Bitlis Eren Üniversitesi Fen Bilimleri Dergisi
BİTLİS EREN UNIVERSITY JOURNAL OF SCIENCE
ISSN: 2147-3129/e-ISSN: 2147-3188
VOLUME: 11 NO: 3 PAGE: 770-778 YEAR: 2022
DOI: 10.17798/bitlisfen.1079499

An Emotion Recognition Model Using Facial Expressions in Distance Learning

Beyza Esin ÖZSEVEN[1*], Naim ÇAĞMAN[2]

[1]*Tokat Gaziosmanpaşa University, Institute of Graduate Studies, Department of Mathematics, Tokat*
[2]*Tokat Gaziosmanpaşa University, Faculty of Science and Letters, Department of Mathematics, Tokat*
(ORCID: 0000-0003-4888-8259) (ORCID: 0000-0003-3037-1868)

Keywords: Distance Learning, Emotion Recognition, Facial Expression, Emotion Classification, Emotional Analysis.

Abstract
The most important factor on the success of the student is the student's readiness for the lesson, motivation, cognitive and emotional state. In face-to-face education, the educator can follow the student visually throughout the lesson and can observe his emotional state. One of the most important disadvantages of distance learning is that the emotional state of the student cannot be followed instantly. In addition, the processing time of emotion detection, in which real-time emotion detection will be performed, should be short. In this study, a method for emotion recognition is proposed by using distance and slope information between facial landmarks. In addition, the feature size was reduced by detecting only those that are effective for emotion recognition among the distance and slope information with statistical analysis. According to the results obtained, the proposed method and feature set achieved 86.11% success. In addition, the processing time is at a level that can be used in distance learning and can detect real-time emotion.

1. Introduction

Student motivation, cognitive and emotional state are important factors on learning performance [1]. In the physical learning environment, the educator and the student are in a face-to-face relationship, and appropriate educational conditions can be provided by responding appropriately according to the emotional state of the student [2]. However, in distance learning, the structure of the learning space and content is often presented in a static manner, without consideration of students' feelings, without interaction and without feedback from the educator. Therefore, determining the student's emotional states and reactions to a given situation is one of the most important elements for any distance learning environment [3]. Information received from the student, speech, facial expression, and vital signals (heart rate, etc.) are used for emotion recognition.

Getting information from the student is the simplest method [4]. The student is asked about his feeling and action is taken according to the answer received from the student. This method will give negative results on students who try to hide their emotions. In addition, students may react negatively when faced with such questions outside of the classroom [2]. Another important point is that emotions are not constant and change in a negative situation. For example, a student who starts the lesson happily may be upset when he cannot answer the question. Therefore, the emotional state should be constantly monitored throughout the learning process. Troung et al. tested the emotional state determined by the participant with the observer. According to the results obtained, the emotional state obtained in the two situations is different [5]. Students' descriptions of their emotions were examined in various papers [5, 6]. According to the results obtained, there are differences between self-assessment and observer evaluation.

It is a known fact that people reflect their emotions in their speech [7, 8]. Verbal communication includes some acoustic cues that aid in emotion recognition [9]. Acoustic feature-based

*Corresponding author: *beyza_esin@hotmail.com* *Received: 26.02.2022, Accepted: 23.08.2022*

mostly used features are pitch, speech rate, formant frequencies, energy and phonetic features. [10]. In recent years, mel-frequency cepstral coefficients (MFCCs) have been used successfully in speech emotion recognition (SER) studies [11]–[21]. SER consists of two main parts. These are acoustic feature detection and classification. For acoustic feature detection, tools such as Praat [22], OpenSmile [23], OpenEar [24], SPAC [25]can be used, as well as coding with the help of software such as MATLAB. Classifiers such as SVM, HMM, GMM, AFDBN, ANN, k-NN, RBF are used for classification [2]. Although there are many studies on SER, there are various factors that negatively affect the success of the system. Acoustic features are affected by factors such as the speaker, sentence, and language of the speaker. Also, some emotion pairs contain similar acoustic attributes. For example, boredom and sadness have similar sound characteristics [26]. Another disadvantage of using SER in distance learning is that the SER system cannot be used actively since the student is mostly in the listening position.

