not display any apparent neutral frames at the beginning or the end of the video, it will be supposed that every frame of a video shares the same label [59].

The classification rate is first obtained per emotion then the average over all five emotions is computed. The classification rate for emotions is calculated as the fraction of the number of videos correctly classified divided by the total number of videos for each emotion in the test set.

The resulting MLP classifier gives a decision $y_{e,j}$ about the presence of emotion e for frame j in a test video. To decide the label Y of a test video composed by n frames, we find the emotion with the largest number of classified frames:

$$Y = \max_{e} \sum_{j=1}^{n} y_{e,j} \tag{7}$$

For the two other databases, the recognition rate is defined by the ratio between the number of true classification by the total number of images in the testing sets.

4 Modified constructive training algorithm

Many researchers have studied the neural network training problem and many algorithms have been reported [36]. Among them, Back-propagation algorithm [36] is one of the most commonly used method. Multi Layer Perceptron neural network, requires the definition of the network architecture, prior to training. Generally this method works well only when the network architecture is appropriately chosen. It is well known that there is no general answer to the problem of defining neural network architecture for a given application.

In fact, the usual way to determine the number of hidden units in Multi Layer Perceptron MLP neural network is by a trial and error procedure. An alternative is to use constructive algorithms [4, 35] which try to solve the problem by building the architecture of the neural network during its training.

Constructive training algorithm incrementally adds hidden neurons and weights to the network during training until stopping criterion is satisfied.

The training pedagogy adopted in the proposed MLP constructive training algorithm is based on the following idea: the presentation of patterns to the classifier is accomplished by dividing patterns corresponding to each class to a given number of subsets. Each subset contains (N_patt) patterns. The training is performed by presenting the first subset of each class then the second subsets and so on. The proposed algorithm starts with a single hidden-layer using a (Nh_1ini) number of neurons. Over the learning process, the hidden neurons number is increased when the MSE threshold on the TD does not reach a predefined parameter called ϵ. Input patterns are learned incrementally subset by subset until all patterns (N_tot) of TD are presented [41]. The pseudo-code for the proposed algorithm is shown by Algorithm 1.

Algorithm 1 Constructive Training Algorithm

```
 1:  N_hid ← Nh_ini;
 2:  N_Input ← N_patt;
 3:  while N_Input ≤ N_tot do
 4:      Train the MLP with Backpropagation algorithm;
 5:      if MSE > ε then
 6:          N_hid ← N_hid + 1;
 7:      else
 8:          N_Input ← N_Input + N_patt;
 9:      end if
10:  end while
```

As presented in Algorithm 1, the proposed algorithm is constituted by eleven procedures. It can be detailed by the following steps:

– *Step 1*: Create a MLP composed by ($N\_hid = Nh\_ini$) hidden neurons.
– *Step 2*: Initialize the neuron connection and bias weights with random values.
– *Step 3*: Select N_patt input patterns from the TD ($N\_Input = N\_patt$) for each class
– *Step 4*: By setting the number of iterations by N_epochs, train the MLP using the backpropagation algorithm with N_Input input patterns to achieve the predefined performance .
– *Step 5*: The final connection and bias weights of the MLP architecture are stored.
– *Step 6*: If the training algorithm can reduce the MSE to ϵ , go to *Step 7*; otherwise, go to *Step 9*.
– *Step 7*: While the $N\_Input \preceq N\_tot$: the total number of patterns from TD, go to *Step 8* otherwise go to *Step 11*.
– *Step 8*: Increase the number of input patterns ($N\_Input = N\_Input + N\_patt$) and go back to *Step 4*.
– *Step 9*: Increase the number of hidden neurons by one ($N\_hid = N\_hid + 1$).
– *Step 10*: Initialize the weights of the new hidden neuron and all connection weights are replaced by the last stored weights obtained from *Step 5* and go back to *Step 4*.
– *Step 11*: Take the found architecture corresponding to an appropriate number of neurons.

The proposed MLP constructive training algorithm has to determine the adequate value of (N_patt), the initial number of hidden neurons (Nh_ini), the number of iterations (N_epochs) during the training step as well as the MSE threshold value ϵ. This paper presents an exploration procedure to determine the suitable values of these parameters giving the best performances.

5 Experimental results and discussion

The proposed approach will be experimentally evaluated using three databases which are the GEMEP FERA 2011 database, the Cohn-Kanade facial expression Database and the FER-2013 Database. All experiments have been developed using MATLAB (version 2010) and its Neural Network toolbox.

5.1 Databases description

– *GEMEP FERA 2011 database:* This database is a partition of the GEMEP corpus [1], developed by the Geneva Emotion Research Group at the University of Geneva. The challenge is divided in two sub-challenges that reflect two popular approaches to facial expression recognition: an action unit detection sub-challenge and an emotion detection sub-challenge. In this work, the emotion detection is considered, which calls for systems to attain the highest possible classification rate for the detection of five discrete emotions: anger, fear, joy, relief, and sadness.

To be able to objectively measure the performance of the participants' entries, the database has been divided into a training set and a test set. A total of 288 portrayals were selected (155 for training and 134 for testing).

The test data is divided in three different partitions. The first one is the partition where the test subjects are not present in the training data (Person Independent partition

Table 1 Number of Videos of Each Emotion used in the training set, the PS, The PI and the Overall test set [59]	Emotion	Train	PS Test	PI Test	Overall Test
	Anger	32	13	14	27
	Fear	31	10	15	25
	Joy	30	11	20	31
	Relief	31	10	16	26
	Sadness	31	10	15	25
	Total	155	54	80	134

PI). The second partition of the test data consists of videos of subjects that are part of the training set (Person Specific partition PS). The third partition is simply the entire (overall) test set [59]. Details about the training and test sets of the GEMEP FERA database can be found in Table 1.

