Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Research article

The use of facial expressions in measuring students’ interaction with distance learning environments during the COVID-19 crisis

Waleed Maqableh a,∗, Faisal Y. Alzyoud b, Jamal Zraqou c

a Luminus Technical University College, SAE Amman Institute, Amman, Jordan
b Isra University, Amman, Jordan
c University of Petra, Amman, Jordan

A R T I C L E    I N F O

Article history:
Received 9 March 2022
Received in revised form 6 October 2022
Accepted 9 October 2022
Available online 24 October 2022

Keywords:
E-learning
COVID-19
Face-to-face learning
Facial expressions
Heart pulse

A B S T R A C T

Digital learning is becoming increasingly important in the crisis COVID-19 and is widespread in most countries. The proliferation of smart devices and 5G telecommunications systems are contributing to the development of digital learning systems as an alternative to traditional learning systems. Digital learning includes blended learning, online learning, and personalized learning which mainly depends on the use of new technologies and strategies, so digital learning is widely developed to improve education and combat emerging disasters such as COVID-19 diseases. Despite the tremendous benefits of digital learning, there are many obstacles related to the lack of digitized curriculum and collaboration between teachers and students. Therefore, many attempts have been made to improve the learning outcomes through the following strategies: collaboration, teacher convenience, personalized learning, cost and time savings through professional development, and modeling. In this study, facial expressions and heart rates are used to measure the effectiveness of digital learning systems and the level of learners’ engagement in learning environments. The results showed that the proposed approach outperformed the known related works in terms of learning effectiveness. The results of this research can be used to develop a digital learning environment.

© 2022 The Authors. Published by Elsevier B.V. on behalf of Zhejiang University and Zhejiang University Press Co. Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Education is considered one of the most important factors for the development and economic sustainability of countries. However, good education could contribute to productivity, higher growth rates, good personal health, and other things. The spread of new technologies contribute to and facilitates the development of educational processes, as educators seek to use these technologies to meet the needs of their students. New technologies and the fourth industrial revolution are facilitating the spread of flexible learning which includes distance education, and face-to-face learning in a fixed space, as this technology facilitates teaching and learning (Van Heerden and Goose, 2020). Distance learning is the ability to learn at a remote location. Various terms have been used to describe new learning techniques such as online learning, e-learning, technological learning, media-based learning, collaborative online learning, virtual learning, and web-based learning (Moore et al., 2011). E-learning has become one of the most important sectors of the global economy as it facilitates the development of knowledge and achievement, the reduction of educational costs and access to educational quality. Therefore, various researchers have been conducted to improve e-learning through the use of new technologies, pedagogies, and organizational frameworks (Chai et al., 2014).

The most important features of e-learning are the ability to access learning material anytime and anywhere, especially with the proliferation of fifth generation telecommunications (5G). E-learning is not tied to computer-based systems, and can be used on smart mobile devices, so it can help increase educational productivity by applying group teaching to students (Benta et al., 2015). Despite the advantages that e-learning offers, there are some limitations and drawbacks that limit its spread. One of these challenges is the time overlap with other schedules, such as the conflict between work and family relaxation. Another limitation is the interruption of network and coverage, and some teachers complain of computer fatigue and physical problems, as well as insufficient knowledge of Internet-based techniques. Recently, e-learning is the backbone of education in all countries of the world, especially in the spread of COVID-19 crisis; as it forces the closure of many activities in the world, and face-to-face teaching is one of the areas that are being closed in most countries of the world to reduce the spread of COVID-19 crisis. Educational
institutions such as schools, colleges, and universities are currently being converted to e-learning systems, whereas before the emergence of the COVID-19 crises, they were only based on traditional learning methods as e-learning systems suffer from many challenges related to lack of IT support, security issues such as cyber-attacks on online platforms, and lack of online education strategies.

The next section provides a literature review on e-learning systems and their limitations. Section 4 presents describes the dataset. Section 5 introduces the proposed approach and shows the results of learning based on facial expression classification. The analysis of the results is presented in Section 6. Finally, in Section 7, the conclusion and future work are explained.

2. Literature review

Most countries in the world are responding to the World Health Organization’s (WHO) situation report on the dangerous virus COVID-19, which was first detected in Wuhan, southern China, in November 2019. The education sector at all levels from school education to higher education, is compelled to shift to virtual classes and e-learning, overriding the physical classroom, despite the limitations of e-learning, as face-to-face and traditional learning systems can provide immediate feedback to teachers, trainers, and students on the quality of teaching, delivery and experience (Mukhtar et al., 2020). In face-to-face learning, teachers can respond more effectively to students by understanding their attention through observing facial reactions. This allows them to immediately adjust their teaching methods to meet the needs of their students (Geng, 2011). To combat student carelessness and intention in e-learning education systems, many studies have been conducted based on various techniques; one of which is the use of facial expressions as an indicator of students' concentration.

