# Relevant Facial Key Parts and Feature Points for Emotion Recognition

Rim Afdhal, Ridha Ejbali and Mourad Zaied
*Research Team on Intelligent Machines, National School of Engineers of Gabès, University of Gabès, Avenue Omar Ibn El Khattab, Zrig Eddakhlania 6029, Gabès, Tunisia*

Keywords: Emotion Recognition, Facial Expressions, Key Parts, Feature Points, Classification Rates.

Abstract: Interaction between people is more than just verbal communication. According to scientific researches, human beings trust a lot on non-verbal techniques of communication, particularly communication and understanding each other via facial expressions. Facial expressions are more descriptive in situations where words fail, such as a surprise or a shock. In addition, lying via spoken words is harder to notice compared to faking expressions. Focusing on geometric positions of facial key parts and well detecting them is the best strategy to boost the classification rates of emotion recognition systems. The goal of this paper is to find the most relevant part of human face which is responsible to express a given emotion using feature points and to define a primary emotion by a minimum number of characteristic points. The proposed system contains four main parts: the face detection, the points location, the information extraction, and finally the classification.

## 1 INTRODUCTION

Emotions are basic characteristics of human beings that play a relevant role in social communication (Foggia.P et al., 2023). People use different manners to express an emotion, like facial expression (Tuncer.T et al., 2023), speech (Pawar and Kokate, 2021) (R.Ejbali et al., 2015) body language (Leong et al., 2023) (I.Khalifa et al., 2019). The human face, as the most expressive and exposed part of the body (R.Afdhal et al., 2021), enables the use of computer vision systems to analyze the images containing faces for emotions recognition. Facial expressions (ELsayed.Y et al., 2023) have a relevant role in recognizing emotions as well as are used in the task of non-verbal communication, and to identify humans. An automatic facial expressions analysis is very relevant for automatic understanding of people (Y.Siwar et al., 2016)(Y.Siwar et al., 2020), their actions as well as their behavior. Facial expression has been a hot topic of research in human behavior (Huang.ZY et al., 2023). It figures surely in research on almost every aspect of emotion such as psychophysiology (Bourke C., 2010), neural correlate , perception, addiction, social processes, depression as well as other emotion disorders. Moreover, it communicates physical pain (Rawal, 2022), alertness, personality as well as well as interpersonal relations. Automated emotions detection using facial expressions is also fundamental for human–robot interaction (F. De la Torre and Cohn., 2015), drowsy driver detection (Fuletra, 2013), analyzing mother–infant interaction (A. R. Taheri and Basiri, 2014), autism, social robotics (So et al., 2018), as well as expression mapping for video gaming.

This paper aims at presenting the following contributions:

- Find the key part of the face which is responsible to express a given emotion using feature points.
- Enhance a classic emotion recognition system by decreasing the number of feature points in order to recognize a basic emotion (joy, fear, disgust, surprise, sadness and anger).

The remainder of this paper is structured as follows: In section 2, we introduce literature related works. Section 3 describes the proposed approach. Section 4 presents the results as well as related discussion. Finally, section 5 gives a conclusion of this work.

## 2 RELATED WORKS

In this section, we show the various works in the literature corresponding to our field of interest. Two popular methods used mostly in the literature for the facial expressions recognition systems are based on geometry and appearance (Iqbal et al., 2023). The recognition process of emotions runs up against the issue



Afdhal, R., Ejbali, R. and Zaied, M.
Relevant Facial Key Parts and Feature Points for Emotion Recognition.
DOI: 10.5220/0012464300003636

In *Proceedings of the 16th International Conference on Agents and Artificial Intelligence (ICAART 2024) - Volume 3*, pages 1270-1277
ISBN: 978-989-758-680-4; ISSN: 2184-433X

of the complexity and the diversity of the human being expression and the relevant flow of data which it generates. The solution of this issue would be to detect feature landmarks of the face and to extract for them the useful data in the operation of recognizing emotions.

The approach presented in (Tarnowski et al., 2017) describes a model (Candid3) which is based on 121 characteristic landmarks of the face. These landmarks are located on characteristic positions on the face namely, the corners of the mouth, the nose, the cheekbones, the eyebrows, etc.

