The Analysis on Student’s Psychologic Status of Online Learning under Extraction Model from Computer Face Features

Lan Wen\(^1\)\(^*\), Shaogang Yang\(^1\), Jiao Zeng\(^2\), Xiaowen Liang\(^1\), Yanping Xu\(^1\)

\(^1\)Institute of Psychological and Education, South China Business College, Guangdong University of Foreign Studies, China
\(^2\)South China Business College, Guangdong University of Foreign Studies, China

*Corresponding author e-mail: windlives@163.com

Abstract. The research aims to analyze students’ psychological mood during online study and enhance the efficiency of learning. In this research, face features are extracted by LBP (local binary pattern) algorithm based on face feature recognition. The facial feature points are tracked by CLM (constrained local model) algorithm, and the face images are normalized by AdaBoost algorithm. In the end, the fatigue level of learners is calculated by P80 method in PERCLOS criterion to verify the accuracy, recognition rate and error rate of the model. The results indicate that through the recognition of the faces in a complex background, the accuracy of the mode is 95%, the recognition rate is 95.53% and the error rate is only 0.53%, the range of aspect ratio in the human eye image is \(0.22 \leq \lambda_{open} \leq 0.27, 0.05 \leq \lambda_{close} \leq 0.1\), and learner’s level of excitement can be estimated according to the range of \(f\). Therefore, this algorithm model can extract learner’s facial features very well and show a good result. Detecting the degree of learner’s excitement accurately can provide an important fundamental for the realization of online learning emotion detection system which also has guiding significance for the development and popularization of the networked education.

1. Introduction
With the rapid development of information technology and network technology, computer and network technology has been widely recognized and applied, and a variety of new learning models have emerged. Among all the learning modes, the most striking one is online learning with the development of network technology, because it conforms to the requirements of interactive learning advocated by the new national curriculum standard \([1]\). In the network teaching environment, teachers can broadcast the teaching after preparing for the class. They can also upload the video to the online teaching platform and conduct one-to-one remote tutoring for students, and students can learn anytime and anywhere. The biggest difference with the traditional classroom teaching is that the teaching place is no longer fixed in a classroom or a training center on the campus, but in the form of network classes distributed in the network teaching system. Students can learn in a variety of ways, from personal computers and laptops to mobile phones and tablets \([2, 3]\). Teaching starts to take students as the core truly and individual needs as the center. Traditional teaching organization forms are impacted, and the time and space limitations of classroom are gradually broken through \([4, 5]\).
At present, online education is more used for adult education, vocational training or professional knowledge learning. However, online learning also has obvious disadvantages, that is, teachers can’t analyze students' learning mood and state by observing their facial expressions, so they can’t timely adjust teaching strategies [6]. In the modern education mode, the main purpose of teaching is to promote the all-round development of students. And the development of students includes both cognitive development and non-cognitive development, especially emotional development. Teaching that neglects students' emotions can easily lead to the frequent occurrence of students' weariness. Therefore, the function of education is not only to cultivate excellent talents, but also to cultivate mental health talents.

Education is a sacred industry, and the development of education determines the future of a country to some extent [7-9]. The talent structure and talent training mode should be combined with the social environment and the current situation of education development. Therefore, based on the above problems, the emotional model of online learning is designed after the in-depth study and analysis of the characteristics of online teaching and teaching psychology. The model can describe the emotions of online learners from three dimensions: cognition, excitement and avoidance. Based on the model, an online teaching system with emotional interaction function is realized with image processing technology, and the feasibility of the system is verified by relevant tests.

2. Methodology

2.1. The overall design of the online learning system
In the design, taking into account some shortcomings of the traditional network teaching platform, under the premise of ensuring compatibility with various electronic devices, the system allows the teacher to understand the learning state of the learner without the need of video monitoring, which can well reduce the pressure on the server storage and network load of the learning platform. In this part of the design of the teacher platform, teachers need to assign learning tasks first, and then learners can choose the learning place to learn. Due to the compatibility of HTML5 with various devices, learning terminals can be computers, laptops, PADs, mobile phones and other electronic products with cameras. For face recognition and learning facial expression recognition, it is no longer necessary to send video information to the server. Instead, it is implemented by JavaScript, and the recognition result is obtained directly from the client. Finally, the recognition result is sent to the server as text information for analysis and processing [10, 11].