Facial expression is considered one of the most important components in many applications, such as virtual online reports and video games. Unfortunately, facial animation in computer graphics is still a challenge because hundreds of muscles are responsible for facial expression, which are difficult to simulate with computer-based techniques. Therefore, motion capture alone is insufficient unless it is complemented by manual or automatic editing tools. (Cao et al., 2005). Various researches have been conducted to develop facial expression recognition by using new technologies. One of these researches using the data extracted from Face SDK in JAFFE database in develop a face expression recognition system, in this research, data was implemented and tested with static and random moving images (Wu, 2016). The transfer of learning from face-to-face to e-learning has a reverse impact on education. Practical education, such as pre-clinical medical education, faces many challenges during the pandemic COVID-19, especially practical education in the pre-clinical and clinical years. Various emerging technologies are being used to adapt e-learning for medical education by using artificial intelligence, virtual simulations, and telemedicine techniques, as students would suffer from with the separation of practical laboratories, pathology specimens, and other necessary practical needs (Gaur et al., 2020; Pather et al., 2020). Blended learning in education is considered a good solution for medical pre-clinical education, as the first three years of medical school deal with basic science subjects and help to shape the competence of clinical doctors (Buja, 2019). Online learning suffers from practical and technical support challenges, especially when the COVID-19 pandemic suddenly hits in 2020, forcing teachers to teach their students from home (Hodges et al., 2020). During the COVID-19 pandemic, teachers used various technologies such as Moodle, which facilitates collaboration of lecture notes, videos, online quizzes, sheets, and others. MS-Teams allows synchronous interaction with students, and can be integrated with MS office tools. Zoom is also used as an application for synchronous video conferencing, as Zoom supports live tutorials with additional features such as screen sharing and chatting. Panopto and MS-PowerPoint can be merged to deliver video lectures after these videos have been split into an appropriate size, so that they can be delivered over the network with acceptable quality of service (Welsen et al., 2020).

E-learning is superior to face-to-face learning in some aspects that are related to information accessibility, the flexibility of time and space; but it does not guarantee the interactivity of face-to-face learning, so hybrid information systems are proposed to combine computer vision combined with machine learning technologies to enhance the interactivity of e-learning systems (Ayvaz et al., 2017), so the proposed system use some detected emotional states based on facial expression for the learners, then these are used by the educator to enhance the interactivity through implementing several classification algorithms. KNN and SVM algorithms are used to determine the best accuracy.

To understand the engagement of students in face-to-face and e-learning virtual environments, experts classify engagement into three categories: automatic, semi-automatic, and manual. Facial expression is used as an indicator in automatic engagement; since facial expression is a promising factor in developing an e-learning virtual environment (Dewan et al., 2019).

The deep learning model is also used to enhance the engagement of students in e-learning through implementing pre-training data based on facial expression and then a model’s weights are used to initialize the base for the deep learning model using histograms of the oriented gradient to support vector machines (Mohamad Nezami et al., 2020). Unfortunately, there is a lack of a dataset to be used for training systems in developing e-learning interactivity, most of the tested data are based on the physical behaviors of the face emotional states which are combined with students’ comprehension, and most of them are based on quantitative observation by recording the behaviors of students (Sathik and Jonathan, 2013). It was noted that facial expression of the students is the main indicator for communication mode which enable the teachers to recognize the interaction of students with their lectures in the virtual classroom, also most of the research is biased toward demographic variables in determining the level of engagement, these demographic variables are related to students’ ages and geographic locations (Vail et al., 2016). It is clear from the literature review that the important issues for e-learning to be effective are: cost, privacy by having the proper security permissions, technologies requirement, and the availability of network communication with sufficient bandwidth. E-learning should be effective and interesting to teachers and students by utilizing the new technologies and enhancing the interactivity between them (Kirschner, 2015).

Extracting features dynamically is a prominent feature that can be used to analyze faces. Several approaches have been developed based on facial action units (AUs) for recognizing expressions, this can be achieved based on extracting dynamic features. Building dynamic characteristics for different facial poses has prominent future research works. Some researchers focus on recognizing prototypic expressions such as sadness, anger, and/or happiness which occur rarely. Hence, researching to further the study to include human emotions and interactions is important to increase the reliability. The facial features are divided into two main categories, the first one is permanent such as eye, brow, and mouth and the second is transient such as wrinkles and furrows. In Tian et al. (2000b), multi-state facial models were used to create several facial features. This led to tracking the facial features in the permanent and transient categories. Also,
Table 1
A comparison between face-to-face and e-learning aspects Hodges et al. (2020).

| Category          | E-learning                                                                 | Face-to-Face learning                                                                 |
|-------------------|----------------------------------------------------------------------------|---------------------------------------------------------------------------------------|
| Learning aspects  | It is based on media, reading, videos, and exercises                       | It is based on direct interactions, discussions, and presentations                    |
| Diversity         | Limited opportunities for communication, and sometimes teachers do not recognize who moves online | Different methods for communication                                                   |
| Model             | Student-center course approach                                             | Teacher-center course approach                                                       |
| Space and Presence| Different space                                                             | Same place                                                                             |

3. E-learning challenges and solutions

When teaching is done remotely, e-learning is an alternative solution to face-to-face learning. Teachers and students use the technologies of the online environment to interact with each other, as shown in Fig. 1.

E-learning differs from face-to-face learning in many aspects related to the learning aspects, diversity of communication, and interaction models as explained in Table 1.

E-learning could be effective and successful if the following factors were adapted: Developing accessibility and autonomy between students and teachers, and developing the content of study materials to improve interaction between all study components. Collaboration can be modified through questions and assessments to extend the sociality of learning; because learning is based on learning from others (Harris, 2012). Teachers can develop e-learning by using their experience in learning effectively planning courses, preparing learners, and enhancing interaction methods in learning by using rich media and synchronous interaction as explained in Fig. 2.

Assessment is considered a complex challenge in e-learning systems as cheating by students can occur (Munoz and Mackay, 2019). Online assessments are divided into synchronous assessments which include quizzes, questions, and answers. Asynchronous assessment focuses on homework, projects and case studies. Continuous assessment models should be emphasized in e-learning. In line with the expectation of self-reflection assessments, continuous e-assessment can be developed to improve student learning outcomes (Shen et al., 2013).

