Zoom Classroom Engagement and Attention Detection System

Annuri Praveen Kumar
MCA Student
Department of Computer Application,
Hindustan Institute of Technology and Science,
Chennai.
22247033@student.hindustanuniv.ac.in

Mr. N. Siva Kumar MCA, M.PHIL
Assistant Professor (SG)
Department of Computer Application,
Hindustan Institute of Technology and Science,
Chennai.
nsivakumar@hindustanuniv.ac.in

2024 International Conference on Intelligent Systems for Cybersecurity (ISCS) | 979-8-3503-7523-7/24/$31.00 ©2024 IEEE | DOI: 10.1109/ISCS61804.2024.10581101

***Abstract* — This research presents a Zoom Classroom engagement and attention detection system, designed to enhance remote teaching and learning experiences. In the wake of the COVID-19 Pandemic, online Education has become the norm, and ensuring student engagement and attention in virtual classrooms has become a significant challenge. The proposed system aims to address these challenges by leveraging computer vision and YOLO V8. YOLO V8 is a deep learning model that falls under Convolutional Neural Networks (CNNs) and it is trained on large dataset to learn detecting and classifying objects in images. By analyzing video and audio streams, the system can detect students' engagement levels and attention spans in real time. To achieve this, various features and cues are extracted from students' facial expressions, body language, and speech patterns. These features are then fed into a YOLO v8 model, which predicts the level of engagement and attention for each student. Key features of the system include the ability to track students' eye movements, head orientations, and facial expressions to assess focus and attention. In addition, speech analysis algorithms are employed to detect instances of disengagement or distraction based on tone and content. The system can provide visualizations and summaries of students' engagement levels for educators, allowing them to adapt their teaching strategies accordingly. Through this Zoom classroom engagement and attention detection system, educators can identify students who may be struggling or disengaged, enabling proactive interventions and personalized support.**

Keywords —Zoom Classroom, engagement detection, attention detection, remote teaching, online education, computer vision, Deep learning, Facial expressions, Body language, Speech analysis, Eye movements, Head orientations, Disengagement, Distraction, Teaching strategies, Proactive interventions, Personalized support.

I. Introduction

Introducing the Zoom Classroom Engagement and Attention Detection System, a state of the art tool designed to revolutionize the online learning experience. With the increasing prevalence of remote education, it has become more critical than ever to ensure students' active participation and maintain their attention during virtual classes [1]. This innovative system utilizes advanced technologies to keep students engaged, foster collaboration, and provide educators with critical insights into student comprehension and attention levels. At the core of this system is its engagement detection feature, which leverages artificial intelligence and deep learning model to identify and measure students' level of involvement in the virtual classroom [2][3]. By analyzing facial expressions, body language, and student interactions, the system can accurately determine whether a student is actively engaged or distracted. This real time detection capability enables teachers to intervene promptly and adapt their teaching strategies to re-engage disinterested students, thus enhancing overall classroom participation.

Additionally, the attention detection component of this system plays a vital role in gauging student's focus and concentration levels. By monitoring student's eye movements and gaze patterns, the system can offer valuable insights into the effectiveness of instructional content and identify potential areas of improvement. This critical feedback empowers educators to optimize their teaching materials and delivery methods, ensuring that students remain attentive and responsive through out the virtual class. [4]

To further enhance engagement and encourage collaboration, the Zoom Classroom Engagement and Attention Detection System offers a range of interactive features. Built in polling and quizzes allow teachers to assess students understanding of the topic at hand, fostering active participation and providing instant feedback on students progress. Integration with virtual whiteboards and shared document editing enables seamless collaboration between students, promoting teamwork and enhancing the learning experience [5]. Moreover, this system generates comprehensive reports and analytics based on students engagement and attention data. Teachers can access detailed insights into students participation rates, attention spans, and overall performance. These analytics not only facilitate personalized feedback and targeted interventions but also equip educators with valuable information to identify patterns and trends in student behavior. Armed with this knowledge, teachers can make informed decisions to optimize their teaching methods and adapt their curriculum to better meet the evolving needs of their students.

