




# Real time emotions recognition through facial expressions

Alisha Fida<sup>1</sup> · Muhammad Umer<sup>1</sup> · Oumaima Saidani<sup>2</sup> · Monia Hamdi<sup>3</sup> ·  
Khaled Alnowaiser<sup>4</sup> · Carmen Bisogni<sup>5</sup> · Andrea F. Abate<sup>5</sup> · Imran Ashraf<sup>6</sup> 

Received: 25 April 2023 / Revised: 1 August 2023 / Accepted: 31 August 2023

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

## Abstract

Human behavior is deeply influenced by emotions. Detection of emotions plays a pivotal role in understanding how individuals respond to various stimuli, such as reading text, encompassing feelings of anger, anxiety, confusion, or nervousness. Real-time facial emotion detection during online text reading represents an innovative approach for receiving immediate feedback based on readers' emotional responses. Real-time emotion detection finds applications in interactive displays and holds immense potential for online learning platforms, where it can be utilized to analyze students' emotional states and gauge their level of comprehension. Despite vast existing literature on emotion detection, real-time emotion detection is not very well studied. This study demonstrates the design and implementation of face emotion detection for students while they are using online learning platforms. The primary objective is capturing human emotions and storing them in the database after five seconds while they are reading online text. The system is implemented using SSD based on VB.NetV1. The proposed system has strong relevance for integration with online web applications to detect learners' real-time emotions. Experiments are performed using CK+ and JAFFE face datasets and results show 96.46% and 98.43% accuracy, respectively. The system not only provides accurate results but also enables high-quality, robust, and real-time feedback based on the facial expressions of readers, facilitating a deeper understanding of students' emotional engagement during their online learning experiences.

**Keywords** Human behavior · Facial expressions · Emotions detection · Online learning · Design science research

## 1 Introduction

Human behavior recognition and action understanding is a complicated task as it involves diverse actions, challenges of capturing appropriate information from videos, and external

---

✉ Imran Ashraf  
imranashraf@ynu.ac.kr

Extended author information available on the last page of the article

factors like light conditions, angle, human appearance, etc. Facial expressions hold a central part in human behavior recognition. Information is perceived through verbal and non-verbal channels, but facial expression is perceived through the non-verbal medium. Facial expression recognition is the technique to detect such expressions and classify them into anger, fear, surprise, sadness, happiness, etc. The final results of the behavior of a person are based on these expressions [1]. The ability to perceive the human's emotional state is critical in recognizing human competence [2]. Today, the learning community focuses on student work and the vision of faculty who are working collaboratively to obtain high learning quality. Hence, digital communication, which leads to learning communication, gives learning a new dimension. "Deep learning (also known as deep structured learning) is part of a broader family of machine learning methods based on artificial neural networks with representation learning" [3].

Education can be supervised, semi-supervised, or unsupervised. As per Hasani's research [4], emotions offer a significant layer of information that can be used for emotional detection in the event of synthetic features' failure. Teachers use facial expressions and technology to assess students' level of understanding, with learning facial expressions playing a crucial role in this process. To detect facial expressions, various deep learning algorithms have been employed, including deep learning networks (DLN) and convolutional neural networks (CNN) [5]. Other studies have also indicated that CNN is one of the most effective methods for emotion detection and is widely utilized [4]. In addition, [6] has focused on determining the best model for emotion detection, while [7] has elaborated on the neural network's role in emotion detection.

CNN is a multi-layer perceptron (MLP) that is specialized in capturing the relationship between pixels from a particular perspective. Its architecture comprises a different number of layers that are tailored for image processing tasks. These layers are responsible for performing segmentation, feature extraction, and classification, requiring minimal preprocessing functions on the input image [8]. The initial CNN architecture for emotion detection was designed for frontal images with facial expressions and was relatively straightforward. Various models have been developed for deep learning architectures to detect emotions, and each model is trained using defined datasets with early specified data. These deep learning architectures, including CNN, have different models, and they may be used in combination with other models for emotion detection. However, all models are trained using defined datasets with early specified data [7].

The 'OS' module in Python provides functions that return the number of images for each emotion type in a dataset. To gain a deeper understanding of the dataset and the types of images it contains, the 'utils' module in Python can be used to plot a few example images. Previous research has indicated that the average achievable accuracy of a training model developed using the FER-2013 dataset is 66.7%. Our goal is to design a CNN model with similar or better accuracy, as stated in various research papers [7].

For this study, we utilized the JavaScript Face API, an open library that is accessible on GitHub. This library includes various models, such as face landmark detection, gender recognition, face detection, age estimation, face recognition, and emotion recognition. In this project, we employed three of these models: face detection, emotion recognition, and face expression recognition model. Specifically, for facial expression detection, we implemented Face-api.js using SSD MobileNet v1, which is based on CNN, as described in Sawyer's research [9]. We make the following contributions

- A real-time emotion recognition framework is designed to capture and demonstrate students' emotions while they are using online learning platforms. This real-time aspect is

particularly valuable as it allows for immediate feedback and analysis of emotions during the learning process, providing new insights into students' emotional responses to online education.

- The proposed model is implemented using SSD MobileNet based on WB NetV1. This choice of model architecture and its application to emotion recognition is a novel approach that may offer advantages in terms of efficiency and accuracy compared to traditional models.
- An open CK+ and JAFFE datasets of facial images are utilized to evaluate the performance of the proposed model.
- The superiority of the proposed model is shown by comparing its performance with previous state-of-the-art approaches. This comparative analysis validates the effectiveness of the model and establishes it as a cutting-edge solution in the field of emotion recognition for online learning.

This research work is structured as follows: Section 2 presents previous work and research conducted in related domains. Section 3 describes the data description and preprocessing steps utilized in the experimental and implementation phases. Section 4 outlines the proposed methodology. Section 5 presents the experimental results and a scientific evaluation of the study. Finally, Section 6 concludes the proposed study.

## 2 Related work

Today, the achievement of digital communication leads students into a new dimension. The scope of virtual schools is increasing worldwide, and computer education is considered necessary for the future [7]. In [10, 11], the authors perform experiments for facial expressions to recognize expressions. They conduct research and do various experiments for virtual learning that use various tools. In this study, we review existing works on emotion recognition [10, 11].

Ekman believes that most facial expressions are blended into several feelings [12]. These are the most challenging things by far to recognize non-verbal expressions and understand them clearly. However, we can use emotional expression as a signal such as wrinkling an eye is considered a gesture of approval. Sticking the tongue out is deemed to be a signal the playful distaste. In [13], Diaz stated that emotional expression is the least controversial in the communication of all non-verbal channels. Hence, emotional expression is the most experimental group of gestures. Therefore, one needs to be more focused on the face than the other parts of the body, as its meaning is widely accepted. The expressions on the face can say a lot, like expressing shock and surprise [13].

Emotional expressions result from muscles and motions. Hence, movement conveys lots of emotions for the individuals, so emotional expressions are considered non-verbal communication with the primary meaning of communicating social information to a human. Diaz states that while considering how many muscles in the human face move, he is surprised to see a wide range of emotional expressions produced [13]. While teaching students, teachers make effective use of these emotions. [14, 15] describes the attention-based modeling for emotion detection by detecting emotion contrasts to understand students' interests. So teachers can easily recognize whether students are satisfied with their teaching method or need more attention [14, 15]. In this way, the proper use of emotion detection makes the teacher understand the students while teaching in the classroom, which eventually leads to better learning outcomes [16].

Virtual learning uses audio, videos, Word, slides, and PDF to simulate the learning and learning environment as closely as possible. For the virtual environment, the plethora used is a pedagogical purpose like distance learning. Worldwide, the benefits of virtual schools are increasing day by day mediated by a computer considered more critical for the future. Some say that virtual education shows instruction in the learning environment where a teacher can interact with a student separated by time, distance, and space. The teacher provides a course management application through multimedia, the internet, conferencing, and student communication with the teacher using technologies [17]. Learning from virtual books, students learn through devices using laptops, mobile phones, iPods, etc. The difference between physical book learning and online learning through soft copies provides apparent flexibility and adaptability. Students can get more information through the internet at the same time by sitting in a single place without going anywhere to search for related books or something else [18].

