Deep Learning Model to Analyze Customer's Satisfaction

Bouzakraoui Moulay Smail, Sadiq Abdelalim, Youssfi Alaoui Abdessamad

Abstract: *Nowadays, measuring customer satisfaction is an important strategic tool for companies; many manual methods exist to measure customer's satisfaction. However, the results have not effective and efficient. In this paper, we propose a new method for facial emotion detection to recognize customer's satisfaction using a deep learning model. We used a convolutional neural network to detect facial key points. These key points help us to extract geometric features from customer's emotional faces. Indeed, we computed distances between neutral face and negative or positive feedback. After that, we classified these distances by using Support Vector Machine (SVM), KNN, Random Forest, and Decision Tree. To evaluate the performance of our approach, we tested our algorithm by using FACEDB and JAFFE datasets. We found that SVM is the most performant classifier. We obtained 96% as accuracy by using FACEDB dataset and 95% by using JAFFE dataset.*

Keywords: Emotion recognition, Customer satisfaction, SVM, KNN, Decision Tree.

I. INTRODUCTION

Measuring customer satisfaction makes it possible to better meet their needs, to personalize the relationship with the customer and the offer of products or services. A satisfied customer is a loyal client, synonymous with additional sales and therefore turnover. Customer satisfaction can be measured using manuals methods such as the satisfaction survey, the interview, and the focus group. These methods are not efficient and effective to the cost, time and reliability of the data. Facial expressions are used to communicate non-verbally. They are a special way to express our emotions and appreciations. In the context of customer satisfaction, a negative feedback emotion is often related to a lower perceived quality of service [1]. The facial expression contributes 55% to communicate with a speech. More than that, 70-95% of negative feedback can be understood verbally [2]. Companies have long been interested in understanding the purchasing decision-making process of consumers, their reactions to products, what attracts their attention and what remains unobserved [3]. In this work, we

Bouzakraoui Moulay Smail*, SIM team of MISC LaboratoryFaculty of Science, University IBN TOFAIL, Kenitra, Morocco

Sadiq Abdelalim, SIM team of MISC LaboratoryFaculty of Science, University IBN TOFAIL, Kenitra, Morocco

Youssfi Alaoui Abdessamad, IRDA Team, ADMIR Laboratory, Rabat IT Center, ENSIAS, Mohammed V University, Rabat 10000, Morocco, abdessamad.youssfialaoui@gmail.com.

aim at detecting the positive or negative emotions of customers from the analysis of facial expressions. In this way, this type of information is useful. For example, we can calculate the statistics about the products as well as the replacement of their exposures. Appreciated products must be brought to the fore, while negatively appreciated products must be replaced or otherwise requested. For this reason, we are so existed to propose a new method for facial emotion detection to recognize customers' satisfaction using a deep learning model. This paper is organized as follows: Firstly, we offer an outline of the related work. Secondly, we introduce our approach that contains a system for customer satisfaction using a deep learning model. To evaluate the performance of our method we reserved a separate section to the experimental results. In the last section, we sum up our work.

II. RELATED WORK

Charles Darwin was the first to provide robust basis for emotions, representing their significance, their usefulness and communication. Already, Darwin argued that many of the facial expressions of emotions have an adaptive significance that serve to communicate something, most emotions are expressed alike on the human face nerveless of culture or race [4], his work "Expression of emotions in humans and animals", remains a reference for many scientists. Psychologists Paul Ekman and Friesen developed in 1978 a method to classify muscle movement to code the facial expression. The facial movements are determined through action units (AU) [5]. Each AU is based on the affinity of the muscles. This method, used mostly in classifying facial movement named Facial Action Coding System (FACS). According to Ekman, we can define six universal basic emotions: happy, sadness, anger, surprise, fear and disgust. Various researchers have supported the universality of the expression of emotions. Generally, in reaction to similar stimuli, people prompt similar expressions with local variations [6]. Facial feature extraction is an important step in the process of emotion recognition. Frequently, we used two methods: motion-based method and deformation of face. The change of face is taken into account for the motion-based method [7, 8]. On the contrary, we take into account a neutral image and another image for the deformation-based method [9,10]. Feature extraction contains geometric and appearance-based models. The first model involves feature extraction about eye, mouth, nose, eyebrow…However, the second model covers the specific part of the face [11].

