



Facilitating Online Collaborative Learning Through Affective Computing

Zhuting Sun

*Department of Computer Science and Technology, Hainan Tropical Ocean University, Sanya, China
sunzhuting1986@163.com*

Abstract

Affective Computing (AC) has immense potential to improve the performance and experience for Collaborative Learning (CL) e-learners by inferring, recognizing and processing e-learners' affective information. Exploring the online CL involves AC algorithm with two groups of identical activities supported by the project of the CSC-UBC. One group implements EmotionSense platform to detect e-learners' emotional states that is based on Convolutional neural networks (CNNs) model, while the other group only employs the traditional online CL. Finally, the results reveal that the better learning performance and experience are from the second group, whereas barely difference of learning outcome between individuals.

Keywords: *Collaborative learning; Affective Computing; CNNs; emotion; e-learner*

1 INTRODUCTION

CL is an umbrella term for a variety of educational approaches involving joint intellectual effort by students, of students and teachers together [1]. It is believed to provide opportunities for meaningful or deep learning. Particularly as e-learners are more likely to perform poorly, that is, learning happens on its own, without direction, without control and without discussion. Undeniably, e-learners desire to learn as group or pair, meanwhile, Collaborative Learning or Cooperative Learning have been proved to improve students' experience and performance and productivity. Barkley (2014) provided CL techniques as a handbook for college faculty, describing authentic onsite and online CL activities respectively. [2] However, the negative mood widely affects e-learners' decision-making and problem-solving that would lead to decreased productivity [6].

Researchers have found that online CL must explore more facial data from e-learners in order to detect emotional states, therefore to provide effective and timely interventions for them. EmotionSense, a mobile built-in platform, is employed to recognize students' facial images as well as proximity and patterns of conversation by processing the outputs from the sensors of commodity smartphones. The accelerometer, Bluetooth and GPS, microphone, touch count and device's shake as well, that can be utilized to detect e-learners' micro expressions [3].

The aim of the paper is composed of four sessions. Firstly, it argues the relationship between e-learners' productivity and affective states. Secondly, it explores CNNs algorithm for AC. Thirdly, it argues for the full potential of EmotionSense for detecting e-learners' facial expressions based on CNNs model. Lastly, it compares the performance of e-learners by implementing traditional CL and CL involving AC approach.

2 E-LEARNER'S PRODUCTIVITY AND AFFECT

Researchers have reached the consensus that the emotional responses and micro expressions have huge impact on motivation. Coping well with emotional awareness before, during and after learning tasks can improve metacognition, strategic action as well. The emotions have been proved to play a critical role on e-learners' ability of decision-making and problem-solving. It is believed that the negative moods would reduce learners' skill and productivity. Precisely, students would be able to achieve the Zone of the Proximal Development in positive attitudes and moods, such as joy, energetic, patient, gratitude, hope, serenity, interest, pride, inspiration, awe and so fourth. Instead, learners who are in negative states are more likely to perform poorly even to fail it, such as anger, disgust, fear, sadness, indifferent, depressive, anxiety, edgy, depressive and surprise as well. The relationship between affective states and performance can be indicated as Fig. 1.

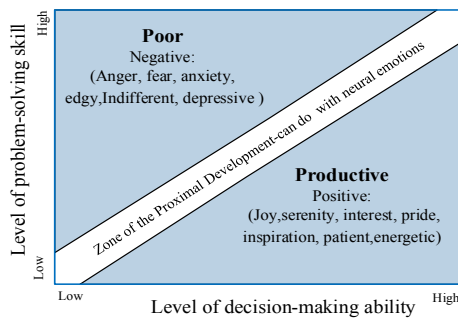


Figure 1: The relationship between emotional states and productivity.

3 PROPOSED METHOD

3.1 Built-in sensor platform for capturing emotional images

The Facial Action Coding System (FACS) has been widely applied as a standardized model which has 44 facial movements that are generated by facial muscles, such as lips, skin texture, nose, eyebrows, eyelids as well [3].

