










## Evaluating the Impact of Emotions and Awareness on User Experience in Virtual Learning Environments for Sustainable Development Education

Amer Ibrahim<sup>1</sup>, Intisar A.M. Al Sayed<sup>2</sup>, Mohanad Sameer Jabbar<sup>3</sup>, Hissah Almutairi<sup>4</sup>, Ravi Sekhar<sup>5\*</sup>,  
Pritesh Shah<sup>5</sup>, Israa Ibraheem Al\_Barazanchi<sup>6</sup>

<sup>1</sup> College of Computer and Information Technology, American University in the Emirates, Dubai 503000, UAE

<sup>2</sup> Faculty of Technical Engineering, Uruk University, Baghdad 10001, Iraq

<sup>3</sup> Medical Instruments Techniques Engineering Department, Technical College of Engineering, Al-Bayan University, Baghdad 10001, Iraq

<sup>4</sup> College of Engineering and Computer Science, Prince Sattam Bin AbdulAziz University, Al-Kharj 11942, Saudi Arabia

<sup>5</sup> Symbiosis Institute of Technology (SIT) Pune Campus, Symbiosis International (Deemed University) (SIU), Pune 412115, Maharashtra, India

<sup>6</sup> Department of Communication Technology Engineering, College of Information Technology, Imam Ja'afar Al-Sadiq University, Baghdad 10001, Iraq

Corresponding Author Email: [Ravi.sekhar@sitpune.edu.in](mailto:Ravi.sekhar@sitpune.edu.in)

Copyright: ©2024 The authors. This article is published by IETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/isi.290108>

### ABSTRACT

**Received:** 13 November 2023

**Revised:** 24 November 2023

**Accepted:** 1 February 2024

**Available online:** 27 February 2024

#### Keywords:

*computer science, network, virtual learning environment, human emotions, and awareness, Fuzzy-based Convolutional Neural Network, histogram equalization, sustainable development education, environmental education*

This study aims to evaluate the impact of emotions and awareness on user experience in Virtual Learning Environments (VLEs). VLEs have become increasingly popular in educational settings, but they have limitations that can negatively affect the learning experience. To address this issue, a novel Fuzzy-based Convolutional Neural Network (FCNN) is proposed for effective emotion evaluation. The study utilizes questionnaire surveys to collect data on awareness of the VLE. The performance of the FCNN method is evaluated based on accuracy, sensitivity, specificity, and precision. The study acknowledges the limitations of VLEs but does not specify them. The sample size and participant details are not mentioned in the abstract. Overall, this study provides insights into the role of emotions and awareness in enhancing the user experience of VLEs.

## 1. INTRODUCTION

The Virtual Learning Environment (VLE) has gained momentum in recent years and has become increasingly integral to college and school activities around the globe as a response to the Covid-19 epidemic. VLEs refer to educational settings where the instructor and the student are detached either by time or distance, or both. The use of Information Technology (IT) solutions, multimedia materials, the Internet, webcam, etc. are used to provide the learning materials. The idea behind virtual learning was originally to offer access to higher education outside of traditional campus settings. VLEs have been shown to have several benefits, including increased flexibility, accessibility, and cost-effectiveness. However, despite these benefits, VLEs have limitations that can negatively affect the learning experience. These limitations include technical issues, lack of social interaction, and difficulty in building trust among students. As a result, students in VLE lectures may struggle to comprehend the fundamental ideas underlying the topic. The specific research gap or problem that this study aims to address is the impact of emotions and awareness on user experience in VLEs. While

existing research has explored the role of emotions in Virtual Learning Environments, there are limitations to this research. For example, some studies have focused on a limited range of emotions or have used self-report measures that may not accurately reflect the emotions experienced by students. This study aims to address these limitations by using a novel Fuzzy-based Convolutional Neural Network (FCNN) method for emotion evaluation and by collecting data on awareness of the VLE. When discussing the merits of Virtual Learning Environments in Figure 1, it would be helpful to include specific references or studies that have demonstrated these benefits. For example, a study by Rashid et al. [1] found that VLEs can improve students' academic performance and satisfaction. Another study by Dung [2] found that VLEs can enhance students' engagement and motivation. Including these references would strengthen the credibility of the claims made in Figure 1. In summary, this study aims to evaluate the impact of emotions and awareness on user experience in VLEs. The proposed FCNN method and questionnaire surveys provide a novel approach to address the limitations of existing research in this field. By improving the user experience of VLEs, this study has the potential to enhance the learning outcomes of

students in Virtual Learning Environments.

Several academics have made an effort to model the association between student accomplishment in conventional and online learning environments and VLE emotions including pleasure, boredom, ambiguity, and stress in recent years. It should be emphasized that the phrases learning-related emotion and achievement-related emotion are used interchangeably in this research. Achievement emotions, as articulated in many frameworks, have a direct impact on user experience by influencing concentration, desire, and self-control [3-5]. Consider bringing awareness-related elements to the virtual learning environment (VLE) it will enhance the user experience in VLE. The subject of instructor and student awareness is crucial to virtual educational perceptions. To successfully assist learning practices, a teacher or student must be aware of the instructing and training process and be aware of nearly all of the other team participants' actions. These awareness processes can be crucial for enhancing group dynamics. The impact of VLEs on academic activities is even

more important for everyone to be aware of VLE [5, 6]. Since educational institutions increasingly place a premium on students' satisfaction and involvement in their academic activities, it is essential to comprehend how emotions and awareness influence students' perspectives and user experience in VLE.

### 1.1 Contribution of the study

- In the virtual learning environment, we developed a novel Fuzzy-based Convolutional Neural Network (FCNN) for effective emotion evaluation. By employing a questionnaire survey, the awareness of the virtual learning environment is evaluated.