Another method used for emotion recognition is facial expressions [27]. Sometimes a facial expression is more effective than speaking too many words. Emotional reactions affect the muscles under the skin of the face and facial expressions are formed. For example, in disgust, the mouth and nose muscles act in such a way as to produce facial disfigurement [28]. Facial expressions can be recognized from still images or camera images with image processing methods. These facial expressions can be used for person recognition or emotional recognition. Facial deformation, facial expression of emotional state is similar for all people, despite differences in gender, ethnicity, and age [29], [30]. Facial features fall into two categories: appearance and geometric features. In geometric-based methods, various points on the face are followed [31]. Accurate finding and tracking of facial features are imperative in geometric-based techniques. In many real applications, finding and detecting facial features is a time-consuming, complex, and error-prone process. Therefore, appearance-based techniques have also been proposed. These techniques model appearance changes such as wrinkles and furrows [2]. Compared to other emotion recognition methods, the use of facial expressions gives more successful results. The weakness for distance learning is that the student's face is mostly static during the learning process.

Physiological sensors can be used to emotion recognition with the help of various vital signals (heart rate, blood pressure, respiration, electrodermal activity, electroencephalograms, electrocardiograms, and electromyograms etc.) [32]. Thus, changes in physiological signals can be associated with emotion [33]. Emotion recognition based on vital signals is not usable for distance learning as it requires special hardware. When wearable technologies become widespread, their usability for distance learning will increase. A comparison of emotion detection methods in distance education is given in Table 1.

Table 1. Advantages and disadvantages of emotion recognition methods in distance education [2].

Method	Advantages	Disadvantages
Getting information from the student	- Easy to apply - No need for additional equipment	- It is boring for the student - Student's emotional changes cannot be detected during the lesson. - The student may not be honest
SER	- Suitability for language learning - Distinguish between high arousal and low arousal feelings	- Not available for all learning content - Inability to distinguish emotions with the same arousal
Facial Expression	- Identifying 7 basic emotional states - Availability of webcams	- Processes are time consuming - There may not be enough change in the faces of the students.
Vital Signals	- Distinguish between high arousal and low arousal feelings - High precision if using many parameters	- Needs some special tools and sensors not available for all students - Connecting sensors may disturb the student

Vital signals from the methods used for emotion recognition as described above are not available for distance learning. According to the information received from the student, emotion recognition will not give accurate results for distance learning as it will be both misleading and stable. Although emotion recognition through speech has achieved high success in various studies, its usability in distance learning is low as the student will mostly be in the listening position in the course. Emotion recognition through facial expressions has been determined as the most suitable method that can provide efficient results for distance learning systems. In this study, a model based on facial points for emotion recognition was created using image processing, analytical geometry, statistics and SVM classifier, and the success of the model was analyzed. In the next part of the study, the relevant studies in the literature, the materials and methods used in the third part, the analysis of the proposed model in the fourth part, and the results and discussion are given.

2. Related Works

Emotion recognition from facial expressions consists of image acquisition, preprocessing, feature detection, classification and postprocessing processes. Most systems based on facial expressions classify them into categories of joy, sadness, surprise, fear, anger, disgust, and nervousness [34]. Feature extraction is the process of obtaining features that will be used to distinguish emotions from images. For this process, the features obtained by filtering methods over the model or face image that refer to the change of various points on the face are used. Gabor filters [35], principal component analysis [36], linear discriminant analysis [37] and facial action coding (a system for classifying human facial movements according to their facial appearance) [38] are used to obtain the features [39]. After the features are obtained, a classifier is needed for emotion recognition. The most widely used classifier at this stage are SVM [39].

Gabor wavelet [39], SVM [40, 41], k-NN [41], deep learning [42–46] methods were used in studies using image processing for emotion recognition using facial expressions in distance learning. According to the results of the study, in which gabor wavelet, principal component analysis and linear discriminant analysis were used for filtering and feature extraction, the highest success was obtained with gabor wavelet [39]. In addition, it has been found that deep learning methods increase the success rate between 1% and 5% compared to traditional methods [42]. Image processing-based methods operate on the image of the entire face. Another method used to detect facial expressions is the use of images of the eye and mouth region [47].

