



# Framework for identifying and visualising emotional atmosphere in online learning environments in the COVID-19 Era

Fei Yan<sup>1</sup> · Nan Wu<sup>2</sup> · Abdullah M. Ilyasu<sup>3,4</sup> · Kazuhiko Kawamoto<sup>2</sup> · Kaoru Hirota<sup>4</sup>

Accepted: 12 October 2021 / Published online: 6 January 2022

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

## Abstract

In addition to the almost five million lives lost and millions more than that in hospitalisations, efforts to mitigate the spread of the COVID-19 pandemic, which that has disrupted every aspect of human life deserves the contributions of all and sundry. Education is one of the areas most affected by the COVID-imposed abhorrence to physical (i.e., face-to-face (F2F)) communication. Consequently, schools, colleges, and universities worldwide have been forced to transition to different forms of online and virtual learning. Unlike F2F classes where the instructors could monitor and adjust lessons and content in tandem with the learners' perceived emotions and engagement, in online learning environments (OLE), such tasks are daunting to undertake. In our modest contribution to ameliorate disruptions to education caused by the pandemic, this study presents an intuitive model to monitor the concentration, understanding, and engagement expected of a productive classroom environment. The proposed apposite OLE (i.e., AOLE) provides an intelligent 3D visualisation of the classroom atmosphere (CA), which could assist instructors adjust and tailor both content and instruction for maximum delivery. Furthermore, individual learner status could be tracked via visualisation of his/her emotion curve at any stage of the lesson or learning cycle. Considering the enormous emotional and psychological toll caused by COVID and the attendant shift to OLE, the emotion curves could be progressively compared through the duration of the learning cycle and the semester to track learners' performance through to the final examinations. In terms of learning within the CA, our proposed AOLE is assessed within a class of 15 students and three instructors. Correlation of the outcomes reported with those from administered questionnaires validate the potential of our proposed model as a support for learning and counselling during these unprecedented times that we find ourselves.

**Keywords** Online learning environment · Classroom atmosphere · Facial expression recognition · Emotion space · COVID-19

## 1 Introduction

The spread of COVID-19 pandemic swept the globe with incendiary events that transformed not only economies and health, but also affected education at all levels, in all nations,

and for all people [1]. One of the most oft-used terms signifying changes brought by pandemic is the phrase “new normal”, which in education entails the emerging learning paradigms. Notably, there has been a significant surge in usage of language applications, virtual tutoring, video conferencing tools, and online learning software [2]. Prior to this “new normal” online and distance learning were the terms connoting knowledge transfer through the internet usually from anywhere in the globe to targeted audiences [3]. However, since the advent of the pandemics, lexicons such as virtual learning, remote learning blended learning, etc. have become popular and are used to connote today's alternative to the traditional physical or face-to-face (F2F) classroom instruction the world over. These platforms are intended to provide some semblance of traditional F2F interaction and communication. It is without doubt that some form of virtual learning will be retained in the post-pandemic era.

✉ Abdullah M. Ilyasu  
a.ilyasu@psau.edu.sa

<sup>1</sup> School of Computer Science and Technology, Changchun University of Science and Technology, Changchun, China

<sup>2</sup> Graduate School of Science and Engineering, Chiba University, Chiba, Japan

<sup>3</sup> College of Engineering, Prince Sattam Bin Abdulaziz University, Al-Kharj, Saudi Arabia

<sup>4</sup> School of Computing, Tokyo Institute of Technology, Tokyo, Japan

Online learning environments (OLE) offer their own merits and demerits many of which have been unveiled by the prevailing learning models [4]. In addition to helping mitigate the spread of COVID, OLEs provide accessibility in terms of location, cost effectiveness, better time management, new technical skills, etc. On the downside, limitations of OLE include prolonged screen time and its attendant health implications, exclusion of students from poor backgrounds (as a result of cost of devices, internet, etc.), heightened sense of isolation, etc. [5]. Despite all this, in most cases, OLE potentially increases time for social engagement with family after or in-between classes.

In terms of pedagogy, learning theory describes how students receive, process, and retain knowledge during learning. Cognitive, emotional, and environmental influences, as well as prior experiences, all play a part in how understanding, or a world view, is acquired or changed and how knowledge and skills are retained. As a popular learning theory, constructivism learning theory is based on the idea that students actually create their own learning based on their previous experiences [6]. In other words, it presupposes that the students take what they are being taught and add it to their previous knowledge and experiences, creating a reality that is unique to them. This learning theory focuses on learning as an active process, which is personal and individual for each student. In this regard, constructivism is crucial to helping many kinds of students take their own experiences and include them in their learning.

Moreover, it has been observed that no matter the effort by instructors to design courses and material, there is no guarantee that comprehension will be successful. Notwithstanding this, the multitude of emotional and psychological disruption to their lives and to the learning process brought about by COVID make it necessary for instructors to constantly monitor their students and initiate timely adjustments to teaching strategies. Prior to COVID, most research associated with online learning were focused on technology used, pedagogy, resources, and efficiency [7]. Therefore, little effort is invested in identifying the emotional stress involved in such learning paradigms. Now, faced with COVID pandemic whence everyone is traumatised by the changes in daily life, emotional and psychological states of learners cannot be overlooked. Moreover, studies across different areas of neuroscience, education and psychology have showcased that emotional wellbeing of learners is as important as the curriculum [8].

Emotions stimulate learners' attention and interest; they influence comprehension and retention, which trigger the learning process. In F2F instruction, emotions are easier to read and address, whereas in OLE instructors must do more to promote emotional presence and engagement. Despite their best efforts, instructors may not be able to stimulate and sustain concentration, engagement, and ultimately

an effective learning process. In most of the present systems, each student is monitored via the combination of a screen and learning management system (LMS) platform that are grossly inadequate for superintending the dynamics of students' emotions on online learning platforms. Consequently, a new tool or system is necessary.

Presently, most studies on emotion recognition are focused on automating the recognition based on facial expressions from images and video [9], spoken words from audio recordings [10], written expressions from text [11] and physiological variables measured by wearable devices [12]. In the COVID era, online instruction includes the use of videos, which provides substantial data for emotion recognition. In this regard, the advances recorded in facial emotion recognition could be very useful. For example, in [13], a framework that integrates a facial recognition algorithm into online learning was proposed. Cameras on the framework were used to capture facial images wherefrom facial expressions were deduced, analysed, and classified into 8 categories of emotions using a facial emotion recognition (FER) algorithm. However, the general classification of the emotion types is not suitable for deciphering and understanding the status of learners in a classroom. For example, in the experiments reported in [13], there are 27 faces, but none exhibited emotions such as anger or fear. In [14], a learning engagement framework was proposed based on students' behaviours that are extracted from facial images and the computer mouse on the OLE platform. The aim of that study was to investigate whether the use of mouse-movement data can enhance the effect of learner participation as detected via a camera. However, the extent of student understanding of the course content was not evaluated using that model. In their contribution, [15] examined changes in emotions of a group students during classes. The students' facial expressions were analysed and digitised to discern their respective emotional states. However, this study did not provide any analysis of each student's emotion over entire study periods (for example, a whole lesson, or even a semester) which limits its use in making conclusions regarding targeted solutions for each student. Similarly, in [16], an intelligent adaptive e-learning environment was modelled by integrating learners' responses to questions and their emotional states. Following that, it introduces a method whereby a group of facial expressions is aggregated into a single representative emotion that provides a new learning resource matching the next learning level. Whereas it was impressive to monitor the learning performance for each learner, the model did not provide a real-time evaluation of the learning cycle based on which the instructor could observe the student's learning status and adjust his/her teaching strategy as required. Most importantly, throughout the available literature, there are very few studies that

provide interactive visualisation frameworks equipped with necessary interfaces for instructors to monitor and engage the whole class and offer timely interventions to individual learners or the entire class.

Considering the unplanned but indispensable COVID imposed transition to online and virtual learning, which many argue is here to stay, the shortcomings enumerated above can no longer be overlooked. This requires integration of some traditional mechanisms into new ones. Among others, we highlight two areas we envision such integration. First, some form of online curriculum and assessment criteria needs to be formulated mainly to monitor learners' emotions in classes throughout the semester and relative to their performance in different assessments. Second, real-time communication and interaction are increasingly required between the instructor and learners to improve the quality of teaching and learning. Therefore, the instructor should be capable of observing the collective status of all the learners in the class at any time and then adjust his/her teaching strategies as necessary.

This study contributes towards attainment of the enumerated solutions via an apposite online learning environment (AOLE) for automatic identification and visualisation of classroom emotional dynamics of individual students. Therefore, the core contributions of the study could be viewed in three areas:

- (1) Adopting a facial expression recognition (FER) technique that utilises convolutional neural networks to generate emotional states of individual students during real-time online sessions and label their emotions within an intuitive emotion coordinate system.
- (2) Utilising a fuzzy inference system for real time identification and visualisation of the emotion status of the whole class, which is composed of emotions

of all the students. This provides instructors with an impression of classroom emotional atmosphere and a mechanism to track the learning context as classes progress.

- (3) Supporting course assessment by using a scoring mechanism to evaluate and analyse subjacent emotional cues of each student, such as emotions during daily classes, and their dynamics over the semester as well as their correlation with students' performance in exams and other assessments.