Blended learning and hybrid learning are considered good solutions for learning in universities, because blended learning is a mixture of face-to-face learning and e-learning. It is therefore strongly recommended to be used in universities, especially; since synchronous face-to-face learning between students and teachers is necessary in practical applications and laboratories (Oźadowicz, 2020).

Affective reasoning is used to unlock significant potential for interactive e-learning by effectively analyzing physiological responses, i.e. facial expression and heart pulse in the interactive learning environments enabling teachers to create a customized expression index that can be used to test changing levels of engagement, interest and emotional state. The next section presents the proposed solution for studying physiological response by testing facial expressions and observing learners behavior in interactive learning environments.

4. Dataset

A benchmark dataset called CMU Multi-PIE Face Database was used to evaluate the proposed method. It contains about 750,000
images of 337 people taken over the course of five months. Different facial expressions were considered in each captured image, as shown in Table 2.

Various faces and poses with different facial expressions were also recorded for our participants. The focus was to analyze the students’ behavior during e-learning based on their facial expressions. The total time of the experiments was set at 60 min, divided into three sessions, 15 min for each session and 5 min for a short quiz. Then, 15 students from Luminus University University/Multimedia Department were randomly selected to participate in this study. After conducting the experiment, more than 150 snapshots were taken of the participants’ faces, some of which were sorted out due to their poor quality, while the others were selected and classified into different categories that benefit from the new trends in facial expression recognition (Wu et al., 2012).

The following steps took place to work through the experimental scenarios:

1. The presentation of the online course to the participants.
2. The instruction of the participants in the use of the system.

Students will learn the entire subject in three sessions, with each session focusing on one topic from the chosen subject. The face tracker will track the faces of the participants and record any differences that may occur as their facial features change. At the end of each session, a short multiple-choice quiz is thrown on the computer screen. At the end of the experiment, the participants’ facial expressions are extracted from the video and then distributed over the three experimental sessions as shown in Table 3.

The study will also analyze the participant’s heart rates using a “TTL SA9308M Heart Rate/BVP” heart-pulse sensor, as shown in Fig. 3. The heart rates are correlated with the recorded facial expressions during the experiments, as it was found that many studies rely on this factor to measure student engagement in the classroom, such as Maier Kj (Maier et al., 2003), and Cranford KN (Cranford et al., 2014). In addition, Peled et al. (Peled et al., 2008) find that blood pressure correlates with an increase in emotional stress. In this way, we can increase the reliability of our interpretation of the effects of facial expressions and measure students’ level of engagement with the topic being learned.

The most important parts of the face that can be analyzed are the analysis factors are (eye direction, eyebrows, lip angle, head position and hand movements) as shown in Tables 4–8. Furthermore, the research presented in Al-Helali et al. (2021), demonstrates a novel technique of image fusion by integrating images with multi-focal points. Two source images are decomposed using the “Discrete Multi-Wavelet Transform” and the “Fast Fourier Transform”, then processed using three different fusion rules such as the gradient rule, the absolute maximum selection rule, and maximum selection rule to merge the coefficients of the low and high frequency sub-bands.

Facial expressions were recorded and divided into three intervals based on the progress time of the experiment, such as:

- Session1: Energy times, from [0–15] min.
- Session2: Distracting times, from [16–30] min.
- Session3: Boring times, from [31–45] min.

Fig. 4, shows that in the first few minutes of the experiments it was found that the participants’ faces smiled a little, the faces were straight and no unusual expressions were seen, and when the hands touched the face, it was most likely closed, and the pupil of the eye was mostly in the center of the eye.

Based on the conducted experiments, the results have shown that all the above observations made in the first moments indicate a more focused expression of the participants, including
Table 4
Variations of eye expressions.

| Eye variations          | Sample of the variations |
|-------------------------|--------------------------|
| Left & Left-Up          | ![Sample Images]          |
| Right & Right-Up        | ![Sample Images]          |
| Up & Center-Up          | ![Sample Images]          |
| Center-Down             | ![Sample Images]          |
| Center                  | ![Sample Images]          |

participation and engagement in the learning experience, than when made later after 15 min.

Fig. 5 shows that the directions of the head are not always straight and the eye movements are not stable. When the hands touched the face, they were most likely differently shaped, and most of the head movements were on the right side. These observations could indicate that the participants lose their concentration during the lecture. It is also affecting the participants’ participation and engagement in the learning experience.

Fig. 6 shows that unstable movements of the head with the eye are clearly shifted between the edges of the eyes, and the participant sometimes puts his fingers in his mouth and makes head movements of 180 degrees. These observations could indicate that the participants have lost their participation and engagement in the lesson and are bored.

5. Proposed method

Facial expression is the significant change in the face in response to a person’s internal emotional states, intentions, or social communication. Facial expression analysis is used by different researchers on different topics to develop computer systems that enable these systems to understand and use this natural form of human communication for system development.

In this research, facial expression analysis is emphasized to create an expression index, and this expression is used to analyze the physiological responses of students in the online learning environment to develop e-learning outcomes, as explained in Fig. 7.

Facial expression can be understood by developing the three main parts of facial analysis: the first part is face detection, the second part is feature recognition, and finally facial expression and emotion classification (Ozdowicz, 2020). Facial expressions can be categorized as happy, sad, anger, neutral, disgust, and surprised (Rizwan et al., 2020). The work presented in Zraqou et al. (2014) is used to automate the process of facial expression recognition. The presented technique works with real-time videos to detect the emotions in each image. The method can recognize 5 important features namely (eyebrows, eyes, nose, mouth and chin) as shown in Fig. 9. The structural algorithm for applying facial expression is performed as shown in Fig. 10. The efficiency was achieved by tracking the points of each feature instead of running the extraction method for each image as shown in Fig. 8.

