Turkish Online Journal of Qualitative Inquiry (TOJQI)
Volume 12, Issue 7 July 2021: 11708 – 11719

Student Behaviour Detection in Eductation Training Institution

Gagana S, Sheba Selvam, Priya G, Preethi H, Seema D
Department of Computer Science and Engineering, BNM Institute of Technology, Bangalore, India.

Abstract-
Student engagement has been a key topic inside the educational training. The three specific styles of engagement of the students in a class are: behavioural, emotional, and cognitive. The time period behavioural engagement is commonly used to describe the scholar's willingness to participate within the getting to know system. Emotional engagement describes a scholar's emotional attitude towards learning. Cognitive engagement is a chief part of overall learning engagement. From the facial expressions the involvement of the students in the magnificence can be decided. Commonly in a lecture room it's far difficult to recognize whether the students are able to understand the lecture or no longer or even whether there is any kind of stress. So that you can know that comments form will be collected manually from the students. However, those feedbacks given by using the students will now not be correct. Hence, they will no longer get proper comments. This hassle can be solved by means of the use of a facial emotion evaluation. From the facial expression the emotion of the students may be analysed. Quantitative observations are achieved in the lecture room wherein the emotion of students might be recorded and statistically analysed. With the aid of the use of facial emotion we will directly get correct information approximately college students understand potential, and determining if the lecture become exciting, boring, or mild for the students. And the apprehend capability of the scholar is recognized by the facial emotions. Also using face data attendance process is made automatic.

Keywords- Emotion, facial expression, student, attendance.

I. INTRODUCTION

The major component of human communication is facial expressions. Facial expressions are used not only to express our emotions, but also to provide important communicative cues during social interaction, such as our level of interest, continuous feedback signalling understanding of the information conveyed. Real time system to recognition the basic emotion from video. The system automatically detects frontal faces in the video stream and codes each frame with respect to 7 dimensions: Neutral, anger, disgust, fear, joy, sadness, surprise. The facial features from the video source is extracted and mapped with the basic emotions. This can be used in any kind of environment. We use this system classroom environment to identify the involvement of students during lecture.

Facial expression is one or more motions or positions of the muscles beneath the skin of the face. Facial expressions are form of nonverbal communication. The important fact that felt emotions is only one source of facial expressions besides others like verbal and non- verbal communication or physiological activities. We use facial expressions to display our emotional states and to manage our interactions. The term “facial expression recognition” often refers to the classification of facial features in one of the six so called basic emotions: happiness, sadness, fear, disgust, surprise, anger and neutral. By these expressions, we can examine the behaviour of the student during lecture. The involvement of students can be identified by their facial emotions. Face speaks more than words, it’s better to convey messages. Non- verbal communication in classroom is more important, as students express their situation or feelings by facial emotion or expression.

II. LITERATURE SURVEY

Aamir et al. [3] made a study and observed 32 studies out of a total of 7,939 reviews searched for the purpose. This study reviews a multitude of factors such as the role of an individual's personal, socio-cognitive, psychological and environmental factors towards cyberbullying and provides a 360-degree view of the factors contributing to cyberbullying behavior instead of the traditional approach of focusing on one or two factors. Tolga Sayata et al. [1] proposed that system is capable of making real-time suggestions to an in-class presenter to improve the quality and memorability of their presentation by allowing the presenter to make real-time adjustments/corrections to their non-verbal behaviour, such as hand gestures, facial expressions, and body language. We base our suggested system components on existing research in affect sensing, deep learning-based emotion recognition, and real-time mobile-cloud computing. We provide a comprehensive study of these technologies and determine the computational requirements of a system that incorporates these technologies. Based on these requirements, we provide a feasibility study of the system. Although the state-of-the-art research in most of the components we propose in our system are advanced enough to realize the system, the main challenge lies in: 1) the integration of these technologies into a holistic system design; 2) their algorithmic adaptation to allow real-time execution; and 3) quantification of valid educational variables for use in algorithms. In this paper, we discuss current issues and provide future directions in engineering and education disciplines to deploy the proposed system. Zhen et al. [2] trained a network using both HR and LR images under the guidance of a fixed network, pretrained on HR face images. The guidance is provided by minimizing the KL-divergence between the output SoftMax probabilities of the pretrained (i.e., Teacher) and trainable (i.e., Student) network as well as by sharing the SoftMax weights between the two networks. The resulting solution is tested on down-sampled images from Face Scrub and Mega Face datasets and shows a consistent performance improvement across various resolutions. We also tested our proposed solution on standard LR benchmarks such as Tiny Face and SC Face. Our algorithm consistently outperforms the state-of-the-art methods on these datasets, confirming the effectiveness and merits of the proposed method.

