Enhancing Online Teaching and Learning Through Machine Learning and Learning Analytics

Sophiya Mathews¹ and Dr. D. John Aravindhar²

Abstract

Online learning environments are gaining popularity as an alternative to traditional learning environments during times of global pandemic. The traditional classroom setting has given way to online learning in recent educational developments. To employ it as a classifier in online education, the goal of this work is to construct a real-time face emotion detection system that recognizes and categorizes emotions of a human. It uses machine learning to identify, forecast, and analyse a learner's facial expressions, and it further maps those expressions to a learning affect that categorizes the emotions of those who are seen on camera. Academic feelings can have a significant impact on learning outcomes. Students typically show their emotions through their facial expressions, voice, and behavior. A spontaneous facial expression database is created in light of the inference algorithm's lack of training samples. It consists of two subsets: a video clip database and a picture database, and it includes the typical emotional facial expressions. The database contains 1,274 video clips and 30,184 photos from 82 pupils. The parameters for student facial detection can be used to gauge each learner's rate of concentration. The performances of the SVM and RF classifier in facial expression for the image are presented. The results of testing this with the RF algorithm and a 90.14% accuracy rate produced extremely good results.

Keywords: Facial Expression Recognition, Deep Learning, Education, Random Forest Classifier, Online Education.

Introduction

The 21st century student is transitioning to an Online Learning, emphasising on the relationships between teachers and students to reach the objective of relevant, excellent, and dynamic education. Online teaching is defined as educational activities conducted using a range of internet-connected electronic devices in synchronous or asynchronous settings. With online learning, the educational process could become more flexible, inventive, and student-centered. Online course delivery is both economical and practical for reaching students in rural and remote areas with content [38]. Online education technology has successfully bridged the gap caused by the current pandemic by facilitating virtual classroom interactions that transcend physical boundaries, enabling students to learn from anywhere and at any time.

Classroom interaction refers to the actions of teachers’ and the students' reaction to the current activity. Recognition of facial expressions in the classroom has the potential to forecast the importance of students' emotions. An interactive feedback system for student conduct during lectures, as opposed to conventional assessment methods, may not only enhance the learning environment but also conserve time and resources [25].

¹ Assistant Professor, Department of Computer Science, SNMIMT. sophiyamathews@gmail.com,
² Professor, Department of CSE, HITS, Padur, Chennai. jaravindhar@hindustanuniv.ac.in
The delivery of knowledge via e-learning has many benefits, but the majority of these benefits can only be properly tapped into if the student's involvement and interest are sustained throughout the online education process [4]. People from all walks of life are given the chance to study and teach themselves devoid of any limitations of time or location because students generally choose to acquire knowledge at their own speed. The need for better online education is growing as an outcome of the advancement of online-learning technology and the rising enrollment of students. It is acknowledged that more thorough research is required to identify the factors that can actually improve the online learning environment [15].

Due to the necessity to offer students a quality education regardless of where they live, e-learning is becoming more popular in higher education institutions [60, 58]. Due to the high dropout rate among e-learning students, and the fast growth in the number of institutions offering e-learning courses, it is crucial to assess the efficacy and sustainability of these learning platforms [54]. In any realistic teaching environment, interaction and communication are essential instruments that boost student motivation and pleasure [33]. Communications among students and teacher occurs orally, through the use of language of body, and through the use of facial expressions in a traditional classroom context. In these interactions, which are typically two-way, the instructor can read the learner's expression of emotion on their face and respond accordingly by asking questions and, of course, altering the manner the content is delivered [44, 39, 35, and 28]. However, communication in an e-learning environment is typically one-way. Since e-learning systems are used in a classroom environment, real-time feedback is not available to e-learners. The most common non-verbal response has been found to be facial emotion expression, and teachers can utilise this to determine the comprehension level of their students. Numerous academics have employed various methodologies to examine how facial emotion expressions affect learning. However, the majority of research studies have concentrated on the conventional classroom setting, despite expanding study in the area of facial expression detection as a method of providing feedback to students or instructors. Chat rooms and video conferencing are used by e-learning platform to allow students to communicate with one another and the instructor. Hence, it is essential to provide a platform that offers learners comments in real-time when they are participating in an online course.

