COMP4801 Final Year Project

Detailed Project Plan

AI-Powered Attention Monitoring System for Enhancing Online Learning Engagement

Tam Cheung Yan Shojin; 3035927673; BENG(CompSc)

Supervisor: Prof. Zhao Hengshuang

1. Introduction

Online classes have become essential since the pandemic, enabling students to participate remotely. Educational institutions have increasingly adopted online learning platforms, which allow for greater flexibility and accessibility [1]. However, research indicates that students are more susceptible to distractions during these sessions, adversely affecting their engagement and learning outcomes [2]. This project proposes an innovative application that leverages advanced AI models to monitor and analyze student attentiveness through camera input, aiming to enhance online learning experiences.

2. Background

The shift to online learning has highlighted the challenges associated with maintaining student engagement. Various studies suggest that the lack of physical presence can lead to a decline in attentiveness [1]. Students may multitask, engage with their phones [3], or even fall asleep during lectures [4]. This project aims to address these issues by implementing a technology-driven solution that provides real-time feedback on student engagement.

2.1 Importance of Engagement in Online Learning

Engagement is crucial for effective learning. Research shows that students who are actively engaged in their learning process are more likely to retain information and perform better academically [5]. By monitoring attention levels, educators can tailor their teaching strategies to better meet the needs of their students.

3. Related works

Previous studies have explored various methodologies for monitoring student attentiveness, including:

Head Pose and Facial Landmark Tracking: Analyzing head orientation and facial features to gauge focus. For example, techniques that track head movement can indicate whether a student is looking at the screen [6].

Phone Detection: Identifying the presence of mobile devices to ascertain distractions. Research has shown that phone use during classes correlates with lower academic performance [7].

Eye Tracking: Monitoring eye movements to determine attention levels. Eye tracking can provide insights into where a student's focus lies during a lecture [8].

These approaches highlight the potential for integrating multiple data sources to create a comprehensive attentiveness monitoring system.

4. Objectives

The proposed application will be built on the OpenCV framework for real-time computer vision. It aims to integrate multiple AI models:

1. L2CS-Net [9]: A gaze detection model to assess where the student is looking.
2. Head Pose Estimation [10]: For detecting head orientation and facial landmarks.
3. Emotion Detection [11]: To analyze the student's emotional state.
4. YOLOv8 [12]: For real-time detection of phones and monitoring student actions.

These models will be trained on diverse datasets to enhance their predictive capabilities. The outputs will be synthesized to calculate a comprehensive attentiveness score for each student. A user-friendly interface will display these results alongside insights derived from the model predictions.

5. Methodology

5.1 L2CS-Net

The model predicts the gaze direction of a student [9]. Attentiveness will be assumed if the student is looking forward; while inattentiveness will be assumed if the student is looking sideways, upward, downward, or backward.

5.2 Head Pose Estimation

This model detects head pose, and facial landmarks [10]. For the facial landmarks, the model predicts 68 facial landmarks as proposed by this database [13]. Eye aspect ratio and yawn aspect ratio will be calculated, which are indicators the drowsiness of the person, and will be used to calculate attentiveness.

As for head pose, similarly for gaze direction, attentiveness will be determined for the direction of the head pose.

5.3 Emotion Detection

This model predicts the emotions of the student [11]. The model categorizes emotions as one of the seven classes: angry, disgusted, fearful, happy, neutral, sad, and surprised. Attentiveness will be calculated based on the affective model showed in figure 1 [14].