– *The Cohn-Kanade facial expression database:* This database is one of the most comprehensive database in the current facial-expression-research community [26]. It consists of 97 classes. For each subject, one neutral face and six expressive faces have been presented in the database. These facial expressions are happiness, anger, sadness, surprise, fear, and disgust. Figure 2 shows the samples of six expressions and the neutral for the Cohn-Kanade Database.

– *The FER-2013 database:* The dataset of the Facial Expression Recognition FER-2013 Challenge [54] consists of 48×48 pixel gray-scale images of faces representing 7 categories of facial expressions: anger, disgust, fear, happiness, sadness, surprise, and neutral. There are 28709 examples for training, 3589 examples for public testing, and another 3589 examples for private testing. The faces have been automatically registered from the web so that each face is approximately centered and occupies about the same amount of area within each image. The training data may also contain labeling noise, meaning that the labels of some faces do not indicate the right facial expression. Another issue is that the images have some variation of poses and the presence of external occlusions caused by eyeglasses, hand and hair.

Fig. 2 Examples of six facial expressions of the Cohn-Kanade facial expression database

5.2 Pre-processed database

In this work, the OpenCV face detector is used to extract the face location in each image [32]. The detected face is scaled to be 200 by 200 pixels. This pre-processed step has been applied in each image from the FERA 2011 and the Cohn-Kanade databases. For the FER-2013 database, the face is initially detected for each image. Figure 3 presents the obtained image when applied the OpenCV face detector.

5.3 Results using GEMEP FERA 2011 database

After pre-processing, the Perceived Facial Image PFI in eight directions is applied for each frame of a video in training and test data. For each orientation, a MLP network using constructive training algorithm is developed. After applying the PCA to reduce the dimensionality, feature vectors of dimension 12 are obtained.

The recognition rates are calculated using the Person Specific partition PS, Person Independent partition PI and the entire overall test set, respectively.

5.3.1 Experimental setup

The proposed constructive training algorithm is used in order to determine the number of trained input patterns (N_patt), the initial number of hidden neurons (Nh_ini), and the number of epochs (N_epochs) as well as the appropriate value of the MSE (ϵ). The exploration of these parameters has been made for each direction of PFI. So, the next experiments have been accomplished using the PFI on the first direction.

In the first step, the number of input patterns (N_patt) for each subset corresponding to each class used in the training step is varied and the other parameters are fixed. In this work, 30 videos for each face expression are used and each video is formed by 20 frames. Therefore, in this experiment 600 images are used in each face expression. Then the training set is formed by 3000 images.

For ϵ equal to 0.01, we chose to start the algorithm with 20 hidden neurons ($Nh\_ini = 20$), 1000 epochs ($N\_epochs = 1000$) and the learning rate η is equal to 0.01. The performances are measured by determining the recognition rate on each partition of test set (PS, PI and Overall), the final number of hidden neurons (Nh_end), which has been obtained at the end of learning step and the training time.

Fig. 3 The original face image and the detected face image

Some simulation studies have been done by varying the value of N_patt which is equal to 50, 100, 150, 200 and 300. For each value of N_patt, the performances have been calculated. So, the appropriate value of the number of input patterns presented for each class has determined, so $N\_patt = 300$. The obtained recognition rates are PS $= 85.18\%$, PI $= 64.55\%$ and Overall $= 72.93\%$. At the end of training procedure, the algorithm converges to 38 hidden neurons ($Nh\_end = 38$).

In the second step, the variation of the initial number of hidden neurons (Nh_ini) is considered. The algorithm starts with 300 input patterns ($N\_patt = 300$) for each class and 1000 epochs ($N\_epochs = 1000$). Different values of Nh_ini have been presented to the MLP network. The value of Nh_ini is varied from 10 to 40 neurons and for each value the performances have been computed. To conclude, the best performances (PS $= 81.48\%$, PI $= 73.41\%$ and Overall $= 76.69\%$) have been obtained when ($Nh\_ini = 21$). At the end of training procedure, the proposed constructive algorithm converges to 36 hidden neurons ($Nh\_end = 36$).

The last procedure is to determine the number of epochs (N_epochs) using in the training. The proposed algorithm is trained using 300 input patterns for each class and an initial number of hidden neurons equal to 21. Several sets of experiments have been made with different values of N_epochs which are varied from 100 to 1000. The best rates have been obtained when ($N\_epochs = 450$). At the end of training procedure, the algorithm converges to 29 hidden neurons ($Nh\_end = 29$).

To conclude, for ϵ equal to 0.01, the best performances are obtained for 300 input patterns for each class in the training step, 21 initial hidden neurons and 450 epochs: $N\_patt = 300$, $Nh\_ini = 21$ and $N\_epochs = 450$. Table 2 illustrates the confusion matrix for Overall partition test using the PFI 1.

Based on these results, the obtained rates are PS $= 100\%$, PI $= 98.45\%$ and Overall $= 99.25\%$. The classification rate on the person specific partition PS is better than the rate on the Person Independent partition PI.

This is can be explained by the difficulty of the facial expression recognition of a subject who has not been used in the training set.

It is to be noted, that the chosen value of ϵ has been done based on several simulations. For ($\epsilon < 0.01$), the training algorithm could not converge. This is due to the trade-off that should be respected, between the number of examples which is limited in the database and the structure of neural network.

