

An adaptive system for predicting student attentiveness in online classrooms

Bhavna Gupta¹, Richa Sharma¹, Roli Bansal¹, Gagan Kumar Soni¹, Parul Negi¹, Paawan Purdhani²

¹Department of Computer Science, Keshav Mahavidyalaya, University of Delhi, Delhi, India

²Department of Electronics, Keshav Mahavidyalaya, University of Delhi, Delhi, India

Article Info

Article history:

Received Dec 26, 2022

Revised Mar 30, 2023

Accepted Apr 2, 2023

Keywords:

Adaptive eye aspect ratio

Attentiveness

Camshift algorithm

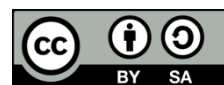
Histogram of oriented gradient

Viola Jones algorithm

ABSTRACT

The effectiveness of teaching methods is determined by whether a student is attentive in a lecture or not. In face-to-face classroom teaching, a teacher is able to judge whether students are understanding the subject, based on their facial expressions. However, since the uprise of COVID-19 pandemic, virtual classrooms have found a holding in the field of education and detecting attentiveness of students is a challenge in the same. This paper proposes a student attentiveness model that would detect and monitor a student's eye state to determine their level of attentiveness and provide a real-time feedback mechanism to the teacher. The proposed model employs a histogram of oriented gradient (HOG) method in conjunction with support vector machine (SVM) algorithm for face recognition. It then computes an adaptive eye aspect ratio (AEAR) for each individual student to determine their level of attentiveness. The model is tested on a real-time dataset and validated using classifiers (SVM, decision tree, and random forest). The results of the classifiers verify that the model produces an accuracy of more than 92%.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Richa Sharma

Department of Computer Science, Keshav Mahavidyalaya, University of Delhi

Delhi, India

Email: richasharma@keshav.du.ac.in

1. INTRODUCTION

In recent years, the thriving online education sector has resulted in a transition from physical to online classrooms. This resulted in providing access to students in remote areas, hence increasing the class size and allowing for more flexible class scheduling. Also, the entire closure of educational institutions to prevent the spread of COVID-19 produced a substantial disruption in the academic year and a significant gap in student learning. The outbreak compelled the traditional classes to shift to the online lecture mode. Although online classes ensure that students continue to learn even when institutions close down, thereby not disrupting their studies or jeopardizing their careers, the environment is relatively new and so has several flaws. The major challenge of online education compared to conventional education is to keep the students engaged throughout the class. Students who take classes online are no longer in a setting that is solely focused on learning and are instead surrounded by everyday activities like using their phones or other gadgets. To take online classes, one needs to be connected to the internet, this allows students to readily access numerous social media and other sites, which becomes a huge source of distraction for them. Lack of direct interaction with the teacher and peers' limits and diminishes feedback causing social isolation, and may make students less attentive.

In a traditional setting, it is easier for the teacher to monitor the students' attention through their facial gestures or expressions; however, this becomes difficult in an online lecture because the teacher finds difficulty in viewing everyone at the same time [1]. Online learning demands more self-attention than classroom

instruction, to understand the syllabus or course content. Further, learners need to be active and interested throughout the online class. Therefore, it becomes paramount for the teachers to determine the attentiveness of the students during the class so that they can improvise the teaching method or pay more attention to the students who are not attentive. The level of attentiveness can be determined with the help of physiological features and expressions.

Attentiveness is the state of being observant and focused on a specific topic. It is viewed as a cognitive process that allows one to position oneself towards relevant stimuli and respond to them, as it is necessary for remaining engaged in the topic. Building real-world systems capable of reliably evaluating students' attentiveness levels by monitoring their facial expressions has various obstacles. Attentive behavior is a complex phenomenon with various manifestations in different people's faces. The intricacy of such behavior is also incompatible with models that rely on the simple, rule-based classification of facial expression patterns, making it an issue that requires training a machine learning model [2]. A substantial amount of training data covering a wide spectrum of attention behavior is required for a machine learning-based attentiveness classification model. The annotation of collected data is another problem linked with modelling attentiveness. Attentiveness is a temporal phenomenon characterized by constant transitions between attentive and inattentive states. It is unclear how many attentive states should be marked and what are the distinguishing properties of each state [3]. Furthermore, the transition between the two states is ambiguous, making it difficult to determine when a particular state begins and ends. As a result, various annotators may annotate distinct states differently, resulting in a considerable shift in the model's forecast [2].

There have been numerous attentiveness detection methods put forth in the literature. Depending on the data they utilize, how they analyze it, and the features they employ, these strategies can be divided into various classes. By predicting the eye states (open/closed) from a single image, a number of techniques have been developed to assess attentiveness. The fundamental step in this estimation is to extract various sets of features, like the histogram of oriented gradients (HOG) and local binary patterns, which are then used to train various machine learning approaches to distinguish between the two eye states. Several convolutional neural network (CNN) architectures have recently been used by academicians to improve the accuracy of state estimate results, although these techniques are challenging to spot in real-time. Currently, a significant amount of research has been done on attentiveness detection. However, we only list a select group of significant and pertinent literary works here.

OpenPose, a real-time multi-person system that jointly detects the human body, hand, facial and foot key points on the image which are connected to generate a skeleton figure, was used for gauging the attention of students [3]. Stiefelhagen *et al.* [4] evaluated the attentiveness of participants in a meeting on the basis of both acoustics and visual cues. Using an omnidirectional camera, faces were captured and neural networks were employed to determine each participant's head pose. This study did not include eye gaze data; instead, a Bayesian approach was used to determine each participant's focus of attention based on estimates of their head altitude. However, the head pose may not always be the correct representation of one's attentiveness.