Fifteen students were randomly selected from a computer course to participate in this study. After the experiment was conducted, more than 150 snapshots were taken of the participants’ faces. Some of them were sorted out, the others were selected and classified into different categories using the Euclidean distance transformation algorithm (Fabbri et al., 2008). The most important parts of the face that can be analyzed or have a relation to
Table 5
Variations of eyebrows expressions.

| Eyebrows variations | Sample of the variations |
|---------------------|--------------------------|
| **Up**              | ![Images of eyebrow variations](image1) |
| **Center**          | ![Images of eyebrow variations](image2) |
| **Down**            | ![Images of eyebrow variations](image3) |

Table 6
Variations of mouth expressions.

| Mouth variations | Description | Mouth/lips |
|------------------|-------------|------------|
| **Closed**       | ![Images of mouth variations](image4) |
| **Hand with mouth** | ![Images of mouth variations](image5) |
| **Open mouth**   | ![Images of mouth variations](image6) |
Table 7
Variations of head expressions.

| Head variations | Description | Head pose |
|-----------------|-------------|-----------|
| Tilt to right-side | Tilt to left-side | Head in center position |
| Move up and down | Nothing has been observed | |

the factors of analysis are listed below. All expressions and their variations are listed in Tables 4-8.

- Direction of the eyes
- Eyebrows
- Angle of the lips
- Position of the head
- Movements of the hands

5.1. Data set validation

The first phase of the proposed method that was tested is face recognition. The EmotionNet database was used to retrieve and create a set of faces for different poses and people. The set includes 250 images to conduct the experiment, as described in Table 9. The average of correct recognition is 86%. The discrepancies were avoided by ignoring all results that yielded a threshold of less than 45%. This threshold is based on the confidence value of the detection. The value of 45% was the best value that distinguishes between the correct and incorrect matches.

6. Results and analysis

6.1. Participant test analysis

In order to determine the level of engagement and involvement depending on the facial expression categories explained earlier, participants were asked to answer several questions. The students’ responses are categorized into the following factors: (Timely response, response over time, correct responses and incorrect responses). Table 10 shows the distribution of the participants’ responses and their distribution according to the four factors, while Table 11 shows the percentages of the students’ average responses in each of the experimental phases.

The results in Table 11 show that the average response rate of the students started to decrease in relation to the first factor, as the rate of responses was highest in the first session at (60%), then decreased to (47%) in the second session, and the lowest rate was reached in the third session at (23%) (Fig. 10 shows this decrease in the participants’ responses). For the second factor (responses exceed time), the average student response rate for this factor is
Table 8
Variations of head movement.

| Hand movements   | Description          | Head pose       |
|------------------|----------------------|-----------------|
| Moving hand      |                      |                 |
| Head reset on hand |                    |                 |

Fig. 4. Facial expressions captured between (0–15) minutes’ interval.

Fig. 5. Face expression during [16–30] min intervals.
Table 9
The accuracy of the facial expression method.

| # | Face pose | EmotionNet database | #Correct matches | Accuracy |
|---|-----------|---------------------|------------------|----------|
| 1 | Happy     | 50                  | 43               | 86%      |
| 2 | Sad       | 50                  | 41               | 82%      |
| 3 | Anger     | 50                  | 45               | 90%      |
| 4 | Neutral   | 50                  | 46               | 92%      |
| 5 | Disgust   | 50                  | 39               | 78%      |
| **Average** | **50** | **42.8** | **86%** | |

The discrepancy in rates obtained by the four factors can be explained by the following:

1. Participants lost their commitment to the lesson over time.
2. Participants lost their connection to the lesson over time.
3. Participants could not recall information discussed earlier in the lesson.
4. Participants could no longer engage with the lesson because they did not understand it.

During class time, more than 150 images of facial expressions were taken. Some were eliminated due to their low quality or they were found duplicated. The remaining have been accepted and classified as shown earlier. In the next sections, the research will discuss the facial expressions observed during the three sessions of the experiment.

### 6.2. Eyes expressions recorded at the experiment

During the different experimental sessions, it was observed that the eyes have different variations as the eyes were observed in different positions and there was a clear movement of the eyes. Also, there are a number of variations that were observed more in the first session and less as the session progressed.

For example, in the first session of the experiment, the eyes were observed to move in three main directions (left, center, and right). In this session, the eyes in the center position scored the highest, 80%, compared to the other variations, but in the second and third sessions, the percentage of this variation begins to score less, 20% in the second session and 6.67% in the third session.

Other variations were also observed, such as the variation in the center of dawn, which is 0% in the first session and 53.33% in the last session, suggesting that this variation may be related to the level of engagement of the participants. Variations like (Left, Right and Top) were observed to achieve some observations as well, these variations achieved between 67% and 13% during the experimental sessions. Other variations such as (Bottom Left and Bottom Right) were not observed during the experiment. This suggests that these variations do not have a strong relationship with the level of engagement of the participants. Table 12 shows the results of the eye expressions in the three sessions of the experiment.