To deliver the enumerated outcomes, the remainder of the study is structured as follows. We present a general outline and composition of the proposed AOLE system in Section 2. Following this, in Section 3, we introduce the intrigues of each unit of the system and its deployment in actual OLEs. This includes discussions on the identification and visualisation of emotional states estimated via administered questionnaires. Finally, in Section 4, we present insights for further improvement of the proposed system.

## 2 General framework of the proposed AOLE system

The architecture of the proposed system, which outlines the facial expression recognition (FER) layer, classroom atmosphere construction layer, and data visualisation as the three interconnected layers of the proposed AOLE system is presented in Fig. 1. The layers interact through an interface that supports replacement and upgrading of their units when required.

At the bottom of the pipeline of our proposed system is the facial expression recognition (FER) layer, where specific emotions of each student are extracted and extrapolated onto their respective emotion coordinates.

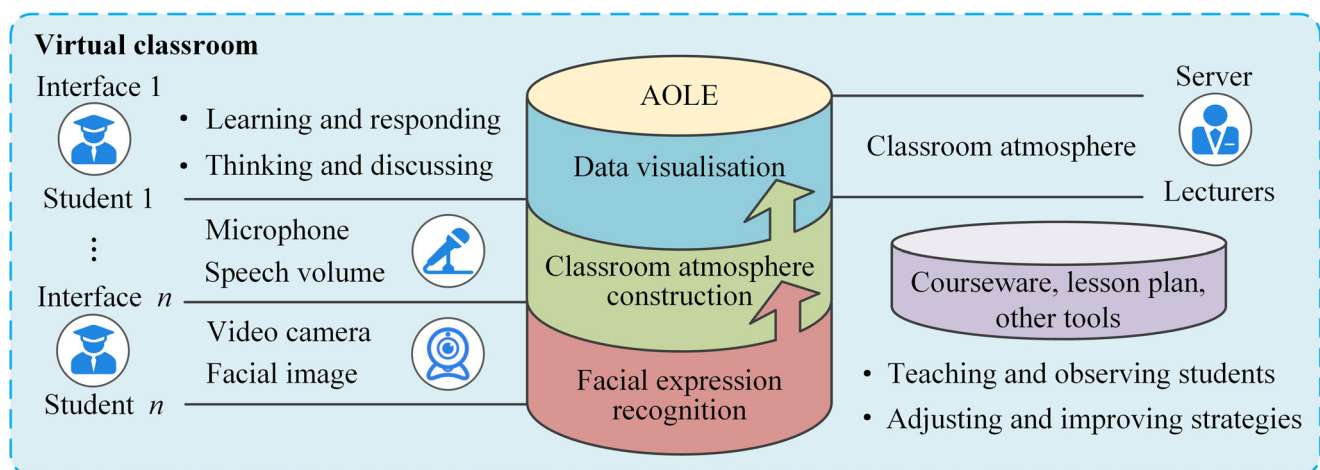


Fig. 1 Architecture of proposed AOLE system

Present FER systems are limited to basic classification of basic emotions, such as sadness, happiness, and the like. However, we note that formulating a competitive FER algorithm is outside the purview of this study. Instead, we adopt a recent, astute FER deep learning model (i.e., the Deep-emotion from [17]) as the algorithm of our FER layer. Further reasoning in support of this choice is presented later in Section 3. Therefore, in AOLE, the classroom atmosphere construction layer fuses the emotion states from the FER layer to build a unified state that reflects the learning atmosphere of the classroom. Here, we note that presently few studies have considered virtual states in OLEs and existing OLEs do not provide real-time visual information to support required adjustments to teaching strategies.

In our construction, the visualisation layer provides an intuitive graphic that reflects the learning status of students in the classroom. Furthermore, in each of the AOLE layers, different approaches and upgrades can be used to improve functionality and upgrades. For example, students' emotion curve during a specific learning period can be built and analysed relative to that period as well as in terms of how it relates to test scores. The remainder of the section outlines technologies in each layer and their use to realise the objectives of the AOLE system.

### 2.1 Generation of emotional coordinate

In order to apply facial expression for emotion recognition and OLEs, we need to start by understanding the causes of different students' emotions during class. Research has shown that students experience many emotions during lessons, while studying, and when taking tests and

examinations [18]. These emotions range from being delighted, relaxed, bored, tired, frustrated, or tense. Furthermore, these emotions can be positive or negative, and they can be intense and/or frequent. Additionally, these emotions can be affected by classroom factors (e.g., curriculum content, environment), individual differences between students (e.g., genetic factors, mood swings, general tendencies), and external factors (e.g., social interactions and home environment). Therefore, these emotions can each affect students and their learning in a variety of ways [19]. For example, tests, examinations, homework, and deadlines are associated with different emotional states that encompass frustration, anxiety, and boredom. Even aesthetic pleasantness or lack thereof in an environment is deemed to influence emotions, which in turn affect one's ability to concentrate, learn and remember. Given the number of students, variety of emotions and their causes, it is nearly impossible for instructors to manage the enumerated experiences effectively.

Based on the highlighted specifications, we target a model capable of recognising students' facial expressions and mapping them onto emotion coordinates of a Valence-Arousal (VA) emotion space. In order to save cost and improve accuracy, we make use of a transfer learning algorithm that includes a spatial transformer network (STN) [20] as composed in Fig. 2.

In the STN, the features extraction unit consists of four convolutional layers (i.e.,  $3 \times 3 \times 10$  kernels), each pair followed by a max-pooling layer and rectified linear unit (ReLU) activation function and, finally, a dropout layer. Subsequent outcomes of the localisation network, which regresses the transformation parameters, are transformed to the sampling grid  $\tau(\theta)$  that produces warped data. For this,

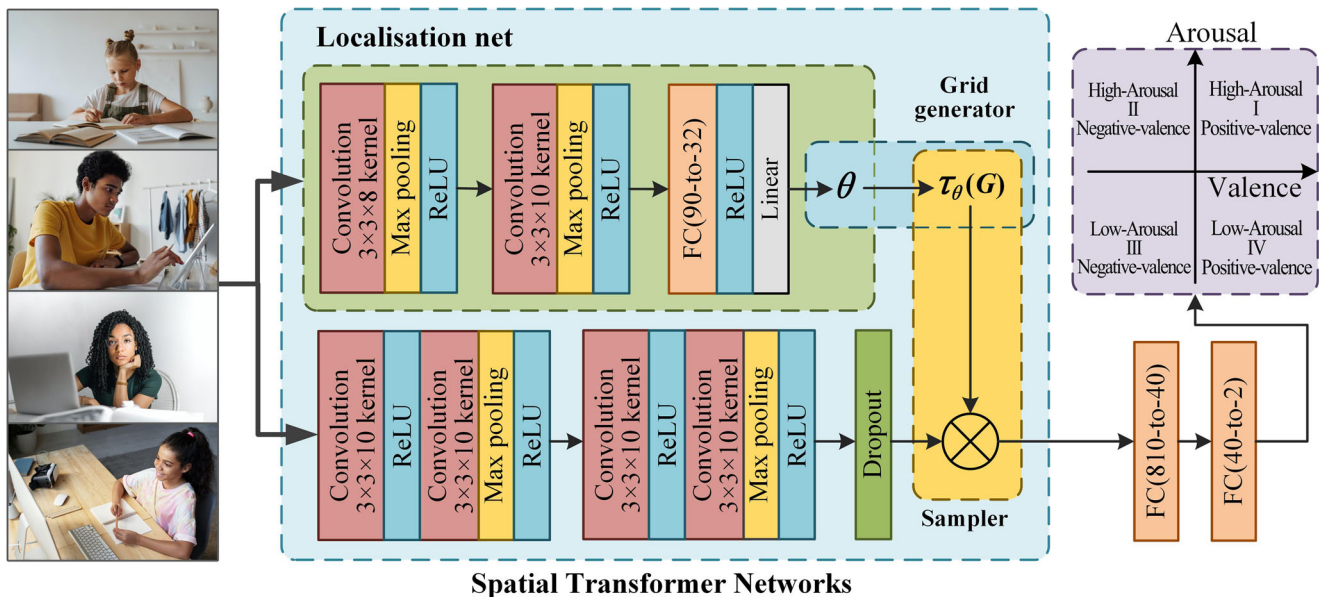


Fig. 2 Facial expression recognition unit for generation of coordinates in the VA emotion space

images in the FER 2013 [17] dataset are used as training model for expression recognition. This dataset is widely used in studies on facial emotion recognition, such as its use to analyse the psychological condition of patients in [21]. FER 2013 facial expression dataset consists of 35886 facial greyscale images each  $48 \times 48$  in size with 28708 as training images, while the remaining 7178 images are further divided equally for verification and testing (i.e., 3589 images each). Therefore, with STN, two fully connected layers with 40 and 2 nodes, respectively, are used to specify a point on the VA scale. Furthermore, manual annotation is used to delineate a new dataset comprising of 300 images and a further 10% (i.e., 30 images) for the training needed to fine tune the model. The resulting model is then used to predict a person's emotion according to his/her facial expression. Additionally, when building the model, an atmosphere can be represented as a two, three or multi-dimensional space like the emotional coordinate system [22]. However, three-dimension (3D) space commonly used because it best simulates our 3D lives. In this context, in our AOLE model, three coordinates of Understanding (U), Concentration (C), and Engagement (E) are used to represent emotions expected in a typical classroom environment.