Sukno *et al.* [5] were able to identify eye blinks in a video series taken with a regular camera by using Active Shape Models with Invariant Optimal Features and 98 landmarks. They used the typical separation between vertically parallel eyes. Their analysis of the Av@Car database [6] revealed that the average blink lasts about 3107.3 milliseconds. They were highly accurate for open-eye frames, but only to 80.5% for closed-eye frames. However, using a 98-point Active Shape model takes a long time and is not suitable for real-time applications where the minimum blink period is 75 milliseconds [7].

Matlovic *et al.* [8] identified emotions using a dual method of electroencephalography (EEG) and facial expression recognition. Despite its utility, this technique is intrusive and infeasible in online classes. Another strategy devised by Zaletelj and Košir [9] made use of sensors to monitor student behavior both in and out of the classroom. This data analysis yielded a distinct set of practical aspects, such as body tilting, head angle, and displacement. However, using a sensor in a classroom is not an ideal choice. Saif and Mahayuddin [10] performed eye state recognition for detecting driver drowsiness using deep learning. However, their model is computationally expensive and thus difficult to implement in a scenario with a large number of online participants.

Melo *et al.* [11] had made use of heart rate variability, another intrusive metric linked to emotional and cognitive regulation. However, extreme exhaustion, stress, or sleepiness can affect this measurement and give inaccurate results [12]. Kawamura *et al.* [13] collected multimodal data from the online session with recording software (Bandicam, 30fps) which included heart rate and seat pressure measurements along with facial expressions. However, asking a student to wear a heart rate sensor at home and using physical devices to monitor seat pressure is rather unfeasible for a wide collection of reasons, one of which is the expensive cost of sensors and physical devices.

The researches [14]–[18] performed face detection for finding the landmarks on the face which was further used to calculate eye aspect ratio (EAR). Then the eye was classified, based on a fixed EAR threshold, into different states. Abate *et al.* [19] suggested an attention monitoring system that used gaze tracking, blink

detection, expression detection and yawn detection. The researchers calculated EAR to count the eye blink. Also computed yawn duration and amplitude. They found that 60-70 centimeters (i.e., the regular distance from the laptop screen) is a good ‘point of observation’ for students and it well balanced the image resolution and gaze detection. But restricting a person within a specific range of distance and not taking eye dimensions into account, is impractical. To overcome these limitations and provide support to teachers during online classes, our proposed system focuses on the eye traits of individual students and adaptively calculates adaptive eye aspect ratio (AEAR) for each of them from their captured video streams. Irrespective of the student’s distance from the camera and varied eye dimensions, the model successfully predicts the attentiveness level of the students. The goal of the proposed study is to predict the attentiveness of the students in an online class. In this study, we collected video data of students sitting in a virtual classroom with their device cameras switched on. Subsequently, different frames of the captured videos were processed for face detection and tracking region of interest (ROI). Further, eye features are extracted and AEAR is computed with the objective to categorize the eyes as “open”, “semi-open” and “close”. Lastly, the annotated data of all the frames of each video is compiled to classify a student as attentive or non-attentive. The following are the contributions of this study: i) implemented a model which is able to predict the attentiveness of all the students based on their AEAR irrespective of their demographic diversity. The novelty of the model lies in its adaptive nature; and ii) the model was trained and tested on a dataset comprising real-time videos of the students attending an online class.

The rest of the paper is organized as follows. In section 2 presents the proposed method. In section 3 presents the results and discussion. In section 4 concludes the paper.

2. METHOD

The paradigm shift from offline to online mode of teaching was the need of the hour during the COVID-19 pandemic. Looking at the advantages of the online classroom, many educational institutions are imparting instruction in either fully online or blended mode even after the subsiding of the pandemic [20], [21]. Hence in online teaching, the major challenge for a teacher to detect student attentiveness still remains. To address this, the proposed work utilizes eye traits to determine an individual’s attentiveness. Figure 1 presents a block diagram of the proposed student attentiveness system comprising various modules that include pre-processing and face detection, tracking ROI, finding facial landmarks, and classification of attentiveness states.

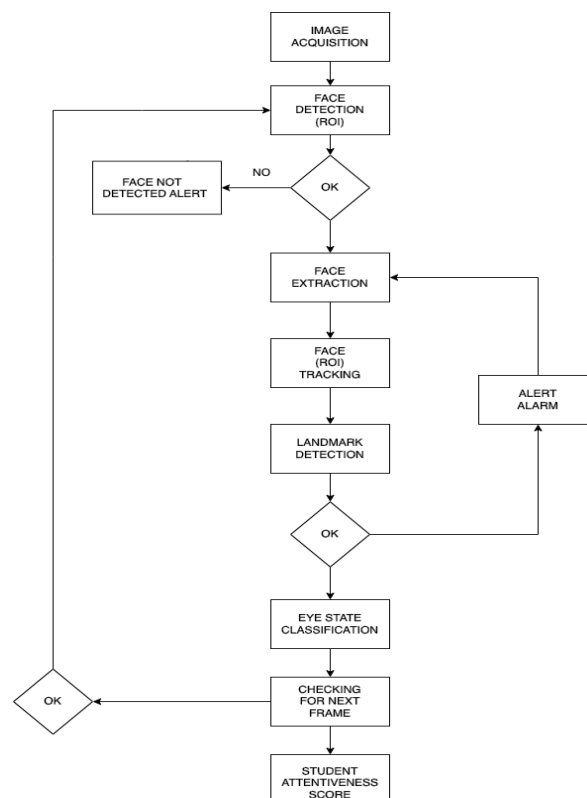


Figure 1. Block diagram of the proposed approach

The detailed workflow of the same has been depicted in Figure 2. As depicted in the figure, the video frame is provided to the model in the bulk of frames. The first frame has been used to determine the ROI using Module 1 and other modules have been looped to the rest of the leftover n-1 frames (taking the total number of frames to be n). All the modules have been thoroughly explained in the subsections that follow.

