Amina KINANE DAOUADJI [0000-0002-1897-1900]*,
Fatima BENDELLA [0000-0002-7568-5177]*

IMPROVING E-LEARNING BY FACIAL EXPRESSION ANALYSIS

Abstract

Modern technology has become a vital part of our daily lives, and the world has undergone remarkable advancements in various scientific and technological fields. The advancement of technology presents a variety of opportunities for students to promote academic development and make it easier to access education through online learning systems. The most difficult and most demanding task during learning is to be aware of and support the emotional side of students. Recognizing one's emotions is easy for humans, but it is a challenging task for computers due to the specific features of the human face. However, recent advances in computing and image processing have made it possible and easy to detect and categorize emotions in images and videos. This paper focuses on detecting learners' emotions in real time during synchronous learning. In this regard, a video/chat application has been developed for the tutor to detect the emotions of the learners while presenting his lesson. The emotions detected are separated into three states (Satisfied, Neutral and Unsatisfied); each state is made up of two or three distinct emotions. The objective is to assist teachers in adapting teaching methods in virtual learning settings according to the emotions of learners.

1. INTRODUCTION

Technology is a vital component of every field, and one area where it is most prominent is in education (Budhwar, 2017). The authors distinguish the emergence of new education methods based on the use of technology, including the use of the computer and its innovations, satellites, satellite channels, and the international information network, to make learning available to everyone, at any time. The evolution of digital technologies and multimedia have increased the use of E-learning (EL) systems. E-learning has several advantages such as ubiquity, flexibility, and availability. The use of multimedia supports, have made it very popular and widely used. According to Engelbrecht, E-learning is a concept that uses electronic media represented by the Internet, CDs, mobile phones or even television, in order to provide distance learning and teaching (Engelbrecht, 2005), and involves the use of computers and information technologies and systems to construct and design educational experiences (Benadla & Hadji, 2021), in the field of education and its

---

* Université des Sciences et de la Technologie d'Oran Mohamed Boudiaf, Faculté Des Mathématiques Et Informatique, Département d'informatique, Algerie, amina.kinane@univ-usto.dz, fatima.bendella@univ-usto.dz

methods. E-learning is one of the most important technological applications. It represents an educational process, using information and communication technology to create training courses, distribute learning content, communicate between students and teachers and for study management (Harandi, 2015). We can cite two types of E-learning synchronous and asynchronous. Synchronous learning is an online learning activity led by the teacher in real time, in which all students and the teacher can connect simultaneously and have direct communication (Memari, 2020). While asynchronous learning allows learners to respond and study in their free time without being constrained by a specific time.

In the field of education, E-learning assumes that all learners have the same behavior and the same degree of concentration throughout a learning session, without taking into account their varied emotional and psychological adaptation during learning.

The covid-19 pandemic has led to the suspension of several activities and the closure of many industries and establishments, particularly educational establishments (Dhawan, 2020), which has imposed remote working, where videoconferencing software have become essential for teamwork, including teaching and learning. Since it is considered a means of communication allowing teachers as well as students from different locations to establish personal and direct contact with audio and video, as if they were in the traditional classroom. There are many platforms to use for video conferencing such as Google Meet, Zoom and Microsoft Teams. According to SINGH, and AWASTHI almost all applications offer the same functionality, but there are often some exceptions (Singh & Awasthi, 2020). Online learning mainly focuses on the course, because the latter can be modified and added with different media and it is uploaded to a networked environment for online access, so online learning has distinct similarities with classroom learning (Mahanta & Ahmed, 2012).

Acquiring information during an online course, and learning sessions is an essential step for every learner, because in the classroom, the behavior of the tutor, the structure of the course to be taught, the personal motivation, the learning style, the interaction and the animation have a direct influence on the learner's satisfaction (Gray & DiLoreto, 2016).

Distance learning can cause several acquisition problems due to the physical and moral state of the learner. According to a study that was carried out on 20 students, to collect information on the problems encountered in distance learning, all students mentioned different health issues such as back pain, neck pain, headaches and various eye problems (Yücelsin-Taş, s. d.) that will affect the acquisition of information of the learner who will translate it with facial expression. According to Eva-Maria Seidel, emotional faces communicate both an individual's emotional state and behavioral intentions (Seidel et al., 2010). Therefore, in the case of teaching, the tutor will be able to adapt his lesson according to the information he receives about the status of his students through their facial expressions.

Estimating the success percentage is an essential step for any type of system, taking as an example human learning (teaching) is always completed by an evaluation step (exam) to estimate the learner's degree of acquisition and success. While failure in the exam is always related to the evaluation of the learner, neglecting that their concentration during the reception of information, or learning, is in reality the key to acquisition of the information itself. In this case, a very important question to ask is how to avoid the student's lack of concentration during the course to ensure that he or she assimilates the information more specifically?