(Salmam et al., 2018) have suggested an emotion recognition technique using facial expression. The technique was based on four facial key parts (eyebrows, eyes, nose, and mouth). First, they detected the face, then, they located and tracked 49 feature points using a Supervised Decent Method. Then, they calculate the distances between each pair of points. Finally, to classify emotions they used a neural network classifier.

In this paper, we use an improved approach which aims at decreasing the number of feature points and boosting the classification rates in order to recognize an emotion.

# 3 PROPOSED APPROACH

In this section, we present an overview of our suggested system for identifying primary emotions, followed by a detailed description of the key components of our proposed methodology. The figure 1 shows the different steps of a classic emotion recognition system. We enhanced this system by reducing the number of feature points at the level of location points step.

## 3.1 Overview of Our System

In this paper, we proposed an improved approach aiming at developing an emotion recognition system able to recognize the six primary emotions with high classification rates and minimum number of characteristic points.
The proposed system presents the importance of every human face elements' on the classification phase or the recognition of primary emotions. The system contains four main steps. To detect the human face, we used Viola-Jones detector. Then, we located the feature points. Besides, we extract information. Finally, we classified the variations of biometric distances in order to have the appropriate emotion using KNN (K-Nearest Neighbor) and SVM classifiers.

We used two methods in order to acheive our goal. The first method is shown at figure2. We started our work by a pre-treatment.
The different sub steps of the first method are as follows:

- We divide the used data sets on subdatasets based on the key parts of every emotion.
- We prepare the first category of dataset which based on three key parts (eyes,eyebrows and mouth). We have two subdatasests (surprise and disgust).
- We prepare the second category of dataset which based on two key parts (eyes and mouth). We have three subdatasests (sadness, fear and anger).
- We prepare the third category of dataset which based on one key part (mouth). We have one subdatasest (joy).
- We detect the face using viola and jones detector.
- We locate the feature points on the face.
- We compute the variations of biometic disatnces.
- We use the KNN and the SVM for the classification.

The second method is shown at figure 3. The different sub steps of this method are as follows:

- We detect the face.
- We hide the down part of the face ( mouth and nose).
- We hide the upper part of the face (eyes and eyebrows).
- We locate the feature points on the not hidden parts of the face.
- We compute the variations of biometic disatnces.
- We classify the obtained distances using SVM and KNN.

## 3.2 Face Detection

The first fundamental stage in all facial analysis techniques is face detection (R.Afdhal et al., 2015). The computer system needs some training to effectively detect faces so that it can quickly determine if anything is a face or not. Several threshold values have been established in order to detect faces. A system can identify human faces based on this value (Salmam et al., 2018).
To complete this phase, we employed the Viola-Jones detector (R.Afdhal et al., 2017). One of the most often used techniques for object detection is the Viola-Jones approach. This approach can offer results with high