The relationships between facial expressions and learning styles in distance learning were also analyzed by image processing methods. In the study using deep learning, a database was created showing which emotion is associated with which learning style [43]. Facial expressions are also used to determine the current education level of the student. To determine the education level of the student, besides facial expressions, exam results, current level and time spent in the test were used [48]. In the study, in which the distance education environment was modeled using facial expressions and fuzzy logic, the answers given by the students to the questions and their emotional states were used. According to the facial expressions obtained here and the answers given by the student, the next learning level of the student was determined [44]. In another deep learning-based study, facial expressions were used to suggest suitable materials to students [46].

There are studies that refer to specific points on the face, facial muscle movements and moving points on the face for emotion recognition. Ayvaz et al. (2017), the coordinates of the points determined on the percentage were determined as features, and the highest success in classification was obtained with SVM and k-NN [41]. The Facial Action Coding System (FACS) [49], which references facial muscle movements, and the facereader [50], which detects moving points on the face, were used in the study aiming to evaluate the emotional state of students during the course in distance learning [51].

According to the results of studies on emotion detection in distance learning, machine learning methods and image processing have been used in many studies. Depending on the feature size, these methods require both a lot of workload and hardware. In addition, emotion recognition for distance learning systems is a process that requires continuity and needs to be done in real time. Therefore, emotion recognition system with low processing complexity and high success rate will increase the performance of distance learning systems.

3. Materials and Methods

A subset of the FACES [52] dataset, which is available for use and consists of 72 images, was used to test the proposed emotion recognition model. This dataset contains 6 emotional expressions (Anger, Disgust, Fear, Happiness, Neutral, Sadness). For each image in this data set, 70 points on the face were used. These points were obtained with faceSDK [53].

FaceSDK can be used in many programming languages (Microsoft Visual C++, C#, Objective C, Swift, Java, VB, Delphi, and Python) is a plugin. An example of face points obtained with FaceSDK is given in Figure 1.

Figure 1. Face dots obtained with the faceSDK for an example image in the FACES dataset.

Before the details of the proposed emotion recognition model, definitions of some basic concepts are given below.

Definition 1 (Emotional State)**.** $E = \{e_1, e_2, \dots, e_n\}$ be the set containing n different facial expressions. Emotional state $\sigma = (\epsilon_1, \epsilon_2, \dots, \epsilon_n)$ is defined as a probability distribution over E with [44]

$$\epsilon_1 = P(E = e_i), 0 \leq \epsilon_i \leq 1 \; for \; 1 \leq i \leq n \; and \; \sum_i \epsilon_i = 1 \qquad (1)$$

According to Definition 1, σ contains n facial expressions, and in the proposed approach, the emotional state includes probability degrees of six facial expression classes: anger (ϵ_1), disgust (ϵ_2), fear (ϵ_3), happiness (ϵ_4), sadness (ϵ_5) and neutral (ϵ_6).

Definition 2 (Facial Landmarks)**.** Facial landmarks (*fp)*, contain specific points on the face to be derived from each facial expression. These points are obtained with faceSDK [53] and defined as

$$fp = \{p_1, p_2, \dots, p_k\}, 1 \leq k \leq 70, for \; \forall E \qquad (2)$$

Definition 3 (Distance between landmarks). The distance between landmarks (*dist*) expresses the distance between each landmark obtained on the percentage and the other point. Analytical geometry is used to calculate these distances and is defined as

$$\text{Distance: } |AB| = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \; for \; A = (x_1, y_1) \; and \; B = (x_2, y_2) \qquad (3)$$

$$\text{Distance between landmarks: } dist = \{|p_1p_2|, |p_1p_3|, \dots, |p_ip_j|\}, 1 \leq i \leq 69, 1 \leq j \leq 70 \; for \; \forall fp \; and \; i \neq j \qquad (4)$$