Technically, the fast evolved wearable techniques based on Artificial Intelligent (AI) have provided considerable opportunities to improve AI’s understanding of human distinct-states, such as the Electrocardiogram (ECG) and the Electrodermal activity (EDA). Yet, researchers have concerned about the validity of recognition of a person’s emotions. Inevitably, the wearable devices cause physical pain that direct to decreased learning experience and outcome. Paredes (2013) provided sensor-less sensing for AC in the detection of stress by computer mouse, keyboard, camera, voice and smartphone’s mice. [5]

For the purpose of inferring more accurate affective data from e-learners. In the research, involving Ekman’s seven basic emotions: anger, disgust, fear, happiness, sadness and surprise, plus one neutral [3]. An application of build-in sensor smartphone platform based on CNNs

model is employed to detect participants’ affective states. Technically, the EmotionSense platform captures emotional images through sensor, and detects emotions by digital image processing involving joint CNNs model and normalized facial images together.

3.2 CNNs algorithm for AC

Artificial Neural Network (ANN) is a commonly used algorithm for AC, CNNs is one of several kinds of ANN which is essentially utilized for image recognition. In particular, CNNs has been applied to identify emotions, objects, skulls etc. The work proposes CNNs model to detect e-learners’ facial images with the normalized emotional images introduced by FACS. The main procedures of the algorithm are provided below.

3.2.1 Input e-learners’ emotional images

By tensors we understand that an image is nothing but a matrix of pixel values, when it is displayed as mode of Red-Green-Blue (RGB), it can be represented by three-dimensional tensors (3D tensors). They are width(pixels), height (pixels), and depth (3 color channels) [4]. After representing the pixels as tensors, those tensors would be filtered to generate feature map, such operation is called convolution.

3.2.2 Convolution

Convolution is a mathematical operation that is performed on two functions to produce a third one which is usually called a filtered version of the original functions. In digital image processing, applying images of pixels $[x, y]$, and a filter of size $K_1 \times K_2$. The convolution is shown below [4]:

$$h[x, y] = \sum_{k_1=0}^{k_1-1} \sum_{k_2=0}^{k_2-1} h[k_1, k_2] \times [x + k_1, y + k_2] \tag{1}$$

According to the expression above, the filter uses $h[x, y]$ to describe the convolution, and the output matrix results from the convolutional operation is called the feature map or activation map. To fully understand the algorithm, an example is provided as Fig. 2 [7].

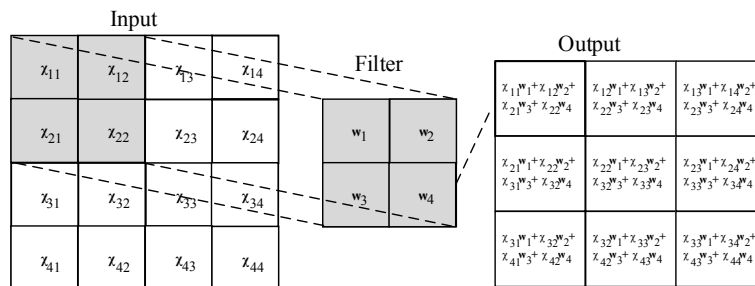


Figure 2: Example of a 2D convolution of stride of 1.

As shown in Fig 2, the input image of emotion can be represented by a matrix of 4×4 dimension, and the filter

size of 2×2 is commonly used in CNNs, the output is a convolved feature. Larger strides lead to fewer steps and

result in a smaller activation map matrix. So far, the size of the feature map matrix still costs too much computation, thus, the pooling perform is indispensable.

3.2.3 Pooling

The pooling operation is essential to diminish the size of the output activation map $O \in R^{W_o H_o D_o}$ for reducing computational cost. In this way, we can obtain high-level abstractions from input by stacking multiple convolutional layers in a CNNs [7]. The kernel(filter) size $k \times k$ and stride S are used to perform the pooling operation. Finally, the width and the height of the activation map are decreased to the below equations [7]:

$$w_o = \frac{(w_o - K)}{S} + 1 \tag{2}$$

$$H_o = \frac{(H_o - K)}{S} + 1 \tag{3}$$

The max pooling and average pooling are generally performed in CNNs. an example is provided as Fig.3 [4].

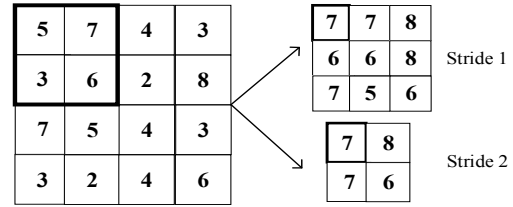


Figure 3: Max pooling with 2×2 filters and stride of 1 and 2.