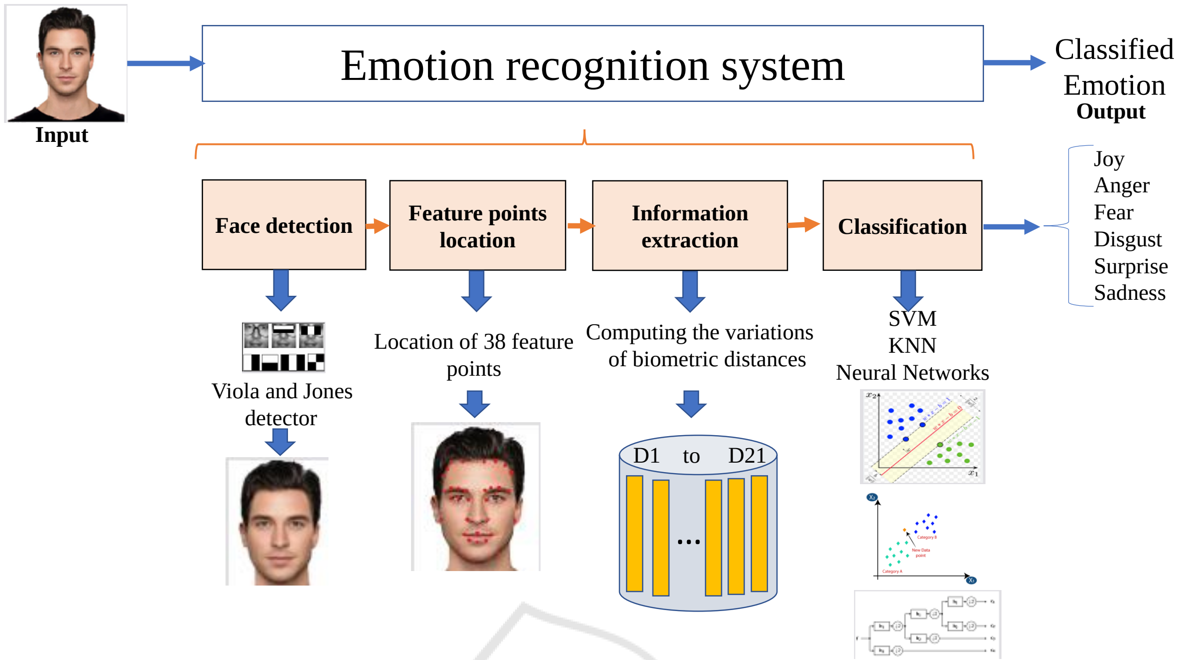


Figure 1: Classic emotion recognition system.

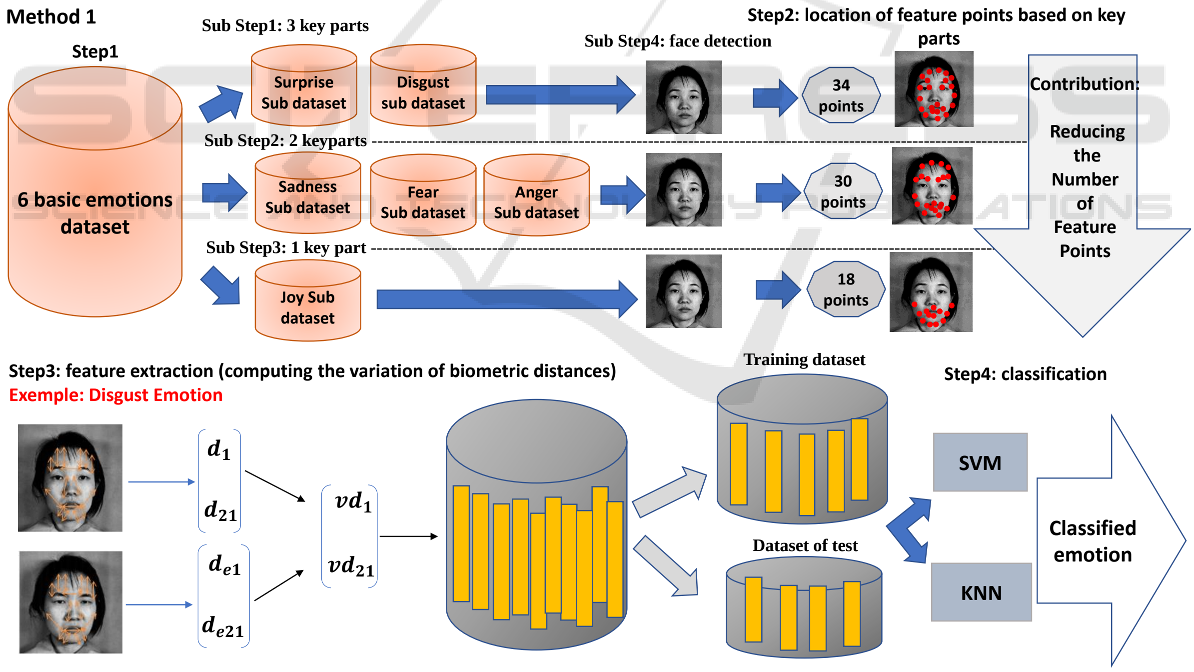


Figure 2: The different steps of the first method.

accuracy of about 93.7% with a speed of 15 times faster. Viola-Jones detector uses rectangular Haar-like features.