Definition 4 (Slope between landmarks). The slope between landmarks (*slp*) expresses the slope between each landmark obtained in the percentage and the other point. These slopes are defined as

$$\text{Slope: } m = \frac{y_2 - y_1}{x_2 - x_1} \; for \; A = (x_1, y_1) \; and \; B = (x_2, y_2) \qquad (5)$$

$$\text{Slope between landmarks: } slp = \left\{\frac{p_{y_2} - p_{y_1}}{p_{x_2} - p_{x_1}}, \frac{p_{y_3} - p_{y_1}}{p_{x_3} - p_{x_1}}, \dots, \frac{p_{y_j} - p_{y_i}}{p_{x_j} - p_{x_i}}\right\}, 1 \leq i \leq 69, 1 \leq j \leq 70, for \; \forall fp \; and \; i \neq j \qquad (6)$$

Definition 5 (Detection of emotion-based landmarks). Detection of emotion-based landmarks is the detection of emotion with the help of features obtained using the information obtained from Definitions 3 and 4. However, the distance and slope data between points over 70 points in each face image increase the feature size. Therefore, those features that do not vary according to emotion can be statistically detected and removed from the feature set. Since the distance and slope data do not show normal distribution, the Kruskal-Wallis test can be used for this purpose.

The Kruskal-Wallis test (KW) is defined as

$$KW = \left[\frac{12}{N(N+1)} \sum t\bar{R}^2\right] - 3(N+1) \qquad (7)$$

where N is the total number of people/observations in the groups, g is the number of groups, $\bar{R}$ is the group mean rank, t is the number of people in each group.

Detection of emotion-based landmarks: Detection of emotion-based landmarks are defined as

$$DEL = \{KW(dist, slp)\}, p \leq 0.05 \qquad (8)$$

The information given from Definition 1 has been created for the system to learn in the emotion recognition process. In the learning process, emotion-based facial points are obtained in line with the definitions given above by using the face points obtained with the faceSDK in the ready-made facial expression dataset. The resulting face points are used for training the SVM. The block diagram of the emotion recognition model is given in Figure 2.

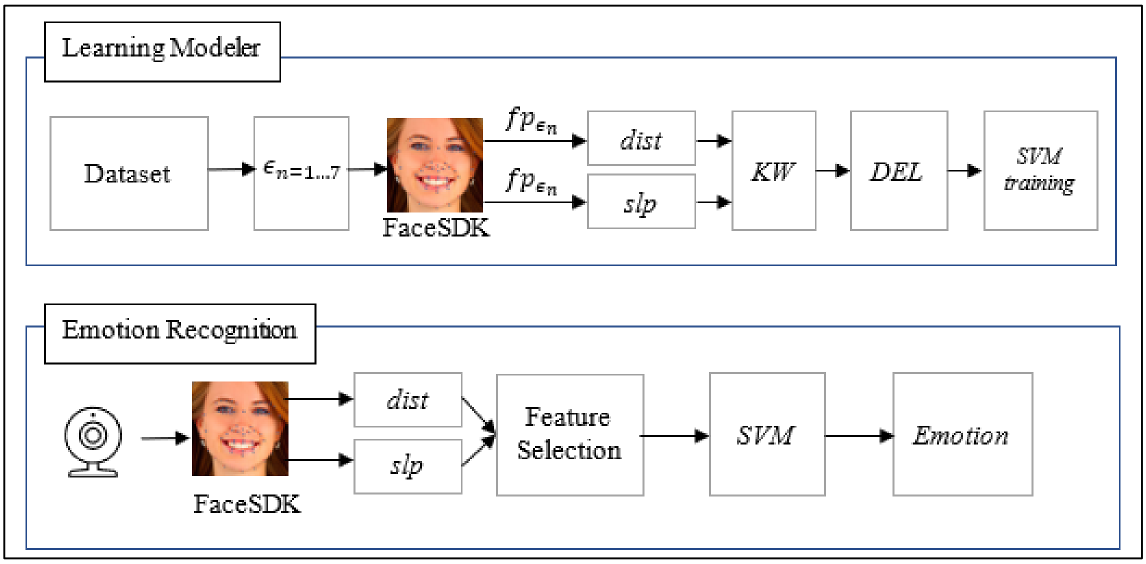

Figure 2. Emotion recognition model

The model given in Figure 2 has two stages: learning modeling and emotion recognition. The learning modeler must be activated once the system is first used and is used to detect DELs on ready datasets and train the classifier. The parameters to be used for SVM should be determined at the learning modeler stage. The emotion recognition stage should be used for real-time emotion detection when using the distance learning system. Here, emotion recognition is done by reducing the feature set and classifying it with the trained SVM.