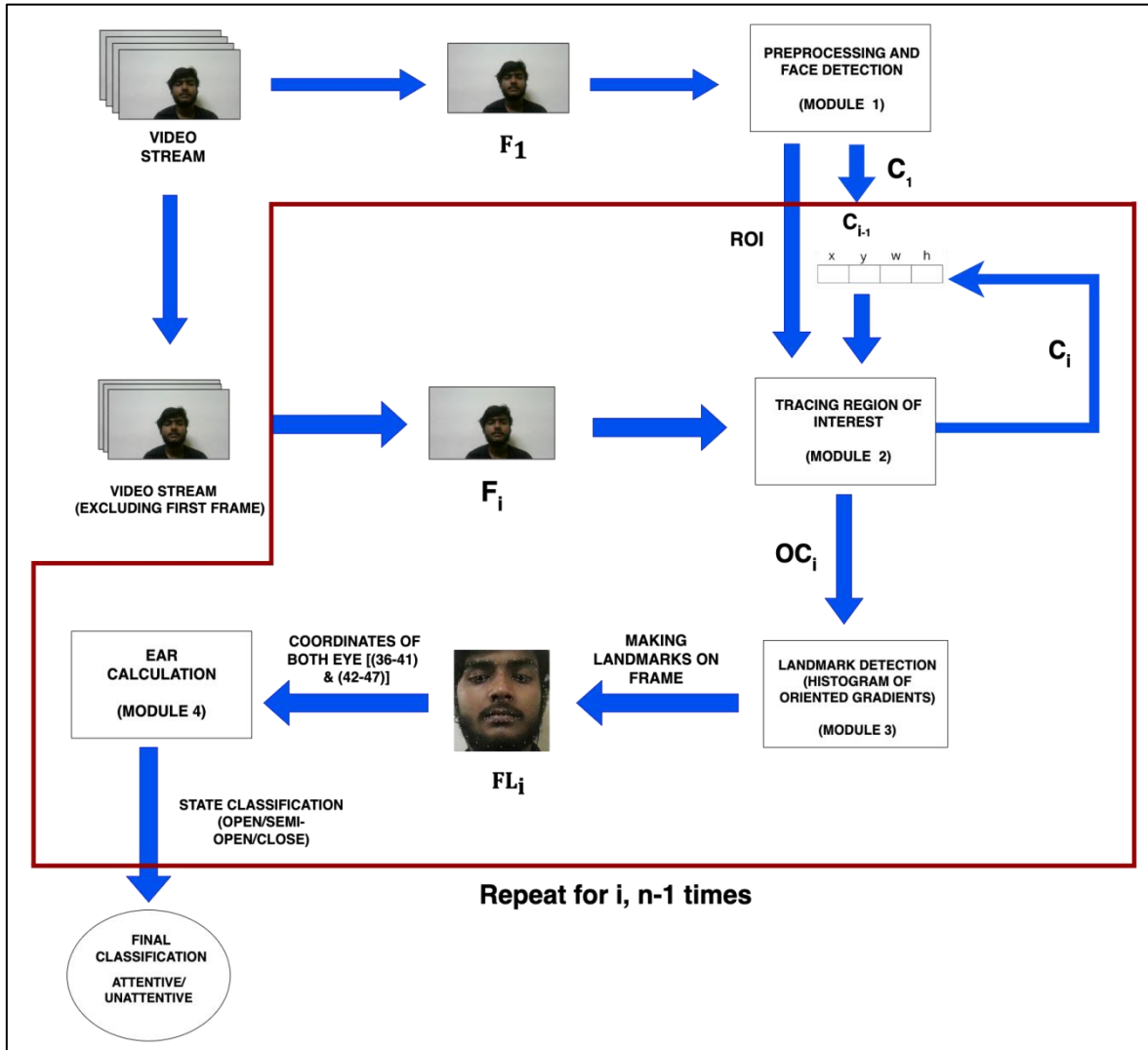


Figure 2. Workflow diagram of proposed approach

2.1. Preprocessing and face detection (module 1)

The first step of the model is to locate the student’s face. We have used the standard Viola Jones algorithm [22], [23] to find $C_1(x, y, w, h)$, the coordinates of the first frame F_1 , where x, y represent upper left corner of the frame and w, h represent the width and height of the frame respectively. C_1 is used to extract out the facial ROI which has been further used in Module 2. The Viola Jones algorithm works in four major steps—choosing haar-like features, calculating the integral image, executing adaboost learning algorithm and implementing a cascade structure [24], [25].

In our study, video captured from the camera linked to each student’s computer or laptop (assuming the student’s camera is active throughout the class), is provided as a video stream from which F_1 is given as input to the Viola Jones algorithm. Figure 3 illustrates the extraction of the face ROI and coordinates C_1 from F_1 as explained above.