- The preprocessing of facial images makes utilization of histogram equalization. With the help of the Discrete Cosine Transform (DCT), the high-level features are retrieved.

The suggested approach is examined using several parameters and compared to conventional methods.

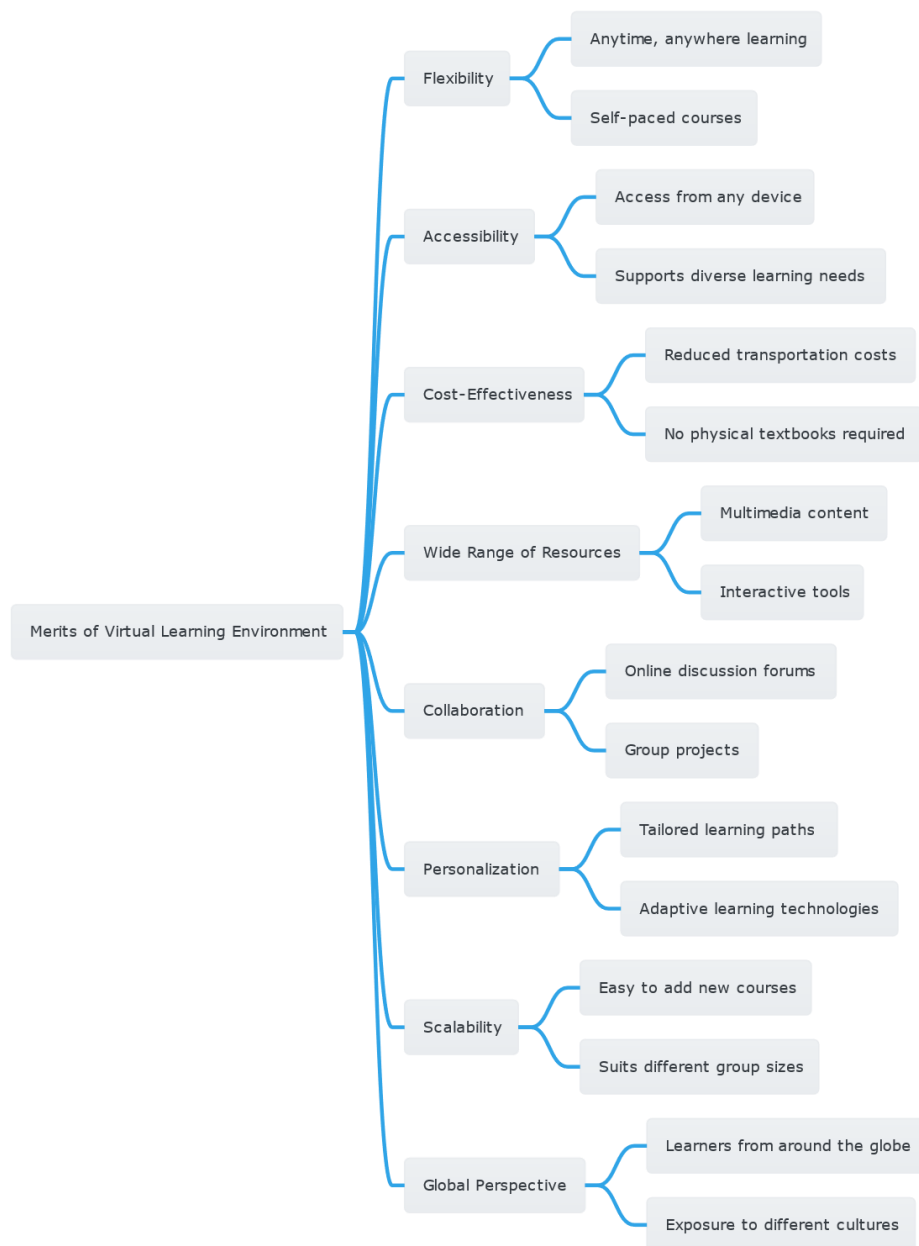


Figure 1. Merits of virtual learning environment

## 2. LITERATURE SURVEY

Aggarwal et al. [7] suggested using a Semantic Graph-based Dual-Stream Network (GBDSN) for face emotion identification. This network designs a graph depiction to simulate key image and dimensional facial adjustments as well as their conceptual connections with layered graph Convolutional attention blocks. It has low recognition accuracy. Liao et al. [8] presented FedFace, Federated Learning (FL) framework for jointly developing face emotion recognition systems in a privacy-conscious way, along with a feature initializing strategy for the class immersion and a spreading out regularize to guarantee the class embedding segregation. Training requires high-quality images; otherwise, it will produce poor recognition. The study by Miao et al. [9] recommended using a Deep Face Spatiotemporal Network (DFSTN) to identify the emotions conveyed by facial photographs in the context of online education. The extreme difficulty is inherent in the process of modal training for face image. Zhang et al. [10] developed a face detection technique to be employed in face emotion detection technology in an addition to identifying the face and intercepting the emotion data. Additionally, it estimates the growth rate in virtual learning. The analysis of a person's facial emotions requires a significant amount of time. Verulkar and Bhurchandi [11] advocated reinforcement learning (RL) and domain knowledge for multi-modal emotion detection of facial images. Domain knowledge utilizes emotion pairs to update recognition results, while RL uses face images to recognize emotions. An excessive amount of reinforcement learning may result in an excessive number of states, which can lead to a reduction in the quality of the outcomes. The study by Padmashree et al. [12, 13] recommended using autoluminance as a strategy for improving the emotion identification performance of face classification systems using images of children's faces. It requires a significant amount of time. The study by Kim et al. [14] offered Haar classifiers because of their high recognition performance and local binary pattern (LBP) classifiers owing to their stable structure under different circumstances to identify emotions based on face photographs. Both of these classifiers were developed by him. They generate histograms that are rather lengthy, which slows down the identification performance, particularly when used in large-scale facial image databases. You [15] introduced local phase quantization and support vector machines as methods for classifying face images and determining the emotion being expressed. It has greater complexity than the little abnormalities that occur in the network. Kuruvayil and Palaniswamy [16] recommended the utilization of enhanced characteristics of pretrained facial emotion identification and training derived from the Xception algorithm of facial images for facial image threshing. It contains a significant number of layers, which results in the overlap of data and contributes to the complexity of the performance. Yadav [17] developed the Discrete Lion Swarm Optimization Algorithm (DLSA) and the Deep Belief Network (DBN) to recognize facial expressions of emotion using images of the face. Because of its intricate data models, DBN requires a significant expenditure to train. As a result, the disadvantages that were discussed before are the reason why the emotion identification system is not very accurate. As a result of the aforementioned challenges associated with identifying emotions based on face images, we developed a unique Fuzzy-based Convolutional Neural Network (FCNN) for efficient emotion assessment [18].

