A Framework for the Development of Computational Model of Emotions Through Facial Expression Recognition

Deepak Singh Chouhan[1*] and Meena Sharma[2]

[1,2] Institute of Engineering and Technology, Devi Ahilya Vishwavidyalaya, Indore, India
[*1]deepak.schouhan@yahoo.com,[2]msharma@ietdavv.edu.in

Abstract. Emotions play a vital role in human existence as they reflect an individual's internal state. Currently, emotion recognition is extensively employed to obtain automated feedback from individuals. Among the various methods available, facial expression recognition is considered the most effective means of capturing human emotions. Computational models of emotions are automated systems designed to accomplish the task of recognizing emotions and deriving meaningful insights from them. This paper compares the various algorithms used for emotion recognition and presents a framework that can serve as a foundation for developing a computational model of emotions, providing a concise overview of all its components.

Keywords: Emotion, Emotion Recognition, Facial Expression Recognition, Face Detection, Computational Model of Emotions.

1 Introduction

A Computational model of emotions is the software system that attempts to explain the phenomena of emotions. Such models implement the computational procedures for the evaluation of emotional impulses, elicitation of emotions and generation of emotional behaviors [1]. It can have so many applications like:

1. **Security:** It can be used for the security of places like schools to prevent violence.
2. **Education:** For the development of effective educational tools by observing students' expressions.
3. **HR Assistance:** Companies can use computational models of emotions in determining whether the candidate is truly interested in the position.
4. **Differently Abled Children:** It can also be used for the assistance of differently abled children like the Google Glass Smart Glass Project that help autistic children to interpret the feeling of people around them.
5. **Usability Testing:** Emotional expressions can indicate the ease of use and efficiency of user interfaces.
6. **Audience Engagement:** Companies can also use it in determining business outcomes in terms of the audience's emotional responses.

S. Bhalerao et al. (eds.), *Proceedings of the International Conference on Recent Advancement and Modernization in Sustainable Intelligent Technologies & Applications (RAMSITA-2025),* Advances in Intelligent Systems Research 192,
https://doi.org/10.2991/978-94-6463-716-8_32

Emotions are the reactions of humans that they experience due to either a reaction of some internal stimuli or an event occurring in their environment. Emotions are present in many aspects of our lives, and they are also a fundamental factor in education as well [2]. We can use emotions to analyze the students' interest and their mood in the classroom lectures [3]. By analyzing the students' emotions, we can identify whether the students feel interested or bored at a particular time of lecture or on a particular content [4]. We can utilize this analysis in deciding whether we need some improvements in our lecture content or not. Overall, it can help us in the improvement of the teaching and learning process [5].

The most common ways through which emotions are captured can be analysis of body posture and movements, analysis of speech patterns, analysis of gesture, and analysis of facial expressions. Among all such methods of emotion recognition, facial expression recognition is a quick way to understand the emotional symptoms [6]. The different expressions of eyes, lips, and face can be detected in real-time or can be recorded for further processing. Based on the detection, they can be compared to the predetermined map of shapes and compared to the data stored in the database to recognize the emotions [7]. The recognized emotions can be used to infer meaningful information and this inferred information can be used in many ways like:

- To develop a feedback system for the improvement in course content and lecture delivery [8].
- In the analysis of the impact of various teaching aids (Visualization, Presentations, etc.) in teaching.
- Analysis of the impact of learning in online and offline classes [9].
- Analysis of student behavior.
- It can be used in effective career guidance for students.
- Through negative emotions, we can identify the students who are inclined to depression and alerts can be generated for further counseling of such students[10].

This paper compares the various algorithms used for emotion recognition based on algorithms used for face detection and emotion classification and suggests a framework designed to aid developers in creating a computational model of emotions (CME). The resulting system has the potential to extract valuable insights from students' emotions by employing facial expression recognition, contributing to enhancements in the teaching and learning experience.