Learning in a real classroom with face-to-face communication, students can see each other and recognize the expressions. Accordingly, many virtual classes have implemented regular online courses and chat with classmates through video conferring interaction. Virtual business is considered analog, so reading is referred to as virtual communication used for learning and teaching in a virtual environment. Students consciously and unconsciously receive and send facial clues several times per day during lectures, classrooms, and online reading. Thus, reading through online learning in classrooms has two primary reasons: to receive student messages better and get the ability to send a positive signal that supports learning more skillfully. So, it is essential to use non-verbal communication and detect nonverbal expressions for better online learning [9].

Interactions with one another, whether student to student or student to teacher, play a vital role in the class environment, so the impact of communication through facial expression is powerful [9]. Those facial signs are rich in expressions representing lots of information about individual identity, mental state, and mood. Several studies reveal that facial expressions are the most prominent and expressive way to display emotions [19]. Facial expressions are the primary source for detecting the information next to words and determining internal human feelings. All people and students use facial expressions during lectures and get one another's impressions [20]. A study reveals that students can stay motivated and interested while taking lessons or learning online through facial expressions [21]. Thus, a lecturer giving a lecture can also use students' facial expressions as a most valuable source of useful feedback, whether students understand the course. According to Sathik, a lecturer can detect the students' facial expressions and recognize whether they have to slow down, speed up, or implement another way to improve or change the presentation [22]. To optimize the student's behavior in class while learning, the basic strategy is that the teachers can feel the student's mind, so they must be a reasonable observer of the student's expression, movement, and actions. This strategy helps to understand the student's strengths and weaknesses to adapt any changes to suit their learning [23].

A facial expression that shows emotions includes muscle movements like wrinkling, eyebrow-raising, curling lips, and rolling eyes. According to medical studies, when students feel uncomfortable, they lower their eyebrows, shrink their brows, wrinkles on the forehead in vertical and horizontal directions, and take time to maintain eye contact [19]. To detect the student's expression correctly, humans have to be familiar with lots of subtle nonverbal clues that most students exhibit. Studies evaluate that the emotional state of a student is expressed by the specific human behavior which can be detected automatically, so grabbing it through the forehead, mouth, eyes, and nose plays an essential role in recognizing the facial expression through an automated system. Automated systems can extract the expression through

the underlying emotional state and categorize the spontaneous facial expression [19]. Since the discussed research works report different approaches and results, a critical summary of these works is presented in Table 1.

In recent years, computer vision has seen remarkable advancements, and numerous research works have leveraged its potential to analyze and understand human emotions. Several state-of-the-art research studies involve the application of computer vision using libraries like OpenCV, Dlib, TensorFlow, or PyTorch to detect and recognize facial expressions, gestures, and emotional cues. While most state-of-the-art audio-visual fusion methods use recurrent networks or conventional attention mechanisms, they often fail to effectively leverage the complementary nature of these modalities. To address this, the paper [24] proposes a joint cross-attention fusion model that efficiently exploits the inter-modal relationships between facial and vocal modalities extracted from videos. Another paper [25] focuses on deep learning techniques, particularly convolutional neural networks, for facial expression recognition tasks. It proposes an occluded expression recognition model that comprises two modules: occluded face image restoration and face recognition.

Table 2 presents the summary of current research works on education and emotions. By presenting a variety of research topics, methodologies, and findings, the summary table demonstrates the multidimensional importance of research works in education, as they contribute to a deeper understanding of various factors influencing teaching, learning, and academic outcomes. These studies collectively enrich the knowledge base and provide valuable insights for enhancing educational practices and student experiences.

The scope of facial expressions among students needs to be further investigated. Several researchers apply different methods for the correct identification of emotional detection. For example, the CNN deep learning model has been used widely for this task. However, it is found that the accuracy of such approaches is low and they lack real-time emotion detection. Real-time emotion detection is important for online learning platforms. Similarly, detecting a reader's emotion while reading books online is not well investigated. Therefore, this study develops a JavaScript-employed web-based emotion detection application. JavaScript language was used because it is the most important language in building web development tasks, enterprise-class applications, and Android development. Recently the information technology industry has been growing rapidly, and the demand for JavaScript is growing, so for experiments, we considered JavaScript language more suitable to use. This research holds substantial importance in advancing emotion detection technology, enhancing online learning platforms, and providing valuable insights into learners' emotional experiences. Its high accuracy, real-time feedback capabilities, and potential for educational interventions make it a valuable contribution to the field of education and technology.

### 3 Design science and research methodology

The systematic study of developing practical solutions is known as design science research (DSR) that emerges from real-life settings and solutions [38]. The DSR uses a rigorous process to design artifacts to solve practical problems evaluate designs and make the research more contributive and communicative. The framework of DSR is presented in Fig. 1.

**Table 1** Summary of research works on emotion detection

Ref.	Short description	Results
[10]	Description of antisocial behaviour in text automatically.	They show that emotional information with a load of words is 90% accurate with ASB classification in text with preliminary experiments showing positive implications in the fast detection of potentially harmful behavior.
[26]	Through gender classification binary classification problems, and its discriminating feature has taken at vector space, including emotion base feature.	Their experimental results show that by exploiting the information, the text has 80% cross-validation accuracy and supports the vector machine algorithm. Their finding shows positive implications in the feature of emotions.
[27]	The research has influenced various writing styles with an extensive array of factors through author gender and text genre.	With the analysis of its linguistic features like a stylistic metric variable, work categories, and n-grams, the authors present an exploratory study of emotion.
[28]	Learning diaries are very informative for feedback generation and become a task to analyze the study manually.	The proposed system automatically detects the emotions and changes the emotions while the student's learning experience is expressed with journals.
[29]	For the automatic detection of feeling, sentiment, and emotion while learning text for proper differentiation among its subjective terms and understanding how these emotions related to one another.	These determinations clarify these differences with five subjective terms and reveal the significant concepts of the computational linguistic community.
[31]	Machine learning technology is used for the detection of sentiments while learning any text on stories.	By identifying peers, school staff, and domestic as the primary influence on student academics, students are used as a class label for supervised classes.
[30]	The researcher proposed an adaptive CNN that determines the CNN structure without comparison.	Their experimental results show a better tradeoff between the recognition rate of ACNN and training time consumption.

**Table 2** Summary table of the research works on education and emotion

Ref.	Focus	Importance
[31]	Investigates linkages between academic resilience, emotional intelligence, and academic performance in Nigeria	Understanding the importance of resilience and emotional intelligence for academic success
[32]	Examines the Link between Emotional Intelligence and Educators' Performance in Higher Education	Understanding the Impact of Emotional Intelligence on Teaching Effectiveness
[33]	Proposes a Deep Learning Method to Detect Student Emotions in Online Education	Understanding the Potential Understanding the Potential of Deep Learning for Emotional Impact Detection in Education
[34]	Examines the reciprocal associations between school grades and students' achievement emotions	Investigating the within-person relations of grades and emotions in the context of longitudinal data
[35]	Explores the relationship between personality, academic confidence, and academic emotions	Understanding the impact of personality on academic confidence and emotions
[36]	Examines the relevance of classroom management and instructional clarity to academic emotions and achievement in different educational contexts	Understanding the role of classroom management and instructional clarity in shaping students' academic emotions and achievement in different cultural settings
[37]	Explores the relationship between online and offline learning motivation, academic emotions, and learning performance	Understanding the mediating effect of positive academic emotions in the context of online and offline learning and its implications for educational management in online courses.