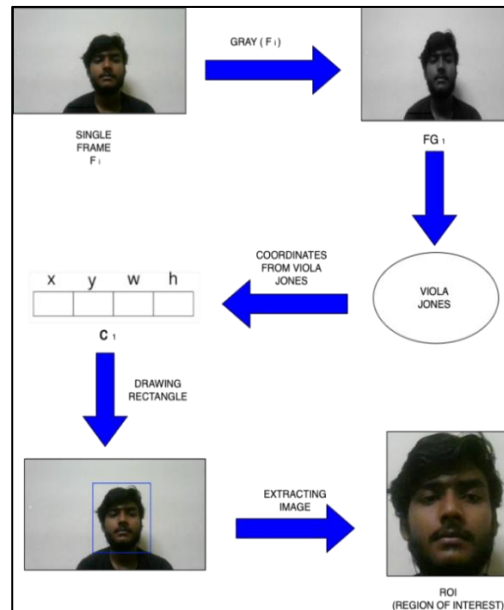


Figure 3. Finding ROI using Viola-Jones algorithm in module 1

2.2. Tracking region of interest (module 2)

Module 2 uses the Camshift algorithm [26], [27] to provide accurate head and facial tracking in a perceptual user interface. As F_1 has already been pre-processed for face detection in Module 1, so Module 2 will track ROI from the remaining N-1 frames, assuming N is the total number of frames in the video stream. The Camshift algorithm has been chosen for its robustness, invariance with respect to the scale and the rotation, treatment of obstructions and insensitivity to the deformations in objects in real-time.

Further, we calculate a probability distribution image (PDI) which associates a pixel value with a probability that the given pixels belong to the target. PDI is being calculated using the histogram back projection method, which is a basic procedure that associates an image's pixel values with the value of the relevant histogram bin [27]. To generate the PDI, we compute an initial histogram from ROI and calculate the back projection with F_i , the i^{th} frame which is in consideration.

PDI of F_i and C_{i-1} (coordinates of ROI in the previous frame) are provided to Camshift as shown in Figure 4. Finally, we get new position of ROI in F_i which is used to get the coordinates of a rectangle, C_i as (x_i, y_i, w_i, h_i) . This is used as input coordinates for the next frame and to extract the output of the Camshift for i^{th} frame, OC_i , that is being provided for landmark detection and classification as discussed in the next subsection.

2.3. Finding facial landmarks (module 3)

The next step is to acquire the facial landmarks. The main idea behind this method is to locate 68 precise spots on the face, such as the corners of the mouth, brows, eyes, nose, and so on [28]. To locate these 68 coordinates on any face, *dlib* a python-based library based on HOG that implements the linear support vector machine (SVM) method for object detection [29] is used. This library creates landmarks for the complete face, which is depicted as white dots in Figure 5. The landmarks are adaptable to varied human faces and are unaffected by extremely quick movements. However, in our context, taking into account other face features may provide redundant information and raise computing expenses, thus jeopardizing real-time monitoring with no consistent advantage in attention detection. As a result, only eye-related landmarks are considered in our model. We first convert OC_i into the grey image and then apply it to get the eye contour with only the eyeball and eyelid. Based on output landmarks, EAR, the proportion between the width and height of the eye is computed as given in (1). All of the landmarks used in the EAR computation are depicted in Figure 6. The coordinates detected for eyes are labelled as p1, p2, p3, p4, p5, and p6. The Euclidean distances between vertically opposite points (p2-p6 and p3-p5) and horizontally opposite points (p1-p4) are calculated. The ratio between the distance of the vertically and horizontally opposite points is termed as EAR [30].

$$EAR = \frac{||p2-p6||+||p3-p5||}{2||p1-p4||} \quad (1)$$

EAR is used as an estimate to detect the state of eye-opening in our model. EAR is calculated for each frame of the input video stream. The traditional method employed a predetermined threshold to identify the eye condition. If the user’s EAR falls below the threshold for a specific number of frames (e.g., five consecutive frames for short blinks; twenty consecutive frames for extended blinks), a blink is identified. However, a common threshold does not suit every user because each individual has a ‘natural EAR’ due to his or her own facial traits. Therefore, we propose an adaptive EAR (AEAR) calculation method, which calculates a personalized EAR for each individual student. To determine the AEAR, we calculated EAR for the first 60 frames and then averaged them to get the AEAR threshold for classifying eye (open/semi-open/close) for each individual student.

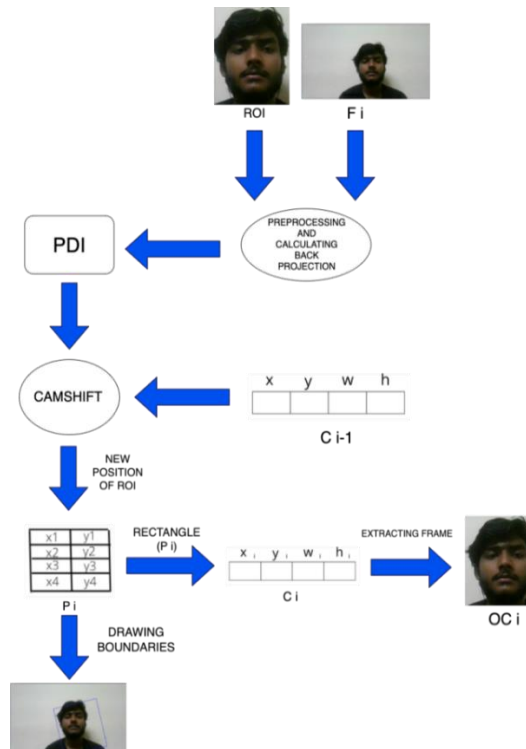


Figure 4. Tracking ROI using Camshift algorithm in module 2



Figure 5. Plotting 68 facial landmarks using dlib library

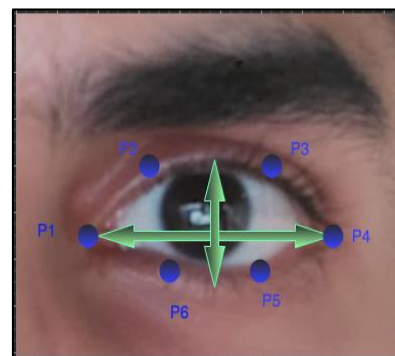


Figure 6. Plotting of landmarks across eyes for EAR computation

2.4. Classification of eye states (module 4)

The blink of an eye is the rapid closing and reopening of a human eye. Each person has a somewhat different blink pattern. Our proposed classifier takes as input a larger temporal window of frames in order to

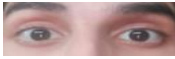
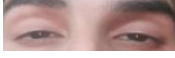
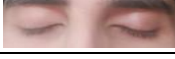
capture the blink pattern effectively and bring out better results. The model classifier performs in two steps, namely- identifying eye states and deciphering student attentiveness. The identified eye states of the student are *open*, *semi-open* and *close* respectively. The student attentiveness is classified into two states-*attentive* and *inattentive* respectively. The details of each step are discussed in sub-sections.

