science, engineering and health studies

Research Article

Science, Engineering and Health Studies
https://li01.tci-thaijo.org/index.php/sehs
ISSN (Online): 2630-0087

Emotion recognition of students during e-learning through online conference meeting

Chulapong Panichkriangkrai[1*], *Piyarat Silapasuphakornwong*[2] *and Thirachit Saenphon*[3]

[1] Department of Imaging and Printing Technology, Faculty of Science, Chulalongkorn University, Bangkok 10330, Thailand
[2] Human Media Research Center, Department of Information Network and Communication, Kanagawa Institute of Technology, Kanagawa 243-0292, Japan
[3] Faculty of Information and Communication Technology, Silpakorn University, Phetchaburi 76120, Thailand

***Corresponding author**: Chulapong Panichkriangkrai chulapong.p@chula.ac.th

Received: *27 January 2021*
Revised: *1 April 2021*
Accepted: *7 April 2021*
Published: *23 December 2021*

Citation: *Panichkriangkrai, C., Silapasuphakornwong, P., and Saenphon, T. (2021). Emotion recognition of students during e-learning through online conference meeting. Science, Engineering and Health Studies, 15, 21020010.*

ABSTRACT

Due to the outbreak of COVID-19, online learning has become a way of life. The objective of this study was to propose techniques to detect students' emotions while studying via online video conferencing. This proposed technique, which updates the facial emotion image of the current class-member, enables the system to achieve a highly accurate performance for facial emotion recognition. This proposed technique can be applied to online teaching systems. As a result, instructors can identify the interest levels of each learner using the interest assessment system, which measures and monitors the period of tiresomeness of each learner. The results showed that our techniques achieved a high percentage of accuracy for each emotion, that is, sleepy/bored = 93.3%, confused = 94.3%, neutral = 92.6%, and happy = 97.2%, which was higher than the convolutional neural network-based emotion recognition system. The proposed system was applied to a real class and satisfactory overall results of 88.7% were achieved. This study proved the feasibility of the proposed technique.

Keywords: emotion recognition; face recognition; e-learning; online conference meeting; class interesting assessment; video conferencing for education

1. INTRODUCTION

Online meeting video conference platforms have been developed for having face-to-face conversations—as if the participants are in the same room despite being in different continents. In the past, this required a lot of specialized equipment, but it can now be done using a general computer with a webcam or via a smartphone with an application program. Furthermore, online meeting video conference platform development has taken a giant leap after the COVID-19 pandemic.

The crisis and the pandemic of COVID-19 has bought a big change in humanity (Cucinotta and Vanell, 2020). Every field has globally been disrupted. People have had to change their lives: working from home, buying and selling everything through online shops, and studying online. This is due to the national policies that encourage everyone to stay at home. This also directly affects the education system (Dwivedi et al., 2020). In Thailand, children could not attend schools from March to June 2020 and online learning platform systems have become solutions to learning problems.

Online classrooms and e-learning are a new normal in the educational system (Marinoni et al., 2020). Two-thirds of the reported classrooms have shifted from real-time face-to-face to distance teaching and learning. Students must study from home through the online classes that the teachers have prepared and posted to the cloud in advance.

With distance and independent learning approaches, friendship and social relationships among people in the class seem to have disappeared. Hence, online meeting video conference platforms, such as Zoom™, Facebook™, Skype™, Microsoft Team®, and Google Meet®, have been applied as educational tools for the class, to hold live interactions between teachers and students (Gonzales-Zamora et al., 2020).

However, teachers using online meeting video conference tools cannot gauge the interest levels of each student (Al-Samarraie, 2019). This limitation is related to several factors. First, online applications may not support interactive learning. The functions of these tools, for instance, muting the microphone and closing and limiting the functions of capturing from the camera, were designed to protect the privacy of the learners during meetings. Thus, when these applications are applied to the online classroom, most of the students choose to use these options. Second, there are many types of devices that students use such as laptops, smartphones, and tablets, which might affect the quality of images. Third, the limitation of these tools is their limited view angle. It is difficult for the teacher to sense students' feelings because they can see only faces, and not their postures and environment. Hence, teachers cannot understand the students' real feelings as well as they can in a real classroom. This leads to the problem of managing the study plan for the learners in future.

The efficiency of the class can be evaluated from active learning, where teachers can examine the feelings and behaviors of the students in the class (Chick et al., 2020). Naturally, most students' faces are "neutral" when they are paying attention to the class. Furthermore, if they smile or laugh, it means that they are happy, feel a part of the class, and are interested in the lessons, then the results of this study are sufficient. Alternatively, if they look bored and sleepy, it means that the students do not understand the lesson or the class is too slow. Then, teachers should change their teaching methods or suspend the class to invigorate the students. Krithika and Lakshmi Priya (2016) proposed the student emotion recognition system, which can evaluate the concentration of the students from their faces in three levels. Hence, it is necessary for teachers to know the feelings and emotions of the learners during the class to adjust and develop a suitable study plan to match their emotions (Rehn et al., 2017). For online studies, there is a need for a system that detects depression in students during the online classes (Gillies, 2007).