### 3.3 Feature Points Location

The feature points location is the second step of the proposed system. We achieved this step using the five rectangles technique (R.Afdhal et al., 2020). It's a very simple technique. It is based on 5 boxes detected

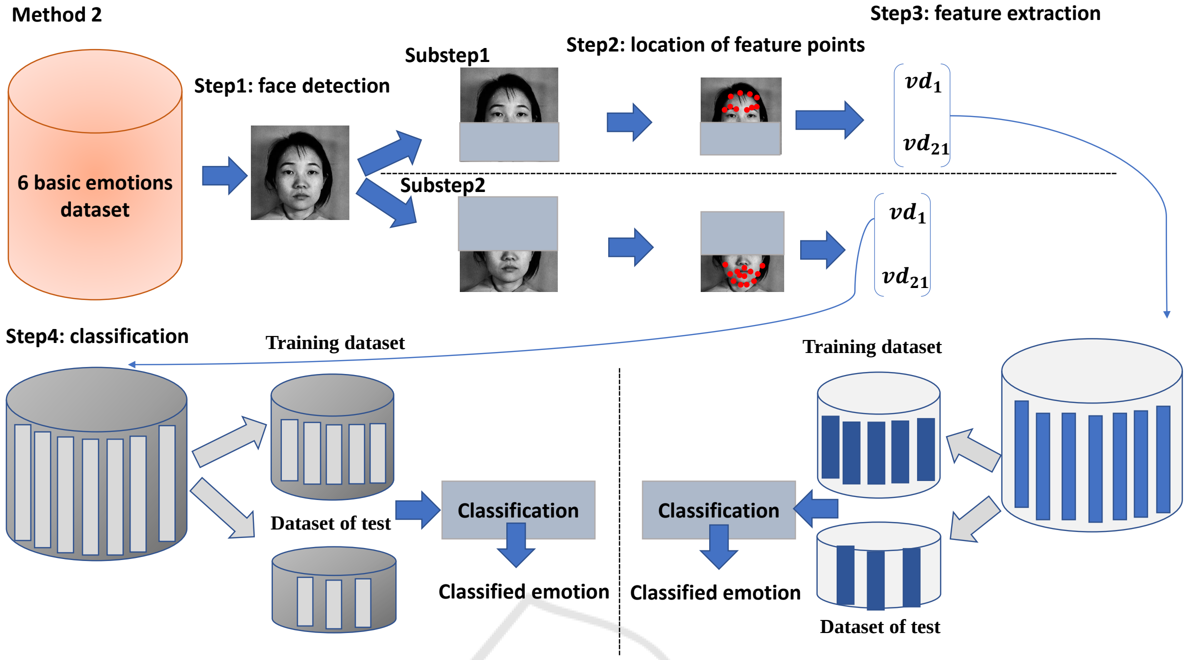


Figure 3: The different steps of the second method.

at the level of facial elements using Viola-Jones detector. To precise the position of each landmark in the human face, we focus on the result of the selection boxes (BBOX) of the face's elements detection (R.Afdhal et al., 2014). The first box was located on the face, the second on the right eye, the third on the left eye, the fourth on the nose, finally the last box was located on the mouth. The feature points were located with the coordinates of the points characterizing each box. The BBOX gives an M-by-4 matrix. It defines M bounding boxes containing the detected objects. This technique provides multi-scale object detection on the input image. The output matrix row, BBOX, contains a vector, [x y width height], that gives in pixels, the upper left corner as well as the size of the box.

$$box_{righteye} = [x_{reye} y_{reye} width_{reye} height_{reye}] \quad (1)$$

$$box_{lefteye} = [x_{leye} y_{leye} width_{leye} height_{leye}] \quad (2)$$

$$box_{face} = [x_{face} y_{face} width_{face} height_{face}] \quad (3)$$

$$box_{nose} = [x_{nose} y_{nose} width_{nose} height_{nose}] \quad (4)$$

$$box_{mouth} = [x_{mouth} y_{mouth} width_{mouth} height_{mouth}] \quad (5)$$

We will precise in this step the number of feature points to every emotion. We used two methods as mentioned above:

method1 : In this method, we are based on the definitions of facial expressions with MPEG 4 Norm and the facial key parts of every emotion mentioned in (Plutchik, 1980). Table 1 presents the key parts of each primary emotion and Table 2 shows the different distances corresponding to each facial key part.Table 3 summarizes the distances as well as the feature points of emotions.