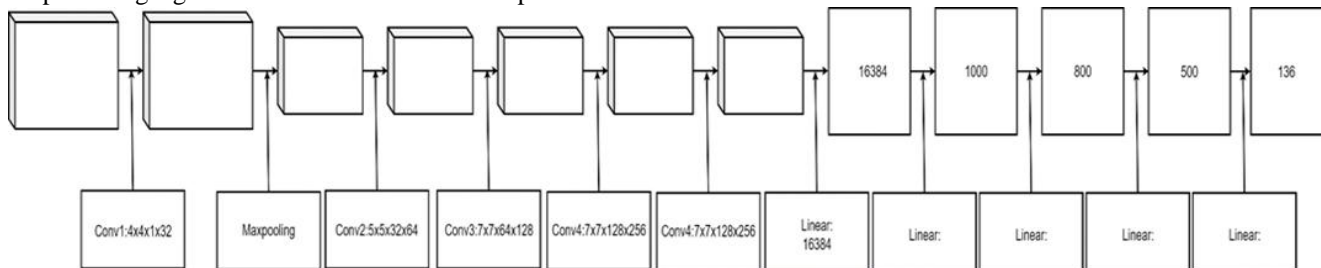Deep learning algorithms are also used to extract pattern from

Fig. 1.Architecture of CNN model used to detect key points of the face.

facial expressions. More specifically, we develop a convolutional neural network to detect features from faces, and to classify these features to different emotions. This type of network contains two parts. The first part contains convolutional layers, which are based on applying mathematic convolution operation. The result of this part is features extracted from faces. The second part contains feed-forward network. This part of network is able to classify features extracted to different facial expression. Shan, Rashid et al. [12, 13] proposed based results researches of this algorithm. Affdex by Affectiva [14] is one of the most implements for visual emotional analysis.

Which permits giving the emotional trend of a subject, through detection of Ekman & Keltner's emotions [15], and the Microsoft Cognitive Services based on the Azure platform [16]. These two studies are able to perceive age, sex and ethnicity and depend on Convolutional and/or Recurrent Neural Network [17]. A multimodal affect recognition system was developed to determine if the customer exhibits negative affect such as being unhappy, disgusted, frustrated, angry or positive affect such as happy, satisfied and content with the product being offered [18].

To appreciate customer satisfaction M Slim et al.[19] proposed a technique founded to special facial emotion namely" happy"," surprised" and "neutral". In the last years, eye-tracking system has been developed to analyze the research of customer [20]. Companies invest intensely within the advancement and publicizing of products, the return on investment (ROI) must be justified by the organization. Hence, the specific observing of the fascination of a product by the customer amid its promotion, and how to develop their promotion campaigns. Eye tracking is very promising in navigation advertising. Shergill et al. [21] proposed an algorithm to compute sales assistant, which depends specially on emotion detection. Desmet et al. [22] published a study, which contains products features impact the costumer's emotional, and they used 21 emotional categories of products.

III. OUR APPROACH

We analyzed facial emotion to measure the satisfaction of a client. Firstly, we achieved the image of the face. Secondly, we detected features by computing a geometric feature from the face. Precisely, we converted the input image to a geometric primitive like points and curves. After detecting faces into an image, we used a deep learning architecture, which contains convolution layers. We utilized a convolution neural network to extract features from image, which contains face. This type of models contains convolution layers followed by a max pooling operation, and the fully

connected layer to reshape the output as vector. In our model, we developed a CNN architecture, which contains five convolutional layers. The first layer is connected with 32 weights, the dimension of these weights is 4x4. After that, we applied a max pooling operation. In second layer, we applied a convolution operation with 64 weight, where each weight is 5x5 filter. After that, we applied the third convolution layer, which contains 128 weight, and each weight contains 7x7 pixels. The fourth layer is realized with 256 weight. In addition, each weight contains 7x7 pixels. The last layer contains 256 weight, and each weight contains 7x7 pixels. The follow equation explains the convolution operation:

$$(A*F)_{sr} = \sum_{p=0}^{k_1-1}\sum_{q=0}^{k_2-1} A(s-p, r-q)F(p,q)$$

$$= \sum_{p=0}^{k_1-1}\sum_{q=0}^{k_2-1} I(s+p, r+q)K(-p,-q)$$

Where A is the input image, F is the weight using to compute convolution, and $k_1 * k_2$ is the dimension of the weight.