There are various approaches to evaluate the mood and attention of students during online classes. Examples include facial expressions, gestures and postures, eye movements, and feedback checklists and exercises from students (Dewan et al., 2019). However, with current online class tools, the teacher only sees the students' faces (Soltani et al., 2018). Hence, the facial recognition of students can help identify their emotions during e-learning. Putra and Arifin (2019) proposed real-time emotion recognition tools, which used a convolution neural network (CNN) and fuzzy unordered rule induction algorithm to predict emotions on students' computers and send the results to the teacher's server. However, it would be less complicated if students were not required to install additional applications.

The concept of applying facial emotion recognition to online video conference meeting tools is to use it to monitor the interest level of each student in an online class through a video conference platform, such as Zoom™, based on facial emotion recognition (Figure 1).

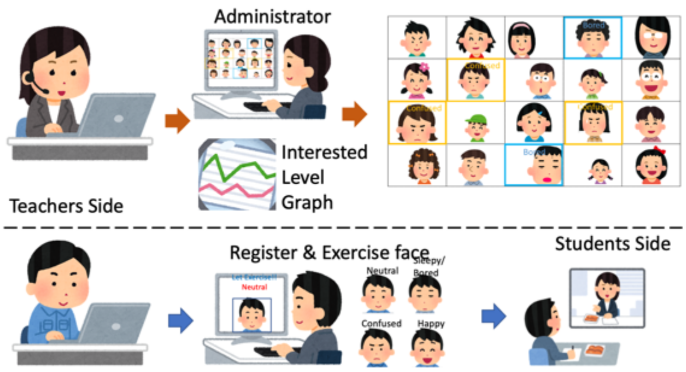

Figure 1. Conceptual system that applies facial emotion recognition to online conference meeting

Here, one class was separated into two sides: the teacher side and the student side. Our system was used from the teacher's perspective to collect data from the students' side by monitoring. Therefore, the teachers are the administrators of the system. When the students logged in to the online class, they were asked to express four emotions with their faces: neutral, bored, confused, and happy. This indicated if the students were awake, active, and ready for the class. The teachers or administrator could view all the students, as in the normal scenario of a video conference platform. When some students were confused, bored, or felt sleepy and did not pay attention to the lesson, the system triggered some notifications to the teacher side by marking them with specific colors to notify the teacher about the tired students. The system could identify these students by evaluating the tiresomeness period from the interest level graph where the emotions of each student in real time were plotted. Hence, our proposed system can help the teacher to detect the level of interest of the learner and their activeness in real time in online class learning. Thus, teachers can manage their teaching plan based on the students' attention and study environment. For example, if most of the students in the class (i.e., 80% of the class) feel sleepy, the teachers can have a break of about 10 min so that students can refresh themselves. This would lead to a more effective and efficient learning process.

The overall workflow system focusing on the process is shown in Figure 2. The input is a video scene from an online classroom meeting. In one image scene, there are many faces representing the students who attend the class. In this study, face recognition boosting part was applied to promote the efficiency of emotion recognition. Thus, the proposed system achieved a high accuracy. The main function of our system is to recognize emotions from facial images. However, applying the emotion classification results to the assessment graph of interest is suggested for evaluating the tiresomeness period of the students, which directly relates to the level of interest in the lesson. The system can notify the teacher about who (student) is paying less attention in class.

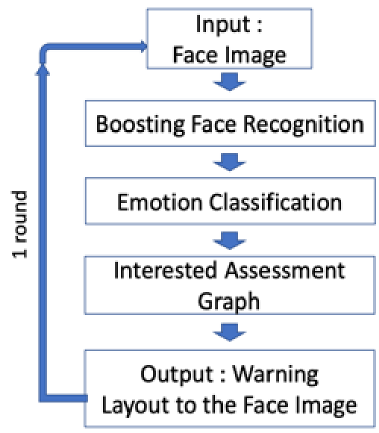

Figure 2. Flowchart of our ideal proposed system

The aim of this study was to propose a new technique to recognize facial emotion, which can increase the accuracy of emotion recognition compared to the widely used current technique. The performance and efficiency of the proposed system were tested using three experiments.

2. MATERIALS AND METHODS

2.1 Proposed technique of boosting face recognition

This study proposed a boosted face recognition technique that aimed to increase the efficiency of emotion recognition. This technique can achieve a higher level of accuracy than the traditional method. Traditional face recognition techniques are commonly used in identifying a person. These systems must be set to recognize each person's emotions. This means training and classifying must be repeated two or more times, which is ineffective owing to its high cost and computation time. Hence, this paper proposed a technique where the traditional methods were applied to directly recognize emotions only once. The proposed technique was initially trained with many students' face emotions. This saves the cost and time required for computation and achieves higher accuracy. Figure 3 illustrates this process. Initially, the system can detect the face from the input scene, based on the normal based face detection (NBFD) system. The well-known technique, Haar face detection, was applied here (Viola and Jones, 2001). In the first few minutes, the system predicts emotions based on the existing database, during which the updated current class-member face (UCCF) part is working. When students first login to the class, the UCCF requests them to set their standard emotion face (SEF), which is collected as the registered face for each emotion for each student in the database. SEFs are composed of four main emotions: 1: sleepy or bored, 2: confused, 3: neutral, 4: fun or active. This is a new and updated database to be used for the current class. Then, the system uses these SEFs for training the local binary patterns histogram (LBPH) directly and ready recognition of facial emotion with high accuracy (Zhao and Wei, 2017). This process, which allows the system to recognize the emotion for each person faster and more accurately, is the key to our success.

