Learner's Engagement Analysis for E-Learning Platform

[1]K. Keerthana, [2]D.Pradeep, [3]Dr. B. Vanathi

[1]PG Scholar, [2]Assistant Professor, [3]Professor & Head
Department of Computer Science,
[1,3]SRM Valliammai Engineering College, Kattankulathur, India
Department of Information Technology,
[2]Chrisitan College of Engineering and Technology, Oddanchatram, India

***Abstract*: Adaptive learning is an educational method that utilizes computers as an interactive teaching device. In existing most educational agents do not monitor engagement explicitly, but rather assume engagement and feedback by interaction based on the student's responses to questionnaires, quiz and tasks. Thus, dynamic learner engagement analysis is a first step towards an automated course feedback tool helpful to the instructor for measuring student engagement. In our system, we propose a hybrid architecture system invoking student facial emotion recognition, eye gaze monitoring, head movements identifications based analyzing dynamic student engagement in classroom and towards a specific course at e-learning platforms. Our proposed architecture uses feature extraction algorithms for facial emotion recognition, Haar Cascade for pupil detection and Local Binary Patterns for recognizing head movements. For a machine learning approach and to provide accurate results we propose a Convolutional Neural Network. Experimental results have been implemented using PyCharm IDE.**

Keywords: Video processing, e-learning, emotion recognition, eye tracking, head pose identification

I. INTRODUCTION

Recognizing faces and distinguishing between faces is an inborn ability of humans. By introducing that ability to computers is feasible by means machine learning and computer vision techniques. This can open mountain of applications like distracted driver detection, access and security to personalized devices, human robot interaction for youngsters with autism, entertainment industries, E-Learning feedback, allow payments to be processed without physical cards, enable criminal identification and permit personalized healthcare like patient monitoring and other services. Emotions on an individual's face by using most advanced image processing tools is read using emotion recognition phenomenon. The proposed model tries to recognize emotions based on facial expressions such as happy, anger, disgust, fear, joy, surprise as shown in Fig. 1.

Fig. 1. Different Emotions from fer2013 [9] Dataset

Facial expression detection in images and videos is the most impetuous and ordinary way to identify emotions of an individual. As the cost of education like tuition, fees and living expenses has skyrocketed over the past few decades, prolonged graduation time has become a crucial contributing factor to the ever-growing student graduation. Recent studies show that only 50 of more than 580 public four-year institutions in the United States have on-time graduation rates at or above 50 percent for their full-time students.

To make college cheaper, it's thus important to ensure that several students graduate on time through prior involvement on students whose performance are unlikely to satisfy the graduation criteria of the degree program on time. A critical step towards effective intervention is to develop a system which will continuously keep track of student's attention levels and accurately predict their mood of listening and supply teachers with information on progress and achievements of their students. Thus, teachers can assess the impact of their instructional strategies they use.

II. LITERATURE SURVEY

A. E-Learning

Online teaching and e-learning methodologies have transcended to new levels after the boom of the information technology age. As a result, the quality of education and number of online learners has increased substantially. Still, the present and recent methods of e-learning arise problems which disturbs a student's learning curve due to lack of availability of any direct supervision.

Fabri, M., Moore and D.J., Hobbs [1] discuss Online teaching and e-learning methodologies in their experimental study about Mediating the Expression of Emotion in Educational Collaborative Virtual Environments. This study gives an insight on how the virtual learning platforms have transcended to new levels after the boom of information technology age. As a result, the quality of education and number of online learners has increased substantially. Still, the modernized way of e-learning creates problems that affect a student's learning curve due to unavailability of any direct supervision.

An instructor can provide some insight into student's satisfaction during lectures, therefore student's involvement in class has direct correlation with the professional aptitude of the instructor. [2] Direct supervision makes learning possible as well as aids the student to keep in sync with the course contents and objectives due to instant communication with the instructor at any time during the lecture. High levels of vexations are seen to be affecting students. Due to communication unavailability.

Interestingness of the content and engagement of students in the e-learning course are measured using facial expressions as a tool. Automatic real-time feedback on the subscribed e-learning course content can be taken from learners [3].