The fundamental issue with e-learning is that there is no supervisor to monitor how the provided content is affecting the students' physical and emotional well-being. When taking a course online, students frequently lose focus and concentration, which negatively affects their academic achievement. By addressing this problem, the e-learning method can be advanced significantly since it will be possible to gauge each pupil's interest level and make the required material changes to keep the user interested during the online lecture. This study is being conducted with the goal of analysing the relationship between a student's facial expressions while using an online-learning scheme and methods to enhance the studying approach of such student by means of data extrapolated from these features to get around the difficulty of analysing students on online-learning platforms. These systems can be viewed as a good thing and a significant IT asset. It is a medium for delivering learning whenever and wherever it is needed. Since the learning environment supplied to the student during e-learning offers sophisticated insights on the pupil's learning curve, transfer of knowledge through information technology tools demands careful preparation and implementation [16]. Important factors that directly affect both the studying curve of student and the online-learning objective include the material delivery, exams, and student feedback. Even so, the time period needed for relating and watching all of these metrics must be sufficient to take into account every conceivable aspect [37].

To as closely replicate the conventional classroom and studying environments as achievable, online-teaching involves the utilisation of text, audio, and video. There are several educational uses for e-learning environments. According to current advancements, e-learning-based teaching will soon be on par with conventional education approaches. In
an e-learning setting, the instructor provides content through online platforms employing multimedia and software interfaces without the student and teacher interacting directly. Since there are no instantaneous communication channels, the device knows only how to comprehend what it records through conventional human-device interfaces [6] one of the typical mood patterns for pupils displaying puzzled expressions may include one or more of the facial traits listed below, such as depressed or drawn-together eyebrows, horizontal or vertical forehead wrinkles, inconsistent eye contact, etc. A teacher can detect the subtle nonverbal signs displayed by the students' expressions in order to determine whether they understand what is being said [18].

**Basics of Facial Expression**

Facial Expression [14] is the technique of examining facial muscle and skin movement that is more indicative of a person's inner feelings and emotional activities. People use their expressed emotions to communicate their sentiments and information to others as well as to express their ideas and beliefs. In other words, the face expression, which is communicated through the mouth, eyes, and brows, is also known as non-verbal communication. Eye contact is one of the many communication techniques that help to enhance interpersonal communication. It aids in the communications procedure to create judgments about people based on their movements of eye. The other method of communication is spoken, and it allows them to express their emotions whether or not they use sign language. Even though there are additional characteristics that help individuals communicate effectively, face-to-face interaction still plays a crucial part.

Since the human face is the most abundant source for emotional information, facial expressions are crucial in social communication [12]. We also respond to facial expressions, some of which are even unintentional [9, 10]. As a type of feedback for learning, emotions are crucial since they let the teacher know how the student is feeling [29]. This is crucial while studying online because a fully automated system can be adjusted to the learner's emotional state [8]. The level of a student's engagement in online learning is strongly related to his ability to pay attention and pay attention to the information being presented. Facial expressions during brief periods of time can be deceptive, and a time-based scrutiny to determine states of emotion can produce compelling findings. In the sphere of education, automatic FER can be used to determine the affective state of a student. In order to give effective teaching tactics and increase student learning, it aids teachers in comprehending their students' areas of interest. For instance, FER developed an intelligent computer-assisted system to comprehend children's affective state and then deliver suitable help to enhance their social communication abilities.

The majority of human interaction in daily life involves facial expression, which is utilised to communicate information through emotions. In many applications and research projects aimed at creating efficient interaction systems; these emotions are given a significant role [27]. The subcortical brain area, which is linked to emotional expressions, is responsible for fully expressing a person's innermost thoughts, sensations, and ideas. Even if the facial expression can be seen on their face, it might be challenging to determine their exact sentiments. In order to reduce erroneous emotion identification and eradicate various emotion disorders in future research applications, a number of automatic emotion recognition systems have been developed. With the aid of machine learning, this work intends to create a system that can recognise, anticipate, and analyse a learner's facial emotions while also offering a platform for feedback.