## 2.1 Problem statement

In many different types of educational settings, the utilization of Virtual Learning Environments is advantageous [19-22]. Despite this, it has drawbacks such as an absence of trust, decreased group interaction, social separation, and concealed scientific costs. Because of this, the students in VLE lectures are unable to comprehend the fundamental ideas underlying the topic. We must work to improve the user experience of the VLE in an attempt to address problems of this kind. Therefore, it is necessary to evaluate human emotions and awareness in an attempt to improve the overall quality of the user experience in Virtual Learning Environments.

## 3. PROPOSED METHOD

Virtual Learning Environments (VLE) are emerging in the educational system. Hence, it is essential to comprehend how emotions and awareness influence students' perspectives and user experience in VLE. Therefore, we presented a novel Fuzzy-based Convolutional Neural Network (FCNN) for effective emotion evaluation. Figure 2 depicts the flow of the proposed method.

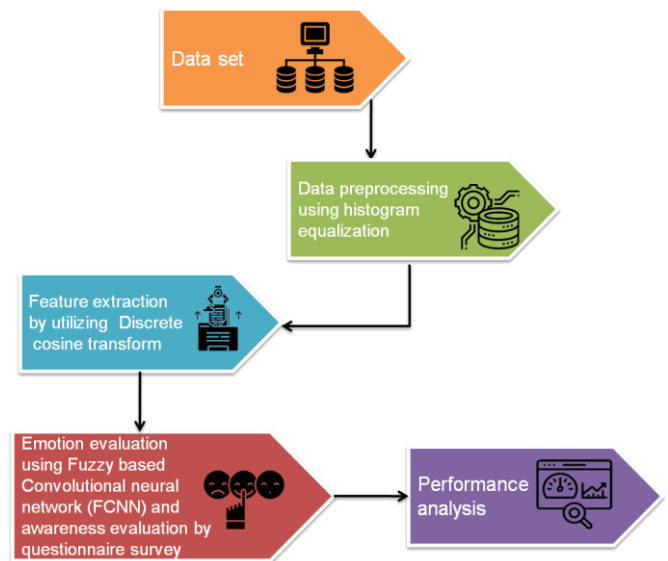


Figure 2. Flow of the proposed method

### 3.1 Data collection

This database contains facial images that were acquired for four separate sessions; the participants displayed a range of emotions while adopting a variety of head postures and exposure settings. There are almost 1,000 different images in all, at a resolution of 640 pixels by 640 pixels. In all, it covers the feelings that were conveyed by 300 participants who participated in any one of the lecture sessions. Smile, surprise, squint, contempt, and fury are some of the facial emotions that have been caught. The images were taken from a variety of tilt angles, and the quantity of light that hit the face was adjusted during the process [23].

### 3.2 Data preprocessing using histogram equalization

Assume that the input face pictures are made up of  $I$  distinct

levels. The following Eq. (1) may be used to write the transforming mechanism.

$$T_k = S(q_k), \quad k = 0, 1, \dots, I - 1 \quad (1)$$

resulting in level  $T_k$  as the outcome of the source, normalized facial input level  $q_k$ . The transformation function  $T(r)$  for histogram equalization [1, 2] is defined by the connection.

$$T_k = S(q_k) = \sum_{i=0}^k R_r(r_i) = \sum_{i=0}^k \frac{n_i}{n}, \quad 0 \leq q_k \leq 1 \text{ and } k = 0, 1, \dots, I - 1 \quad (2)$$

In Eq. (2),  $n$  is the input level  $I$  pixel count and  $R_r(r_i)$  is the probability density function (pdf) of the input picture. Consequently, the transforming function  $T_k$  reflects the initial facial image's continuous dispersion function (cdf). This technique alters the contrast of a facial image by increasing the contrast of big items and decreasing the contrast of minor things, such as the background. While this outcome is suitable for generic pictures like images, it is less ideal if we desire a backdrop with a stark comparison. Figure 3 shows the procedure of histogram equalization [24, 25].

Histogram equalization may provide general brightness augmentation, despite the position in the input facial image. A sub-block is defined by histogram equalization, which also obtains its histogram data. Then, using the cdf of that sub-block, histogram equalization is carried out for the sub-center block's pixel. Up to the ending of the source image, the sub-block is then pushed forward by one pixel and the sub-block histogram equalization is reiterated. As a consequence of employing the nearby sub-block to histogram equalizes each pixel; the outcome is ideally suited to the local lighting conditions. As a result, the local contrast is maximized [26].