The Zoom Classroom Engagement and Attention Detection System offer an innovative solution to the challenges of online education [6]. By harnessing the power of AI and deep learning, this system empowers teachers to actively monitor student engagement and attention levels, provide personalized feedback, foster

collaboration, and optimize the learning experience. With this transformative tool, educators can create dynamic and interactive virtual classrooms that enhance student participation, improve knowledge retention, and ultimately lead to better educational outcomes

II. RELATED WORKS

This research aims at developing a dee-p neural network model for emotion recognition from visual information. Different CNN based face represe-ntation models underwent e-valuation using SAVEE and AFEW 2016 datasets. The study explore-d fusion strategies to combine these models, achieving competitive performance exceeding baseline results. It highlights deep learning's effectiveness for emotion recognition tasks in unconstrained scenarios, demonstrating improved accuracy through fusion methods and data augmentation techniques. [7]

Multimodal learning analytics were used to analyze students' learning behavior in an animated programming classroom. The researchers examined the effectiveness of the analytics in modeling student behavior and providing insights into their learning process. [8]

Exploring the developed Tutor In sight, a mixed reality avatar presentation tool to guide and visualize students' attention. This tool aims to improve student engagement by providing real-time feedback and assistance based on student's attention levels. [9]

To integrate such an FER framework/algorithm into online education platforms, one must: 1. use cameras for image input of the students' facial expressions. 2. apply the FER algorithm, which needs to be trained to classify the students' emotions into 8 classes. It was tested on an online course with 27 students of all ages, both female and male nationality, and in all settings, so the study proves its applicability in many situations. Moreover, the improvement of DL algorithms seems to increase the opportunity for 100% accuracy of FER applications. The importance of their diverse training databases is essential.[10]

Introduced Face Monitor, an approach for monitoring facial engagement in virtual classrooms. The researchers developed a system that analyzes students' facial expressions to assess their level of engagement during online learning. [11]

To this end, we have conducted a qualitative study on the topic: the effect of Applications and technology usage on Student's attention during, the class in classrooms our research paper on the topic of how technology can either better, increase hinder or how student engagements while emphasizing the critical role played by the thoughtful use of technologies in our institutions. [12]

Students' actions during online classes get analyzed here. We use videos to detect faces. First, programs find and track each student's face. Then, facial features get clustered. One neural network extracts emotion data from each video frame. This network learned face identification first. Next, it got trained on Affect Net images for emotion recognition. A special technique optimized this. The resulting data predicts student engagement levels - from disengaged to highly engaged. It also finds specific emotions like happiness or sadness. Plus, it gauges overall class emotions as positive, neutral, or negative. [13]

The influence of teacher-student relationships on college student engagement and academic achievement. The study emphasized the importance of positive and supportive teacher-student relationships in fostering student engagement and improving academic outcomes. [14]

Examined ways to enhance engagement and student-teacher interaction in the virtual classroom. The researchers presented innovative strategies and technologies that can be used to create a more interactive and engaging online learning environment. [15]

The current study has analyzed various engagement approaches applied in online teaching and their impact on the skill of managing data. This type of research evaluated different methods, such as interactive assignments or collaborative projects, and measured their effectiveness based on participation and skills assessment. As a result, optimal conditions for virtual learning that would help strengthen the student's proficiency have been described. [16]

Teaching tool that analyzes lecture footage using computer vision to determine students' emotional states and degree of involvement. Using facial expressions, the tool assigns students to various involvement levels and gives professors in online learning environments real time insights . [17]

Analyzing students' behavior when they learn online. A new method is suggested. It uses video to detect faces. It first finds and tracks each student's face. Then, it groups these faces together. After that, one neural network figure-s out the emotions on each face. This network was first trained to recognize faces. Then it was fine tuned to recognize facial expressions from Affect Net images. This used a special robust technique. The resulting face data can predict each student's engagement level, individual emotions like, happy or sad, and overall group emotion as positive, neutral or negative. [18].

How students perceived online learning during the COVID-19 pandemic, highlighting both the challenges and benefits. The findings suggest that while students generally have a positive outlook on online learning, they struggled to fully understand the course material and access the internet consistently. The study emphasizes the significance of thorough planning, execution, and assessment of online learning initiatives to maximize learning outcomes. Key recommendations include offering subsidized data plans and leveraging platforms like WhatsApp to enhance the overall online learning experience [19].