**Emotions**

Emotion, as opposed to thinking, is the component of a person's personality that consists of their feelings. Emotions are intricate [50]. There are six main facial emotions, including anger, disgust, fear, happiness, sadness, and surprise, that can be expressed through the facial expressions face [56]. In terms of student emotions, such categorization is overly specific; for instance, the large amount of emotion is not always relevant when students are
watching video lessons in front of computers. As a result, we categorise the emotions that are relevant to and utilised in academic learning. Figure 1 depicts a typical face expression.

Figure 1: Typical Expressions on the Face

In the field of Human Computer Interaction, automatic face expression recognition and emotion computing are crucial. The feelings of an individual can be used to determine their emotional condition. Facial expression data is therefore frequently used in automatic FER systems. Both digital entertainment and police enforcement can make use of it. The six universal emotions are represented by the six facial expressions in Table 1; these expressions are categorised by the facial features they are displayed on, such as the eyes, brows, nose, lips, forehead, etc. These small, fleeting facial gestures are known as micro expressions since they don't last for very long.

Table 1: Different Emotional Expressions on the Face

<table>
<thead>
<tr>
<th>EMOTION</th>
<th>FACIAL DESCRIPTIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fear</td>
<td>Eyes Open, Mouth Open, Lips Retracted, Eyebrows Raised.</td>
</tr>
<tr>
<td>Anger</td>
<td>Eyes Wide Open, Mouth Compressed, Nurtures Raised.</td>
</tr>
<tr>
<td>Surprise</td>
<td>Mouth Open, Lower Lip down, Upper Lip raised.</td>
</tr>
<tr>
<td>Happy</td>
<td>Eyes Sparkle, Mouth drawn back at corners, Nose neutral, Skin under eyes wrinkled.</td>
</tr>
<tr>
<td>Surprise</td>
<td>Eyes Open, Mouth Open, Eyebrows Raised, Lip Protruded.</td>
</tr>
<tr>
<td>Sad</td>
<td>Mouth Corner depressed, Eyebrows Raised.</td>
</tr>
</tbody>
</table>

Techniques and Methods for Learning Analytics that Personalise Instruction

Learning analytics involves the measurement, collection, analysis, and reporting of data related to learners and their learning environments, aiming to better understand and enhance the learning process and the contexts in which it occurs. They offer a variety of fresh options to aid pupils in their learning [61]. It presents a wealth of information regarding student behaviour and learning needs, providing teachers and education designers with a fresh and useful source of data to supplement their own observations and evaluations. The study arrived to the ending that basic LA approaches should be used in this situation after carefully analysing related studies on the use of learning analytics in education.
The steps involved in learning analytics are shown in Figure 2. In order to obtain the optimal learning outcomes, teachers and students should be able to organize their work using trustworthy technologies that can generate specific, individualized recommendations for what needs to be done. Students' profiles should be adjusted in accordance with the data gathered using the information on actual students' behaviors in a learning environment that was obtained using LA method and technique. The majority of the illustrative information comes from only a few distinct regions of the face, such as the mouth and eyes and components, such as the ears and hair, play a considerably smaller role. This implies that the recognition model should only concentrate on components that are crucial to facial expressions [46]. The system we built in this study will assist the instructor in determining the best ways to improve the infrastructure for both the practical work scenario and the course scenario.

The Importance of Facial Expressions in Online Learning

Lower levels of student engagement are still a problem in online learning, but they can be easily fixed to a large amount in face-to-face teaching settings because the instructor can watch when student end paying attention to the material. Adaptive e-learning platforms have been introduced in online settings as a solution to this issue. Essentially, these systems adjust themselves to a learner based on their behaviour and a series of inferred assumptions about them [13]. In light of each student's goals, preferences, degree of expertise, and chosen learning method, these systems assist in customising learning materials for individual learners [11]. Systems for adaptive online learning change the material and how it should be delivered to each student. The creation and application of AES have grown over the past ten years and have become a key component of online learning platforms [43].