### 6.3. Eyebrows expressions recorded at experiment

In the experiment, three types of eyebrow variations were observed (bottom, middle, and top), as shown in Table 15. In the first session, the highest percentage for eyebrow expressions was obtained by the eyebrows in "middle position" with a percentage of 53.33%, and there was only one observation for the eyebrow in "high position" with a percentage of 6.67%. In the second session of the experiment, the "middle position" expression decreased to 26.67%, and the "high position" expression increased to 26.67%. In the last session of the experiment, these two expressions were

---

Table 10
Participant answers results.

| Q-No | Questions                                      | Answers on time | Answers exceed the time | Correct answers | Wrong answers |
|------|------------------------------------------------|-----------------|-------------------------|-----------------|---------------|
| 1    | List the types of programming languages.      | 10              | 5                       | 12              | 3             |
| 2    | What is the main difference between compiling and scripting? | 8               | 7                       | 9               | 6             |
| 3    | What are the main characteristics of good programmers? | 7               | 8                       | 6               | 9             |
| 4    | What the programmers can do?                  | 4               | 11                      | 8               | 7             |
| 5    | The term API stands on?                       | 3               | 12                      | 5               | 10            |
observed with the same percentage of the previous occurrence, but with a small decrease for the expression “upper position”, which reached 13.33%. Based on these results, the researchers believe that these two expressions of the eyebrows could also be related to the level of engagement of the participants.

6.4. Mouth expressions recorded at the experiment

During the experiment, it was observed that the mouths of the participants showed different variations as the mouths were observed in different positions such as (turned left, no movement, turned right, and open mouth). As shown in Table 14.

---

### Table 11
Participants’ questions answering percentages.

| Q-No. | Answers on Time | Answers exceed the Time | Correct answers | Wrong answers |
|-------|------------------|-------------------------|-----------------|--------------|
| 1     | 67%              | 33%                     | 80%             | 20%          |
| 2     | 53%              | 47%                     | 60%             | 40%          |
| Average | 60%           | 40%                     | 70%             | 30%          |
| 3     | 47%              | 53%                     | 40%             | 60%          |
| Average | 47%           | 53%                     | 40%             | 60%          |
| 4     | 27%              | 73%                     | 53%             | 47%          |
| 5     | 20%              | 80%                     | 33%             | 67%          |
| Average | 23.5%          | 76.5%                   | 43%             | 57%          |

### Table 12
Eye expression results during different interval times.

| Eye expressions | Center-Up | Left-Up | Right-Up | Center | Down center | Down left | Down right | Not applicable |
|-----------------|-----------|---------|----------|--------|-------------|-----------|------------|----------------|
| Experiment time (0–15) minutes | | | | | | | | |
| Observed | 2 | 1 | 0 | 12 | 0 | 0 | 0 | 0 |
| Percentage    | 13.33% | 6.67% | 0 | 80% | 0 | 0 | 0 | 0 |
| Experiment time (16–30) minutes | | | | | | | | |
| Observed | 1 | 2 | 1 | 3 | 8 | 0 | 0 | 0 |
| Percentage    | 6.67% | 13.33% | 0 | 20% | 53.33% | 0 | 0 | 0 |
| Experiment time (31–45) minutes | | | | | | | | |
| Observed | 2 | 1 | 3 | 1 | 8 | 0 | 0 | 0 |
| Percentage    | 13.33% | 6.67% | 20% | 6.67% | 53.33% | 0 | 0 | 0 |

---

Fig. 7. Architecture of the proposed e-learning using facial expression.

Fig. 8. The detected vital features of the face.

Fig. 9. The structure of facial expression detection.
In the first session, the closed mouth variation was the highest with a percentage of 60%, and no other variations were observed. In the second and third sessions, the score for closed mouth expression decreased to 13.3% and 26.6%, respectively, and new mouth variations were observed, i.e., (turned to the left with a percentage of 26.6% in the second session, and turned to the right with a percentage of 13.3% in the third session). Based on these results, the researchers assume that the expression “closed mouth” is positively related to the level of engagement of the participants, since it reached a high value in the first session. During the experiment sessions, it was observed that the participants' mouths have different variations since the mouths were observed to be in different positions such as (Twisted left, no move, twisted right, and open mouth). As listed in Table 14.

In the first session, the variation of the closed mouth scores the highest with a percentage of 60%, and no other variations were observed, while in the second and third sessions the score of the closed mouth expression was decreased to score 13.3%, 26.6% respectively, and new mouth variations were observed i.e. (twisted left with a percentage of 26.6% in the second session, and twisted right with a percentage of 13.3% in the third session). From these results, the researchers believe that the closed mouth expression may have a positive relation to the degree of participants' engagement since its score high in the first session.

6.5. Hands expressions recorded at the experiment

During the analysis of the experiment results, it was not expected to see any hand movements that might have a relationship with the degree of participants engagement, as the results of the first session shows that (no hand movements) expression scores the highest percentage by (60%), and “Closed hand” by (33.33%), as shown in Table 13.

But, in the second sessions, the hand expressions began to be observed, for example, the “Fingers and mouth” expression were observed by (26.67%), and the “Closed hand” by (13.33%). In the third session, a new hand expression “Headrest on a hand” was observed with a percentage of (26.67%), and “Fingers and mouth” score a percentage of (20%).