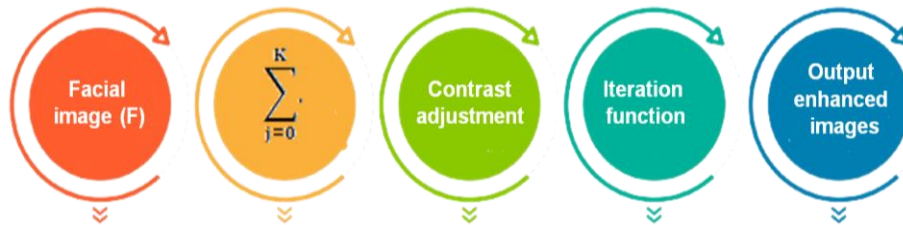


Figure 3. Procedure of Histogram equalization

### 3.3 Feature extraction by using Discrete Cosine Transform (DCT)

Features are an image's representative measurements that set it apart from other data. The chosen attributes need to enhance the contrast between diverse images. By converting the image from the temporal domain to the spectral domain, features are retrieved. The image frequency describes variations in intensity or color throughout an image; it is not a consequence of time, but rather of spatial coordinates. To view content that is not evident in the temporal domain, the frequency data of the image is required. The characteristics are described briefly below. Energy is concentrated into lower-order coefficients through the discrete cosine transform (DCT). The DCT is entirely genuine. The DCT represents a series of infinitely many data elements as the sum of cosine functions that oscillate at various frequencies in an attempt to maintain the most crucial characteristics. The DCT coefficients for the converted output image,  $A_{xy}$  in Eq. (3), are calculated using an input face image,  $B_{nm}$ , by the equation presented below. In the Eq. (4),  $A$  stands for the  $N$ -by- $M$ -pixel face image,  $A_{nm}$  for the intensity values in row  $n$  and column  $m$ , and  $A_{xy}$  for the DCT coefficients in a row  $a$  and column  $b$  of the DCT matrices. The feature extraction of a facial image using DCT is as follows in Algorithm 1. The process of DCT is depicted in Figure 4.

$$A_{xy} = \alpha_a \alpha_b \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} B_{nm} \cos \frac{\pi(2n+1)p}{2N} \cos \frac{\pi(2m+1)q}{2M}, \quad (3)$$

$$0 \leq a \leq N - 1$$

$$0 \leq b \leq M - 1$$

$$\alpha_a = \begin{cases} 1/\sqrt{N}, & a = 0 \\ \sqrt{2/N}, & 1 \leq a \leq N - 1 \end{cases} \quad \alpha_b = \begin{cases} 1/\sqrt{M}, & 1 \leq q \leq M - 1 \end{cases} \quad (4)$$

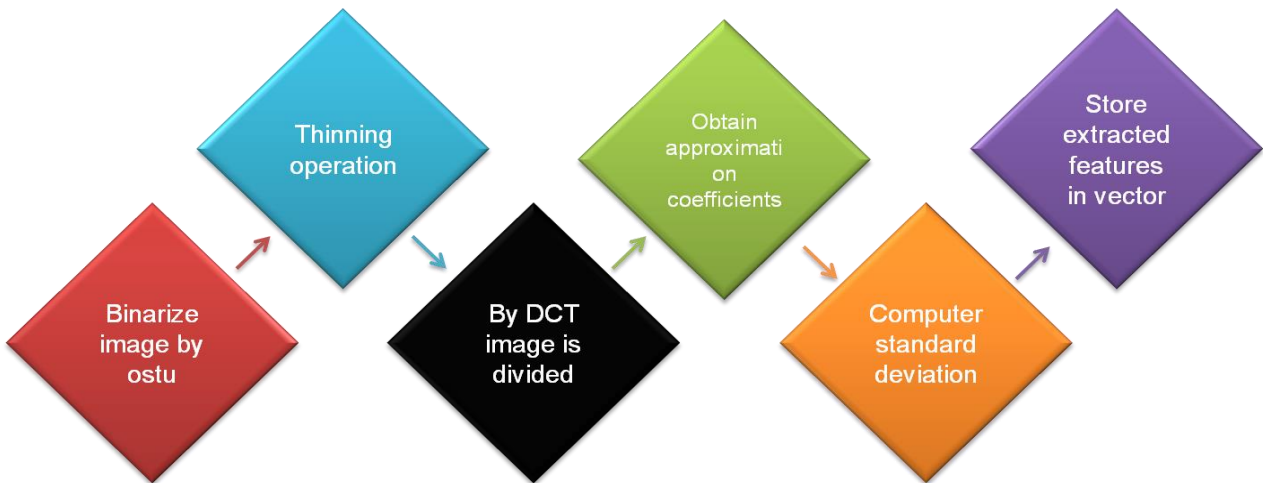
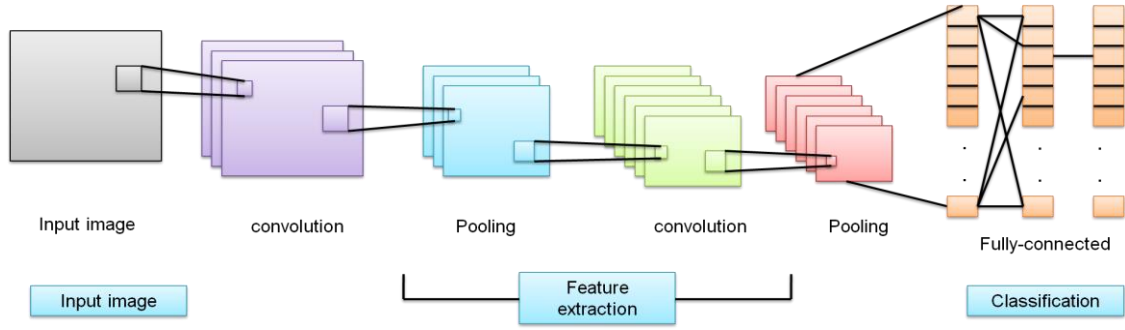


Figure 4. Process of DCT



**Figure 5.** Architecture of FCNN

**Algorithm 1: Discrete cosine transform**

Gray-scale image blocks with a resolution of 640x640 pixels.

