A New Facial Expression Recognition Algorithm Based on DWT Feature Extraction and Selection

Fatima Zohra BOUKHOBZA
Department of Electrical Systems Engineering, LIST Laboratory, M'hamed Bougara University, Algeria
f.boukhoubza@univ-boumerdes.dz

Abdenour HACINE GHARBI
Department of Electronics, MSE Laboratory, Mohamed El Bachir El Ibrahimi University, Algeria
abdenour.hacinegharbi@univ-bba.dz

Khaled ROUABAH
Department of Electronics, ETA Laboratory, Mohamed BOUDIAF University of M'sila, Algeria
khaled.rouabah@univ-msila.dz

Abstract: *In this paper, we propose an efficient framework to improve accuracy and computational cost of a Facial Expression Recognition (FER) system. This framework is carried out in three stages. In the initial one, corresponding to feature extraction, three descriptors, derived from Discrete Wavelet Transform (DWT), are introduced to extract distinct feature types. In the second stage, focused on feature selection, a Wrapper approach is adopted to carefully select the most relevant features from the previously extracted pool. Following feature selection, the Support Vector Machine (SVM) classifier is employed, in the final stage, to determine an individual's affective state. The experiments were conducted in person-independent mode using both the Japanese Female Facial Expression (JAFFE) and extended Cohn-Kanade (CK+) databases which included the following emotions: anger, disgust, contempt, fear, happy, sad, surprise, and neutral. The obtained results demonstrated the effectiveness of the proposed framework in increasing recognition rate and decreasing response time compared to other state-of-the-art methods. A comparative study between our proposed framework and that based on the Local Binary Patterns (LBP) method demonstrated that our framework outperforms the latter for most emotions. In fact, our proposed framework converges rapidly and achieves good performance, thus allowing us to develop a real-time Facial Expression Recognition (FER) system in person-independent mode. Average recognition rates of 89.66% and 87.76% were obtained using our method with the JAFFE database and the CK+ database, respectively.*

Keywords: *Facial expression recognition, 2D-DWT, person-independent system, feature extraction, wrapper feature selection, dimensionality reduction.*

Received February 9, 2024; accepted June 29, 2024
https://doi.org/10.34028/iajit/21/4/6

1. Introduction

A Facial Expression Recognition (FER) system is defined as a system that uses a set of algorithms to identify and recognize a person's emotional state through her facial expressions. It can be widely applied in several research fields, such as the analysis of mental diseases and the recognition of human social/physiological interactions [35, 51]. In particular, several applications of the FER task have the objective of identifying the emotional state of a person by classifying their facial image in a targeted emotion class C_i:{anger, disgust, contempt, fear, happy, sad, surprise, neutral}; otherwise, it is classified in a class $\overline{C_i}$ (in the opposite case). For example, taking into account fear in a public place (including airports, banks, hospitals, etc. is an essential component of detecting stress and thus ensuring safety (civil security and surveillance). Likewise, the detection of an emotion of joy makes it possible to measure customer satisfaction, which helps individuals adapt to their needs (marketing), while the detection of a disgust emotion during a distance learning session can improve the quality of educational performance (education). In addition, detecting contempt or sadness helps individuals find certain psychological illnesses (medicine) [7, 21, 32].

Recognizing an emotional state is natural for humans but difficult for machines, especially when the system has no prior knowledge of the test subject (in person-independent mode). An FER system usually consists of three main processes, namely, face detection, feature extraction and classification [17]. According to the literature, feature extraction, which is the most important process [11, 49], consists of transforming the image into a set of features allowing separation of different classes. In this context, various methods have been proposed in the scientific literature. The most commonly used methods for facial expression analysis include Scale Invariant Feature Transform (SIFT) [27], Gabor filters [13], log-Gabor filters [22], Local Binary Patterns (LBP) [38] and Histogram of Oriented Gradient (HOG) [8]. All these methods have been proven to be effective for image representation. However, due to the high dimensionality of the extracted features, the direct use of these methods will increase the computation time and memory required for storage [1, 18, 37]. For example, the Gabor wavelet representation method is capable of detecting multi-scale and multi-direction texture changes [51]. Nevertheless, it generates a highdimensional Gabor coefficient matrix composed of redundant features. This latter feature redundancy is responsible for increasing confusion and consequently reducing the accuracy of the FER system [12]. Regarding the use of log-Gabor filters for feature extraction, the same limitations, in terms of complexity, can be observed. Indeed, according to the study realized in [22], the use of this type of filters in FER systems has demonstrated that, despite their high recognition accuracy, the log-Gabor filters have a major drawback which consists of complexity and calculation time. This high computational load makes them less efficient compared to other methods, especially for certain facial expressions [22]. Similarly, the LBP method, which assigns a binary code to a pixel according to its neighborhood, overcomes the problems of disequilibrium displacement, rotation and illumination in an image. In addition, it extracts enough texture information allowing tahe classifier to handle the FER problem well [36]. However, according to [1], a small neighborhood cannot capture the dominant features with large-scale structures, whereas a large neighborhood augments the size of the feature histogram and subsequently increases the computational cost of each LBP image [36]. In another example, the HOG method was proven to be an effective descriptor for preserving local information using the orientation density distribution and gradient of the edge [37]. Nevertheless, the HOG method is limited by its high computational cost and large feature dimension [18].

To remedy all these problems, in this paper, we propose a new feature extraction method that is more compact than many existing methods. Our method, which is capable of decreasing computational time and increasing recognition accuracy, consists of the use of three features extracted by means of the Discrete Wavelet Transform (DWT). The procedure used herein has been applied in other fields of one-Dimensional (1-D) signal processing, such as speech processing [19, 20], electrical appliance identification [15], and healthy and Parkinsonian subject identification [6, 46], but has never been used as a solution for two-Dimensional (2D) image processing problems, notably the FER problem.

To enhance the efficiency, in terms of the response time and learning performance of our proposed method, it is used in conjunction with the feature selection process. This process consists of selecting the feature subset that best discriminates between different emotion classes. The procedure used herein has two main purposes. The first is to retain only the most relevant features and avoid redundancies to increase the speed, in terms of response time, of the FER system. The second is to remedy the curse of dimensionality, which often has negative consequences on the behavior and performance of learning algorithms. It is worth noting that in machine learning, feature selection consists of choosing a subset of original features containing the most relevant and essential information. This operation can be performed by two different methodologies, namely, the "Filter" approaches and the "Wrapper" approaches. Filter approaches, which are independent of any machine learning algorithm, perform feature selection using intrinsic attribute properties. In contrast, "Wrapper" approaches, which use attributes to evaluate the classification rate, are based on a specific learning algorithm [14]. A comparison of these two methodologies revealed that the "Filter" approaches are much faster than the "Wrapper" ones because computational methods are not involved unlike the case of the "Wrapper" approaches which are computationally intensive. However, the "Wrapper" approaches have the advantage of being more efficient than the "Filter" approaches because they take into account the classifier hypothesis [45].

In the field of feature selection for FER, Lajevardi and Hussain [23], adopted a global optimization Genetic Algorithm (GA) process to select the optimal feature, generated using log-Gabor filters. This process reduced the number of features from 36000 features/image to 100 features/image and achieved an accuracy of 78% using the Cohn-Kanade (CK) database with the Naïve Bayesian (NB) classifier. Siddiqi *et al*. [42], developed a random forest-based "Wrapper" approach for feature selection. The performance of this method was evaluated on the basis of the extracted features using a large Gabor filter bank, in which a multi-class Support Vector Machine (SVM) classifier was used for FER. The results obtained from the CK+ database showed that this method increases the recognition rate for angry, sad and neutral expressions. Abdulrazaq *et al*. [2], used the Relief-F method, which ranks features based on their variance values to evaluate six classifiers namely Multi-Layer Perceptron (MLP), Random Forest (RF), Decision Tree (J48), SVM, K-Nearest Neighbor (KNN), and Radial Basis Function (RBF). Experimental results demonstrate that KNN performs the best with the selected features, highlighting the effectiveness of the Relief-F feature selection method in improving classification accuracy for FER system. Soyel *et al*. [43], proposed a FER feature selection methodology based on Non dominated Sorting Genetic Algorithm-II (NSGA-II). This methodology is one of the latest GA developed for resolving problems with high accuracy. The obtained results, using 3D facial expression database BU-3DFE demonstrated the effectiveness and the flexibility of NSGA-II-based methodology when compared with those reported in the literature.