2 Related Work

Bartlett et al developed an automatic facial expression recognition system with the help of a modified version of the Viola-Jones algorithm for face detection and support vector machine for facial expression recognition [11]. They also highlighted the need for a spontaneous facial expression database and also developed a dataset of spontaneous facial expressions coded using Facial Action Coding System (FACS). Whitehill et al used this automatic facial expression recognition system to measure the difficulty of a delivered lecture perceived by the students and to determine the speed at which the lecture material should be delivered [12]. In another work,

Whitehill et al. performs a study on students' engagement recognition, performs a comparative study of various computer vision techniques for automatic engagement recognition and develops a system to automate the process of engagement recognition [13]. A system to automatically analyze students' expressions in the classroom environment was proposed by Tang et at that uses AdaBoost for face detection, Uniform Local Gabor Binary Pattern Histogram Sequence (ULGBPHS) for facial expression recognition and KNN for classification [5]. Sahala and Kumar performed an analysis on the emotions of students in a classroom environment and proposed a method based on deep learning for the emotion recognition of students [3]. The recognized emotions are then mapped with classroom events like sleeping, laughing, watching outside, thinking, etc. Krithika and LaxmiPriya developed a system to monitor the concentration level of students in an e-learning environment that works by detecting the eye movement and head rotation of students during an online lecture [6]. Li and wang proposed an emotion recognition model for online learning that uses facial expressions and eye blink frequency to understand the overall emotional state of learners [4]. In another work a system to detect the emotional state of learners in an e-learning environment and give feedback to the educator was proposed by Avyaz et al [14]. Prasaad et al proposed a model of emotion recognition using facial expression recognition that works in a distance learning environment [15]. They add mouth features with eye features to improve the accuracy of the proposed solution. A modified version of Viola-Jones algorithm for face detection as well as emotion recognition was proposed by ChandraKirana et al to capture students' emotions in the classroom environment [16]. Alkabbany et al proposed a framework for automatic engagement classification that can work in both online as well as offline learning environments [9]. Putra et al performed a study to monitor students' mood through classroom activity using facial expression analysis that can be used to detect negative emotions like depression [10]. Lasri proposed a method to recognize the facial emotions of students using Viola-Jones and CNN [8]. Tonguc et al examined how the emotions of the students changed during a lecture and also analyzed the effect of factors such as department, genders, class time, location of computers, etc. on student emotions [20]. A comparative study of the existing approaches for facial expression recognition was performed by Bouhlal et al and an architecture that can be used to develop facial emotions recognition systems [19]. Wang et al proposed a framework for real-time monitoring of students' emotions in an online environment [1]. In their work Osuna analyzed various computation models of emotions from the software engineering perspective and identified some challenges that are to be addressed in order to take advantage of software engineering principles for the development of Computational Models of Emotions [20]. In another work Osuna et al proposed an architectural design pattern for the development of Computational Models of Emotion that can be used to categorize and describe the components of Computational models of Emotions [21]. Kuntzler et al performed a comparative study on the performance of different facial expression recognition systems and evaluated them on acted standardized and non – standardized emotional facial expression recognition datasets [16]. Swadha et al introduces a system for the detection of engagement that will autonomously identify learner participation in online educational settings in real-time through facial emotion recognition. The system categorizes student emotions using deep learning architectures such as Inception-V3, VGG19, and ResNet-50 and

computes an engagement index to measure levels of engagement [22]. Aly et al. presented an advanced facial expression recognition system for online learning, using the optimized ResNet-50 with attention mechanisms and refined residual modules. Their method achieved very high accuracy on RAF-DB (87.62%) and FER2013 (88.13%), which allows dynamic teaching adjustments to enhance the quality of online education [23]. Tang et al. proposes a novel machine learning method that helps to measure students' emotional learning engagement in real classrooms. The proposed facial expression recognition system, based on the perception network, effectively predicts academic achievements and provides fine-grained tracking of students' emotional engagement during different teaching activities [24]. Following table summarizes the work done already in the field:

Table 1. Summary of related work

Title (Author(s))	Algorithm Used for Face Detection	Algorithm Used for Emotion Classification	Dataset(s) Used
Automatic Recognition of Facial Actions in Spontaneous Expressions (Bartlett et al) [11]	Modified Viola-Jones	SVM	CK+
Automatic Facial Expression Analysis of Students in Teaching Environments (Tang et al) [5]	Viola-Jones	KNN	FER2013
Classroom Teaching Assessment Based on Student Emotions (Sahla et al) [3]	Viola-Jones	CNN	AffectNet
Research on Learner's Emotion Recognition for Intelligent Education System (Li et al) [4]	OpenCV (Viola-Jones) and Dlib	CNN, SVM	JAFFE, AffectNet
Emotion Recognition Model for Virtual Learning Environment (Yang et al) [7]	Viola-Jones	Neural Network	CK+, FER2013
Facial Emotion Recognition Based on Viola-Jones Algorithm in Learning Environment (Kirana et al) [16]	Modified Viola-Jones	Modified Viola-Jones	CK+, JAFFE

Title (Author(s))	Algorithm Used for Face Detection	Algorithm Used for Emotion Classification	Dataset(s) Used
Measuring Student Engagement Level Using Facial Information (Alkabbany et al) [9]	Viola-Jones	SVM	RAF-DB
Facial Expression Recognition for Probing Students' Emotional Engagement (Gupta et al) [22]	Multiscale Perception Network	CNN	FER2013, RAF-DB
Enhancing Facial Expression Recognition System in Online Learning Context Using Efficient Deep Learning Model (Aly et al) [23]	Deep Learning (Inception-V3, VGG19)	Emotion Recognition Models	AffectNet, FER2013
Facial Expression Recognition for Probing Students' Emotion Engagement in Science Learning (Tang et al) [24]	ResNet-50	Complex Emotion Analysis	EmoReact, CK+

2.1 Commonly Used Datasets

Here is a list of some commonly used datasets in the area of emotion recognition that fall in line with studies conducted in educational settings and facial expression analysis:

- FER2013 (Facial Expression Recognition 2013) [25]

- Description: This dataset contains 35,887 grayscale, 48x48 pixel face images categorized into seven emotion classes: anger, disgust, fear, happiness, sadness, surprise, and neutral.
- Usage in Studies:
  -Highly used for training and benchmarking deep learning models for facial expression recognition.
  -Applicable to education-related researches for which actual-time identification of various emotional conditions is indispensable.

-CK+ (Cohn-Kanade Extended Dataset) [26]

- Description: 593 video sequences recorded from 123 subjects have been added to CK+, which are labelled for seven basic emotions. Sequences start with a neutral expression and progress to a peak expression.
- Usage in Studies:
  -Often utilized in testing algorithms intended for classrooms because it has good, well-labelled data.
  -Supports training emotion recognition systems to identify transitions in emotions while teaching.

-AffectNet [27]

- Description: AffectNet is annotated for basic emotions (happy, sad, angry, etc.), valence, and arousal, with more than 1 million facial images collected from the internet.
- Usage in Studies:
  -Useful for developing systems that associate emotional expressions with levels of engagement in virtual and physical classrooms.
  -Offers a wide range of real-world data that can be generalized.

-RAF-DB (Real-world Affective Faces Database) [28]

- Description: This dataset comprises 29,672 real-world facial images annotated with seven basic emotion labels and compound emotions like "happily surprised" or "sadly fearful."
- Usage in Studies:
  -It can be used to detect complex emotional states in students, for example, when they feel mixed emotions regarding the lecture content.

-EmoReact Dataset [29]

- Description: The dataset primarily deals with video-based emotion recognition with focus on expressions relevant to the classroom environment, such as confusion, excitement, or boredom.
- Usage in Studies:
  -Suitable for systems targeting engagement monitoring in virtual learning environments.

-JAFFE (Japanese Female Facial Expressions) [30]

- Description: Consists of 213 images of 10 Japanese female subjects displaying six basic emotions plus a neutral expression. Images are rated on a 5-point scale regarding the intensity of the emotional state.
- Usage in Studies:
  -Often used as a pilot study to test emotion recognition systems before implementation in real-world applications for education.