B. Emotion Recognition and Eye-tracking for Student Monitoring

Hend et al. [4] mentioned that eye tracking devices issue data recorded to identify an individual's level of concentration and the focus of an individual's attention. From eye position tracking and such indirect measures as fixation numbers and duration, gaze position, and blink rate, information can be drawn about user levels of attention, stress, relaxation, problem solving, learning success, and fatigue.

Ismail and Mohamed [5] integrated eye tracking technology to measure and analyze learner behaviors on an E-Learning platform. This study is fixated on user emotion attention, stress, relaxation, problem solving, and fatigue from the compulsive parts reflected by the courses.

Pushpaja V. Saudagare and D.S Chaudhari [6] came forward with a technique to detect expression from emotions through neural networks. This study reviews the numerous methods of facial emotion detection with the use of MATLAB which is a toolbox for neural networks.

Facial expressions can provide critical information on student's interest and participation in online educational learning. Faces provide detailed information about an individual's state of mind, mood and also emotional state [7]. Studies throughout history have shown that facial expressions are the prime representation of human emotions. Facial expressions can be considered as the main source of information, after words, in estimating an individual's thoughts and state of mind.

Facial features (Forehead, eyes, nose, mouth, etc.) [8] are the fundamental attributes that are extensively used in face recognition systems as their movements help determine the construction of expression on a human face.

III. PROPOSED WORK

A. Proposed Methodology

The implementation of the proposed system architecture as shown in Fig. 2 is done in python console with functions for face detection and features detection using OpenCV. OpenCV object detection framework has been used for detecting the face and other facial features.

The following procedure has been proposed for detecting the engagement level mainly based on two categories namely distracted and focused and access id grant according for e-learning course content which are as follows:

To detect the level of concentration, it's to determine whether the face is at the right position i.e. whether the user is properly facing the webcam so that their face can be detected. To determine the face status, the face should be facing the screen.

- Distracted: Learner is looking away from the computer and obviously not listening to the lecture. If the student is not properly facing the screen and the concentration level is given as medium or low hence that user is not engaged so the course URL will not be activated instead documentation link is provided.
- Concentrated: Student could be commended for his/her level of engagement in task. If the user is properly facing the webcam their face can be detected from and classified as engaged actively so the course URL will be activated.