In the process of operation, this system first obtains the learner's learning video through the camera, preprocesses the video image frame (including image enhancement, grayscale, local binary pattern feature extraction), uses the trained AdaBoost classifier for face recognition, and obtains the approximate position of the face. In this step, the number of faces determines whether the learner's behavior is normal. If it is abnormal, the warning message will be displayed. If it is normal, the next step will be carried out. Next, Constrained local model (CLM) algorithm is used for face feature point tracking. Since the approximate position of the face has been obtained in the previous step, the accuracy of tag tracking in this step has been greatly improved. The multi-class Support Vector Machine (SVM) classifier based on probability is used to classify the learning expressions of CLM-labeled faces, and the four learning expressions of learners are obtained. The aspect ratio of the learner's right eye is obtained, and the PERC value of the P80 standard is calculated, thereby obtaining the degree of fatigue of the learner. Finally, the classified text information is sent to the server for analysis and archiving [12].

2.2. The extraction of face features of learners in the online learning environment
The online learning emotion detection system requires high practicability and robustness, and the result of feature extraction is directly related to the accuracy of the entire online learning emotion detection system [13]. In this study, based on the LBP feature face recognition algorithm, face detection is performed on the learner image collected by the webcam.

Local binary pattern (LBP) is a simple and effective feature extraction algorithm for texture classification proposed by Ojala et al in 1996 [14]. The idea of LBP is to use the gray value of a pixel
as the threshold and compare the gray value of its neighboring pixels with this threshold. If the gray value of adjacent pixels is greater than or equal to this threshold, it is marked as 1, and if it is less than this threshold, it is marked as 0. If the marked value is encoded in binary in a clockwise direction, the LBP encoding of the pixel point can be obtained. As shown in the center pixel in Figure 1, its LBP code is 00010011, and its decimal value is 19.

![LBP Algorithm](image)

The calculation formula of the basic LBP operator is as shown in equation (1).

\[ LBP(x_c, y_c) = \sum_{p=0}^{P-1} 2^p s(i_p - i_c) \]  

(1)

Among them, \((x_c, y_c)\) is the center pixel, \(i_c\) is the gray value of the center point, \(P\) is the number of adjacent sample points, \(i_p\) is the gray value of the adjacent \(p\)th pixel, \(s\) is a flag operation function, and the definition of \(s\) is shown in formula (2).

\[ s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \]  

(2)

LBP texture features can adapt to various application scenarios. The most important feature of LBP texture features is that it is insensitive to factors such as uneven illumination and rotation, and it has good robustness. Another reason why the LBP algorithm is selected as the feature algorithm for real-time video image extraction is that the LBP algorithm is efficient and has good real-time performance. A learner image is selected and the LBP value of each point can be calculated through the above method. Then the LBP histogram of the whole image can be counted as the texture feature of the image.

2.3. Design of learner face recognition algorithm based on AdaBoost in online learning environment

Since the online learning environment is affected by factors such as face position, face size, lighting conditions, and face posture, the face is inconsistent, so the face image needs to be normalized.

First, geometric normalization is performed. Assume that the width of face is 2d and the midpoint is O. Based on O, the facial features need to be cut 0.5d in the vertical direction, 1.5d in the vertical direction, and d in the left and right directions.

Then, the gray level normalization is performed, and the face image is grayed out by different illumination intensity and direction. It is assumed that the face color image is determined by three channel components R, G, and B. When R=G=B, it indicates for the same grayscale color. The gray value used in this study is as shown in equation (3):

\[ Gray = 0.229 \times R + 0.578 \times G + 0.114 \times B \]  

(3)

The size of the image is set to be M×N and the matrix of the grayscale image be L (x, y), then the mean and variance of the grayscale of the obtained image are calculated from equation (4) and equation (5):

\[ \bar{\mu} = \frac{1}{M \times N} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} I(x, y) \]  

(4)

\[ \sigma^2 = \frac{1}{M \times N} \sum_{y=0}^{N-1} \sum_{x=0}^{M-1} (l(x, y) - \bar{\mu})^2 \]  

(5)
When $\mu_0 = 0, \sigma_0 = 1$, the pixel is calculated by equation (6), and the calculated illumination is a uniform grayscale image of the face.

$$I(x, y) = \frac{\sigma_0}{\sigma} = (I(x, y) - \bar{I}) + \mu_0 (0 \leq x < M, 0 \leq y < N) \quad (6)$$

By normalizing the face by the AdaBoost algorithm described above, it is possible to obtain that each person's eyes are on the same $y$ coordinate of the picture, thereby ensuring that each person's face can be aligned.

