International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-11 Issue-9, August 2022

Real-time Attention Span Tracking in Online Education

Kavyashree N, Shailaja L K, Anitha J, Sindhu P

***Abstract:** E-learning has changed how students grow over the past ten years by allowing them access to high-quality education whenever and wherever they need it. In any case, understudies frequently get occupied in light of different reasons, which influence the learning ability by and large. Numerous experts have been striving to address the nature of online education, but we really need a comprehensive solution to this problem. This essay aims to present a method for monitoring students' continuing attention during online classes using the surveillance camera and oral input. We investigate different picture handling strategies and AI calculations all through this review. We suggest a framework that makes use of five specific non-verbal cues to calculate an understudy's consideration score during computer-based tasks and generate continual feedback for both the association and the understudy. The output can be used as a heuristic to investigate both the speakers' and understudy' general methods of exhibiting themselves.*

***Keywords:** AI, Attention, Blink rate, drowsiness, Eye gaze tracking, Emotion classification, Face identification, Posture assessment, and Noise recognition.*

I. INTRODUCTION

Online learning is becoming increasingly popular and necessary. Practically every one of the schools and universities all through the world have moved to the web-based method of talks and tests because of the new Covid-19 episode, and this pattern will in all probability go on in the forth coming years. Growing demand for online education creates a path for field mechanisation.

One major problem with the internet-based style of lectures is that students frequently lose interest after a certain amount of time so there is no automated component to monitor their activities during the lessons. A few understudies will quite often begin a talk on the web and get away from the spot, or could try and utilize an intermediary to compose online tests for them. This current situation also occurs in online learning environments, such as EdX and Coursera, in which the student tries to bypass addresses just for fulfilment and accreditation. The misfortune in focus influences the understudy's information level as well as damages the general public by creating low-gifted workers. In our work, we offer a solution to this problem. The following is how the paper is set up: The literature is reviewed in Section 2, the attention of span tracking is illustrated in Section 3, the Proposed Methodology Execution evaluation is demonstrated in Section 5, and the paper is concluded in Section 6 with future ideas works.

II. LITERATURE REVIEW

The reason for ability to focus location during on the web classes is to assemble information and break down the condition of the understudy, to assess his presentation in light of fixation level, rather than simply scholastic scores. As indicated by [1], the typical squint pace of an individual is between 8 to 21 flickers each moment, yet when the individual is profoundly centred around a particular visual errand, the pace of flickering has altogether decreased to a normal of 4.5 flickers each moment. Similarly, the flicker rate heightened to over 32.5 squints each moment when the singular's fixation level is low. The review [2] investigates how the close to home condition of understudy's changes during the growing experience and how profound input can further develop learning encounters. Profound states like satisfaction, happiness, shock, and unbiased mean a positive helpful opportunity for growth, though, feelings like bitterness, dread, outrage, and loathing address a pessimistic encounter. The review [3] talks about how mind-meandering can adversely affect execution and how eye stare information, gathered utilizing a committed eye tracker, can naturally identify deficiency of consideration during PC based errands. The eye's behaviour patterns were observed, and it was discovered that the instances frequently changed when the mind was meandering. The paper [4] explains what a lack of regard actually entails for students' ability to learn. The audit paper [5] recommends that sluggishness, brought about by rest, fretfulness, and mental tension, is one of the central points that lead to deficiency of consideration. Innovative methods for identifying fatigue were examined, and it was discovered that the Haar classifier and Support Vector Machine (SVM) provide higher precision in progressively challenging scenarios. The review [6] lays out proof that the focusing ability understudies is dependent upon the natural commotion conditions, and the outcomes propose that clamor levels more prominent than 75dB truly affect the precision of the understudies. A computerised facial recognition model using Convolutional Neural Networks (CNN) and Principle Component Analysis is suggested in the publication [7]. (PCA).

Manuscript received on 05 July 2022 | Revised Manuscript received on 11 July 2022 | Manuscript Accepted on 15 August 2022 | Manuscript published on 30 August 2022.
* Correspondence Author

Kavyashree N*, Assistant Professor, Department of Master of Computer Applications, Dr. Ambedkar Institute of Technology Bangalore (Karnataka), India.