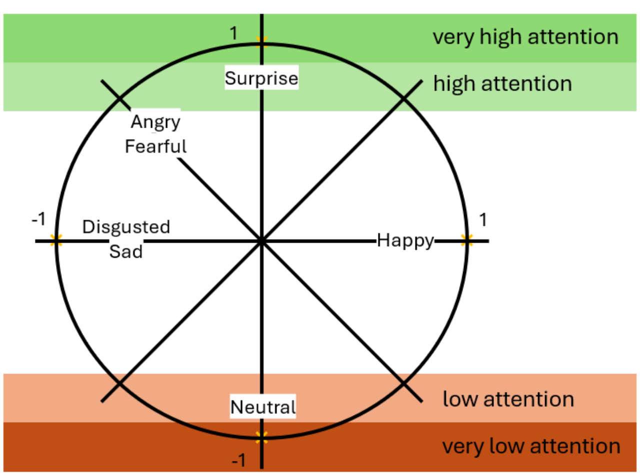

Figure 1 Affection Model modified from [14] to match with classes of Emotion Detection

We calculate the attentiveness as follows [15]:

$$S_1 = S_{Happy} + (-1) \cdot S_{Sad}$$

Equation 1

$$S_2 = S_{Surprise} + (-1) \cdot S_{Neutral}$$

Equation 2

$$S_3 = S_{Disgusted} + S_{Fearful}$$

Equation 3

Where $S = \{S_{Happy}, S_{Sad}, S_{surprise}, S_{Neutral}, S_{Disgusted}, S_{Fearful}\}$ are the percentages of each of the emotions predicted by the model. The algorithm then calculates attentiveness x:

$$x = \sqrt{\left(\overline{S_1} + \frac{\overline{S_3}}{\sqrt{2}}\right)^2 + \left(\overline{S_2} + \frac{\overline{S_3}}{\sqrt{2}}\right)^2} \cdot \frac{\overline{S_2} + \frac{\overline{S_3}}{\sqrt{2}}}{\left|\overline{S_2} + \frac{\overline{S_3}}{\sqrt{2}}\right|} + 1$$

Equation 4

5.4 Yolov8

Yolov8 is a detection model suitable for real time detection [12]. It is selected due to its performance. The base model will be used to train two detection models: one for detecting phones, and the other for detecting student actions.

The phone detection model will be trained mainly with MSCOCO [16]. The student action detection model will be trained with EduNet dataset [17] , pending permission from the dataset owner.

5.5 Overall attentiveness calculation

The overall attentiveness is the mean of all the parameters calculated with the above models, as suggested in [8].

$$Attentiveness = \frac{\Sigma\, parameters}{n} * 100$$

5.6 Dataset preparation

Most AI models come with pretrained versions, except for the student action detection model, which will be trained using the EduNet dataset [17]. Training with additional datasets online can enhance accuracy. However, finding datasets that match the output format of the pretrained models may prove challenging. For instance, the Head Pose Estimation model predicts 68 facial landmarks, making it difficult to find new datasets that have not already been used in training the fine-tuned model.

5.7 Program coding

The program will be implemented using Python within the OpenCV framework, which is known for its efficiency and flexibility in handling computer vision tasks. Python will be used for ease of programming.

5.8 UI

The UI will be developed using Qt, providing a user-friendly experience. It will display model predictions, overall attentiveness scores, and visual feedback on student engagement. If inattentiveness is detected, the UI will offer actionable tips and strategies for improving focus, such as reminders to take breaks or suggestions for reducing distractions in the environment.

6. Schedules and milestones

Semester 1

Oct 1, 2024	Deliverables of Phase 1 • Detailed project plan
Nov 1, 2024	Acquire permission to use EduNet dataset Prepare other datasets
Dec 1, 2024	Finish training the ai models
Jan 1, 2025	Finalize first presentation and interim report

Semester 2

Jan 13-17, 2025	First presentation
Jan 26, 2025	Deliverables of Phase 2 • Preliminary implementation • Detailed interim report
Feb 1, 2025	Complete implementation of attentiveness calculation
Mar 1, 2025	Complete UI implementation
Mar 10, 2025	Finish testing
Apr 1, 2025	Complete final report
Apr 20, 2025	Finish presentation preparation and video
Apr 21, 2025	Deliverables of Phase 3 • Finalized tested implementation • Final report

Apr 22-26, 2025	Final presentation
Apr 29, 2025	Prepare for project exhibition
Apr 30, 2025	Project exhibition • 3-min video

References

[1] C. B. Hodges, S. Moore, B. B. Lockee, T. Trust and M. A. Bond, "The difference between emergency remote teaching and online learning," 2020.