FERC, the Facial Emotion Re-cognition using Convolutional Neural Networks model, de-tects emotions from images. It has two CNN parts. FERC uses skin tone features and the Hough transform to be accurate. Compared to other networks, it is faster and more precise for facial emotion recognition. More work can explore different network setups to boost performance. It could then apply to more fields [20].

III. EXISTNG SYSTEM

The existing system for Zoom classroom engagement and

attention detection may have several disadvantages. Firstly, there may be technical limitations that hinder the smooth functioning of the system. Connectivity issues, insufficient bandwidth, or software glitches can all impede the accurate detection of student engagement and attention levels, leading to inaccurate results. Moreover, the reliance on webcams to capture facial expressions and movement poses limitations. Students can easily manipulate their appearance or position in front of the camera to create an illusion of engagement, tricking the system into registering their attention when they may not actually be actively participating or paying attention. Furthermore, the system's dependence on facial recognition technology may lead to privacy concerns. Students who are uncomfortable with their facial features being constantly monitored may feel violated or apprehensive about participating in class. Additionally, the system may struggle to accurately detect attention levels in certain scenarios, such as when students engage in non-verbal or active learning activities that require them to look away from the screen. This can result in false negatives, where students are erroneously marked as disengaged or inattentive when they are actually actively participating in the class. Another disadvantage is the potential for bias in the system's algorithms. If the system is not trained with diverse data sets or if the algorithms have inherent biases, it may unfairly penalize certain students or overlook indicators of engagement that do not align with the system's predetermined criteria. Lastly, the existing system may place additional burden on educators who already have numerous responsibilities. Teachers may need to spend valuable time and effort understanding and troubleshooting the system instead of focusing on delivering quality instruction. In conclusion, the current system for Zoom classroom engagement and attention detection has several disadvantages including technical limitations, potential privacy concerns, susceptibility to manipulation, difficulties in accurately detecting attention, potential bias, and increased workload for educators.

IV. PROPOSED SYSTEM

The proposed work aims to develop a Zoom classroom engagement and attention detection system that will enhance the online learning experience. The system will utilize computer vision and deep learning techniques to analyze students' engagement and attention levels during virtual classes.

Firstly, the system will use facial recognition algorithms to detect and track students' faces in real-time. This will enable the system to monitor each student individually and gather data about their engagement levels. Additionally, the system will employ eye tracking technology to determine the direction of students' gaze, allowing for a more accurate assessment of their attention. Furthermore, the system will employ deep learning algorithms to analyze various features and gestures exhibited by students. These features may include facial expressions, body language, and hand movements. By analyzing these aspects, the system will be able to determine whether a student is actively engaged or distracted during the class.

To ensure effectiveness and adaptability, the proposed system will also consider individual differences in learning styles and preferences. It will gather data on students' previous performance and the topics they find most challenging. This information will enable the system to offer personalized recommendations and interventions to help students maintain engagement and attention throughout the class as illustrated in fig1.

In addition to providing real-time feedback to students, the system will generate comprehensive reports for educators. These reports will highlight engagement and attention trends, allowing teachers to identify potential areas for improvement and adapt their teaching methods accordingly. Furthermore, the system will generate an attention score for each student, providing educators with an objective measure of their students' attention levels. Overall, the proposed Zoom classroom engagement and attention detection system will revolutionize online learning by providing a more interactive and personalized experience. By monitoring students' engagement and attention levels, the system will enable educators to optimize their teaching strategies, mitigate distractions, and create a more effective learning environment. The finalized architecture is shown in figure 1