Shailaja L K, Assistant Professor, Department of Master of Computer Applications, Dr. Ambedkar Institute of Technology, Bangalore (Karnataka), India.

Anitha J, Assistant Professor Department of Master of Computer Applications, Dr. Ambedkar Institute of Technology, Bangalore (Karnataka), India.

In order to familiarise oneself with the overwrite meagre and specific face component maps, the review [8] suggests a twolayer CNN. Use of a sparsely chosen include extractor is how the Scanty Representation Classifier (SRC) operates on the exhibition. By identifying instances of conduct that is related to consideration, the review [9] makes a beeline for anticipating the client's level of consideration. From the writing audit, we have distinguished that the five boundaries - squint rate, look, eye stare, foundation commotion, and body pose make a decent list of capabilities to evaluate the consideration level of the understudies.

III. ATTENTION SPAN TRACKING

- **Attention span in Education**

By identifying instances of conduct that is related to consideration, the review [9] makes a beeline for anticipating the client's level of consideration.

In general, older children tend to have a lot more patience than younger children, and girls tend to have it more frequently than boys. When given tasks that suit their skills and interests, children typically have a longer attention span. In order for students to organise and consolidate key elements of the material they are studying; they need to have an attention span that is adequate. The majority of kids acquire the necessary degree of concentration during their regular school years. However, learning issues can occasionally arise for people with attention spans, including excessive focus on unimportant task details, general restlessness, and excessive activity that diverts attention. Children with learning difficulties, such as those who have been labelled as attention deficit disorder (ADD) and attention deficit hyperactivity disorder (ADHD), typically report having issues with their attention span (ADHD).

- **Online and distracted**

According to research conducted by Kent University, the effects of a reduced attention span and a lower concentration levels are more severe for online courses. Researchers contend that because online learners have more freedom to multitask, they are more susceptible to distractions. Students who took classes in a classroom with a today's lesson, on either hand, performed better.

The teacher, who serves as an anchor to keep students more grounded than they would be if they're on their own, sets several conditions that might be used to illustrate this. It's critical that we develop new strategies to maintain students' span of attention and focus at the best level possible in an educational environment that leans more and more toward digital content. Changing how we perceive them both could be a starting step toward achieving this. Instead of viewing them as concepts which need to be taught, start thinking of them as abilities that should be developed and supported. In this manner, pupils can start focusing in class on their own and stop relying on a teacher's authority. They can take charge of their very own learning experience, not only at school but in all area of their lives, when they're doing it voluntarily and cons consciously rather than as a result of being told to.

- **Online learning battles with students' attention spans**

In the Fig 1: explains the learning battles with student's attention span in online courses is a persistent problem that parents, students, and educators have all experienced before the COVID-19 school closures. Here are some details to help us comprehend the issue and how to assist our pupils in meeting this obstacle.

Fig. 1. Learning battles with Students' Attention Spans

It is a valid concern shared by many parents given the growing dependence on screen use for learning. The brain was built to learn from human connection and to explore the natural environment, according to Dr. Alice Holland, a clinical neuropsychologist at Child's welfare and UT Southwestern. 2020 (NBC 5 News). For many educators, the battle to capture children' attention was always an ongoing battle. According to a Cornell University study, "evidence reveals [that] learner focus starts to waver during lectures every 10–20 minutes." (2018) Smart Sparrow Many people have offered teachers tools like "active learning," for example. Recent studies have demonstrated that listening and active learning are more beneficial as "students perform better." (2018) Smart Sparrow In the Fig 2 explain the active listening skills which has been illustrated below

- **Active Learning**

A "wide variety of teaching practises which involve students' active stakeholders in their studying throughout class hours with their instructor" is mentioned. 2020; University of Minnesota. Students' learning potential will rise as long as they use "various sensory, cognitive, emotional, and social processes." (2018) Smart Sparrow What teaching strategies are open to students themselves now that Active Learning has given teachers some?