Linear model follows the convolution layers. Which contains six layers with linear operations. In the first layer, we resulted a vector with 16i 384 values. Secondly, we applied linear layer to product a vector with 1000 values. In the fourth layer, we applied a linear operation to calculate a vector with 136 values; these values are the coordinates of 68 Key points of the face. Fig.1 illustrates the architecture of our model.

A. Features extraction

After computing key points, we have to extract features from these key points. We have to evaluate the deformation of the face. For this reason, we computed the variation of key points into two different images of the same face, which explain two different expressions. Firstly, we detected 68 key points using CNN model of neutral facial expression and positive or negative facial expression. Then, we computed the ratio CNN model of neutral facial expression and positive or negative facial expression. Then, we computed the ratio between the same key points into the same face of two different face expression. The follow equation and figure demonstrate the operation to detect ratios:

$$I_{neutrel}(x_1, y_1, ..., x_{68}, y_{68})$$

$$I_{positive}\left(x_1^{'}, y_1^{'},...,x_{68}^{'}, y_{68}^{'}\right)$$

$$I_{negative}\left(x_1^{''}, y_1^{''},...,x_{68}^{''}, y_{68}^{''}\right)$$

$$D\left(I_{neutrel}, I_{neutrel}\right) =$$
$$\left(\sqrt{\left(x_1 - x_1\right)^2 + \left(y_1 - y_1\right)^2} ,..., \sqrt{\left(x_{68} - x_{68}\right)^2 + \left(y_{68} - y_{68}\right)^2}\right)$$

$$D\left(I_{neutral}, I_{positive}\right) =$$
$$\left(\sqrt{\left(x_1 - x_1^{'}\right)^2 + \left(y_1 - y_1^{'}\right)^2} ,..., \sqrt{\left(x_{68} - x_{68}^{'}\right)^2 + \left(y_{68} - y_{68}^{'}\right)^2}\right)$$

$$D\left(I_{neutral}, I_{negative}\right) =$$
$$\left(\sqrt{\left(x_1 - x_1^{''}\right)^2 + \left(y_1 - y_1^{''}\right)^2} ,..., \sqrt{\left(x_{68} - x_{68}^{''}\right)^2 + \left(y_{68} - y_{68}^{''}\right)^2}\right)$$

Fig. 2.The results of detecting face key points by using CNN shown in Fig. 1.

B. Training and classification

- **SUPPORT VECTOR MACHINE (SVM)**: SVM is a discriminative algorithm, which used to develop a supervised model. It used to classify or to predict regression [23]. In this algorithm, we aim to find a separator, which can classify neural, positive, and negative emotion class. This separator is baptized hyperplane. To find a very performant hyperplane, we tested kernel methods. We found that RBF Kernel is the best kernel for our model. We utilized xi, as a vector support, and σ, as positive float, which is, specify a priori, to compute RBF Kernel. Alternatively, RBF $K(x_i, x) = exp\left(\frac{-\|x_i - x\|^2}{2\sigma^2}\right)$. We used also S as a set of vectors support, which contains x_i,b as bias, and Lagrangian coefficient α_i, is a vector support, to compute hyperplane equation:

$$H(x) = \sum_{i \in S} \alpha_i y_i K(x_i, x) + b.$$

- **RANDOM FOREST:** Random forest [25] is an algorithm constructed by combination between trees. This algorithm used decision trees concept to classify. In decision Tree, we calculate Gini index or Gain ratio to construct the tree. In Random Forest, we compute just Gini index to find the most popular attributes founded by Decision Trees [26]. To compute Gini index, we used the probability, which link between class k and selected case $p\left(q/p\right)$. Or, Gini index can be written as follow: $Gini(p) = 1 - \sum_{q=1}^{C} p\left(q/p\right)^2$.

- **K NEAREST NEIGHBORS (KNN):** KNN [27] used k closest neighbors to classify. It can be used for classification or clustering. The concept of this algorithm is finding the class, which have the majority of k closest neighbors. Precisely, we compute distances between examples and we take the examples, which have the smaller distances.