Meanwhile, according to the dimensionality theory of emotion space [23], all human emotions are distributed in a certain dimensional space and various emotions are distributed in different positions according to attributes of the used dimensions. In this study, we use the VA space where valence (V) and arousal (A) dimensions represent the degree of pleasant-unpleasant and excited-calm emotions, respectively [24]. Furthermore, the model is divided into four quadrants (based on positive and negative valence as well as the high and low arousal), each representing a human's emotional states according to the combination of values in the Valence and Arousal dimensions. Using this intuition, the emotional state of each student can be fused to produce an atmosphere. This process is further expatiated in the next subsection.

## 2.2 Atmosphere identification from emotional states

Instinctively, an atmosphere is invisible, yet it is supposed to exist by occupying space and percolating energy [25]. It is generated by the emotion of each individual during an interactive communication, and it also influences the emotional state of each individual as well as the collective emotions in the atmosphere [26]. In the context of our AOLE model, an atmosphere refers to the learning situation that can reflect the learning states of all the students in a virtual class on an OLE platform. Usually, in a class, by observing the overall learning atmosphere together combined with the knowledge setting (such as the degree of difficulty of the course or topic), an instructor can measure

the learning situation and then adjust his/her teaching strategy as appropriate.

In AOLE, we define Understanding as the degree of comprehension and interpretation of content relative to content read in lecture material, on-going lesson with instructor and previous knowledge related to these two. On its part, Concentration is considered as the extent the learner exerts effort and attention on absorbing the various content disseminated without distraction. This entails focusing on the instructor as well as the hard and soft material related to the course or topic. Finally, the third dimension of our AOLE system, Engagement refers to the degree of the learner's curiosity, interest, and passion towards comprehending the content being taught. This extends to the underlying motivation for learning and progress in their studies.

As previously mentioned, in OLE setups, each learning station is equipped with an internet-ready device, a camera, and a microphone. Therefore, facial expressions and spoken features can be calculated and analysed. Based on the 3D space espoused earlier and the outlined relationship between emotion and learning, the learning station could be used to extract the learner's understanding and concentration during classes. However, on the surface, these attributes cannot be easily perceived on standard OLEs. For example, fatigue (i.e., state of tiredness) could affect concentration, which could impede understanding. Similarly, frustration can affect understanding, which could, to some extent, stall concentration. Meanwhile, using the microphone (i.e., speech volume) the third dimension of our AOLE, i.e., engagement, could be integrated to effectively assess the emotional composition of the system. Here, both individual and collective engagement could be considered based on number of learners participating in a discussion (i.e., utterances) as well as the overall pitch and volume in speech. Consequently, all the three components of the AOLE could be recognised as separate axis of our emotional atmosphere.

The process of interpreting the combined emotion state in a classroom atmosphere (CA) are discussed in the sequel. Before that, in order to construct the CA in an OLE, we note that:

- (1) Changes in CA are continuous, so its value could be retrieved instantaneously at time  $t-1$ ,  $t$ , and  $t+1$  as CA ( $t-1$ ), CA ( $t$ ) and CA ( $t+1$ ), respectively. At any instance, such CA is influenced by the current emotional state of whole class as well as the atmosphere at an earlier instance.
- (2) CA is percolated by energy that changes dynamically with time. This means that when new emotions are not added, the strength of the CA weakens. However, it does not disappear instantaneously but stabilises at some neutral state over time.

- (3) Every expression of the learner contributes to the CA. However, theoretically, the contribution, i.e., weight, of each expression to the the cumulative CA should be the same. Nevertheless, instructors could preset these weights for one or more students depending on intricacies and peculiarities of the situation.

In the meantime, since it is difficult to establish an exact mathematical model of an atmosphere [27], then it becomes necessary that the uncertainty and vagueness of the atmosphere are also considered. Such consideration requires subjective comprehension to determine its attributes and the use of mathematical vagueness in fuzzy logic to determine the reasoning from atmosphere-related factors to the CA [28]. Consequently, the definition of CA could be formalised as:

$$CA(t) = (1 - \lambda) \sum_{i=1}^n \omega_i \tilde{E}_i(t-1)\gamma + \lambda \sum_{i=1}^n \omega_i \tilde{E}_i(t),$$

$$t = 1, 2, 3, \dots, n,$$
(1)

where  $\tilde{E}_i(t)$  is the function of emotional states at time  $t$  after the fuzzy inference,  $n$  is the number of individuals involved in the communication scenario, and  $\gamma$  is a monotonically decreasing function in the form  $0 \leq \gamma \leq 1$ , such as an exponential function  $exp(-kT)$ ,  $k$  is a positive parameter defining the decreasing speed of the CA( $t-1$ ),  $T$  is the sampling period for calculating the CA. Additionally,  $\lambda$  is the correlation factor in the range  $0 \leq \lambda \leq 1$ ; where at  $t=0$ , the CA is at the origin (i.e., as the initial state) and when  $t=1$ ,  $\lambda$  is set as 1, signifying that the CA acquired its first set of data.

The illustration of the whole procedure explained so far is presented in Fig. 3, where, following the formulation in Eq. (1), fuzzy inference is used to generate CA from emotional states. Fuzzy membership functions are used to transform individual emotional states into the CA. Moreover, in the VA emotion space, the affect grid is used to set five levels of Valence and Arousal as illustrated in Table 1. Evenly distributed membership functions where the fuzzy domain of each axis is defined in the range 0 to 1 is adopted. Similarly, the Valence is graduated in terms of five linguistic variables namely: Very Negative Valence (VNV), Negative Valence (NV), Average Valence (AV), Positive Valence (PV), and Very Positive Valence (VPV). Consequently, five linguistic variables are employed for the Arousal axis. These are Very High Arousal (VHA), High Arousal (HA), Average Arousal (AA), Low Arousal (LA) and Very Low Arousal (VLA). Finally, for defuzzification of the extreme input values, we utilise the trapezoidal membership function instead of the triangular membership function.

The two sets of outputs from the CA are defined in seven levels as presented in Table 1. Using the evenly distributed triangular membership function whose fuzzy domain is defined in the range 0 to 1, we obtain several linguistic variables to describe the understanding component of our AOLE. These variables are Extremely Low Understanding (ELU), Very Low Understanding (VLU), Low Understanding (LU), Average Understanding (AU), Very High Understanding (VHU), and Extremely High Understanding (EHU). Correspondingly, the concentration component could be described using linguistic variables Extremely Low

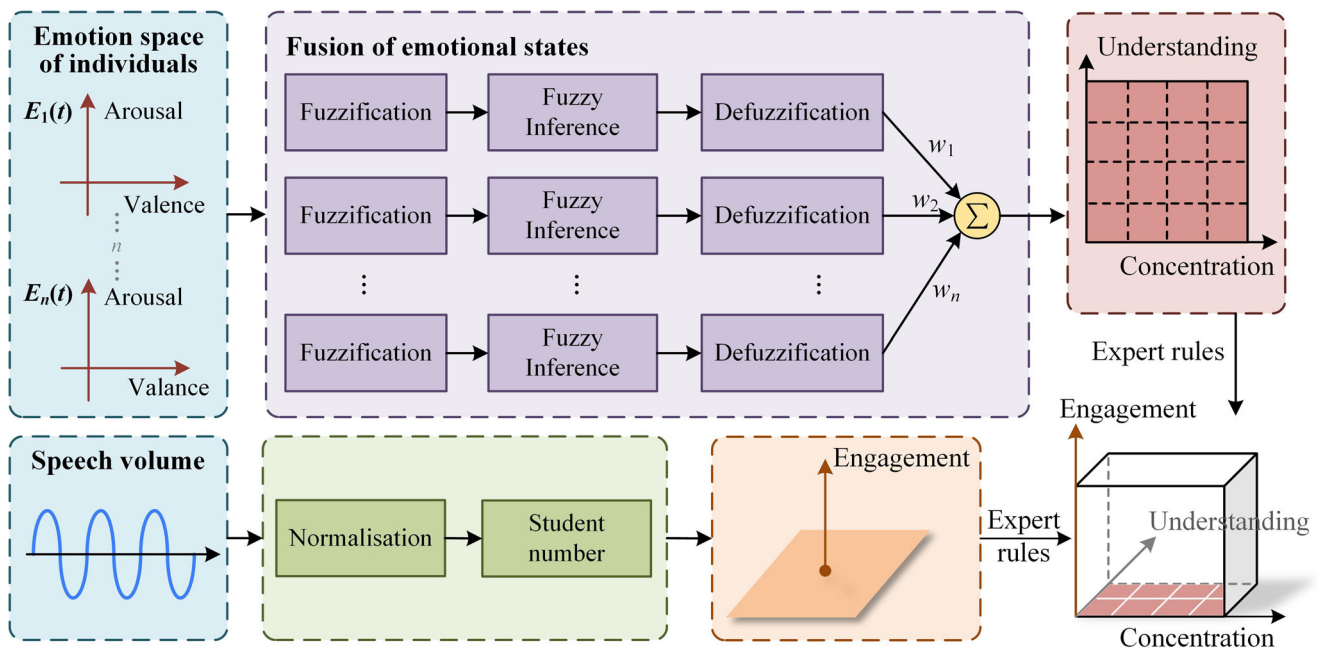


Fig. 3 Illustration of fusion of emotional states to generate classroom atmosphere (CA) in AOLE system