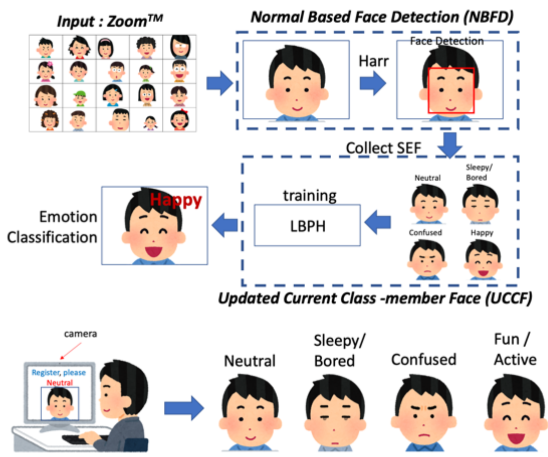

Figure 3. Two parts of boosting face recognition technique: NBFD and UCCF (our proposed technique)

Emotion recognition for each student occurs when they are studying in the class. The students' captured faces are compared and classified into four SEF classes. These four-face expressions cover the basic emotions that every student feels during learning. The teacher can evaluate the achievement of the class when the students feel the emotion #4-happy/fun/active. In addition, emotion #2-confused and #1-sleepy/bored have effects on the learning efficiency of students. It is then necessary for teachers to be notified so that they can manage the teaching approach.

Algorithm 1 explains the algorithm of the boosted face recognition technique used for training the emotion cascade. First, a normal-based face detection using the Haar feature-based cascade classifier is applied to the input image $I(i)$ to detect the student face f. Then, the system collects the face image f of the emotion e when the student is requested to perform each emotion. After correcting 20 images for each of the four emotions, the emotion cascade *C* will be created using the LBPH face recognition training method. Algorithm 2 is used for emotion recognition. First, normal-based face detection using a Haar feature-based cascade classifier is applied to input image *I(i)* to detect the student face *f*. Then, normal-based face detection is applied using the emotion cascade *C* to detect emotion *e*.

Algorithm 1. Emotion training
```
input: image I(i)
Normal based face detection using Haar face detection cascade
if I(i) is face,
for each face f in I(i) do
for each emotion e = 1:4 do
for i = 1:20
correct face image f of emotion E(e)
Training emotion E using LBPH face recognition
return emotion cascade C
```

Algorithm 2. Emotion recognition
```
input: image I(i)
Normal based face detection using Haar face detection cascade
if I(i) is face,
Normal based face detection using emotion cascade C
return emotion e
```

2.2 Students' emotion monitoring process

Here, the class was separated into small groups. Each group comprised three to five persons. The system recognized the facial emotion of each person in a group parallelly. The time taken for each student's emotion monitoring was approximately 30 s. For example, as shown in Figure 4, there were 25 students in this class. The system divided the students into four groups. Each group had five students, and the emotions of each person in a group were monitored simultaneously and in parallel; it took approximately 30 s. This means that in one round (20 persons), The system took two minutes to monitor all the members in the class in one round. It also means that the students' emotions would be updated every two minutes. This monitoring can notify the teachers in real time.

2.3 Interest assessment system

After the system classified the emotions of the students into four classes, a real-time graph was plotted. This emotion graph describes the interest level of each student in the lesson, as shown in Figure 5.

The tiresomeness period is the area of the graph in state 1 (bored/sleepy) and state 2 (confused). This area refers to the inattentive or inactive state of students. When the student is in state 2, it means that they do not understand the lesson. Then, if they continue in state 2 for a longer period, they have a high chance of changing from an active to an inactive state, that is, state 1 (sleepy or bored). This means that they will stop paying attention to the class. Moreover, this phenomenon leads to the failure of the study and affects the understanding level of the students.

However, this graph might not be suitable for direct use by teachers because it is too complex, and there is insufficient time to monitor all the student data in the class. Hence, a warning system was used as a notification display on the system interface to communicate with the teacher (Figure 6). The specific notification colors are marked in the original scene of the video conference. In Figure 6, specific colors were used to identify emotions such as yellow, which represented the students who were in the emotion state 2 and blue, which represented the students in the emotion state 1, respectively. With these clear notifications, the teachers can immediately recognize the students who are in the tiresomeness period or those who are active in the class.