- **DECISION TREE:** Decision tree is a discriminative algorithm, which we used to classify neural, negative and positive satisfaction class. It is a connection between attributes as nodes and tests as branches (i.e. each branch represents just one test of an attribute). Classification in decision tree is computed by cumulating the results of each testing branch. To construct Decision Tree, we can use many algorithms. In this paper, we utilized algorithm called ID3 [24]. The concept of ID3 is choosing the most contributive attribute by computing Entropy and Gain index (i.e. we choose the attribute, which contains the maximal Gain). The equation of Entropy is as follow: $Entropy(E) = \sum_{i=1}^{C} p_i log_2(p_i)$. p_i is the proportion link between E and the class i (Neutral, positive, negative). On the other hand, the equation of Gain is as follow:

$$Gain(E, A) = Entropy(E) - \sum_{q \in vlaue(A)} \frac{|Eq|}{|E|} Entropy(E_q)$$

Where $vlaue(A)$ is all the values of A, E_v is a subset of E, this subset is values of the attribute A.

IV. PERFORMANCE

To evaluate our proposed method, we have to apply our algorithm by using a benchmark dataset. We did not find a dataset correspond to customer's satisfaction. On the other hand, our method is used faces expression as criteria to detected satisfaction of a customer. More specifically, we computed the position changes of facial key point between neutral face and other expression. After that, we applied a classifier to classify between neutral, satisficed, and non-satisfied customer. For this reason, we applied our algorithm by using two facial expression datasets. We categorized the seven expression to three classes. We used neutral faces as a neutral satisfaction, the surprise and joy faces expression as a satisfied customer, and others faces expression as a non-satisfied customer. We applied our algorithm by using the following datasets.

A. Datasets

- **IMPA-FACE3D (Instituto de MatematicaPura e Aplicada Face 3D)**

IMPA-FACE3D is a dataset developed by Instituto de MatematicaPura e Aplicada Face 3D in 2008. They take the six universal expressions proposed by Ekman: happiness, sadness, surprise, anger, disgust and fear. They made images contain geometric information with color (i.e., the geometry and texture are correlated). This dataset contains neutral face sample, and sample corresponding to six facial expressions. They made images of 22 man and16 women. The majority of these persons are aged between 20 and 50 years.

▪ **JAFFE Dataset**

JAFFE dataset is a facial expression dataset generated by Michael Lyons, Miyuki Kamachi, and Jiro Gyoba. They made this dataset at psychology department in Kplyushu University. This dataset contains 213images correspond to 10 female Japanese. The images are categorized to seven classes; neutral and six basic facial expressions.

B. Training process

Our algorithm is divided to two steps. The first step is features selection. We applied a convolution neural network to detect features extraction. More specifically, we extracted key points of an input face. We calculated 68 key points. Since, a vector of 68 elements (Fig. 3) will represent each face. Where each element corresponds to the coordinate of a key point into the face. We applied this method to all images of JAFFE and FACEDB datasets. After that, we computed the variation of each key points detected. More specifically, we calculated the variation of these points between the neutral face and satisfied or non-satisfied face. We baptized these features distances between neutral and satisfied or non-satisfied faces. The second step is applied a classifier to classify faces to three classes; the faces correspond to neutral satisfaction customer, faces correspond to satisfaction customer, and non-satisfaction customer. We used four classifiers to test the performance of our proposed method. We utilized SVM, KNN, Decision Tree, and Random Forest The section2 contains more details about our method.
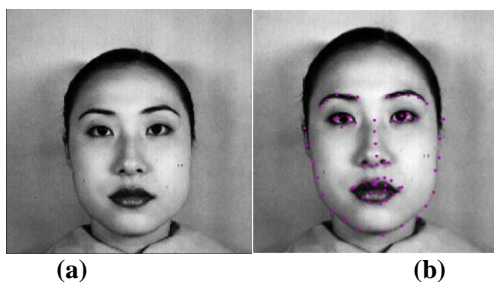

(a)   **(b)**

Fig. 3.The results of detecting key points: the first image (a) is the input image the second (b) is the image contains key points detected.