B. Proposed system Architecture

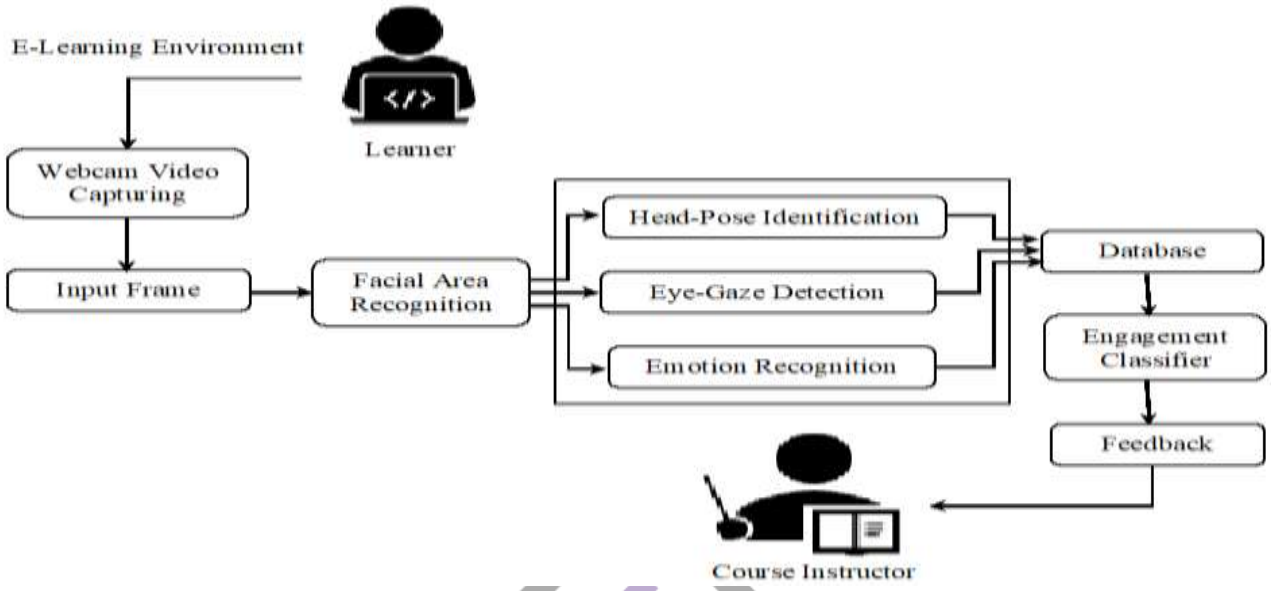


Fig. 2. Proposed System architecture

C. Abbreviations and Acronyms

CNN-Convolutional Neural Network, IDE-Integrated Development Environment.

IV. IMPLEMENTATION

The proposed system is implemented in real time e-learning platform using the OpenCV framework in PyCharm IDE.

A. Pre-Processing

Pre-Processing is an important segment in the complete process. Enhancement of input frame from real-time video and locating region of interest is done in the Pre-Processing stage by smoothing the image and eliminating unwanted segments of frame. Thus, redundant information from the real-time input frame is eliminated without removing the needed details from the captured frame.

To obtain uniform dimensions and rotated frame techniques like filtering and normalization of selected frames from video input are involved in pre-processing. The image read from the loaded database is cropped, resized and normalized. The important step in pre-processing is converting color frames to grayscale images using OpenCV libraries.

Regions of information are separated by the process called segmentation. Input image frames are divided into homogenous, informative regions corresponding to different objects in the image based on edge, intensity and texture. during the process of segmentation.

Dataset is split into three parts. Training dataset which is data used to fit the model. Validation dataset to validate the generalization ability of the model or for early stopping, during the training process. Testing dataset that can be used for other purposes other than training and validating.

B. Feature Extraction

Students Facial Emotion Recognition towards the lecture is determined by the facial image obtained from the face detection stage forms an input to the feature extraction stage. To obtain real time performance and to reduce time complexity, for the intent of expression recognition, only eyes, lips and mouth are considered.

The combination of three features is adequate to convey emotions accurately. We propose Haar Cascade algorithm obtained from the OpenCV python library for face detection in the input video and pre trained CNN model for the emotion feature classification from the input dataset. In eye extraction the eyes display strong vertical edges and horizontal transitions due to its iris and eye white. Thus, the Sobel mask is applied to an image and the horizontal projection of vertical edges can be obtained to determine the Y coordinate of the eyes.

Eyebrow Extraction is done by selecting the region of interest in the edge image that is present just above each of the eye regions and indicated by two rectangular regions. The edge images obtained from the above regions of interests are improvised for gathering useful information using further refinement. Numerous edges are extracted from the edge image and these extracted edge frames are then dilated and the holes are closed by filling. The resulting edge frames are used in further refinement of the rectangularly marked eye regions.

C. CNN Emotion Classifier

CNN classifier model shown in Fig. 2 is used for emotion classification.