Wang et al. [4], proposed an online method on modelling the position-behaviour feature of many people scene is presented. First, a position-based model of individual behavioural feature transformation is proposed. The behaviour features in other positions are generated by the behaviour of the individual in a particular position. Second, a many people behaviour feature generation method based on noise reduction is proposed to generate the same behaviour feature in an online pattern. Finally, taking advantage of the models, a many people fitness coaching system is designed and implemented, named multiuser fitness coach. The system can identify the irregular behaviour of individuals in the many people environment. The performance of the system is evaluated in different scenarios, and the results show that the precision of feature generation can be effectively applied to the decision of irregular behaviour in many people scenarios.

Mark John et al. [5] addresses both of these issues by examining teacher-student learning within a sequence-level framework, and assessing the flexibility that these approaches offer. Various sequence-level teacher- student criteria are examined in this work, to propagate sequence posterior information. A training criterion based on the Kullback-Leibler (KL)-divergence between context- dependent state sequence posteriors is proposed that allows for a diversity of state cluster sets to be present in the ensemble. This criterion is shown to be an upper bound to a more general KL-divergence between word sequence posteriors, which places even fewer restrictions on the ensemble diversity, but whose gradient can be expensive to compute. These methods are evaluated on the augmented multi-party interaction (AMI) meeting transcription and MGB-3 television broadcast audio tasks. Rafik et al. [6] discussed that classroom attendance check is a contributing factor to student participation and the final success in the courses Tests were done with both iOS and ANDROID. Forty different attendance monitoring tests were performed in a real classroom, including 11 students, and 264 students faces were detected. Subhrobhattacharya et al. [7] suggests a significant role in the assessment and quality monitoring. An automatic attendance management system aims at solving the issues of manual methods of existing systems. We have used the concept of face recognition to implement a system that marks the attendance of a particular person by detecting and recognizing the face. Individual student's picture is stored in the database and using MATLAB software GUI is created for each student by Marko et al. [8]. When the program is run a tick will be sounded which indicates the user to hold student's photo. After certain delay the image captured is the test image. Again, on one more tick, the user must hold the internal marks of a student and it is captured in the image format.

Attendance automation using face recognition has become one of the most important needs in educational institutional and work places across the world. Since its saves time and accurate too Shwetha et al. [9]. Manisha et al. [10] starts with preprocessing of face images using Haar classifier followed by facial feature extraction using LBP algorithm and then classification is done by training and testing of dataset using support vector machine classifier. The basic emotion stated by using MTCNN and Keras, MB-LBP operator is defined by comparing the central rectangles average intensity with those of its neighborhood rectangles by Neha et al. [11]. Here, sadness places a critical role. Sheng et al. [12] proposed Haar cascade, LBPH algorithms are used efficiently. Classification is done by training and testing of dataset using support vector machine classifier. Putra et al. [13] used CNN to detect emotions, the system performs Euclidean distance computation It compares original dataset with normalized shape feature benchmark. The dataset will be tested on canonical appearances benchmark, finally, the dataset will be compared with both benchmarks for the validation of dataset.

Humans use facial expressions successfully for conveying their emotional states. However, replicating such success in the human-computer interaction domain is an active research problem Oyebade et al. [14]. In this paper, we propose deep convolutional neural network (DCNN) for joint learning of robust facial expression features from fused RGB and depth map latent representations. We posit that learning jointly from both modalities result in a more robust classifier for facial expression recognition (FER) as opposed to learning from either of the modalities independently. A FER system was proposed based on the combination of Bayesian Belief Net (BBN) and statistical facial features model by Tang et al. [15]. The work described an approach where manual landmarking for facial features extraction is eliminated. For feature-based FER systems, learning highly discriminative and robust facial expression is very critical. This approach is challenging as the features that discriminate the different facial expression are quite subtle, Tian et al. [16].