Using the Otsu approach, binarize the picture to produce text that represents binary 1 and backdrop binary 0.

Utilizing structural aperture, eliminate any nearby small stuff and perform the thinning procedure.

Apply DCT, split the resulting magnitude (facial image) into four equal, non-overlapping blocks, and then compute the standard deviations for the first and second blocks to retrieve the local patterns. That creates two characteristics.

To get approximation parameters (cA), vertical variables (cV), horizontal coefficients (cH), and diagonal coefficients, perform decomposition for the magnitude (image) of DCT (cD).

Calculate the standard deviation independently for the cA, cV, and cH frequency bands. This creates three characteristics and stores all the calculated characteristics in a vector.

**3.4 Evaluation of emotion utilizing Fuzzy-based Convolutional Neural Network (FCNN)**

The Oxford definitions describe the emotion as a powerful feeling that arises from one's surroundings, moods, or interactions with other people. It might also be described as a sensation made up of behavioral and psychological responses to interior and exterior situations. By analyzing facial expressions, machines can automatically discern emotions. Speaking, minor variations in facial expressions that convey various emotions may be utilized to communicate between individuals. The identification rate and the efficiency of facial expression identification have both been improved through the use development of several methodologies and approaches. CNN, which are artificial neural networks proficient in completing a variety of tasks including image categorization, is one of these methods. The essential premise of fuzzy logic is to substitute the set of real numbers  $\{0, 1\}$  with a member value in the range  $[0, 1]$ . Using a narrative phrase based on fuzzy logic with fuzzy definitions, systems based on fuzzy rules are utilized to construct a fuzzy inference system. The fuzzy information base, which comprises fuzzy If-Then rules, and a fuzzy system, that is utilized to apply the fuzzy logical procedure in the input to generate the output, make up the two essential parts of networks based on fuzzy rules. Hence, we proposed a combined novel technique Fuzzy-based Convolutional Neural Network (FCNN) to evaluate facial emotion. Convolutional, pooling, and fully connected layers make up the CNN's associated layers. The architecture of the FCNN is shown in Figure 5.

The Convolutional layer is in charge of utilizing a convolution operation with filters to extract a characteristic

mapping from the initial facial image. Sliding the filter over the initial facial image is how the convolution procedure is carried out. Everywhere, a matrix multiplication procedure is carried out, and the output is added to create the feature map. A stride value regulates the convolution around the input space and governs the filter's moving step. The filter typically moves pixels at a time. This model's primary job is to categorize the student's facial expressions using the images that the facial detection system provides. The method also offers the n expression categories' possibility levels. We choose the database that will be used to train and evaluate the FCNN model before building it. During virtual learning, we choose the students' face images. The model is trained on 80% of the images that we choose, with a balance of 20% being used for testing. It has three different kinds of layers: fully connected, pooling, and Convolutional. The suggested CNN model's training hyper-parameters are shown. The number of epochs, batch size, and the learning rate is used for training. Much conventional feature recognition employ convolution, a face recognition method that is focused on the idea of a responsive field. A Convolutional kernels mask of the source matrix's sliding window is used during convolution and is calculated using the following Eq. (5).

$$E_{ij} = \sum_{i=0}^{K_a} K \sum_{j=0}^{K_b} y_{ij} \times ka_{ij} \tag{5}$$

where,  $K_a$  and  $K_b$  are the Convolutional kernel's size and length, respectively, and  $E_{ij}$  is the output matrices. To preserve the identical size and length of the matrices after convolution, the border of the intake matrix is often padded with 0, and 0 is still 0 after the Convolutional process. Typically,  $K_a=K_b$  is set as a square kernel,  $y_{ij}$  is the input matrix, and  $ka_{ij}$  is the weight of the convolution kernel. Thus, padding 0 won't have an impact on convolution. The activation function is used to provide nonlinear outcomes for networks that have been uniformly coupled. However, the sigmoid, which was utilized in previous networks, emerges and the gradients vanish over reverse dispersion when the network is large and the number of layers is huge. As a result, FCNN uses ReLU as its activation function. The following Eq. (6) represents the definition of ReLU.

$$g(y) = \max(0, y) \tag{6}$$

The input facial image is condensed after the convolution procedure to reduce the intricacy of the network processing. Using pooling, the dimension may be decreased. The input matrices are used in pooling, but a sliding window mask is used; the masks do not intersect as it moves. In addition, to

allow an  $N \times N$  mask to minimize  $1/N^2$  instances of the source feature matrices and hence lower the dimension, the step must be equal to the height of the Convolutional network. The two types of pooling are typically maximal and average pooling. In this case, the output is the greatest value from the mask, with all other values being immediately eliminated. Optimal pooling is employed for the Convolutional process since it is often preferable to keep the most significant features and theoretically, less significant features shouldn't alter features. Average pooling employs the outcome layer since it determines the aggregate of all elements in the mask. A huge target might receive a high result during the calculation process due to average pooling, which has a physical significance in emotion recognition. Traditional fuzzy integrals' fuzzy density is often adjusted by user feedback. The formula  $Y=y_i = 1$  states that  $m$  represents a set of  $m$  classifiers or the number of CNNs utilized in this research, and  $g(y_i)$  denotes the fuzzy measure, which is taken into account while determining the value of the subset  $x_i$ . In other words,  $y_i$  is the result of the  $i$ th classifier, and  $g(y_i)$  indicates the level of confidence for this class. If the fuzzy measure value is 1, it means that this output may be completely trusted. The fuzzy measure ranges from  $[0, 1]$ . If the value is 0, there is no reference value for the output. These fuzzy measure characteristics are shown.