[2] K. A. Aivaz and D. Teodorescu, "College Students' Distractions from Learning Caused by Multitasking in Online vs. Face-to-Face Classes: A Case Study at a Public University in Romania," *Int J Environ Res Public Health,* vol. 18, 2022.

[3] R. Junco, "In-class multitasking and academic performance," *Computers & Education,* vol. 59, no. 2, pp. 505-514, 2012.

[4] A. J. Mendez, "The effects of online learning on student engagement: A study on sleepiness," *Journal of Online Learning Research,* vol. 6, no. 1, pp. 45-63, 2020.

[5] J. Fredricks, P. C. Blumenfeld and A. Paris, "School engagement: Potential of the concept, state of the evidence.," *Review of Educational Research,* vol. 74, no. 1, pp. 59-109, 2004.

[6] M. Elbawab and R. Henriques, "Machine Learning applied to student attentiveness detection: Using emotional and non-emotional measures," *Educ Inf Technol,* vol. 28, p. 15717–15737, 2023.

[7] M. K. Hossen and M. S. Uddin, "Attention monitoring of students during online classes using XGBoost classifier," *Computers and Education: Artificial Intelligence,* vol. 5, 2023.

[8] K. Vignesh, V. D. N. Gupta and H. Kumar, "Attentiveness Recognition System for Online Classes using OpenCV," in *2022 2nd International Conference on Intelligent Technologies (CONIT)*, Hubli, India, 2022.

[9] A. A. Abdelrahman, T. Hempel, A. Khalifa, A. Al-Hamadi and L. Dinges, "L2cs-net: Fine-grained gaze estimation in unconstrained environments," in *2023 8th International Conference on Frontiers of Signal Processing (ICFSP)*, 2023.

[10] G. Ying, "Head Pose Estimation," [Online]. Available: https://github.com/yinguobing/head-pose-estimation. [Accessed 30 September 2024].

[11] A. Balaji, "Emotion Detection," [Online]. Available: https://github.com/atulapra/Emotion-detection. [Accessed 30 September 2024].

[12] D. Reis, J. Kupec, J. Hong and A. Daoudi, "Real-Time Flying Object Detection with YOLOv8," 2023.

[13] C. Sagonas, E. Antonakos, G. Tzimiropoulos, S. Zafeiriou and M. Pantic, "300 Faces In-The-Wild Challenge: database and results," *Image and Vision Computing,* vol. 47, pp. 3-18, 2016.

[14] K. Altuwairqi, S. K. Jarraya, A. Allinjawi and M. Hammam, "A new emotion–based affective model to detect student's engagement," *Journal of King Saud University - Computer and Information Sciences,* vol. 33, no. 1, pp. 99-109, 2021.

[15] L. Q. Thao, D. T. Kien, N. C. Bach, D. T. T. Thuy, L. T. M. Thuy, D. D. Cuong, N. H. M. Hieu, N. H. T. Dang, P. X. Bach and L. P. M. Hieu, "Monitoring and improving student attention using deep learning and wireless sensor networks," *Sensors and Actuators A: Physical,* vol. 367, 2024.

[16] T.-Y. Lin, M. Maire, S. Belongie, L. Bourdev, R. Girshick, J. Hays, P. Perona, D. Ramanan, C. L. Zitnick and P. Dollár, "Microsoft coco: Common objects in context," in *Computer Vision–ECCV 2014: 13th European Conference*, Zurich, Switzerland, 2014.

[17] V. Sharma and M. Gupta, "Classroom-Monitoring-Action-Dataset," [Online]. Available: https://github.com/vijetait/Classroom-Monitoring-Action-Dataset?tab=readme-ov-file. [Accessed 30 September 2024].