Emotion Recognition in E-learning Systems

Oussama El Hammoumi, Fatimaezzahra Benmarrakchi, Nihal Ouherrou, Jamal El Kafi, Ali El Hore
Dept. of Computer Science
Faculty of Science
Chouaib Doukkali University,
El Jadida, Morocco
elhammoumi.oussama@gmail.com, fz.benmarrakchi@gmail.com,
n.ouherrou@gmail.com, jelkafi@gmail.com,
aelhore@gmail.com

Abstract - **Technological advances in e-learning systems offer new opportunities for students to reinforce academic development and to improve accessibility to education. Facial expression recognition is an increasingly important area in intelligent e-learning systems. However, supporting the emotional side of students during learning tasks is challenging and demands an awareness of students' emotions. For this purpose, a brief review is presented in order to underlie the importance role of facial expression recognition in e-learning systems. Its role in special education is investigated too. This paper centers on facial expression recognition using convolutional neural networks and its application in e-learning systems, by introducing a new system composed of three main steps: preprocessing, features extraction and classification. We tested the proposed system with students aged between 8 and 12 years old, in an educational game. The results showed that the proposed system reached state of the art results.**

Keywords— Facial Expression Recognition; Artificial Intelligence; Deep Learning; Convolutional Neural Networks; Education; e-learning system.

I. INTRODUCTION

Emotions have an important role on students learning and achievement. Emotions control the student's attention, effect their motivation to learn and influence their self-regulation of learning. As stated by [1], self-regulated learning and motivation mediate the effects of emotions on academic achievement. Especially, positive emotions positively affect academic achievement when they are mediated by self-regulated learning and motivation.

In the last ten years deep learning algorithms dominated the field of Artificial Intelligence (AI), one of the subfields is computer vision which knew some great successes thanks to Convolutional Neural Networks (CNN) [2], which are one of deep learning architectures that have been widely used for image recognition/classification tasks. Recently, researchers have shown an increased interest in facial expression recognition using CNN [3][4] due to their potential applications such as e-learning [5][6], monitoring and security, Health-care and entertainment.

Although some research has been carried out on facial expression recognition, there have been few empirical investigations into its application in e-learning systems, and there is still lack of adequate strategies to address the presence of emotions in learning. Therefore, the main focus of this work is to propose a novel approach that analyses emotions which could be integrated in e-learning systems.

II. RELATED WORKS

A. Emotion and learning

Recent studies [6][7][8][9], have shown that emotions have important effects on students' learning processes. The available evidence seems to suggest that positive emotions influence learning by affecting students' attention, motivation, and self-regulation of learning. Whereas, negative emotions also influence learning by affecting students' achievement and performance. For example, shame and anxiety decrease interest and motivation. Different emotions may invoke different cognitions unrelated to the tasks, and they may differ with respect to the intensity and persistence of such cognitions. Hence, positive and negative emotions have a very important role in the storage and retrieval of information in learning. As mentioned by [9], the activation of emotions promotes the activation of material that is associative in long term memory which facilitates the retrieval.

Further, research [10] has shown that students experience a great variety of self-referenced, task-related, and social emotions in academic settings, such as enjoyment of learning, hope, pride, admiration, anxiety, shame, hopelessness, and boredom. Furthermore, previous research findings into emotions awareness and learning disabilities have shown that students with learning disabilities have emotional difficulties such as depression and anxiety. In fact, the emotional lives and their difficulties problems are much neglected [11]. Besides, the vast majority of children with learning disabilities have

978-1-5386-6220-5/18/$31.00 ©2018 IEEE

some emotional problem associated with the learning difficulty [12]. Although extensive research [13] [14] [15] [16], has been carried out on students with learning disabilities and e- learning systems, there no studies have been found which investigate the emotional state of the target students.

On the basis of the evidence currently available, it seems fair to suggest that learning environment that takes into account learners' emotions awareness could positively affect students' motivation and performance. Therefore, e-learning systems should integrate emotion recognition systems to support students during learning tasks.