Table 1: Key parts of primary emotions.

| Emotion | Key parts |
|---|---|
| Surprise<br>Disgust | Eyebrows, Eyes and Mouth |
| Sadness<br>Fear<br>Anger | Eyebrows and Mouth |
| Joy | Mouth |

Table 2: Key parts Distances.

| Key part movements | Distances |
|---|---|
| Eyebrow movements | D1 to D7 |
| Eye movements | D8 and D9 |
| Mouth movements | D10 to D19 |

We will use for the emotion of surprise and disgust three key parts eyebrows, eyes and mouth. So, to recognize these two emotions we will use 34 feature points and 19 biometric distances as presented at figure 4.

We will locate for the emotion of sadness, fear and anger two key parts eyebrows and the mouth. So, we will eliminate the feature points located on the eyes. To recognize these three emotions we will use 30 feature points and we will eliminate 4 points and 2 bio-

Table 3: Distances and characteristic points of primary emotions.

| Emotion | Distances | Characteristic points |
|---|---|---|
| Surprise Disgust | D1 to D19 | 34 |
| Sadness, Fear, Anger | D1 to D7, D10 to D19 | 30 |
| Joy | D10 to D19 | 18 |

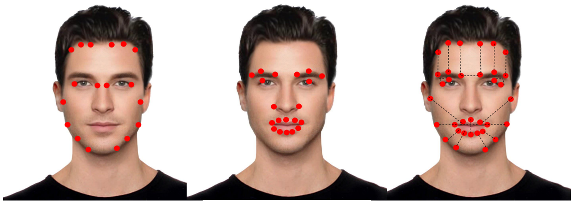

Figure 4: Static points- dynamic points - biometric distances of surprise and disgust emotions.

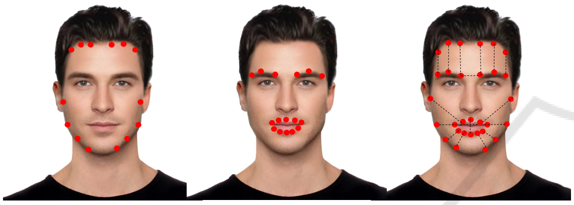

Figure 5: Characteristic points - biometric distances of sadness, fear and anger emotions.

metric distances as presented at the figure 5. We will use for the emotion of joy only the mouth so for this emotion we will locate the feature points of the mouth and we will eliminate all the points located on the eyes and the eyebrows. So, we will use 18 feature points and 10 biometric distances.

method2: The idea is to hide the key parts of the face to each emotion in order to see the influence of each part. Then, we computed the number of points to locate for each emotion. We hid the lower part of the face (nose and mouth). To classify emotions in this case, we used only two key parts (eyes and eyebrows). As well as, we located only the eyes and eyebrows feature points. After that we hid the upper part of the face (eyebrows and eyes). To classify emotions in this case, we used only the mouth.

## 3.4 Information Extraction

An expressive human face is a deformation from the neutral face. The differences between a neutral face and its corresponding expressive face have lots of data related to facial expressions. During the production of a facial expression, a set of transformations in the facial features appears on the face (R.Afdhal et al., 2016). The idea is to classify different expressions using the variation of some distances computed from the coordinates of the points of interest that delimit

Figure 6: The biometric distances.

the eyes, nose, mouth and eyebrows from the neutral state (R.Afdhal et al., 2020). For facial movement processing, we represent each muscle by two points, a static point and a dynamic point. Static points are points that do not move during the production of a facial expression, while dynamic points undergo movement according to the type of expression. Figure 5 shows the different Euclidean distances used to characterize the movement of the facial muscles. D19 is the biometric distance that relates two feature points p26 and p35. It is computed as follows: $D19 = \sqrt{((p_{26}(x) - p_{35}(y))^2 + (p_{26}(x) - p_{35}(y))^2)}$