Fig. 2. Listening skills

International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-11 Issue-9, August 2022

Active Listening is the "practise of paying close attention while somebody else speaks, summarising and reflecting what is heard, and avoiding judgement and counsel." 2020's Very Well Mind. Although this tactic is typically linked with counselling, any student can develop this skill to combat attention span problems.

Active listening in the classroom refers to "students giving complete attention to educators or their peers," and this emphasis is designed to promote deep learning. 2020 (Education Corner). This differs from passive listening since passive listeners cannot retain the information for developing a better comprehension of the subject matter; instead, they merely pay attention long enough to reply. 2020 (Education Corner).

The following are advantages of active listening in education:

- Greater communication abilities;
- Quicker second-language learning;
- Lower levels of stress, anxiety, and despair; better relationship skills;
- Stronger feeling of empathy. 2020 (Waterford).

Students can still practice active listening when enrolled in distance education courses online. Here are some methods that students in grades K–12 can use to enhance their learning.

Techniques to Practice Active Listening

- ***Keep your eyes in contact***. "Making eye contact is a straightforward method of concentrating on a presenter and enhancing attention to the message being given." 2020 (Education Corner). When attending to online lessons, establishing eye contact can be done both in person and through a screen.
- ***At the end, pose questions.*** If listeners take breaks mid-sentence, they run the danger of understanding the subject matter inaccurately. When a learner is attempting to gain a deeper comprehension of the material, this may further confuse them.
- ***Synthesize and repeat after me.*** The ability to reflect back what was said gives the listener the opportunity to check their understanding of what was spoken. 2019 (ThoughtCo.). Students will find it simpler to write notes and to study the information later on as a result.
- ***Determine the main message.*** Students can gain a more thorough comprehension of the subject matter if they can recognize the main points of a teacher's lesson.

Students of all ages can benefit from practising active listening both within and outside of the classroom. Listening skills is one method that any student could use to control their online learning experience, as parents and teachers are concerned about the situation of student attention and concentration in distant distance learning.

This school year, distance learning is being used, so it's important that students learn AND use the habit of active listening. Visit https://www.acetutoring.com/blogs for additional helpful updates and advice. Best of luck to everybody; this is a team effort!

IV. PROPOSED METHODOLOGY

This review utilizes five boundaries to work out the capacity to focus level of the understudy going to the web-based class. Facial acknowledgment is utilized to approve the understudy's participation. Flicker rate, gaze, eye stare, foundation noise, and body movement are used to calculate the capacity to focus score, which is updated continuously for a window of 5 seconds. Once the internet-based address starts, all of the models anticipated to calculate the capacity to focus are performed equally rather than sequentially.

This is accomplished by using multithreading for all of the capabilities, which significantly reduces the amount of time that each model and the overall framework consume. As if on cue, the model will compute a focus score and provide ongoing feedback to the students in the form of live charts that are drawn for each border as well as the calculated focus score.

The supplementary sections will provide in-depth explanations of each model used in this analysis and their significance in determining one's capacity for concentration. Fig 3 shows the proposed framework's overall layout, while Fig 4 shows flowcharts explaining how each module operates (Fig. 4)

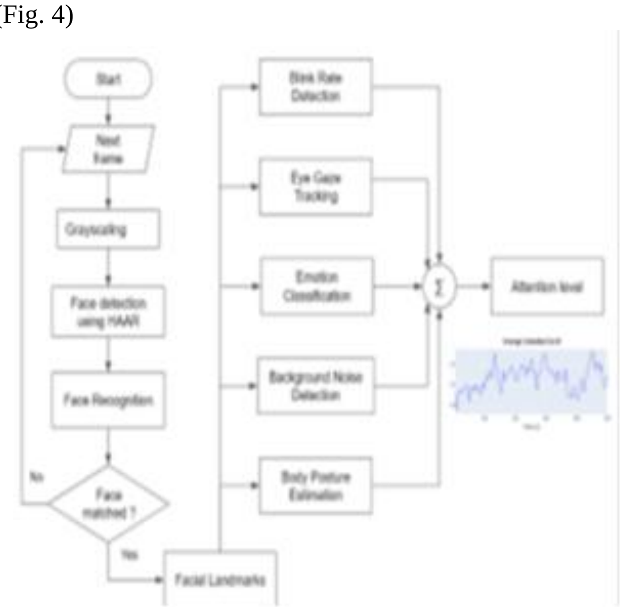

Fig. 3. The architecture of the proposed system.