B. Facial expression recognition techniques

In 1971 Ekman and Friesen stated that there are 6 global facial emotions: anger, disgust, fear, happiness, sadness and surprise [17], these emotions contain signals that are filled with information which can be useful to determine the mental state of someone. Each person has a way of showing the facial emotions, and with the same person the expression of an emotion varies also, this makes it a hard problem for machine learning methods. Another problem is that the data usually is not fit for processing, in this case it means that many factors must be taken in consideration such as illumination, head pose, face detection and other problems.

More recent attention has focused on facial expression recognition. One example is Chih-Hung Wu [18], who proposed a facial expression recognition system based on SVM and decision trees that uses Luxund face recognition software-FaceSDK to detect faces from the JAFFE facial expression database. Other example, is Chen et al.[5], who also proposed a system based on SVM but they used Active Shape Models to detect face and location with the VOSM tool and Gabor wavelets to acquire facial appearance information. Or, Shan et al. [19], who presented a method that use LBP for facial representation. The researchers used SVM for the expression recognition with boosted-LBP features. However, the problem with LBP is that it needs more faces added to the training set so that it can detect faces in low light conditions [20].

Further, there exists a variety of research within the areas of the inclusion of emotions awareness in e-learning systems. Among the studies there is a paper presented by [6], which stated that Intelligent Tutorial Systems (ITS) are privileged with affective feedback capabilities, are able to send appropriate affective or cognitive signals to learners, in response to their affective state detection, ensuring in that way their emotional safety and their engagement or persistence in the learning experience. Similar works, an integrated emotion-aware framework for ITS introduced by [7], that elaborates upon two types of features proactive and reactive systems by mapping out their different components, dependencies, and interrelationships in order to capture and structure the rich and creative variety of ways ITS can support positive emotions.

III. OUR PROPOSED METHOD

In this current study, we develop a facial expression recognition system based on CNN, with the aim to integrate it in e-learning system. For this purpose, we used two publicly available databases CK+ [21] [22] and KDEF [23], to carry out the experiments. We added a preprocessing step of images using OpenCV so as to detect and crop faces. The second step is the extraction of features using the proposed convolution neural network and the final step is the classification of the facial images using a fully connected network. The figure 1 illustrates the main three steps of our proposed facial expression recognition system.

Fig. 1. The three steps of the proposed facial expression recognition system

A. Dataset

The model is trained and tested on a combination of images from the CK+ and the KDEF databases. The CK+ database is widely used for facial expression recognition, it includes 327 video sequences acted out by 123 participants aged between 18-50 years old, 69% female, 81% Euro-American, 13% Afro-American, and 6% other groups, each display began and ended with neutral face. The KDEF is a set of totally 4900 pictures of human facial expressions, it contains of 70 subjects 35 females and 35 males aged between 20-30 years old displaying seven different emotional expression. The combined database took only straight head pose images from the KDEF database and all available images from CK+, it was split into 683 training samples, 81 validation samples and 81 test samples as shown in figure 2.

We also use the Japanese Female Facial Expression (JAFFE) Database to test the proposed model, the database contains 213 images of 7 facial expressions (6 basic facial expressions + 1 neutral) acted out by 10 subjects as shown in figure 3.

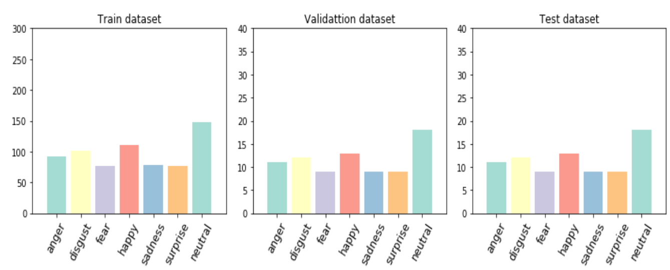

Fig. 2. The custom database (Combination of KDEF and CK+).