Lajevardi and Hussain [24], applied the Mutual Information Feature Selection (MIFS) algorithm for FER system. In fact, after using Log-Gabor filters to extract features from whole face and facial regions images, the most informative features are classified using NB classifier to recognize six expressions: anger, fear, disgust, surprise, sadness and happiness. The experimental results using CK and JAFFE databases showed the effectiveness of the proposed feature selection method in improving the overall performance of the FER system.

In this paper, we propose the use of a "Wrapper" selection method in combination with the proposed feature extraction method, to select the best feature. We have chosen to use the "Wrapper" selection method because, in all situations, with small training sets (which is the case for our system), the process of feature selection is often most useful using this technique [41]. In the literature, several studies in the field of FER have explored the adoption of "Wrapper" approaches for feature selection. For example, Mlakar *et al*. [33], introduce an efficient Wrapper-based feature selection strategy for FER based on the use of a modified multi-objective differential evolution algorithm. Here, two selection strategies were developed and validated on commonly used evaluation databases. The obtained results demonstrated high recognition rates with significantly reduced feature sets, showing that this approach is very promising for improving the performance of FER systems and reducing computational costs in various applications.

Paharia *et al*. [39], proposed a new Wrapper-based feature selection method. The latter one, named improved MOCBGWO, consists of an enhancement of the competitive gray wolf binary optimizer. The selected features are then evaluated on CK+ and JAFFE datasets using SVM and KNN classifiers. The obtained results indicated that the proposed MOCBGWO method significantly enhances recognition accuracy while reducing the size of the feature vector, thus highlighting its effectiveness in FER systems.

Al-Qablan *et al*. [5] proposed a hybridized Binary GWO (BGWO) and adaptive β-hill climbing (AβCH) for the improvement Wrapper-based feature selection approach.

In this approach, the KNN classifier is used to evaluate the performance of the proposed hybrid approach. The obtained results, based on 18 standard feature selection UCI benchmark datasets showed this approach outperforms, in terms of accuracy and selected feature size, various state-of-the-art feature selection methods, thereby demonstrating promising performance for various machine learning and data analysis applications.

Perez-Gomez *et al*. [40] studied the selection of optimal geometric features for classifying facial expressions. Indeed, by exploiting techniques inspired by the Facial Action Coding System (FACS) and the MPEG-4, a first set of 89 features, derived from the geometry of 2D and 3D facial landmarks, was proposed. Here, two feature selection methods, namely PCA and GA which works as a Wrapper feature selection method (using a support vector machine), are employed to reduce the feature set. The results obtained, on the Bosphorus and UIVBFED datasets, demonstrated that this method has a median accuracy of 86.62% and 93.92%, respectively, which highlights the importance of feature selection to obtain a high accuracy with low computational cost.

The remainder of this paper is organized as follows. In section 2, we detail our proposed FER framework. In section 3, we present and discuss the results. Finally, in section 4, we conclude the paper.

2. Proposed Framework

An overview of the proposed framework, which consists of a set of processes and tasks, is shown in Figure 1.

In what follows, we will provide the principle of our proposed FER system. We will therefore explain in detail the principle of each of these tasks.

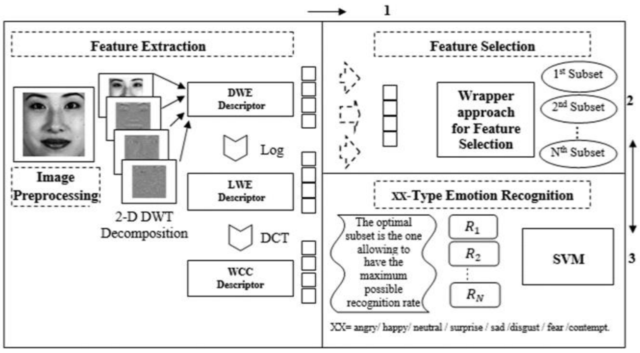

Figure 1. Flowchart of the proposed framework.

2.1. Image Preprocessing

The purpose of the preprocessing task is to improve the quality of an image and to prepare it for further analysis. This step is important for reducing noise and illumination variations. Several algorithms can be used to realize this task. Some of them are numerated in what follows:

- The use of sharpening filter to enhance the edges and fine details of an image.
- Normalization which has the effect of eliminating variations in brightness and contrast of the image.
- Adaptive histogram equalization that helps better handle images with different levels of contrast and brightness in different regions.

In our work, we first convert the original image to grayscale to deal with bad acquisition quality. This operation is often used in image preprocessing for several reasons, including:

- Eliminating color variations can simplify analysis and make relevant features more visible.
- In a grayscale image, each pixel represents only the luminance or light intensity at that location, allowing information about the structure and details of the image to be retained without being affected by color variations.

Then, as a second step, we crop the original image according to the position of both eyes and normalize it to 200×200 pixels. In the literature, there are several methods that can be applied for face detection, such as for example the Viola-Jones method and others based on pre-trained detection algorithms. In our work, as the objective was specially devoted to the development of extraction and selection methods of relevant features, we opted for manual retrieval from noiseless images.

2.2. Feature Extraction

A block diagram of our proposed feature extraction method is shown in Figure 2.

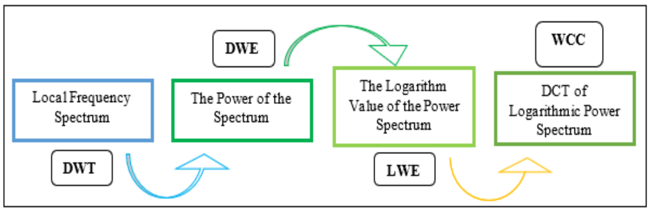

Figure 2. Block diagram of the proposed feature extraction method.

As illustrated in this figure, this method includes four phases. The first one consists of performing a 2D-DWT decomposition of the facial image, using a specific mother wavelet. This last is a windowed function that moves/shifts along the time series signal [19]. Wavelet families vary in terms of several important properties. Examples include Daubechies, Biorthogonal, Coiflets, Symlets, and Morlet. In this work, three wavelet functions, namely Daubechies, Coiflet and Symlet, were used to carry out this feature extraction method.

As shown in Figures (3), (4), and (5), the differences between these wavelet families mainly lie in their properties of regularity, symmetry, and asymmetry, which make them suitable for different types of applications in signal processing and data analysis [31]. In reality, Daubechies wavelets have moments of high regularity, which means that they are well localized in both time and frequency. Coiflet wavelets in turn were developed to be an alternative to Daubechies wavelets with some asymmetry and improved regularity. Symlet wavelets are a symmetric extension of Coiflet wavelets.

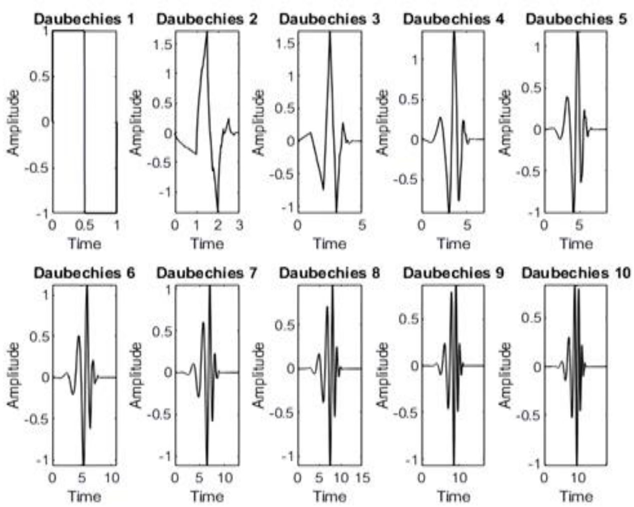

Figure 3. Daubechies wavelets family.