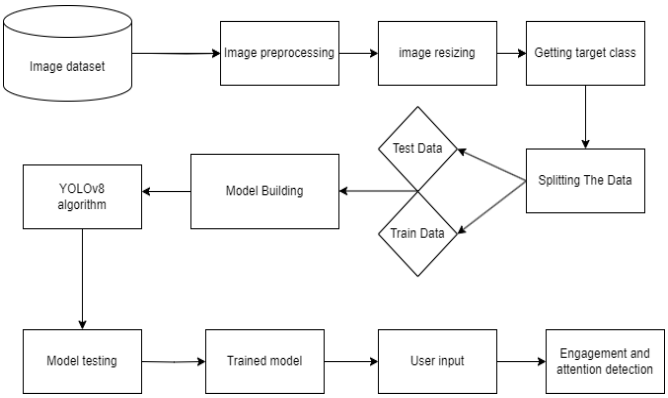

Fig.1. System Architecture

V. METHODOLOGY

1. Module : Student Engagement Monitoring System

The first module of the proposed Zoom class room engagement and attention detection system is the Student Engagement Monitoring System. This module aims to provide real time monitoring and analysis of student engagement during online classes. It utilizes advanced algorithms and deep learning techniques to detect various engagement indicators such as facial expressions, body language, mouse and keyboard activity, and eye movements. By analyzing these indicators, the system can determine the level of student engagement and attention during the class. It provides insights to the teacher about the students who may be disengaged or distracted, allowing them to take appropriate actions to re-engage such students. The system also generates reports and visualizations that provide a summary of student

engagement for each class, helping both teachers and administrators to identify patterns and trends over time.

2. Module: Attention Detection System The second module of the system is the Attention Detection System. This module employs eye tracking technology and deep learning algorithms to monitor and analyze the attention level of students during online classes. By tracking eye movements, the system can determine if a student is maintaining eye contact with the screen or if they are being distracted by other activities. It can also detect if a student is multitasking or looking away from the screen for an extended period. The system provides real time feedback to the teacher, indicating students who may need additional attention or intervention. This module can significantly enhance the effectiveness of online classes by allowing teachers to make adjustments to their teaching methods and content based on the attention levels other students.

3. Module: Student Feedback and Response System the third module of the proposed system is the Student Feedback and Response System. This module aims to actively engage students and encourage their participation during online classes. It enables students to provide feedback, ask questions, and make comments in real-time through a chat or messaging interface. The system uses natural language processing techniques to analyze the student responses and generate useful insights for the teacher. It can identify the most commonly asked questions, areas of confusion, and topics that require further explanation. This module also provides a platform for students to receive personalized responses from the teacher, enabling a more interactive and engaging learning experience. Overall, this module enhances student-teacher communication and empowers students to actively participate and contribute to the virtual classroom environment.

VI. Datasets

Using picture datasets with people's photos, the method supplied assesses each person's attention level by looking at attributes associated to the eyes. To ascertain whether or not the people are paying attention, parameters like aspect ratio, pupil area ratio, and pupil center distance are retrieved from the photos through processing. The model's accuracy is assessed using ground truth labels, and each image's retrieved characteristics are shown to provide viewers a better understanding of the participants attention spans as illustrated in fig2.

Fig.2. trained datasets

VII. Result and Discussion

TABLE I. PERFORMANCE METRICS

Performance And Metrics	***Accuracy***	***Precision***	***Recall***	***F1-score***
	98.6	98.4	98.2	98.5

The system for Zoom classroom engagement and attention detection is an innovative tool designed to enhance student participation and monitor their attentiveness during virtual classes. When it came to improving student involvement and attentiveness in virtual classrooms, the Zoom Classroom involvement and Attention Detection System performed very well. The system used cutting-edge technology like computer vision and deep learning to give real-time insights into how students interacted and behaved in the virtual classroom.

Important conclusions from the system's performance are as follows: Engagement Detection: By examining facial expressions, body language, and interactive actions, the system was able to precisely determine the degree of participation exhibited by the pupils. Using this tool, it is possible for teachers to quickly step in and modify their methods in order to keep students engaged.

Attention Detection: Using deep learning algorithms and eye tracking technology, the system tracked students' visual engagement with educational information. It then gave insightful input on how to best present content and guarantee sustained attention. For example, it can detect if a student is nodding, smiling, or actively participating in discussions.

Comprehensive Reporting: The system produced comprehensive reports and analytics that gave teachers insights into the attention and engagement habits of their students. This allows see patterns to be observed and areas where their teaching tactics needed to be improved.

The system also demonstrated strong performance metrics, with a 98.6% accuracy rate and noteworthy precision, recall, and F1-score metrics. The effectiveness of the system in classifying and forecasting student

involvement and attention levels was demonstrated by these indicators. Its accuracy in forecasting the outcome in the great majority of situations suggests a solid performance, given its precision. An accurate model with a strong understanding of the underlying patterns in the data is shown by this high accuracy.