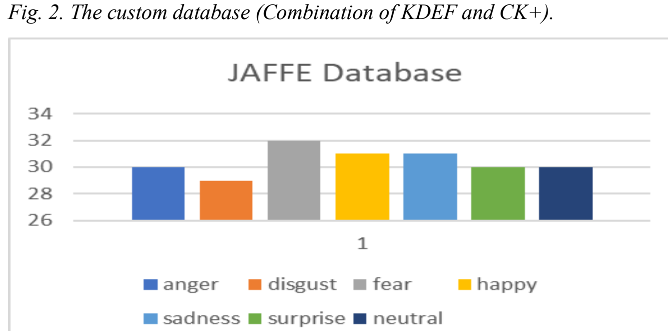

Fig. 3. The Japanese Female Facial Expression (JAFFE) Database.

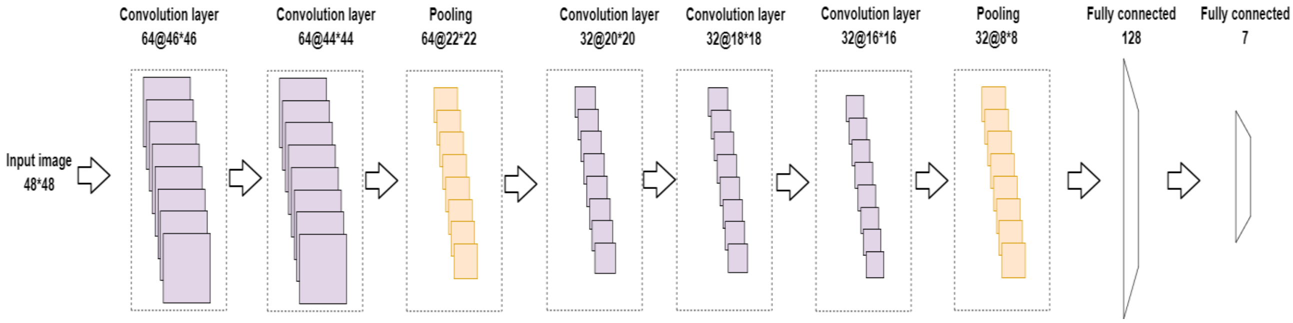

Fig. 4. Structure of the Convolutional Neaural Network (CNN)

B. Preprocessing

This phase consists of four steps:

- First the images are converted to grayscale because color information does not help us identify important edges or other features.
- Face detection using Haar Cascades in OpenCV, it is an effective object detection method proposed by Paul Viola and Michael Jones [24]. This method is based on a machine learning approach where a cascade function is trained on a lot of positive (images of faces) and negative images (images without faces), it is then used to detect faces. There are three categories of haar features, edge features, linear features and four-rectangle features shown in figure 5, each feature is composed of a black rectangle and a white rectangle, by subtracting the sum of all pixels in the white rectangle from the sum of all pixels in the black rectangle we get a single value of the feature, the aim of these features is to indicate the existence of certain characters in the image, finally to distinguish the non-face portion and the face portion.
- Cropping the detected faces.
- Resizing the images to a fixed size of (48,48). Figure 6 demonstrates the preprocessing step.

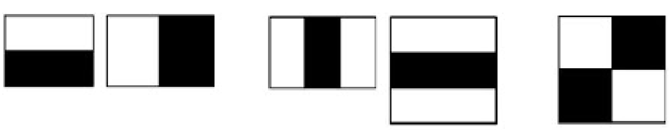

Fig. 5. Haar-like feature templates.

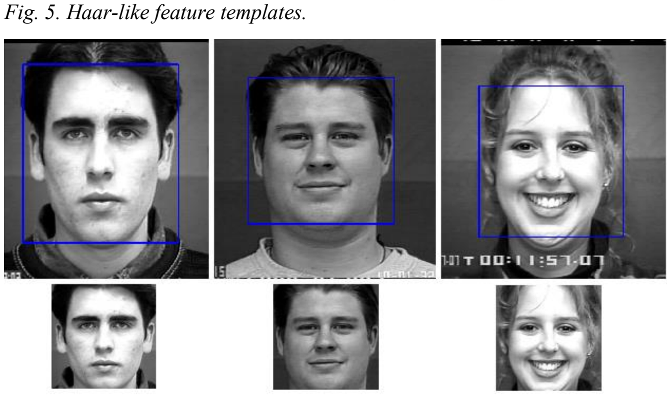

Fig. 6. Preprocessing step demonstration.