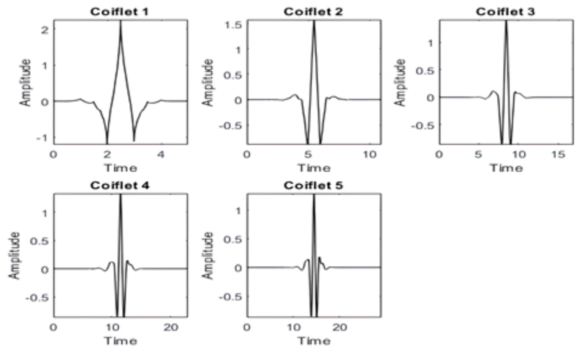

Figure 4. Coiflet wavelets family.

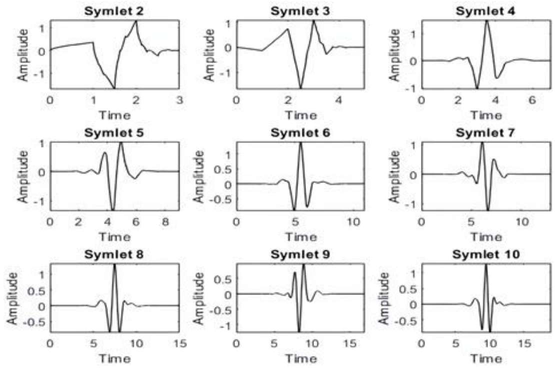

Figure 5. Symlet wavelets family.

Recall that the DWT splits information into its low and high frequency components. Indeed, the low frequency component is a lowpass approximation of the original image (smooth variations) that contains most energy of the image. However, the high frequency component represents edge components (which give the detail) and has three separate unidirectional directions namely horizontal, vertical and diagonal. In addition, DWT can be calculated at diverse resolutions or scales [9, 30].

In 2D-DWT configuration, we have to use one scaling function $\varphi(x,y)$ and three wavelets $\psi^H(x,y)$, $\psi^V(x,y)$ and $\psi^D(x,y)$. Each scaling function or wavelet is the product of a one-dimensional scaling function φ and corresponding wavelet ψ.

The four 2D products produce the scaling function given by Equation (1) and separable directional sensitive wavelets given by Equations (2), (3), and (4) [47].

$$\varphi(x,y) = \varphi(x)\varphi(y) \tag{1}$$

$$\psi^H(x,y) = \varphi(y)\psi(x) \text{ (horizontal)} \tag{2}$$

$$\psi^V(x,y) = \varphi(x)\psi(y) \text{ (vertical)} \tag{3}$$

$$\psi^D(x,y) = \psi(x)\psi(y) \text{ (diagonal)} \tag{4}$$

With:

$\psi^H(x,y)$ measures gray-level variations along the columns (horizontal edges); $\psi^V(x,y)$ measures gray-level variations along the rows (vertical edges); $\psi^D(x,y)$ corresponds to the variations along diagonals.

It is worth noting that the wavelet function acts as a bandpass filter whose bandwidth is reduced to half after each scaling. In addition, for a 2D image, the rows and columns are treated as 1D signals. The block diagram in Figure 6 describes 2D wavelet decomposition on an input image. According to this diagram, during the first step of level 1 decomposition, each row of the image is scaled and wavelets transformed (horizontal transform). The results from this step are two half images, one with scaling coefficients (L) and the other with wavelet coefficients (H). In the next step, the scaling (L) and wavelet transform (H) are applied on each column of the two half images from the previous step (vertical transform). The results from this step are four-quarter sized images, which are: approximation coefficients (LL), vertical details (LH), horizontal details (HL) and diagonal details (HH). At the next level, when the wavelet transform is applied, the above steps will be repeated on the LL image [41]. The scaled and translated basis functions are defined by Equations (5) and (6) given as:

$$\Phi_{j,m,n}(x,y) = 2j/2\varphi(2^j x - m, 2^j y - n) \tag{5}$$

$$\psi^i{}_{j,m,n}(x,y) = 2j/2\psi^i(2^j x - m, 2^j y - n) \tag{6}$$

Where: index *i={H,V,D}* identifies the directional wavelets in terms of values of Horizontal, Vertical, and Diagonal (H, V, D).

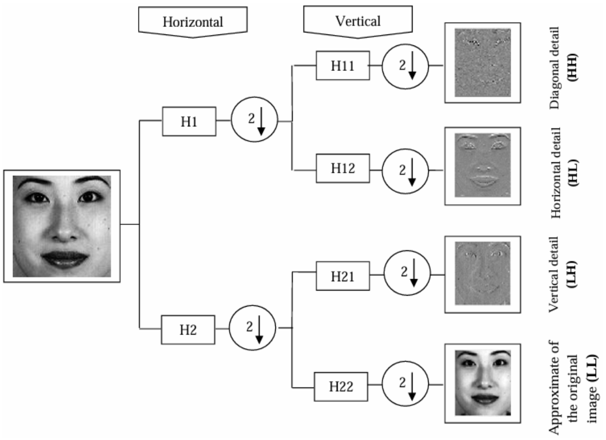

Figure 6. One level 2D-DWT decomposition on an input image.

The approximation and detail coefficients, corresponding to the function *f(x,y)* of size *M* by *N* are formulated as follows:

- Approximation coefficients:

$$W_\varphi(j_0, m, n) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y)\varphi_{j_0,m,n}(x,y) \tag{7}$$

- Detail coefficients**:**

$$W_\psi^i(j, m, n) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y)\psi^i{}_{j,m,n}(x,y) \tag{8}$$

Where:

j_0 is the starting value of scale j ($j \geq j_0$);
$j = 0,1,2,\dots,J-1$;
2^J is the size of the selected data;
$n = m = 0,1,2,\dots,2^j - 1$;

The coefficients given by Equations (7) and (8) are then used to propose, through the following three phases of our feature extraction method, three types of descriptors. The first one, called Discrete Wavelet Energy (DWE), consists of converting each facial image into a feature vector. The DWE is composed of the energy of the wavelet coefficients computed at each decomposition level from an image. The second one, called the Log Wavelet Energy (LWE), consists of applying the logarithm to the DWE features. The last descriptor, called the Wavelet Cepstral Coefficient (WCC), consists of applying the Discrete Cosine Transform (DCT) to the LWE feature vector.

As we know, in signal processing, the output signal of a time-invariant linear system is the result of the convolution between the input signal and the system impulse response. Convolution in time domain means normal product in frequency domain and vice versa. In the cepstral analysis method, the logarithm of the spectrum is used to convert multiplication in the frequency domain into addition (sum) thereby simplifying the separation of the input signal spectrum and the frequency response of the system. Then, by applying the cosine transform, we can separate the input signal from the impulse response of the system. This transform, which is only the real cepster representing the reconversion of the log-spectrum in the time domain, has a decorrelation effect between the characteristics of the logarithm of the energy spectrum. Therefore, in our proposed framework, we applied the cepstral method to the filter bank of dyadic wavelet decomposition in order to obtain the features for classifying the emotional state of a person.