An interactive system that customises and adapts e-learning content, pedagogical models, and interactions between participants in the environment to match the unique preferences and needs of users, as and when they arise, is referred to as an adaptive e-platform, according to [47]. The teacher and the students' contact is the most crucial one in any classroom [36]. Faces have the capacity to convey details about a learner's frame of mind or current emotional and mental state of being, and therefore to some limit, their inner mind-set. As a result, communication through facial expressions plays a very crucial function during this encounter. [51] Asserts that the lecturer's facial expressions are crucial in maintaining the students' motivation during the entire session. A professor may also draw inspiration for upcoming sessions from the students' expressions. It can also be a sign to the teacher of the common information the student want to convey, such as whether the pace is appropriate for their learning and whether they are unclear about a particular subject. In summary, a lecturer should be able to gauge the audience's degree of understanding from their facial expressions and adjust their delivery accordingly.

Literature Review

A feed-forward training approach was created by Yusra et al. in [59] for a teacher's facial expression recognition technique in a classroom. The face was initially identified from collected lecture recordings, and key frames were selected, eliminating all extraneous frames. Deep feature were then collected and provided to a classifier utilising a number of convolution neural networks and parameter optimization. Modern techniques, conventional classifiers, and CNN models were contrasted with the suggested strategy. The results of the studies showed a significant increase in recall, F1-score, and accuracy. In [21], Hanusha and Varalatchoumy investigated how CNN could use student facial expressions to construct a facial expression recognition model for online teaching platforms. Major problem in the e-learning environment was identified to be the listener's poor level of engagement. The ideal learning environment for online learners with the maximum level of involvement in educational activities was the responsibility of educational institutions and teachers.

In [40], Moutan et al. developed a model utilising CNN to evaluate students' mental health. The findings showed 65% and 62% accuracy for classifying emotions and identifying states.
of mind, respectively. A new approach for measuring learning engagement was developed by Junge et al. [25]. An attentional mechanism was introduced to the network for feature extraction to address the issue of web camera images with complex backgrounds, lighting, and resolutions. A compact framework designed for real-time applications was used to develop CNN for categorisation. Additionally, domain adaptation was used to solve the issue of lacking labelled data. The suggested technique may accurately discern emotions with little labelled data, according to experiments in the paper.

Nazia et al. [41] captured facial expressions from a variety of subjects using a Panasonic camera with a 5mm focal length. Each person's six fundamental expressions were photographed with the camera four feet away from them. Following the retrieval of facial features, classification using K-NN was used. It was 100% correct to feel happy, 80% accurate to feel angry, 80% accurate to feel sad, 100% accurate to feel scared, 80% accurate to feel disgusted, and 100% accurate to feel surprised. 90% of the time was accurate overall. Using the LBP technique, Turabzadeh et al. [55] focused on real-time facial emotion recognition. LBP features were extracted from the video footage and utilised as input for a K-NN regression using dimensional label. Using MATLAB Simulink, the accuracy of system reached 51.2%, while it was 47.4% in the Xilinx simulation.

With the aid of an automated gaze system, Bidwell and Fuchs [3] evaluated students' engagement. Using recorded video from classrooms, they created a classifier for measuring student attention. To gather the students' focus, they employed a face tracking device. For the purpose of training an HMM, the automatic gaze model that was produced and the patterns generated by the observations of a panel of experts were compared. They attempted to create seven discrete behaviour categories using HMM, but were only able to categorise whether a pupil is "not engaged" or "engaged." Students' concentration is checked via eye and head movements by Krithika [32], who also generates an alarm for low concentration. The video was split up into frames before being subjected to analysis.