**Table 1** Membership functions, linguistic variables, and terms used in the AOLE with their descriptions

Term	Variable	Membership	Description
Valence	Very negative Negative Average Positive Very positive		VNV:=trape -1 -1 -0.8 -0.4; NV:=trian -0.8 -0.4 0; AV:=trian -0.4 0 0.4; PV:=trian 0 0.4 0.8; VPV:=trape 0.4 0.8 1 1
Arousal	Very low Low Average High Very high		VLA:=trape -1 -1 -0.8 -0.4; LA:=trian -0.8 -0.4 0; AA:=trian -0.4 0 0.4; HA:=trian 0 0.4 0.8; VHA:=trape 0.4 0.8 1 1
Understanding	Extremely low Very low Low Average High Very high Extremely high		ELU:=trian 0 0 0.2; VLU:=trian 0 0.15 0.3; LU:=trian 0.2 0.35 0.5; AU:=trian 0.35 0.5 0.65; HU:=trian 0.5 0.65 0.8; VHU:=trian 0.65 0.8 0.95; EHU:=trian 0.8 1 1
Concentration	Extremely low Very low Low Average High Very high Extremely high		ELC:=trian 0 0 0.2; VLC:=trian 0 0.15 0.3; LC:=trian 0.2 0.35 0.5; AC:=trian 0.35 0.5 0.65; HC:=trian 0.5 0.65 0.8; VHC:=trian 0.65 0.8 0.95; EHC:=trian 0.8 1 1

Concentration (ELC), Very Low Concentration (VLC), Low Concentration (LC), Average Concentration (AC), High Concentration (HC), Very High Concentration (VHC) and Extremely High Concentration (EHC).

Subsequently, in this study, we use the VA model to record the emotional states of students in the learning process, where Valence records the values of pleasant and unpleasant experiences that reflect the students' preferences during learning; while Arousal records the values of excitement and calmness, which reflect the students' liveliness (i.e., alertness) during the learning process. In AOLE, when all similar emotional states of the students are fused together, it should reflect the overall concentration and understanding of what they have learned. Consequently, similar to the formulations in [29] and [30], custom fuzzy rules are generated to map the levels of Valence and Arousal to the levels of Understanding and Concentration, according to the correlation mapping between each attribute (i.e., each axis) of the VA emotion space and attributes of the CA. From this mapping, 25 fuzzy rules (i.e., IF and THEN statements) are realised in the fuzzy inference system as presented in Appendix 1. Very importantly, the establishment of these rules builds on the assumption that

students are at their best behaviour devoid of any unruliness, i.e., no heightened emotions, as would be expected in normal classroom settings.

### 2.3 Visualisation of emotion atmosphere

The most important function of the proposed AOLE system is to provide a real-time reference for instructors while of teaching, and to monitor the level of students' understanding, concentration, and engagement in the class and their interaction with its contents. Therefore, in order to provide a more convenient and efficient display for instructors to make sense of the classroom atmosphere quickly and intuitively (even while immersed in their lesson) we need to visualise the identified classroom atmosphere. Generally, in engineering, visual discrimination is accomplished in terms of shape, colour, size, position, and direction [31]. Here, we adopt the shape-colour-length (SCL) visualisation to measure and monitor the three attributes of the classroom atmosphere as defined earlier. These attributes of our proposed visualisation of the emotion atmosphere are further explained in the rest of this section.

### 2.3.1 Shapes as cues for Understanding in AOLE

Intuitively, it has been suggested that there is a relationship between geometry and emotional atmosphere [22]. Generally, a circle (o) creates the impression of completeness, which connotes a positive emotion. Similarly, a cross (+) portends little feeling, which, for an optimist (i.e. one always expecting more) this could be deemed negative. Therefore, we adopt the use of “+” to indicate absence of understanding, while, at the other end, a “o” is used to indicate complete or total understanding. In between these extremes, the emotion representation varies as though edges of the “+” are progressively squished or rounded until a circle is realised as depicted in the top row of Fig. 4. Therefore, a “+” represents no understanding, i.e., a state 0, while “o” represents complete understanding denoted as state 1.

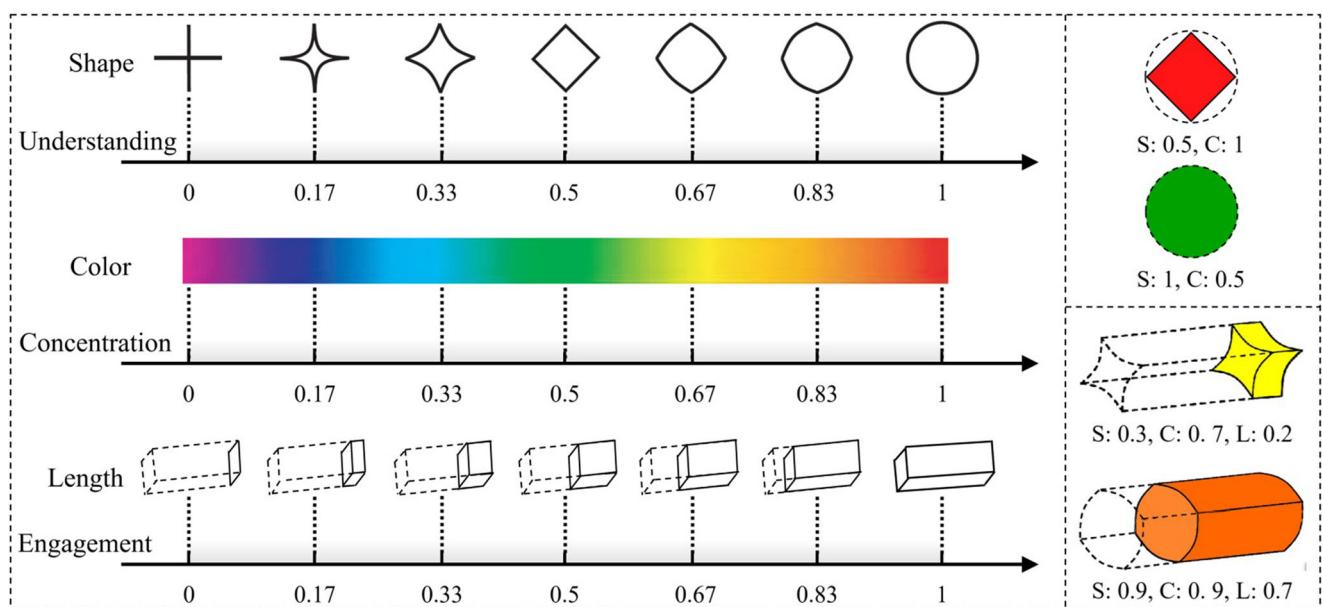
### 2.3.2 Colour as cues for Concentration in AOLE

Like shapes, colours and emotions are intricately linked. Warm colours can evoke different emotions than cool ones and bright colours can create different feelings than muted colours would. In this regard, red is generally viewed as an indication of heat while blue creates a sense of energy dispersion or coldness. Guided by this intuition, we adopt a nomenclature where the two extremes of a learners’ concentration are denoted using red (full concentration) also state 1, while values in the middle ground indicate transition in concentration and are denoted as state 0.

### 2.3.3 Length as cues for Engagement in AOLE

Intuitively, length is used to denote presence or absence of something. This intuition is widely used in emotional intelligence where the length of a shape indicates the magnitude of an emotion [32]. Here, we use length to denote the extent of learners’ engagement with activities in the classroom. A short length denotes low or no participation (i.e., values of 0) while a full length indicates maximum participation denoted by a value of 1.

Having described the notations used to represent the three emotional dimensions of our CA, in subsequent discussions we explore their use to represent real-time and actual emotional settings in OLE. First, we emphasize that, while engagement clearly monitors the learners’ engagement during a lesson, the subtlety of monitoring the learners’ understanding and concentration makes them more difficult to perceive. Furthermore, whereas using interactions with the learners, instructors could easily adjust engagement to increase learners’ participation, it is not that easy to adjust their concentration and understanding. Using our model, however, this can be alleviated via 2D graphics to simultaneously visualise degrees of concentration (colour) and understanding (shape). By adding the third dimension (i.e., engagement) mentioned earlier, our model provides an intuitive visualisation of understanding, concentration, and engagement in a classroom atmosphere (CA) as illustrated in the rightmost column in Fig. 4. Using this model, in the next section we report outcomes of efforts to monitor



**Fig. 4** Visualisation of three attributes of CA, i.e., understanding (S), concentration (C), and engagement (L), and illustrations of their use in two and three-dimensional spaces



learners' emotions in a classroom atmosphere typical of today's COVID influenced learning environments.