3 Proposed Framework

Evaluation of teaching through emotions in a teaching-learning environment is gaining interest in recent years. We can capture emotions through facial expression recognition. Although lots of methods have been proposed for this purpose still there is a need for further research in this field and there is a requirement for a comprehensive framework of emotions. A comprehensive computational framework of emotions through facial expression recognition is proposed in the figure 1. The major components are given as follows:

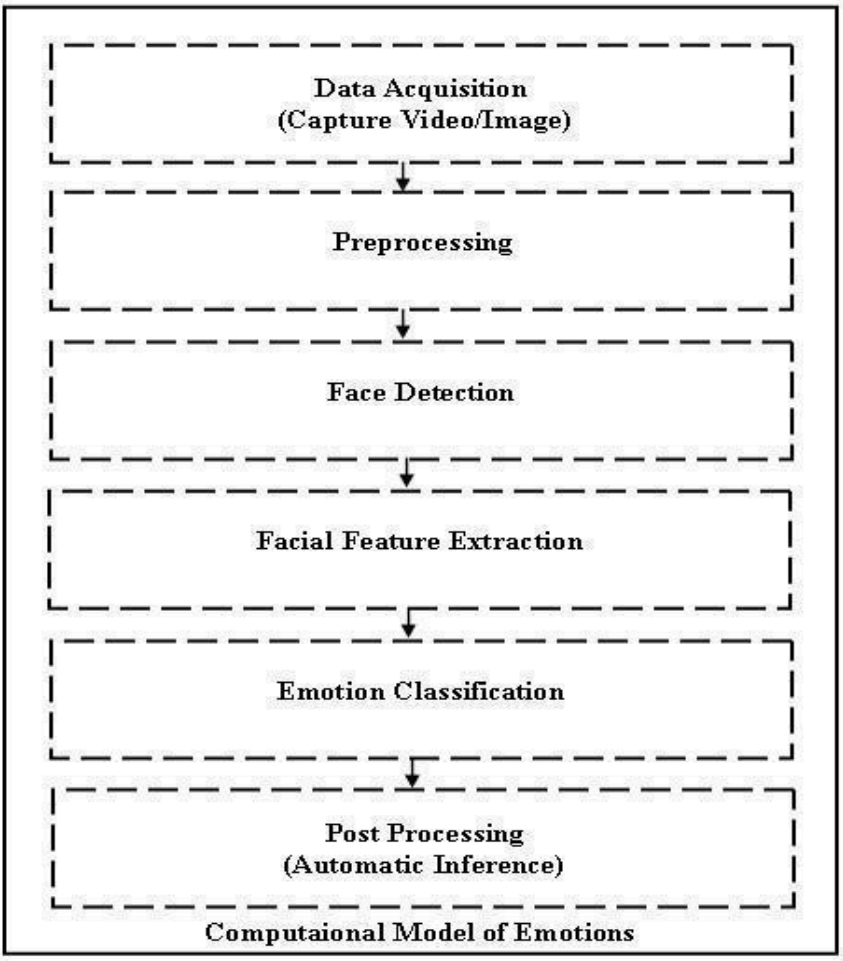

Fig. 1. Proposed Framework

3.1 Data Acquisition (Capture Video/Image):

It refers to the capture of the video flow or images from the environment during learning. This is normally regarded as a source of visual input through its primary input. In normal live or physical classroom settings, capturing clear visuals is normally via a high-resolution HD camera. In online settings, however, users make use of device-integrated webcams through these devices to transmit images. This is a key feature that serves as the basis for recording the detail of the learning environment to foster effortless engagement and monitoring.

3.2 Preprocessing:

It would be seen that this, in addition to proper standardization of facial images guided by heuristic principles, will improve the system significantly with the incorporation of appropriate preprocessing methods. It eliminates objects from the background during face detection of an image as objects resembling faces, in turn, will be enhanced through the quality of preprocessed data. In such a way, the technique lets the system augment its capability to precisely identify and examine the faces that appear within images, thereby enhancing overall effectiveness and reliability in several applications, including facial recognition and analysis.