In an e-learning environment, Sharma et al. [49] introduced a real-time system based on students' expressed facial emotions during a lesson to assess their level of concentration. The system automatically adjusts the course equipment based on the student level of concentration by trying to analyse the student's emotional responses. To get the final concentration index, the emotions are analyzed. The findings demonstrated a relationship between the students' reported emotions and their levels of concentration, and they developed three unique levels of concentration (high, medium, and low). In their proposal [17], Guojon Yang et al. used vectorized facial characteristics to create a DNN model. Since human facial expressions are represented as vectors, highly accurate DNN training is possible. Observations show that CNN is the most effective advanced machine learning technique in terms of automatic extraction of feature, little input, and accuracy.

emotion recognition is done on a static image, but it might be difficult to identify emotions from facial expressions in videos. According to Zhang et al. [48], a hybrid model that combines two CNN models, a DBN model, and other models is particularly effective in extracting facial expression from moving video. Marsh [20] uses a self-recorded speech recognition system to automate instructor feedback while the lecturer is giving lectures. By giving instructors unbiased feedback for development, this strategy makes use of the discourse factors used by the instructor and enhances student learning.

Author [45] Presents a real-time student engagement system that offers teachers' individualised support to students who are at risk of disengaging. It aids in enhancing the teacher's classroom procedures and allocating attention to pupils who need it most. Rani et al. [26] proposed a classification theory of educational goals.
Methodology

Dataset

In order to infer academic emotions, an online learning spontaneous facial expression database (OL-SFED) is constructed in this research. All database users have signed a consent form authorising the usage of their facial photographs for research. We want the OL-SFED to be accessible to everyone. Research on educational emotional computing will benefit from it, particularly when examining the applicability of educational emotion inference in e-learning.

The OL-SFED contains both video and still pictures of natural face expressions. The experiment involved 82 healthy Ocean University of China students between the ages of 17 and 26 (mean age = 20.09, standard deviation = 2.26) who willingly agreed to take part. There were 29 males and 53 females in all. All volunteers signed informed consent forms after being made conscious of their ability to leave from the learning at any occasion. Only those participants’ videos and pictures are included in the database that authorised the usage of their facial image for study. The study setting, which is depicted in Fig. 3, is designed to mimic a genuine online learning environment. Researchers in another room use the Internet to oversee the entire procedure, which takes place in a learning hall. A working computer is set up to play online course in the study area. Above the screen is a webcam that records video at a frame rate of 30 frames per second at a resolution of 1280 by 720. Throughout the experiment, participant recordings will be stored.

The standard facial expression database was used to train the facial expression recognition algorithm, which was used to identify faces and categorise expressions of face into anger, disgust, fear, happiness, sad, surprise, contempt. Each image has a label for the various emotions. The data set, which consists of 48 by 48 pixel grayscale photographs of faces of human with each one labelled with one of seven expressions, is used to train the network.

![Figure 3: setup of the recording system. While the investigator observes the participant facial expression remotely from another room, the participants view the online course by themselves in the study hall.](image)

Data Processing Techniques

![Figure 4: Structure of Facial Expression Detection](image)
Figure 4 shows the overall layout of the suggested facial expression recognition system. Our FER methodology is broken down into three basic phases: (1) taking photographs of students’ faces from their learning devices; (2) extracting features of facial expressions; and (3) categorising the characteristics into expressions of faces to assess learning outcomes. The method of facial expression emotion analysis and recognition is described as follows, according to the discussions.

**Image Acquisition**

The initial stage of emotion recognition is image acquisition, when subjects are evaluated independently in a calm environment using a setup consistent with the experimental technique. The participants are then asked to respond to the specific questions; during this time, their reactions are successfully assessed and images of their facial expressions are taken for use in a later recognition procedure.

Because colour photographs in this situation are difficult to handle, the acquired image must be turned into a grayscale version. The transformed grayscale image has binary values (0 and 1) that the automatic system can process quickly. Utilizing computer technology, the face has been identified after converting the facial expression photographs. The psychological process of face detection is when a person’s face is recognised in accordance with visual settings. Additionally, the frontal human face is the main focus of the method used to detect the face according to object class. The photographs are inspected bit by bit to see if they match the images that are recorded in the database throughout the face detection process.