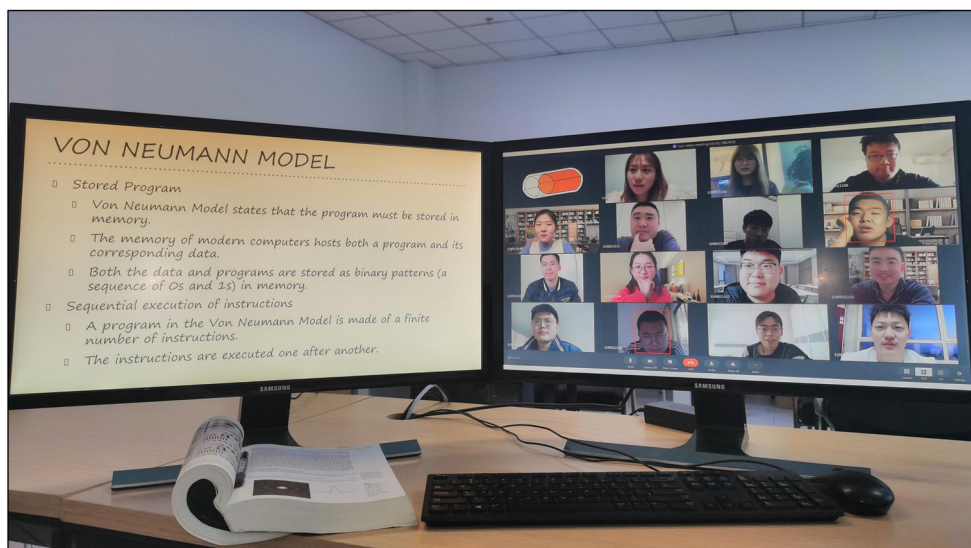
### 3 Deployment of proposed AOLE system

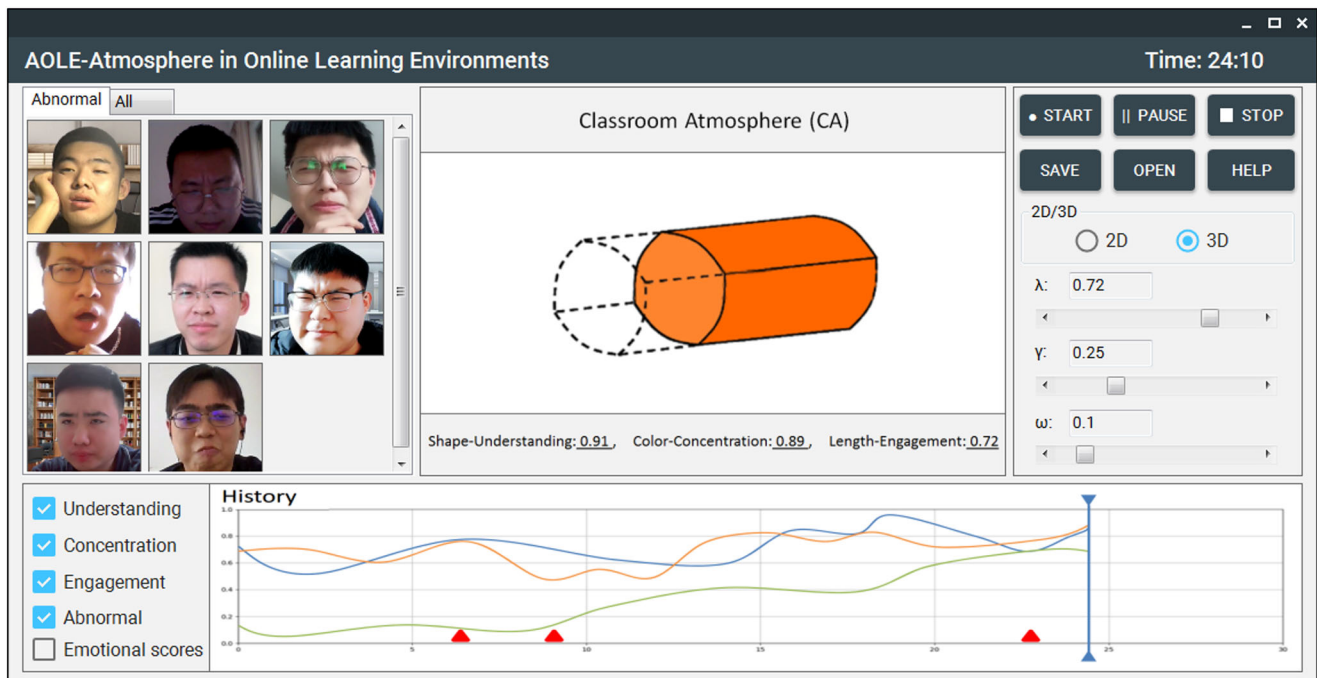
This section illustrates the deployment of the proposed AOLE model in a real-time classroom environment. By its design, AOLE is an independent system that can be integrated into other learning or conference management systems. Figure 5 presents an outline of the actual deployment of the proposed framework based on which the intrigues discussed in the remainder of this section were realised. In this example, Ding Talk, which is an enterprise-level intelligent mobile workspace for organisational management and operations (similar to Skype and Zoom) powered by Alibaba Group is used. However, since the advent of the COVID-19 pandemic, Ding Talk has gained widespread, seamless integration as an online education platform for different sizes and types of educational organisations [33]. As depicted in Fig. 5, and its integration as an intelligent, efficient, stable, and secure e-learning platform for classroom instruction is noted, where, for convenience, the screen is divided to simultaneously view all learners and the lecture materials. However, it is also noteworthy that in the AOLE model, face recognition is only used to identify the position of each learner's face from the screen and to further extract the facial data required for subsequent computation of emotions. Therefore, these image data are unaffected by changes in platforms used (i.e., if we change to other platforms, the algorithm is still applicable). As a result, the outcomes reported, and the general deployment of the proposed system are not affected by the choice of platform.

On their ends, the instructors load the AOLE system and simply click the start button whenever a computation and/or observation of the classroom atmosphere is required. On booting, the facial recognition unit locates the face of each learner and determines their facial expressions. As explained earlier in Section 2, these expressions facilitate generation of the classroom atmosphere (CA), which will be shown in the centre of the AOLE interface (in Fig. 6). Theoretically, the content comprehension, and classroom participation should show minimum deviation in the beginning and as the lesson progresses. Subsequently, the system discriminates learners exhibiting emotions that deviate from the class average. Such students are flagged and identified in the red box as presented in Fig. 5. Since this process is executed in real time, it is expected that this could change as the lesson progresses.

Moreover, these deviations could potentially affect the visualisation and computation of the overall emotion in the CA so their deviations should attract the attention of the instructor. The AOLE interface provides observable tools for tracking these deviated or abnormal expressions as presented in Fig. 6. The AOLE is user-friendly and supports easy-to-see representations of the individual emotions as well as the combined CA. Panes of the visualised atmosphere model could be minimised and dragged anywhere on the screen. For example, since there is no need for the instructor's facial expression in the calculation, it is expedient to place his/her visualised graphic at the top-left corner, i.e., which, by the design of the interface, is supposed to be the instructor's video window. Furthermore, by using the space bar to control the recognition process a one-click access is provisioned so that the 3D curves of the CA and details of the students exhibiting abnormal expressions are easily viewed. Furthermore, instantaneous

**Fig. 5** Real-time deployment of AOLE using Ding Talk platform to determine classroom atmosphere





**Fig. 6** Main interface of proposed AOLE system showing the visualised CA, faces of learners with abnormal expressions, the settings interface, and emotion curves of the atmosphere. The size and location of each pane or area can be adjusted as required

changes in these expressions can be visualised and saved for future comparative analysis. These functions are important as the instructor gets engrossed in the lesson.

Additionally, as observed earlier in Section 2, instructors could adjust individual learner parameters according to specific situations and needs. Examples of these scenarios include when it is feared that the individual emotion of one learner could affect the rest of the class or the combined CA (in such a case, his/her weight  $\omega$  needs to be reallocated), when continuity is threatened, such as the current atmosphere relying more on the current emotional states of all learners and less on the previous state of the atmosphere (whence, the parameter  $\lambda$  in (1) should be similarly adjusted) and so on. By using these functions and interfaces, the instructor has in his/her hands visualisation tools to monitor and control the classroom strategy as needed at any time, observe changes in the atmosphere, and readjust it when needed to enhance the teaching process.

### 3.1 Analysis of changes in emotion of individual learners

The visualisation of the classroom atmosphere (CA) can provide real-time guidance to help instructors regarding necessary adjustments to teaching strategies. From the instructor’s perspective, it is important to track the learner’s emotional changes as a lesson progresses or even throughout the duration of the semester, i.e., the learning cycle. The emotion curves of individual learners and visualisation

of the collective classroom atmosphere provide important reference points regarding expectations in performance assessment. This could be integrated into course reports and learning outcomes at the end of each semester. For example, reference to a learner’s progressive emotion curve could be used to support validation of eventual performance in assessments and examinations. To elucidate, it would be no surprise to see a learner performing below expectation in Understanding, Concentration, and Engagement throughout a course end up doing badly in the examination. In fact, the opposite should perplex the instructor. More importantly, instructors could use the emotional curve to identify such learners and tailor specific tasks such as additional lessons, homework, etc. to help them improve. In other cases, the emotional curves could provide indications of a learner’s psychological state, in which case further counselling and intervention could be sought for such learners.