The segmented facial surfaces into expressive regions and applied histogram statistics on the segmented regions. Finally, the features obtained from the histogram statistics were used to train a linear discriminant classifier for FER Sun et al. [17]. Happy et al. [18] proposes an elaborate FER system was proposed composing low pass filtering, eye, nose, lip corner and eyebrow corner detection; features were extracted from detected regions of interests using local binary pattern (LBP) encoding. Later, support vector machine (SVM) based classifier is trained on a feature space estimated with Principal Component Analysis (PCA) from LBP encoding. Several studies were performed on facial expression recognition (FER). Many of these works essentially rely on two stages of information processing: (1) features extraction Zhang et al. [21] (2) classification of extracted features Zhang et al. [19][20]. The already existing system focuses on either attendance generation or emotion generation but not both.

III. PROPOSED METHODOLOGY

In the proposed work, face image and facial landmark detection is performed first for stress recognition using Convolution Neural Network (CNN) algorithm. In the proposed network, the face images and expression detected earlier are inputted to output stress recognition results. The results of face recognition are composed of students present in the class.

Step 1: Convolution

A convolution is a joined integration of two methods that demonstrates to you how one method changes the other.

Step 2: Apply the RLU (Rectified Linear Unit)

In this step, the corrective function is used to increase nonlinearity on CNN. The data set is made up of different objects which are not linear to one another. Under this function, the grouping of information can be seen as a linear problem, although it is a non-straight problem.

Step 3: Pooling

Spatial invariance is a term that does not influence the neural network's ability to detect its particular feats when finding an item in the data collection. Pooling helps CNN to detect swimming pools, such as max and min pools, for example.

Face Recognition based Attendance:

The total system is divided into 3 modules- Database creation, Training the dataset, Testing, sending alert messages as an extension.

1. Database creation
   a) Initialize the camera and set an alert message to grab the attention of the students.
   b) Get user id as input
   c) convert the image into grey scale, detect the face and
   d) Store it in database by using given input as label up to 20 frames.
2. Training
   a) Initialize LBPH face recognizer.
   b) Get faces and Id's from database folder to train the LBPH face recognizer.
   c) Save the trained data as xml or yml file.

3.Testing

Load Haar classifier, LBPH face recognizer and trained data from xml or yml file.
   a) Capture the image from camera,
   b) Convert it into grey scale,
   c)Detect the face in it and
   d)Predict the face using the above recognizer.

This proposed system as shown in fig.1, uses Haar cascade algorithm for face detection which uses modified Haar Cascades for detection. Raspberry Pi is the main component in the project. We will be using USB webcam to capture photos. We can access System console either by using SSH in laptop. Firstly, the algorithm needs a lot of positive images and negative images to train the Haar cascades classifier. Positive images are images with clear faces where negative images are those without any faces.