V. RESULT AND DISCUSSION

Our work contains two steps. The first step is extracting features by applying a convolutional neural network. We detected the key points of an input face. After that, we computed the variation of these key points between neutral and satisfied or non-satisfied face. The second step is the classification step. In this step, we utilized four classifiers. We tested SVM, KNN, Random Forest, and Decision Tree. On the other hand, we computed confusion matrix of each classifier. We had also calculated accuracy, sensitivity, and specificity of each classifier. The following equations present the formula of specificity, sensitivity, and accuracy:

- Specificity is the ratio between non-satisfied and neutral faces, which are not classified as a satisfied face and the number total of non-satisfied and neutral faces.

$$Specificity = \frac{non\text{-}satisfied \& neutral faces correctly classified}{All non\text{-}satisfied \& neutral faces}.$$

- Sensitivity is the ratio between the number of satisfied faces, which are classified correctly, and the number of all satisfied faces (i.e. the satisfied

faces, which are classified correctly, satisfied faces, which are classified neutral faces, and satisfied faces, which are classified as non-satisfied faces).

$$Sensitivity = \frac{satisfied faces correctly classified}{Total of satisfied faces}.$$

- Accuracy is the ratio between the number of satisfied, neutral, and non-satisfied faces are correctly classified and the number of all faces.

$$Accuracy = \frac{Faces are correcty classified}{Total of all faces}.$$

Table- I: The sensitivity, specificity, and accuracy of applying our algorithm by using SVM, KNN, Random Forest, and Decision Tree, and by utilizing JAFFE and FACESDB datasets.

Dataset		Accuracy	Sensitivity	Specificity
FACEDB	SVM(RBF Kernel)	96%	97%	95.50%
	KNN	95%	97%	92.20%
	Decision Tree	96%	97%	95.50%
	Random Forest	92%	95%	91%
JAFFE	SVM(RBF Kernel)	93%	100%	89.50%
	KNN	90%	95%	87.50%
	Decision Tree	93%	100%	95.50%
	Random Forest	92%	97%	92%

We also calculated the ROC curve (Receiver Operating Characteristics curve). On the other hand, ROC curve determinate the probability of distinguishing capability of our algorithm, i.e. how much our algorithm is able to separate between satisfied, neutral, and non-satisfied faces.

Fig. 4 illustrate confusion matrices of our algorithm. This figure contains two plots. The first plot shows the confusion matrix of using FACEDB dataset and SVM with RBF kernel.
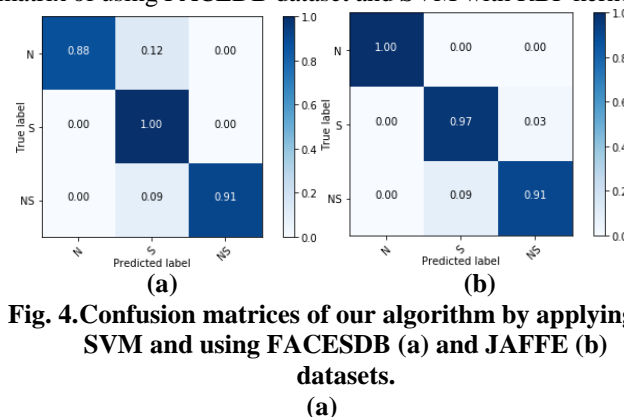

(a)   **(b)**

Fig. 4.Confusion matrices of our algorithm by applying SVM and using FACESDB (a) and JAFFE (b) datasets.
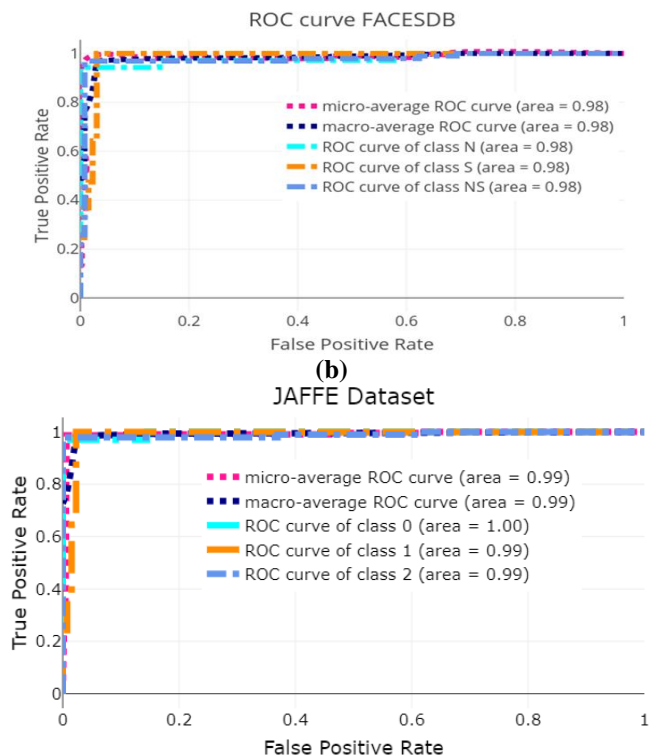
(a)