The proposed AOLE system provides instructors with easy-to-use visualisation of each learner’s emotional curve which can be analysed at any instance within the learning cycle. This is accomplished via a scoring mechanism built to score each learner’s emotions during a class. To compute this emotional score, using the fuzzy rule sets in Appendix 1, we divide the VA model into 49 subareas, each composed of a range of emotion scores. Subsequently, for each subarea, the distance between the coordinate point and its origin is measured as its concrete score (e.g., when the point is located in the first quadrant, the longer distance indicates a higher score, while when it falls in the third

quadrant, the longer distance will conversely produce a lower score). The score is set in the range between 0 and 100 or graded as A, B, C, and so on. Since it is expected that the students will be in their best behaviour, exaggerated expressions (such as extremely scared) will not be expected; therefore, in most cases the emotions will be located in the middle areas of the circle rather than the four corners of the VA space. In other words, the emotional scores will usually not reach as high as 100 or low as 0. Therefore, correlating the emotion score with learner’s performance in the assessments, such as final exam, could provide inference regarding the learning state of different learners.

For our illustration of the deployment of the AOLE model, a class comprising of 15 learners (i.e., students) and three instructors is used. This size is typical of average university classrooms, which is also considered manageable for an average instructor even in F2F instruction [34, 35]. As presented earlier in Fig. 6, seven learners exhibited deviated or abnormal expressions. The emotion curves depicting some of these expressions are plotted in Fig. 7. Reading these curves, we can deduce that Learner S190522101 had low emotional score (in the range 10 to 20) throughout the class and the fluctuation in that learner’s emotional score is low, which suggests that the student may be distracted by some external factors leading to loss of focus on the lesson. In contrast, Learner S190522103 has an emotion score of around 40 during the interval 1000 to 1400 seconds, which corresponds to the period the instructor was explaining difficult parts of the lesson. This heightened or increased focus may be an indication of lack of comprehension of those difficult parts of the lesson. Further, the consistent emotional score of Learner S190522107 and decreasing scores for Learner S190522112 could indicate the need to flag such students for further attention from the instructor.

### 3.2 Estimation of identified classroom atmosphere

To assess CA computed using our proposed AOLE, the 3 instructors in our experiment were asked to evaluate the lesson (40 minutes) explained earlier as it progressed. To suppress the impact of disparities arising from learner’s behaviour at the beginning and end of the lesson (e.g., settling down at the beginning of the class and anxiousness towards the end) only the middle 30 minutes of the lesson was used in the assessment. All the instructors are experienced in terms of teaching evaluation, and they had prior experience with the course content, including difficult areas of the lesson. Furthermore, to familiarise themselves with its use, the instructors had practiced using the AOLE system prior to the reported lesson and they were introduced to rudiments of emotion recognition and atmosphere generation in the AOLE context.

While not delivering the lesson, two of the three instructors were requested to listen to the lesson and observe changes in learning status of the learners such as changes in their facial expressions and engagement. Further, they were asked to observe the learning status at specific periods, notably, when the teaching instructor was explaining difficult areas of the lesson, giving easy examples, and engaging the students via questions. Table 2 presents the questionnaire administered to these none-teaching “observant” instructors. They were asked to assess the communication atmosphere on a scale 1 through 7. To quantitatively analyse the outcomes from the observant instructors, each response was assigned a numerical value 1(0), 2(0.17), 3(0.33), 4(0.5), 5(0.67), 6(0.83) and 7(1). Considering the difficulty of aggregating these scores with those of the CA, correlation analysis using Pearson’s correlation was employed. Relative to the three key emotions used in AOLE,



Fig. 7 Illustration of emotional scores in VA space

**Table 2** Questionnaire for evaluation of communication atmosphere

Questions	Answers
How much do you think the students understood the contents of the lesson?	1-Extremely low, 2-Very low, 3-Low, 4-Average, 5-High, 6-Very high, 7-Extremely high
How much do you think the students concentrated during the lesson?	1-Extremely low, 2-Very low, 3-Low, 4-Average, 5-High, 6-Very high, 7-Extremely high
How much do you think the students engaged during the class?	1-Extremely low, 2-Very low, 3-Low, 4-Average, 5-High, 6-Very high, 7-Extremely high

correlation scores of  $\Gamma_U=0.75$ ,  $\Gamma_C=0.82$ , and  $\Gamma_E=0.96$  were obtained for Understanding, Concentration, and Engagement, respectively. These outcomes are indicators that the CA in AOLE is consistent with the subjective assessment via the administered questionnaire.

Possible reasons for the above results could be: first, the correlation coefficient for the “Engagement” axis is greater than the others, since engagement in discussions and answering questions, etc. is easier to observe by an observing instructor. Second, the correlation coefficient for the “Understanding” axis is the least because sometimes it is difficult to read the learners’ mind to discern whether they comprehended contents of the lesson or gained knowledge based on their facial expression (moreover, they may also pretend to understand what the instructor explained (such as by smiling or nodding) even when they do not). Notwithstanding, in all the above cases, the analysis suggests that AOLE system is an effective tool capable of assisting the instructor in teaching and training students on the OLE platforms.

### 3.3 Comparison between proposed AOLE model and affective tutoring systems

An intelligent tutoring system (ITS) is a computer-based educational system that aims to provide immediate and customised instruction or feedback to learners, usually without requiring any intervention from a human instructor [36]. Similar to a personal tutor, the ITS continuously interacts with the tutor and makes assessments of the learners’ progress to enhance effectiveness. In the past, researchers’ main criticism of ITS is that they were devoid of emotional awareness and empathy, which they argued limits the effectiveness of the tutoring provided. It is this significance of incorporating emotional states into the learning process that motivated the development of affective tutoring systems (ATS) is an extension of the ITS. Therefore, with “affectiveness” suffused into ITS, the

resulting ATSs supposedly sense the emotional state of a learner and then intelligently suggests appropriate strategies that can guide the learning process and ultimately shift the negative attitude of students toward enhanced course and content learning.

As the most advanced system similar to our proposed AOLE model, we provide below a rudimentary comparison between our AOLE system and typical ATS platforms, such as those listed in Table 3.

Guided by the feature comparisons in Table 3 and the discussions in earlier sections, we note that:

- (1) Fundamentally, whereas the development of ATS was motivated by the need to infuse or supplement traditional F2F teaching platforms with capability to recognise personal emotions and develop personalised tutoring programmes, on its part, interest in AOLE is motivated by the inevitable transition to remote, online, and blended learning. It equips instructors with a tool for real-time monitoring of the emotional states of all students during lessons. In this manner, by identifying and visualising the classroom atmosphere (CA), in AOLE, instructors should be able to adjust their teaching strategies as and when needed.
- (2) From users’ perspectives, ATS aims to provide immediate and customised guidance or feedback without need for the intervention of teachers, which makes it student focused. In contrast, AOLE is an add-on system that requires an online learning platform (such as Skype, Zoom, or Ding Talk) to deliver teaching activities specified in the syllabus. Therefore, it could be considered more teacher-centred with focus on delivering instruction especially during current COVID restrictions to regular F2F classroom interactions.
- (3) From the application viewpoint, although the ultimate aim of both the proposed AOLE and the ATS platforms reported in Table 3 is to improve learning efficiency, ATS is more tailored towards reducing learners’ negative emotions, promoting their interest

**Table 3** Comparison between proposed AOLE model and ATS systems

Feature Model	Data type	Target users	Usage	System feedback	Target
EDTS [37]	Text, Speech, Facial expressions	Single student	Emotional design course	Transform students' emotional changes to a series of expressive cartoon pictures (cat) for teacher's analysis	Assisting instructors in teaching and improving students' learning motivation
Fwa's model [38]	Text, Speech, Facial expressions	Single student	Programming exercise	Provide solution tips to guide students to correct mistakes in assessments administered	Regulating the negative emotions of students
ATS-JP [39]	Text, Speech, Facial expressions	Single student	Japanese learning	Offer specific feedback when learners complete a stage of the learning	Improving the students' interest in learning
Lin's model [40]	Text, Facial expressions	Single student	Account remedial curriculum	Display expressive cartoon pictures (fog) to ease and relax learners when their negative emotions are detected	Regulating the negative emotions of students
ATEN [41]	Text, Speech, Physiological signals	Single student	Built environment management	Integrate the student's personal learning style with available materials for appropriate learning adaptation	Increasing the understanding and productivity of students
PAT2Math [42]	Mouse action, Facial expression	Single student	Mathematics learning	Show the "thumb up" or "thumb down" symbols and give students the tips when they encounter difficulties	Giving students timely feedback when solving problems
Proposed AOLE	Speech, Facial expressions	Instructor	As an add-on unit that is independent of the type of course or online platforms	Visualisation of classroom atmosphere, recording and monitoring of emotion scores of individual learners each student's emotional scores	Evaluating the emotional atmosphere of the classroom to provide reference to instructors for enhanced teaching and learning

in learning, or enhancing their learning experience via computer applications that are usually course-specific. In contrast, AOLE provides a visualisation of the classroom atmosphere for the instructor’s reference. It is mainly intended to provide them with the guidance needed to adjust their teaching strategies, timely identification of learners with emotional distress that may require intervention, in and out of the classroom.