## 3.5 Classification

In order to classify emotions, We choose the KNN method. It is a simple Machine Learning algorithm based on Supervised Learning technique, to categorize emotions. The KNN technique places the new instance in the category that is most similar to the available categories by assuming similarity between the new data and existing cases. A new data point is classified using the KNN algorithm depend on similarity after all the existing data has been stored. This means that as fresh data is generated, it may be rapidly and correctly categorized using the KNN technique. These are the different steps of the algorithm:

- Choose the number K of the neighbors.
- Compute the Euclidean distance of K number of neighbors.
- Take the K closest neighbors, determined by the Euclidean distance.
- Compute the amount of data landmarks in each category among these k neighbors.

Table 4: Classification rates of primary emotions of the Chon-kanade data base.

| Emotion | Method1 using SVM | Method1 using KNN | (L.Nwosu et al., 2017) | (R.Afdhal et al., 2020) |
|---|---|---|---|---|
| Surprise | 99% | 98.5% | 98.4% | 99.3% |
| Disgust | 97% | 96% | 97.2% | 99.2% |
| Anger | 94% | 93% | 93.7% | 98.5% |
| Fear | 80% | 90% | 92.2% | 95% |
| Sadness | 87% | 85.5% | 91.8% | 92.31% |
| Joy | 99% | 98.5% | 98.1% | 98.5% |

- Assign the additional data landmarks to the category where the neighbor count is at its highest.
- The model is complete.

In order to enhance the emotions' classification rates, we used the Support Vector Machine classifier. It is a preferment machine learning technique used for nonlinear and linear classification, regression, as well as outlier detection issues. It can be used for many applications, including face detection, spam detection, handwriting identification, gene expression analysis, text classification as well as picture classification. SVMs can handle high-dimensional data and nonlinear interactions, making them flexible in a wide range of applications.

# 4 RESULTS AND DISCUSSION

This section shows the results obtained by the proposed approach. As mentioned above, we used two methods in order to precise the most relevant facial key parts of every primary emotion. If an emotion is defined by a few number of key parts the number of feature points will decrease.

We have used two datasets: the Chon-kanade and JaFFE (Japanese Female Facial Expression) data basis. The Chon-kanade data basis has 593 images from 123 persons. The JaFFE data basis contains 213 images of different facial expression from 10 japanese female persons.

Table 4 shows the classification rates of our first method using SVM and KNN classifiers and the other approaches (L.Nwosu et al., 2017) and (R.Afdhal et al., 2020). This method (L.Nwosu et al., 2017) is based on a two-channel convolutional neural networks. The extracted eyes is the input of the first convolutional layer and the mouth is the input of the second convolutional layer. Information from the two channels converge in fully connected layer which is used to learn global information as well as is used for classification. The second method (R.Afdhal et al., 2020) contains three phases: pre-treatment, feature extraction as well as classification. This method used the deep learning for the feature extraction and the fuzzy logic for the classification. This comparison is conducted to validate the performance of the suggested method.

Table 5: Overall accuracy of JaFFE dataset.

| Approach | Classification rates |
|---|---|
| (V.Chernykh et al., 2017) | 73% |
| (Abate1 et al., 2022) | 90,42% |
| (Chung et al., 2019) | 72.16% |
| (Salmam et al., 2018) | 93.8% |
| Our method | 95% |

As shown in table 5, obviously, our proposed method outperforms the cited approaches by as much as 22%, 4.58%, 22.84%, and 1.2% respectively.

The tables 6 and 7 present the results of the second method. Table 6 presents the classification rates when we hid the lower part of the face (nose and mouth). To classify emotions in this case, we used only two key parts (eyes and eyebrows). As well as, we located only the eyes and eyebrows feature points. The results show the influence of the elimination of the feature points located at the level of the mouth.