5.3.2 Results evaluation

Previous results are made using the PFI on the first direction. Then, the same procedure is repeated with the eight orientations. Subsequently, the proposed constructive training

	Anger	Fear	Joy	Relief	Sadness
Anger	27	0	0	0	0
Fear	0	25	0	0	0
Joy	0	0	31	0	0
Relief	0	0	0	26	1
Sadness	0	0	0	0	24

Table 2 Confusion matrix for Overall emotion Recognition using the PFI 1 and the proposed constructive training algorithm

algorithm is applied on eight directions separately. Table 3 presents the classification rates for emotion recognition using the PFI on eight directions and the proposed algorithm with ($\epsilon = 0.01$).

The obtained parameters corresponding to the training algorithm are illustrated on the Table 4. These parameters are the number of patterns presented for each class N_patt, the initial number of hidden neurons Nh_ini, the number of epochs used in the training N_epochs, the final number of hidden neurons Nh_end achieved at the end of learning and the training time.

Observing Tables 3 and 4, the best rates are PS = PI = Overall = 100 %. These rates have been obtained using the PFI on direction 3. At the end of learning procedure, the algorithm converges to 39 hidden neurons.

After that, fusion which is the sum of the obtained results of each neural network corresponding in each orientation of the PFI, is computed. The obtained rates using fusion are PS = PI = Overall = 100 %.

Table 3 Classification rates for emotion recognition on the PS, PI and Overall test set using the proposed constructive training algorithm and the PFI on eight directions separately.

		Anger	Fear	Joy	Relief	Sadness	Average
PFI 1	PS	100	100	100	100	100	100
	PI	100	100	100	100	90	98.45
	Overall	100	100	100	100	96	99.25
PFI 2	PS	100	100	100	100	100	100
	PI	100	100	100	0	100	81.01
	Overall	100	100	100	40	100	88.72
PFI 3	PS	100	100	100	100	100	100
	PI	100	100	100	100	100	100
	Overall	100	100	100	100	100	100
PFI 4	PS	100	100	100	100	90	98.14
	PI	100	100	100	0	100	81.01
	Overall	100	100	100	40	96	87.96
PFI 5	PS	100	100	100	100	100	100
	PI	42.85	100	100	100	53.33	81.01
	Overall	70.37	100	100	100	72	88.72
PFI 6	PS	100	100	100	100	100	100
	PI	100	100	100	80	0	77.21
	Overall	100	100	100	88	40	86.46
PFI 7	PS	100	100	100	100	100	100
	PI	100	100	100	60	100	92.40
	Overall	100	100	100	76	100	95.48
PFI 8	PS	100	100	100	100	100	100
	PI	0	100	100	100	100	82.27
	Overall	48.14	100	100	100	100	89.47

Table 4 The obtained parameters from the proposed constructive training algorithm for each orientation

	PFI 1	PFI 2	PFI 3	PFI 4	PFI 5	PFI 6	PFI 7	PFI 8
N_patt	300	300	300	300	300	300	300	300
Nh_ini	21	20	27	20	21	20	21	20
Nh_end	29	38	39	27	37	36	27	25
N_epochs	450	400	400	100	300	200	100	250
$Training$	0.45	0.68	0.33	0.24	1.47	1.01	0.28	0.44
$Time$	mn	mn	mn	mn	mn	mn	mn	mn

5.3.3 Comparison with a fixed MLP architecture

In order to show the advantage of the proposed approach, a fixed MLP architecture is applied on facial expression recognition using the PFI 3. The MLP is trained using the same number of neurons in the hidden layer ($Nh\_end = 39$) obtained at the end of the proposed constructive training algorithm. The learning rate η is equal to 0.01. The stop criterion is the value of MSE which is equal to 0.01. The fixed MLP architecture needs a high number of iterations to converge to ϵ. Should be noted that the initial weights values used in the constructive training algorithm have been considered in the training of the fixed structure MLP neural network. The confusion matrix for Overall partition test is presented on the Table 5.

The obtained classification rates are PS = 98.14 %, PI = 55.69 % and Overall = 72.93 %. To compare to the proposed training algorithm results presented on Table 2, the best rates have been obtained using the constructive training algorithm. Also, it can be noted that a large portion of anger, fear and sadness are mis-classified to fear, anger and relief respectively with the fixed MLP architecture and but they are successfully classified with the proposed method. This is explained by the fact that the fera-2011 database is particularly difficult to treat due in particular to the existence of a strong intra-class confusion (between joy/relief on one side, and anger/fear/sadness of the other), which makes it particularly difficult recognition even for a human [59]. Despite these difficulties, the proposed algorithm gives good performances.

Next, a fixed MLP architecture using the PFI on eight orientations is developed. The classification rates are computed on the PS, PI and Overall test set. Table 6 presents the obtained results on each direction separately.

The obtained number of hidden neurons at the end of the learning of the proposed algorithm is considered in the structure of the fixed MLP neural networks ($N\_hid = Nh\_end$).

The learning of the fixed MLP neural network has been ended when the MSE value reaches ϵ. Comparing the obtained results in Table 6, the best rates (PS = 100 %,

Table 5 Confusion matrix for Overall emotion recognition using the PFI 3 and the fixed MLP architecture

	Anger	Fear	Joy	Relief	Sadness
Anger	16	15	0	0	0
Fear	11	10	0	0	0
Joy	0	0	31	0	0
Relief	0	0	0	26	10
Sadness	0	0	0	0	15

Table 6 Classification rates on the PS, PI and Overall test set using the fixed MLP architecture and the PFI on eight orientations separately.