2.4.1. Identifying eye states

We found out through experimental results that 80% AEAR is a good threshold to segment an eye in open, close and semi-open states which is depicted in Table 1. If EAR is greater than or equal to AEAR, the state of the eye is taken as open. If EAR is greater than or equal to 80% of AEAR, the state of the eye is taken as semi-open. Otherwise, the state of the eye is taken as closed. In (2) suggests the criteria of classification for our model.

$$State_{Eye} = \begin{cases} Open & \text{if } EAR \geq AEAR \\ Semi - Open & \text{if } EAR \geq 0.8 * AEAR \\ Close & \text{if } EAR < 0.8 * AEAR \end{cases} \quad (2)$$

Table 1. EAR for different eye states

Eye state	Eye image	EAR	Eye open area
Open		$\geq AEAR$	High
Semi-open		$\geq 80\% \text{ of } AEAR$	Medium
Closed		$< 80\% \text{ of } AEAR$	Low

2.4.2. Deciphering student attentiveness

The blink pattern of each person differs in the speed with which their eyes close and open, the degree of squeezing and the duration of the blink. The eye blink lasts between 100 and 400 milliseconds [23] whereas the normal eye blink rate of a person is 12.55 blinks per min [31], according to the study, thus any period greater than that for closed eyes indicates sleepiness. Using the AEAR method, we are effectively able to determine whether the student is blinking or not. Traditional computer vision approaches compute blinks by looking for the whites of the eyes and detecting if they disappear for a period of time. Unlike this, in our approach it was possible to extract an eye-based variable, AEAR and draw inference based on its value.

To classify the final state of a student, our model accepts a temporal window of 24 frames as input. Eye states from 24 consecutive frames are concatenated to create a student's state feature depicted by $State_{student}$. The $State_{student}$ feature has been annotated and provided to our classification model for training. The same model is used further for the classification of the final state of a student into two categories: inattentive (0) and attentive (1) respectively. The overall combination of these two categories throughout the session has been used to provide a final percentage of attentiveness of each student at the end of the session represented by attention percentage, att_perc as depicted in (3).

$$State_{student} = \begin{cases} Attentive & \text{if } att_perc > att_thres \\ Inattentive & \text{otherwise} \end{cases} \quad (3)$$

where attention threshold, att_thres is to be decided by the teacher.

3. RESULTS AND DISCUSSION

We tested our algorithm on a system with a Ryzen 3 processor and 8 GB of RAM, using real-time data collected for this research. The objective of data collection was to collect examples of real-world attentive and inattentive behavior of students in online classes. Next, the data used in the study was annotated as attentive or inattentive based on eye states and their ratios. This dataset included videos of 10 online sessions each containing 40 participants ranging in age from 19 to 45 years old, with the majority (80%) falling between the age group of 19 and 21. All participants were instructed to sit in front of the system with their cameras switched on, and their faces were recorded. The data was saved with the consent of all participants to use their data for research and development. The camera object specific properties are shown in Table 2.

Next, the AEAR was computed for each participant in our model using (2). The graphs in Figure 7 illustrate varying EAR values of five students. The x-axis shows the frames, and the y-axis shows the EAR at a specific frame. After implementing the model to calculate the AEAR with a real-time camera, we compared it to the commonly used fixed threshold of the EAR (0.25) for classification [24], [25]. Table 3 presents the comparison of both the approaches (fixed and adaptive) on the sample data (84 frames) of one participant. The participant has an AEAR of 0.21 calculated using our model, which has been compared to a fixed EAR of 0.25. The table shows high variability in the prediction values, which can adversely affect the final results/classifications.

Table 2. Camera object-specific properties

Camera name	Integrated Webcam
Quality rating	720
Frame rate	30 fps
Stream type	Video
Image mode	RGB
Camera megapixel	0.9 MP
Camera resolution	1,280x720
Video standard	HD
Flicker reduction	60 Hz