The three proposed descriptors can be given as follows:

- DWE descriptor:

The first descriptor consists of converting the approximation and detail coefficients, extracted at each level *j* of DWT decomposition, into energy coefficients. Here, $E_\varphi(j_0)$, which is the energy of the approximation coefficients at scale j_0, is given as:

$$E_\varphi(j_0) = \sum_m \sum_n \left|W_\varphi(j_0, m, n)\right|^2 \tag{9}$$

Similarly, the energy of the detail coefficients, at scale *j* and orientation *i*, is given as:

$$E_\psi^i(j) = \sum_m \sum_n \left|W_\psi^i(j, m, n)\right|^2 \tag{10}$$

- LWE descriptor:

The second descriptor consists of calculating the log of the energy coefficients $E_\varphi(j_0)$ and $E_\psi^i(j)$ at each level of wavelet decomposition. The resulting coefficients are given as follows:

$$LWE(W_\varphi(j_0, m, n)_K) = log[E_\varphi(j_0)] \tag{11}$$

$$LWE\left(W_{\psi}^{i}(j,m,n)_{L}\right) = log\left[E_{\psi}^{i}(j)\right] \quad (12)$$

- WCC descriptor:

The last descriptor is obtained by applying the DCT to the logarithm of the resulting energy values. This transformation operates by decomposing a finite sequence of data points into a sum of cosine functions oscillating at different frequencies [4]. In our work, the DCT is applied to obtain coefficients that are weakly correlated. Hence, the WCC approximation coefficients are given as [48]:

$$WCC_s\left(W_{\varphi}(j_0,m,n)\right) = \sqrt{\frac{2}{K}}\sum_{q=1}^{K}\left(log[E_{\varphi}(j_0)]\right)_q cos\left(\frac{\pi s}{K}(q-0.5)\right) \quad (13)$$

The WCC detail coefficients are given as:

$$WCC_s\left(W_{\psi}^{i}(j,m,n)\right) = \sqrt{\frac{2}{K}}\sum_{q=1}^{K}\left(log[E_{\psi}^{i}(j)]\right)_q cos\left(\frac{\pi s}{K}(q-0.5)\right) \quad (14)$$

Where:
K is the number of filter bank channels.

$$s = 1, \dots, L \; (L = 3j+1) \quad (15)$$

In addition to the features resulting from the use of these three descriptors, we have also integrated, in our study, an additional energy feature Et given by Equation (16). This feature that is also subject to the application of the different aforementioned transformations (*log* and *DCT*).

$$Et = \sum(a(n), h(n), v(n), d(n), h(n-1), v(n-1), d(n-1), \dots, h(1), v(1), d(1))^2 \quad (16)$$

After extracting suitable features, we can perform feature selection to improves the system performance.

2.3. Feature Selection

To find a compromise between improving performance and reducing the size of the feature subsets, we adopted a feature selection algorithm whose principle is shown in Figure 7. This algorithm makes it possible to identify the most important features in a dataset and to remove irrelevant information.

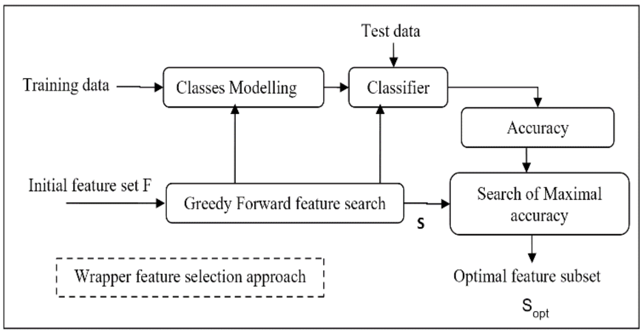

Figure 7. The general concept of the feature selection algorithm adopted in our proposed framework.

As illustrated in this figure, the adopted feature selection algorithm is based on the Wrapper approach. The latter starts from an empty set of features, and in each iteration, the feature is added (using greedy forward search), whose addition improves performance. In other words, the different feature subsets are generated successively by adding the features one by one. The classifier, illustrated in the same figure, is subsequently used to evaluate the relevance of the different generated subsets. The process stops when there are no additional variables to add. The optimal subset is then the one that gives the best recognition precision. In what follows, we will give the principle of the application of this algorithm for our proposed framework.

Let F be the starting set and S be the subset of relevant features. Initially, the subset of relevant features S is empty, while the starting set F contains all the features. The selection algorithm proceeds to determine the recognition rate by considering each of the features f_i of the starting set, and after having determined the different recognition rates corresponding to the different features, the algorithm removes the feature f_{s1} (from the starting set F), making it possible to have the maximum recognition of the starting set F and adds it to the subset of relevant features S. The algorithm repeats the same procedure on the remaining features in the starting set F. Indeed, by comparing the recognition rates obtained by using the features of subset S already selected and combined each time with one of the remaining features in the starting set F, the algorithm removes (from the set F), at each iteration, the feature that allows us to give the maximum recognition rate of the set F and adds it to the set S (until it completes the test on all the features' departure). Several candidate subsets with different numbers of features are therefore formed. Finally, the optimal subset S_{opt} represents the subset allowing the best recognition rate among those obtained by the different tested subsets (candidates) [15, 16, 20].

The proposed algorithm can be summarized as follows:

Algorithm 1: Proposed feature selection algorithm.

Step 1. Initialize the starting set F with all features and the subset of relevant features S with no features (empty subset).
Step 2. Calculate the recognition rate using each feature $f_i \in F$.
Step 3. Find the first feature f_{s1} *that maximizes the recognition rate.*
Step 4. Remove the feature f_{s1} *of the set F and add it to the subset of relevant features S.*
Step 5. Repeat the process (from Step 2 to Step 4) until it completes testing all features. This is given as follows:

1) $\forall f_i \in F$, *determine the recognition rate using* $\{S, f_i\}$.
2) *Choose the feature* $f_i \in F$ *that maximizes the recognition rate using* $\{S, f_i\}$ *at step j;*
3) *Affect* $F \leftarrow F - \{f_{sj}\}$; $S \leftarrow S \cup \{f_{sj}\}$.

Step 6. Remove the subset S of the selected features that maximizes the recognition rate.

2.4. Emotional Recognition

Recall that the objective of our FER system is to recognize a person's emotional state by verifying whether their facial image corresponds to a targeted emotional class. The validation of the proposed method is carried out by considering the different types of emotions (angry, happy, contempt, neutral, surprise, sad, disgust, and fear) using a two-class SVM. As illustrated in Figure 8, the SVM works to find an optimal hyperplane to linearly separate two classes.

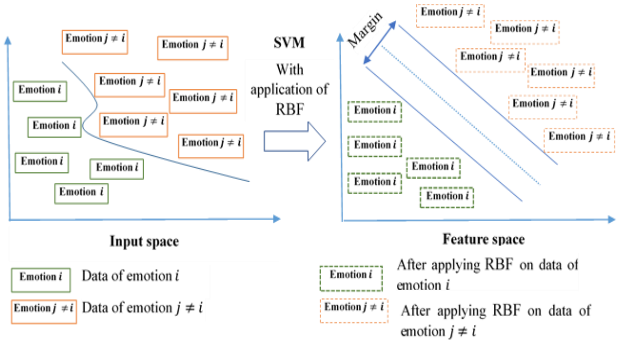

Figure 8. Representation of the concept of the RBF kernel in SVM.

When the data are non-linearly separable, different kernels such as polynomial, sigmoid and RBF kernels, can be used to construct a linear decision function in the feature space so that the database becomes separable with the maximum margin [34]. In our experiments, the RBF kernel is used. The decision function of the SVM using the RBF kernel is given as:

$$f(x) = sgn\left(\sum_{i=1}^{N} \alpha_i\, y_i\, K(x, x_i) + b\right) \quad (17)$$

where:

$sgn(\cdot)$: is the signum function;

x: the input data;

α_i: are the Lagrange multipliers obtained during the training phase;

y_i : are the class labels of the training data points;

b: is the bias term;

$K(x, x_i)$: is the RBF kernel function, defined as:

$$K(x, x_i) = \exp(-\gamma\, \|x - x_i\|^2) \quad (18)$$

With:

γ: is a hyperparameter that controls the kernel's width; $\|.\|$: is the Euclidean distance.

$f(x)$ predicts the class label of the input data point x.

During training, the SVM learns the optimal values of α_i and b by solving the optimization problem to maximize the margin between support vectors. The decision function then uses these parameters to classify the new data points based on their proximity to the decision boundary defined by the hyperplane [26].

3. Experimental Results and Discussion

Experiments are conducted to test the performance of the proposed framework. For this reason, several situations were considered.