A.Detecting Facial Landmarks

Face identification is carried out applying the Viola-Jones computation [13], which utilises a windowing approach to analyse photographs for differentiating characteristics of human countenances. The paper [8] provides a reliable, efficient method for handling 68 key concerns from the detected facial image using the OpenCV dlib package, as seen in (Fig 4).

To separate Haar highlights from the image, we use rectangular regions. The milestones are arranged into five categories of facial features: the brows, eyes, nose, mouth, and jaw. These categories are denoted sequentially using the central problems. These specific milestone highlights will be used as additions to other modules. By processing more central problems, we can improve the face identification module's accuracy. But doing so lengthens the handling time.

Retrieval Number: 100.1/ijitee.G91910811922
DOI: 10.35940/ijitee.G9191.0811922
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

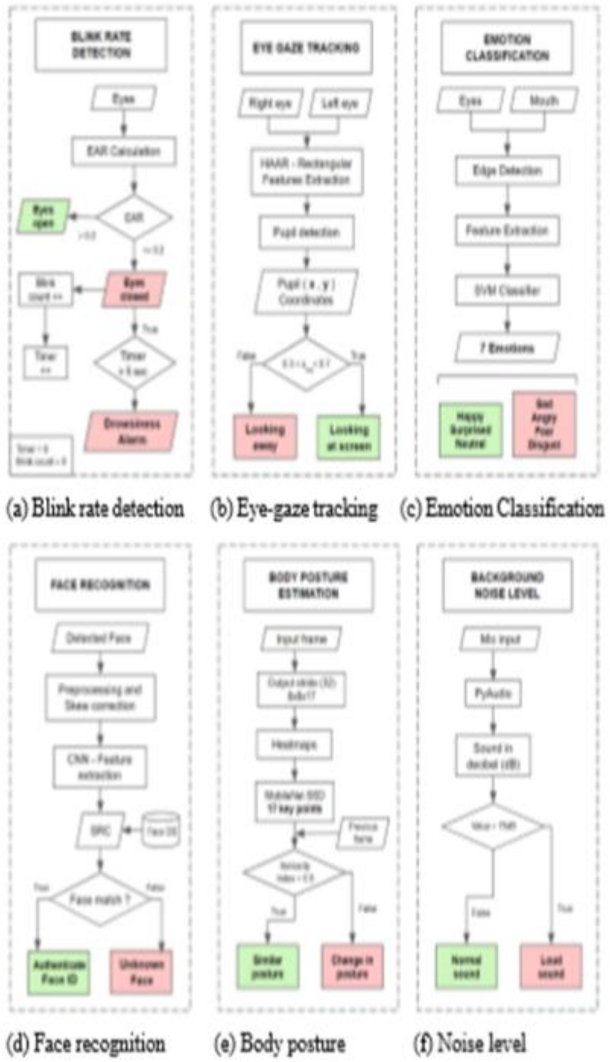

Fig 4. Diagrams showing each module's flow

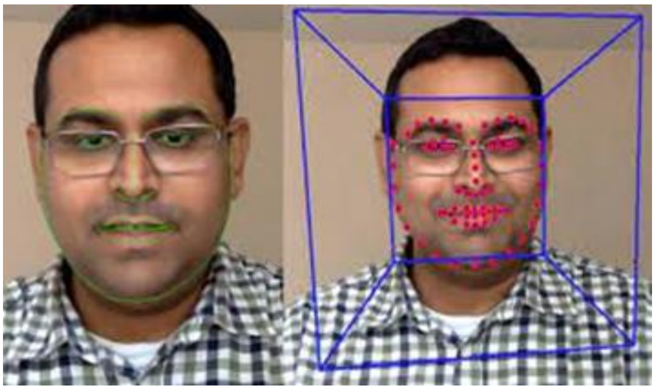

Fig. 5. Facial Landmarks and Face Detection.