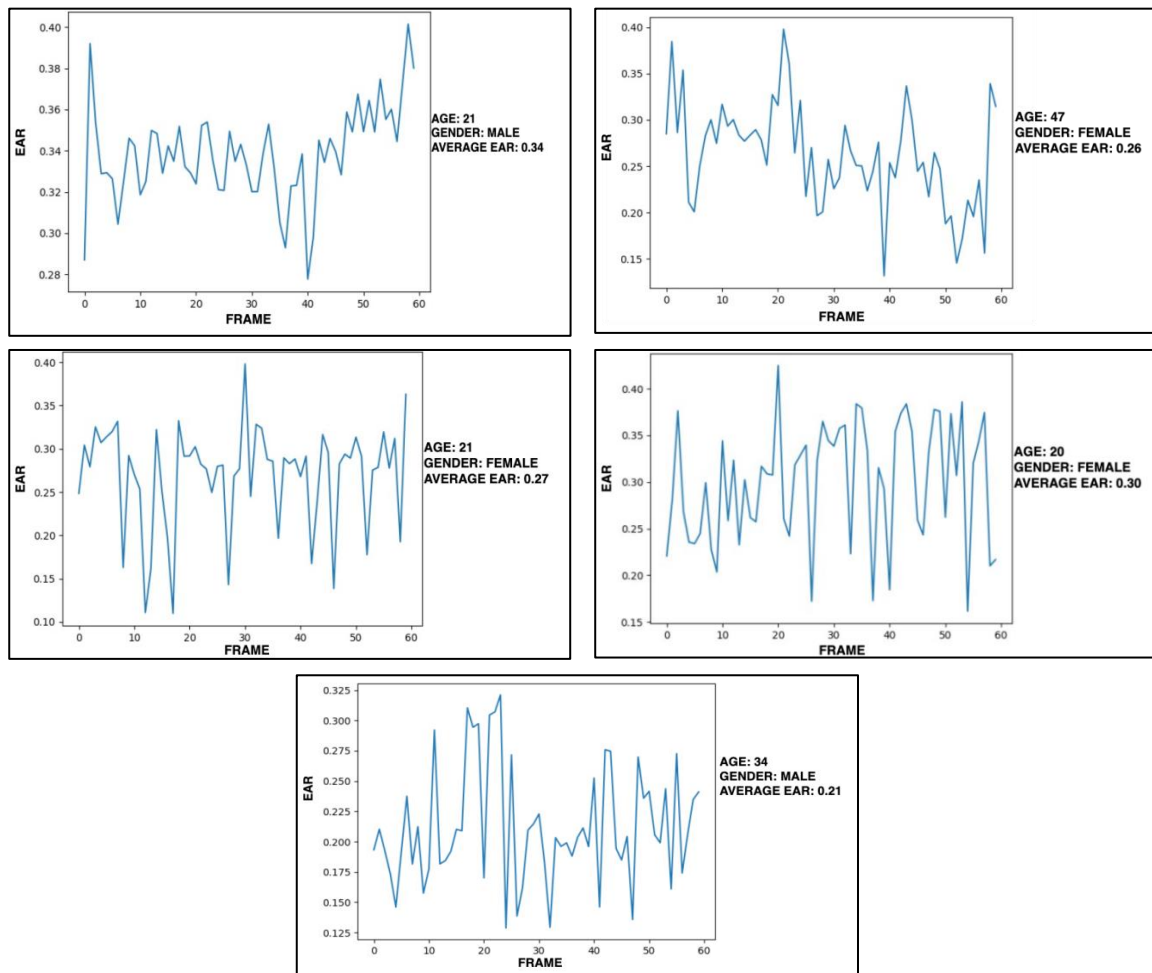


Figure 7. Graphs depicting variability in EAR

Table 3. Comparison of eye-state classification using fixed and adaptive approaches

No. of observed frames (84)	Eye state		
	Open	Semi-open	Close
Actual (annotated)	45	37	2
AEAR approach (0.21)	48	35	1
Fixed EAR approach (0.25)	6	68	10

The results in Table 3 show that a fixed EAR threshold approach yielded inaccurate predictions as compared to the proposed AEAR threshold approach. The mean absolute error and mean square error between fixed EAR and AEAR are 0.0304 and 0.0013, respectively. This error results in the misclassification of a student's state, which has been resolved in the proposed approach.

We performed evaluation analysis of our final classification model on SVM, decision tree, and random forest classification models. The data was split into 3:1 training and validation sets. We found that all three models had high accuracy, precision, and recall, show in the Table 4.

Table 4. Model accuracy table

Classifier	Inattentive (0)		Attentive (1)		Accuracy
	Precision	Recall	Precision	Recall	
SVM	0.93	0.93	0.94	0.94	93.37
Decision tree	1.00	0.91	0.92	1.00	95.36
Random forest	0.97	0.88	0.89	0.97	92.71

4. CONCLUSION AND FUTURE WORK

The culture of online classes, which got popularised due to the spread of COVID-19 pandemic, is still here to stay as it supports the paradigm of anywhere, anytime learning. However, the environment is new and has a number of flaws. One of the most difficult issues for teachers is not knowing whether or not their students are actively engaged throughout the session. In this study we presented a model to determine the different eye states (open, semi-open, and close) of students and classified them as attentive or inattentive. The model keeps track of a student's eye in the video stream to compute AEAR, the adaptive eye aspect ratio personalized for every student. This model can be used to monitor the student's attentiveness state and alert the teacher if the student seems to get inattentive.

Further, we validated our model with SVM, decision tree, and random forest classifiers which gave an accuracy of more than 92%. The results verify that the proposed approach effectively classifies the overall attentiveness of students in real-world scenarios. In future, the proposed model can be employed as an extension to online video conferencing platforms to keep a check on the diligence of a user to make the session interactive and efficient. Other facial features can also be analysed to make the model more versatile and inclusive.