**Pre-processing**

We employ a variety of image processing techniques for pre-processing photos, such as face detection and cropping, to acquire the correct areas containing faces and subsequently improve image quality. In real-world applications, input photos are typically captured from webcams or cameras and include both the student’s image and any object-filled backgrounds. As a result, we must use face detection techniques before removing the backdrop from the photos. The phrase “pre-processing” refers to input images with minimal levels of abstraction. It is crucial for feature extraction and normalising since it helps make data more interpretable or perceptible and gets rid of distortions. Therefore, this step improves the image and gets it ready for additional image processing. The following pre-processing methods were applied in this project. The pre-processing methods applied to the raw image samples were follows: wavelet transformation, colour conversion reduction of noise, normalisation, and feature engineering.

**Feature Engineering**

This method of picture augmentation reduces overfitting in the architecture by increasing the quantity of input samples for training and testing.

**Normalization**

A crucial technique for lowering feature variability while keeping features’ salient strength is feature normalisation. The process of changing the sizes of intensity values is known as contrast stretching. Histogram normalisation is the most typical kind of normalisation. Normalization, though, can happen at various phases.

**Noise Reduction**

Images are taken and stored in real-world circumstances in actual or virtual locations under various lighting and illumination conditions. Noise reduction strategies must be used in order to decrease or eliminate the noise because this issue seriously affects the recognition accuracy rate.
Binarization
It is the method used to achieve gray-level modification. This method's excellent temporal complexity and ability to easily extract features make it a vital component of image analysis.

Feature Extraction
A pre-processing technique known as feature extraction can be thought of as removing distracting variation from a dataset in order to improve the performance of classifiers or regression estimators that come later. It is impossible to clearly define where feature extraction stops and classification, or regression, starts. Geometric and visual extraction techniques are the two categories. Geometric approaches are ways to draw forth unique geometric details from pictures. The characteristics of an object that are made up of geometric elements like points, lines, curves, or surfaces are known as geometric features. In visual-based approaches, a feature vector is extracted from a face image by applying image filters to either the entire face or just a few selected areas. In this project, both methodologies were used to derive the components.

Facial Expression Classification
In the FER job, feature categorization has been completed. The feature classification phase begins with the outcomes of the feature extraction step as its input. Salient features from the trained data are preserved in the algorithm's learned features, which are then tested on our algorithm to determine how well it recognizes objects. The outcomes are used to gauge the effectiveness of the algorithm. The usage of an effective classification algorithm is the next crucial step after the dataset has been prepared utilizing the necessary attributes. Nearly all instances of multi-class categorization of human expressions use SVM and Random forest.

Support Vector Machine (SVM)
It is a classification technique that uses supervised learning. By employing a hyperplane, it distinguishes between the classes. The approach generates an ideal hyperplane that categorises the classes exactly. This approach has drawn particular interest from academics due to its strong compatibility to be used in machine learning applications of big amounts data, such as computer vision and pattern recognition [1]. It tries to construct the best hyperplane-like margins while being heavily employed. It performs the task of maximizing the separation among the closest training data samples and the hyper plane. Numerous studies have confirmed that the ideal hyperplane would offer greater accuracy with all types of datasets in a linearly dividable situation [52]. To effectively make decisions on a specific topic, the algorithm aids in the analysis of the data gathering [5]. During this procedure, the features or data that were collected were used to develop an efficient model that could accurately predict the data.

Random Forest (RF)
It is a supervised computer algorithm, that operates by constructing and amalgamating multiple decision trees to create a forest-like structure. This approach is commonly employed to tackle classification and regression problems, serving as a machine-learning tool for predicting new data based on historical datasets. By using this technique, operators can benefit from applying various learning models to enhance accuracy to a new level. The fact that root nodes are redundantly linked in this approach sets it apart from other learning machines [57]. A decision tree ensemble learning technique called an RF classifier consists of many decision trees, each of which is randomly built using bootstrap aggregation or bagging during training [42]. An RF is known as a classifier that is robust to over-fitting and produces superior performance than SVM or AdaBoost-based methods because it is based on randomizing techniques with regards to subset and feature selection while
building the trees [30, 31]. In some instances, RF has also shown to be superior to SVM [23].