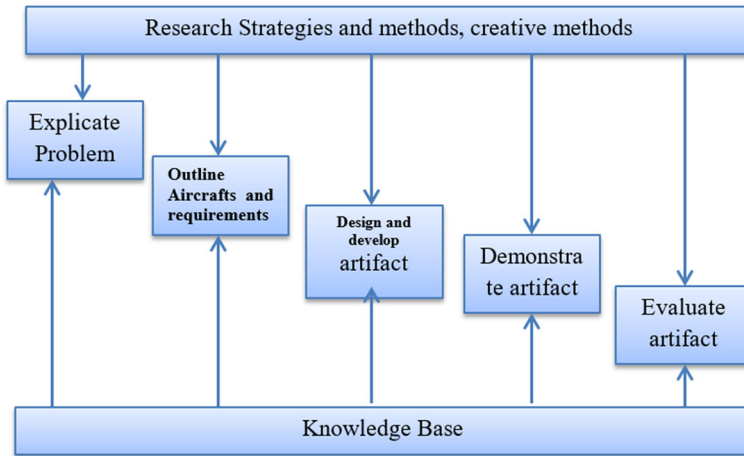


Fig. 1 Design science research framework

### 3.1 Problems with emotion detection system

Various technical solutions are observed to detect the human emotion used by firms to improve responsiveness, decide regarding the candidates during interviews, and optimize the emotional impact on the system by human and customer service improvement. Studies have shown that many solutions in this field rely on outdated psychological theories and cannot always be trusted [39]. While reading PDF books, teachers and students face many challenges in the learning activities due to ineffective achievement goals involving cognitive skills, low morale, motivation towards reading, and self-esteem.

### 3.2 Design and develop an artifact

By analyzing the evaluated designs and models, we know that the main functional requirement that can support the system to detect human emotion continuously while reading online text is compulsory for the teachers to understand the students' emotions. According to [40], communication advancement in the network and social media users use massive feelings for which the extraction of emotions is necessary and extracted frequently. Therefore, an analysis of the requirements for emotion detection can be very important and have diverse applications [40]. One requirement is that a device must continuously detect emotion because emotions are a continuous process. [41] emphasize that a system is required that can accurately capture emotions while people are reading and the human association with them. An adaptive system towards the lexical variations provides a fine-grained quantitative assessment to detect accurate human emotions [1]. Another requirement is detecting human feelings when they raise any question or have any problem while reading text from a PDF. When students learn, they generate some questions with irregular intervals of time. A system must detect the feelings and clues whenever students have any questions, so sentiment analysis is needed.



### 3.3 Demonstrate artifact

For this purpose, machine learning is needed to detect human emotions like aggressiveness, anger, sadness, happiness, surprise, or any questions. Another requirement found in the literature is using the CNN model with the use of images only, so deep learning is required to detect human emotions continuously and change feelings/emotions when they are reading any text [42]. In previous studies, we observed the vital role of machine learning in the interaction of emotions, representing a text with a sparse bag of words. A method is required that combines sparse and dense behavior of human representation [43].

### 3.4 Evaluate artifact

So, in this paper, a model for emotion detection is proposed with the use of JavaScript. We used the JavaScript library because it makes the functions and applications that can be internet indispensable coded in JavaScript form. According to the survey, it is found that 94.5% of software is using JavaScript, either for reading material or other. We also found from the literature that for emotion detection while reading text from PDF, JavaScript is not very well studied. Therefore, it is desirable to investigate the use of JavaScript for real-time emotion detection [44].

### 3.5 System architecture

The study conducted emotion recognition tests on students while they read a short article using the JavaScript face API. The videos are captured using Nikon D5300, a digital single-lens reflex (DSLR) camera with an 18/140 lens. The dimension of recording video is  $11903 \times 13096$  pixels (width  $\times$  height). An emotion detection algorithm captured students' emotions after each paragraph with a 5-second delay. The face recognition application recorded emotions once users allowed camera access, securely storing the data. The system for real-time emotion recognition offered valuable insights into the emotional responses of students while engaging in online learning. To evaluate the proposed model, the researchers also employ open CK+ and JAFFE datasets of facial images. Using these datasets for emotion recognition in the context of online learning is a novel application, showcasing the model's performance in realistic scenarios and ensuring transparency in the evaluation process.

The system architecture diagram for the emotion detection system is shown in Fig. 2. In this research, the target people are all types of learners, so the detection of emotions is done quickly while reading the online text. To cover every expression after every five seconds, we use a small face detector and a real-time face detector is much faster, smaller, and consumes fewer resources. Using Face API-JS, emotion detection is done in this research by following a model to recognize the face and its expressions. SSD, which is based on Web NetV1 is implemented to detect emotion after every 5 seconds. At each location, each face image will return its boundary boxes together with each face probability. This research aims to obtain very high-accuracy data for students in education systems who are reading online texts. We use this model to detect every facial expression every five seconds to get higher accuracy [16]. The model is small-sized, lightweight, and web-friendly so it should be the GO-TO face detector on limited resources and the web for clients.

This model takes the space of 190kb only so the face detector can be trained with the same custom dataset of 14K images labeled with the boundary boxes. With oriented boundary boxes, the face's full features can be detected so generally it produces better results with

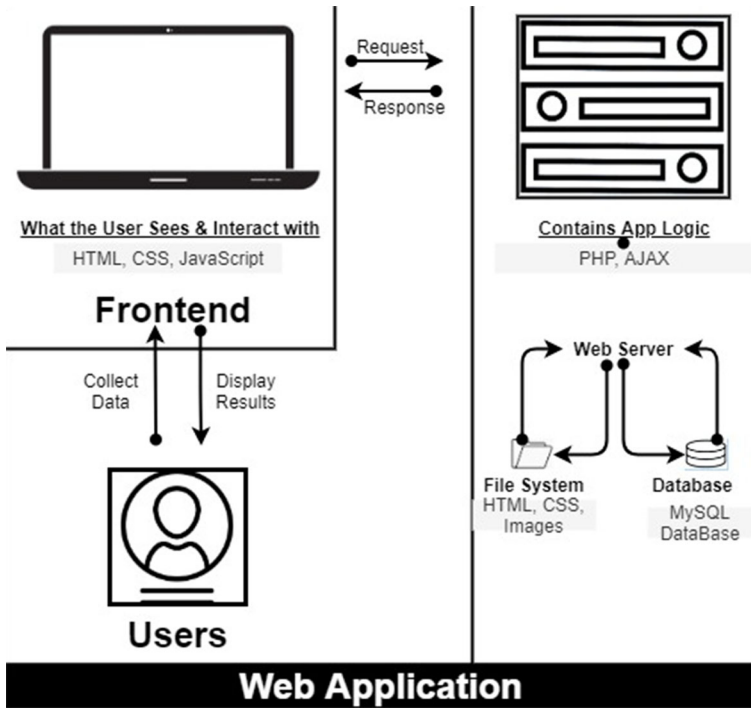


Fig. 2 System Architecture diagram for the emotion detection system

the combination of subsequent landmark detection. We know that lots of people have small beginnings and round or other shapes of the face so for the detection of such expressions the use of tiny face detection model is used so that each type of look can be detected. Their expression can be seen more precisely [16].

For the recognition of the face, an architecture known as ResNet-34 is implemented to compute face descriptor with the given image of a face. With this model's use, one benefit is evaluated, which is the detection of two arbitrary faces through a comparison of face descriptors. For example, the Euclidean distance is computed with the use of a personal choice of the classifier. For face recognition.js, the use of face recognizer Net is done which is equivalent to the neural net and is used in Idlib face recognition.

For the next step, which is the detection of change in emotions, to implement these characteristics, the employment of a feature extraction layer as a regression layer and recognition layer is done. This model classifies the expressions based on their implementation. The model gains the size of 420kb only, so its feature and emotion detection category employ tinier architecture changes. This model's performance is done with the use of some databases like UTK, Chalearn, Wiki, IMDB, CACD, MegaAge-Asian, and FGNET. From all above the main focus and functions of this architecture with input and output have the following points.