3.3 Face Detection:

The goal of face detection is to detect and localize one or more faces in photos or videos. In the case of finding a face, this requires localizing its coordinates within the image, whereas the localization of a face is done through a definition of its boundary usually expressed as a bounding box. A detected face works as a preliminary stage for applications including tracking of the face, facial analysis, and recognition of the facial expression.

Different face detection methodologies have been proposed by researchers, but most of them are unique, and every methodology has unique pros and cons. Most common methodologies are Haar Cascade, which is one of the machine learning type methods, and Multi-task cascade Convolutional Neural Networks-based Deep Learning strategy, or MTCNN. The success of any application is quite dependent on the reliability of an operation such as face detection. Consequently, attaining excellent accuracy in face detection is considerably the factor determining the overall system performance success.

This particular component will perform the necessary function of face detection with excellent accuracy. It detects and locates faces in images or video feeds using sophisticated algorithms and methodologies as a basis for further analysis and applications characterized by increased efficacy and dependability.

3.4 Facial Feature Extraction:

From the identified faces, facial features are derived. The whole process of extracting the facial features from the entire facial image means to draw out the particular elements in it such as the eyes, nose, and mouth, among others. The features drawn out have been significant for input into the following stages of the classification activities.

Historically, many techniques have been used for facial feature extraction, each with its specific approach. These techniques include geometry-oriented approaches, template-driven methodologies, and appearance-centric strategies. In geometry-oriented methods, facial features are obtained based on the spatial relationships between various facial elements. On the other hand, template-driven methodologies rely on predefined templates for the identification and extraction of facial characteristics. On the other hand, appearance-centric strategies rely on the visual characteristics of facial components to extract relevant features.

The extraction of facial features is considered one of the most essential steps in facial recognition systems as well as their applications. Accurate isolation and extraction of face elements in this aspect enable a system to classify facial images with maximum accuracy and to perform identification, emotion analysis, and analysis of face attributes with a high degree of accuracy and dependability.

3.5 Emotion Classification:

In this part, emotional states are categorized into different classes, such as fear, anger, disgust, happiness, sadness, surprise, and others, which are derived from the facial features obtained in the previous phase. Different methods, such as SVM, KNN, and CNN, can be used for this purpose.

With advanced algorithms and machine learning frameworks, the system analyses the minute differences in facial expressions to identify and classify the emotions correctly as presented by individuals.

Among the methodologies commonly used for emotion classification are SVM, KNN, and CNN because of their ability to deal with complex datasets and identify subtle patterns within facial features. These techniques enable the system to learn from the extracted facial features and make informed decisions regarding the emotional states of individuals depicted in the images or video streams.

3.6 Post Processing (Automatic Inference):

This section analyzes findings from the previous phase, focusing on general positive and negative emotions. These are the emotions attached to particular classroom events, such as laughter, thought, or boredom, because of which emotional dynamics in the learning environment are understood. Based on this analysis, automatic suggestions or alerts are generated to adjust teaching

strategies to enhance participation or maintain positive conditions. The final interpretation stage converts emotional data into actionable insights, allowing for real-time adjustments to promote positive learning experiences and improve educational outcomes.

4 Conclusion and Discussion

In the current study, we present a framework that aims at providing a basis for designing a computational model that has been focused on the recognition of learner emotions in educational contexts. The framework consists of several components. Briefly, each is stated. The components could be applied in a flexible way according to personal choice and the specific needs of the environment in which the system would be used.

Throughout the implementation phase, users must consider a number of factors. For example, the effectiveness of the face detection algorithm is sensitive to lighting conditions and has problems when many faces are in close proximity. Additionally, the availability of real-time classroom datasets for training the system is still limited.