REFERENCES

- [1] L. B. Krithika and G. G. L. Priya, "Student emotion recognition system (SERS) for e-learning improvement based on learner concentration metric," *Procedia Computer Science*, vol. 85, pp. 767–776, 2016, doi: 10.1016/j.procs.2016.05.264.
- [2] A. Joshi, S. Kyal, S. Banerjee, and T. Mishra, "In-the-wild drowsiness detection from facial expressions," in *IEEE Intelligent Vehicles Symposium, Proceedings*, Oct. 2020, pp. 207–212, doi: 10.1109/IV47402.2020.9304579.
- [3] A. Revadekar, S. Oak, A. Gaddekar, and P. Bide, "Gauging attention of students in an e-learning environment," in *4th IEEE Conference on Information and Communication Technology, CICT 2020*, Dec. 2020, pp. 1–6, doi: 10.1109/CICT51604.2020.9312048.
- [4] R. Stiefelhagen, J. Yang, and A. Waibel, "Estimating focus of attention based on gaze and sound," in *ACM International Conference Proceeding Series*, Nov. 2001, vol. 15-16-November-2001, p. 22, doi: 10.1145/971478.971505.
- [5] F. M. Sukno, S. K. Pavani, C. Butakoff, and A. F. Frangi, "Automatic assessment of eye blinking patterns through statistical shape models," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 5815 LNCS, 2009, pp. 33–42, doi: 10.1007/978-3-642-04667-4_4.
- [6] A. Ortega et al., "AV@CAR: A Spanish multichannel multimodal corpus for in-vehicle automatic audio-visual speech recognition," in *Proceedings of the 4th International Conference on Language Resources and Evaluation, LREC 2004*, 2004, pp. 763–766.
- [7] P. P. Caffier, U. Erdmann, and P. Ullsperger, "The spontaneous eye-blink as sleepiness indicator in patients with obstructive sleep apnoea syndrome - A pilot study," *Sleep Medicine*, vol. 6, no. 2, pp. 155–162, Mar. 2005, doi: 10.1016/j.sleep.2004.11.013.
- [8] T. Matlovic, P. Gaspar, R. Moro, J. Simko, and M. Bielikova, "Emotions detection using facial expressions recognition and EEG," in *Proceedings - 11th International Workshop on Semantic and Social Media Adaptation and Personalization, SMAP 2016*, Oct. 2016, pp. 18–23, doi: 10.1109/SMAP.2016.7753378.
- [9] J. Zaletelej and A. Košir, "Predicting students' attention in the classroom from Kinect facial and body features," *Eurasip Journal on Image and Video Processing*, vol. 2017, no. 1, p. 80, Dec. 2017, doi: 10.1186/s13640-017-0228-8.
- [10] A. F. M. S. Saif and Z. R. Mahayuddin, "Robust drowsiness detection for vehicle driver using deep convolutional neural network," *International Journal of Advanced Computer Science and Applications*, vol. 11, no. 10, pp. 343–350, 2020, doi: 10.14569/IJACSA.2020.0111043.
- [11] H. M. Melo, A. A. Hoeller, R. Walz, and E. Takase, "Resting cardiac vagal tone is associated with long-term frustration level of mental workload: ultra-short term recording reliability," *Applied Psychophysiology Biofeedback*, vol. 45, no. 1, pp. 1–9, Mar. 2020, doi: 10.1007/s10484-019-09445-z.
- [12] F. Rundo, S. Conoci, F. Trenta, and S. Battiato, "Car-driver drowsiness monitoring by multi-layers deep learning framework and motion analysis," in *Lecture Notes in Electrical Engineering*, vol. 629, 2020, pp. 169–175, doi: 10.1007/978-3-030-37558-4_25.
- [13] R. Kawamura et al., "Detecting drowsy learners at the wheel of e-learning platforms with multimodal learning analytics," *IEEE Access*, vol. 9, pp. 115165–115174, 2021, doi: 10.1109/ACCESS.2021.3104805.