**Performance Metrics**

The performance parameters used to evaluate the facial expression emotion classifiers are detailed below. Accuracy, recall/specificity, precision, and F1-score metrics obtained from confusion matrix were utilised to assess the performance of the models, and the formulations of these metrics are as follows:

**Precision**

\[
\text{Precision} = \frac{\text{True positive}}{\text{True positive} + \text{False negative}}
\]

**Recall**

\[
\text{Recall} = \frac{\text{True negative}}{\text{True negative} + \text{False positive}}
\]

**F1-score**

\[
\text{F1-score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}
\]

**Accuracy**

\[
\text{Accuracy} = \frac{\text{number of true positive} + \text{number of true negative}}{\text{number of true positive} + \text{false positive} + \text{false negative} + \text{true negative}}
\]

**Result and Discussions**

The research is conducted on a workstation using The system has a 12-core Intel Core i7-6850K processor clocked at 3.60 GHz and a single 2760 MHz NVIDIA GeForce GTX 1080 Ti GPU. Python 3.6/3.7, the Tensor Flow 2.0 framework, and Windows 10 are all installed on the deep learning workstation platform.

**SVM Model**

The created confusion matrix is used to identify emotions from the face points that have been detected. This confusion matrix is evaluated along with the system's efficiency. Using the SVM training and testing process, several facial emotions are predicted from the identified facial points. The confusion matrix, which is illustrated in fig. 4, illustrate the potential outcomes of this procedure, which calculates the genuine emotion anticipated information in relation to the testing label.

![Figure 4: Support Vector Machine Confusion Matrix](image)

These values were acquired through the training and testing processes for each emotion because only a small number of emotions are interdependent. Therefore, the confusion matrix is created in an efficient method in accordance with the specific expected emotion, and they subsequently received the dataset's validated emotions.

The obtained accuracy is 54.67% overall. The SVM classifier's various expression classification parameters are shown in Table 2.

<table>
<thead>
<tr>
<th>Expression</th>
<th>Precision</th>
<th>Recall</th>
<th>F1-Score</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anger</td>
<td>0.35</td>
<td>0.22</td>
<td>0.27</td>
<td>34.7</td>
</tr>
<tr>
<td>Contempt</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Disgust</td>
<td>0.54</td>
<td>0.48</td>
<td>0.51</td>
<td>54</td>
</tr>
<tr>
<td>Fear</td>
<td>0.16</td>
<td>1.00</td>
<td>0.28</td>
<td>16</td>
</tr>
<tr>
<td>Happy</td>
<td>1.00</td>
<td>0.85</td>
<td>0.92</td>
<td>100</td>
</tr>
<tr>
<td>Sadness</td>
<td>0.38</td>
<td>0.23</td>
<td>0.29</td>
<td>37.5</td>
</tr>
<tr>
<td>Surprise</td>
<td>1.00</td>
<td>0.88</td>
<td>0.94</td>
<td>100</td>
</tr>
</tbody>
</table>
Table 2: Expression Classification parameters for SVM classifier

<table>
<thead>
<tr>
<th></th>
<th>Macro average</th>
<th>Weighted average</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.49</td>
<td>0.65</td>
</tr>
<tr>
<td></td>
<td>0.52</td>
<td>0.55</td>
</tr>
<tr>
<td></td>
<td>0.46</td>
<td>0.58</td>
</tr>
</tbody>
</table>

Figure 5 shows graphical representation of various expressions while using SVM classifier.

The primary explanation of the decreased accuracy for facial expressions was that they entail identical facial muscle movements or that multiple significant local elements were lost due to incorrect landmark localization. The impact of lighting variations on the image frames and image deterioration over time due to the presence of Gaussian noise, which leads in the camouflaging of salient features during feature extraction, lead to less accurate results. Similar to this, the dataset's high affinity of inter-class differences contributed to the misclassification issue. Last but not least, the choice of image size may have stretched and altered the geometry of the features, resulting in a lack of parameter characteristics for categorization.