- Recognize human emotions while reading online text
- Highlight the face type
- Give accurate results

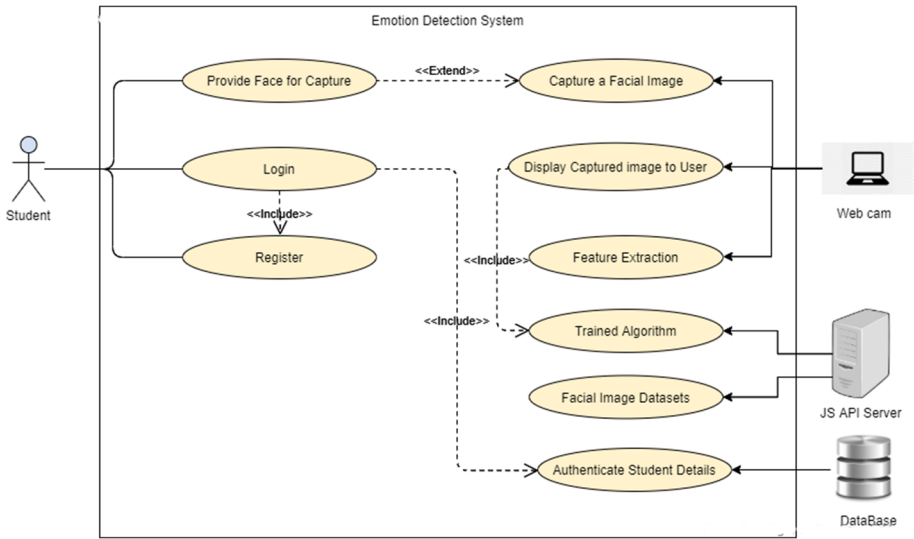


Fig. 3 Use case diagram for Emotion Detection System

Tracking human emotion while reading online text, is an essential target of this research. The second target is to recognize the changing expressions of the person reading online text so that recognition of emotion would be more comfortable. The third function is the detection of face type; for example, if a human has a small face, then recognition with Euclidean distance would be done. The last function is getting highly accurate results. The whole process is repeated every five seconds to achieve accuracy and detect every expression of a human [45].

### 4 Model implementation

The evaluated framework works on the online text file, read by the specific human in a web environment. In a web environment, the algorithm for the face emotion detection of the interest points for localization of feature extraction is performed. The whole architecture is implemented on a web platform [18]. The use case diagram for the emotion detection system is shown in Fig. 3

#### 4.1 JavaScript face-API

With the use of Face-ap.js, the detection of emotions and recognition of face landmarks is implemented using CNN. Several modules are available on the face-api.js like face detection, recognition of facial emotions, etc.

#### 4.2 Algorithm

The proposed approach aims at emotion detection while the reader interacts with online reading. We want to calculate the time spent on a single paragraph or a single paragraph with the online text page’s start and end time and capture the image to detect emotions.

Face API is used to capture the face's appearance so that recognition of emotion would be more accurate.

When the user opens the particular text from the platform and starts reading, his appearance is captured and downloaded. Face API supports the image for emotion recognition.

That first image is saved onto the database with the recorded start time. When the user starts reading text, its time is recorded, and the image is captured. After every five seconds, the image would be captured. The captured image goes into the database. That image is used for the detection of emotion. After calculating the end time of the page when the user turns to the next page, time is also recorded. The total time spent on a single page is recorded like the start and end time of reading and reading. The total time spent on that page would be calculated as:

$$\text{Total time spend on a single page} = \text{number of paragraphs on page} / \text{time spent}$$

After applying this equation, we evaluate that the calculated time would be 4 minutes which means that the total time spent on a single paragraph is 4 minutes. We conclude that we can easily find how much time a user spends while reading single sections on a page. All the images that are detected every five seconds are captured for emotion detection. Now we have a dataset that consists of face images and spends time reading online text, so annotation is applied to the face images by focusing on various face regions to detect emotions. Focusing on different face coordination points like the mouth, nose, eyes, lips, and other features, the feelings are detected. The aim of facial coordination points is the detection of the reader's emotions. After comparing reference images, emotion is detected through the facial recognition model with a descriptive image from the facial appearance. The height and width of the image are detected and added with the detection of images. After that, unique features are obtained through which the exact emotion is recognized.

### 4.3 API to integration

We can integrate this system with any other online platform for emotion detection. But for now and for testing purposes, we integrated this platform with an online learning platform. Learning platform provides opportunities for creating courses of various types taking into consideration different corresponding standards and interactions among the stimulating and students' creativity regarding tutor and students. When the student reads any particular text from the platform, the working of the system is started.

In the proposed architecture, the detection of emotion needs a preprocessing algorithm in JavaScript code using the web page. The general steps for the interface are

- During the run time, the detection and accessing of the camera hardware if present.
- Creating a preview of the human who is reading the online text through the surface which displays the live image
- The preview of the layout is built, which considers the preview as user interface control according to user choice.
- A setup is created to capture the listeners so that user action response is recorded during reading.
- Capturing and saving the detected emotion when a preview of the image through the layout is built, so it becomes ready to capture the next expression or emotion.

## 5 Evaluation of the emotion detection application

To evaluate the emotion detection application, students were recruited to test the application. The primary purpose of the evaluation is to access how users feel and to derive data from the real-life situations of online learners by showing their emotions while reading standard text [46]. We evaluated whether students found this platform interesting because they could see their expressions through live graphs while reading. In addition, we intend to conduct a further evaluation to obtain a deeper analysis of students' behavior while studying online and how teachers found the application useful in terms of emotion detection. It is hoped that the application will provide real-time analysis for teachers to support their students while studying online.

### 5.1 Learning activity and experiment design

In this research, different phases describe the whole activity. In the first phase, we provided an overview of the emotion detection system, and then students logged into the system and read a short article. While the student was reading, the system recorded the student's facial expressions and stored them in the database. In the second phase, we plotted some charts, graphs, and a table on the stored databases.

### 5.2 Data collection

We collected data from students through the emotion detection application while students were testing the application online. The application recorded their emotions and stored them in the database. These data were analyzed in the data analysis section.

### 5.3 Data analysis

In this section, we discuss the findings of the analysis of data collected from students who read an online text. To detect emotions, we utilized the JavaScript Face-API and publicly available datasets, including CK+ Facial Expression, JAFFE, and KDEF. These datasets contain more than 2,000 images featuring various facial expressions.

For this process, the user allows the camera while reading then the application starts streaming video of the user's face. After that system detects the face and indicates the facial landmarks of the user's face. Facial landmarks are the points used to indicate the position of muscles. Those students, who logged in to the system, read a short article, and the system recorded their facial expressions (see Table 3).

### 5.4 Results

The system has been tested with different students/learners while reading a short article through the system. We got different emotions from every student. When they read any paragraph, their emotions change. In this study, we have detected the emotions of students after each paragraph by using JavaScript face-API. After 5 seconds, an emotion detection algorithm starts detecting students' emotions, i.e., what they are feeling after reading any particular paragraph.

**Table 3** Student's expressions while reading a short article

Students	Page	Para. #	Expression	Count	Time	Data
Student1	2	1	neutral	7	21:44:47 - 21:45:12	I once asked a clinical colleague to describe what it had been like working as a trainee with a world-famous surgeon...
Student1	2	2	neutral	4	21:45:12 - 21:45:37	Although infrequent, this type of desensitization, or inurement, unfortunately, does occur in clinical settings...
Student1	2	3	neutral	3	21:45:37 - 21:46:02	Hannah Arendt, the German-American philosopher, coined the term "banality of evil" in her New Yorker series covering the Nazi war criminal trial of Adolph Eichmann
Student1	2	4	neutral	3	21:46:02 - 21:46:27	Despite the fact that desensitization and the "Milgram" effect are real, what I am really concerned about in health care.
Student1	3	1	neutral	1	21:46:49 - 21:47:14	Could the use of technologies that distance health care providers from the first-hand experience of the suffering...
Student1	3	2	neutral	2	21:47:14 - 21:47:39	I recently spoke with Dr. Eric Cassell - a pioneer in the development of the specialty of palliative care.
Student1	3	3	neutral	5	21:47:39 - 21:48:04	"They are related," he said. "The problem that physicians have now is not that they get too close to their patients, but that they do not get close enough... The suffering has special therapeutic ..."
Student1	3	4	neutral	2	21:48:04 - 21:48:29	"The thing that heals people is relationships - the problem is that technology has the ability to actually facilitate relationships."
Student1	4	1	neutral	3	21:48:40 - 21:49:01	"Death becomes an abstraction, except for those at the receiving end," he says. "What is at stake is not only the psyche of each young fighter but our humanity"
Student1	4	2	neutral	2	21:49:01 - 21:49:22	Mars was the Roman god of war, and the Romans believed that Mars sanctioned war only as a means...