C. Feature extraction and classification.

Images from the preprocessing step are fed to the CNN in figure 4. Major advantage of CNN is that it accepts 2D images directly as input data. Therefore, it has an advantage in the field of computer vision. The proposed architecture is composed of five convolution layers, two pooling layers and two fully connected layers.

The convolution layer is used to extract features such as edges, corners, and shapes to generate feature maps. The convolution output is calculated as follows.

$$Output\ width = (W - F_w + 2P)/S_w + 1$$

$$Output\ height = (H - F_h + 2P)/S_h + 1$$

Where (W, H) are the input size, (F_w, F_h) are the filter's size, P is zero padding and (S_w, S_h) are the stride of the convolution.

Pooling layer is a dimension reduction technic used to reduce the size and form a new layer. It accepts an input of $W_1 \times H_1 \times D_1$ and produces $W_2 \times H_2 \times D_2$, where D is the depth. It is calculated as follows.

$$W_2 = (W_1 - F_w)/S_w + 1$$

$$H2 = (H_1 - F_h)/S_h + 1$$

$$D_2 = D_1$$

Hence, these layers are applied multiple times till the system extracts enough features, then the 2D output of the final layer is rasterized into 1D vector that will be fed to the traditional fully connected layer network classifier, at the final layer a Softmax function used for multi-classification.

IV. Experiments and Results

Pre-processing and feature extraction steps were performed using python libraries such as OpenCV and TensorFlow on GPU Nvidia Tesla K80 (12GB), we split the dataset into 90:10

for (training, validation) and testing to evaluate the effectiveness of the system. In this study, we used all emotions from the CK+ database except contempt and we generated the neutral images by taking the first face from each display. All images were preprocessed and fed to the CNN for features extraction, a final step for classification used two fully connected layers for multi-classification. The proposed system has a test accuracy of 97.53% as shown in figure 7.

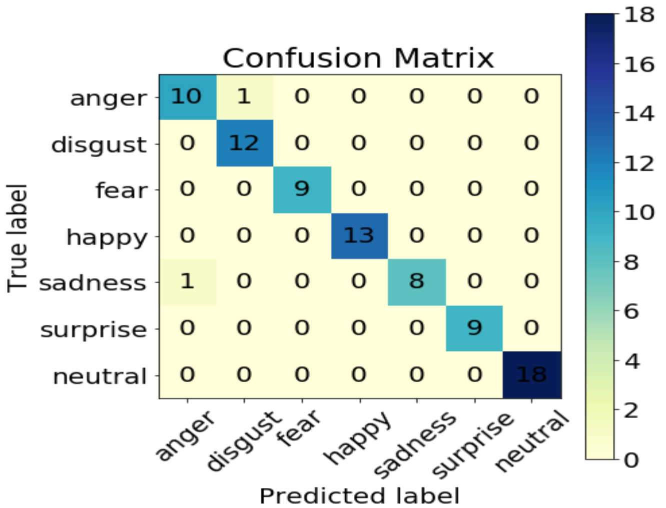

Fig. 7. Confusion matrix for test data of the combined dataset.

To test the proposed system on a new dataset, we used all images in the JAFFE database to test the proposed system and it has an accuracy of 97.18% as shown in figure 8.

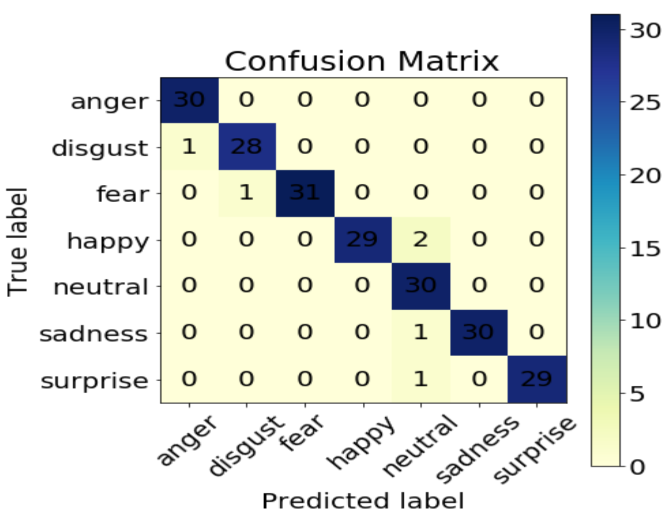

Fig. 8. Confusion matrix for JAFFE dataset.