B. Squint Rate Detection

One of the key factors in determining whether an understudy is paying attention or dozing off during class is blink rate. In the this module, we divide every eye into equal pieces by cropping the localities that include the eye matches. To determine if the eyes are open or closed, we calculate the Eye Aspect Ratio (EAR) using Euclidean distances for each edge (Fig 4a) in accordance with Formula (1). Additionally, to keep track of how long the eyes remain closed, we have a commencing clock that starts soon a squint is detected. If the eyes are seen as closed for more than two seconds [5], it is highly likely that the client feels tired (lack of consideration), and an alert will be sent both externally as shown in (Fig. 4b) and externally in the form of cautionary alert noises. To determine the typical squint speed of the client, we calculate the number of flickers occurring continuously over a time period of seconds. In light of test results, the EAR edge value is set at 0.2. On 306 images with 156 eyes closed and 150 open eyes, we used the flickering recognition module to arrange the squints with 91.02 percent accuracy and the open eyes with 92.66 percent precision.

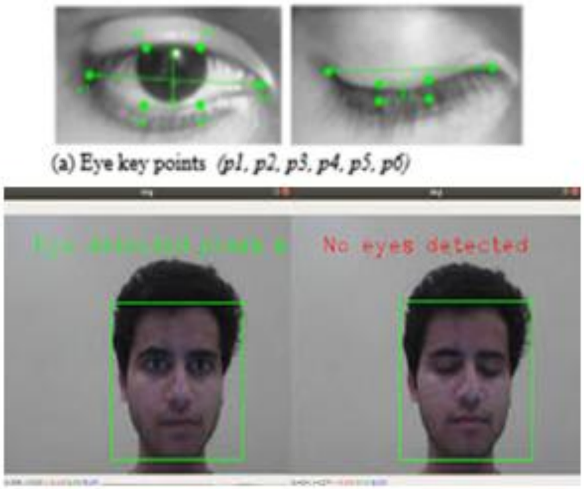

Fig. 6. Blink rate and sleepiness monitoring

C. Eye-stare Tracking

An understudy's eye-stare can be used to determine where he is looking, and it is frequently strongly correlated with his level of interruption. We analyse the excised eye region using rectangular parts, as suggested in [10], to identify the student's eye regions. The eye look course is calculated and planned using the understudy facilites (x, y) of each eye (Fig. 5a). We have outlined two potential classifications of eye stares based on the purpose of the screen: looking at a screen (Fig. 5c) and averting one's gaze (Fig. 5b) (right or left). We gathered 150 images in two categories: looking directly at the camera and looking away. For 113 images, our system had the opportunity to place the eyes precisely with an accuracy of 75.33 percent.

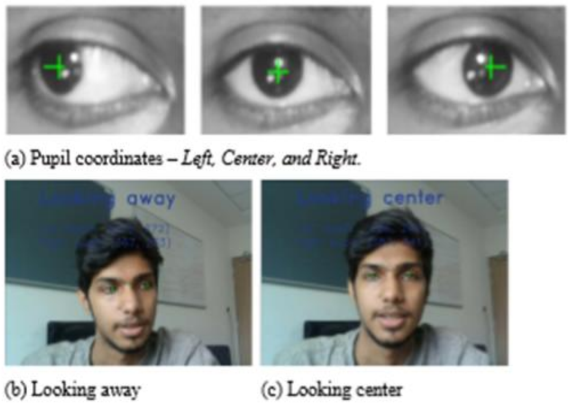

Fig.7. Eye-gaze monitoring

Retrieval Number: 100.1/ijitee.G91910811922
DOI: 10.35940/ijitee.G9191.0811922
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