Table 3 continued

Students	Page	Para. #	Expression	Count	Time	Data
Student1	4	2	surprised	1	21:49:01 - 21:49:22	Although the latter may be achieved with war from a distance, the former simply cannot...
Student1	4	3	neutral	2	21:49:22 - 21:49:43	The adoption and deployment of next-generation technologies made possible by the digital...
Student1	4	4	neutral	5	21:49:43 - 21:50:04	However, as we use these new tools, let us take care to not allow them to create "walls difficult to traverse" between those that are providing care, and those that are suffering.
Student2	2	1	neutral	1	13:36:35 - 13:37:13	I once asked a clinical colleague to describe what it had been like working as a trainee with a world-famous surgeon.
Student2	2	1	happy	3	13:36:35 - 13:37:13	His response was surprising and troubling - "Oh... he's brilliant and talented, of course, but to be honest - he's dangerous.
Student2	2	1	neutral	8	13:36:35 - 13:37:13	He has become more or less indifferent to patient suffering, and even death - I guess because he has seen so much of it."
Student2	2	2	neutral	8	13:37:13 - 13:37:51	Although infrequent, this type of desensitization, or inurement, unfortunately, does occur in clinical settings...
Student2	2	3	neutral	6	13:37:51 - 13:38:29	Hannah Arendt, the German-American philosopher, coined the term "banality of evil" in her New Yorker series covering the Nazi war criminal trial of Adolph Eichmann
Student2	2	4	neutral	8	13:38:29 - 13:39:07	Despite the fact that desensitization and the "Milgram" effect are real, what I am really concerned about in health care.
Student2	3	1	neutral	10	13:39:08 - 13:39:58	Could the use of technologies that distance health care providers from the first-hand experience of the suffering...
Student2	3	2	neutral	10	13:39:58 - 13:40:48	I recently spoke with Dr. Eric Cassell - a pioneer in the development of the specialty of Palliative Care.

Table 3 continued

Students	Page	Para. #	Expression	Count	Time	Data
Student2	3	3	neutral	11	13:40:48 - 13:41:38	"They are related," he said. "The problem that physicians have now is not that they get too close to their patients, but that they do not get close enough... The suffering has special therapeutic...."
Student2	3	4	neutral	10	13:41:38 - 13:42:28	"The thing that heals people is relationships - the problem is that technology has the ability to actually facilitate relationships,
Student2	4	1	happy	1	13:42:33 - 13:43:17	"Death becomes an abstraction, except for those at the receiving end," he says. "What is at stake is not only the psyche of each young fighter but our humanity.
Student2	4	1	neutral	9	13:42:33 - 13:43:17	The chances for transformational psychological experiences are decreased enormously when you wage war with all the comforts of home. You have no way of knowing if you are in the sacred space of Mars and death, or coming home from it."
Student2	4	2	happy	1	13:43:17 - 13:44:01	Mars was the Roman god of war, and the Romans believed that Mars sanctioned war only as a means to achieve peace. In the spirit of this ancient viewpoint, Marliantes implies that for one to "enter the Temple of Mars"
Student2	4	2	neutral	8	13:43:17 - 13:44:01	That he must understand fully (experience) the consequences of his actions, and only seek war as a means to more humane ends. Although the latter may be achieved with war from a distance, the former simply cannot.
Student2	4	3	neutral	8	13:44:01 - 13:44:45	The adoption and deployment of next-generation technologies made possible by the digital....
Student2	4	4	neutral	10	13:44:45 - 13:45:29	However, as we use these new tools, let us take care to not allow them to create "walls difficult to traverse" between those that are providing care, and those that are suffering.
Student3	2	1	neutral	3	10:28:16 - 10:28:38	I once asked a clinical colleague to describe what it had been like working as a trainee with a world-famous surgeon....



Table 3 continued

Students	Page	Para. #	Expression	Count	Time	Data
Student3	2	2	neutral	1	10:28:38 - 10:29:00	Although infrequent, this type of desensitization, or inurement, unfortunately, does occur in clinical settings...
Student3	2	3	neutral	5	10:29:00 - 10:29:22	Hannah Arendt, the German-American philosopher, coined the term "banality of evil" in her New Yorker series covering the Nazi war criminal trial of Adolph Eichmann
Student3	2	4	neutral	4	10:29:22 - 10:29:44	Despite the fact that desensitization and the "Milgram" effect are real...
Student3	3	1	neutral	2	10:29:46 - 10:30:11	Could the use of technologies that distance health care providers from the first-hand experience of the suffering...
Student3	3	2	neutral	4	10:30:11 - 10:30:36	I recently spoke with Dr. Eric Cassell - a pioneer in the development of the specialty of Palliative Care.
Student3	3	3	neutral	4	10:30:36 - 10:31:01	"They are related," he said. "The problem that physicians have now is not that they get too close to their patients, but that they do not get close enough... The suffering has special therapeutic..."
Student3	3	4	neutral	4	10:31:01 - 10:31:26	"The thing that heals people is relationships - the problem is that technology has the ability to actually facilitate relationships,
Student3	4	1	neutral	5	10:31:36 - 10:32:01	"Death becomes an abstraction, except for those at the receiving end," he says. "What is at stake is not only the psyche of each young fighter but our humanity..."
Student3	4	2	neutral	2	10:32:01 - 10:32:26	Mars was the Roman god of war, and the Romans believed that Mars sanctioned war only as a means to achieve peace.
Student3	4	3	neutral	4	10:32:51 - 10:33:16	The adoption and deployment of next-generation technologies made possible by the digital revolution - those that move the locus of care closer to the patient.

Table 3 continued

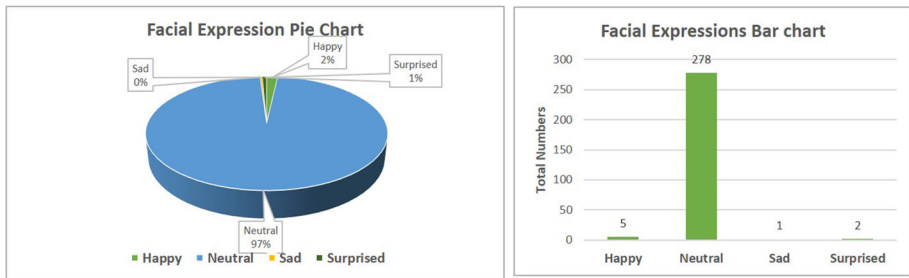
Students	Page	Para. #	Expression	Count	Time	Data
Student4	2	1	neutral	9	19:42:51 - 19:43:14	I once asked a clinical colleague to describe what it had been like working as a trainee with a world-famous surgeon...
Student4	2	2	neutral	4	19:43:14 - 19:43:37	Although infrequent, this type of desensitization, or inurement, unfortunately, does occur in clinical settings...
Student4	2	3	neutral	2	19:43:37 - 19:44:00	Hannah Arendt, the German-American philosopher, coined the term "banality of evil" in her New Yorker series covering the Nazi war criminal trial of Adolph Eichmann
Student4	2	4	neutral	3	19:44:00 - 19:44:23	Despite the fact that desensitization and the "Milgram" effect are real...
Student4	3	1	neutral	4	19:44:26 - 19:44:49	Could the use of technologies that distance health care providers from the first-hand experience of the suffering of those they care...
Student4	3	2	neutral	4	19:44:49 - 19:45:12	I recently spoke with Dr. Eric Cassell - a pioneer in the development of the specialty of Palliative Care. He wrote a landmark article published in the New England Journal of Medicine in 1982, entitled
Student4	3	2	surprised	1	19:44:49 - 19:45:12	the fact that many physicians are "burning out" in a time when technology should actually be making their jobs easier
Student4	3	3	neutral	4	19:45:12 - 19:45:35	"They are related," he said. "The problem that physicians have now is not that they get too close to their patients, but that they do not get close enough..."
Student4	3	4	neutral	5	19:45:35 - 19:45:58	"The thing that heals people is relationships - the problem is that technology has the ability to actually facilitate relationships,