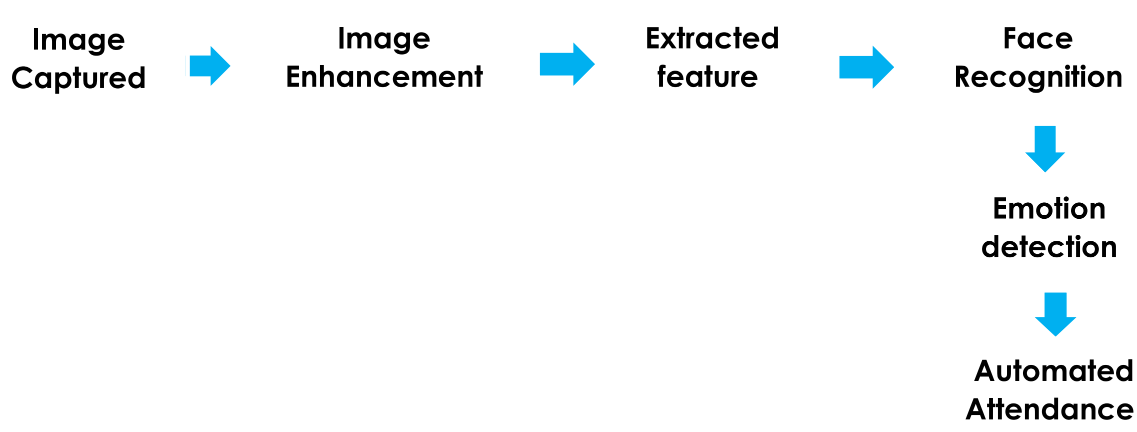

Fig 1: Proposed Model

Algorithm Used:

1. Haar-Cascade Algorithm

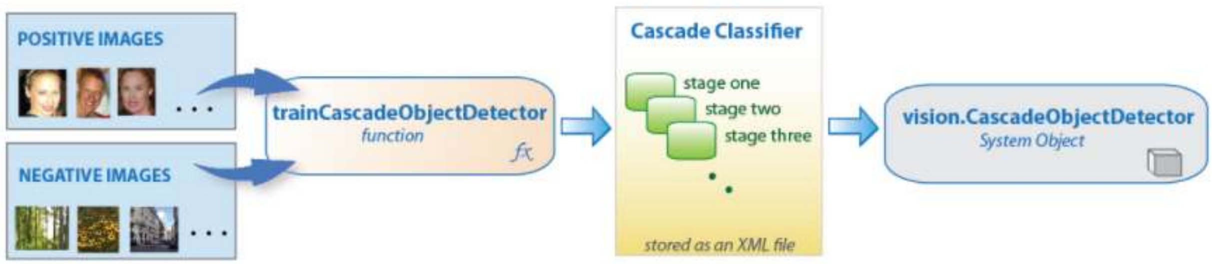

Fig 2: Haar Cascade algorithm

Haar Cascade is a machine learning object detection algorithm used to identify objects in an image or video. It is a machine learning based approach where a cascade function is trained from a lot of positive and negative images as shown in fig.2. It is then used to detect objects in other images. It is well known for being able to identify almost any object. First step is to collect the Haar Features. A Haar feature considers adjacent rectangular regions at a specific location in a detection window, sums up the pixel intensities in each region and calculates the difference between these sums.

2. CNN (Convolutional neural Network)

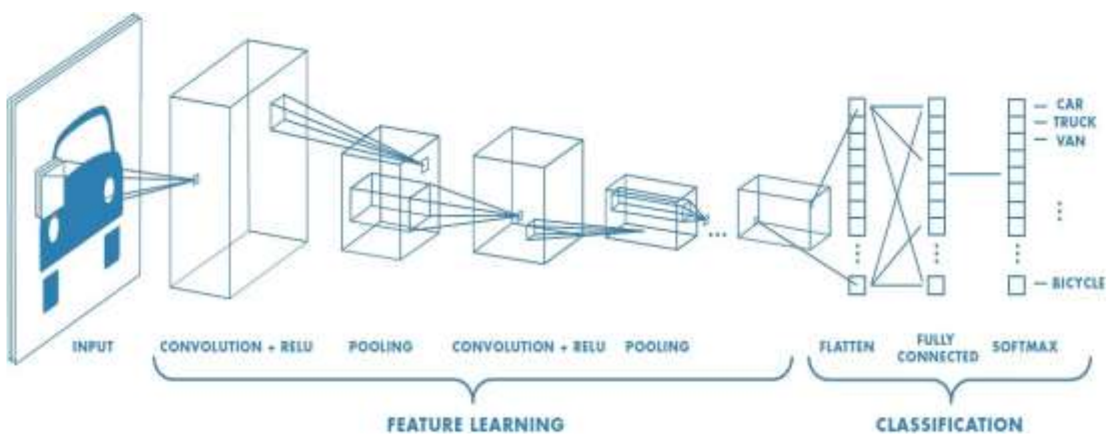

Fig 3: Working of Convolutional Neural Network

In the proposed work CNN classifier is used as shown in fig.3. The classifier is utilized for the classification of the Image Fusion and Recursive Filtering features. A convolution multiplies a matrix of pixels with a filter matrix or 'kernel' and sums up the multiplication values. Then the convolution slides over to the next pixel and repeats the same process until all the image pixels have been covered.