The proposed solution is to present to the tutor an estimate of the learner's engagement state in real time, based on their facial expressions during the online course session. This will allow the tutor to detect any flaws overlooked during his presentation, and to improve his way of teaching. Deep learning is very important for making effective decisions (Sridharan et al., 2021), so the authors used a convolutional neural network to detect and analyses learner emotions.

2. FACE DETECTION AND ITS CHALLENGES

Facial detection is a computer technology that detects facial features and their location in a digital image or video, ignoring any other object or shape found in the same image, like trees, buildings, cars (Sati et al., 2021), etc. According to (Kumar et al., 2019) face detection from a single image is a difficult task, due to the variation in scale, location, orientation, facial position, facial expressions, occlusions and conditions lighting. All these criteria changes the overall appearance of faces, for example pose variation causes significant problems in face detection, this may be due to a change in viewing angle and also due to rotation in head position (Rizvi et al., 2011). Another criteria like facial expression which directly affects on the appearance of faces, so the human face in an image may have expressions unlike normal, which poses a challenge for face detection, also include facial occlusions where faces can be partially obscured by other objects such as glasses, a scarf, a hand, a mask, etc. We also have the problem of lighting (Illuminations). Part of the image may have very high lighting than others, which affects the appearance of the face, without forgetting the image resolution which plays an important role for face detection.

2.1. Facial expressions

Facial expressions are facial changes that reflect a person's internal emotional states and intentions. According to (Schmidt & Cohn, 2001) people's social intentions and motivations are visualized by the face. Facial expressions play an important role in emotion recognition. Human beings use it to convey different types of meanings in various contexts (Elliott & Jacobs, 2013). They are also used for non-verbal communication (sign language), and represent an indicator of feelings, which allows a person to express an emotional state, that's why they are crucial for both daily emotional communication and automatic emotion recognition systems (Tarnowski et al., 2017; Tian et al., 2011). Facial expressions are generally labeled with these six basic emotions: joy, surprise, sadness, anger, disgust, fear and neutrality.

2.2. Emotions detection

Decision-making, and communication between people, is generally affected by emotions, which represent a natural process based on visual, textual, vocal and other physiological means (Garcia-Garcia et al., 2017; Heredia et al., 2022). The different emotions such as joy, sadness, anger, surprise, etc. are represented by facial expressions (Azcarate et al., 2005). Facial emotion detection is a technology that draws on the field of artificial intelligence. It belongs to the family of technologies called "computing affective", a multidisciplinary field

of research on the abilities of computers to recognize and interpret human emotions and affective states.

The facial emotion recognition system includes three steps:
- – Face detection;
- – Facial expression detection;
- – Classification of expression to an emotional state.

The first step is to detect the facial region from the image. Then the input image is preprocessed to obtain an image which has a normalized size or intensity. The second step is to extract expression features from the facial or image sequence in question. The third and last step, the extracted features are presented to the classifier which will provide the recognized expression (Farkhod et al., 2022).

3. RELATED WORKS

Several works have been carried out on the detection of emotions, like the work of Farkhod and his team (2022) who proposed a method for recognizing feelings, based on graphs. The results obtained showed an accuracy of 91.2/100 in the recognition of the seven emotional classes (joy, sadness, fear, disgust, surprise, anger and neutral), real-time emotion detection showed less results accurate due to noise in expressions. Another work proposed by Hussain and AlBalush is designed to develop a real-time system for detection, recognition and classification of human faces. Classified expressions were represented in seven cases (joy, sadness, fear, disgust, surprise, anger and neutral) (Hussain & Salim Abdallah Al Balushi, 2020). The performance measures were validated using an integrated CNN model with an accuracy of 88/100. The following work (Keshri et al., 2022) aims to identify, perceive and characterize the human face. In this work, facial expressions were treated in seven states. The Viola Jones algorithm was used for the detection of face. The databases, VGG16 and KDEF were used also. Another study conducted by Al-Hazaimeh & Al-Smadi (2019) suggests a deep learning technique for automatic pedestrian traffic based on CNN networks, the latter relies on representing pedestrians in an adaptive manner and obtaining effective recognition with higher accuracy and reduced pre-processing time.

Tab. 1. Comparative table of related works

Title	Years	Methods	Data Base	Accuracy
Automated Pedestrian Recognition Based on Deep Convolutional Neural Networks.	2019	CNN	MIO-TCD	94%
Development of Real-Time Landmark-Based Emotion Recognition CNN for Masked Faces	2022	HaarCascade CNN	FER-2013	91,2 %
A real time face emotion classification and recognition using deep learning model	2020	HaarCascade openCV CNN	VGG 16 KDEF	88%
Automatic Detection and Classification of Human Emotion in Real-Time Scenario	2022	CNN OpenCV	Not mentionned	Not mentionned

4. OUR CONTRIBUTION

Our objective is to present a system that makes it possible to detect learners' emotions in real time such as joy, surprise, sadness, disgust and neutrality; during the online course and sent them to the tutor in the form of a message indicating their engagement status.