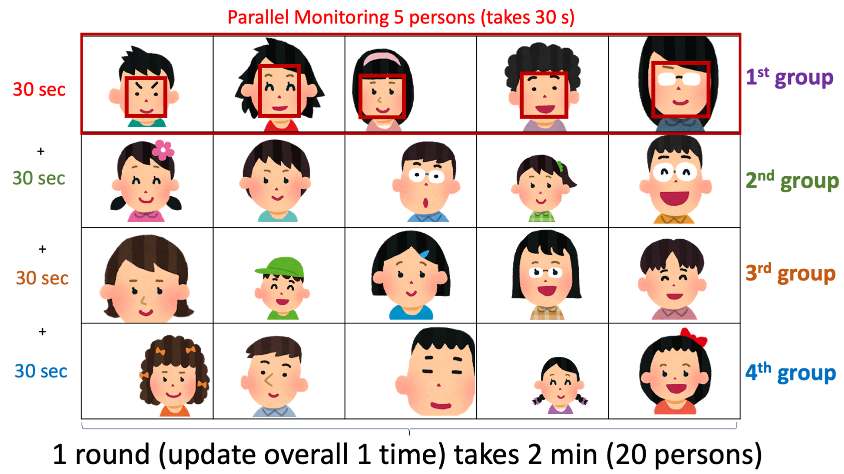

Figure 4. Students' emotion monitoring which students are divided into small groups and their emotions can be recognized in parallel

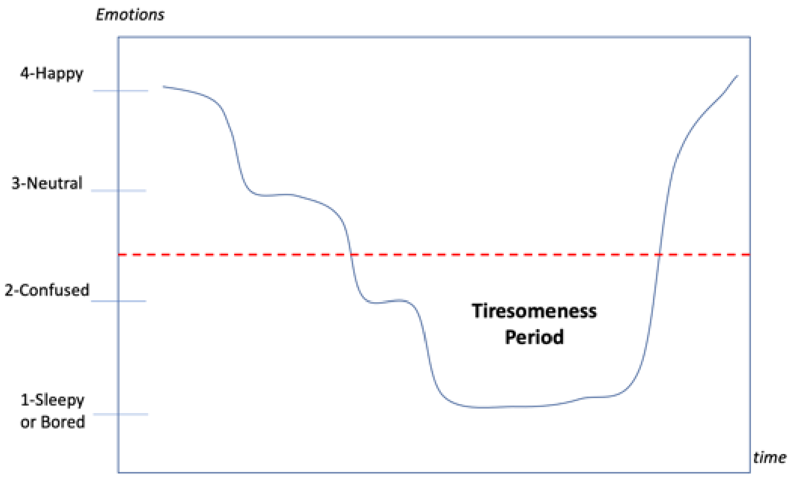

Figure 5. Tiresomeness period in the emotion graph during the system monitoring of a student, presenting the period during which the student lost his/her interest in the lesson

Figure 6. Notification in the video conference tools

3. RESULTS AND DISCUSSION

Our proposed system is directly based on facial emotion recognition. Hence, to achieve high performance, it is necessary that the part of the facial emotion recognition should achieve a high level of accuracy.

This paper proposed a new technique for boosting facial recognition—a system for updating the class members and recognizing their emotion directly instead of identifying the person's face prior to recognizing their emotion. The LBPH, which is simple but very effective, was used as the recognition algorithm. To evaluate its performance, three experiments were conducted to show the emotion recognition results which were achieved with a higher accuracy.

3.1 UCCF Performance

The UCCF work once the students login to the class. The system collected the current feelings of the students as their registered faces for that class. Another issue to explore here was the number of training sets that can be sufficient for the system to recognize the face better.

A dataset was created containing 2,400 face images from 10 Thai university students (aged 17-20 years), half of whom were men, and the other half were women. The system had already learned to recognize their emotion before the class. This supported the NBFD part that would then recognize emotions initially while the UCCF was working.

The results (Figure 7) showed a comparison of the performance (emotion recognition results) of the system with and without the UCCF for boosting. When the NBFD started working, there were zero image for training the system. Then, when the UCCF began working, the system collected (capture) many new face images for training. In this experiment, 20 pictures were collected as training images because the performance was stable (that is, additional photos cannot increase the accuracy level).

The accuracy of the recognition increased from 30% to 50% for every emotion after the first training data was included (after our proposed UCCF had worked). Then, the accuracy increased to 80% after the 9th image and reached 88% after the 16th image. After this, it became stable. The results showed that when the system updated the training data with the current images, the accuracy increased. These results indicated that the UCCF needed sample face images of around 16 per emotion to update the dataset of the system and boost the performance of recognition of the system by approximately 50%. Hence, the proposed technique, the UCCF, was efficient.

In practice, one can design a method for the face image collection using the UCCF technique by asking the students to express their emotions in four faces when they first login. These trial activities were recorded in the system as a video file at a normal rate of 24 frames/s. The system can randomly collect an image file in every five frames. Thus, it can obtain sufficient sample capturing images in a few seconds. Hence, it is confirmed that the proposed technique was efficient in practical application.