	N_hid	N_epochs	PS	PI	Overall	Training Time (mn)
PFI 1	29	1800	87.03	44.03	62.90	0.87
PFI 2	38	2200	75.92	50.63	60.90	1.03
PFI 3	39	2100	98.148	55.69	72.93	0.92
PFI 4	27	3500	87.03	44.30	61.65	1.34
PFI 5	37	6000	57.40	62.02	60.15	2.02
PFI 6	36	3500	90.74	43.03	62.40	1.44
PFI 7	27	6500	100	75.94	85.71	1.71
PFI 8	25	6000	59.25	43.03	49.62	1.47

PI = 75.94 % and Overall = 85.71 %) have been obtained using the PFI on direction 7. Also, the fixed MLP architecture requires more number of iterations than the proposed constructive algorithm so that it can converge to ϵ.

For each orientation, it is clear that the obtained rates using the proposed algorithm are better than those obtained by the fixed MLP. Moreover, the training time for the constructive training algorithm is lower than the one of the fixed MLP architecture. So, the achieved results show the effectiveness of the proposed constructive training algorithm compared to the fixed MLP architecture.

5.3.4 Comparison with the literature

FERA database has been used for automatic facial expression recognition by many researches [12]–[67]. Table 7 gives performances obtained by the proposed constructive training algorithm and the existing systems on emotion recognition using GEMP FERA database.

Table 7 Comparison with the literature: classification rates for emotion Recognition using FERA 2011 database

Systems	Recognition rate (%)		
	PS	PI	Overall
Dhall et al. [12]	83.8	64.9	73.4
Dahmane et Meunier [8]	87	57.9	70
Littlewort et al. [37]	83.7	71.4	76.1
Meng et al. [42]	83.7	60.9	70
Srivastava et al. [52]	73	63.6	67.2
Tariq et al. [56]	100	65.5	79.8
Yang et Bhanu [67]	96.2	75.2	83.8
Baseline [59]	73	44	56
Boughrara et al. [5]	100	60	76.11
Proposed approach	100	100	100

Using the proposed constructive training algorithm, the best obtained rates (PFI 3) are considered on Table 7. To conclude, the obtained classification rates on the PS, PI and Overall test set using the proposed approach are better than other methods.

5.4 Results using Cohn-Kanade database

The second database used in this study is the Cohn-Kanade database. In this experiment, we evaluate the proposed algorithm for the person-independent expression recognition, which requires that the same persons with the same expressions should not simultaneously appear in both the training set and the testing set.

The classification performance has been evaluated using a 10-fold cross validation method and the average recognition results were reported. In this work, 2700 images have been selected from the database corresponding to six expressions. So, 450 images have been presented in each expression.

After pre-possessing step applied in each image from the database, the PFI on eight orientations has been computed. PCA algorithm has been executed to reduce the dimensionality. The proposed constructive training algorithm has been developed for each orientation separately.

5.4.1 Results evaluation

The same procedures applied in the first database (FERA 2011) have been repeated. The value of ϵ has been chosen by the experiences which is equal to 0.001. After some experiment studies, the appropriate value of the number of input patterns initially presented for each subset N_patt, the adequate number of hidden neurons Nh_ini used initially and the number of epochs N_epochs have been determined. For each PFI, the recognition rates using the constructive training algorithm have been calculated and illustrated in Table 8. The obtained parameters of the proposed algorithm have been also presented in the table.

Table 8 demonstrates that the best recognition rate is obtained for the PFI on direction 6. When fusion has been calculated, the recognition rate reaches 96.66 %.

Next, the fixed MLP architecture has been developed in each direction separately. The MLP network is constructed using the same number of neurons in the hidden layer (Nh_end) obtained at the end of the proposed learning algorithm. The number of epochs is

Table 8 Classification rates on the Cohn-Kanade database using the proposed constructive training algorithm for each orientation

	N_patt	Nh_ini	Nh_end	N_epochs	Recognition Rate	Training Time
PFI 1	200	32	35	350	94.07 %	0.52 mn
PFI 2	200	28	29	450	90.74 %	0.40 mn
PFI 3	200	21	23	300	90 %	0.34 mn
PFI 4	200	37	43	200	91.11 %	0.54 mn
PFI 5	200	39	40	400	89.25 %	0.30 mn
PFI 6	200	28	38	250	95.18 %	0.86 mn
PFI 7	200	39	43	150	94.81 %	0.30 mn
PFI 8	200	25	25	450	90 %	0.55 mn

Table 9 Classification rates on the Cohn-Kanade database using the fixed MLP architecture and the PFI on eight orientations separately.

	N_hid	N_epochs	Recognition Rate (%)	Training Time (mn)
PFI 1	35	1400	82.53	0.67
PFI 2	29	2500	79.25	0.78
PFI 3	23	1200	76.46	0.56
PFI 4	43	3000	85.55	1.03
PFI 5	40	3500	82.59	1.11
PFI 6	38	3500	86.66	1.12
PFI 7	43	4000	89.62	1.38
PFI 8	25	3500	80.64	1.02

determined by the experiment when the fixed MLP network reaches the convergence. The results using the fixed MLP architecture are presented in the Table 9.

Observing the Table 9, the best result (89.62 %) has been achieved using the direction 7. The training time of the fixed architecture is higher than the one of the proposed algorithm. By carrying out the fusion, the recognition rate is equal to 82.78 %. To conclude, the constructive algorithm is better than the fixed MLP architecture in terms of recognition rate and the training time.

5.4.2 Comparison with the literature

The proposed framework is compared with the other references using the same database (Cohn-Kanade database) and the obtained results are presented in Table 10. The obtained rate when applying fusion is presented in Table 10.

Seeing Table 10, the proposed approach gives a best recognition rate compared to the other methods in the literature such as the works in [28, 51, 60, 69]. This demonstrates the efficiency of the proposed constructive training algorithm. But the obtained rate is somewhat lower than the one of Khan et al. [27].