From these results, it cannot be considered that the hand expression is a form of the facial expressions, and here it is difficult to interpret hand expression as a measurement of the participants’ engagement, but it was observed through the results that the hand expressions were associated with the head expressions and mouth expressions, and this is clearly appears in the second and third sessions.
Table 14
Mouth expression results during different interval times.

| Mouth expressions | Twisted left | Closed | Twisted right | Not applicable |
|-------------------|-------------|--------|---------------|---------------|
| **Experiment time (0–15) minutes** |             |        |               |               |
| Observed          | 0           | 9      | 0             | 6             |
| Percentage        | 0           | 60%    | 0             | 40%           |
| **Experiment time (16–30) minutes** |             |        |               |               |
| Observed          | 4           | 2      | 0             | 9             |
| Percentage        | 26.67%      | 13.33% | 0             | 60%           |
| **Experiment time (31–45) minutes** |             |        |               |               |
| Observed          | 0           | 4      | 2             | 9             |
| Percentage        | 0           | 26.67% | 13.33%        | 60%           |

Table 15
Hand expressions results during different interval time.

| Hands expressions | Closed Hand | Headrest on a hand | Fingers & mouths | No hand movements | Not applicable |
|-------------------|-------------|--------------------|------------------|-------------------|---------------|
| **Experiment time (0–15) minutes** |             |                    |                  |                   |               |
| Observed          | 5           | 0                  | 1                | 9                 | 0             |
| Percentage        | 33.33%      | 0                  | 6.67%            | 60%               | 0             |
| **Experiment time (16–30) minutes** |             |                    |                  |                   |               |
| Observed          | 2           | 0                  | 4                | 3                 | 6             |
| Percentage        | 13.33%      | 0                  | 26.67%           | 20%               | 40%           |
| **Experiment time (31–45) minutes** |             |                    |                  |                   |               |
| Observed          | 1           | 2                  | 3                | 0                 | 9             |
| Percentage        | 6.67%       | 13.33%             | 20%              | 0                 | 60%           |

Table 16
Head expression results during different interval times.

| Head expressions | Tilt-Right | Tilt-Left | Center | Up  | Down | Not applicable |
|------------------|------------|-----------|--------|-----|------|---------------|
| **Experiment time (0–15) minutes** |             |           |        |     |      |               |
| Observed          | 0          | 0         | 15     | 0   | 0    | 0             |
| Percentage        | 0          | 0         | 100%   | 0   | 0    | 0             |
| **Experiment time (16–30) minutes** |             |           |        |     |      |               |
| Observed          | 9          | 2         | 2      | 0   | 2    | 0             |
| Percentage        | 60%        | 13.33%    | 13.33% | 0   | 13.33%| 0             |
| **Experiment time (31–45) minutes** |             |           |        |     |      |               |
| Observed          | 9          | 2         | 1      | 0   | 3    | 0             |
| Percentage        | 60%        | 13.33%    | 6.67%  | 0   | 20%  | 0             |

Table 17
Distribution of high expressions occurrences in the first session [0–15].

|                  | First session [0–15] |
|------------------|----------------------|
|                  | Eyes in center Eyebrows-center Mouth-closed Hand-No Movement Head-Center |
|                  | 80% 53.33% 60% 60% 100% |

6.6. Head expressions recorded at the experiment

From the analysis of the head variations as shown in Table 16, we found that the head has the highest degree of expressing the psychological state of the participants in the experiment.

It was observed that the participants' heads have three variations (Right, left, and center). In the first session, the variation of the head in the center position scores the highest with a percentage of 100%, and no other variations were observed. While in the second session there was a significant drop in “center position” percentage to 13.2% and some of new head variations were observed i.e., "tilt to right-side" with a percentage of 60%, and "tilt to left-side" with a percentage of 13.33% and also to downside with a percentage of 13.33%, which means that these variations may also have a relation to the degree of participants' engagement. The results in the last session of the experiment came to enhance the results of the second phase where the same percentage of all variations are the same in both sessions.

6.7. Participant's electrocardiograms analysis

As mentioned earlier, the pulse measurements in this study served to increase the reliability of our interpretation of the effects of facial expressions and to measure the level of student engagement in the learning environment. Tables 17–19 summarize the facial expressions with the highest frequency during the experiment, Figs. 14a, 15a and 16a show an example of the participants' electrocardiogram, and Figs. 14b, 15b and 16b show an example of the facial expressions associated with each electrocardiogram.

The above heart rate graph in Fig. 14a shows that on average, most participants have a normal heart rate, which is within the normal range. When we compare this result with the participants' exam results (Q1 and Q2), we find that about 70% of the participants answer the exam questions correctly, and since we
In the first session of the experiment, the participants are at the highest level of attention, which indicates that the facial expressions (i.e., eye in the center, head in the center, eyebrows straight, and no hand movements), as shown in Fig. 14b, have the highest frequency, directly related to the level of engagement in the lesson.

In the second session, the participant’s pulse became high and not stable as shown in Fig. 15, this can be explained by the presence of changes in the mental state of the participants compared to the first session, in other words, a new physical activity began to appear in the participants, where there was a change in the facial expressions and their frequency, the frequency of some
To give meaning to these facial expressions and by comparing this result with the participants’ exam scores (Q3), we found that the percentage of participants who gave correct answers to the exam questions decreased to 40% and the percentage of those who gave incorrect answers increased to 47%. This suggests that the facial expressions (head left and right, eyes down to the center, fingers and mouth) may mean that the level of engagement decreased.

In the last session of the experiment, the result confirmed our expectation that the participants reached the lowest level of engagement. 16 shows a significant change in auditory stimuli compared to the first two sessions. Moreover, new facial expressions appear very frequently, while some facial expressions disappear or reach the lowest levels in this session.

Comparison of these results with participants’ final exam scores (Q4 and Q5) revealed that approximately 57% of participants could not answer the exam questions. In addition, the significant changes in participants’ mental state indicate low levels of participant engagement, and the facial expressions (eyes down to center, fingers and mouth, eyebrows down, and right and left head) that are most common are directly related to the level of engagement in class.