### 2.4. Face feature tracking based on CLM algorithm

The CLM algorithm includes two parts of the shape model and the texture model near the feature points, and the fitting equation is as shown in equation (7):

$$x = \bar{x} + P_s b_s \quad (7)$$

$$g = \bar{g} + P_g b_g \quad (8)$$

In the equation, $\bar{x}$ is the average shape, $P_s$ is the principal component matrix of the shape model, $b_s$ and $b_g$ distribution is the different non-rigid change produced the shape model and the average shape of the local texture model. The shape model and texture model are transformed into equation (9) by the PCA algorithm:

$$b = P_c c \quad (9)$$

In the equation, $P_c = (P_{cs} P_{cg})^T$ and $b = (W_s b_s \ b_g)^T$. $b$ not only represents the link shape but also represents the texture parameter vector, $W_s$ is the adaptive weight, which represents the difference between the shape unit and the texture unit, and $c$ is the combined parameter, $P_c$ represents the orthogonal matrix obtained by the PCA algorithm from the shape parameter PCS and the texture parameter PCS.

### 2.5. Real-time learning emotion recognition in online learning environment

This section mainly introduces real-time learning emotion recognition in the online learning environment, and Figure 2 is a framework diagram of learning emotion recognition algorithms. The expression recognition module is introduced from three aspects: feature point tracking, learning expression classification of recognition and avoidance and excitation detection.

![Figure. 2 Learning Emotion Recognition Algorithm Framework.](image)

In the three-dimensional learning emotion model, the judgment of recognition degree and avoidance is based on the learner's expression, while the main determination of excitement degree is based on the degree of open and close of eyes. In the learning process, the open and close motions of the left and right eyes can be regarded as synchronous. In this study, PERCLOS standard is used to detect the excitability of the learner's right eye. PERCLOS was first proposed by Walt Wiewille in 1994, defining...
the proportion of time that the eye is in a closed state for a certain period of time. The core idea is to study the relationship between human eyes, optical variables and fatigue, analyze the relationship between human eyes' open-close condition and fatigue degree, and draw a conclusion that PERCLOS criterion based on the proportion of human eyes' closing time can accurately and quickly judge the fatigue state of learners.

In order to detect learner excitement, firstly, the fatigue degree of learners should be calculated. P80 method in PERCLOS criterion is selected in this research. Suppose that within a certain period, the opening and closing time of eyes is shown in figure 3. Among them, t1 represents the time taken from the full opening of the eye to the opening of 80%, and t2 represents the time taken from the full opening to the opening of 20%, t3 is the time taken from fully opening to 20% opening and then to 20% opening again, t4 indicates the time taken from fully opening to 80% opening and then to 80% opening again.

\[ f = \frac{t_3 - t_2}{t_4 - t_1} \times 100\% \]  

(10)

In the formula, t3-t2 represents the duration of the learner's closed eyes, and t4-t1 represents the duration of change of the learner's eye state.

3. Results and discussion

3.1. The algorithm test of learner's face feature extraction

In this study, the test face database contains a total of 1500 184 × 286 pixel grayscale images of 25 people, all of which are frontal face images, and the background difference is large. This experimental algorithm is implemented by OpenCV2.4.10 and written in Python2.7. The test results are shown in table 1.

| Number of faces detected | The right ratio | Number of residual | Miss rate | Number of faces detected by mistake | Error detection rate |
|--------------------------|-----------------|--------------------|-----------|-------------------------------------|---------------------|
| 1425                     | 95%             | 67                 | 4.47%     | 8                                   | 0.53%               |

As can be concluded from the test results, even in the case of multi-face recognition in a complex background, this algorithm is able to extract learners' facial features well and show good results, providing an important basic condition for the implementation of online learning emotion detection system.
3.2. Test of learner's excitement degree
Currently, there are three methods commonly used to identify human eye state: grayscale template matching, Hough transform detection iris and basic pattern classification. These three methods have their own advantages in different application scenarios. In this study, CLM algorithm is used to track various facial organs, so that the eye state can be calculated by the aspect ratio of the human eye. As shown in figure 4, six key points p1, p2, p3, p4, p5, and p6 around the human eye are selected, and the aspect ratio $\lambda$ of eye is calculated, as shown in the formula (11).

$$\lambda = \frac{||p_2-p_6|| + ||p_3-p_5||}{2||p_1-p_4||}$$

(11)

According to the statistics of the aspect ratio in a large number of human eye images, the range when eyes are opened and closed is $0.22 \leq \lambda_{\text{open}} \leq 0.27$, $0.05 \leq \lambda_{\text{close}} \leq 0.1$, respectively. The change in the aspect ratio $\lambda$ of a blink is shown in figure 5.