Table 10 Comparison with the literature: classification rates using CohnKanade database

Systems	Recognition rate (%)
Zhao et al. [69]	95.85
Villegas et al. [60]	88.5
Shan et al. [51]	92.6
Gu et al. [20]	91.51
Liu et al. [38]	82
Fan et al. [16]	86.28
Bejani et al. [2]	71.8
Khan et al. [27]	96.7
Lajevardi et al. [28]	91.9
Proposed Approach	96.66

Table 11 Classification rates on the Public and the Private testing set using the proposed constructive training algorithm for each orientation

	N_patt	Nh_ini	Nh_end	N_epochs	Public Test (%)	Private Test (%)	Training Time (mn)
PFI 1	500	28	36	250	81.01	79.63	1.23
PFI 2	500	29	36	150	77.45	78.30	0.69
PFI 3	500	25	37	200	79.79	79.14	1.48
PFI 4	500	29	36	250	75.26	76.97	1.08
PFI 5	500	29	32	200	79.63	76.77	0.49
PFI 6	500	28	36	150	80.97	78.58	0.73
PFI 7	500	30	44	200	78.75	76.17	1.85
PFI 8	500	26	33	200	79.92	79.43	1.02

5.5 Results using FER-2013 database

The last database used in this work is the FER-2013 database. In this study, 1000 images have been used for each expression in the training set, 200 images for each expression have been selected in the public and private testing set.

PFI in eight directions is applied for each image in training and testing data. Then, PCA algorithm has been applied to reduce the dimensionality. Next, a neural network has been built based on the constructive training algorithm in each direction.

5.5.1 Results evaluation

For the value of ϵ equal to 0.01, appropriate values of the number of input patterns initially presented for each class N_patt, the adequate initial number of hidden neurons Nh_ini and the number of epochs N_epochs have been identified based on exploration. For each PFI, the recognition rates using the constructive training algorithm have been calculated and illustrated in Table 11.

Comparing the obtained results for each orientation, the best recognition performances (Public test = 81.01 % and Private test = 79.63 %) are obtained on direction 1. At the end

Table 12 Classification rates on the Public and the Private testing set using the fixed MLP architecture and the PFI on eight orientations separately.

	N_hid	N_epochs	Public test (%)	Private test (%)	Training Time (mn)
PFI 1	36	600	77.06	76.30	0.35
PFI 2	36	450	77.15	74.21	0.29
PFI 3	37	500	76.78	74.83	0.31
PFI 4	33	600	71.89	72.74	0.37
PFI 5	32	600	76.01	74.43	0.35
PFI 6	36	500	75.92	74.34	0.29
PFI 7	44	750	77.06	74.67	0.5
PFI 8	33	600	76.87	77.98	0.35

Table 13 Comparison with the literature: classification rates using FER-2013 database

Systems	Recognition rate (%)	
	Public Test	Private Test
RBM [55]	69.769	71.162
Unsupervised [54]	69.072	69.267
Maxim Milakov [54]	68.153	68.821
Radu + Marius + Cristi [58]	67.317	67.484
Proposed approach	84.58	82.76

of training procedure, the algorithm converges to 36 hidden neurons ($Nh\_end = 36$). By computing the fusion, the recognition rates are (Public test = 84.58 % and Private test = 82.76 %). So, these rates are better than those obtained using the eight directions.

On the other hand, the fixed MLP architecture is applied on facial expression recognition using the PFI on eight directions. The fixed MLP is trained used the same number of neurons on the hidden layer (Nh_end) obtained at the end of the proposed learning algorithm. Table 12 presents the results using the fixed MLP architecture.

Based in the Table 12, it can be concluded that the achieved rates using the proposed constructive training algorithm for each orientation are better than the ones obtained by the fixed MLP architecture. So, the suggested algorithm is more efficient.

5.5.2 Comparison with the literature

Compared to the literature review, Table 13 gives performances obtained by the proposed constructive training algorithm and the existing ones using FER-2013 database on public and private testing set. The existing systems have used all the images in the database. Then, the obtained scores are computed on all the images in the test set.

The recognition rates obtained by the proposed approach on public an private sets are slightly better than the one obtained in literature. It should be noted that obtained results have considered reduced training and testing sets. Due to training time, we cannot use all the images in the FER-2013 database. This objective will be under consideration as a future work.

6 Conclusion

The MLP neural network based on backpropagation algorithm is one of the most popular neural networks topologies. One of the difficulties of using the MLP neural network is to determine the optimal number of hidden neurons before the training process. Two approaches have been developed in the literature to solve this problem which are the constructive and the pruning algorithms.

In this paper, a constructive training algorithm for MLP neural networks has been proposed. Starting with a neural network containing a given number of hidden neurons and a small number of training patterns, the MLP neural network using the back-propagation algorithm is trained. The hidden neurons grow during the training when the MSE on the TD is not reduced to a predefined value. Input patterns are trained incrementally until all patterns of TD are selected and learned.

The suggested constructive learning algorithm for MLP neural networks has been applied in the classification step for the facial expression recognition system. A biological vision-based facial description, namely Perceived Facial Images PFI is applied in the feature extraction stage. This paper uses the PFI on eight directions applied to extract features from facial expression images. Suitable parameters for the training of the neural classifier for each direction are determined. The final decision the facial expression recognition system is computed by the fusion of those obtained by the neural networks corresponding to the eight directions.

The GEMEP FERA, the Cohn-Kanade and the FER-2013 databases have been used for the experiment. Compared to the fixed MLP architecture, the best recognition rate has been obtained using the constructive training algorithm.