6.8. Results and analysis

After analyzing the participants’ facial expressions and their distribution across the experimental phases, the students’ test scores were correlated with their facial expressions to understand the meaning of these expressions and to determine the extent to which these expressions are relevant to measuring the participant’s level of engagement with the learned topic.

In the experimental sessions, some facial expressions have different repetition rates (high, medium, low), and by comparing these repetitions with participants’ scores on the test questions in Section 6.1 and with the electrocardiograms in Section 6.7, we
can conclude that the group of facial expressions that achieved the highest repetition rate in the first session can be considered as an indicator of a high level of engagement, while the group of expressions that achieved the highest repetition rate in the second session is considered as an indicator of a decrease in participants’ engagement, and the group of expressions that achieved the highest repetition rate in the third session is considered as an indicator of a low level of participants’ engagement. Figs. 17, 18, and 19 show the distribution of facial expressions that occurred most frequently in the experiments, and Table 20 summarizes the most frequently observed expressions and their possible indications.

7. Conclusion

E-learning is an indispensable solution and could be the best alternative to face-to-face learning, especially in the pandemic COVID–19 and in any kind of crisis that may occur.

The first stage aimed to capture the facial expressions of the participants and distribute them to three time intervals, i.e. [0–15], [16–30] and [31–45].

In the second phase, the most frequent facial expressions in each time interval were counted. In the third phase, the facial expressions were linked to the results of the students’ brief examination, which took place at the end of each time interval of the lesson.

At the end of the study, three groups of facial expressions were identified, which are believed to be indicative of the participants’ level of engagement in the learning environment.

To verify this result, the participants’ heart rates were recorded during the three time intervals of the experiment, as the electrocardiograms showed significant differences between the participants’ heart rates during these time intervals, confirming that the three groups of facial expressions observed during the time intervals of the experiment were related to the differences in the participants’ psychological state.

Based on these results, we conclude that the level of students’ engagement in distance learning environments can be measured by facial expressions. The figure shows the proposed scale for engagement based on facial expressions. Fig. 20, shows the proposed mimic scale for measuring engagement.
Y. Alzyoud: Data curation, Investigation, Visualization.
Faisal – original draft, Supervision, Writing – review & editing.

CRediT authorship contribution statement

Waleed Maqableh: Conceptualization, Methodology, Writing – original draft, Supervision, Writing – review & editing. Faisal Y. Alzyoud: Data curation, Investigation, Visualization. Jamal Zraqou: Visualization, Investigation, Software, Writing – review & editing.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Dr. Maqableh is an assistant proof. In multimedia and virtual Reality, Dr. Zaraqo is an associated proof. In multimedia, Dr. Alzyoud is an associated proof. In Computer Science

Acknowledgments

The researchers would like to thank Abdoun Medical Center for their cooperation in this study by providing the heart rate sensors (TTL SA9308M Heart Rate/BVP). Special thank to Dr. Hatem Al-Tarawneh for his valuable comments on the electrocardiograms of the students who participated in this experiment. Finally, we also would like to thank each of Luminus, Isra, and Petra Universities for their help and support.

Table 20
Distribution of facial expressions with high frequency rate.

| Expression                  | Percentage | Indications          |
|-----------------------------|------------|----------------------|
| Eyes in center position     | 80%        | High level of engagements |
| Eyebrows-center             | 53.33%     | High level of engagements |
| Mouth-closed                | 60%        | High level of engagements |
| Hand-No Movement            | 60%        | High level of engagements |
| Head-Center                 | 100%       | High level of engagements |

| Expression                  | Percentage | Indications          |
|-----------------------------|------------|----------------------|
| Eyes looking down           | 53.33%     | Engagements decreased |
| Eyebrows-up                 | 26.67%     | Engagements decreased |
| Mouth-twisted left          | 26.67%     | Engagements decreased |
| Hand-with Movement          | 20%        | Engagements decreased |
| Fingers & mouths            | 26.67%     | Engagements decreased |
| Head Tilt-Right             | 60%        | Engagements decreased |

| Expression                  | Percentage | Indications          |
|-----------------------------|------------|----------------------|
| Eyes looking down           | 53.33%     | Low level of engagements |
| Eyes Right-Up               | 20%        | Low level of engagements |
| Fingers & mouths            | 20%        | Low level of engagements |
| Head Tilt-Right             | 60%        | Low level of engagements |
| Head down                   | 20%        | Low level of engagements |

Fig. 20. Facial expressions scale.

Ethical Approval

This study does not contain any studies with human or animal subjects performed by any of the authors.