i. **Loading the Dataset**

We will plot the first image in our dataset and check its size. By default, the shape of every image in the minimum dataset is 28 x 28, so we will not need to check the shape of all the images. When using real-world datasets, you may not be so lucky. 28x 28 is also a fairly small size, so the CNN will be able to run over each image pretty quickly.

ii. Data pre-processing

We need to 'one-hot-encode' our target variable. This means that a column will be created for each output category and a binary variable is inputted for each category. Our first layer also takes in an input shape. This is the shape of each input image, 28,28,1 as seen earlier on, with the 1 signifying that the images are greyscale.

iii. Compiling the model

Next, we need to compile our model. Compiling the model takes three parameters: optimizer, loss and metrics. The optimizer controls the learning rate. We will be using 'Adam' as our optimizer. Adam is generally a good optimizer to use for many cases. The Adam optimizer adjusts the learning rate throughout training. After loading of the hyperspectral images the hyperspectral images are spectrally partitioned into 4th subsets of adjacent bands that is, Band 1, Band 2, Band 3,Band 4.After fusing hyperspectral image by using Transformation and Principle Component Analysis Image Fusion Techniques then I used Transform domain Recursive Filtering for filtering the fused image.

iv. Training the model

Now we will train our model. To train, we will use the 'fit()' function on our model with the following parameters: training data (train_X), target data (train_y), validation data, and the number of epochs. For our validation data, we will use the test set provided to us in our dataset, which we have split into X_test and y_test. Now we will be using our model to make predictions.

3. Facial Landmark

Face landmark detection is the process of finding points of interest in an image of a human face. For example, we have shown the ability to detect emotion through facial gestures, estimating gaze direction, changing facial appearance (**face swap**), augmenting faces with graphics, and puppeteering of virtual characters.

To achieve this, the landmark detector must find dozens of points on the face, such as corners of the mouth, corners of eyes, the silhouette of the jaws, and many more. Many algorithms were developed and implemented in OpenCV. To run the face-mark detector, a pre-trained model is required. This pre-trained model which we have used is shape_predictor_68_face_landmarks. The indexes of the 68 coordinates can be visualized on the image below.