The importance of the proposed constructive algorithm is explained by the fact that learning patterns are presented sequentially to the neural network. Adopting this training procedure allows to reduce considerably the training time. Therefore, the obtained results have been compared favorably with those obtained from the fixed MLP architecture and with other state-of-the-art works that prove the effectiveness of the proposed method.

References

1. Banziger T, Scherer KR (2010) Introducing the geneva multimodal emotion portrayal (gemep) corpus, Blueprint for Affective Computing, A Sourcebook, Series in affective science, chapter 6.1 pp 271–294. Oxford University Press, Oxford
2. Bejani M, Gharavian D, Charkari NM (2012) method, Audiovisual emotion recognition using ANOVA feature selection networks, multi-classifier neural Neural Computer & Application. Springer-Verlag London
3. Boughrara H, Chen L, Ben Amar C, Chtourou M (2012) Face Recognition under varying Facial Expression Based on Perceived Facial Images And Local Feature Matching, International Conference on Information Technology and e-Service, pp. 1–6, 2nd edition. Sousse Tunisia
4. Boughrara H, Chtourou M, Ben Amar C (2012) MLP Neural Network Based Face Recognition System Using Constructive Training algorithm, International Conference on Multimedia Computing and System ICMCS, pp. 233–238, 3rd edition. Tangier Morocco
5. Boughrara H, Chen L, Ben Amar C, Chtourou M (2013) Facial Expression Recognition Based on Perceived Facial Images and Local Feature Matching, International Conference on Image Analysis and Processing, ICIAP, Part II, LNCS 8157. Springer-Verlag Berlin, Heidelberg, pp 591–600
6. Chakrabarti D, Dutta D (2013) Facial Expression Recognition Using Eigenspaces Procedia Technology vol 10, pp 755–761
7. Chumkamon S, Hayashi E (2013) Facial expression recognition using constrained local models and Hidden Markov models with consciousness-based architecture IEEE/SICE International Symposium on System Integration (SII), pp 382–387
8. Dahmane M, Meunier J (2011) Emotion recognition using dynamic gridbased hog features IEEE Int'l Conf. Automatic Face and Gesture Analysis, pp 884–888
9. Danisman T, Bilasco IM, Martinet J, Djeraba C (2013) Intelligent pixels of interest selection with application to facial expression recognition using multilayer perceptron. Signal Process 93(6):1547–1556
10. Danisman T, Bilasco IM, Martinet J, Djeraba C (2013) Intelligent pixels of interest selection with application to facial expression recognition using multilayer perceptron. Signal Process 93(6):1547–1556
11. Danisman T, Bilasco IM, Martinet J, Djeraba C (2013) Intelligent pixels of interest selection with application to facial expression recognition using multilayer perceptron. Signal Process 93(6):1547–1556
12. Dhall A, Asthana A, Goecke R, Gedeon T (2011) phog Emotion recognition using, features, lpq IEEE Int'l Conf Automatic Face and Gesture Analysis, pp 878–883
13. Edelman S, Intrator N, Poggio T (1997) Complex cells and object recognition, unpublished manuscript. http://kybele.psych.cornell.edu/edelman/archive.html
14. Ekman P (1992) Facial expressions of emotion: an old controversy and new findings. Philos Trans: Biol Sci 335:63–69

15. Eskil MT, Benli KS (2014) Facial expression recognition based on anatomy Computer Vision and Image Understanding vol 119, pp 1–14
16. Fan W, Bouguila N (2013) Face detection and facial expression recognition using simultaneous clustering and feature selection via an expectation propagation statistical learning framework Multimed Tools Appl
17. Fang H, Parthalïn NM, Aubrey AJ, Tam GKL, Borgoa R, Rosin PL, Grant PW, Marshall D, Chend M (2014) Facial expression recognition in dynamic sequences: An integrated approach. Pattern Recognit 47(3):1271–1281
18. Farajzadeh N, Pan G, Wu Z (2013) Facial expression recognition based on meta probability codes, Pattern Analysis Application. Springer-Verlag London
19. Gizatdinova Y, Surakka V (2006) Feature-based detection of facial landmarks from neutral and expressive facial images. IEEE Trans Pattern Anal Mach Intell 28:135139
20. Gu W, Xiang C, Venkatesh YV, Huang D, Lin H (2012) Facial expression recognition using radial encoding of local Gabor features and classifier synthesis. Pattern Recognit 45:80–91
21. Honggui H, Junfei Q. (2009) A Novel Pruning Algorithm for Self-organizing Neural Network, Proceedings of International Joint Conference on Neural Networks, Atlanta, Georgia USA, pp 1245–1250
22. Huang D, Ben Soltana W, Ardabilian M, Wang YH, Chen L (2011) Textured 3D face recognition using biological vision-based facial representation and optimized weighted sum fusion Proc. CVPR Workshop on Biometrics, Colorado Springs, CO USA
23. Huang D (2011) Robust Face Recognition Based on Three Dimentional Data,thesis informatic, LRIS. university of Lyon
24. Jaimes A, Sebe N (2007) Multimodal human-computer interaction: a survey. Comput Vis Image Underst 108:116–134
25. Kaburlasos VG, Papadakis SE, Papakostas GA (2013) Lattice computing extension of the FAM neural classifier for human facial expression recognition. IEEE Trans Neural Netw Learn Syst 24(10):1526–1538
26. Kanade T, Cohn J, Tian YL (2000) Comprehensive database for facial expression analysis, International Conference on Automatic Face and Gesture Recognition, 4th edition, pp 46–53
27. Khan RA, Meyer A, Konik H, Bouakaz S (2013) Framework for reliable real-time facial expression recognition for low resolution images. Pattern Recognit Lett 34:11591168
28. Lajevardi SM, Hussain ZM (2012) Automatic facial expression recognition: feature extraction and selection. SIViP 6:159–169
29. Lee SH, Kim H, Man Ro Y, Plataniotis KN (2013) Using color texture sparsity for facial expression recognition, 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG), pp 1–6
30. HC, Wu CY, Lin TM (2013) Facial Expression Recognition Using Image Processing Techniques and Neural Networks Advances in Intelligent Systems & Applications. Springer-Verlag Berlin Heidelberg, pp 259–267
31. Lee C, Huang S, Shih C (2010) Facial affect recognition using regularized discriminant analysis-based algorithms, EURASIP J Adv Signal Process vol 10
32. Lienhart R, Maydt J (2001) An extended set of haar-like features for rapid object detection. Int Conf Image Process 1:900–903
33. Liu S, Ruan Q, Wang C, An G (2012) Tensor rank one differential graph preserving analysis for facial expression recognition. Image Vis Comput 30:535–545
34. Liu S, Ruan Q, Wang C, An G (2012) Tensor rank one differential graph preserving analysis for facial expression recognition. Image Vis Comput 30:535–545
35. Liu D, Tsu-Shuan C, Zhang Y (2002) A constructive algorithm for feedforward neural networks with incremental training. Trans Circ Syst-Fundam Theory Appl 49:1876–1879
36. Lippmann RP (1989) Pattern classification using neural networks, IEEE Communications Magazine, vol. 27, pp 47–50 59–64
37. Littlewort G, Whitehill J, Wu T, Butko N, Ruvolo P, Movellan J, Bartlett M (2011) The motion in emotion-a cert based approach to the fera emotion challenge IEEE Int'l Conf. Automatic Face and Gesture Analysis, pp 897–902
38. Liu S, Ruan Q, Wang C, An G (2012) Tensor rank one differential graph preserving analysis for facial expression recognition. Image Vis Comput 30:535–545
39. Long F, Wu T, Movellan JR, Bartlett MS, Littlewort G (2012) Learning spatiotemporal features by using independent component analysis with application to facial expression recognition. Neurocomputing 93:126132
40. Ma L, Khorasani K (2004) Facial expression recognition using constructive feedforward neural networks. IEEE Trans Syst Man Cybern B: Cybern 34:1588–1595