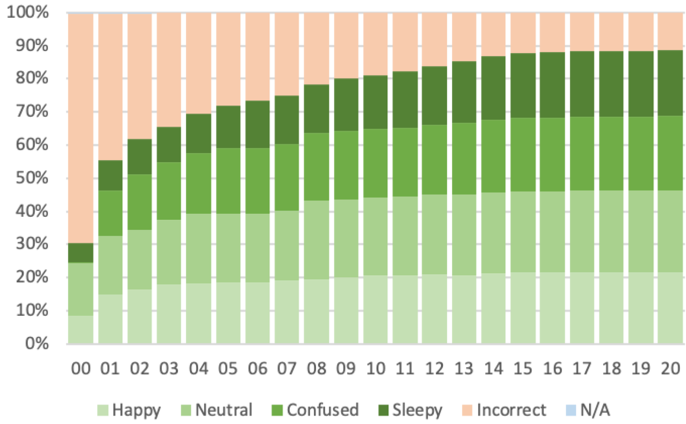

Figure 7. Emotion recognition accuracy: without UCCF (0 training image) and after the application of UCCF in boosting the recognition system (1-20 training images)

3.2 Robustness of technique

To prove the robustness of our proposed technique for any face image, different facial expression images were tested, namely our created dataset and the general database or FER2013. Our dataset had a prominent point of extremely the same face images but with different emotions. However, in the FER2013 dataset, there were diverse face images and large differences in gender, age, race, and facial expressions. Hence, the results of our study were compared with the deep learning model to confirm the efficiency of our model in deep and wide recognition.

The FER2013 dataset (Goodfellow et al., 2013) was used to test system's wide proficiency in facial recognition. This dataset is well known for facial emotion recognition and is used in the recognition of students' faces in online education (Wang et al., 2020). It was created by Pierre-Luc Carrier and Aaron Courville. There are many examples of faces with several feelings that vary with gender, age, and race. The data of FER2013 consists of 48 × 48 pixel grayscale images of faces, with a total training set of 28,709 examples, and a testing set of 3,589 examples in 7 categories such as 0-anger, 1-disgust, 2-fear, 3-happy, 4-sad, 5-surprise, and 6-neutral. In this experiment, 3 emotions were selected from the FER2013 database which consisted of 16,280 training images and 1,998 testing images. The selected emotions were fear(F), neutral(N), and happy(H). The results were compared with three classes in our proposed system, namely, confused(C), neutral(N), and happy(H). Sample images of each emotion are shown in Figure 8.

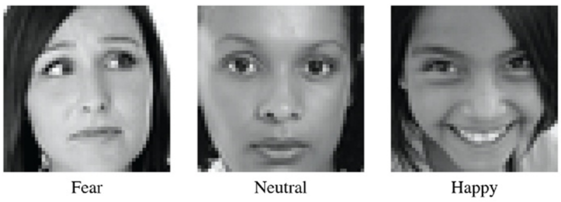

Figure 8. FER2013 database: sample images from three emotions that match with our proposed system

Moreover, our facial expression dataset was used to test the in-depth proficiency in recognition. A dataset from 10 Thai students with diverse feelings (of the same emotion) on different days was generated. This demonstrated what really happens in a real class where the same students in different classes on different days might have different facial expressions. For this experiment, a three-class dataset, that is, confused(C), neutral(N), and happy(H) was compared to the FER2013 dataset. Our dataset contained 450 training images and 1,854 testing images.

The results were compared with those of the deep learning model. The CNN was used because it is a well-known effective method for recognition (Krizhevsky et al., 2017). Both techniques were used to recognize the emotions in both datasets. MATLAB R2019b by MathWorks, Inc., Natick, Massachusetts with a deep learning toolbox was used for the CNN processing. We trained the CNN with our dataset with 20 images from each student per emotion, totaling approximately 200 images for each emotion. The architecture of the CNN used in this study is shown in Figure 9. The input images were normalized to 48 × 48 pixels grayscale images. Two convolutional layers were used. The first convolutional layer had 32 filters of size 5-by-5 and 2-by-2 padding with a learning rate for the biases in the layer, which was twice that of the current global learning rate. The second convolutional layer had 64 filters of size 5-by-5 and 2-by-2 padding with a learning rate for the biases in the layer, which was twice that of the current global learning rate. After each convolutional layer, an average pooling layer was applied with a pooling size of 3-by-3 and a stride of 2-by-2. All the results were presented as confusion matrices, as shown in Figure 10.

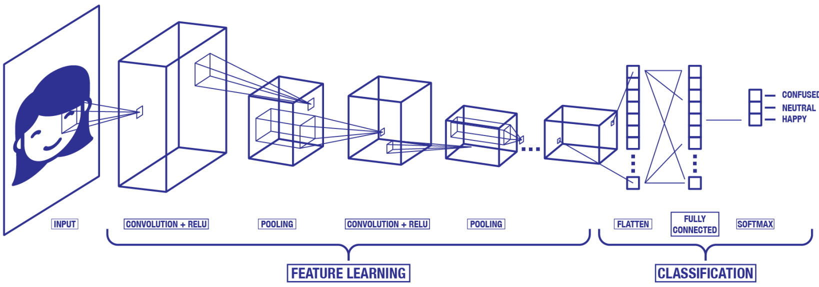

Figure 9. Example of a network with many convolutional layers (adapted from) The MathWorks, Inc., 2021)

As seen in Figure 10, it can be said that our proposed technique achieved the best accuracy with our dataset at 75.24%, whereas the CNN achieved only 70.01% accuracy. On the other hand, our proposed technique achieved an accuracy of 50.91% with the FER2013 dataset, the CNN achieved an accuracy of 68.37% with the same. This means that our proposed technique has deep recognition proficiency whereas the CNN has wide recognition proficiency.