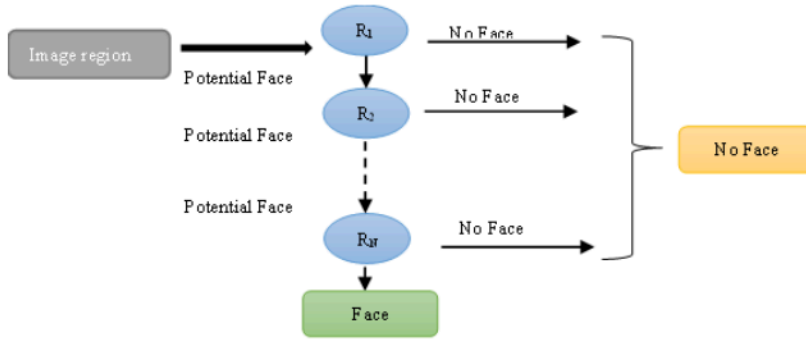In order to achieve our objective, this work was divided into three steps:

1. Capturing student faces based on detection model of viola and Jones face from the Open CV library;
2. Detecting the learner's facial emotions based on convolutional neural networks (CNN) a Deepface of python library, the database used is "fer2013", it contains 35,887 images, resolution 48×48.
3. Making an output pipeline that can send the information back to the tutor.

4.1. Face detection steps

In order to recognize facial emotions, it is necessary to detect only the face of the person in the image/video, because there may be other objects on the image/video, such as cars, buildings, trees, etc., that are considered unnecessary and must eliminated. This step is known as noise removal. The noise can be created at the time of shooting or at the time of image transmission. This work uses the Viola and Jones method algorithm from the OpenCV library. This method is the most widely used and provides satisfactory results for face detection. OpenCV (Open Source Computer Vision Library) is an open source library that has over 2500 libraries for image processing, computer vision and machine learning (Perwej et al., 2022). It mainly focuses on capturing and analyzing videos and image processing, including features such as face detection and object detection. It can be associated with other libraries, such as Numpy, a very efficient library for numerical operations. It also includes the multidimensional table Mat, which is used to store images (Chandrakala et al., 2022).

4.2. Viola and Jones' facial detection method and CNN

The Viola-Jones algorithm is used to detect the human face from an image. The system takes facial images or images without face as input. Face detection includes two phases. The first one is the training phase. In this phase, two types of datasets are included, a positive and a negative image set. Positive images are images composed of faces and negative images are faceless. All features related to face images are collected and stored in a file. In the second phase which represents the test phase, all stored features are applied to an image input and classified, whether face or not. The characteristic of the Viola and Jones algorithm is the existence of a cascade classifier. In this algorithm, there is a classification level to determine whether or not there are facial object features in the selected features (Muhammad et al., 2023). Each sub-image is compared to every feature in every step. If the characteristic value result does not meet the desired criteria, then the result is rejected and the sub-image will move to the next sub-image to perform the same calculations than the previous process.
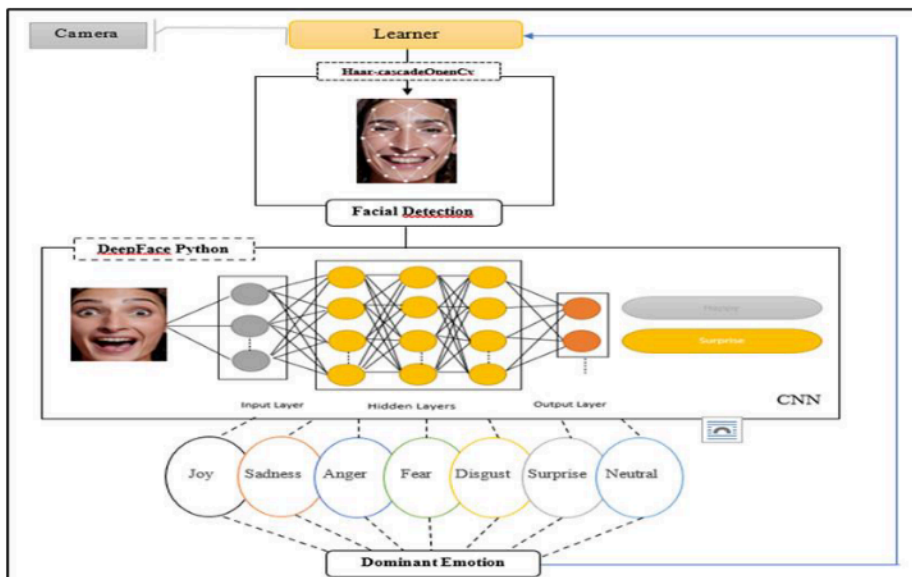
Fig. 1. Cascade classifier