41. Masmoudi S, Frikha M, Chtourou M, Hamida A, Efficient MLP (2011) constructive training algorithm using a neuron recruiting approach for isolated word recognition system. Int J Speech Technol 14:1–10

42. Meng H, Romera-Paredes B, Berthouze N (2011) Emotion recognition by two view svm 2k classifier on dynamic facial expression features IEEE Int'l Conf. Automatic Face and Gesture Analysis, pp 854–859

43. Majumder A, Behera L, Subramanian VK (2014) Emotion recognition from geometric facial features using self-organizing map. Pattern Recognit 47(3):1282–1293

44. Mohseni S, Kordy HM, Ahmadi R (2013) Facial expression recognition using DCT features and neural network based decision tree, 55th International Symposium ELMAR, pp 361–364

45. Niese R, Al-Hamadi A, Farag A, Neumann H, Michaelis B (2012) Facial expression recognition based on geometric and optical flow features in colour image sequences. IET Comput Vis 6(2):79–89

46. Owusu E, Zhan Y, Mao QR (2014) A neural-AdaBoost based facial expression recognition system. Expert Syst Appl 41(7):3383–3390

47. Pratama M, Anavatti SG, Angelov PP, Lughofer E (2014) PANFIS: A novel incremental learning machine. IEEE Trans Neural Netw Learn Syst 25(1):55–68

48. Puma-Villanueva WJ, dos Santos EP, Von Zuben FJ (2012) A constructive algorithm to synthesize arbitrarily connected feedforward neural networks. Neurocomputing 75(1):14–32

49. Sadeghi H, Raie A-A, Mohammadi M-R (2013) Facial expression recognition using geometric normalization and appearance representation, 8th Iranian Conference on Machine Vision and Image Processing (MVIP), pp 159–163

50. Sharma SK, Chandra P (2010) Constructive neural networks: A Review. Int J Eng Sci Technol 2(12):7847–7855

51. Shan C, Gong S, McOwan PW (2009) Facial expression recognition based on local binary patterns: A comprehensive study. Image Vis Comput 27:803–816

52. Srivastava R, Roy S, Yan S, Sim T (2011) Accumulated motion images for facial expression recognition in videos IEEE Int'l Conf. Automatic Face and Gesture Analysis, pp 903–908

53. Sridhar SS, Ponnavaikko M (2011) Improved adaptive learning algorithm for constructive neural networks. Int J Comput Electr Eng 3(1):1793–8163

54. https://www.kaggle.com/c/challenges-in-representation-learning-facial-expression-recognition-challenge/data

55. Tang Y (2013) Machines, Deep learning using linear support vector Workshop on Challenges in Representation Learning ICML

56. Tariq U, Lin K-H, Li Z, Zhou X, Wang Z, Le V, Huang T, Lv X, Han T (2011) Emotion recognition from an ensemble of features IEEE Int'l Conf. Automatic Face and Gesture Analysis, pp 872–877

57. Tian Y, Kanade T, Cohn J (2005) Facial expression analysis Handbook of Face Recognition. Springer. (Chapter 11)

58. Ionescu RT, Popescu M, Grozea C (2013) Local learning to improve bag of visual words model for facial expression recognition. Workshop on Challenges in Representation Learning ICML