3.1. Preparation of the Data

To evaluate our proposed framework, two databases were used. These are given as follows:

- JAFFE database [29]: the JAFFE database consists of a set of facial images, each with a resolution of 256×256 pixels. The images represent seven facial expressions (happy, sad, anger, fear, surprise, disgust, neutral) from 10 different Japanese female subjects, which simulated 3-4 examples for each emotion.
- Extended Cohn–Kanade (CK+) database [28]: this database is among the most widely used databases for emotion detection. It includes a total of 593 image sequences from 123 different subjects. All participants were between 18 and 50 years of age; 69% were female, 81% were European American, 13% were Afro-American, and 6% were in other groups. Typically, the last frame of the sequences is used for image-based facial expression recognition. In our study, the CK+48 version was used. The latter consists of a set of 981 facial images each with a resolution of 48×48 pixels. The images represent seven facial expressions, namely, anger, contempt, disgust, fear, happy, sad and surprise.

Table 1 summarizes the emotion distribution of each database.

Table 1. Number of images (per emotion) used in our experiments from both the CK+ and JAFFE databases.

Database	Anger	Contempt	Fear	Disgust	Sad	Happy	Surprise	Neutral	Σ
CK+	135	54	75	177	84	207	249	/	981
JAFFE	30	/	32	29	30	31	31	30	213

In our experiments (for both databases), the number of images, for each emotion, was divided into two subsets: 50% for the training phase and 50% for the tests. Note here that the choice to split the data equally between training and testing has several advantages and justifications. In fact, this distribution ensures a balanced evaluation of the model's performance. This means that the model is evaluated on a substantial amount of data that was not used during training, thereby allowing a more reliable estimate of its ability to generalize to new data. On the other hand, allocating an equal portion of data to testing helps avoid overfitting, where the model learns too specifically from the training data and does not generalize well to new data.

3.2. Experiments

As previously explained, for feature extraction, the DWT was applied to each image to decompose it into its respective approximation and detail coefficients. Practically, two criteria are necessary to ensure a better decomposition. These are focused on the optimal mother wavelet function and the level of wavelet decomposition. As we discussed previously, in our experiments, three different families of wavelets were used to perform DWT analysis on each image. The latter are given as follows:

- The Daubechies family with an order that varies from 1 to 8 (Db1, Db2, ..., Db8).
- The Coiflet family with an order that varies from 1 to 5 (Coif1, Coif2..., Coif5).
- The Symlet family has an order that varies from 1 to 8 (Sym1, Sym2, ..., Sym8).

The wavelet decomposition of an image at level n, returns a vector C that contains the approximation and detail coefficients organized by level ($a(n)$, $h(n)$, $v(n)$, $d(n)$, h (n-1), v (n-1), $d(n$-1),..., $h(1)$,$v(1)$, $d(1)$). Here, a, h, v, and d, are vectors containing the approximation coefficients, the horizontal detail coefficients, the vertical detail coefficients and the diagonal detail coefficients, respectively.

After 2D wavelet decomposition, the energy percentages, corresponding to the approximation Ea, the horizontal details Eh, the vertical details Ev and the diagonal details Ed, are determined at each decomposition level. The resulting feature set, combined with energy feature Et, forms the first DWE descriptor given as follows:

$$X1 = [Ea(n), Eh(n), Ev(n), Ed(n), Eh(n-1), Ev(n-1), Ed(n-1), \ldots, Eh(1), Ev(1), Ed(1), Et] \quad (19)$$

Then, the second LWE descriptor is obtained by calculating the logarithm of the values belonging to the $X1$ vector. This is given as:

$$X2 = log(X1) \quad (20)$$

Finally, the WCC descriptor is obtained by applying DCT to the feature belonging to the second descriptor:

$$X3 = DCT(X2) \quad (21)$$

After the feature extraction phase, a feature selection algorithm is used to test the resulting feature subsets and thereby select the most relevant one. The choice of the optimal subset is made by the classifier, which evaluates all the candidate subsets to choose the one ensuring better recognition for each emotion.

3.3. Results and Discussion

In the following, we present the results and discussions of the proposed FER framework. Firstly, we will give the effect of the mother wavelet and decomposition of the associated levels on the performances.

According to the different orders of the mother wavelet and the associated maximum decomposition level, the SVM classifier is trained and tested using the selected features. The performance results, for each emotion, were evaluated by computing the recognition rate R given as follows [50]:

$$R = \frac{Number\ of correctly\ recognised\ images}{Number\ of\ test\ images} \times 100 \quad (22)$$

The best results, in terms of the recognition rate of the different emotions, corresponding to each of the three descriptors (DWE, LWE and WCC), are presented in Table 2. These results were obtained after testing the different possible cases according to the three types of features that result from using the three descriptors.

The analysis of the obtained results shows that the proposed framework achieves a good recognition rate exceeding 81% for all types of expression. In fact, our proposed framework, for features extracted from the DWE descriptor, yields a higher recognition rate for expressions of surprise and fear. For angry and neutral expressions, this rate is greater when using the features extracted from the LWE descriptor. Finally, for the emotions of contempt, disgust, happiness and sadness, the same rate is greater when using the features extracted from the WCC descriptor.

Table 2. Performance results obtained for each emotion.

Emotion	Best descriptor	Optimal mother wavelet	Decomposition level	Recognition rate (JAFEE)	Recognition rate (CK+)
Anger	LWE	Sym1	7	85.71%	86.34%
Contempt	WCC	Db8	3	/	94.41%
Fear	DWE	Db2	6	90.48%	90.06%
Disgust	WCC	Db1	7	91.43%	81.37%
Sad	WCC	Sym7	3	87.62%	90.68%
Happy	WCC	Sym2	6	90.47%	85.51%
Surprise	DWE	Sym6	4	96.19%	85.92%
Neutral	LWE	Db6	4	85.71%	/

3.4. Performance Evaluation

The results presented in Tables 3 and 4 show the contributions of the proposed features selection approach to improving the overall performance of the recognition framework. Indeed, as shown by these two tables, the proposed selection approach significantly increases the recognition rate by using only a reduced number of relevant features. This is because the selection method solves the problem of the dimensionality curse. For example, for the case of the emotion of anger, the use of the proposed selection approach made it possible to exploit only 4.35% (a reduction of 95.65%) of the initial total number of features (used before the selection operation) to ensure the recognition task with an increase of 17.14% in the recognition rate.

Table 3. Recognition rate and number of features obtained for each emotion for the JAFFE database before and after adoption of the selection approach.

Emotion	Recognition rate			Number of features		Rate of reduction in the number of features
	Before Selection (R_b)	After Selection (R_a)	Improvement Rate ($R_a - R_b$)	Before Selection (N_{fb})	After Selection (N_{fa})	$(1-(N_{fa}/N_{fb})) \times 100$
Anger	68.57%	85.71%	17.14%	23	1	95.65%
Contempt	/	/	/	/	/	/
Fear	74.29%	90.48%	16.19%	20	9	55%
Disgust	81.90%	91.43%	9.53%	23	8	65.22%
Sad	72.38%	87.62%	15.24%	11	1	90.91%
Happy	84.76%	90.47%	5.71%	20	7	65%
Surprise	89.52%	96.19%	6.67%	14	9	35.71%
Neutral	45.71%	85.71%	40%	14	1	92.86%

Table 4. Recognition rate and number of features obtained for each emotion in the CK+ database before and after adoption of the selection approach.

Emotion	Recognition rate			Number of features		Rate of reduction in the number of features
	Before Selection (R_b)	After Selection (R_a)	Improvement Rate ($R_a - R_b$)	Before Selection (N_{fb})	After Selection (N_{fa})	$(1-(N_{fa}/N_{fb}))\times 100$
Anger	75.57%	86.34%	10.77%	23	1	95.65%
Contempt	82.61%	94.41%	11.8%	11	1	90.91%
Fear	87.99%	90.06%	2.07%	20	17	15%
Disgust	73.08%	81.37%	8.29%	23	1	95.65%
Sad	70.39%	90.68%	20.29%	11	1	90.91%
Happy	82.19%	85.51%	3.32%	20	12	40%
Surprise	75.98%	85.92%	9.94%	14	5	64.29%
Neutral	/	/	/	/	/	/

Generally, for all other emotions, the proposed feature selection approach reduces the number of features and increases the recognition rate, leading to improved performance in terms of response time. In fact, the increase in the recognition rate is between 5.71% and 40 % for the JAFFE database and between 2.07% and 20.29% for the CK+ database. Likewise, the reduction rate in the number of features was between 35.71% and 95.65% for the JAFEE database and between 15% and 95.65% for the CK+ database.