4. Experimental Results

The slope and distance information of 70 facial landmarks obtained from each image were calculated and 4830 (*slp: slope, dist: distance*) features were obtained for each image. While calculating the slope value at the overlapping landmarks, infinite values depending on the 0 value were obtained. The features containing these values were removed from the data set and the number of features was reduced to 4618.

The Kruskal-Wallis test was applied to determine the remaining features related to emotional expression. Since the significance level is generally 95% in statistical analysis, the significance level was determined as 95% for Kruskal-Wallis. As a result of the analysis, 3328 of 4618 features were found to be related to emotional expressions.

SVM classifier was used to test the emotion recognition success of the dataset containing all features and reduced features, and the tests were carried out with WEKA [54]. The LibSVM library in WEKA was used and linear kernel was used. In addition, the normalization results are compared to eliminate measurement differences and reduce the classifier processing time. In the normalization process, the values of the features are normalized between 0-1. The results obtained are given in Table 2.

Table 2. Success rates of the proposed emotion detection model

Features	Normalization	Success Rate	Processing Time
4618	No	80.55%	0.27 seconds
3328	No	80.54%	0.08 seconds
4618	Yes	86.11%	0.18 seconds
3328	Yes	8472%	0.05 seconds

According to the results given in Table 2, distance and slope information can be used for emotion recognition. Reducing the feature set with Kruskal-wallis reduces the processing time by approximately 70%. Since emotion recognition will be performed in real time in the distance learning system, the processing time is as important as the success rate. Facial landmarks detection has a processing time of 0.00027 seconds and emotion recognition has a processing time of 0.05 seconds. In other words, facial emotion recognition takes a total of 0.05027 seconds. However, 0.05 seconds here is the time for both the learning modeler and emotion recognition. The processing time here will vary depending on the hardware used. Analyzes were made with a laptop computer with standard hardware (i5 processor, 8 GB RAM).

5. Conclusion and Discussion

The most important disadvantage of distance learning systems is the low interaction between the educator and the student. Although this disadvantage is tried to be eliminated with live lessons, the educator cannot fully follow the cognitive and emotional state of the student. For this reason, the emotional states of students in distance learning systems are also tried to be followed by the educator. For emotion recognition, information from the student, voice analysis, facial analysis and vital signals are used. Among these methods, the most suitable method for the student in distance learning is the use of facial expressions. Emotion recognition speed is as important as emotion recognition success in detecting emotion from facial expressions. Because instant emotion detection is necessary in live lessons. Therefore, the methods and models used for emotion recognition are important.

The aim of this study is to perform real-time and high-speed emotion recognition from facial expressions in distance learning systems. For this purpose, distance and slope information between facial landmarks were used. Effective ones from these landmarks were obtained by statistical analysis and the feature size was reduced. In addition, the processing time was shortened by normalization.

In the literature in recent years, deep learning-based methods have come to the fore. However, the computational complexity of these methods is higher than traditional methods such as SVM, k-NN. For this reason, the SVM classifier was used in the proposed model and the emotion recognition success was 86.11% with 0.18 seconds processing time and 84.72% with 0.05 seconds processing time. These times include the sum of both training and testing time of the model. Since the distance learning system will be carried out on server computers with special hardware, this period will be shortened even more. The learning modeler will be used only once after the distance learning system is activated, and emotion recognition will be performed with the training data there. Therefore, the processing time for emotion recognition will be approximately 0.00327 (0.00027 facial landmarks, 0.003 emotion recognition) seconds while the system is running. Since the Kruskal-Wallis test is used only for the learning modeler, the processing time is not considered. In addition, the normalization process also has very low processing time, which can be ignored.

Standard cameras have 30 fps and capture 30 frames per second. In other words, each frame is obtained in 0.033 seconds. Considering the image formation time and the time spent for emotion recognition, emotion recognition can be performed on real-time images simultaneously and without delay with the methods used.

The proposed model can be used for continuous monitoring of students' emotional states in live lessons, as well as exams in distance education, determination of readiness level and identification of students with the help of face landmarks.