- When all classifier outputs are consistent, as they are when  $g(Y) = 1$ , the findings may be believed.
- All classifier outputs are not taken into account when  $g() = 0$ .
- The function of a fuzzy measure should be incrementally monotonic is represented as in Eq. (7).

$$X \subset Y \subset Z, \text{ then } 0 \leq g(X) \leq g(Y) \leq 1 \quad (7)$$

Starting with a fuzzy density that only has one element in  $g(A)$ , such as  $g(a_1)$  and  $g(a_2)$ , will allow one to get the fuzzy measure. It is shown how fuzzy measurements and fuzzy density relate to one another. Fuzzy density is defined as  $g(a_1, a_2, \text{ and } a_3)$ , where  $g(a_1, a_2, \text{ and } a_3)$  is the certainty level of the particular class created by merging three CNNs.

### 3.5 Evaluate awareness by questionnaire survey

In a VLE, awareness evaluation is vitally important for enhancing both instructors' and students' user experiences. A questionnaire survey approach could be employed to increase awareness. To measure awareness, several questions are presented to both students and instructors and their responses were validated. Table 1 shows the measure awareness, several questions and the questions as below:

- (1) Is a virtual learning environment beneficial to education?
- (2) Has the implementation of VLE in school education improved your skills?
- (3) Should the VLE's user interface be changed for e-learning?
- (4) Whether instructors and students can quickly get interaction when using a VLE?
- (5) Do you find it difficult to comprehend concepts in a virtual learning environment?

The interface of a virtual learning environment may be customized, according to this survey. Implementing advanced applications in a virtual learning environment could significantly facilitate interactions between students and teachers.

**Table 1.** the measure awareness

S. No	Survey Questions	Yes	No
1	Is a virtual learning environment beneficial to education?	30	20
2	Has the implementation of VLE in school education improved your skills?	22	28
3	Should the VLE's user interface be changed for e-learning?	35	15
4	Whether instructors and students can quickly get interaction when using a VLE?	20	30
5	Do you find it difficult to comprehend concepts in a virtual learning environment?	26	24

## 4. RESULT AND DISCUSSION

In this work, a Fuzzy-based Convolutional Neural Network (FCNN) is used to evaluate emotion and awareness in a virtual learning environment. In this part, we evaluate the performance of the proposed FCNN method for face image-based emotion recognition. As a result, the important criteria to evaluate the effectiveness of the suggested technique are Accuracy, Sensitivity, Specificity, and Precision. The outcomes were compared to those obtained by the use of established techniques. The typically used conventional methods include Support Vector Machines (SVM), Multimodal Approaches and Recurrent Neural Networks (MA-RNN), Joint Adaptive Fine-Tuning (JAFT), and Spatial and Spectral Graph Neural Networks (SSGNN).

### 4.1 Accuracy

To assess the success of the measurement, the accuracy metric will demonstrate the system's efficiency when emotions are correctly identified from face images. Table 2 demonstrates the accuracy of FCNN. SVM achieved 65% accuracy, MA-RNN achieved 85% accuracy, JAFT achieved 74% accuracy, SSGNN achieved 58% accuracy, and the recommended approach achieved 95% accuracy. As a result, the recommended method is more effective than the existing methods.

**Table 2.** Comparison of accuracy of existing and proposed techniques

Technique	Accuracy (%)
SVM [17]	65
MA-RNNJAFT [18]	85
JAFT [19]	74
SSGNN [20]	58
FCNN [Proposed]	95

### 4.2 Sensitivity

When a test is utilized in research, the actual positive proportion, also known as sensitivity, indicates how well the system can identify emotion from face images and provide a favorable result. The sensitivity of SVM obtained 55%, MA-RNN obtained 65%, JAFT obtained 77%, SSGNN obtained 87%, and the suggested approach obtained 93%. So, compared to the existing systems, the recommended technique is more effective. Table 3 displays the sensitivity of the suggested

strategy.

**Table 3.** Comparison of the sensitivity of existing and proposed techniques

Technique	Sensitivity (%)
SVM [17]	55
MA-RNNJAFT [18]	65
JAFT [19]	77
SSGNN [20]	87
FCNN [Proposed]	93

### 4.3 Specificity

A methodology's specificity is its potential to correctly extract emotion from face images. While the proposed method acquired 97%, the Specificity of SVM acquired 53%, MA-RNN acquired 63%, JAFT acquired 74%, and SSGNN acquired 82%. Table 4 illustrates the Suggested Technique's Specificity. As a result, the preferred technique is more specific than the traditional one.

**Table 4.** Comparison of Specificity of existing and proposed techniques

Technique	Specificity (%)
SVM [17]	53
MA-RNNJAFT [18]	63
JAFT [19]	74
SSGNN [20]	82
FCNN [Proposed]	97

### 4.4 Precision

Precision is the proportion of relevant emotions identified from facial expressions. Precision is the average likelihood of recovery. The precision of SVM scored 75%, MA-RNN scored 67%, JAFT scored 55%, SSGNN scored 84%, and the recommended technique scored 97%. The recommended method is thus more precise than the ones previously in use. Table 5 displays the precision of the suggested technique.

**Table 5.** Comparison of Precision of existing and proposed techniques

Technique	Precision (%)
SVM [17]	75
MA-RNNJAFT [18]	67
JAFT [19]	55
SSGNN [20]	84
FCNN [Proposed]	97

## 5. DISCUSSION

Yadav [17] presented a support vector machine (SVM) for facial emotion identification and the ability to recognize seven various outward appearances of two individuals, such as anger, sorrowful, pleased, disgusted, indifferent, fearful, and surprised expressions in the dataset. When there is a large data collection, it does not operate very effectively since the necessary training period takes a longer amount of time. The study by Ackerson et al. [18] recommended using a multimodal approach and recurrent neural network (MA-RNN) to emotion detection, which integrates many different modes into the RNN architecture in an attempt to increase the

accuracy of emotion detection. When the gradient is too low, the changes to the parameters are rendered meaningless and unimportant. Because of this, the processing of lengthy data patterns is made more complicated. Hu et al. [19] developed an approach known as joint adaptive fine-tuning, or JAFT, to combine localized and global information in addition to enhancing the effectiveness of facial emotion classification. This technique can recursively alter the weights of the network. There is a correlation between lowering the dimension of weight in the network and lowering its accuracy. Zhang et al. [20] suggested spatial and spectral graph neural network (SSGNN), is intended to retrieve spatial and spectral context characteristics from face images in recognizing emotions based on both macro- and micro-expressions. Efficiency suffers if there is less processing of graph edges according to the properties and linkages they contain.