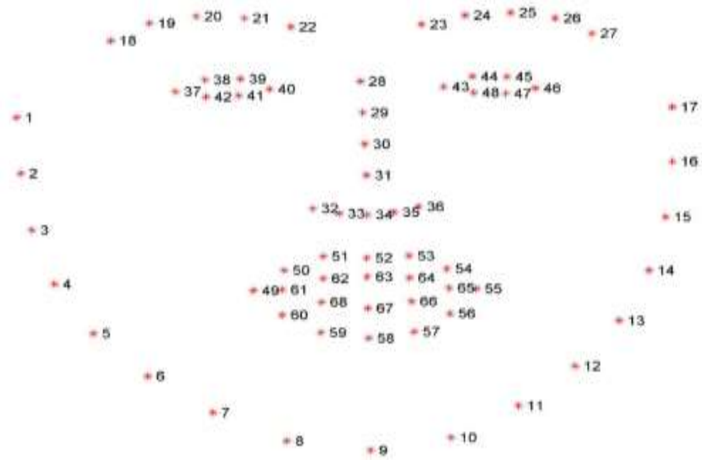

Fig 4: Visualizing the 68 facial landmark co-ordinates

IV. DATASET

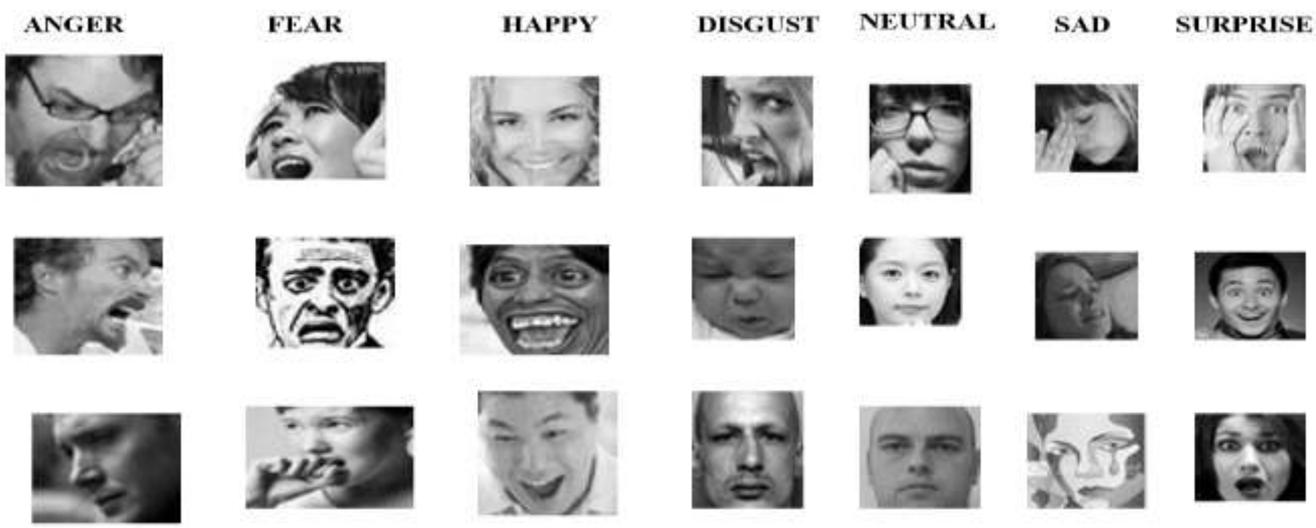

Fig 5: Dataset

The data which we get from the extraction of facial features has very high dimensionality therefore classification is used to reduce the data dimensionality. This process is done by CNN that is convolution neural network. CNN produces high classification accuracy even with the availability of moderate level of training data.