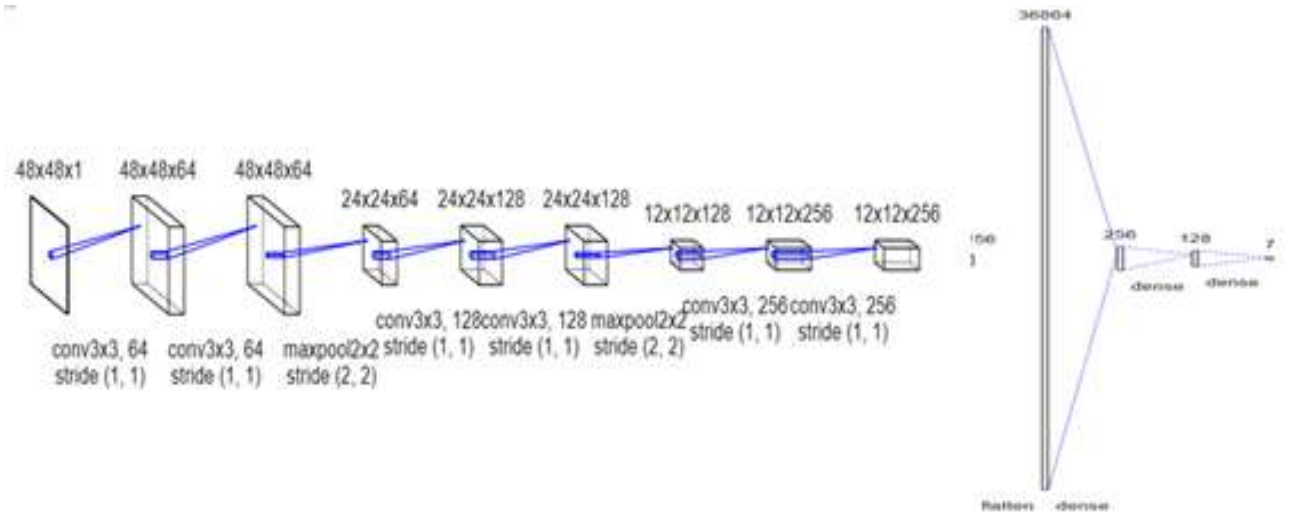


Fig. 3. CNN Emotion Classifier Model

Convolution neural network is one of the representative network structures in deep learning and has become a hotspot in the field of speech analysis and image recognition. CNN can take real-time raw image frame input, thus eliminating feature extraction and data reconstruction process in the standard deep learning methodologies. Weight network layer's structure of CNN architecture matches the biological neural network of humans, which reduces the difficulty of the network model and lowers the number of weights.

D. Algorithm

- Step 1: The learner logins into the e-learning portal and the web-camera begins frame acquisition.
- Step 2: The face of the learner is detected and then processed.
- Step 3: The eyes region and other regions of interest are detected and cropped.
- Step 4: Student's head position is identified
- Step 5: Students eye position is tracked whether left, right or up.
- Step 6: Student emotion is predicted.
- Step 7: The student's attention state is classified into concentrated or distracted based on the valid combinations of head, eye and student's emotion.
- Step 8: Finally, if the student's engagement level is engaged course URL is activated for video lecture otherwise an URL is activated for course documentation.

E. List of Hyperparameters

The below table TABLE I. shows the list of hyper-parameters and their corresponding variants involved in the proposed convolutional neural network of the emotion classifier module.

TABLE I. HYPER-PARAMETER LIST

SI. NO	***Hyper-parameter***	***Variants***
1	Non-linearity	ReLU
2	Batch Normalization	after non-linearity
3	Pooling	Max
4	Pooling window size	3x3 with same-padding
5	Color space & Pre-processing	Grayscale
6	Classifier design	Conv2d-BatchNormalization-MaxPooling-Dense
7	Input frame size	48X48
8	Dataset size	294MB
9	Batch size	64
10	Epochs	100

V. RESULTS DISCUSSION

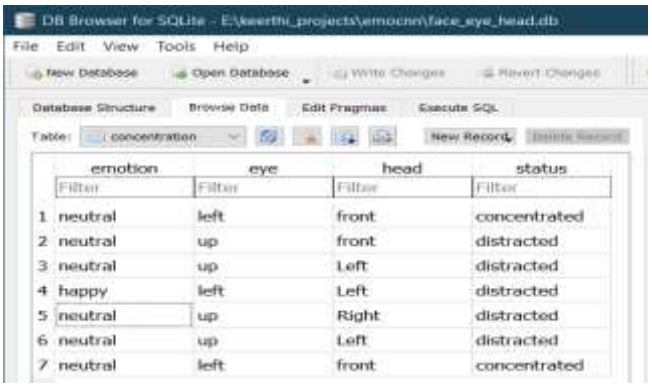
The results obtained from the implementation of the proposed model has about 63.8% of test accuracy. The output from the three basic modules of the proposed system are shown below in Fig. 4.