D.Feeling Classification

In fig 8, the individual's attitude about taking an online course plays a big role in how seriously he takes it. This study makes use of face features such the eyes, nose, and mouth, which are differentiated using the Haarcascade classifier and the facial milestone finder. The understudy' emotions are classified into seven different classifications using a Support Vector Machine (SVM) calculation: anger, disgust, dread, happiness, misery, shock, and neutral. A score is given for every inclination relying upon its impact on the consideration level of the client. The conventional strategy for feeling characterization pre-handling utilizes just the trimmed eye. In any case, [11][12] proposed an elective answer for incorporate mouth highlights for better precision. As this technique just purposes Haar fountains to group the inclination, the handling speed was a lot quicker when contrasted with the Sobel edge eye discovery. The JAFFE dataset, which consists of 213 images and 7 feelings expressed by 10 distinct Japanese women, was used to approve the model. The practise set had 42 pieces and 7 different feeling classes. There were 70 images in the test set. On our test set, we attained an usual exactness of 82.55 percent (Fig. 6).

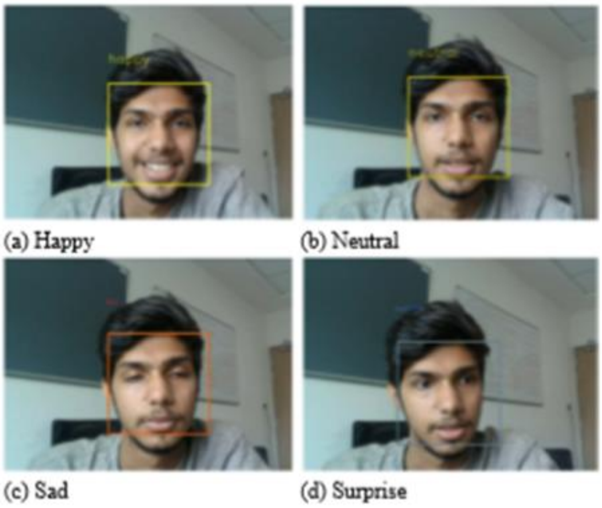

Fig 8: Classifying Emotions (Happy, Neutral, Sad, Surprise)

E.Face Recognition

A hearty facial acknowledgment framework is fundamental to verify the understudy in light of biometrics to keep away from understudy intermediaries and to robotize the participation the executive's interaction utilizing the webcam feed during classes. In this module, we modified [8]'s engineering and implemented a three-layer Convolutional Neural Network (CNN) made up of three convolution layers with max-pooling, a totally associated layer with dropout, and an output layer for the Sparse Representation Classifier (SRC). The dropout layer aids in lowering the cost of computing. 500 photos with four distinct faces made up the dataset we created. The dataset was divided into 200 photos for validation and 300 images for training. We achieved an accuracy rate of 94.8 percent and a validation data of 90 percent after 15 training iterations (Fig. 9).

Fig 9: Face Recognition

F.Estimating body posture

In the fig 10 Many scientists have used CNN or R-CNN to precisely assess the posture. However, the main goal of our exam is to continually calculate the understudy's consideration level without settled on handling time. Subsequently, we utilize the TensorFlow posture assessor (PoseNet), in light of Mobilenet SSD to appraise the stance of the understudy. The posture difference between one case and the past edge is measured using this model's intensity guidance. PoseNet can identify 17 key problems, including those involving the lower leg, hip, shoulders, elbows, and wrists. By comparing the differences in head body posture attitude across successive cases, we may predict whether the understudy would be anxious or calm during the internet-based address and assign a pixel similarity score accordingly.

Fig 10: Estimating body posture

G. Foundation Noise Detection

To identify the information sound coming from the device's mouthpiece, we use the Python package PyAudio. The understudy's level of fixation may be impacted by the fundamental turmoil during class. A school's normal noise level is 50 dB, and 75 dB is the limit for clearly audible commotion [6]. Anything that is louder than 75 dB is deemed to be an uproarious environment, and the results will be in direct opposition to the background noise. The model will continuously monitor the noise level at the foundation, and we can predict the usual clamour level with clockwork precision.