The system performed well compared to Shan et al.[25] their method reached 76.4742% on the JAFFE database, Mayya et al.[26] reached a recognition accuracy of 98.12% on the JAFFE, they used a SVM network at then end of the system to do the classification.

We tested our proposed system in an educational game, with 4 students aged between 8 and 12 years old (2 participants with learning difficulties), selected from primary school and Speech-Language Pathology Service-Health center, El Jadida Morocco. The choice of our target users was tended to show the important role of emotion recognition in e-learning for students with and without learning difficulties.

To carry out this experiment we integrated our proposed system in an educational game developed by [8] that provides a learning environment which foster learning and help students with their learning difficulties by improving some of their elementary skills, such as reading and writing. The participants were invited to play in the game while the system analyses their emotion in real time, at the end of the game the system saves the results.

We used Plotly, an online framework that provides online solutions for data scientists to make graphs, analytics and statistics. The following figures (9, 10, 11 and 12) show a timeseries of all the results of the system, each one represents emotion probabilities of student during learning.

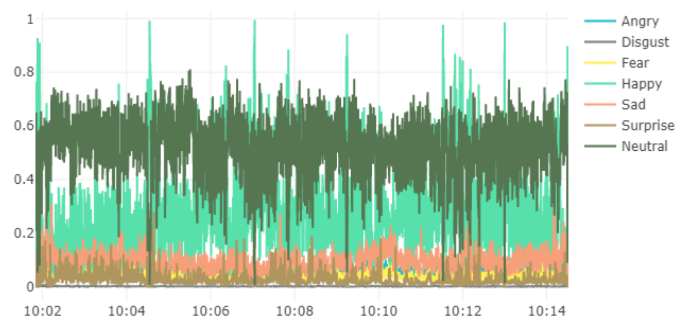

Fig. 9. Emotion probabilities of participant 1.

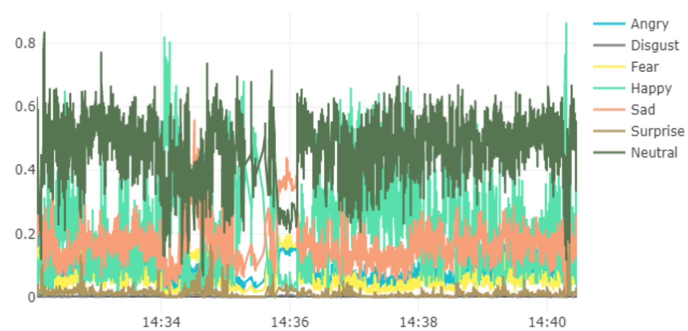

Fig. 10. Emotion probabilities of participant 2.

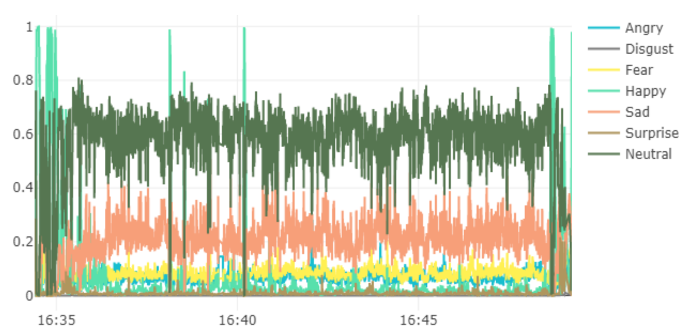

Fig. 11. Emotion probabilities of participant 3 (with learning difficulties)