Table 6: The classification rates using the upper part of the face.

| Emotion | Rates with KNN | Rates with SVM |
|---|---|---|
| Surprise | 15% | 25% |
| Disgust | 10% | 12% |
| Anger | 60% | 65% |
| Fear | 50% | 55% |
| Sadness | 10% | 13.5% |
| Joy | 19% | 20% |

Table 7: The classification rates using the down part of the face.

| Emotion | Rates with KNN | Rates with SVM |
|---|---|---|
| Surprise | 45% | 50% |
| Disgust | 42% | 55% |
| Anger | 10% | 15% |
| Fear | 30% | 35% |
| Sadness | 35% | 40% |
| Joy | 80% | 83% |

As shown at table 6 the surprise and disgust were classified with low classification rates 15% with KNN classifier and 25% with SVM classifier, 10% with KNN classifier and 12% with SVM classifier, respectively. These low rates show the importance of the mouth to classify these two emotions. We observed also if we use the down part of the face the classification rates of these emotions increase. So, to perfectly classify surprise and disgust emotions, we must use three facial key parts (mouth, eyebrows and eyes) which confirm our first method. Anger, fear and sadness were defined according to our first method by two key parts (eyes and eyebrows).

Table 7 shows high classification rates of anger and fear emotions which prove the first method. However, the sadness emotion was classified by low classification rates. So, for this emotion the mouth is relevant to better classify this emotion.

Finally, joy emotion was defined by one key part (the mouth). The second method confirms this definition. When we hid the mouth, the emotion was classified with low classification rates(19% with KNN and 20% with SVM). However, table 7 shows high classification rates of this emotion when, we didn't eliminate the mouth.

## 5 CONCLUSION

The goal of this study is to define a primary emotion (joy, fear, sadness...) by a minimum number of characteristic points which will be very useful especially for real time applications. Our proposed system consists of four phases: face detection, characteristic points localization, information extraction and classification. We contribute to two levels: to assign the most important facial key parts in order to perfectly recognize a basic emotion and to reduce the number of characteristic points to define an emotion. The rates in the experimental results present the effectiveness of the suggested system.

## ACKNOWLEDGEMENTS

The authors would like to acknowledge the financial support of this work by grants from General Direction of Scientific Research (DGRST), Tunisia, under the ARUB.

## REFERENCES

A. R. Taheri, M. Alemi, A. M. H. R. P. and Basiri, N. M. (2014). *Social robots as assistants for autism therapy in Iran: Research in progress*. International Conference on Robotics and Mechatronics (ICRoM), Tehran, Iran, 2nd edition.

Abate1, A. F., Cimmino, L., Narducci, B.-C. M. F., and Pop, F. (2022). The limitations for expression recognition in computer vision introduced by facial masks. In *Multimedia Tools and Applications*.

Bourke C., Douglas K., P. R. (2010). Processing of facial emotion expression in major depression: a review. In *Aust. N.Z. J. Psychiatry*.

Chung, C.-C., Lin, W.-T., Zhang, R., Liang, K.-W., and Chang, P.-C. (2019). Emotion estimation by joint facial expression and speech tonality using evolutionary deep learning structures. In *IEEE Global Conference on Consumer Electronics (GCCE)*, pages 12–14.

ELsayed.Y, ELSayed.A, and Abdou.M.A. (2023). An automatic improved facial expression recognition for masked faces. In *Neural Computing and Applications*, page 14963–14972.

F. De la Torre, W. S. Chu, X. X. F. V. X. D. and Cohn., J. (2015). *IntraFace*. IEEE International Conference and Workshops on Automatic Face and Gesture Recognition, Ljubljana, Slovenia,, 11th edition.

Foggia.P, Greco.A, Saggese.A, and Vento.M" (2023). Multi-task learning on the edge for effective gender, age, ethnicity and emotion recognition. In *Eng. Appl. Artif. Intell*.

Fuletra, J.D.; Bosamiya, D. (2013). Int. j. recent innov. trends comput. commun. In *Int J of Soc Robotics*.

Huang.ZY, Chiang.CC, and Chen.JH (2023). A study on computer vision for facial emotion recognition. In *Sci Rep*.

I.Khalifa, R.Ejbali, and , M. (2019). *Body gesture modeling for psychology analysis in job interview based on deep spatio-temporal approach*. Parallel and Distributed Computing, Applications and Technologies: 19th International Conference, PDCAT, Jeju Island, South Korea.