Convolutional neural networks (ConvNets or CNNs) are used for classification and computer vision tasks. It works with three main layers, which are: a) convolutional layer: it is the main building block of a CNN, where the majority of calculations occur; b) pooling layer helps to reduce the dimensionality and number of parameters in the input of the system and c) fully-connected (FC) layer. These layers are placed at the end of the CNN architecture and are fully connected to all output neurons.
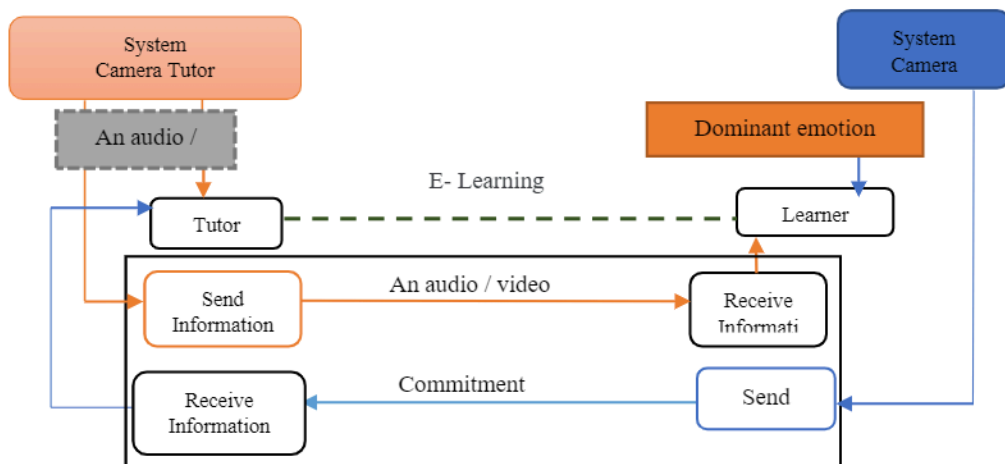
4.3. System architecture

The architecture of the system is divided into two parts; the first concerns the detection of the learner's face and emotions in real time, and the second part concerns sending the engagement status of the learner to the tutor (Figure 2).

Fig. 2. The architecture of the proposed system

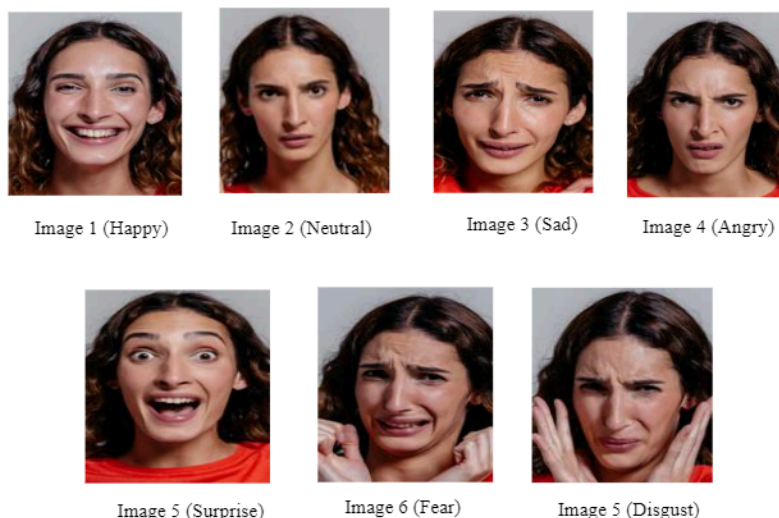The data exchange between the tutor and the learner is represented in the following architecture (Figure 3).

Fig. 3. The process of communication between the Tutor and the Learner

4.4. Test and results

To test the effectiveness of our emotion detection system, two trials were carried out.

4.4.1. Fixed facial expression

In this trial, the authors present some tests carried out on fixed emotions taken from images chosen randomly from the internet (the images used for this trial are chosen from Pexels site), of different emotions and of which we know their emotional state beforehand, we will fix each image towards the camera to visualize the emotion detected by our system.