The dataset used for training and testing the FCNN model was selected based on the availability of publicly available datasets that contain facial images with labeled emotions. The dataset used in this study is the Japanese Female Facial Expression (JAFFE) dataset, which contains 213 grayscale images of Japanese female facial expressions. The criteria for choosing emotions to capture were based on the six basic emotions (happiness, sadness, anger, fear, surprise, and disgust) and the neutral expression. However, it would be beneficial to provide more details on the diversity of participants in terms of demographics, such as age, gender, and ethnicity. DCT is chosen as the feature extraction technique because it is a widely used method for image compression and has been shown to be effective in capturing facial features for emotion recognition. DCT is relevant to emotion recognition from facial images because it can capture the spatial frequency information of the image, which is important for identifying facial expressions. The algorithm for DCT is provided in the paper, and it involves dividing the image into non-overlapping blocks and applying the DCT to each block to obtain the frequency coefficients. The training process for the FCNN model involved dividing the dataset into training and testing sets, with a ratio of 80:20. The model was trained using the Adam optimizer with a learning rate of 0.001, a batch size of 32, and a total of 100 epochs. The fusion of features from different frequency bands is achieved by calculating the standard deviation independently for the cA, cV, and cH frequency bands and storing all the calculated characteristics in a vector. While the paper does not provide a comparison of the proposed FCNN method with other existing methods, it does evaluate the performance of the FCNN method based on accuracy, sensitivity, specificity, and precision.

## 6. CONCLUSION

The latest trend in education is toward the use of Virtual Learning Environment (VLE) technology in the classroom to engage students who are 21st-century learners. By influencing focus, motivation, and self-control, the evaluation of emotions, as expressed in different frameworks, directly affects user experience. Enhancing teamwork may be achieved by analyzing awareness processes, and the influence of VLEs on academic activities makes it even more necessary that everyone is aware of VLE. Therefore, for improving the user experience in VLE, this article suggested a novel Fuzzy-based Convolutional Neural Network (FCNN) for recognizing human emotions in a virtual learning environment. Awareness

of VLE was evaluated by a questionnaire survey and it results that enhancing the interface improves user experience in VLE. These results indicate the accuracy (93%) sensitivity (95%) specificity (97%) and precision (97%) of the recommended strategy. Examples of existing methods that may be compared to the proposed FCNN include Support Vector Machines (SVM), Multimodal Approaches and Recurrent Neural Networks (MA-RNN), Joint Adaptive Fine-Tuning (JAFT), and Spatial and Spectral Graph Neural Networks (SSGNN). The suggested study may then focus on enhancing computation time performance metrics and accurately identifying emotions utilizing optimization methodologies.

## ACKNOWLEDGEMENT

Authors would like to take the opportunity to thank the grants provider, Prince Sattam bin Abdulaziz University for their support and cooperation in this research.