Authors' Contributions

The authors' contributions to the paper are equal.

Statement of Conflicts of Interest

There is no conflict of interest between the authors.

Statement of Research and Publication Ethics

The authors declare that this study complies with Research and Publication Ethics

References

[1] N. Saberi and G. A. Montazer, 'A new approach for learners' modeling in e-learning environment using LMS logs analysis', in *6th National and 3rd International conference of e-Learning and e-Teaching*, 2012, pp. 25–33.

[2] M. Imani and G. A. Montazer, 'A survey of emotion recognition methods with emphasis on E-Learning environments', *Journal of Network and Computer Applications*, vol. 147, p. 102423, 2019.

[3] C. Villiger *et al.*, 'Effectiveness of an extracurricular program for struggling readers: A comparative study with parent tutors and volunteer tutors', *Learning and Instruction*, vol. 60, pp. 54–65, 2019.

[4] A. A. Kardan and Y. Einavypour, 'Multi-Criteria Learners Classification for Selecting an Appropriate Teaching Method', in *Proceedings of the World Congress on Engineering and Computer Science*, 2008, pp. 22–24.

[5] K. P. Truong, D. A. Van Leeuwen, and F. M. De Jong, 'Speech-based recognition of self-reported and observed emotion in a dimensional space', *Speech communication*, vol. 54, no. 9, pp. 1049–1063, 2012.

[6] C. Busso and S. S. Narayanan, 'The expression and perception of emotions: Comparing assessments of self versus others', presented at the Ninth annual conference of the international speech communication association, 2008.

[7] D. Morrison and L. C. De Silva, 'Voting ensembles for spoken affect classification', *Journal of Network and Computer Applications*, vol. 30, no. 4, pp. 1356–1365, 2007.

[8] N. Sadoughi and C. Busso, 'Speech-driven animation with meaningful behaviors', *Speech Communication*, vol. 110, pp. 90–100, 2019.

[9] E. Mendoza and G. Carballo, 'Vocal tremor and psychological stress', *Journal of Voice*, vol. 13, no. 1, pp. 105–112, 1999.

[10] M. Pantic and L. J. Rothkrantz, 'Toward an affect-sensitive multimodal human-computer interaction', *Proceedings of the IEEE*, vol. 91, no. 9, pp. 1370–1390, 2003.

[11] H. Cao, R. Verma, and A. Nenkova, 'Speaker-sensitive emotion recognition via ranking: Studies on acted and spontaneous speech', *Computer speech & language*, vol. 29, no. 1, pp. 186–202, 2015.

[12] F. Chenchah and Z. Lachiri, 'A bio-inspired emotion recognition system under real-life conditions', *Applied Acoustics*, vol. 115, pp. 6–14, 2017.

[13] W. Dai, D. Han, Y. Dai, and D. Xu, 'Emotion recognition and affective computing on vocal social media', *Information & Management*, vol. 52, no. 7, pp. 777–788, 2015.

[14] C.-C. Lee, E. Mower, C. Busso, S. Lee, and S. Narayanan, 'Emotion recognition using a hierarchical binary decision tree approach', *Speech Communication*, vol. 53, no. 9–10, pp. 1162–1171, 2011.

[15] K. Mannepalli, P. N. Sastry, and M. Suman, 'A novel adaptive fractional deep belief networks for speaker emotion recognition', *Alexandria Engineering Journal*, vol. 56, no. 4, pp. 485–497, 2017.

[16] S. Mariooryad and C. Busso, 'Compensating for speaker or lexical variabilities in speech for emotion recognition', *Speech Communication*, vol. 57, pp. 1–12, 2014.

[17] V. V. Nanavare and S. K. Jagtap, 'Recognition of human emotions from speech processing', *Procedia Computer Science*, vol. 49, pp. 24–32, 2015.

[18] C. S. Ooi, K. P. Seng, L.-M. Ang, and L. W. Chew, 'A new approach of audio emotion recognition', *Expert systems with applications*, vol. 41, no. 13, pp. 5858–5869, 2014.

[19] T. Özseven, 'A novel feature selection method for speech emotion recognition', *Applied Acoustics*, vol. 146, pp. 320–326, 2019.