H. The general detection of attention level

To determine the consideration score in accordance with Formula, all of the scores from the aforementioned border scores (squint rate detection, eye gaze following, sensation grouping, body act evaluation, and basis clamour location) are standardised (2). With the understudy's projected consideration level and the scores for each boundary gradually updated, we construct live diagrams as shown in Fig. 11. Face recognition is not a component of our scoring system because it doesn't help us determine the understudy's comprehension level; rather, they use it for biometric verification and computerised involvement of the understudies. $\sum *100$

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

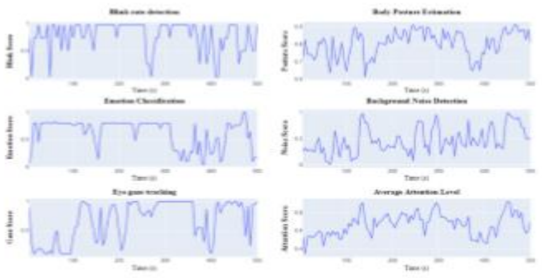

Fig 11: Real-time graphing of the attention level

V. EXECUTION EVALUATION

The demonstration of the framework was investigated using a data of 15 freshman students, comprised of nine men and six women. Three human witnesses were asked to rate the attention paid by the students based on the webcam video that was recorded, which served as the basis for the ground truth-value. The students were asked to participate in online conversations on diverse topics for 500 seconds each. To evaluate the overall presentation of our framework, we compared the noticed scores with the anticipated consideration ratings. The association between expected and noticed scores is shown in (Fig. 12), and the framework's exhibition metrics are shown in Table 1.

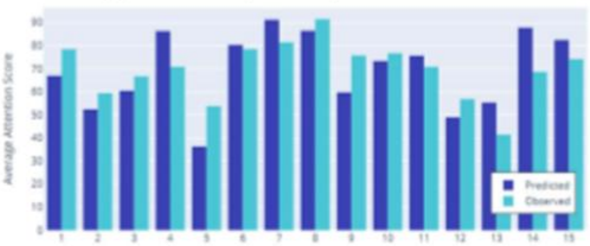

Fig 12: Comparison of the predicted and observed attention scores.

Table I: Performance Metrics

Metric	RMSE	MAE	R2	MAPE
Value	11.152	9.837	0.154	15.248

Given the scant data utilised to train the models, our system managed to perform fairly well. By averaging the accuracy of each module, we were able to determine the total accuracy of our attention mechanism. Our models performed better and required less time to infer conclusions when Caffe and OpenCV's DNN module were compiled with CUDA support, as demonstrated in Table (2). Our overall accuracy rate was 84.6233 percent.

Table II: System Performance

Module	*Accuracy*	*Inference time*
Facial Landmarks	89.67 %	0.033 ms
Blink rate detection	91.02 %	0.026 ms
Eye gaze tracking	75.33 %	0.032 ms
Emotion classification	82.55 %	0.057 ms
Facial recognition	90.11 %	0.052 ms
Body posture	79.06 %	0.048 ms
Overall System	**84.6233 %**	**0.258 ms**

VI. CONCLUSION AND FUTURE WORKS

Given the scant data utilised to train the models, our system managed to perform fairly well. By averaging the accuracy of each module, we were able to determine the total accuracy of our attention-tracking model. Our models performed better and required less time to infer conclusions when Caffe and OpenCV's DNN module were compiled with CUDA support, as demonstrated in Table 2. Our overall accuracy rate was 84.6233 percent. We picture the scores as a live chart and produce computerized reports. The input created can be utilized for:

1) Assessing understudy performance
2) Increasing exhibiting standards
3) Preventing carelessness during online evaluations

We can improve the presentation of our framework as a part of future projects by building our models using more data. Additionally, a similar consideration following component can be additionally advanced to at the same time work with different subjects in a homeroom utilizing video film from the CCTV cameras. Because we now lack a dataset for measuring the ability to focus while on the web addresses, we also stand out sufficiently to be considered scores as ground truth-values. A common dataset can make evaluating the framework's exhibition more reliable.