### 3.4 Comparison of different algorithms applied in the FER layer

Although this study is not intended as an addition to the remarkable literature on facial emotion recognition (FER), its use of FER as one of the three layers of our AOLE platform makes it worthwhile to present a comparison between different FER algorithms. Nevertheless, we reiterate that AOLE is not simply an FER algorithm, but, instead, its three layers coalesce into an intuitive model to monitor the concentration, understanding, and engagement expected of a productive online classroom environment. Therefore, from the perspective of the whole system, to the best of our knowledge, there is no like-for-like system to base our comparison. However, as presented in Table 3, empirically, in terms of learning platforms, our AOLE can be compared with other emotion-based online learning systems, i.e., affective tutoring systems (ATSS). The table highlights differences and advantages of our AOLE system in terms of design motivation, users’ perspectives, and applications viewpoint.

This empirical analysis, which is a way of gaining knowledge by means of direct and indirect observation or experience provides basis to use empirical evidence (i.e., the record of one’s direct observations or experiences) in quantitative or qualitative analysis [39]. Quantifying the evidence or making sense of it in qualitative form, a researcher can answer empirical questions, which should be clearly defined and answerable from the evidence collected (usually, called data). Moreover, many researchers

combine qualitative and quantitative forms of analysis to better answer questions that cannot be studied in laboratory settings. Therefore, similar to the studies enumerated in Table 3, to establish the performance of our AOLE system, we presented (in Section 3.2) a rudimentary evaluation using a case study (i.e., the empirical analysis from a class environment comprising of 15 students and 3 instructors).

Notwithstanding the clarification of the main scope of our study presented here, to further illustrate the feasibility and effectiveness of our proposed model, we present a quantitative comparison of our FER layer relative to other face recognition algorithms. As we discussed earlier, the main purpose for our use of FER is to map facial expressions to values in the VA coordinate system. More specifically, this process consists of two steps, i.e., facial expression recognition and emotional coordinate mapping (as presented earlier in Fig. 2). Table 4 presents a comparison between our adopted FER technique, i.e., the Deep-emotion deep learning model (in [17]) and other FER algorithms (in the first column of Table 4).

As emphasized in Section 2, in order to realise the interface communication between layers of our AOLE system, the outcome from the FER layer is not the classification of emotion, but the coordinate points in the Valence-Arousal (VA) axes. Therefore, the most effective quantitative comparison of the FER algorithms would be in terms of their effectiveness to provide the required mapping of the facial emotions to the values of the VA coordinate system. Consequently, to establish the accuracy (ACC) of each FER algorithm in mapping facial expression into VA coordinate system in our model, we compute the Euclidean distance (denoted as DIS) of the prediction data and ground truth including the distance along the X-axis (Valence) and Y-axis (Arousal) as presented in Table 4. All reported results are based on the data and experimental settings highlighted earlier in Section 2.1.

We complete this analysis by noting that, from Table 4, the adopted Deep-emotion FER algorithm exhibits better average performance in mapping the facial expressions into

**Table 4** Comparisons of the use of different algorithms in the FER layer of proposed the AOLE platform (i.e., as FER layer in Fig. 2)

Method	Valence		Arousal		Length		Average accuracy (%)
	DIS	ACC	DIS	ACC	DIS	ACC	
AlexNet [43]	0.194	80.6%	0.206	79.4%	0.301	69.9%	76.6
VGG [44]	0.156	84.4%	0.195	80.5%	0.273	72.7%	79.2
GoogleNet [45]	0.145	85.5%	0.164	83.6%	0.242	75.8%	81.6
ResNet-18 [46]	0.210	79.0%	0.147	85.3%	0.276	72.4%	78.9
ResNet-152 [47]	0.182	81.8%	0.113	88.7%	0.232	76.8%	82.4
Adopted Deep-emotion FER algorithm [17]	0.116	88.4%	0.116	88.4%	0.184	81.6%	86.1

the VA coordinate system of our AOLE model. Moreover, results reported in [17] attribute increases of 4 and 5% in classification accuracy for the FER 2013 dataset than the VCG and GoogleNet. The authors attribute this to the focus of the model on parts of the face. Consequently, based on the outcomes in Table 4, we can conclude that the choice and adoption of the Deep-emotion FER algorithm as the FER layer in our AOLE model is validated.

## 4 Concluding remarks and future perspectives

Online education has continued to benefit from advances in engineering infrastructure, internet and information technology, and, more recently, artificial intelligence. Despite its recorded growth in size and efficiency, the unplanned transition to online and virtual learning necessitated by the COVID-19 pandemic has exposed many areas of online education that need to be enhanced. The proposed AOLE system is conceived as an add-on feature aimed at improving existing online learning environments (OLEs), especially in the face of the unprecedented challenges caused by COVID. AOLE provides a bridge between the flexibility of online learning in that by using an Internet connection, camera, and microphone, learners can join classes remotely from anywhere in the world as well as interactions that facilitate learner-centered engagements that are available in traditional F2F learning. The latter entails understanding nuances around how instructors read, monitor, and adjust lessons based on individual peculiarities of the learners as well as the collective learning atmosphere of the class. In this regard, AOLE integrates essential features of both the online and F2F learning paradigms.

In order to further summarise the contributions emanating from our study, we present the following conclusions. First, AOLE is modeled as an intuitive add-on tool to monitor the emotional interplay underpinning learners' concentration, understanding, and engagement in online learning environments. Second, the proposed model supports tracking the progress of an individual learner via visualisation of his/her emotion curve at different stages of the learning cycle. Third, emotion curves are used to progressively monitor learner emotions as the learning cycle progresses and guided by it, instructors could adjust their content delivery. Fourth, experimental results reported suggest the potential of the proposed model in supporting learning and counselling during these unprecedented times.

While we have demonstrated both the practicality and utility of the proposed AOLE system, many aspects of it still need improvement. Consequently, in short- and long-term efforts, we plan to overhaul the system in the following directions. First, each of the three layers of the system

is being improved for enhanced functionality. For example, the FER algorithm is being updated to improve results of the emotion recognition. We are also exploring how to reformulate the visualisation mode with a view towards providing enhanced experiences for all users. Second, similar to the intuition in quantum mechanics, where momentum of microparticles and their position cannot be simultaneously measured [48], it is possible that recognition errors could be involved in the emotion recognition system when both facial and spoken information are considered simultaneously. Therefore, for the enhanced AOLE model, we are exploring optimal adjustments to preset the compensation function targeted at decreasing the likelihood of such errors. Third, with improvement in quality and quantity of data, machine learning and empirical analysis could be exploited to enhance the performance of the proposed AOLE model. Fourth, as illustrated in this study, we envision AOLE as a plug-in system that can smoothly coalesce with existing OLE and learning tools [49]. Therefore, it is important to enhance seamlessness of data collection and overall user-friendliness of the system. Fifth, among the major shortcomings of existing OLE systems is the difficulty around its use in laboratory and practical instruction for students of engineering, medicine, and other physical science disciplines [50]. However, it is envisioned that by fusing other emotion recognition methods that are based on other non-verbal and written inputs (such as gestures, text, etc.) as well as virtual and augmented reality, our AOLE system could potentially provide learners with better experiences in such courses than is available presently. Sixth, while it will mostly be used for the good it is designed for (such as in providing identified learners with timely needed counselling based on their emotional states) the gathered emotion data could also be exploited in many negative ways. As private information of the learners, their emotional states must be treated confidentially [51]. Therefore, as OLE develops and platforms such as AOLE are deployed, protocols similar to those between doctors and patients, lawyers and clients, etc. must be developed to safeguard the confidentiality of learners' emotional wellbeing. Seventh, instructors must be encouraged to see the classroom atmosphere (CA) as reflections of physical and traditional learning environment and process. Therefore, they should use it to complement course material, especially as it pertains to identifying, isolating, and explaining important and difficult areas of the curriculum.

While optimistically looking forward to the post-COVID era, we share the view held by many that the educational landscape has been altered forever and that some form of OLE or the other will always be part of the teaching-learning process. With this in mind, as outlined in this study and our concluding remarks, we envision roles for platforms such as the proposed AOLE in future learning landscapes.

## Appendix 1

The twenty-five (25) rules in Table 5 are used to map “Valence” and “Arousal” into “Understanding” and “Concentration” in accordance with the membership functions described earlier in Table 1.

**Table 5** Rules used in the fuzzy inference system of proposed our AOLE system

Rule	Valence		Arousal		Understanding		Concentration	
1		VHV		VHA		EHU	EHC	
2		VHV		HA		VHU	VHC	
3		VHV		AA		VHU	AC	
4		VHV		LA		VHU	LC	
5		VHV		VLA		HU	LC	
6		HV		VHA		VHU	HC	
7		HV		HA		HU	HC	
8		HV		AA		HU	AC	
9		HV		LA		HU	LC	
10		HV		VLA		AU	LC	
11		AV		VHA		HU	HC	
12	IF	AV	AND	HA	THEN	AU	AND	HC
13		AV		AA		AU		AC
14		AV		LA		AU		LC
15		AV		VLA		LU		LC
16		LV		VHA		AU		HC
17		LV		HA		LU		HC
18		LV		AA		LU		AC
19		LV		LA		LU		LC
20		LV		VLA		VLU		LC
21		VLV		VHA		LU		HC
22		VLV		HA		VLU		HC
23		VLV		AA		VLU		AC
24		VLV		LA		VLU		VLC
25		VLV		VLA		ELU		ELC

**Funding** This study is sponsored in full by the Prince Sattam Bin Abdulaziz University, Saudi Arabia via the Deanship for Scientific Research funding for the Advanced Computational Intelligence and Intelligent Systems Engineering (ACIISE) Research Group Project number 2020/01/12173.