Table 3 continued

Students	Page	Para. #	Expression	Count	Time	Data
Student4	4	1	neutral	3	19:46:01 - 19:46:21	"Death becomes an abstraction, except for those at the receiving end," he says. "What is at stake is not only the psyche of each young fighter but our humanity ..."
Student4	4	2	neutral	5	19:46:21 - 19:46:41	Mars was the Roman god of war, and the Romans believed that Mars sanctioned war only as a means to achieve peace.
Student4	4	3	neutral	5	19:46:41 - 19:47:01	The adoption and deployment of next-generation technologies made possible by the digital revolution - those that move the locus of care closer to the patient.
Student4	4	4	neutral	5	19:47:01 - 19:47:21	However, as we use these new tools, let us take care to not allow them to create "walls difficult to traverse" between those that are providing care.
Student5	2	1	neutral	2	20:16:44 - 20:16:58	I once asked a clinical colleague to describe what it had been like working as a trainee with a world-famous surgeon...
Student5	2	2	neutral	1	20:16:58 - 20:17:12	Although infrequent, this type of desensitization, or inurement, unfortunately, does occur in clinical settings...
Student5	2	3	neutral	2	20:17:12 - 20:17:26	Hannah Arendt, the German-American philosopher, coined the term "banality of evil" in her New Yorker series covering the Nazi war criminal trial of Adolph Eichmann
Student5	2	3	sad	1	20:17:12 - 20:17:26	Although she was aggressively criticized by many for implying an apology for the Third Reich's role in the Holocaust
Student5	2	4	neutral	4	20:17:26 - 20:17:40	Despite the fact that desensitization and the "Milgram" effect are real...

Table 3 continued

Students	Page	Para. #	Expression	Count	Time	Data
Student5	3	1	neutral	4	20:17:45 - 20:18:02	Could the use of technologies that distance health care providers from the first-hand experience of the suffering of those they care...
Student5	3	2	neutral	3	20:18:02 - 20:18:19	I recently spoke with Dr. Eric Cassell - a pioneer in the development of the speciality of Palliative Care. He wrote a landmark article published in the New England Journal of Medicine in 1982, entitled
Student5	3	3	neutral	4	20:18:19 - 20:18:36	"They are related," he said. "The problem that physicians have now is not that they get too close to their patients, but that they do not get close enough..."
Student5	3	4	neutral	3	20:18:36 - 20:18:53	"The thing that heals people is relationships - the problem is that technology has the ability to actually facilitate relationships,
Student5	4	1	neutral	5	20:18:55 - 20:19:18	"Death becomes an abstraction, except for those at the receiving end," he says. "What is at stake is not only the psyche of each young fighter but our humanity..."
Student5	4	2	neutral	5	20:19:18 - 20:19:41	Mars was the Roman god of war, and the Romans believed that Mars sanctioned war only as a means to achieve peace.
Student5	4	3	neutral	4	20:19:41 - 20:20:04	The adoption and deployment of next-generation technologies made possible by the digital revolution - those that move the locus of care closer to the patient.
Student5	4	4	neutral	4	20:20:04 - 20:20:27	However, as we use these new tools, let us take care to not allow them to create "walls difficult to traverse" between those that are providing care.



**Fig. 4** Illustration of Emotions through charts

Face recognition application only starts recording their emotions after the user has allowed the application to access the camera. Through video streaming, the system starts detecting students' emotions. These emotions are securely stored in the database. Figure 4 presents the descriptive analysis of students' emotions obtained during the reading of the short text. This short text was used to conduct a preliminary evolution of the emotion detection application.

Table 4 describes the scenario of students who read paragraph 1 of page 2, We intend to know the situation of students studying online by presenting the analysis of their emotions in a single instance. For example, students' emotions while reading paragraph 1 page 2 show varied expressions. While Student 1 recorded 7 neutral sentiments reading a particular text "I once asked a clinical colleague to describe what it had been like working as a trainee with a world-famous surgeon", Student 3 and Student 5 recorded 3, 9, and 2 neutral emotions, respectively. This result suggests that the same test had a different impact on students.

## 5.5 Discussion

The proposed system's emotion detection has proven to be highly efficient, as evidenced by the results we obtained. Previous studies have utilized various techniques, such as vector descriptors for facial motion, active contours for the mouth, eye shape retrieval, and 2D deformable mesh models [3]. In contrast, the proposed system uses a model based on CNN for feature extraction and recognition, which is a cutting-edge approach to emotion detection [3]. Additionally, this face expression recognition model was trained on a diverse set of images from publicly available datasets [16]. However, certain physical factors such as hardware limitations, camera quality, low light conditions, and wearing glasses may reduce the accuracy of results.

The proposed system is quite helpful for teachers to understand the learner's emotions. This application detects learners' emotions every five seconds, meaning teachers can instantly understand the learner's mood or emotion and adjust his/her teaching method according to the learner's behavior. Concerning the students who participated in the experiments, most of the students gave neutral expressions while reading. We can see the bar of neutral expression is far higher than the other ones (4), but on the other hand, we also got other expressions i.e., happy, sad, and surprised. Most of the students only change expressions when they find something different while reading. Also depending on the content, if the content is funny then most probably we can get more happy expressions.

Table 4 lists expressions of five students while reading a particular paragraph 1 of page 2. We can see the count of every student is different from each other. Student 1, Student 3, Student 4, and Student 5 read the same text but the count of their expression varies. On the

**Table 4** Student's emotions while reading a particular paragraph of a short article

Students	Page #	Para. #	Expression	Count	Time	Data
Student1	2	1	neutral	7	21:44:47 - 21:45:12	I once asked a clinical colleague to describe what it had been like working as a trainee with a world-famous surgeon...
Student2	2	1	happy	3	13:36:35 - 13:37:13	His response was surprising and troubling - "Oh... he's brilliant and talented, of course, but to be honest - he's dangerous.
Student3	2	1	neutral	3	10:28:16 - 10:28:38	I once asked a clinical colleague to describe what it had been like working as a trainee with a world-famous surgeon...
Student4	2	1	neutral	9	19:42:51 - 19:43:14	I once asked a clinical colleague to describe what it had been like working as a trainee with a world-famous surgeon...
Student5	2	1	neutral	2	20:16:44 - 20:16:58	I once asked a clinical colleague to describe what it had been like working as a trainee with a world-famous surgeon...

other hand, Student 2 read a different passage “His response was surprising and troubling - “Oh... he’s brilliant and talented, of course, but to be honest - he’s dangerous” and gave a different expression. Moreover, if we look at the time duration in Table 4 of each student we can see that those students who spent more time on paragraph 1 gave more emotions. For instance, Student 1 and Student 4 spent almost 1:30 minutes on paragraph 1, and the application recorded 7 and 9 neutral expressions respectively. Student 3 and Student 5 spent a few seconds and the application recorded 3 and 2 neutral expressions respectively.

## 5.6 Comparison with the existing state-of-the-art studies

An ideal system for real-time facial emotion recognition would typically consist of the following components:

**Face detection:** A face detection algorithm that can accurately detect the presence and position of faces in real-time video or image streams.

**Feature extraction:** A feature extraction algorithm that can extract relevant features from the face, such as the position and shape of facial landmarks, facial textures, and the movement of facial muscles, in real-time.

**Preprocessing:** A preprocessing stage that normalizes the images, removes noise and enhances contrast to improve the accuracy of feature extraction in real-time.

**Classification model:** A machine learning model, such as a deep neural network or a support vector machine, that can learn from the extracted features to classify different facial expressions in real-time.

**Training data:** A large and diverse dataset of annotated facial expressions can be used to train the classification model.

**Real-time data processing:** A data processing framework that can handle the real-time streaming of video or images and process the extracted features and classification results in real-time.

**Output and feedback:** A real-time output and feedback system that can display the classification results and provide feedback to the user in real-time, such as using audio, visual, or haptic signals.

Overall, an ideal real-time facial emotion recognition system should be accurate, fast, and able to detect and classify a wide range of facial expressions under different lighting conditions, poses, and occlusions in real time. We have devised some parameters in Table 6 in order to analyze the quality of a good facial emotion recognition model. The comparison of the proposed model with state-of-the-art research is shown in Table 5.