There still exists a number of shortcomings of this area, despite making much progress in the domain of emotion recognition. This deficiency contains sparsity of datasets related to spontaneous facial expressions, the requirement for precise techniques in facial detection and facial feature extraction, efficient approaches related to classification into different emotions, robust models concerning developing models for creating computational models of emotions, and development of holistic models which should perform efficiently in multiple learning environments.

Rectifying these deficiencies is essential for the progress of the emotion recognition domain and for improving the effectiveness of computational frameworks aimed at interpreting and reacting to the emotional states of learners within educational environments. Recognizing these obstacles and concentrating on additional research and development efforts will drive the discipline forward and reveal new opportunities for individualized and adaptive learning experiences.

References

1. Weiqing Wang, Kunliang Xu, HongliNiu, Xiangrong Miao, “Emotion Recognition of Students Based on Facial Expressions in Online Education Based on the Perspective of Computer Simulation, Hindawi Complexity, vol.2020, 2020.
2. Carmen Lacave, J. Ángel Velázquez-Iturbide, Maximiliano Paredes-Velasco, Ana I. Molina, “Analyzing the influence of a visualization system on students’ emotions: An empirical case study”, Computers & Education, vol. 149, 2020.
3. K. S. Sahla, T. Senthil Kumar, “Classroom Teaching Assessment Based on Student Emotions”, Intelligent System Technologies and Applications 2016, Advances in Intelligent Systems and Computing vol. 530,2016, pp 475-486

4. G. Li and Y. Wang, "Research on leamer's emotion recognition for intelligent education system," 2018 IEEE 3rd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), 2016, pp. 754-758.
5. Tang C., Xu P., Luo Z., Zhao G., Zou T. "Automatic Facial Expression Analysis of Students in Teaching Environments". In: Yang J., Yang J., Sun Z., Shan S., Zheng W., Feng J. (eds) Biometric Recognition. CCBR 2015. Lecture Notes in Computer Science, vol 9428. Springer, Cham.
6. Krithika L.B, LakshmiPriya G G, "Student Emotion Recognition System (SERS) for e-learning Improvement Based on Learner Concentration Metric", Procedia Computer Science, vol. 85,2016, pp 767-776.
7. D. Yang, AbeerAlsadoon, P.W.C. Prasad, A.K. Singh, A. Elchouemi, "An Emotion Recognition Model Based on Facial Recognition in Virtual Learning Environment", Procedia Computer Science, vol. 125, 2018, pp 2-10.
8. I. Lasri, A. R. Solh and M. E. Belkacemi, "Facial Emotion Recognition of Students using Convolutional Neural Network," 2019 Third International Conference on Intelligent Computing in Data Sciences (ICDS), 2019, pp. 1-6.
9. I. Alkabbany, A. Ali, A. Farag, I. Bennett, M. Ghanoum and A. Farag, "Measuring Student Engagement Level Using Facial Information," 2019 IEEE International Conference on Image Processing (ICIP), 2019, pp. 3337-3341.
10. W B Putra, F Arifin, "Real-Time Emotion Recognition System to Monitor Student's Mood in a Classroom", Journal of Physics Conference Series, vol. 1413, 2019.
11. M.S. Bartlett, G. Littlewort, M.G. Frank, C. Lainscsek, I. Fasel, and J.R. Movellan. "Automatic recognition of facial actions in spontaneous expressions." Journal of Multimedia, 2006.
12. J. Whitehill, M. Bartlett, J. Movellan, "Automatic facial expression recognition for intelligent tutoring systems," 2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, 2008, pp. 1-6.
13. J. Whitehill, Z. Serpell, Y. Lin, A. Foster, J. R. Movellan, "The Faces of Engagement: Automatic Recognition of Student Engagement from Facial Expressions", in IEEE Transactions on Affective Computing, vol. 5, no. 1,1 Jan.-March 2014, pp. 86-98.
14. Ayvaz, U., Gürüler, H., &Devrim, M. O. (2017). Use of Facial Emotion Recognition in E-Learning Systems. Information Technologies and Learning Tools, *60*(4), 95–104.
15. D. Yang, AbeerAlsadoon, P.W.C. Prasad, A.K. Singh, A. Elchouemi, "An Emotion Recognition Model Based on Facial Recognition in Virtual Learning Environment", Procedia Computer Science, vol. 125, 2018, pp 2-10.
16. K. CandraKirana, S. Wibawanto and H. WahyuHerwanto, "Facial Emotion Recognition Based on Viola-Jones Algorithm in the Learning Environment," 2018 International Seminar on Application for Technology of Information and Communication, 2018, pp. 406-410.
17. GurayTonguc, BetulOzaydınOzkara, "Automatic recognition of student emotions from facial expressions during a lecture", Computers & Education, Volume 148, 2020.
18. M. Bouhlal, K. Aarika, R. AitAbdelouahid, S. Elfilali, E. Benlahmar, "Emotions recognition as innovative tool for improving students' performance and learning approaches", Procedia Computer Science, vol. 175, 2020, pp 597-602.
19. Enrique Osuna, Luis-Felipe Rodríguez, J. Octavio Gutierrez-Garcia, Luis A. Castro, "Development of computational models of emotions: A software engineering perspective", Cognitive Systems Research, vol. 60, 2020, pp 1-19.