To better explain this phenomenon, we take the example of surprise emotion recognition using the JAFFE database. The resulting variation in the recognition rate as a function of the number of features is shown in Figure 9. As illustrated in this figure, if the classifier uses the entire set of extracted features, then the maximum level of decomposition, corresponding Sym6 wavelet, is equal to 4. This process yielded a total of 14 features (1 approximation coefficient, 4 horizontal detail coefficients, 4 vertical detail coefficients, 4 diagonal detail coefficients plus 1 additional energy coefficient). In this case, the recognition rate reaches a value of 89.52%.

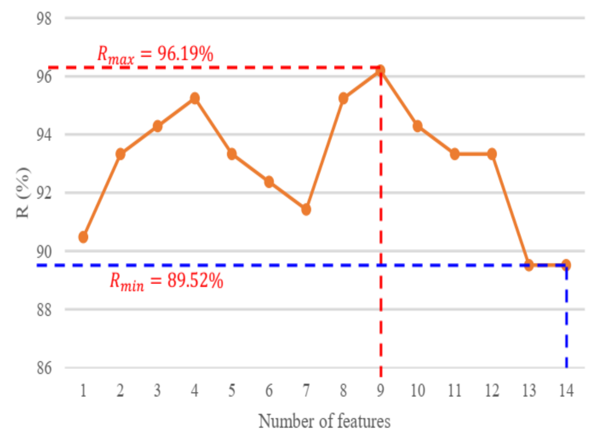

Figure 9. Variation in the recognition rate (surprise emotion) as a function of the number of features.

If the classifier now evaluates the relevance of each of the features and selects only those that increase the performance, the recognition rate will reach 96.19% (an improvement of 6.67%) using only 9 features (a reduction rate in the number of features equal to 35.71%).

3.5. Performance Comparison

To demonstrate the effectiveness of our proposed FER framework, it is compared to that of the LBP method. The latter one has proven to be a simple and effective method for facial expression representation [3]. It consists of a descriptor that can be used in image processing and computer vision for representing the texture of an image [38]. The LBP operator takes a local neighborhood around each pixel, limits the pixels in that neighborhood to the value of the central pixel, and uses the resulting binary valued image patch as a local image descriptor [1].

The LBP descriptor for every pixel in an image is given as [38]:

$$LBP(P,R) = \sum_{p=0}^{P-1} s(g_p - g_c)\, 2^p \tag{23}$$

Where: $s(x)$ represents the signum function corresponding to the LBP. g_p and g_c are the gray levels of a neighboring pixel and of the central pixel respectively.

The resulting binary code is then converted into a decimal number, which represents the texture pattern of that pixel. By applying LBP to each pixel in an image, a histogram of these patterns can be created, providing a compact representation of the image's texture.

In this work, the comparison of the performance of our proposed framework with that of LBP one is carried out under the same conditions using the same data. The results, in terms of facial expression recognition rate, are presented in Figures 10 and 11 for both the JAFFE and CK+ databases, respectively. As shown in Figure 10, for the JAFFE database, our framework outperforms the LBP-based framework for the majority of emotions (fear, disgust, happy, surprise, sad and neutral).

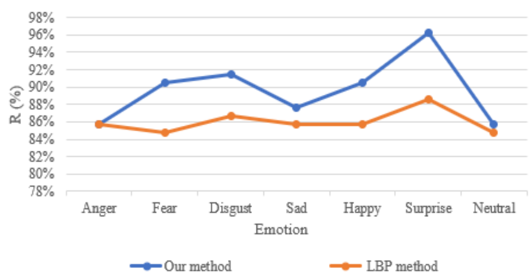

Figure 10. Comparison between the facial expression recognition rates (in %) obtained by our proposed framework and those obtained by the LBP-based framework in the JAFFE database.

However, the recognition rate is similar to that obtained by the LBP-based framework for the emotion of anger. For the CK+ database, as shown in Figure 11, the proposed framework, compared to the LBP-based framework, achieves the best performance for the emotions of anger and contempt. For other emotions, the LBP-based framework slightly outperforms our framework. Despite these last findings regarding the recognition rate of our framework, it remains more efficient in terms of the response time than the LBP-based framework, as we will see in the next section.

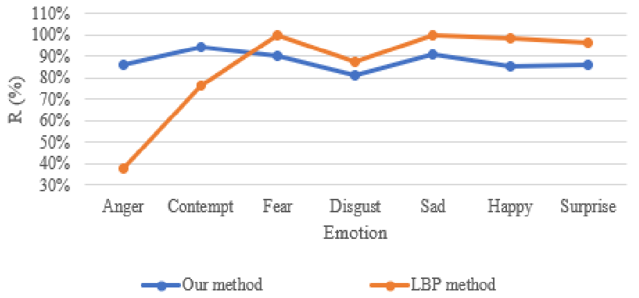

Figure 11. Comparison between the facial expression recognition rates (in %) obtained by our proposed framework and those obtained by the LBP-based framework in the CK+ database.

3.6. Response Time Comparison

To compare the response time of our proposed framework with that of the LBP-based framework, a study was conducted to determine the number of features required to train each framework. The results are given as follows:

- For the LBP-based framework: as shown in Table 5, each original image is transformed into an LBP image by assigning a value among $2^p=2^8=256$ others to each of its pixels. Next, the resulting image is converted into a feature vector by calculating a histogram showing the frequency of occurrence corresponding to each LBP value. The final vector is therefore of length 256.
- For our proposed framework: as shown in the same table, the number of features does not exceed 9 in the JAFFE database (for fear and surprise) or 17 features in the CK+ database (for fear).

Therefore, compared with that of our feature vector, the size of the resulting feature vector based on the LBP framework is large, which proves that our proposed framework outperforms the LBP-based one in terms of computational complexity. To confirm this result, simulations using the MATLAB tool are conducted to test the computation time consumed by the two frameworks mentioned above. The results, provided and illustrated in Table 5, demonstrate the response time efficiency of our proposed method compared to that of the LBP method. In fact, according to this table, our proposed framework has a very low calculation time for the two databases (25% for JAFEE and 10.63% for CK+).

Table 5. Simulation results of the computing time criterion for both the JAFFE and CK+ databases.

Database	Method	Number of required features	Consumed time in %
JAFFE database	Our method	9	25%
	LBP method	256	100%
CK+ database	Our method	17	10.63%
	LBP method	256	100%

4. Conclusions

In this paper, an improved framework, based on the combination of a new feature extraction method and a feature selection scheme, is proposed. This framework, which is realized in several steps, can be used to achieve more efficient recognition of individual facial expressions. In its first step, which consists of feature extraction, three descriptors (based on DWT) are proposed for extracting three types of features. Concerning its second step in characterizing feature selection, a Wrapper approach has been used to choose only the most relevant features among those already extracted. A complete study was carried out according to the different orders of DWTs of three mothers and the maximum level of decomposition, allowing us to find the best decomposition method enabling a better FER. Here, the SVM classifier was used in the last step to determine the recognition rates in each study case. The simulation results, based on the JAFFE database show that the proposed framework yields the best overall performance in terms of the recognition rate compared to the LBP-based framework. For the CK+ database, our proposed method shows better performance for some emotions (with maximum recognition rate degradation of LBP-based system reaching up to 48.25%.) and a fairly similar performance for the rest of the emotions (with maximum recognition rate degradation of our proposed framework reaching only 12.58%.) compared to the LBP framework. For both databases, our proposed FER framework yields a shorter response time than do the other framework. This is because our proposed framework has low complexity in the recognition process, which classifies it among the frameworks that allow a trade-off between complexity and efficiency. In future work, we consider combining the proposed framework with image denoising and automatic face detection algorithms.