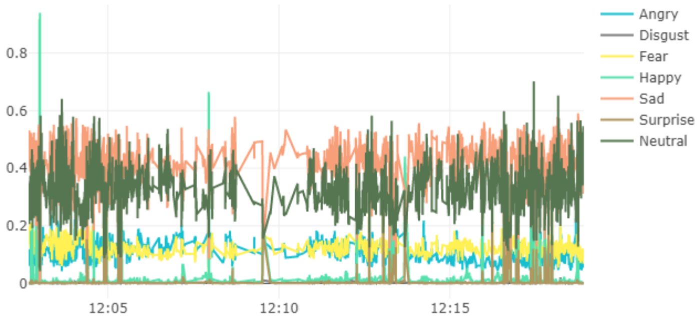

Figure 12: Emotion probabilities of participant 4 (with learning difficulties).

Based on the plots above, we notice that the system detected different emotions during learning, figure 11 and 12 show that at certain times the system did not detect the emotions, that is because the students were not looking directly to the camera. We can notice that the sad emotion in figure 11 and 12 is higher compared to figure 9 and 10 where the happy emotion is higher, probably due to their learning difficulties. Therefore, emotion recognition is important in e-learning systems to help and support students in learning.

V. Conclusion

In this article we proposed a system for facial expression recognition based on CNN, the test was successful and the system was able to detect faces and classify emotions with an accuracy of 97.53% on test data and an accuracy of 97.18% on the JAFFE dataset. And an experiment with students aged between 8 and 12 years old, was conducted to test our proposed system. The outcomes show that emotions were detected and that the system reached state of the art results.

Future studies on the current topic are therefore recommended, especially, to investigate the emotional state of students in e-learning systems.

At this stage, iterative testing with target users is needed to gather feedback on our proposed designs and ensure the resulting design with the goal of identifying the most promising alternatives and best-matching them to e-learning systems.

Acknowledgements

This work was financially supported by an Excellence Grant accorded to Oussama El Hammoumi (11UAE2017) and to Fatimaezzahra Benmarrakchi (2UCD2015) by the National Center of Scientific and Technical Research (CNRST)-Minister of National Education, Higher Education, Staff Training and Scientific Research, Morocco.

The authors would like to acknowledge the president and staff at Speech-Language Pathology Service-Health center, El Jadida Morocco and also the students who have participated in this study.

The authors would like also to thank the speech therapist Ilham ELhousni for her valuable suggestions and recommendations and the stuff at Groupe Scolaire l'Ange Bleu -El Jadida for their cooperation.

References

[1] C. Mega, L. Ronconi, and R. De Beni, "What makes a good student? How emotions, self-regulated learning, and motivation contribute to academic achievement.," J. Educ. Psychol., vol. 106, no. 1, p. 121, 2014.

[2] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in Advances in Neural Information Processing Systems 25, F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, Eds. Curran Associates, Inc., 2012, pp. 1097–1105.

[3] Z. Meng, P. Liu, J. Cai, S. Han, and Y. Tong, "Identity-Aware Convolutional Neural Network for Facial Expression Recognition," in 2017 12th IEEE International Conference on Automatic Face Gesture Recognition (FG 2017), 2017, pp. 558–565.

[4] Z. Yu and C. Zhang, "Image Based Static Facial Expression Recognition with Multiple Deep Network Learning," in Proceedings of the 2015 ACM on International Conference on Multimodal Interaction, New York, NY, USA, 2015, pp. 435–442.

[5] L. Chen, C. Zhou, and L. Shen, "Facial Expression Recognition Based on SVM in E-learning," IERI Procedia, vol. 2, no. Supplement C, pp. 781–787, Jan. 2012.

[6] M. Feidakis, T. Daradoumis, and S. Caballe, "Endowing e-Learning Systems with Emotion Awareness," in 2011 Third International Conference on Intelligent Networking and Collaborative Systems, 2011, pp. 68–75.

[7] J. M. Harley, S. P. Lajoie, C. Frasson, and N. C. Hall, "An Integrated Emotion-Aware Framework for Intelligent Tutoring Systems," in Artificial Intelligence in Education, 2015, pp. 616–619.