## REFERENCES

- [1] Rashid, A.H.A., Shukor, N.A., Tasir, Z., Na, K.S. (2021). Teachers' perceptions and readiness toward the implementation of virtual learning environment. *International Journal of Evaluation and Research in Education*, 10(1): 209-214.
- [2] Dung, D.T.H. (2020). The advantages and disadvantages of virtual learning. *IOSR Journal of Research & Method in Education*, 10(3): 45-48. <https://doi.org/10.9790/7388-1003054548>
- [3] Al Rawashdeh, A.Z., Mohammed, E.Y., Al Arab, A.R., Alara, M., Al-Rawashdeh, B. (2021). Advantages and disadvantages of using e-learning in university education: Analyzing students' perspectives. *Electronic Journal of E-learning*, 19(3): 107-117. <https://doi.org/10.34190/ejel.19.3.2168>
- [4] Tzafilkou, K., Perifanou, M., Economides, A.A. (2021). Negative emotions, cognitive load, acceptance, and self-perceived learning outcome in emergency remote education during COVID-19. *Education and Information Technologies*, 26(6): 7497-7521. <https://doi.org/10.1007/s10639-021-10604-1>
- [5] Collazos, C.A., Fardoun, H., AlSekait, D., Pereira, C.S., Moreira, F. (2021). Designing online platforms supporting emotions and awareness. *Electronics*, 10(3): 251. <https://doi.org/10.3390/electronics10030251>
- [6] Liu, Y., Zhang, X., Zhou, J., Fu, L. (2021). SG-DSN: A semantic graph-based dual-stream network for facial expression recognition. *Neurocomputing*, 462: 320-330. <https://doi.org/10.1016/j.neucom.2021.07.017>
- [7] Aggarwal, D., Zhou, J., Jain, A.K. (2021). Fedface: Collaborative learning of face recognition model. In 2021 IEEE International Joint Conference on Biometrics (IJCB), IEEE, Shenzhen, China, pp. 1-8. <https://doi.org/10.1109/IJCB52358.2021.9484386>
- [8] Liao, J., Liang, Y., Pan, J. (2021). Deep facial spatiotemporal network for engagement prediction in online learning. *Applied Intelligence*, 51: 6609-6621. <https://doi.org/10.1007/s10489-020-02139-8>
- [9] Miao, X., Yu, Z., Liu, M. (2021). Using partial differential equation face recognition model to evaluate students' attention in a College Chinese Classroom. *Advances in Mathematical Physics*, 2021: 1-10. <https://doi.org/10.1155/2021/3950445>
- [10] Zhang, K., Li, Y., Wang, J., Cambria, E., Li, X. (2021). Real-time video emotion recognition based on reinforcement learning and domain knowledge. *IEEE Transactions on Circuits and Systems for Video Technology*, 32(3): 1034-1047. <https://doi.org/10.1109/TCSVT.2021.3072412>
- [11] Verulkar, A., Bhurchandi, K. (2022). Auto-luminance-based face image recognition system. In *Proceedings of First International Conference on Computational Electronics for Wireless Communications: ICCWC 2021*. Springer Singapore, pp. 553-565. [https://doi.org/10.1007/978-981-16-6246-1\\_47](https://doi.org/10.1007/978-981-16-6246-1_47)
- [12] Padmashree, G., Karunakar, A.K. (2022). Improved LBP face recognition using image processing techniques. In *Information and Communication Technology for Competitive Strategies (ICTCS 2021) ICT: Applications and Social Interfaces*. Singapore: Springer Nature Singapore, pp. 535-546. [https://doi.org/10.1007/978-981-19-0095-2\\_51](https://doi.org/10.1007/978-981-19-0095-2_51)
- [13] Tran, C.K., Ngo, T.H., Nguyen, C.N., Nguyen, L.A. (2021). SVM-based face recognition through difference of Gaussians and local phase quantization. *International Journal of Computer Theory and Engineering*, 13(1): 1-8. <https://doi.org/10.7763/IJCTE.2021.V13.1282>
- [14] Kim, J.H., Poulouse, A., Han, D.S. (2021). The extensive usage of the facial image threshing machine for facial emotion recognition performance. *Sensors*, 21(6): 2026. <https://doi.org/10.3390/s21062026>
- [15] You, S. (2021). Discrete lion swarm optimization algorithm for face recognition: Discrete lion swarm optimization algorithm for face recognition. *Multimedia Research*, 4(3): a1. <https://doi.org/10.46253/j.mr.v4i3.a1>
- [16] Kuruvayil, S., Palaniswamy, S. (2022). Emotion recognition from facial images with simultaneous occlusion, pose and illumination variations using meta-learning. *Journal of King Saud University-Computer and Information Sciences*, 34(9): 7271-7282. <https://doi.org/10.1016/j.jksuci.2021.06.012>
- [17] Yadav, S.P. (2021). Emotion recognition model based on facial expressions. *Multimedia Tools and Applications*, 80(17): 26357-26379. <https://doi.org/10.1007/s11042-021-10962-5>
- [18] Ackerson, J.M., Dave, R., Seliya, N. (2021). Applications of recurrent neural network for biometric authentication & anomaly detection. *Information*, 12(7): 272. <https://doi.org/10.3390/info12070272>
- [19] Hu, M., Ge, P., Wang, X., Lin, H., Ren, F. (2022). A spatio-temporal integrated model based on local and global features for video expression recognition. *The Visual Computer*, 38(8): 2617-2634. <https://doi.org/10.1007/s00371-021-02136-z>
- [20] Zhang, J., Sun, G., Zheng, K., Mazhar, S., Fu, X., Li, Y., Yu, H. (2022). SSGNN: A macro and microfacial expression recognition graph neural network combining spatial and spectral domain features. *IEEE Transactions on Human-Machine Systems*, 52(4): 747-760. <https://doi.org/10.1109/THMS.2022.3163211>
- [21] Jawarneh, M., Alshar'e, M., Dewi, D.A., Al Nasar, M., Almajed, R., Ibrahim, A. (2023). The impact of virtual reality technology on Jordan's learning environment and medical informatics among physicians. *International*



- Journal of Computer Games Technology, 2023. <https://doi.org/10.1155/2023/1678226>
- [22] Ghazal, T.M., Hasan, M.K., Wahab, A.N.A., Ibrahim, A., Khan, W.A., Raza, N.Z., Atta, A., Mago, B. (2022). Towards Privacy Provisioning for Internet of Things (IoT). In 2022 International Conference on Cyber Resilience (ICCR), Dubai, United Arab Emirates, pp. 01-07. IEEE. <https://doi.org/10.1109/ICCR56254.2022.9995916>
- [23] Al-Juboori, S.A.M., Almutairi, H., Almajed, R., Ibrahim, A., Gheni, H.M. (2022). Detection of hand gestures with human computer recognition by using support vector machine. Periodicals of Engineering and Natural Sciences, 10(2): 467-476. <https://doi.org/10.21533/pen.v10i2.2866>
- [24] [24] GAMES, P.E.I.E.V. (2020). Using playability heuristics to evaluate player experience in educational video games. Journal of Theoretical and Applied Information Technology, 98(23): 3632-3642.
- [25] Niu, Y., Al Sayed, I.A., Alya'a, R.A., Al\_Barazanchi, I., JosephNg, P.S., Jaaz, Z.A., Gheni, H.M. (2023). Research on fault adaptive fault tolerant control of distributed wind solar hybrid generator. Bulletin of Electrical Engineering and Informatics, 12(2): 1029-1040. <https://doi.org/10.11591/eei.v12i2.4242>
- [26] Jabbar, M.S., Al\_Barazanchi, I.I., Khalaf, A.L., JosephNg, P.S., Radhi, A.D. (2023). Optimizing multi-antenna M-MIMO DM communication systems with advanced linearization techniques for RF front-end nonlinearity compensation in a comprehensive design and performance evaluation study. Periodicals of Engineering and Natural Sciences, 11(3): 124-138. <https://doi.org/10.21533/pen.v11i3.3609.g1296>