59. Valstar MF, Jiang B, Mehu M, Pantic M, Scherer K (2011) The First Facial Expression Recognition and Analysis Challenge, 9th IEEE conference on Automatic Face and Gesture Recognition. Santa Barbara, California

60. Villegas M, Paredes R (2011) Dimensionality reduction by minimizing nearest-neighbor classification error. Pattern Recognit Lett 32:633–639

61. Viola P, Jones M (2004) Robust real-time face detection. Int J Comput Vis 57(2):137154

62. Wan S, Aggarwal JK (2014) Spontaneous facial expression recognition: A robust metric learning approach. Pattern Recognit 47(5):1859–1868

63. Wang Z, Ruan Q, An G (2012) Facial expression recognition based on tensor local linear discriminant analysis IEEE 11th International Conference on Signal Processing (ICSP), pp 1226–1229

64. Wang S, Liu Z, Wang Z, Wu G, Shen P, He S, Wang X (2013) Analyses of a multimodal spontaneous facial expression database. IEEE Trans Affect Comput 4(1):34–46

65. Luo Y, Wu C (2013) Facial expression recognition based on fusion feature of PCA and LBP with SVM. Optik - Int J Light Electron Opt 124(17):2767–2770

66. Yang P, Liu Q, Metaxas DN (2010) Exploring facial expressions with compositional features. IEEE Conf Comput Vis Pattern Recognit:2638–2644

67. Yang S, Bhanu B (2011) Facial expression recognition using emotion avatar image IEEE Int'l Conf. Automatic Face and Gesture Analysis, pp 866–871

68. Zhang Z, Qiao J (2010) A Node Pruning Algorithm for Feedforward Neural Network Based on Neural Complexity, International Conference on Intelligent Control and Information Processing Dalian, China, pp 406–410

69. Zhao X, Zhang S (2012) Facial expression recognition using local binary patterns and discriminant kernel locally linear embedding. EURASIP J Adv Signal Process 10:1687–6180

70. Zavaschi THH, Britto JrAS, Oliveira LES, Koerich AL (2013) Fusion of feature sets and classifiers for facial expression recognition. Expert Syst Appl 40(2):646–655
71. Zhen W, Zilu Y (2012) Facial Expression Recognition Based on Local Phase Quantization and Sparse Representation Eighth International Conference on Natural Computation (ICNC), pp 222–225

Hayet Boughrara was born in Ksour-Essef, in 1982. She received the BS degree in electrical engineering in 2006 and the MS degree in Automatic and Industrial Computing in 2008, both from the Sfax Engineering Scool, University of Sfax, Tunisia. She has been a contractual assistant of electrical engineering at the high Institute of Computer Science and Multimedia of Gabes, Tunisia, since 2008. Mrs Boughrara is a member of the Control and Energy Management Laboratory of the University of Sfax. Her current research interests include learning algorithms, artificial neural networks and their engineering applications, fuzzy systems, and intelligent control.

Mohamed Chtourou was born in Sfax (Tunisia) in 1963. He received the Engineering Diploma in electrical engineering from the National School of Engineers of Sfax-Tunisia in 1989, the Diplme dEtudes Aprofondies in Automatic Control from the Institut National des Sciences Appliques de Toulouse-France in 1990, and the Doctorat in Process Engineering from the Institut National Polytechnique de Toulouse-France in 1993 and the Habilitation Universitaire in Automatic Control from the National School of Engineers of Sfax -Tunisia in 2002. He is currently a professor in the Department of Electrical Engineering of National School of Engineers of Sfax- University of Sfax-Tunisia. His current research interests include learning algorithms, artificial neural networks and their engineering applications, fuzzy systems, and intelligent control. He is author and co-author of more than twenty papers in international journals and of more than 50 papers published in national and international conferences.

Chokri Ben Amar received the B.S. degree in Electrical Engineering from the National Engineering School of Sfax (ENIS) in 1989, the M.S. and PhD degrees in Computer Engineering from the National Institute of Applied Sciences in Lyon, France, in 1990 and 1994, respectively. He spent one year at the University of "Haute Savoie" (France) as a teaching assistant and researcher before joining the higher School of Sciences and Techniques of Tunis (ESSTT) as Assistant Professor in 1995. In 1999, he joined the Sfax University (USS), where he is currently a professor in the Department of Electrical Engineering of the National Engineering School of Sfax. His research interests include Computer Vision and Image and video analysis. His research interests include Computer Vision and Image and video analysis. These research activities are centered on Wavelets and Wavelet networks and their applications to data Classification and approximation, Pattern Recognition and image and video coding, indexing and watermarking. He is a senior member of IEEE, and the chair of the IEEE SPS Tunisia Chapter since 2009. He was the chair of the IEEE NGNS2011 (IEEE Third International Conference on Next Generation Networks and Services) and the Workshop on Intelligent Machines: Theories & Applications (WIMTA 2008) and the chairman of the organizing committees of the "Traitement et Analyse de lInformation : Mthodes et Applications (TAIMA 2009)" conference, International Conference on Machine Intelligence ACIDCA-ICMI'2005 and International Conference on Signals, Circuits and Systems SCS2004.

Liming Chen received the Bsc degree in Mathematics and Computer Science from University de Nantes in 1984. He obtained the Master degree in 1986 and the PhD in computer science from University of Paris 6 in 1989. First served as associate professor at Universit de Technologies de Compigne, he joined in 1998 Ecole Centrale de Lyon as Professor where he has been leading an advanced research team on multimedia analysis. From 2001 to 2003, he also served as Chief Scientific Officer in a Paris based company, Avivias, specialized in media asset management. In 2005, he served as Scientific expect multimedia in France Telecom R&D China. He has been Chairman of the department of Mathematics and Computer science from 2007.