**RF Model**

Figure 6 shows that fear and happiness performed best, whereas anger and disgust performed worst. The explanation for this is that while the facial expressions of anger and disgust in the database are nearly identical, the shift in facial expression caused by fear is rather big. The overall level of accuracy is 90.14%. The expression classification parameters for the RF classifier are shown in Table 3.

![Random forest confusion matrix](image)

<table>
<thead>
<tr>
<th>Expression</th>
<th>Precision</th>
<th>Recall</th>
<th>F1-Score</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anger</td>
<td>0.91</td>
<td>0.72</td>
<td>0.81</td>
<td>91.3</td>
</tr>
<tr>
<td>Contempt</td>
<td>0.75</td>
<td>0.90</td>
<td>0.82</td>
<td>75</td>
</tr>
<tr>
<td>Disgust</td>
<td>1.00</td>
<td>0.97</td>
<td>0.99</td>
<td>100</td>
</tr>
<tr>
<td>Fear</td>
<td>0.76</td>
<td>1.00</td>
<td>0.86</td>
<td>76</td>
</tr>
<tr>
<td>Happy</td>
<td>1.00</td>
<td>0.94</td>
<td>0.97</td>
<td>100</td>
</tr>
<tr>
<td>Sadness</td>
<td>0.75</td>
<td>1.00</td>
<td>0.86</td>
<td>75</td>
</tr>
<tr>
<td>Surprise</td>
<td>1.00</td>
<td>0.84</td>
<td>0.91</td>
<td>100</td>
</tr>
</tbody>
</table>
Table 3: Expression classification parameters for RF classifier

The final results obtained using the feature selection method applied to seven different types of emotional expressions on people’s face show that RF is the most perfect classifier. Therefore, Random Forest is the most effective classifier when it comes to recognising facial expressions.

Figure 7 shows graphical representation of various expressions while using RF classifier. Figure 7: RF Graphical representation of expression

Figure 8: show some sample predictions of both models with random images.

Figure 8 (a): Disgust
Figure 8 (b): Surprise
Figure 8 (c): Anger
Figure 8 (d): Happy

Conclusion

Academic emotions have a significant influence on learning because they are closely related to motivation, controllability, cognition, and other factors. Online learning will undoubtedly benefit from the replacement of the lost emotional interactions through the application of academic emotion inference. A trustworthy facial expression database is
required in order to develop an efficient facial expression-based inference method. This study establishes OL-SFED, a freely accessible database of spontaneous facial expressions in an online learning environment. It covers the natural facial expression made in response to typical academic emotions, and the samples are thoroughly annotated by the participant and outside coders. A system for feedback tracking using facial expression recognition was created and tested in this study. The framework used facial expression recognition to determine a student's emotional state and sent the appropriate feedback to the tutor or teacher so they would know how the student felt about a lesson or discussion. The FER Dataset underwent analysis, preprocessing, facial feature extraction, and input into the SVM and RF models. The output was evaluated, and example photos were then categorised into the many types of emotions, including Sad, Happy, disdain, Fear, Disgust, surprise, and Angry. In order to increase utilisation and acceptability of online teaching platforms, the simulation results of were viable and shown the potential of building a mechanism for feedback monitoring. This mechanism achieved an accuracy of 90.14% using RF classifier.

References


Enhancing Online Teaching and Learning Through Machine Learning and Learning Analytics


Krithika, L.B.: ‘Student emotion recognition system (SERS) for e-learning improvement based on learner concentration metric’, Procedia Computer Science, 2016, 85, pp. 767 –776


Redmond, P. (2011). From face-to-face teaching to online teaching: Pedagogical transitions. In B. Williams, G. Statham, P. Brown, & N. Cleland (Eds.), ASCILITE Hobart 2011 (pp. 1050–1060)