V. PERFORMANCE EVALUATION

Accuracy

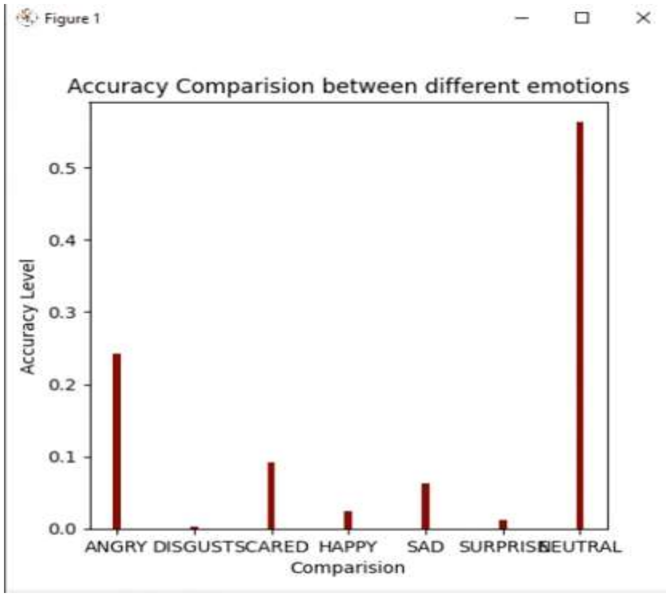

Fig 6: accuracy comparison between different emotions

```
w = int(prob * 300)
cv2.rectangle(canvas, (7, (i * 35) + 5),
(w, (i * 35) + 35), (0, 0, 255), -1)
cv2.putText(canvas, text, (10, (i * 35) + 23),
cv2.FONT_HERSHEY_SIMPLEX, 0.45,
(255, 255, 255), 2)
cv2.putText(frameClone, label, (fX, fY - 10),
cv2.FONT_HERSHEY_SIMPLEX, 0.45, (0, 0, 255), 2)
cv2.rectangle(frameClone, (fX, fY), (fX + fW, fY + fH),
              (0, 0, 255), 2)
dic = {'ANGRY': preds[0],'DISGUST' :preds[1] ,'SCARED' :preds[2] ,'HAPPY':preds[3] ,'SAD':preds[4],'SURPRISE':preds[5],'NEUTRAL':preds[6]}

algm = list(dic.keys())
accu = list(dic.values())

fig = plt.figure(figsize = (5, 5))

# creating the bar plot
plt.bar(algm, accu, color ='maroon',
        width = 0.1)

plt.xlabel("Comparision")
plt.ylabel("Accuracy Level")
plt.title("Accuracy Comparision between different emotions")
plt.show()
break
```

Fig 7: calculation for accuracy

VI. RESULTS

This application shows the high performance of classifier and feature extraction method that enhances the efficiency of system and improved the accuracy of facial emotion recognition. In this 7 universal emotion from different static images is analysed. This paper shows pre-processing of face images using Haar classifier followed by facial feature extraction using LBPH algorithm. This system is 86 – 90 percent more efficient when compared to existing system.

Fig 8: neutral

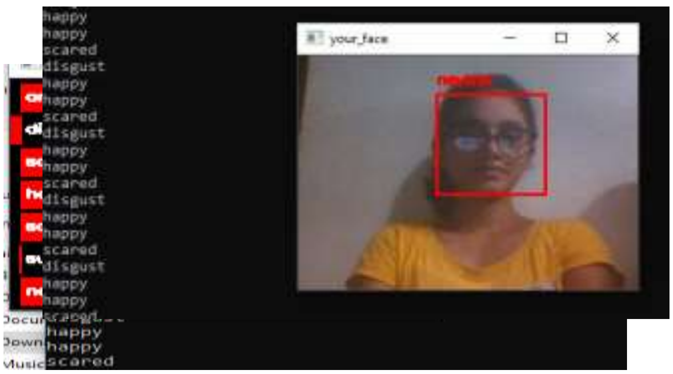

Fig 9: sad

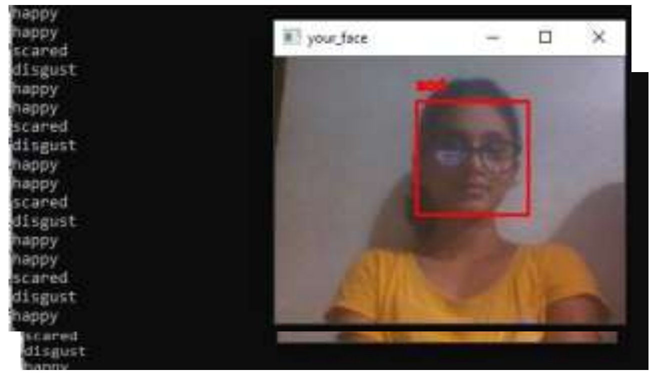

Fig 10: emotion probability

Fig 11: happy

VII. CONCLUSION

We propose a student attendance and stress recognition algorithm using face detection. As a result of the experiment, we confirmed that the attendance of the student and stress recognition performance was further improved when using facial detection. Facial landmarks are better at perceiving stress because they allow you to better understand eye, mouth, and head movements. We also found that the performance was improved by better identifying stress- related information when using a grey face image of the appropriate size and also finding out the unknown data and sending it to authorized person.

REFERENCES

[1] Kim, Y., Soyata, T. and Behnagh, R.F., 2018. Towards emotionally aware AI smart classroom: Current issues and directions for engineering and education. *IEEE Access*, *6*, pp.5308-5331.

[2]. Khalid, S.S., Awais, M., Feng, Z.H., Chan, C.H., Farooq, A., Akbari, A. and Kittler, J., 2020. Resolution invariant face recognition using a distillation approach. *IEEE Transactions on Biometrics, Behavior, and Identity Science*, *2*(4), pp.410-420.

[3]. Shaikh, F.B., Rehman, M. and Amin, A., 2020. Cyberbullying: a systematic literature review to identify the factors impelling university students towards cyberbullying. *IEEE Access*, *8*, pp.148031-

148051.