The problem with online video conference classes is that most students remained with the same expression for one class, but in other classes, they had a different facial expression for the same feeling, which varied in details. Thus, a method that can focus on details and be robust in recognizing the same feeling is required. Our proposed technique works better than the CNN in such situations.

3.3 Our proposed technique performance

Deep learning is an efficient and powerful technique, and it is currently used for solving classification problems. Many studies have applied deep neural networks for recognition as well as in online education systems (Sahla and Kumar, 2016). Here, the efficiency of our proposed technique was compared to that of a deep neural network to prove the superiority of the proposed technique. In this study, CNN, a type of deep neural network, was compared with our technique.

This experiment was conducted with our four-class dataset, which can also be applied to the interest assessment system in future to confirm that the proposed technique is better than the CNN and is suitable for online class face recognition. The efficiency of recognition was evaluated using the receiver operating characteristic (ROC) curve and the area under the curve (AUC). The results of both the methods are shown in Figure 11.

FER-Propos FERTrain
Process Our propos
Train data FER-2013
Test data FER-2013

		True condition						
		F	N	H				
Predicted condition	F	193	88	131	193	46.84	219	412
	N	151	320	260	320	43.78	411	731
	H	152	199	504	504	58.95	351	855
Correct		193	320	504	1017			
%		38.91	52.72	56.31		**50.90**		
Incorrect		303	287	391			981	
Sum		496	607	895				1998

(a)

Our-Propos OurTrain
Process Our propos
Train data Our
Test data Our

		True condition						
		C	N	H				
Predicted condition	C	420	129	30	420	72.54	159	579
	N	135	482	44	482	72.92	179	661
	H	70	51	493	493	80.29	121	614
Correct		420	482	493	1395			
%		67.20	72.81	86.95		**75.24**		
Incorrect		205	180	74			459	
Sum		625	662	567				1854

(b)

FER-CNN FERTrain
Process CNN
Train data FER-2013
Test data FER-2013

		True condition						
		F	N	H				
Predicted condition	F	249	102	72	249	58.87	174	423
	N	158	415	121	415	59.8	279	694
	H	89	90	702	702	79.68	179	881
Correct		249	415	702	1366			
%		50.20	68.37	78.44		**68.37**		
Incorrect		247	192	193			632	
Sum		496	607	895				1998

(c)

Our-CNN OurTrain
Process CNN
Train data Our
Test data Our

		True condition						
		C	N	H				
Predicted condition	C	380	109	79	380	66.9	188	568
	N	196	519	89	519	64.55	285	804
	H	49	34	399	399	82.78	83	482
Correct		380	519	399	1298			
%		60.80	78.40	70.37		**70.01**		
Incorrect		245	143	168			556	
Sum		625	662	567				1854

(d)

Figure 10. Confusion matric results of different datasets experiment (Robustness): our proposed technique tested with (a) FER2013, and (b) our dataset, compared with the CNN method, tested with (c) FER2013, and (d) our data