References

[1] Abdulrahman M., Gwadabe T., Abdu F., and Eleyan A., "Gabor Wavelet Transform Based Facial Expression Recognition Using PCA and LBP," *in Proceedings of the 22nd Signal Processing and Communications Applications Conference*, Trabzon, pp. 2265-2268, 2014. doi:10.1109/SIU.2014.6830717

[2] Abdulrazaq M., Mahmood M., Zeebaree S., Abdulwahab M., and Zebari, R., "An Analytical Appraisal for Supervised Classifiers' Performance on Facial Expression Recognition Based on Relief-F Feature Selection," *in Proceedings of the International Conference of Modern Applications on Information and Communication Technology*, Babylon-Hilla, pp. 012055, 2021. DOI:10.1088/1742-6596/1804/1/012055

[3] Ahmed F., Bari H., and Hossain E., "Person-

Independent Facial Expression Recognition Based on Compound Local Binary Pattern (CLBP)," *The International Arab Journal of Information Technology*, vol. 11, no. 2, pp. 195-203, 2014.

[4] Ahmed N., Natarajan T., and Rao K. "Discrete Cosine Transform," *IEEE Transactions on Computers*, vol. 100, no. 1, pp. 90-93, 1974, doi:10.1109/T-C.1974.223784

[5] Al-Qablan T., Noor M., Al-Betar M., and Khader A., "Improved Binary Gray Wolf Optimizer Based on Adaptive B-Hill Climbing for Feature Selection," *IEEE Access*, vol. 11, pp. 59866-59881, 2023. doi:10.1109/ACCESS.2023.3285815

[6] Bengacemi H., Hacine-Gharbi A., Ravier P., Abed-Meraim K., and Buttelli O., "Surface EMG Signal Classification for Parkinson's Disease Using WCC Descriptor and ANN Classifier," *in Proceedings of the 10th International Conference on Pattern Recognition Applications and Methods*, Porto, pp. 287-294, 2021. DOI:10.5220/0010254402870294

[7] Clavel C. and Ehrette T., "Fear-Type Emotion Recognition and Abnormal Events Detection for an Audio-Based Surveillance System," *WIT Transactions on Information and Communication Technologies*, vol. 39, pp. 471-479, 2008. DOI:10.2495/RISK080461

[8] Dalal N. and Triggs B., "Histograms of Oriented Gradients for Human Detection," *in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, San Diego, pp. 886-893, 2005. doi:10.1109/CVPR.2005.177

[9] Daren H., Jiufen L., Jiwu H., and Hongmei L., "A DWT-Based Image Watermarking Algorithm," *in Proceedings of the IEEE International Conference on Multimedia and Expo*, Tokyo, pp. 313-316, 2001. doi:10.1109/ICME.2001.1237719

[10] Didiot E., Illina I., Fohr D., and Mella O., "A Wavelet-based Parameterization for Speech/Music Discrimination," *Computer Speech and Language*, vol. 24, no. 2, pp. 341-357, 2010. https://doi.org/10.1016/j.csl.2009.05.003

[11] Dino H., Abdulrazzaq M., Zeebaree S., Sallow A., and Zebari R., "Facial Expression Recognition based on Hybrid Feature Extraction Techniques with Different Classifiers," *TEST Engineering and Management*, vol. 83, pp. 22319-22329, 2020.

[12] Dosodia P., Poonia A., Gupta S., and Agrwal S., "New Gabor-DCT Feature Extraction Technique for Facial Expression Recognition," *in Proceedings of the 5th International Conference on Communication Systems and Network Technologies*, Gwalior, pp. 546-549, 2015. doi:10.1109/CSNT.2015.162

[13] Gabor D., "Theory of Communication. Part 1: The Analysis of Information," *Journal of the Institution of Electrical Engineers-part III: Radio and Communication Engineering*, vol. 93, no. 26, pp. 429-441, 1946. DOI:10.1049/ji-3-2.1946.0074

[14] Ghosh M., Kundu T., Ghosh D., and Sarkar R., "Feature Selection for Facial Emotion Recognition Using Late Hill-Climbing Based Memetic Algorithm," *Multimedia Tools and Applications*, vol. 78, pp. 25753-25779, 2019. https://doi.org/10.1007/s11042-019-07811-x

[15] Hacine-Gharbi A., Ravier P., "Wavelet Cepstral Coefficients for Electrical Appliances Identification using Hidden Markov Models," *in Proceedings of the 7th International Conference on Pattern Recognition Applications and Methods*, Madeira, pp. 541-549, 2018. DOI:10.5220/0006662305410549

[16] Hacine-Gharbi A., Petit M., Ravier P., and Némo F., "Prosody based Automatic Classification of the Uses of French 'Oui' as Convinced or Unconvinced Uses," *in Proceedings of the International Conference on Pattern Recognition Applications and Methods*, Lisbon, pp.349-354, 2015. DOI:10.5220/0005293103490354

[17] Huang Y., Yan Y., Chen S., and Wang H., "Expression-Targeted Feature Learning for Effective Facial Expression Recognition," *Journal of Visual Communication and Image Representation*, vol. 55, pp. 677-687, 2018. https://doi.org/10.1016/j.jvcir.2018.08.002

[18] Kar N., Babu K., and Jena S., "Face Expression Recognition Using Histograms of Oriented Gradients with Reduced Features," *in Proceedings of the International Conference on Computer Vision and Image Processing*, Uttarakhand, pp. 209-2019, 2016.

[19] Karataş F., Koyuncu I., Tuna M., Alçın M., and Avcioglu E., "Design and Implementation of Arrhythmic ECG Signals for Biomedical Engineering Applications on FPGA," *European Physical Journal Special Topics*, vol. 231, no. 5, pp. 869-884, 2022. https://doi.org/10.1140/epjs/s11734-021-00334-3

[20] Kohavi R., and John G., "Wrappers for Feature Subset Selection," *Artificial Intelligence*, vol. 97, no. 1-2, pp. 273-324, 1997. https://doi.org/10.1016/S0004-3702(97)00043-X

[21] Kołakowska A., Landowska A., Szwoch M., Szwoch W., and Wrobel M., *Human-Computer Systems Interaction: Backgrounds and Applications*, Springer, 2014. https://doi.org/10.1007/978-3-319-08491-6_5

[22] Lajevardi S. and Hussain Z., "Facial Expression Recognition Using Log-Gabor Filters and Local Binary Pattern Operators," *in Proceedings of the International Conference on Communication, Computer and Power*, Muscat, pp. 349-353, 2009.

[23] Lajevardi S. and Hussain Z., "Feature Selection for Facial Expression Recognition Based on

Optimization Algorithm," *in Proceedings of the 2nd International Workshop on Nonlinear Dynamics and Synchronization*, Klagenfurt, pp. 182-185, 2009. doi:10.1109/INDS.2009.5228001

[24] Lajevardi S. and Hussain Z., "Feature Extraction for Facial Expression Recognition Based on Hybrid Face Regions," *Advances in Electrical and Computer Engineering*, vol. 9, no. 3, pp. 63-67, 2009. DOI:10.4316/aece.2009.03012

[25] Lei L. and Kun S., "Speaker Recognition Using Wavelet Cepstral Coefficient, I-Vector, and Cosine Distance Scoring and its Application for Forensics," *Journal of Electrical and Computer Engineering*, vol. 2016, 2016. https://doi.org/10.1155/2016/4908412

[26] Liu Q., Chen C., Zhang Y., and Hu Z., "Feature Selection for Support Vector Machines with RBF Kernel," *Artificial Intelligence Review*, vol. 36, no. 2, pp. 99-115, 2011. https://doi.org/10.1007/s10462-011-9205-2

[27] Lowe D., "Distinctive Image Features from Scale-Invariant Keypoints," *International Journal of Computer Vision*, vol. 60, pp. 91-110, 2004. https://doi.org/10.1023/B:VISI.0000029664.99615.94

[28] Lucey P., Cohn J., Kanade T., Saragih J., and Ambadar Z., "The Extended Cohn-Kanade Dataset (CK+): A Complete Dataset for Action Unit and Emotion-Specified Expression," *in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition-Workshops*, San Francisco, pp. 94-101, 2010. doi: 10.1109/CVPRW.2010.5543262

[29] Lyons M., Akamatsu S., Kamachi M., and Gyoba J., "Coding Facial Expressions with Gabor Wavelets," *in Proceedings of the 3rd IEEE International Conference on Automatic Face and Gesture Recognition*, Nara, pp. 200-205, 1998. doi: 10.1109/AFGR.1998.670949

[30] Mallat S., "A Theory for Multiresolution Signal Decomposition: The Wavelet Representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 7, pp. 674-693, 1989. doi: 10.1109/34.192463

[31] Mallat S., *A Wavelet Tour of Signal Processing*, Academic Press, 1999.