Iqbal, J. M., Kumar, M. S., G.R., G. M., A.N., S., Karthik, A., and N., B. (2023). Facial emotion recognition using geometrical features based deep learning techniques. In *International Journal of Computers Communications and Control*.

Leong, S., Tang, Y. M., Lai, C. H., and Lee., C. (2023). Using a social robot to teach gestural recognition and production in children with autism spectrum disorders, disability and rehabilitation. In *Assistive Technology*.

L.Nwosu, H.Wang, J.Lu, I.U.X.Yang, and T.Zhang (2017). *Deep Convolutional Neural Network for Facial Expression Recognition Using Facial Parts*. Int. Conf. Dependable, Autonomic and Secure Computing, Orlando, FL, USA, 15th edition.

Pawar, M. and Kokate, R. (2021). Convolution neural network based automatic speech emotion recognition using mel-frequency cepstrum coefficients. In *Multimed. Tools Appl.* Springer.

Plutchik, R. (1980). A general psychoevolutionary theory of emotion. In *Theories of Emotion*. Elsevier.

R.Afdhal, A.Bahar, R.Ejbali, and , M. (2015). *Face detection using beta wavelet filter and cascade classifier entrained with Adaboost*. Eighth International Conference on Machine Vision (ICMV), BARCELONA, SPAIN.

R.Afdhal, R.Ejbali, and , M. (2014). *Emotion recognition using features distances classified by wavelets network and trained by fast wavelets transform*. 14th International Conference on Hybrid Intelligent Systems, Kuwait.

R.Afdhal, R.Ejbali, and , M. (2016). *Emotion recognition using the shapes of the wrinkles*. 19th International Conference on Computer and Information Technology (ICCIT), Bangladesh.

R.Afdhal, R.Ejbali, and , M. (2017). *A Hybrid System Based on Wrinkles Shapes and Biometric Distances for Emotion Recognition*. The Tenth International Conference on Advances in Computer-Human Interactions, Nice, France.

R.Afdhal, R.Ejbali, and , M. (2020). Emotion recognition by a hybrid system based on the features of distances and the shapes of the wrinkles. In *The Computer Journal,*.

R.Afdhal, R.Ejbali, and , M. (2021). Primary emotions and recognition of their intensities. In *The Computer Journal,*.

Rawal, N., S.-H. R. (2022). Facial emotion expressions in human–robot interaction: A survey. In *Int J of Soc Robotics*.

R.Ejbali, O.Jemai, M., and Amar, C. (2015). *A speech recognition system using fast learning algorithm and beta wavelet network*. 15th International Conference on Intelligent Systems Design and Applications (ISDA), Marrakesh, Morocco.

Salmam, F. Z., Madani, A., and Kissi, M. (2018). Emotion recognition from facial expression based on fiducial points detection and using neural network. In *International Journal of Electrical and Computer Engineering*.

So, W.-C., Wong, M. K.-Y., Lam, C. K.-Y., Lam, W.-Y., Chui, A. T.-F., Lee, T.-L., Ng, H.-M., Chan, C.-H., and Fok, D. C.-W. (2018). Using a social robot to teach gestural recognition and production in children with autism spectrum disorders, disability and rehabilitation. In *Assistive Technology*, pages 527–539.

Tarnowski, P., Kołodziej, M., Majkowski, A., and Rak., R. J. (2017). Emotion recognition using facial expressions. In *International Conference on Computational Science*, pages 12–14, Zurich, Switzerland.

Tuncer.T, Dogan.S, and Subasi.A (2023). Automated facial expression recognition using novel textural transformation. In *J Ambient Intell Human Comput*, page 9435–9449.

V.Chernykh, G.Sterling, and P.Prihodko (2017). Emotion recognition from speech with recurrent neural networks.

Y.Siwar, J.Olfa, E.Ridha, and , Z. (2016). *Improving the classification of emotions by studying facial feature*. International Journal of Computer Theory and Engineering.

Y.Siwar, S.Salwa, and , Z. (2020). *A novel classification approach based on Extreme Learning Machine and Wavelet Neural Networks*. Multimedia Tools and Applications.