REFERENCES

1. ORA staff, Andover, Massachusetts; Mark B. Abelson, MD, CM, FRCSC It's Time to Consider the Blink, June 2011, Review of Ophthalmology
2. Minjuan Wang, Ruimin Shen, and Liping Shen, "Affective e-Learning: Using "Emotional" Data to Improve Learning in a Pervasive Learning Environment," International Forum of Educational Technology & Society 2009.
3. R. Bixler and S. D'Mello automatic, user-undependent detection of daydreaming when reading on a computer. 26, 33, and 68 User Model User-Adap Inter 2016. [CrossRef]
4. J. Smallwood, D. J. Fishman, and J. W. Schooler The cost of an absent mind: The underappreciated impact of mind wandering on academic performance. 230–236. Psychonomic Bulletin & Review, 14, 2007. [CrossRef]
5. A Survey of State-of-the-Art Drowsiness Detection Techniques, M. Ramzan, H. U. Khan, S. M. Awan, A. Ismail, M. Ilyas, and A. Mahmood, IEEE Access, vol. 7, no. 6, 2019, pp. 61904-61919. [CrossRef]
6. Zhen Zhang and Yuan Zhang, An Experimental Study on the Influence of Environmental Noise on Students' Attention, conference on noise 2018, EuroNoise.
7. 7."Real-time Smart Attendance System Using Face Recognition Techniques," 2019 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence), Noida, India, 2019, pp. 522-525. S. Sawhney, K. Kacker, S. Jain, S. N. Singh, and R. Garg.
8. Deep Sparse Representation Classifier for Facial Recognition and Detection System by Eric-Juwei Cheng, Kuang-Pen Chou, Shantanu Rajora, Bo-Hao Jin, M. Tanveer, Chin-Teng Lin, Ku-Young Young, Wen-Chieh Lin, and Mukesh Prasad.
9. Stanley, Darren, "Using Microsoft Kinect to gauge attention" (2013). Rochester Institute of Technology thesis
10. "Pupil recognition and tracking utilising various light sources," Image Vis. Comput, vol. 18, no. 4, pp. 331-336, 2008. C. Morimoto, D. Koons, A. Amir, and M. Flickner. [CrossRef]
11. An Emotion Recognition Model Based on Facial Recognition in Virtual Learning Environment by D. Yang, Abeer Alsadoona, P.W.C. Prasad*a, A.K. Singh, and A. Elchouemi.
12. One Millisecond Face Alignment using an Ensemble of Regression Trees, Vahid Kazemi and Josephine Sullivan, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2014, pp. 1867–1874.
13. Accelerating Viola-Jones Face Detection to FPGA-Level with G PUs by Daniel Hefenbrock, Jason Oberg, Nhat Tan Nguyen Thanh, Ryan Kastner, and Scott B. Baden

AUTHORS PROFILE

Kavyashree N, Assistant Professor, Department of Master of Computer Applications at Dr. Ambedkar Institute of Technology Bangalore. 12 years of Teaching experience. Published 10 papers and presented in 2 conferences. Attended the faculty development programs. I aim to motivate and help students to achieve academic excellence as well as extracurricular activities and teach them not only theoretical knowledge but also practical knowledge

Shailaja L K, Assistant Professor, Department of Master of Computer Applications at Dr. Ambedkar Institute of Technology, Bangalore. 14 years of Teaching experience. Published 10 papers. My strengths are disciplined, friendly with new environment and being Motivated. I am a person who is positive about every aspect of life. There are many things I like to do see, and to experience.

Anitha J, Assistant Professor Department of Master of Computer Applications at Dr. Ambedkar Institute of Technology, Bangalore. 14 years of Teaching experience. Published 10 papers. My strengths include eagerness to learn and use the knowledge in real world applications and problem solving. My aim is to gain knowledge regarding my interested fields and create projects.