[20] T. M. Rajisha, A. P. Sunija, and K. S. Riyas, 'Performance analysis of Malayalam language speech emotion recognition system using ANN/SVM', *Procedia Technology*, vol. 24, pp. 1097–1104, 2016.

[21] P. Vasuki and C. Aravindan, 'Improving emotion recognition from speech using sensor fusion techniques', in *TENCON 2012 IEEE Region 10 Conference*, 2012, pp. 1–6.

[22] P. Boersma, 'Praat, a system for doing phonetics by computer', *Glot. Int.*, vol. 5, no. 9, pp. 341–345, 2001.

[23] F. Eyben, M. Wöllmer, and B. Schuller, 'Opensmile: the munich versatile and fast open-source audio feature extractor', in *Proceedings of the 18th ACM international conference on Multimedia*, 2010, pp. 1459–1462.

[24] F. Eyben, M. Wöllmer, and B. Schuller, ‘OpenEAR—introducing the Munich open-source emotion and affect recognition toolkit’, in *2009 3rd international conference on affective computing and intelligent interaction and workshops*, 2009, pp. 1–6.

[25] T. Özseven and M. Düğenci, ‘SPeech ACoustic (SPAC): A novel tool for speech feature extraction and classification’, *Applied Acoustics*, vol. 136, pp. 1–8, 2018.

[26] A. Jaimes and N. Sebe, ‘Multimodal human–computer interaction: A survey’, *Computer vision and image understanding*, vol. 108, no. 1–2, pp. 116–134, 2007.

[27] M. IqbalQuraishi, J. Pal Choudhury, M. De, and P. Chakraborty, ‘A framework for the recognition of human emotion using soft computing models’, *International Journal of Computer Applications*, vol. 40, no. 17, pp. 50–55, 2012.

[28] K. Shanmugarajah, S. Gaind, A. Clarke, and P. E. Butler, ‘The role of disgust emotions in the observer response to facial disfigurement’, *Body Image*, vol. 9, no. 4, pp. 455–461, 2012.

[29] C. Darwin, ‘The expression of the emotions in man and animals (1872)’, *The Portable Darwin*, pp. 364–393, 1993.

[30] P. Ekman and W. V. Friesen, ‘Constants across cultures in the face and emotion.’, *Journal of personality and social psychology*, vol. 17, no. 2, p. 124, 1971.

[31] D. Ghimire and J. Lee, ‘Geometric feature-based facial expression recognition in image sequences using multi-class adaboost and support vector machines’, *Sensors*, vol. 13, no. 6, pp. 7714–7734, 2013.

[32] F. Nasoz, C. L. Lisetti, K. Alvarez, and N. Finkelstein, ‘Emotion recognition from physiological signals for user modeling of affect’, presented at the Proceedings of the 3rd Workshop on Affective and Attitude User Modelling (Pittsburgh, PA, USA, 2003.

[33] S. Koelstra *et al.*, ‘Deap: A database for emotion analysis; using physiological signals’, *IEEE transactions on affective computing*, vol. 3, no. 1, pp. 18–31, 2011.

[34] L. Li, L. Cheng, and K. Qian, ‘An e-learning system model based on affective computing’, in *2008 international conference on cyberworlds*, 2008, pp. 45–50.

[35] J. G. Daugman, ‘Complete discrete 2-D Gabor transforms by neural networks for image analysis and compression’, *IEEE Transactions on acoustics, speech, and signal processing*, vol. 36, no. 7, pp. 1169–1179, 1988.

[36] M. Turk and A. Pentland, ‘Eigenfaces for recognition’, *Journal of cognitive neuroscience*, vol. 3, no. 1, pp. 71–86, 1991.

[37] K. Delac, M. Grgic, and S. Grgic, ‘Independent comparative study of PCA, ICA, and LDA on the FERET data set’, *International Journal of Imaging Systems and Technology*, vol. 15, no. 5, pp. 252–260, 2005.

[38] S. D’Mello, R. W. Picard, and A. Graesser, ‘Toward an affect-sensitive AutoTutor’, *IEEE Intelligent Systems*, vol. 22, no. 4, pp. 53–61, 2007.