[32] Martin D., O'neill M., Hubbard S., and Palmer A. "The Role of Emotion in Explaining Consumer Satisfaction and Future Behavioral Intention," *Journal of Services Marketing*, vol. 22, no. 3, pp. 224-236, 2008. DOI:10.1108/08876040810871183

[33] Mlakar U., Fister I., Brest J., and Potočnik B., "Multi-Objective Differential Evolution for Feature Selection in Facial Expression Recognition Systems," *Expert Systems with Applications*, vol. 89, pp. 129-137, 2017. https://doi.org/10.1016/j.eswa.2017.07.037

[34] Moeini A., Faez K., Sadeghi H., and Moeini H., "2D Facial Expression Recognition Via 3D Reconstruction and Feature Fusion," *Journal of Visual Communication and Image Representation*, vol. 35, pp. 1-14, 2016. https://doi.org/10.1016/j.jvcir.2015.11.006

[35] Mohammadi M., Fatemizadeh E., and Mahoor M., "PCA-based Dictionary Building for Accurate Facial Expression Recognition Via Sparse Representation," *Journal of Visual Communication and Image Representation*, vol. 25, no. 5, pp. 1082-1092, 2014. https://doi.org/10.1016/j.jvcir.2014.03.006

[36] Moore S. and Bowden R., "Local Binary Patterns for Multi-View Facial Expression Recognition," *Computer Vision and Image Understanding*, vol. 115, no. 4, pp. 541-558, 2011. https://doi.org/10.1016/j.cviu.2010.12.001

[37] Nazir M., Jan Z., and Sajjad M., "Facial Expression Recognition Using Histogram of Oriented Gradients Based Transformed Features," *Cluster Computing*, vol. 21, pp. 539-548, 2018. https://doi.org/10.1007/s10586-017-0921-5

[38] Ojala T., Pietikainen M., and Maenpaa T., "Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 971-987, 2002. doi: 10.1109/TPAMI.2002.1017623

[39] Paharia N., Jadon R., and Gupta S., "Feature Selection Using Improved Multiobjective and Opposition-based Competitive Binary Gray Wolf Optimizer for Facial Expression Recognition," *Journal of Electronic Imaging*, vol. 31, no. 3, pp. 033039-033039, 2022. DOI:10.1117/1.JEI.31.3.033039

[40] Perez-Gomez V., Rios-Figueroa H., Rechy-Ramirez E., Mezura-Montes E., and Marin-Hernandez A., "Feature Selection on 2D and 3D Geometric Features to Improve Facial Expression Recognition," *Sensors*, vol. 20, no. 17, pp. 4847, 2020. https://doi.org/10.3390/s20174847

[41] Sanmay D., "Filters, Wrappers and a Boosting-Based Hybrid for Feature Selection," *in Proceedings of the International Conference on Machine Learning*, San Francisco, pp.74-81, 2001.

[42] Siddiqi M., Alruwaili M., and Ali A., "A Novel Feature Selection Method for Video-based Human Activity Recognition Systems," *IEEE Access*, vol. 7, pp. 119593-119602, 2019. doi:10.1109/ACCESS.2019.2936621

[43] Soyel H., Tekguc U., and Demirel H., "Application of NSGA-II to Feature Selection for Facial Expression Recognition," *Computers and Electrical Engineering*, vol. 37, no. 6, pp. 1232-1240, 2011. https://doi.org/10.1016/j.compeleceng.2011.01.010

[44] Story M. and Congalton R., "Accuracy Assessment: A User's Perspective," *Photogrammetric Engineering and Remote Sensing*, vol. 52, no. 3, pp. 397-399, 1986.

[45] Suto J., Oniga S., and Sitar P., "Comparison of Wrapper and Filter Feature Selection Algorithms on Human Activity Recognition," *in Proceedings of the 6th International Conference on Computers Communications and Control*, Oradea, pp. 124-129, 2016. doi: 10.1109/ICCCC.2016.7496749

[46] Xiong L., Zhao Z., Pan J., Yang H., and Wang W., "Recognition of Heart Sound Based on Wavelet Cepstrum Coefficient and Probabilistic Neural Network," *Hans Journal of Biomedicine*, vol. 9, no. 1, pp. 10-16, 2019. DOI:10.12677/HJBM.2019.91002

[47] Ye Z., Mohamadian H., and Ye Y., "Information Measures for Biometric Identification Via 2D Discrete Wavelet Transform," *in Proceedings of the IEEE International Conference on Automation Science and Engineering*, Scottsdale, pp. 835-840, 2007. doi: 10.1109/COASE.2007.4341670

[48] Young S., Kershaw D., Odell J., and Ollason D., *The HTK Book*, Cambridge: Entropic Ltd, 1999.

[49] Youssif A. and Asker W., "Automatic Facial Expression Recognition System Based on Geometric and Appearance Features," *Computer and Information Science*, vol. 4, no. 2, pp. 115-124, 2011. DOI:10.5539/cis.v4n2p115

[50] Zhang D., *Fundamentals of Image Data Mining*, Springer, 2019. https://doi.org/10.1007/978-3-030-17989-2_3

[51] Zhao X. and Zhang S., "A Review on Facial Expression Recognition: Feature Extraction and Classification," *IETE Technical Review*, vol. 33, no. 5, pp. 505-517, 2016. DOI:10.1080/02564602.2015.1117403

Fatima Zohra BOUKHOBZA is a Doctoral student in Telecommunications and Networks at M'Hamed Bougara Boumerdes University. In 2017 she received her Master degree in Telecommunications, Networks and Multimedia from the University of Science and Technology Houari Boumediene (Algiers, Algeria). She joined the Department of Electrical Systems Engineering. She is a member of the Engineering system and telecommunications Laboratory, University M'Hammed Bougara Boumerdes, Algeria since 2020. Her research interests are: Image Processing, Computer Vision and Pattern Recognition.

Abdenour HACINE GHARBI received his Engineering and Magister degrees from University of Setif, Algeria in 1995 and 2002, respectively. He received his PhD from the University of Setif, Algeria and the University of Orlèans, France in 2012. In 2021, he received his HDR degree from the university of Bordj Bou Arréridj, Algeria. He is currently a lecturer at the Electronics Department of Bordj BouArreridj University, Algeria and member of LMSE Laboratory. His main research interests are Pattern Recognition, Classification and Feature Selection.

Khaled ROUABAH received his engineering degree in electronics from the University of Farhat Abbas, Setif, Algeria, in 1999. He obtained his Master's degree in telecommunications and networking in 2001 from Higher Institute of Aeronautics and Space, Toulouse, France, his Magister degree in communications and his Ph.D. degree in electronics engineering from the University of Farhat Abbas, Setif, Algeria, in 2005 and 2010 respectively. In 2013, he obtained his HDR degree in electronics in Mohamed BOUDIAF University of Msila. Between July 2018 and October 2022, he worked as a professor in the Department of Electronics at Mohamed EL Bachir El Ibrahimi University, Bordj Bou Arreridj. He is currently holding the same position at the electronics department of Mohamed BOUDIAF University of Msila. His research interests include Communication Signals, Geolocation, Parallel Processing, Pattern Recognition, Hardware Implementation, Signal Structure Design, Mobile Computing and Telecommunications Network.