Image 1 (Happy)   Image 2 (Neutral)   Image 3 (Sad)   Image 4 (Angry)

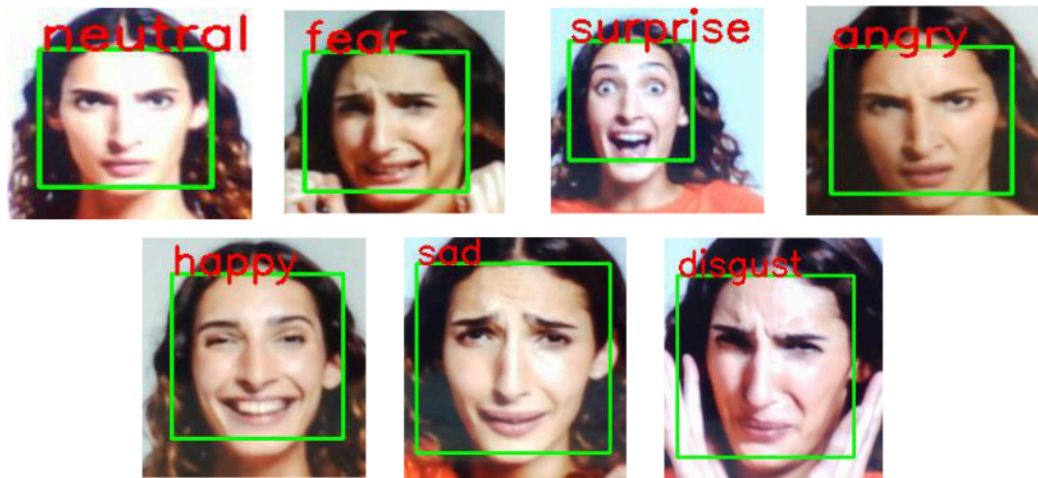Image 5 (Surprise)   Image 6 (Fear)   Image 5 (Disgust)
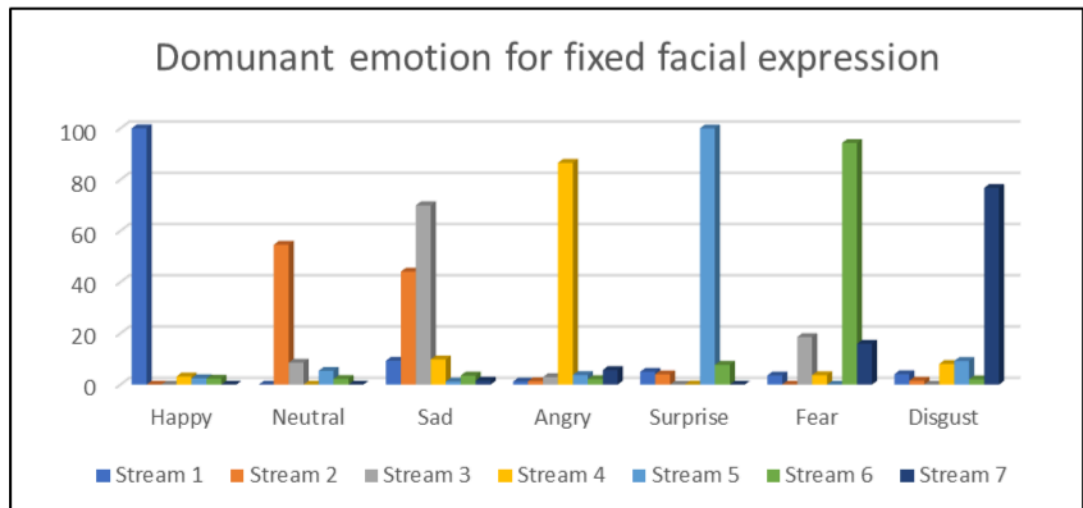
Fig. 4. The seven facial expression images

The results obtained are represented in the following table:

Tab. 2. Representative table of Fixed Facial expression

	Happy	**Neutral**	**Sad**	**Angry**	**Surprise**	**Fear**	**Disgust**
Stream 1	99,99	0	9,31	1,17	4,97	3,55	4,07
Stream 2	0	54,57	44,09	1,29	4,01	0	1,53
Stream 3	0	8,54	69,95	2,91	0	18,55	0
Stream 4	3,15	0	9,79	86,5	0,03	3,66	8,06
Stream 5	2,47	5,37	1,2	3,76	99,98	0	9,22
Stream 6	2,3	2,18	3,52	2,13	7,7	94,28	2,06
Stream 7	0	0	1,55	5,68	0	15,9	76,75