[8] F. Benmarrakchi, J. E. Kafi, A. Elhore, and S. Haie, "Exploring the use of the ICT in supporting dyslexic students' preferred learning styles : A preliminary evaluation," Educ. Inf. Technol., pp. 1–19, Oct. 2016.

[9] G. H. Bower, "Mood and Memory American Psychologist 36: 129–148," CrossRef Google Sch., 1981.

[10] R. Pekrun, T. Goetz, W. Titz, and R. P. Perry, "Academic Emotions in Students' Self-Regulated Learning and Achievement: A Program of Qualitative and Quantitative Research," Educ. Psychol., vol. 37, no. 2, pp. 91–105, Jun. 2002.

[11] A. R. Arthur, "The emotional lives of people with learning disability," Br. J. Learn. Disabil., vol. 31, no. 1, pp. 25–30, Mar. 2003.

[12] J. C. Abrams, "On Learning Disabilities: Affective Considerations," J. Read. Writ. Learn. Disabil. Int., vol. 2, no. 3, pp. 189–196, Jan. 1986.

[13] F. Benmarrakchi, J. E. Kafi, and A. Elhore, "Communication Technology for Users with Specific Learning Disabilities," Procedia Comput. Sci., vol. 110, no. Supplement C, pp. 258–265, Jan. 2017.

[14] F. E. Benmarrakchi, J. E. Kafi, and A. Elhore, "User Modeling Approach for Dyslexic Students in Virtual Learning Environments," Int. J. Cloud Appl. Comput. IJCAC, vol. 7, no. 2, pp. 1–9, Apr. 2017.

[15] F. Benmarrakchi, J. E. Kafi, and A. E. Hore, "A Different Learning Way For Pupils With Specific Learning Disabilities," Int. J. Comput. Technol., vol. 14, no. 10, pp. 6157–6162, Oct. 2015.

[16] F. Benmarrakchi, J. E. Kafi, and A. Elhore, "Supporting dyslexic's learning style preferences in adaptive virtual learning environment," in 2016 International Conference on Engineering MIS (ICEMIS), 2016, pp. 1–6.

[17] P. Ekman and W. V. Friesen, "Constants across cultures in the face and emotion.," J. Pers. Soc. Psychol., vol. 17, no. 2, p. 124, 1971.

[18] C. H. Wu, "New technology for developing facial expression recognition in e-learning," in 2016 Portland International Conference on Management of Engineering and Technology (PICMET), 2016, pp. 1719–1722.

[19] C. Shan, S. Gong, and P. W. McOwan, "Facial expression recognition based on Local Binary Patterns: A comprehensive study," Image Vis. Comput., vol. 27, no. 6, pp. 803–816, May 2009.

[20] J. Chang-yeon, Face Detection using LBP features. 2008.

[21] T. Kanade, J. F. Cohn, and Y. Tian, "Comprehensive database for facial expression analysis," in Proceedings Fourth IEEE International

Conference on Automatic Face and Gesture Recognition (Cat. No. PR00580), 2000, pp. 46–53.

[22] P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar, and I. Matthews, “The Extended Cohn-Kanade Dataset (CK+): A complete dataset for action unit and emotion-specified expression,” in 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops, 2010, pp. 94–101.

[23] “Lundqvist, D., Flykt, A., & Öhman, A. (1998). The Karolinska Directed Emotional Faces - KDEF, CD ROM from Department of Clinical Neuroscience, Psychology section, Karolinska Institutet, ISBN 91-630-7164-9.” .

[24] P. Viola and M. Jones, “Rapid object detection using a boosted cascade of simple features,” in Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001, 2001, vol. 1, p. I-511-I-518 vol.1.

[25] K. Shan, J. Guo, W. You, D. Lu, and R. Bie, “Automatic facial expression recognition based on a deep convolutional-neural-network structure,” 2017, pp. 123–128.

[26] V. Mayya, R. M. Pai, and M. M. Manohara Pai, “Automatic Facial Expression Recognition Using DCNN,” Procedia Comput. Sci., vol. 93, pp. 453–461, 2016.