Fig. 5. ROC curve computed by applying our algorithm using SVM and JAFFE (b) FACEDB (a) dataset.

We found that our algorithm detected 97% of satisfied faces correctly. In addition to that, our method extracted 91% non-satisfied faces correctly. Finally, it detected 100% of neutral faces correctly. The second plot corresponds to applying our algorithm by using Jaffe dataset. We found that, our algorithm detected 100% of satisfied faces correctly. Moreover, it extracted 88% of neutral faces correctly. Finally, it detected 91% non-satisfied faces correctly.

After computing confusion matrices of our algorithm, we computed the sensitivity, specificity, and sensitivity of our algorithm. We found the results showing into table 1. We calculate these statistic criteria for all classifier. We calculate the sensitivity of applying our algorithm by using SVM with RBF kernel, KNN, Decision Tree, and Random Forest. We found that the most performant classifier is SVM with RBF Kernel. Table 1 shows the results of each classifier by using FACESDB and JAFFE datasets.

VI. CONCLUSION

In this paper, we have proposed a new method to recognize customer's satisfaction based of a deep learning model. The main idea is to calculate distances between facial key points into neutral and non-neutral face, and to classify these distances by using Support Vector Machine (SVM), KNN, Random Forest, and Decision Tree. Furthermore, we extract geometrics features that represent distances between neutral face and negative or positive feedback; we used a convolutional neural network to detect facial key points. After that, we computed the distance between each key point into neutral face and negative or positive face. Next, we generated a vector of 68 values that represent customer's face expression. To evaluate the performance of our approach, we tested our algorithm by using FACEDB and JAFFE datasets and we obtained a performants result. The future work will focused our work to develop an adequate customers' shopping behavior dataset. We well also enrich our approach

by developing a multimodal algorithm, which contains speech recognition, eyes gaze, eyes tracking, and motion recognition.