Fig. 5. Fixed facial expression

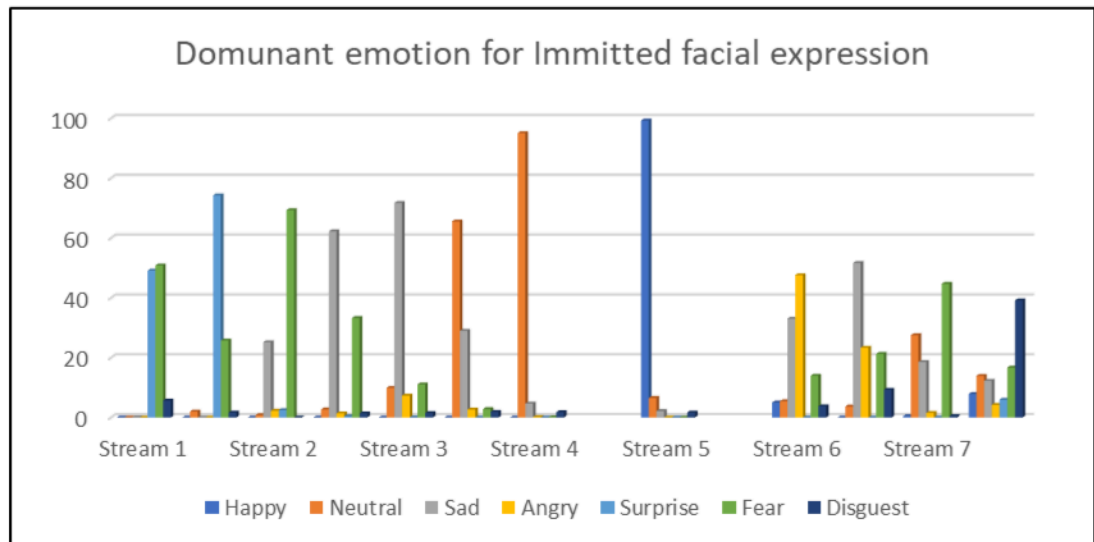Fig. 6. The seven different emotions detected by the system

4.4.2. Immitted facial expression

This trial focused on camera and tried to mimic different expressions of seven different emotions (Happy, Neutral, Sad, Surprise, Fear, Angry, Disgust).
The following table illustrates the different results obtained:

Tab. 3. Representative table of Immitted Emotions

Emotion imitation	Happy	Neutral	Sad	Angry	Surprise	Fear	Disgust
Stream 1 (Surprise)	0	0	0.05	0	49.10	50.83	5.65
	0	1.98	0	0	74.26	25.72	1.63
Stream 2 (Fear)	0	0.79	25.14	2.28	2.46	69.30	0
	0	2.69	62.24	1.38	0.39	33.26	1.40
Stream 3 (Sad)	0	9.91	71.71	7.29	0.021	11.06	1.51
	0	65.54	28.97	2.66	0	2.81	1.87
Stream 4 (Neutral)	0	95.05	4.66	0.16	0	0.11	1.78
Stream 5 (Happy)	99.25	6.56	2.16	0	0	0	1.66
Stream 6 (Angry)	5	5.43	33.02	47.57	0	13.96	3.75
	0	3.67	51.64	23.32	0	21.26	9.25
Stream 7 (Disgust)	0.47	27.52	18.57	1.52	7 .08	44.70	0.48
	7.86	13.91	12.26	4.21	5.93	16.66	39.14

Fig. 7. Immitted Facial expression

4.5. Discussion and proposed system

It was observed from two trials that the system succeeds in detecting different emotions. In the first, the system was tested on images exposed in front of the camera, different facial expressions were fixed, and the system detected emotions well. In the second, the system

was tested by trying to mimic different facial expressions, because it is impossible to look in front of the camera with the same facial expression, so there is always a slight movement, which affects the detection of the dominant emotion, in this case the system can give two different emotions at the same time. Difficulty was encountered in the second test, in which the system was tested by volunteers. Two students from the department tested the work (second test), and the results obtained are shown in the tables (Table. 2, Table. 3).
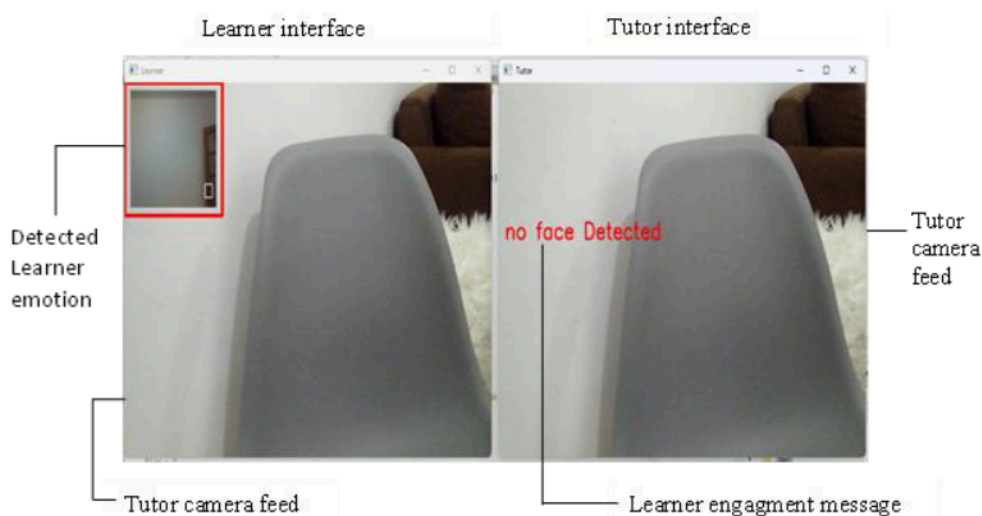
Based on the results of the second trial it was found that the emotional detection system can consider a neutral emotion as sad emotion, and a Disgust emotion as angry emotion, so it was decided to separate the emotions into three different states of engagements.