## 5.7 Limitation of the study

The proposed model for real-time emotion detection during online learning shows promising results with high accuracy. However, it does have certain limitations that need consideration. One limitation is the challenge of accurately capturing and interpreting complex emotional

**Table 5** Performance evaluation of the proposed system against previously proposed systems for automatic facial emotion recognition

Reference	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
<b>Comparison with static facial images.</b>																				
[47]	✓	-	✓	✓	✓	✓	✓	✗	✓	-	✓	✗	✓	✓	✓	✗	7	✗	✗	✗
[48]	✓	1	-	✓	✓	✗	✓	✓	-	✓	✗	✓	-	✗	✓	6	✗	✗	✗	✓
[49]	✓	-	✗	✓	✓	✓	✓	✓	✓	✓	✗	✓	✓	✗	✓	7	✗	✗	✗	✓
[50]	✓	1	✗	✓	✓	✗	✓	✓	-	✓	✗	✓	✓	✗	✓	6	✗	✗	✗	✗
[51]	✗	✓	✗	✓	✓	-	✗	✗	-	✓	✗	✓	✓	✗	✓	7	✗	✗	✗	✗
[52]	✗	✓	✗	✓	✓	-	✗	✗	-	✓	✗	✓	✓	✗	✓	7	✗	✗	✗	✗
[53]	✓	3	✗	✓	✓	✓	✓	✓	✓	✓	✗	✓	✓	31	✓	6	✗	✓	✓	✗
[54]	✓	1	-	✓	✓	✗	✓	✓	-	✓	✗	✓	✓	✗	✓	4	✗	✗	✗	-
[55]	✗	✓	✗	✓	✓	-	✗	✗	-	✓	✗	✓	✓	✗	✓	7	✗	✓	✓	✗
[56]	✗	✓	✗	✓	✓	-	✗	✗	-	✓	✗	✓	✓	✗	✓	7	✗	✓	✓	✗
<b>Comparison with real time sequence of images.</b>																				
[57]	✓	-	✓	✓	✓	✓	✗	✓	✗	✓	✗	✓	✓	-	✓	6	✗	✓	✗	✗
[58]	✓	3	✗	✓	✓	✗	✗	✓	✗	✓	✗	✓	✓	15	✓	-	✗	✗	✗	-
[59]	✓	✓	✓	✓	✓	-	✓	✓	✓	✓	✗	✓	✓	2	✓	4	✗	✗	✗	✓
[60]	✓	✗	✓	✓	✓	✗	✓	✓	✓	✓	✗	✓	✓	✗	✓	3	✗	✓	✗	-
[61]	✓	-	-	✓	✓	✓	-	✓	✗	✓	✗	✗	✓	✗	✓	6	✗	✗	✗	✗
[62]	✓	1	✗	✓	✓	✗	✗	✓	-	✓	✗	✓	✓	✗	✓	3	✗	✓	✗	✗
<b>Proposed</b>	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

**Table 6** Characteristics of a novel facial emotion recognition tool

1	Automatic facial image acquisition
2	People of any age, outlook, and ethnicity
3	Covers variation in lighting
4	No special filters required
5	Deals with partially occluded faces
6	Deals with head motions
7	Autonomous face detection
8	Autonomous facial feature extraction module
9	Deals with inappropriate facial expression data
10	Deals with unilateral facial changes
11	Differentiate all possible facial expressions
12	Autonomous facial expression classification module
13	obeys anatomical rules
14	Distinguishes 44 types of facial expressions
15	Quantifies facial actions codes
16	Quantifies interpretation categories
17	Adaptive feature learning facility
18	Assigning quantified interpretation labels
19	Assigning multiple interpretation labels
20	Real-time processing of features



nuances. Emotions are intricate and can vary greatly based on individual experiences and cultural backgrounds, which may impact the precision of the model's emotional recognition.

Furthermore, ethical concerns arise when dealing with students' emotional data, as it involves privacy issues and the responsible use of personal information. Safeguarding the collected emotional data and ensuring its proper use is of utmost importance to protect the students' well-being and rights.

Another aspect that requires attention is the generalization of the model across diverse populations. The accuracy achieved with specific datasets, such as CK+ and JAFFE, may not translate uniformly to other demographics and contexts, leading to potential biases and reduced reliability when applied in real-world scenarios.

To maximize the effectiveness of the real-time emotion detection system, ongoing research, and development are needed to refine and expand the model's emotional recognition capabilities. Additionally, comprehensive validation with a broader and more diverse set of datasets will be crucial to ensure its adaptability and accuracy across various learning scenarios.

Considering these limitations and addressing them in the implementation and future iterations of the model will contribute to a more robust and reliable system, enabling valuable insights into students' emotional responses during online learning and supporting their overall educational experience.

## 6 Conclusion

Emotion detection while reading online text is a thriving research area and with the rise of the online learning system, the need to detect human expression/emotions has increased. Existing methods used vector descriptions for face motion along with contours for mouth and eye shape which do not provide good accuracy. This study presents an emotion detection approach to obtain high-quality, robust, and accurate results. The emotion detection is carried out every five seconds. Experiments were performed on the CK+ database and JAFFE face database and the obtained results show the accuracy of 96.46% and 98.43%, respectively. It is observed that the number of expressions depends on the time duration that a student spends while reading a particular text. Different type of expressions also depends on the content and student's behavior.

In the future, we intend to include more participants in this research. Also, we can extend this research with teachers on how teachers can get useful results in the form of feedback from students. Dealing with other challenges including camera position, lighting conditions, and noise is also left for the future.

**Acknowledgements** Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2023R125), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

**Funding** This research was funded by Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2023R125), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