Fig. 4. Emotion, eye-gaze and head-pose detection

The output combinations from the above three modules are used to detect the status of the learner as concentrated or distracted as represented in TABLE II. If the student is classified as concentrated then the course URL for the video lecture can be activated. Otherwise the course documentation URL will be activated for the distracted learner.

TABLE II. DATABASE ENTRIES FOR TEACHER REFERENCE

Emotion	***Eye-Gaze***	***Head-Pose***	***Status***	***Activate URL for***
neutral	left	front	concentrated	video lecture
surprise	up	front	distracted	course documentation
happy	right	left	concentrated	video lecture
neutral	left	front	concentrated	video lecture
neutral	up	right	distracted	course documentation
neutral	up	left	distracted	course documents
neutral	right	right	concentrated	video lecture
surprise	right	left	distracted	course documents
sad	up	front	distracted	course documents
happy	left	front	concentrated	video lecture
sad	left	front	distracted	course documents
happy	right	front	concentrated	video lecture
fear	up	right	distracted	course documents
disgust	left	left	distracted	course documents
sad	up	left	distracted	course documents
happy	up	front	concentrated	video lecture



	emotion	eye	head	status
1	neutral	left	front	concentrated
2	neutral	up	front	distracted
3	neutral	up	Left	distracted
4	happy	left	Left	distracted
5	neutral	up	Right	distracted
6	neutral	up	Left	distracted
7	neutral	left	front	concentrated

Fig. 5. Resultant attributes are uploaded to the database

VI. CONCLUSION

In this paper, we have proposed a system that uses a progressive and repetitive method to monitor and record the learning individual's engagement level thereby analyzing the learner's concentration to the instructional content and initiates a response to redirect the learner's attention when they become distracted. The quality of the course content can be enhanced forward based on the level of engagement recognized using eye tracking and head movement detection. The proposed system is highly effective in identifying facial expression, emotions, eye-gaze and head-pose in real-time e-learning environments and detects the level of focus or distraction of the student. Moreover, the common methodology can be used for other aspects like eye-gaze and head-pose movement identifications. Future improvements can be made by adding facial area recognition to uniquely identify the individual student and provide personal feedback and evaluation of learner focus during the e-learning period of time. Also, the same proposed system can be modified to evaluate individual engagement in a classroom environment and provide the teacher with this information in real-time or as a summarized report at the end of a lecture. Further, this method can also be applied to any kind of real-time learning situation.

REFERENCES

[1] Fabri, M., Moore, D.J., Hobbs, D.J "Mediating the Expression of Emotion in Educational Collaborative Virtual Environments: An Experimental Study" (2014)
[2] M. Feidakis, T. Daradoumis, S. Caballé and J. Conesa "Measuring the Impact of Emotion Awareness on e-learning Situations" (2013)
[3] J. Yu "An Infrastructure for Real-Time Interactive Distance E-Learning Environment" (2010)
[4] S. Al. Hend, G. K. Remya "Eye Tracking and e-Learning: Seeing Through Your Students" (2010)
[5] I.E. Haddioui, and M. Khaldi "Learner Behavior Analysis through Eye Tracking" (2011)
[6] Pushpaja V. Saudagar, D.S. Chaudhari "Facial Expression Recognition using Neural Network –An Overview "(2012)
[7] Mohamed Sathik M, Sofia G "Identification of student comprehension using forehead wrinkles" (2011)
[8] Bailenson J, Beall A, Blascovich J, Raimundo M, Weishbush M "Intelligent agents who wear your face: User's reactions to the virtual self (2010)
[9] Goodfellow, I.J., Erhan, D., Carrier, P.L., Courville, A., Mirza, M., Hamner, B., Cukierski, W., Tang, Y., Thaler, D., Lee, D.H., et al.: Challenges in representation learning: A report on three machine learning contests. In: ICONIP. pp. 117–124. Springer (2013)