The following table illustrates the three distinguished states of engagement:

Tab. 3. States of learner engagement

States of engagement	Emotions
Satisfied	Happy, neutral
Neutral	Sad , Angry , Suprise
Unsatisfied	Fear, Disgust

The operation process of the system is as follows:

Fig. 8. System interface

If the learner's camera does not detect any face, the tutor will receive the following message "No face detected".

If the camera detects the learner's face, the emotion will automatically be detected. The dominant emotion will be displayed on the learner's window, and the tutor will receive a message about the learner's engagement state.

The presented system allows the tutor to adapt his teaching method, through the degree of engagement which is initially based on emotional detection, so the tutor can make decisions at the appropriate time. He can for example change the subject, make a break, transmit the information according to the degree of importance, etc.

5. CONNLUSION AND PERSPECTIVES

Online learning is developing in various forms. Assessing interactions between teachers and distance learners has become a challenge for teachers. This work aims to improve communication and engagement between teachers and students during the online course through the use of facial expression recognition technology. A student's learning behavior reflects their learning acquisition, so behavioral analysis through emotion detection is of great interest in distance learning.

The developed system provides teachers with information about the states of engagement of students in real time, which can be used to adapt teaching methods and materials to better meet student needs.

As a perspectives, and to improve their work, the authors plan to take into consideration the following points: Since their system represents a prototype made for the detection of the emotions of a single learner, they plan to add other learners, and improve the tutor interface, to visualize the different emotional states of all connected learners during the course session.

Add another parameter to help the tutor improve his teaching method such as estimating the degree of concentration learners during the course session (in real time).

Add a speech detection and recognition system, to help learners with hearing problems to properly assimilate the tutor's speech.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

REFERENCES

Al-Hazaimeh, O. M., & Al-Smadi, M. (2019). Automated pedestrian recognition based on deep convolutional Neural Networks. *International Journal of Machine Learning and Computing, 9*(5), 662-667. https://doi.org/10.18178/ijmlc.2019.9.5.855

Azcarate, A., Hageloh, F., Sande, K., & Valenti, R. (2005). Automatic facial emotion recognition. *Universiteit van Amsterdam.*

Benadla, D., & Hadji, M. (2021). EFL Students Affective Attitudes towards Distance E-Learning Based on Moodle Platform during the Covid-19the Pandemic : Perspectives from Dr. MoulayTahar University of Saida, Algeria. *Arab World English Journal,* 55-67. https://doi.org/10.31235/osf.io/4xepz

Budhwar, K. (2017). The role of technology in education. *International Journal of Engineering Applied Sciences and Technology, 2*(8), 55-57.

Chandrakala, P., Srinivas, B., & Anil, K. M. (2022). Real time face detection and face recognition using OpenCV and Python. *Journal of Engineering Sciences, 13*(06), 696-706.

Dhawan, S. (2020). Online learning : A Panacea in the time of COVID-19 crisis. *Journal of Educational Technology Systems, 49*(1), 5-22. https://doi.org/10.1177/0047239520934018

Elliott, E. A., & Jacobs, A. M. (2013). Facial expressions, emotions, and sign languages. *Frontiers in Psychology, 4,* 115. https://doi.org/10.3389/fpsyg.2013.00115

Engelbrecht, E. (2005). Adapting to changing expectations : Post-graduate students' experience of an e-learning tax program. *Computers & Education, 45*(2), 217-229. https://doi.org/10.1016/j.compedu.2004.08.001

Farkhod, A., Abdusalomov, A. B., Mukhiddinov, M., & Cho, Y.-I. (2022). Development of real-time landmark-based emotion recognition CNN for masked faces. *Sensors, 22*(22), 8704. https://doi.org/10.3390/s22228704

Garcia-Garcia, J. M., Penichet, V. M. R., & Lozano, M. D. (2017). Emotion detection : A technology review. *Proceedings of the XVIII International Conference on Human Computer Interaction* (pp. 1-8). https://doi.org/10.1145/3123818.3123852

Gray, J. A., & DiLoreto, M. (2016). The effects of student engagement, student satisfaction, and perceived learning in online learning environments. *International Journal of Educational Leadership Preparation*, *11*(1).