**Data Availability** The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

## Declarations

**Conflicts of interest** The authors declare no conflict of interest.

## References


1. Altameem T, Altameem A (2020) *Image and Vision Computing* 103:104044
2. Jin H, Wang X, Lian Y, Hua J (2019) *The Visual Computer* 35:535
3. Breuer R, Kimmel R (2017) A deep learning perspective on the origin of facial expressions
4. Hasani B, Mahoor MH (2017) In: *Proceedings of the IEEE conference on computer vision and pattern recognition workshops* pp. 30–40
5. Syafeeza A, Khalil-Hani M, Liew S, Bakhteri R (2014) *International Journal of Engineering & Technology* 6(1):0975
6. Georgakopoulos SV, Tasoulis SK, Vrahatis AG, Plagianakos VP (2018) In: *Proceedings of the 10th hellenic conference on artificial intelligence*, pp 1–6
7. Wang X, Jiang W, Luo Z (2016) In: *Proceedings of COLING 2016, the 26th international conference on computational linguistics: Technical papers*, pp 2428–2437
8. Krizhevsky A, Sutskever I, Hinton GE (2017) *Communications of the ACM* 60(6):84
9. Sawyer R, Smith A, Rowe J, Azevedo R, Lester J (2017) In: *Proceedings of the 25th conference on user modeling, adaptation and personalization*, pp 192–201
10. Munezero M, Montero CS, Mozgovoy M, Sutinen E (2013) In: *Proceedings of the 13th Koli Calling International Conference on Computing Education Research* , pp 145–152
11. Kolog EA, Montero CS, Toivonen T (2018) In: *Proceedings of the International Conference on Information Technology & Systems (ICITS 2018)* (Springer), pp 453–463
12. Ekman P (2006) Darwin and facial expression: A century of research in review (Ishk, 2006)
13. Díaz-Aristizabal U, Valdés-Vilches M, Fernández-Ferreras T, Calero-Muñoz E, Bienzobas-Allué E, Moracén-Naranjo T (2019) *Neurología (English Edition)* 34(7):423
14. Ragheb W, Azé J, Bringay S, Servajean M (2019) [arXiv:1906.07020](https://arxiv.org/abs/1906.07020)
15. Shivhare SN, Khethawat S (2012) [arXiv:1205.4944](https://arxiv.org/abs/1205.4944)
16. Muhler, V (2020) [justadudewhohacks/face-api.js](https://github.com/justadudewhohacks/face-api.js) . <https://github.com/justadudewhohacks/face-api.js/#models-face-detection> (May 2020). Accessed 10 February 2023
17. Yang H, Ciftci U, Yin L (2018) In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp 2168–2177
18. Zhang M, Meng W (2020) In: *Proceedings of the 28th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering*, pp 38–49
19. Kawulok M, Celebi E, Smolka B (2016) *Advances in face detection and facial image analysis* (Springer)
20. Zen G, Porzi L, Sanginetto E, Ricci E, Sebe N (2016) *IEEE Transactions on Multimedia* 18(4):775
21. Refat CMM, Azlan NZ (2019) In: *2019 7th International Conference on Mechatronics Engineering (ICOM) (IEEE)*, pp 1–6
22. Sathik M, Jonathan SG (2013) SpringerPlus 2:1
23. Yu Z, Zhang C (2015) In: *Proceedings of the 2015 ACM on international conference on multimodal interaction*, pp 435–442
24. Praveen RG, de Melo WC, Ullah N, Aslam H, Zeeshan O, Denorme T, Pedersoli M, Koerich AL, Bacon S, Cardinal P et al (2022) In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp 2486–2495
25. Ge H, Zhu Z, Dai Y, Wang B, Wu X (2022) *Computer Methods and Programs in Biomedicine* 215:106621
26. Ahmed M, Rasool AG, Afzal H, Siddiqi I (2017) *Expert Systems with Applications* 85:158
27. Alluqmani A, Shamir L (2018) *Scientometrics* 115(2):1071
28. Gleaves A, Walker C, Grey J (2008) *Assessment & Evaluation in Higher Education* 33(3):219
29. Perez-Rosas V, Kleinberg B, Lefevre A, Mihalcea R (2017) [arXiv:1708.07104](https://arxiv.org/abs/1708.07104)
30. Zhang Y, Zhao D, Sun J, Zou G, Li W (2016) *Neural Processing Letters* 43:389
31. Ononye U, Ndudi F, Bereprebofa D, Maduemezia I (2022) *Knowledge and Performance Management* 6(1):1
32. Khassawneh O, Mohammad T, Ben-Abdallah R, Alabidi S (2022) *Behavioral Sciences* 12(12):511
33. AlZu'bi S, Abu Zitar R, Hawashin B, Abu Shanab S, Zraiqat A, Mughaid A, Almotairi KH, Abualigah L (2022) *Electronics* 11(18). <https://doi.org/10.3390/electronics11182964>. <https://www.mdpi.com/2079-9292/11/18/2964>
34. Pekrun R, Marsh HW, Suessenbach F, Frenzel AC, Goetz T (2023) *Learning and Instruction* 83:101626
35. Sander P, de la Fuente J (2022) *Current Psychology* 41(7):4329
36. Chen X, Lu L (2022) *Learning and Individual Differences* 98:102183
37. Zhu Y, Xu S, Wang W, Zhang L, Liu D, Liu Z, Xu Y (2022) *Education and Information Technologies* 27(7):8921
38. Oyelere SS, Suhonen J, Wajiga GM, Sutinen E (2018) *Education and Information Technologies* 23:467
39. Seyeditabari A, Tabari N, Zadrozny W (2018) [arXiv:1806.00674](https://arxiv.org/abs/1806.00674)

40. Al-Saqqa S, Abdel-Nabi H, Awajan A (2018) In: 2018 8th International Conference on Computer Science and Information Technology (CSIT), pp 136–142 (IEEE)
41. Bandhakavi A, Wiratunga N, Massie S, Padmanabhan D (2017) IEEE intelligent systems 32(1):102
42. Chatterjee A, Narahari KN, Joshi M, Agrawal P (2019) In: Proceedings of the 13th international workshop on semantic evaluation (2019), pp. 39–48
43. Herzig J, Shmueli-Scheuer M, Konopnicki D (2017) In: Proceedings of the ACM SIGIR international conference on theory of information retrieval , pp 269–272
44. Muyltermans D (2019)
45. Krumhuber EG, Skora L (2016) Handbook of human motion pp. 2271–2285
46. Smythe, R (2023) Delivering Health Care By Drone. <https://www.forbes.com/sites/roysmythe/2014/09/22/delivering-health-care-by-drone/?sh=1d97c0004917> (May 2021). Accessed 10 February 2023
47. Edwards GJ, Cootes TF, Taylor CJ (1998) In: Computer Vision'98: 5th European Conference on Computer Vision Freiburg, Germany, June 2–6, 1998 Proceedings, Volume II 5 (Springer, 1998), pp 581–595
48. Kobayashi H, Hara F (1997) In: 1997 IEEE international conference on systems, man, and cybernetics. Computational cybernetics and simulation, vol. 4, pp 3732–3737 (IEEE)
49. Hong H, Neven H, Von der Malsburg C (1998) In: Proceedings third IEEE international conference on automatic face and gesture recognition, pp 354–359 (IEEE)
50. Huang CL, Huang YM (1997) Journal of visual communication and image representation 8(3):278
51. Lyons MJ, Budynek J, Akamatsu S (1999) IEEE transactions on pattern analysis and machine intelligence 21(12):1357
52. Padgett C, Cottrell G (1996) Advances in neural information processing systems 9
53. Pantic M, Rothkrantz LJ (2000) Image and Vision Computing 18(11):881
54. Yoneyama M, Ohtake A, Iwano Y, Shirai K (1997) In: Proceedings of international conference on image processing, vol. 1, pp 117–120 (IEEE)
55. Zhang Z, Lyons M, Schuster M, Akamatsu S (1998) In: Proceedings Third IEEE International Conference on Automatic face and gesture recognition, pp 454–459 (IEEE)
56. Zhao J, Kearney G (1996) In: International Conference on Neural Information Processing, vol. 1, pp 454–457
57. Black MJ, Yacoob Y (1997) International Journal of Computer Vision 25:23
58. Cohn JF, Zlochowicz AJ, Lien JJ, Kanade T (1998) In: Proceedings third IEEE international conference on automatic face and gesture recognition, pp 396–401 (IEEE)
59. Essa IA, Pentland AP (1997) IEEE transactions on pattern analysis and machine intelligence 19(7):757
60. Kimura S, Yachida M (1997) In: Proceedings of IEEE computer society conference on computer vision and pattern recognition, pp 295–300 (IEEE)
61. Otsuka T, Ohya J (1998) In: Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition, pp. 442–447 (IEEE)
62. Wang M, Iwai Y, Yachida M (1998) In: Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition, pp 324–329 (IEEE)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

## Authors and Affiliations

Alisha Fida<sup>1</sup> · Muhammad Umer<sup>1</sup> · Oumaima Saidani<sup>2</sup> · Monia Hamdi<sup>3</sup> ·  
Khaled Alnowaiser<sup>4</sup> · Carmen Bisogni<sup>5</sup> · Andrea F. Abate<sup>5</sup> · Imran Ashraf<sup>6</sup> 

Alisha Fida  
Alisha.fida@iub.edu.pk,umer.sabir@iub.edu.pk

Oumaima Saidani  
ocsaidani@pnu.edu.sa

Monia Hamdi  
MSHamdi@pnu.edu.sa

Khaled Alnowaiser  
k.alnowaiser@psau.edu.sa

Carmen Bisogni  
cbisogni@unisa.it; abate@unisa.it

- <sup>1</sup> Department of Computer Science, The Islamia University of Bahawalpur, Bahawalpur, Pakistan
- <sup>2</sup> Department of Information Systems, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, Riyadh 11671 84428,, Saudi Arabia
- <sup>3</sup> Department of Information Technology, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, Riyadh 11671 84428,, Saudi Arabia
- <sup>4</sup> Department of Computer Engineering, College of Computer Engineering and Sciences,Prince Sattam Bin Abdulaziz University, Al-Kharj 11942, Saudi Arabia
- <sup>5</sup> Department of Computer Science, University of Salerno, Fisciano, Italy
- <sup>6</sup> Department of Information and Communication Engineering, Yeungnam University, Gyeongsan 38541, South Korea