## References

- Quezada RL, Talbot C, Quezada-Parker KB (2020) From bricks and mortar to remote teaching: a teacher education programme’s response to COVID-19. *J Educ Teach* 46(4):472–483
- Anthony R, Artino J (2012) Emotions in online learning environments: introduction to the special issue. *Internet Higher Educ* 15(3):137–140
- Dhawan S (2020) Online learning: a panacea in the time of COVID-19 crisis. *J Educ Technol Syst* 49(1):5–22
- Sadeghi M (2019) A shift from classroom to distance learning: advantages and limitations. *Int J Res Engl Educ* 4(1):80–88
- Chen T, Peng L, Jing B, Wu C, Yang J, Cong G (2020) The impact of the COVID-19 pandemic on user experience with online education platforms in China. *Sustainability* 12(18):7329
- Bada S (2015) Olusegun, Constructivism learning theory: A paradigm for teaching and learning. *IOSR J Res Method Educ* 5(6):66–70
- Salmon G (2004) *E-moderating: the key to teaching and learning online*. Routledge Falmer, London
- King RB, Chen J (2019) Emotions in education: asian insights on the role of emotions in learning and teaching. *Asia-Pacif Educ Res* 28:279–281
- Dores AR, Barbosa F, Queirós C, Carvalho IP, Griffiths MD (2020) Recognizing emotions through facial expressions: a largescale experimental study. *Int J Environ Res Publ Health* 17(20):7420



10. Ooia CS, Seng KP, Ang L.-M., Chew LW (2014) A new approach of audio emotion recognition. *Expert Syst Appl* 41(13):5858–5869
11. Erenel Z, Adegboye OR, Kusetogullari H (2020) A new feature selection scheme for emotion recognition from text. *Appl Sci* 10(15):5351
12. Saganowski S, Dutkowiak A, Dziadek A, Dziezyc M, Komoszynska J, Michalska W, Polak A, Ujma M, Kazienko P (2020) Emotion recognition using wearables: A systematic literature review-work-in-progress. In: *IEEE International Conference on Pervasive Computing and Communications Workshops*, pp 1–6
13. Wang W, Xu K, Niu H, Miao X (2020) Emotion recognition of students based on facial expressions in online education based on the perspective of computer simulation. In: *Complexity problems handled by advanced computer simulation technology in smart cities, complexity*
14. Zhang Z, Li Z, Liu H, Cao T, Liu S (2020) Data-driven online learning engagement detection via facial expression and mouse behavior recognition technology. *J Educ Comput Res* 58(1):63–86
15. Tonguc G, Ozkara BO (2020) Automatic recognition of student emotions from facial expressions during a lecture. *Comput Educ* 148:103797
16. Megahed M, Mohammed A (2020) Modeling adaptive E-learning environment using facial expressions and fuzzy logic. *Expert Syst Appl* 157:113460
17. Minaee S, Minaei M, Abdolrashidi A (2021) Deep-emotion: facial expression recognition using attentional convolutional network. *Sensors* 21(9):3046
18. Mukhopadhyay M, Pal S, Nayyar A, Pramanik PKD, Dasgupta N, Choudhury P (2020) Facial emotion detection to assess learner's state of mind in an online learning system. In: *Proceedings of the 5th International Conference on Intelligent Information Technology*, pp 107–115
19. Pekrun R (2014) Emotions and learning. *Int Acad Educ* 24(1):1–31
20. Jaderberg M, Simonyan K, Zisserman A, Kavukcuoglu K (2015) Spatial transformer networks, in: *Proceedings of the 28th International Conference on Neural Information Processing Systems*, pp 2017–2025
21. Meryl CJ, Dharshini K, Juliet DS, Rosy JA, Jacob SS (2020) Deep learning based facial expression recognition for psychological health analysis. In: *Proceedings of the International Conference on Communication and Signal Processing (ICCSP)*, pp 1155–1158
22. Yan F, Iliyasa AM, Hirota K (2021) Emotion space modelling for social robots. *Eng Appl Artif Intell* 100:1–25
23. Sokolov EN, Boucsein W (2000) A psychophysiological model of emotion space. *Integr Physiol Behav Sci* 35(2):81–119
24. Tim T, Hofling A, Gerdes ABM, Fohll U, Alpers GW (2020) Read my face: automatic facial coding versus psychophysiological indicators of emotional valence and arousal. *Front Psychol* 11:1388
25. Takac P, Mach M, Sincak P (2016) Cloud-based facial emotion recognition for real-time emotional atmosphere assessment during a lecture. In: *IEEE International Conference on Systems, Man, and Cybernetics*, pp 001696–001701
26. Liu Z, Wu M, Li D, Chen L, Dong F, Yamazaki Y, Hirota K (2013) Concept of fuzzy atmosphere for representing communication atmosphere and its application to humans-robots interaction. *J Adv Comput Intell Intell Inf* 17(1):3–17
27. Yan F, Iliyasa AM, Hirota K (2021) Conceptual framework for quantum affective computing and its use in fusion of multi-robot emotions. *Electronics* 10:100
28. Hirota K, Dong F (2013) Concept of fuzzy atmosphere and its visualization. In: *On Fuzziness*, vol 298. Springer, Berlin, pp 257–263
29. Vashishtha S, Susan S (2019) Fuzzy rule based unsupervised sentiment analysis from social media posts. *Expert Syst Appl* 138:112834
30. Zhang X, Yang X, Zhang W, Li G, Yu H (2021) Crowd emotion evaluation based on fuzzy inference of arousal and valence. *Neurocomputing* 445:194–205
31. Gallagher RS (1994) *Computer visualization: graphics techniques for engineering and scientific analysis*. CRC press, London
32. Yan F, Iliyasa AM, Jiao S, Hirota K (2019) Quantum structure for modelling emotion space of robots. *Appl Sci* 9:3351
33. Alibaba G Ding Talk for online Education, <https://www.alibabacloud.com/zh/solutions/dingtalk/online-education>
34. Wang W, Xu K, Niu H, Miao X (2020) Emotion recognition of students based on facial expressions in online education based on the perspective of computer simulation. *Complexity*
35. Yang D, Alsadoon A, Prasad PC, Singh AK, Elchouemi A (2018) An emotion recognition model based on facial recognition in virtual learning environment. *Procedia Comput Sci* 125:2–10
36. Graesser AC, Chipman P, Haynes BC, Olney A (2005) Autotutor: an intelligent tutoring system with mixed-initiative dialogue. *IEEE Trans Educ* 48(4):612–618
37. Wang C, Lin HK (2018) Emotional design tutoring system based on multimodal affective computing techniques. *Int J Distance Educ Technol* 16(1):103–117
38. Fwa HL (2018) An architectural design and evaluation of an affective tutoring system for novice programmers. *Int J Educ Technol Higher Educ* 15(1):38
39. Lin HK, Chao C, Huang T (2015) From a perspective on foreign language learning anxiety to develop an affective tutoring system. *Educ Technol Res Dev* 63(5):727–747
40. Lin HK, Wu C, Hsueh Y (2014) The influence of using affective tutoring system in accounting remedial instruction on learning performance and usability. *Comput Hum Behav* 41:514–522
41. Kaklauskas A, Kuzminskas A, Zavadskas EK, Danišnas A, Kaklauskas G, Seniut M, Raistenskis J, Safonov A, Kliukas R, Juozapaitis A, Radzeviciene A (2015) Affective tutoring system for built environment management. *Comput Educ* 82:201–216
42. Jaques PA, Seffrin H, Rubi G, de Morais F, Ghilardi C, Bittencourt II, Isotani S (2013) Rule-based expert systems to support step-by-step guidance in algebraic problem solving: The case of the tutor pat2math. *Expert Syst Appl* 40(14):5456–5465
43. Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems (NIPS)*, pp 1106–1114
44. Simonyan K, Zisserman A Very deep convolutional networks for large-scale image recognition, arXiv:1409.1556
45. Giannopoulos P, Perikos I, Hatzilygeroudis I (2018) Deep learning approaches for facial emotion recognition: a case study on FER-2013. *Adv Hybrid Intell Methods* 85:1–16
46. Meng D, Peng X, Wang K, Qiao Y (2019) Frame attention networks for facial expression recognition in videos. 2019 IEEE International Conference on Image Processing (ICIP), pp 3866–3870
47. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp 770–778
48. Robertson HP (1929) The uncertainty principle. *Phys Rev* 34(1):163
49. Wang S, Chen Y (2018) Rain classroom: a tool for blended learning with MOOCs. In: *Proceedings of the Fifth Annual ACM Conference on Learning at Scale*, pp 1–2
50. Boettcher JV, Conrad R-M (2016) *The online teaching survival guide: simple and practical pedagogical tips*. Wiley, Hoboken
51. Cole EV, Cawthon SW (2015) Self-disclosure decisions of university students with learning disabilities. *J Postsecond Educ Disab* 28(2):163–179

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.