Harandi, S. R. (2015). Effects of e-learning on students' motivation. *Procedia - Social and Behavioral Sciences*, *181*, 423-430. https://doi.org/10.1016/j.sbspro.2015.04.905

Heredia, J., Lopes-Silva, E., Cardinale, Y., Diaz-Amado, J., Dongo, I., Graterol, W., & Aguilera, A. (2022). Adaptive multimodal emotion detection architecture for social robots. *IEEE Access*, *10*, 20727-20744. https://doi.org/10.1109/ACCESS.2022.3149214

Hussain, S. A., & Salim Abdallah Al Balushi, A. (2020). A real time face emotion classification and recognition using deep learning model. *Journal of Physics: Conference Series*, *1432*, 012087. https://doi.org/10.1088/1742-6596/1432/1/012087

Keshri, A., Singh, A., Kumar, B., Pratap, D., & Chauhan, A. (2022). Automatic detection and classification of human emotion in real-time scenario. *Journal of IoT in Social, Mobile, Analytics, and Cloud*, *4*(1), 5. https://doi.org/10.36548/jismac.2022.1.005

Kumar, A., Kaur, A., & Kumar, M. (2019). Face detection techniques : A review. *Artificial Intelligence Review*, *52*, 927-948. https://doi.org/10.1007/s10462-018-9650-2

Mahanta, D., & Ahmed, M. (2012). E-Learning objectives, methodologies, tools and its limitation. *International Journal of Innovative Technology and Exploring Engineering*, *2*(1), 46-51.

Memari, M. (2020). Synchronous and asynchronous electronic learning and EFL learners' learning of grammar. *Iranian Journal of Applied Language Studies*, *12*(2), 89-114. https://doi.org/10.22111/ijals.2020.6043

Muhammad, N., Ariyanto, E., & Yudo, Y. (2023). Improved face detection accuracy using Haar cascade classifier method and ESP32-CAM for IoT-based home door security. *Jurnal Ilmiah Penelitian dan Pembelajaran Informatika*, *8(1)*, 154-161. https://doi.org/10.29100/jipi.v8i1.3365

Perwej, Y., Trivedi, A., Tripathi, C., Srivastava, A., & Kulshrestha, N. (2022). Face recognition based automated attendance management system. *International Journal of Scientific Research in Science and Technology*, *9*(1), 261-268. https://doi.org/10.32628/IJSRST229147

Rizvi, Q. M., Agarwal, B. G., & Beg, R. (2011). A Review on face detection methods. *Journal of Management Development and Information Technology*, *11*.

Sati, V., Sánchez, S. M., Shoeibi, N., Arora, A., & Corchado, J. M. (2021). Face detection and recognition, face emotion recognition through NVIDIA Jetson Nano. In P. Novais, G. Vercelli, J. L. Larriba-Pey, F. Herrera, & P. Chamoso (Eds.), *Advances in Intelligent Systems and Computing* (pp. 177-185). Springer International Publishing. https://doi.org/10.1007/978-3-030-58356-9_18

Schmidt, K. L., & Cohn, J. F. (2001). Human facial expressions as adaptations: Evolutionary questions in facial expression research. *American journal of physical anthropology*, *33*, 3-24. https://doi.org/10.1002/ajpa.2001

Seidel, E.-M., Habel, U., Kirschner, M., Gur, R. C., & Derntl, B. (2010). The impact of facial emotional expressions on behavioral tendencies in women and men. *Journal of Experimental Psychology. Human Perception and Performance*, *36*(2), 500-507. https://doi.org/10.1037/a0018169

Singh, R., & Awasthi, S. (2020). Updated comparative analysis on video conferencing platforms - Zoom, Google Meet, Microsoft Teams, WebEx Teams and GoToMeetings. *EasyChair Preprint*, *4026*. https://easychair.org/publications/preprint/Fq7T

Sridharan, M., Arulanandam, D. C. R., Chinnasamy, R. K., Thimmanna, S., & Dhandapani, S. (2021). Recognition of font and tamil letter in images using deep learning. *Applied Computer Science*, *17*(2), 90-99. https://doi.org/10.23743/acs-2021-15

Tarnowski, P., Kołodziej, M., Majkowski, A., & Rak, R. J. (2017). Emotion recognition using facial expressions. *Procedia Computer Science*, *108*, 1175-1184. https://doi.org/10.1016/j.procs.2017.05.025

Tian, Y., Kanade, T., & Cohn, J. F. (2011). Facial expression recognition. In S. Z. Li & A. K. Jain (Eds.), *Handbook of Face Recognition* (pp. 487–519). Springer London. https://doi.org/10.1007/978-0-85729-932-1_19

Yücelsin-Taş, Y. T. (2021). Difficulties encountered by students during distance education in times of confinement in Turkey. *Educational Research and Reviews*, *16*(3), 87-92.