[39] A. K. Oryina and A. O. Adedolapo, ‘Emotion recognition for user centred e-learning’, in *2016 IEEE 40th Annual Computer Software and Applications Conference (COMPSAC)*, 2016, vol. 2, pp. 509–514.

[40] K. Bahreini, R. Nadolski, and W. Westera, ‘Data fusion for real-time multimodal emotion recognition through webcams and microphones in e-learning’, *International Journal of Human–Computer Interaction*, vol. 32, no. 5, pp. 415–430, 2016.

[41] U. Ayvaz, H. Gürüler, and M. O. Devrim, ‘Use of facial emotion recognition in e-learning systems’, *Information technologies and teaching aids*, no. 60, вип. 4, pp. 95–104, 2017.

[42] Y. Guo, D. Tao, J. Yu, H. Xiong, Y. Li, and D. Tao, ‘Deep neural networks with relativity learning for facial expression recognition’, in *2016 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)*, 2016, pp. 1–6.

[43] F. L. Gambo, G. M. Wajiga, and E. J. Garba, ‘A Conceptual Framework for Detection of Learning Style from Facial Expressions using Convolutional Neural Network’, in *2019 2nd International Conference of the IEEE Nigeria Computer Chapter (NigeriaComputConf)*, 2019, pp. 1–5.

[44] M. Megahed and A. Mohammed, ‘Modeling adaptive E-Learning environment using facial expressions and fuzzy logic’, *Expert Systems with Applications*, vol. 157, p. 113460, Nov. 2020, doi: 10.1016/j.eswa.2020.113460.

[45] A. Pise, H. Vadapalli, and I. Sanders, ‘Facial emotion recognition using temporal relational network: an application to E-learning’, *Multimedia Tools and Applications*, pp. 1–21, 2020.

[46] S.-Y. Lin, C.-M. Wu, S.-L. Chen, T.-L. Lin, and Y.-W. Tseng, ‘Continuous Facial Emotion Recognition Method Based on Deep Learning of Academic Emotions’, *Sensors and Materials*, vol. 32, no. 10, pp. 3243–3259, 2020.

[47] D. Yang, A. Alsadoon, P. C. Prasad, A. K. Singh, and A. Elchouemi, ‘An emotion recognition model based on facial recognition in virtual learning environment’, *Procedia Computer Science*, vol. 125, pp. 2–10, 2018.

[48] H. Hesham, M. Nabawy, O. Safwat, Y. Khalifa, H. Metawie, and A. Mohammed, ‘Detecting Education level using Facial Expressions in E-learning Systems’, in *2020 International Conference on Electrical, Communication, and Computer Engineering (ICECCE)*, 2020, pp. 1–6.

[49] P. Ekman and W. V. Friesen, *Facial action coding system: Investigator’s guide*. Consulting Psychologists Press, 1978.

[50] M. J. Den Uyl and H. Van Kuilenburg, ‘The FaceReader: Online facial expression recognition’, in *Proceedings of measuring behavior*, 2005, vol. 30, no. 2, pp. 589–590.

[51] M. Soltani, H. Zarzour, M. C. Babahenini, M. Hammad, A.-S. Mohammad, and Y. Jararweh, ‘An emotional feedback based on facial action coding system for MOOCs with computer-based assessment’, in *2019 Sixth International Conference on Social Networks Analysis, Management and Security (SNAMS)*, 2019, pp. 286–290.

[52] N. C. Ebner, M. Riediger, and U. Lindenberger, ‘FACES—A database of facial expressions in young, middle-aged, and older women and men: Development and validation’, *Behavior research methods*, vol. 42, no. 1, pp. 351–362, 2010.

[53] ‘Luxand - Detect and Recognize Faces and Facial Features with Luxand FaceSDK’. https://www.luxand.com/facesdk/ (accessed Jun. 16, 2021).

[54] I. H. Witten, E. Frank, L. E. Trigg, M. A. Hall, G. Holmes, and S. J. Cunningham, ‘Weka: Practical machine learning tools and techniques with Java implementations’, *(Working paper 99/11). Hamilton, New Zealand: University of Waikato, Department of Computer Science.*, 1999.