[4]. Wang, P., Dong, L., Liu, W. and Jing, N., 2020. Clustering-based emotion recognition micro-service cloud framework for mobile computing. *IEEE Access*, *8*, pp.49695-49704.

[5]. Wong, J.H.M., Gales, M.J.F. and Wang, Y., 2019. General Sequence Teacher–Student Learning. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, *27*(11), pp.1725-1736.

[6]. Samet, R. and Tanriverdi, M., 2017, September. Face recognition-based mobile automatic classroom attendance management system. In *2017 International conference on cyberworlds (CW)* (pp. 253-256). IEEE.

[7]. *Bhattacharya, S., Nainala, G.S., Das, P. and Routray, A., 2018, July. Smart attendance monitoring system (SAMS): a face recognition based attendance system for classroom environment. In* 2018 IEEE 18th International Conference on Advanced Learning Technologies (ICALT) *(pp. 358-360). IEEE.*

[8]. Arsenovic, M., Sladojevic, S., Anderla, A. and Stefanovic, D., 2017, September. FaceTime—Deep learning based face recognition attendance system. In *2017 IEEE 15th International Symposium on Intelligent Systems and Informatics (SISY)* (pp. 000053-000058). IEEE.

[9]. Soniya, V., Sri, R.S., Titty, K.S., Ramakrishnan, R. and Sivakumar, S., 2017, March. Attendance automation using face recognition biometric authentication. In *2017 International Conference on Power and Embedded Drive Control (ICPEDC)* (pp. 122-127). IEEE.

[10]. Singh, M., Tuli, H. and Singh, N., 2020. Emotion Recognition System through Facial Expressions Using Machine Learning.

[11]. Udeshi, N., Shah, N., Shah, U. and Correia, S., 2020. DETECTION AND ANALYSIS OF STRESS LEVELS. *IJRAR-International Journal of Research and Analytical Reviews (IJRAR)*, *7*(1), pp.595-599.

[12]. Classroom teaching feedback system based on emotion detection Authors: Sheng Chen 2019

[13]. Chen, S., Dai, J. and Yan, Y., 2019. Classroom teaching feedback system based on emotion detection. In *9th International Conference on Education and Social Science (ICESS 2019). https://doi. org/10.25236/icess*.

[14]. Oyedotun, O.K., Demisse, G., El Rahman Shabayek, A., Aouada, D. and Ottersten, B., 2017. Facial expression recognition via joint deep learning of rgb-depth map latent representations. In *Proceedings of the IEEE international conference on computer vision workshops* (pp. 3161-3168).

[15]. Tang, H. and Huang, T.S., 2008, June. 3D facial expression recognition based on automatically selected features. In *2008 IEEE computer society conference on computer vision and pattern recognition*

workshops (pp. 1-8). IEEE.

[16]. Tian, Y., Kanade, T. and Cohn, J.F., 2011. Facial expression recognition. In *Handbook of face recognition* (pp. 487-519). Springer, London.

[17] Yin, L., Wei, X., Sun, Y., Wang, J. and Rosato, M.J., 2006, April. A 3D facial expression database for facial behavior research. In *7th international conference on automatic face and gesture recognition (FGR06)* (pp. 211-216). IEEE.

[18]. Happy, S.L. and Routray, A., 2014. Automatic facial expression recognition using features of salient facial patches. *IEEE transactions on Affective Computing*, *6*(1), pp.1-12.

[19]. Luo, Y., Wu, C.M. and Zhang, Y., 2013. Facial expression recognition based on fusion feature of PCA and LBP with SVM. *Optik-International Journal for Light and Electron Optics*, *124*(17), pp.2767-2770.

[20]. Owusu, E., Zhan, Y. and Mao, Q.R., 2014. A neural-AdaBoost based facial expression recognition system. *Expert Systems with Applications*, *41*(7), pp.3383-3390.

[21]. Beaudry, O., Roy-Charland, A., Perron, M., Cormier, I. and Tapp, R., 2014. Featural processing in recognition of emotional facial expressions. *Cognition & emotion*, *28*(3), pp.416-432.