20. Enrique Osuna, Luis-Felipe Rodriguez, J. Octavio Gutierrez-Garcia, “Toward integrating cognitive components with computational models of emotion using software design patterns”, Vol 65, 2021, pp 138-150.
21. Kuntzler T, Hofling TTA and Alpers GW, “Automatic Facial Expression Recognition in Standardized and Non-Standardized Emotional Expressions.” Frontiers in Psychology,2021.
22. Gupta, S., Kumar, P. & Tekchandani, R.K. Facial emotion recognition based real-time learner engagement detection system in online learning context using deep learning models. Multimed Tools Appl 82, 11365–11394, 2023.
23. M. Aly, A. Ghallab and I. S. Fathi, "Enhancing Facial Expression Recognition System in Online Learning Context Using Efficient Deep Learning Model," in IEEE Access, vol. 11, pp. 121419-121433, 2023.
24. Tang, Xiaoyu, Yayun Gong, Yang Xiao, Jianwen Xiong, and Lei Bao. "Facial Expression Recognition for Probing Students’ Emotional Engagement in Science Learning." Journal of Science Education and Technology (2024): 1-18.
25. Goodfellow, Ian J., Dumitru Erhan, Pierre Luc Carrier, Aaron Courville, Mehdi Mirza, Ben Hamner, Will Cukierski et al. "Challenges in representation learning: A report on three machine learning contests." In Neural information processing: 20th international conference, ICONIP 2013, daegu, korea, november 3-7, 2013. Proceedings, Part III 20, pp. 117-124. Springer berlin heidelberg, 2013.
26. P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar and I. Matthews, "The Extended Cohn-Kanade Dataset (CK+): A complete dataset for action unit and emotion-specified expression," *2010* IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops, San Francisco, CA, USA, 2010, pp. 94-101
27. Mollahosseini, Ali, David Chan, and Mohammad H. Mahoor. "AffectNet: A database for facial expression, valence, and arousal computing in the wild." IEEE Transactions on Affective Computing 10.1 (2017): 18-31
28. S. Li, W. Deng and J. Du, "Reliable Crowdsourcing and Deep Locality-Preserving Learning for Expression Recognition in the Wild," *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, 2017, pp. 2584-2593
29. B. Nojavanasghari, T. Baltrusaitis, C. E. Hughes, and L.-P. Morency, “EmoReact: a multimodal approach and dataset for recognizing emotional responses in children.,” in *ICMI*, 2016, pp. 137–144
30. M. Lyons, S. Akamatsu, M. Kamachi, and J. Gyoba, "Coding facial expressions with Gabor wavelets," Proc. 3rd IEEE Int. Conf. Automatic Face and Gesture Recognition, Nara, Japan, 1998, pp. 200-205.