References

Al-Helaili, A.H.M., et al., 2021. A new hybrid multi-focus image-fusion using dmmf with fft transforms. J. Theor. Appl. Inform. Technol. 99 (10).
Ayvaz, U., Gürüler, H., Devrim, M.O., 2017. Use of facial emotion recognition in e-learning systems.
Benta, K.-L., Cremene, M., Vaida, M.-F., 2015. A multimodal affective monitoring tool for mobile learning. In: 2015 14th RoEduNet International Conference-Networking in Education and Research. RoEduNet NER, IEEE.
Buja, L.M., 2019. Medical education today: all that glitters is not gold. BMC Med. Educ. 19 (1), 1–11.
Cao, Y., et al., 2005. Expressive speech-driven facial animation. ACM Trans. Graph. 24 (4), 1283–1302.
Chai, C.S., et al., 2014. Deepening ICT integration through the multilevel design of technological pedagogical content knowledge. J. Comput. Educ. 1 (1), 1–17.
Cranford, K.N., Tiettmeyer, J.M., Chuprinko, B.C., Grove, N.P., 2014. Measuring load on working memory: The use of heart rate as a means of measuring chemistry students’ cognitive load. J. Chem. Educ. 91, 641–647.
Devan, M., Murshed, M., Lin, F., 2019. Engagement detection in online learning: a review. Smart Learn. Environ. 6 (1), 1–20.
Fabbri, R., et al., 2008. 2D euclidean distance transform algorithms: A comparative survey. ACM Comput. Surv. (CSUR) 40 (1), 1–44.
Gaur, U., et al., 2020. Challenges and opportunities of preclinical medical education: COVID-19 crisis and beyond. SN Compr. Clin. Med. 2 (11), 1992–1997.
Geng, G., 2011. Investigation of teachers’ verbal and non-verbal strategies for managing attention deficit hyperactivity disorder (ADHD) students’ behaviours within a classroom environment. Australian J. Teach. Educ. 36 (7), 17–30.
Harris, P.L., 2012. Trusting What You'Re Told: How Children Learn from Others. Harvard University Press.
Hodges, C.B., et al., 2020. The difference between emergency remote teaching and online learning.
Kirschner, P.A., 2015. Do we need teachers as designers of technology-enhanced learning? Instr. Sci. 43 (2), 309–322.
Maier, K.J., Waldstein, S.R., Synowski, S.J., 2003. Relation of cognitive appraisal to cardiovascular reactivity, affect, and task engagement. Ann. Behav. Med. 26, 32–41. http://dx.doi.org/10.1207/S15324796AMB2601_05.
Mohamad Nezami, O., et al., 2020. Automatic recognition of student engagement using deep learning and facial expression. In: Joint European Conference on Machine Learning and Knowledge Discovery in Databases. Springer.
Moore, J.L., Dickson-Deane, C., Galen, K., 2011. e-learning, online learning, and distance learning environments: Are they the same? Internet Higher Educ. 14 (2), 129–135.
Mukhtar, K., et al., 2020. Advantages, Limitations and Recommendations for online learning during COVID-19 pandemic era. Pakistan J. Med. Sci. 36 (COVID19-54), S27.
Munoz, A., Mackay, J., 2019. An online testing design choice typology towards cheating threat minimization. J. Univ. Teach. Learn. Pract. 16 (3), 5.
Ożadowicz, A., 2020. Modified blended learning in engineering higher education during the COVID-19 lockdown—Building automation courses case study. Educ. Sci. 10 (10), 292.
Pather, N., et al., 2020. Forced disruption of anatomy education in Australia and New Zealand: An acute response to the COVID-19 pandemic. Anat. Sci. Educ. 13 (5), 284–300.

CRediT authorship contribution statement

Waleed Maqableh: Conceptualization, Methodology, Writing – original draft, Supervision, Writing – review & editing. Faisal Y. Alzyoud: Data curation, Investigation, Visualization. Jamal Zraqou: Visualization, Investigation, Software, Writing – review & editing.
Peled, R., Carmil, D., Siboni-Samocha, O., Shoham-Vardi, I., 2008. Breast cancer, psychological distress and life events among young women. BMC Cancer 8 (245).

Rizwan, S.A., Jalal, A., Kim, K., 2020. An accurate facial expression detector using multi-landmarks selection and local transform features. In: 2020 3rd International Conference on Advancements in Computational Sciences. ICACS, IEEE.

Sathik, M., Jonathan, S.G., 2013. Effect of facial expressions on student’s comprehension recognition in virtual educational environments. SpringerPlus 2 (1), 1–9.

Shen, D., et al., 2013. Unpacking online learning experiences: Online learning self-efficacy and learning satisfaction. Internet High Educ. 19, 10–17.

Tian, Y.-L., Kanade, T., Cohn, J.F., 2000a. Recognizing upper face action units for facial expression analysis. In: Proceedings IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2000. Cat. No. PR00662, IEEE.

Tian, Y.-L., Kanade, T., Cohn, J.F., 2000b. Recognizing lower face action units for facial expression analysis. In: Proceedings Fourth IEEE International Conference on Automatic Face and Gesture Recognition. Cat. No. PR00580, IEEE.

Vail, A.K., et al., 2016. Predicting learning from student affective response to tutor questions. In: International Conference on Intelligent Tutoring Systems. Springer.

Van Heerden, D., Goose, L., 2020. Promoting the growth of fourth industrial revolution information communication technology students: The implications for open and distance E-learning. In: Promoting Inclusive Growth in the Fourth Industrial Revolution. IGI Global, pp. 118–147.

Welsen, S., Morgan, P., Walker, J., 2020. Flexible learning during educational disruption: a case study of teaching integrated circuits design. In: 2020 IFEES World Engineering Education Forum–Global Engineering Deans Council. WEEF-GEDEC, IEEE.

Wu, C.-H., 2016. New technology for developing facial expression recognition in e-learning. In: 2016 Portland International Conference on Management of Engineering and Technology. PICMET, IEEE.

Wu, T., Fu, S., Yang, G., 2012. Survey of the facial expression recognition research. In: International Conference on Brain Inspired Cognitive Systems. Springer.

Zraqou, J., Alkhadour, W., Al-Nu’aimi, A., 2014. Robust real-time facial expressions tracking and recognition. Int. J. Technol. Educ. Mark. (IJTEM) 4 (1), 95–105.