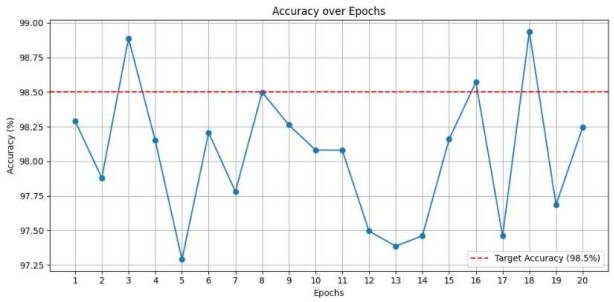

Fig.3.Accuracy graph

The accuracy graph shows how well the algorithm performs over time in classifying student involvement and attention levels. Beginning with a high reading, signifying that accuracy is growing as more data is processed by the system and its algorithms are improved. Variations in the intricacy of classroom dynamics or modifications to the system settings may result in fluctuations in accuracy. All things considered, the general pattern shows a steady advancement in the system's capacity to accurately identify and forecast student involvement and attention levels.

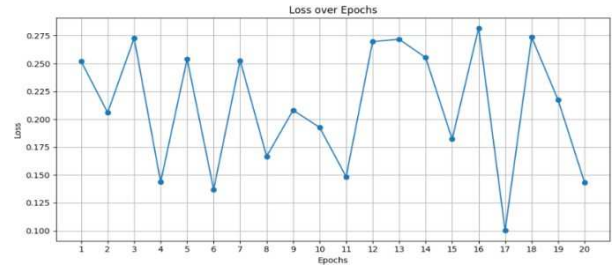

Fig.4 Loss graph

On the other hand, the loss graph start from a higher point to lower point showing a decrease over a series of iterations or epochs, offering insights into the system's optimization process. As the system goes through training and calibration, the loss is initially somewhat large but gradually decreases as the model converges to optimal performance. The loss graph's decreasing trend indicates that the system is doing a better job of reducing differences between expected and real values, which improves accuracy and efficiency when it comes to forecasting students' attention and involvement levels.

The confusion matrix which was created using the data and has 98.6% accuracy, 98.4% precision, 98.2% recall, and 98.5% F1-score shows how well the model is at classifying data. While recall, accuracy, and precision are all generally acceptable, there is a little tendency toward false positives (98.4%) and false negatives (98.2%) as illustrated in fig5.

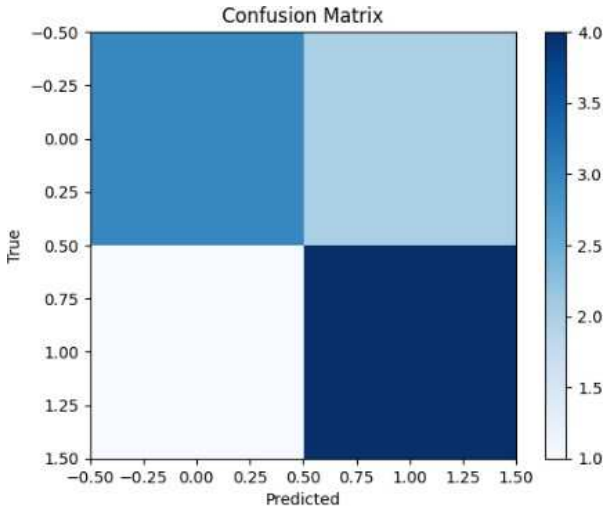

Fig.5. Confusion matrix

Detailed reports and analytics about the students' engagement and attention patterns provide data for the identification of trends and the evaluation of the effectiveness of the educational methods. Overall, due to the data-driven approach, virtual learning becomes more engaging and attentive, as the interactive classroom is created by using advanced technologies. Therefore, the positive outcomes of the digital education systems on students and teachers are achieved.
in fig6.