In Fig 3, with the stride of 1, the maximum value of each filter can be illustrated as $\max(x_1, x_2, x_3, x_4)$, thus, both the results of the max $(5,7,3,6)$ and $\max(7,4,6,2)$ equal 7. While with the stride of 2, the max $(5,7,3,6)$ and the max $(4,3,4,6)$ equal to 7 and 6 separately. We move the filter across the image in steps of one and two coordinating to stride of 1 and 2 respectively until we reach the last corner. After performing the stride of n , the width and the height are considerably diminished without relevant loss of image information.

3.2.4 Motivation and matching

The CNNs use tensors to represent the information of emotional images. Through applying the convolution and pooling mathematical performs, and after adding multiple convolutional hidden layers, the output of the last layer is flattened into a long vector. Such vector then is developed to automatic classification or filter matching. The motivation of CNNs can be indicated as Fig. 4 [4].

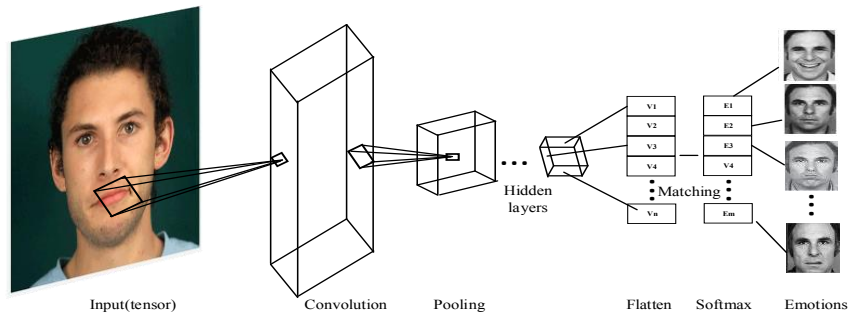


Figure 4: The motivation of CNNs.

By applying this model to generate the value of matching result from the input image and the normalized emotional images. The input vector is defined as x_i , weights are defined as w_i , which are also described as filters above, the output is indicated as $f(x)$ and a bias as well. The result is ultimately transformed to binary values. Each position of the filter matching is done by computing the $f(x)$ with the equation below:

$$f(x) = \begin{cases} 1 & \text{if } \sum_{i=1}^n x_i w_i + bias \geq 0 \\ 0 & < 0 \end{cases} \tag{4}$$

Through the equation above, the output $f(x)$ is the predicted outcome of the matching value. As we train the model, the cost function is customarily used to evaluate its accuracy.

3.2.5 Cost and information loss of image

In the equation given below, i represents the index of the filter, m is the number of the filter, y is the real value, \hat{y} is the output of $f(x)$, whereas $f(c)$ is defined as the loss of the image information.

$$f(c) = \frac{1}{2m} \sum_{i=1}^m (\hat{y} - y)^2 \tag{5}$$

In CNNs, the initial hidden layers are applied to classified simple features of image, whereas the multiply later hidden layers are used to recognize more complex shapes. To improve the accuracy of the computing, we need to tune:(1) the size of the pixels and filters and stride,

the value of threshold(bias), the standardized dataset of facial images.

4 THE ONLINE CL IMPLEMENTATION

There are three phases are designed and implemented in the work: Preparation, Activity design and Implementation. Since the participants are required to pass the final exam in order to graduate in the project of CSC-UBC, therefore, all of them paly the role as e-learner. The Faculty of Education at UBC adopts the online CL activity called two-stage-exam, which is comprised of two stages, the individual test and retaking the test in collaborative groups. Finally, the learning experience, performance and outcome of two groups would be assessed through Collaborative Learning Peer Evaluation Form and the marks of the two-stage-exam.

4.1 Preparation

(1) Participants

The first experimental group involves 27 participates who are from Science class in the project, called Group A. They were grouped as 5 groups along with one leader respectively. While the Group B are from Art class, 21 in total. They also are divided into 5 groups and one scholar is appointed as leader for each group.