- [14] A. Balasundaram, S. Ashokkumar, D. Kothandaraman, S. Kora, E. Sudarshan, and A. Harshaverdhan, "Computer vision based fatigue detection using facial parameters," *IOP Conference Series: Materials Science and Engineering*, vol. 981, no. 2, p. 022005, Dec. 2020, doi: 10.1088/1757-899X/981/2/022005.
- [15] T. Ghani, N. Jahan, M. M. Khan, S. M. T. Rahman, and S. T. A. Islam, "Development and analysis of a machine learning based software for assisting online classes during COVID-19," *Journal of Software Engineering and Applications*, vol. 14, no. 03, pp. 83–94, 2021, doi: 10.4236/jsea.2021.143006.
- [16] T. Soukupova and J. Cech, "Real-time eye blink detection using facial landmarks," in *21st Computer Vision Winter Workshop Luka Cehovin, Rok Mandeljc, Vitomir Struc (eds.)*, 2016.
- [17] S. Mehta, S. Dadhich, S. Gumber, and A. Jadhav Bhatt, "Real-time driver drowsiness detection system using eye aspect ratio and eye closure ratio," *SSRN Electronic Journal*, 2019, doi: 10.2139/ssrn.3356401.
- [18] U. Lahoti, "Drowsiness detection system for online courses," *International Journal of Advanced Trends in Computer Science and Engineering*, vol. 9, no. 2, pp. 1930–1934, Apr. 2020, doi: 10.30534/ijatcse/2020/158922020.
- [19] A. F. Abate, L. Cascone, M. Nappi, F. Narducci, and I. Passero, "Attention monitoring for synchronous distance learning," *Future Generation Computer Systems*, vol. 125, pp. 774–784, Dec. 2021, doi: 10.1016/j.future.2021.07.026.
- [20] O. Chamorro-Atalaya *et al.*, "Technological tools for virtual teaching and their effect on the satisfaction of online learning," *Indonesian Journal of Electrical Engineering and Computer Science (IJECS)*, vol. 25, no. 3, pp. 1634–1643, Mar. 2022, doi: 10.11591/ijeecs.v25.i3.pp1634-1643.
- [21] S. F. Omar, H. S. A. Nawi, R. W. M. Mee, L. S. Pek, T. S. T. Shahdan, and A. Woo, "Readiness in using online interactive platforms for remote teaching," *Indonesian Journal of Electrical Engineering and Computer Science (IJECS)*, vol. 24, no. 2, pp. 1047–1053, Nov. 2021, doi: 10.11591/ijeecs.v24.i2.pp1047-1053.
- [22] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2001, vol. 1, pp. 1-511-1-518, doi: 10.1109/cvpr.2001.990517.
- [23] H. Ma, Z. Yang, Y. Song, and P. Jia, "A fast method for monitoring driver fatigue using monocular camera," *Proceedings of the 11th Joint Conference on Information Sciences (JCIS 2008)*, 2008, doi: 10.2991/jcis.2008.52.
- [24] K. D. Ismael and S. Irina, "Face recognition using Viola-Jones depending on Python," *Indonesian Journal of Electrical Engineering and Computer Science (IJECS)*, vol. 20, no. 3, pp. 1513–1521, Dec. 2020, doi: 10.11591/ijeecs.v20.i3.pp1513-1521.
- [25] T. K. Soon *et al.*, "The utilization of feature based Viola-Jones method for face detection in invariant rotation," *International Journal of Advanced Computer Science and Applications*, vol. 9, no. 12, pp. 73–78, 2018, doi: 10.14569/IJACSA.2018.091211.
- [26] B. Akrouf and W. Mahdi, "Spatio-temporal features for the automatic control of driver drowsiness state and lack of concentration," *Machine Vision and Applications*, vol. 26, no. 1, pp. 1–13, Jan. 2015, doi: 10.1007/s00138-014-0644-z.
- [27] J. G. Allen, R. Y. D. Xu, and J. S. Jin, "Object tracking using camshift algorithm and multiple quantized feature spaces," *Reproduction*, vol. 36, pp. 3–7, 2006, doi: 10.5555/1082121.1082122.
- [28] S. E. Bekhouche, I. Kajo, Y. Ruichek, and F. Dornaika, "Spatiotemporal CNN with pyramid bottleneck blocks: application to eye blinking detection," *Neural Networks*, vol. 152, pp. 150–159, Aug. 2022, doi: 10.1016/j.neunet.2022.04.010.
- [29] I. Gupta, N. Garg, A. Aggarwal, N. Nepalia, and B. Verma, "Real-time driver's drowsiness monitoring based on dynamically varying threshold," in *2018 11th International Conference on Contemporary Computing, IC3 2018*, Aug. 2018, pp. 1–6, doi: 10.1109/IC3.2018.8530651.
- [30] C. B. S. Maior, M. J. das C. Moura, J. M. M. Santana, and I. D. Lins, "Real-time classification for autonomous drowsiness detection using eye aspect ratio," *Expert Systems with Applications*, vol. 158, p. 113505, Nov. 2020, doi: 10.1016/j.eswa.2020.113505.
- [31] L. G. Carney and R. M. Hill, "The nature of normal blinking patterns," *Acta Ophthalmologica*, vol. 60, no. 3, pp. 427–433, May 1982, doi: 10.1111/j.1755-3768.1982.tb03034.x.

BIOGRAPHIES OF AUTHORS







Dr. Bhavna Gupta     is currently working as an Associate Professor in the Department of Computer Science, Keshav Mahavidyalaya College, University of Delhi. She did her doctorate in computer science from the Department of Computer Science, University of Delhi in 2014. She has over 22 years of teaching experience. She has published various research papers in international journals and conferences. Her research interests include trust-based recommender systems, cloud computing and deep learning techniques. She can be contacted at email: bgupta@keshav.du.ac.in.







Dr. Richa Sharma     is working as an Associate Professor in the Department of Computer Science, Keshav Mahavidyalaya, University of Delhi, Delhi. She has 25 years of teaching experience. She obtained Ph.D. in Computer Science from the Department of Computer Science, University of Delhi, India. She has published several papers in international journals and presented papers in international conferences. Her research interests include natural language processing, computer vision, sentiment analysis, machine learning, and social opinion modeling. She can be contacted at email: richasharma@keshav.du.ac.in.







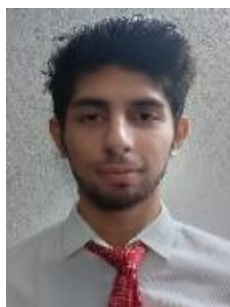
Dr. Roli Bansal     is working as an Associate Professor in the Department of Computer Science, Keshav Mahavidyalaya, University of Delhi. She obtained her Masters from Rani Durgavati Vishwavidyalaya, Jabalpur and Ph.D. from Department of Computer Science, University of Delhi. Having 25 years of teaching experience. She has publications in numerous international journals and conferences. Her research interests are image processing, biometrics, computer forensics and computational intelligence. She can be contacted at email: roli.bansal@keshav.du.ac.in.







Mr. Gagan Kumar Soni     graduated in Computer Science from Keshav Mahavidyalaya, University of Delhi in 2022. He is currently pursuing post graduation in Computer Science from University of Delhi. He has actively participated in and organized various co-curricular events at the undergraduate and postgraduate levels. His research interests are artificial intelligence, data mining and deep learning. He can be contacted at email: gagansoni4321@gmail.com.



Ms. Parul Negi     graduated in Computer Science from Keshav Mahavidyalaya, University of Delhi in 2022. Her area of specialization included machine learning and artificial intelligence and digital image processing. Her research interests are neural networks, practical application of AI and cloud computing. She can be contacted at email: parulnegi141@gmail.com.



Mr. Paawan Purdhani     is a B.Sc.(H) Electronics 3rd year student of batch 2020-2023, from Keshav Mahavidyalaya, University of Delhi. He completed his schooling from Delhi Public School, Dwarka, Delhi in 2020. He has keen interest and knowledge in the field of machine learning, cloud computing and data science and is looking forward to explore more about these topics. He can be contacted at email: paawanpurdhani@gmail.com.