REFERENCES

1. G. O. Young, "Synthetic structure of industrial plastics (Book style with paper title and editor)," in *Plastics*, 2nd ed. vol. 3, J. Peters, Ed. New York: McGraw-Hill, 1964, pp. 15–64.
2. Mehrabian.A, "Communication without words," Psychology today, vol. 2, pp. 53–56, (1968).
3. Wells W.D., Lo Sciuto L.A. Direct Observation of Purchasing.
4. Darwin C., "The expression of the emotions in man and animals", D. Appleton and Company, London, (1913).
5. P.Ekman and W.V.Friesen."Facial Action Coding System" .Consulting Pshychologists Press Inc., 577 College Avenue, PaloAlto,California 94306,(1978).
6. Abigail A Marsh, H.A.E., NaliniAmbay, NonVerbal "Accents: Cultural Differences in Facial Expressions of Emotion," Psychological Science, 2003. 14.
7. Masahiro Nishiyama, H.K., Takatsugu Hirayama, Takashi Matsuyama. "Facial Expression representation based on timing structures in faces," in IEE Int' workshop on Analysis Modeling of faces and Gestures. 2005.
8. Lee, K.K.C., Human expression and Intention via motion analysis: Learning," recognition and ystem implementation, in Graduate School. 2004, The Chinese University of Hong Kong
9. Pantic M, R.L.J.M. "An expert system for multiple emotional classification of facial expressions," in IEEE International Conference on Tools with Artificial Intelligence, 1999.
10. Zhihong Zeng, J.T., Ming Liu, Tong Zhang, Nicholas Rizzoto, Zhenqiu Zhang. Bimodal HCI-related Affect Recognition. in International Conference Multimodal Interfaces. 2004. State College, PA, USA.
11. Zhao, X., Zhang, S., 2016. "A review on facial expression recognition: feature extraction and classification," IETE Tech. Rev. 33, 505–517. https://doi.org/ 10.1080/02564602.2015.1117403
12. Shan, K., Guo, J., You, W., Lu, D., Bie, R., 2017. "Automatic Facial Expression Recognition Based on a Deep Convolutional-Neural-Network Structure," IEEE 15th Int. Conf. Softw. Eng. Res. Manag. Appl. 123–128
13. Rashid, T.A., 2016. "Convolutional neural networks based method for improving facial expression recognition,"Intell. Syst. Technol. Appl. 73–84. https://doi.org/ 10.1007/978-3-319-47952-1.
14. Affdex,https://www.affectiva.com/product/affdex-for-marketresearch/ [Retrieved April 5, 2018]
15. P. Ekman and D. Keltner, "Universal facial expressions of emotion," California Mental Health Research Digest, Vol.8, No.4, pp. 151- 158, 1970
16. Microsoft Azure. Cognitive Services, https://azure.microsoft.com/en-us/services/cognitive-services/ [Retrieved May, 21, 2018].
17. Cui, R., Liu, M., Liu, M., 2016, ''Facial expression recognition based on ensemble of mulitplecNNs,'' Chinese Conf. Biometric Recognit. 511–518. https://doi.org/ 10.1007/978-3-319-46654.
18. Amol S. Patwardhan and Dr. Gerald M. Knapp MIE, "Multimodal Affect Analysis for Product Feedback Assessment," LSU,(SSD 2018), March (2018),Hammamet ,Tunisia.
19. M Slim, R Kachouri, AAtitallah. "Customer satisfaction measuring based on the most significant facial emotion" - 15th IEEE International Multi-Conference on Systems, Signals & Devices,2018,
20. M Wedel, R Pieters - Review of marketing research, 2008.
21. Shergill et al. "Computerized sales assistants: the application of computer technology to measure consumer interest – a conceptual framework," Journal of Electronic Commerce Research, VOL 9, NO 2, 2008
22. Desmet P. M. A. and P. Hekkert, "The basis of product emotions", In W. Green and P. Jordan (Eds.), pleasure with product,beyond usability,pp60-68,London.
23. C. Cortes and V. Vapnik, "Support-Vector Networks," Machine Learning, vol. 20, pp. 273-297, 1995.
24. Quinlan, J.R., "Induction of Decision trees," Machine Learning, 1986.
25. BREIMAN, L., "Random forests-random features," Technical Report 567, Statistics Department, University of California, Berkeley, 1999
26. Raileanu, Laura Elena and Kilian Stoffel, "Theoretical Comparison between the Gini Index and Information Gain Criteria," Annals of Mathematics and Artificial Intelligence 41 (2002): 77-93.

27. Padraig Cunningham, Sarah Jane Delany. "k-NearestNeighbour Classifiers." Technical Report UCD-CSI-2007-4March 27, 2007.

AUTHORS PROFILE

Moulay Smail Bouzakroui graduated a B.S in software engineering from Sciences and technologies Faculty Moulay Ismail University - Errachidia in 1999, The M.Sc. in IT governance in 2014 in University International Rabat. In 2015, He prepares a PhD student in Faculty of science, Ibn Toufail, Kenitra, Morocco. His works research interest: Machine learning, Deep learning, facial expression and emotion recognition applied to customer satisfaction.

Abdelalim SADIQ graduated a B.S in software engineering from Sciences and technologies Faculty Moulay Ismail University - Errachidia in 1999, DESA degree in computer network and telecommunication, in 2002 Ph.D. degree in Computer Science from ENSIAS, Mohammed V University - Rabat, Morocco, in 2007. He is currently an Associate Professor in computer science, Department of Sciences Faculty, Ibn Tofail University - Kenitra, Morocco and Team Leader Information System and Multimedia (SIM). His research interests include multimedia information retrieval and processing, sentiments analysis, IoT and data science.

Abdessamad Youssfi Alaoui was born in Errachidia, Draa-Tafilalet, Morocco in 1992. He is a Ph.D. student at the higher national school of computer science and system analysis (ENSIAS), University of Mohamed V, Rabat, Morocco. In 2014, he got the bachelor in software engineering from faculty of techniques and sciences Errachidia-Morocco. In 2016, he got the Master degree in Business Intelligent and computer vision from Sciences Faculty Fez-Morocco.