(2) Devices and platform

Two experimental groups experience the same online CL activity, but then with different mobile devices. With supporting of the Center for Teaching, Learning and Technology (CTLT) and the Emerging Media Lab (EML Project) at UBC, the Art class (Group B) are allowed to borrow pertinent mobile devices with built-in EmotionSense platform, which can detect participants' various micro expressions.

It is worthy to highly mention that there was no prior manual annotation to participants that those devices had been imbued in the EmotionSence platform.

4.2 Activity design

(1) Test design

The instructor designs one test for two-stage-exam with 10 choice questions, which is allocated as 50 scores to 50 scores for two stages averagely, posting the test on Canvas on two deferent webpages with exam time for individuals and groups respectively.

(2) Learning environment preparation

The assistant (of the instructor at UBC) prepares the Canvas and Zoom for CL, specifically, creating enough breakout rooms on Zoom and guiding students to join appropriately.

4.3 Implementation

(1) Taking the test individually

At the beginning of the learning on Zoom and Canvas, each group member is asked to take the test and submit the answer individually on Canvas in 15 minutes, downloading the digital answer sheet of his or her answers for the upcoming group discussion.

(2) Retaking the test in collaborative groups

After finishing the test individually, participants meet in their CL groups in small rooms on Zoom and retake the test with the online IF-AF "scratch-and-win" answer sheet on Canvas. The sample of the IF-AF answer sheet is shown as Fig. 5.

IF-AF scratch-and-win
 IMMEDIATE FEEDBACK ASSESSMENT TECHNIQUE(IF-AT)
 SCRATCH OFF COVERING TO EXPOSE ANSWER

Name	Group #	A	B	C	D	Score
Week #	___					Total _____
1.		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2.			★		

Figure 5: The part sample of the IF-AF answer sheet.

The rule of the IF-AF is set with 5 to 2 to 1 to 0, which means if the group scratch the star at the first time, they would get 5 scores, the cases of calculating other scores are repeated by the same rule. It is composed of six steps:

(1) Comparing their answers on the first question.

(2) If there is agreement, one member explains the rationale or procedure underlying the question and the group moves on to the second question.

(3) If there is disagreement, members discuss in the small rooms until they reach the consensus. The group moves on to the third question on the IF-AF answer sheet.

(4) The learning group repeat this procedure until they have covered all test questions.

(5) The EmotionSense platform sends messages to users so as to improve the learning performance.

5 RESULTS AND DISCUSSION

In terms of the CNNs algorithm in equation 4, if $f(x) \geq 0$, the emotional image would be defined as filter matching completely and successfully. Various reports would be initially generated and would be sent to the targeted users (Group B) instantly. Based on the big data provided by the project, during the procedure of the online two-stage-exam of the Art class (Group B). It was reported that each student had received message from the EmotionSense platform. Some of them receive the same notice with the same color icon, such as green icon, yellow icon, blue icon, red icon as well.

At the end of the the online two-stage-exam activity, participants are encouraged to provide feedback by filling evaluation forms responsibly, which will be used to discuss the level of collaborative experience and performance in the research, while learning outcome in accordance with the two-stage-exam grade.

5.1 The Collaborative Learning Peer Evaluation Form

The Collaborative Learning Peer Evaluation Form (Barbara, 1998) considers four criteria: attendance, participation, preparation, cooperation, and attitude. Undeniably, as a group member, his or her learning experience and performance can be indicated through these criteria. The form is shown as Fig .6:

Collaborative Learning Peer Evaluation Form

	1-never	2-rarely	3-sometimes	4-usually	5-always
Has the student attended your collaborative learning group on a regular basis?	1	2	3	4	5
Has the student made an effort at assigned work?	1	2	3	4	5
Does the student attempt to make contributions and/or seek help within the group when he/she needs it?	1	2	3	4	5
Does the student collaborate with group effort?	1	2	3	4	5
Name of the student being evaluated: _____	Total: _____				
Group name: _____					

Figure 6: Collaborative Learning Peer Evaluation Form.

Each participant is asked to fill the above evaluation form responsibly through an online link, the final marks of the two groups are shown as Fig. 7.

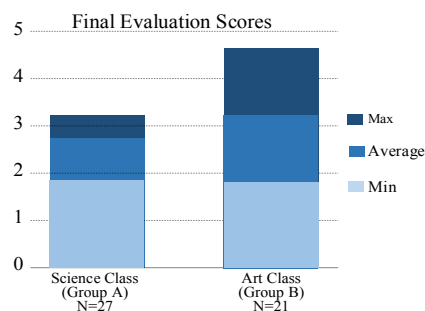


Figure 7: The scores of Collaborative Learning Peer Evaluation.