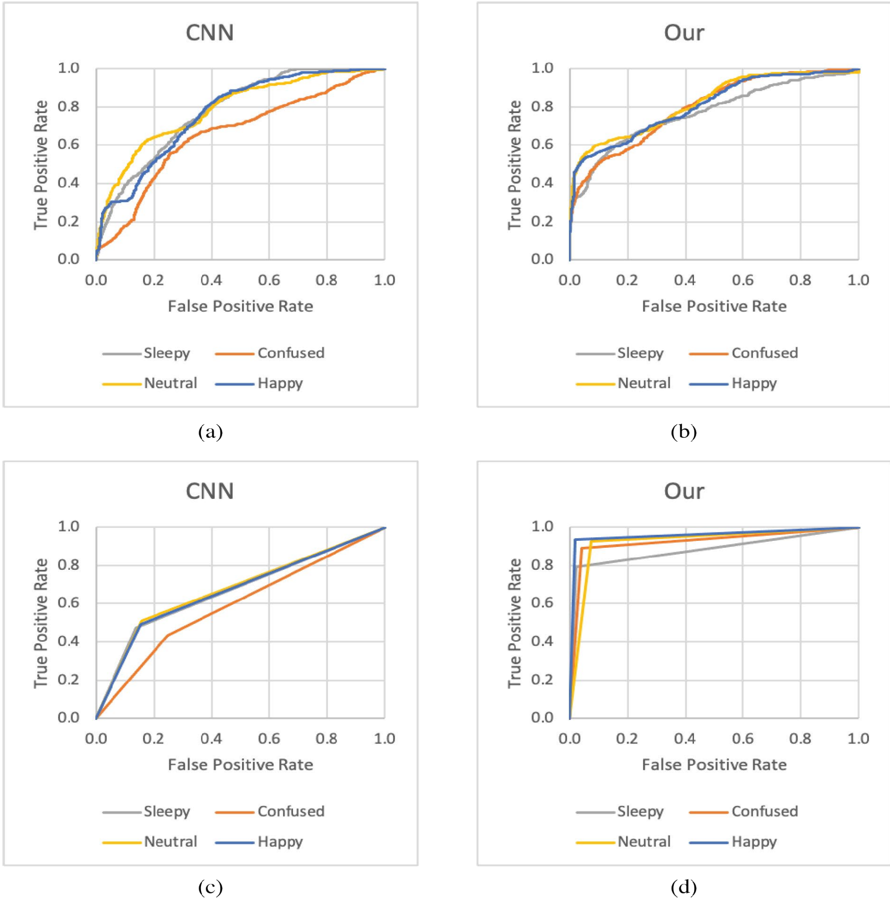

Figure 11. Receiver operating characteristic graphs from the recognition results

The true positive rate of the ROC in Figures 11a and 11b was simulated by testing it with the false positive rates varying from 0 to 1. On the other hand, in Figures 11c and d, the real values from the experiment were obtained, which indicated the average performance of the overall system. The AUC is equal to that of the probability of a classifier ranking a randomly chosen positive instance higher than a randomly chosen negative instance. Hence, if the ROC graph is close to 1, the performance is much better (Marzban, 2004).

The AUC results of the recognition of the four emotions, calculated from Figures 11a and 11b, are reported in Table 1. The criteria for a rough guide for classifying the accuracy of a diagnostic test is the traditional academic point system (Tape, 2013), which was applied as shown in Table 2. The AUC results of the CNN method were 0.76 (sleepy/bored), 0.64 (confused), 0.77 (neutral), and 0.75 (happy). For our technique, the values were 0.77 (sleepy/bored), 0.80 (confused), 0.82 (neutral), and 0.81 (happy).

Table 1. Area under the curve (AUC) results of the classifier

Class	**Our**		**CNN**	
	AUC	**Grading**	**AUC**	**Grading**
Sleepy/Bored	0.77	Fair	0.76	Fair
Confused	0.80	Good	0.64	Poor
Neutral	0.82	Good	0.77	Fair
Happy	0.81	Good	0.75	Fair

Table 2. Area under the curve (AUC) results of the system

Class	**Our**		**CNN**	
	AUC	**Grading**	**AUC**	**Grading**
Sleepy/Bored	0.89	Good	0.67	Poor
Confused	0.93	Excellent	0.59	Fail
Neutral	0.93	Excellent	0.68	Poor
Happy	0.96	Excellent	0.67	Poor

Figures 11c and d show that the AUC results of our proposed technique for each class were 0.89 (sleepy/bored), 0.93(confused), 0.93(neutral), and 0.96(happy). In comparison, the results of the CNN were 0.67, 0.59, 0.68, and 0.67, respectively which are not good.

In conclusion, our proposed technique achieved better performance in the recognition of the four classes of emotion compared to CNN, reported as the ROC and AUC. Moreover, the accuracy and sensitivity of our system were compared to those of the CNN, as shown in Table 3. The equations for the accuracy and sensitivity were calculated using Equation 1 and 2, respectively.

$$Accuracy = (TP + TN) / (TP + TN + FP + FN), \quad (1)$$

where TP is true positive, TN is true negative, FP is false positive, *and* FN is false negative.

$$Sensitivity = TP/(TP + FN), \quad (2)$$

Our system achieved an efficiency of 88.7% (the CNN achieved only 47.8%). It is reiterated that the proposed system is more effective. The results confirmed that our proposed technique is suitable for facial emotion recognition in online video conferences.

Table 3. Accuracy of the proposed system

Class	Our Accuracy	Sensitivity	CNN Accuracy	Sensitivity
Sleepy/Bored	0.93	0.79	0.76	0.47
Confused	0.94	0.89	0.67	0.43
Neutral	0.93	0.93	0.75	0.51
Happy	0.97	0.94	0.76	0.49
Overall	0.89	0.89	0.48	0.48

The UCCF is efficient and sufficient for online classrooms because of the instability of a person's image in the videoconference system in each classroom. The results showed that for every video conference class, the scenario of the students attending the class always changes even for the same students. They change in personality, dressing, level of emotion (during or even before they attend the class), environment (for example, room, lighting, tools, and platforms (desktop or mobile). These factors have a significant effect on the feelings of the same person in each class. Hence, because our system requires UCCF in every new class, this is collecting the face image in update. Thus, the performance of our system is better.

The interest assessment system is still in the process of being build-up as an application for the near future. This experiment should be conducted in a real class.

4. CONCLUSION

This study proposed an effective system for monitoring the level of interest of students in an online class through an online video conference platform. A conceptual system that can detect the stress of the students through the tiresomeness period, reported via the color change in the student's scene in the original platform, was proposed. Moreover, the boosting face recognition, an important part of our proposed system for achieving high accuracy (88.7%) of facial emotion recognition, was tested. The results showed that optimizing the number of face images for boosting the system was only 16 face images per emotion which is sufficient to improve accuracy significantly. The experiments also proved that the performances of our technique for classifying emotions into 4 classes were 0.933 (sleepy/bored), 0.943 (confused), 0.926(neutral), and 0.972 (happy), respectively, which is higher than the CNN method, although this is a simple technique. Furthermore, our technique is robust and can be applied to various datasets. This confirms that the proposed technique is suitable and ready for use in real-time application.