![Figure 4 Key point selection](image)

![Figure 5 Variation of height to width ratio $\lambda$](image)

The PERCLOS value is the ratio of the duration of the blink process to the time to close the eyes. Assuming that the time interval collected during online learning is equal, the number of frames of the learner's image is used to calculate the learner's fatigue state. Therefore, the learner fatigue status detection can be finally obtained by formula 5, where $N\lambda_{0.1}$ is the frame number when $\lambda$ is less than or equal to 0.1, and $N\lambda_{0.22}$ is the frame number when $\lambda$ is less than or equal to 0.22.

$$f = \frac{N\lambda_{0.1}}{N\lambda_{0.22}} \times 100\%$$

(12)

According to the definition of PERCLOS, the value range of $f$ is [0, 1]. Studies by Walt Wiewille show that PERCLOS values are much lower in excitation than in fatigue. When people are excited, the $f$ value is generally between 0 and 0.15, while when they are in the state of fatigue, the $f$ value will exceed 0.5, and when they sleep, $f=1$. Therefore, the degree of excitement of the learner can be judged according to the range of values of $f$.

Finally, based on the collected data combined with the PERCLOS standard, it can be concluded that the system can accurately detect the learner's excitement.

4. Conclusion
In order to realize the emotion detection system of online learning and ensure the high practicability and robustness of the system, a large amount of data in the field of face recognition is studied and verified.
in the specific environment of online learning. Through a large number of experiments, the selection, design and system implementation of the algorithm are completed. Finally, based on the 3d emotion model, the online learning emotion detection system is realized through face detection technology, expression recognition technology, excitement detection technology, etc., and the advantages and uniqueness of the system are verified through relevant tests. Among them, face detection is the premise and basis of all research work.

Although the shortcomings of the current online teaching at home and abroad are summarized, and the online learning system is improved and established based on it, there are still some areas to be improved in the system. For example, the extraction of facial expression features is related to the speed and accuracy of facial expression recognition, and the algorithm proposed in this research needs to be further improved when dealing with non-frontal faces. Therefore, the system will be further improved in the future to make it play a better role and enable students to have higher learning efficiency.

Acknowledgement
The article is supported by education reformation project (Code Number: 2019JG13).

References
[1] Fakhrmoosavy S H, Setayeshi S, and Sharifi A. A modified brain emotional learning model for earthquake magnitude and fear prediction. Engineering with Computers, 2018, 34(2), pp. 261-276.
[2] Mazitov T, Bregin A, Philips M A, et al. Deficit in emotional learning in neurotrimin knockout mice. Behavioural Brain Research, 2017, 317, pp. 311-318.
[3] Bollegala D. Dynamic feature scaling for online learning of binary classifiers. Knowledge-Based Systems, 2017, 2012(1), pp. 53-58.
[4] Jansen R S. Validation of the self-regulated online learning questionnaire. Journal of Computing in Higher Education, 2017, 29(1), pp. 1-22.
[5] Wu W C V, Hsieh J S C, and Yang J C. Creating an online learning community in a flipped classroom to enhance EFL learners’ oral proficiency. Journal of Educational Technology & Society, 2017, 20(2), pp. 142-157.
[6] Thaker P, Tenenbaum J B, and Gershman S J. Online learning of symbolic concepts. Journal of Mathematical Psychology, 2017, 77, pp. 10-20.
[7] Bahrami S, Wong V W S, and Huang J. An online learning algorithm for demand response in smart grid. IEEE Transactions on Smart Grid, 2017, 99, pp. 1-1.
[8] Cuniah C. To mooc or not to mooc: how can online learning help to build the future of higher education? Innovations in Education & Teaching International, 2017, 54(5), pp. 521-522.
[9] Kim T D, Yang M Y, Bae J, et al. Escape from infinite freedom: effects of constraining user freedom on the prevention of dropout in an online learning context. Computers in Human Behavior, 2017, 66, pp. 217-231.
[10] Zhou Q. Multi-layer affective computing model based on emotional psychology. Electronic Commerce Research, 2017, 18(1), pp. 109-124.
[11] Perugini M L L, Iglesia G D L, Solano A C, et al. The mental health continuum--short form (mhc–sf) in the argentinean context: confirmatory factor analysis and measurement invariance. Eur J Psychol, 2017, 13(1), pp. 93-108.
[12] Morgan S D, and Labowe R. Discrete ratings and dimensional judgments of emotional speech: a preliminary look at gender differences. Journal of the Acoustical Society of America, 2017, 142(4), pp. 2707-2707.
[13] Motamed S, Setayeshi S, and Rabiee A. Speech emotion recognition based on a modified brain emotional learning model. Biologically Inspired Cognitive Architectures, 2017, 19, pp. 32-38.
[14] Fakhrmoosavy S H, Setayeshi S, and Sharifi A. A modified brain emotional learning model for earthquake magnitude and fear prediction. Engineering with Computers, 2018, 34(2), pp. 261-276.