As the figure shown above, the maximum, average and minimum scores of the Group A are illustrated as 3.22, 2.73 and 1.85 respectively; whereas the Group B is indicated as 4.65,32.2 and 1.80 separately. In other words, participants in the second group have better collaborative experience and performance.

5.2 The two-stage-exam grades of two groups

The instructor gets the results and feedback about all submitted tests from both individual and group through Canvas, the results is illustrated as Fig. 8.

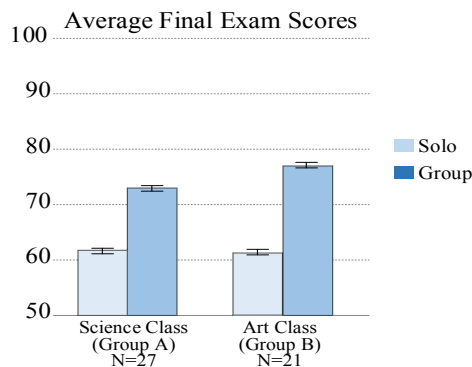


Figure 8: The grade of solo and group in two groups.

Through Fig 8, the solo grade in group A and group B are illustrated as 62 and 61.2 respectively, while the group grade of them are shown as 73.4 and 77 separately. It reveals that there is merely difference between the solo scores from two groups, however, the participants from Group B have better learning outcome, in particular the long duration of CL activity.

6 CONCLUSION AND CHALLENGE

In this experiment, the participants of two groups are different, and such constraints can barely be avoided in the project, resulting in data inaccuracy. Other concerns such as the duration of the online learning activity also should be considered in next research. Nevertheless, somehow, the implementation of the AC approaches had been proved to facilitate the experience of online CL.

With the supporting of the constantly evolved AI technologies in recent years, the AC approach would be endorsed and become more popular within the wider academic and technological circles [3]. However, the discussion for technical and ethical challenges deriving from AC faced are still infancy [3]. Some of its key aspects are as (1) The validity and the accuracy;(2) The privacy and informed consent;(3) the physical pain caused by the wearable devices;(4) Data misuse;(5) Trust and engagement;(6) the duration of the online CL experience.

ACKNOWLEDGEMENTS

This work was sponsored by China Scholarship Council (Higher Education Visiting Scholar Project, Grant No. [2017]5087.

Project supported by Education Department of Hainan Province, project number. Hnjg2019-70; project number. Hnjg2022-90; project number. Hnjg2017-46.

Project supported by Education Department of Hainan Tropical Ocean University, project number: RHYjg2022-09.

Project supported by Education Scientific Planning of Hainan Province, project number: QJZ20181008.

Project supported by Academy and Local Government of Sanya, project number.2017YD12.

REFERENCES

- [1] David W. Johnson, Robert T. Johnson. Learning together and alone: cooperation, competition and individualization. 5th ed, Boston: Allyn and Bacon. 1999, p.4-111.
- [2] Elezabeth F.Barkley, Claire Howell Major, and K.Patricia Cross. Collaborative Learning and Techniques [M], SF: Jossey-Bass, 2014, pp18-139.
- [3] E. Politou, et al., A survey on mobile affective computing, Computer Science Review (2017), <http://dx.doi.org/10.1016/j.cosrev.2017.07.002>.
- [4] Osva A Montesinos-López, et al, 2022. Multivariate Statistical Machine Learning Methods for Genomic Prediction, Springer Cham, Colima, 1st edition.
- [5] Pablo Paredes, et al. Sensor-less Sensing for Affective Computing and Stress Management Technology (C), January 2013, DOI: 10.4108/icst.pervasivehealth.2013.252380
- [6] Rawin Assabumrungrat, et al. Ubiquitous Affective Computing: A Review [J] IEEE Sensors Journal. December 2021, DOI: 10.1109/JSEN.2021.3138269
- [7] Wen Xu, Jing He, Yanfeng Shu and Hui Zheng. 2020. Advances and Applications in Deep Learning, Intech Open, Melbourne.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