Our technique will support class management for teachers in e-learning and online teaching and solve problems and respond to the "new normal" of the online class due to the disruption caused by COVID-19.

REFERENCES

Al-Samarraie, H. (2019). A scoping review of videoconferencing systems in higher education: Learning paradigms, opportunities, and challenges. *International Review of Research in Open and Distributed Learning,* 20(3), 121-140.

Chick, R. C., Clifton, G. T., Peace, K. M., Propper, B. W., Hale, D. F., Alseidi, A. A., and Vreeland, T. J. (2020). Using technology to maintain the education of residents during the COVID-19 pandemic. *Journal of Surgical Education*, 77(4), 729-732.

Cucinotta, D., and Vanelli, M. (2020). WHO declares COVID-19 a pandemic. *Acta Bio-medica: Atenei Parmensis,* 91(1), 157-160.

Dewan, M. A. A., Murshed, M., and Lin, F. (2019). Engagement detection in online learning: a review. *Smart Learning Environments,* 6(1), 1-20.

Dwivedi, Y. K., Hughes, D. L., Coombs, C., Constantiou, I., Duan, Y., Edwards, J. S., Gupta, B., Lal, B., Misra, S., Prashant, P., Raman, R., Rana, N. P., Sharms, S. K., and Upadhyay, N. (2020). Impact of COVID-19 pandemic on information management research and practice: transforming education, work and life. *International Journal of Information Management,* 55, 102211.

Gillies, D. (2007). Student perspectives on videoconferencing in teacher education at a distance. *Distance Education,* 29(1), 107-118.

Gonzales-Zamora, J. A., Alave, J., De Lima-Corvino, D. F., and Fernandez, A. (2020). Videoconferences of infectious diseases: an educational tool that transcends borders. a useful tool also for the current COVID-19 pandemic. *Le Infezioni in Medicina,* 28(2), 135-138.

Goodfellow, I. J., Erhan, D., Carrier, P. L., Courville, A., Mirza, M., Hamner, B., Cukierski, W., Tang, Y., Thaler, D., Lee, D. H., Zhou, Y., Ramaiah, C., Feng, F., Li, R., Wang, X., Athanasakis, D., Shawe-Taylor, J., Milakov, M., Park, J., . . . Bengio, Y. (2013). Challenges in representation learning: a report on three machine learning contests. In *Neural Information Processing. International Conference on Neural Information Processing* (Lee M., Hirose A., Hou ZG. and Kil R. M., eds.), pp. 117-124. Berlin: Springer-Verlag Berlin Heidelberg.

Krithika, L. B., and Lakshmi Priya, G. G. (2016). Student emotion recognition system (SERS) for e-learning improvement based on learner concentration metric. *Procedia Computer Science,* 85, 767-776.

Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2017). Imagenet classification with deep convolutional neural networks. *Communications of the Association for Computing Machinery,* 60(6), 84-90.

Marinoni, G., Van't Land, H., and Jensen, T. (2020). IAU global survey report. In *The impact of Covid-19 on higher education around the world*, pp. 1-50. Paris: International Association of Universities.

Marzban, C. (2004). The ROC Curve and the Area Under it as Performance Measures. *Weather and Forecasting,* 19(6), 1106-1114.

Putra, W., and Arifin, F. (2019). Real-time emotion recognition system to monitor student's mood in a classroom. *Journal of Physics: Conference Series*, 1413(1), 012021.

Rehn, N., Maor, D., and McConney, A. (2017). Navigating the challenges of delivering secondary school courses by videoconference: Navigating challenges of videoconference. *British Journal of Educational Technology,* 48(3), 802-813.

Sahla, K., and Kumar, T. S. (2016). Classroom teaching assessment based on student emotions. In *Proceedings of the International Symposium on Intelligent Systems Technologies and Applications,* pp. 475-486. Jaipur, India.

Soltani, M., Zarzour, H., and Babahenini, M. C. (2018). Facial emotion detection in massive open online courses. In *Proceedings of the World Conference on Information Systems and Technologies*, pp. 277-286. Napies, Italy.

Tape, T. G. (2013). The area under an ROC curve. *Interpreting Diagnostic Tests*. [URL: https://www. mathworks.com/hardware-support/home.html?s_tid=srchtitle_MATLAB%2520%2526%2520Simulink_1] accessed on March 18, 2021.

The MathWorks, Inc. (2021). *Convolutional neural network. MATLAB & Simulink*. [URL: https://www.mathworks.com/discovery/convolutional-neural-networkmatlab.html] accessed on March 18, 2021.

Viola, P., and Jones, M. (2001). Rapid object detection using a boosted cascade of simple features. In *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 511-518. Hawaii, USA.

Wang, W., Xu, K., Niu, H., and Miao, X. (2020). Emotion recognition of students based on facial expressions in online education based on the perspective of computer simulation. *Complexity,* 2020, 4065207.

Zhao, X., and Wei, C. (2017). A real-time face recognition system based on the improved LBPH algorithm. In *Proceedings of 2017 IEEE 2nd International Conference on Signal and Image Processing (ICSIP)*, pp. 72-76. Singapore.