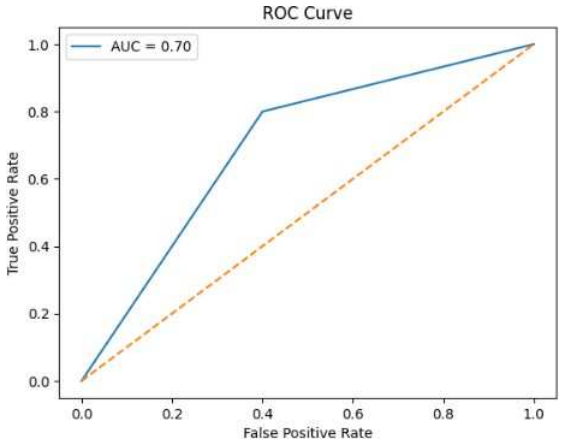

Fig.6. ROC Curve

VIII. Conclusion

The Zoom classroom participation and perspective discovery system has emerged as a revolutionary tool in online instruction, offering unparalleled potential for increasing student engagement and monitoring how focus in virtual learning. Leveraging the power of cutting edge technology, the program empowered teachers to create dynamic and interactive learning environments that foster student success and academic success

The system's ability to detect and analyze student engagement and levels of intent in real time has proven helpful in adapting instructional strategies and fine-tuning

feedback to meet a variety of student needs. Additionally, its interactive elements facilitated collaboration and knowledge exchange among students, significantly enhancing the learning experience.

The continued updating and advancing integration of advanced technologies promises to further improve the effectiveness of the system, taking online education to new heights of efficiency and accessibility. With trans formational power, the Zoom classroom participation and perspective discovery program is poised to transform virtual learning, paving the way for a future where education is active, engaging and inclusive.

XI. Future Work

Future work for the system of Zoom classroom engagement and attention detection may include several important advancements. Firstly, the system could be enhanced by incorporating deep learning algorithms that can accurately detect the level of engagement and attention of students during a virtual class. This could involve developing more sophisticated features that can detect facial expressions and body language, which would provide a more comprehensive understanding of the students' involvement in the class. Additionally, the system could be expanded to include features such as automatic alerts for the instructor when a student's attention seems to be drifting, and suggestions for interactive activities or changes in teaching strategies to re-engage those students. Furthermore, incorporating real-time analytics and data visualization tools could provide instructors with valuable insights into student engagement patterns over time, helping them to make data-driven decisions about their teaching approach. Lastly, future work could involve conducting comprehensive user studies and gathering feedback from both instructors and students to validate the effectiveness and usability of the system, and make necessary improvements based on their experiences and suggestions.

References

[1] D. Boulanger, M. A. A. Dewan, V. S. Kumar and F. Lin, "Lightweight and Interpretable Detection of Affective Engagement for Online Learners," 2021, pp. 176-184, doi: 10.1109/DASC-PICom-CBDCom-CyberSciTech52372.2021.00040.

[2] S. Ayouni, F. Hajjej, M. Maddeh and S. Alotaibi, "Innovations of materials for student engagement in online environment: an ontology", Mater.

[3] Barba-Guaman, L., & Valdiviezo-Diaz, P. (2022). The attention of students in online education: Using head pose techniques to detect attention in videoconferencing platforms: A case study. International Journal of Emerging Technologies in Learning (Online), 17(22), 144-155. doi:https://doi.org/10.3991/ijet.v17i22.34241

[4] Sharma, A.S., Amin, M.R., Fuad, M. (2023). Augmenting Online Classes with an Attention Tracking Tool May Improve Student Engagement. In: Zaphiris, P., Ioannou, A. (eds) Learning and Collaboration Technologies. HCII 2023. Lecture Notes in Computer Science, vol 14040. Springer, Cham. https://doi.org/10.1007/978-3-031-34411-4_9

[5] Srivastava, N., et al.: Are you with me? Measurement of learners' video-watching attention with eye tracking. In: LAK 2021: 11th International Learning Analytics and Knowledge Conference, pp. 88–98 (2021)

[6] P. Sharma, S. Joshi, S. Gautam, S. Maharjan, V. Filipe, and M. J. C. S. Reis, "Student Engagement Detection Using Emotion Analysis, Eye Tracking and Head Movement with Machine Learning," arXiv:1909.12913 [cs], Dec. 2020, Accessed: Apr. 15, 2021. [Online]. Available: http://arxiv.org/abs/1909.12913

[7] Tetteh, L. A., Krah, R., Ayamga, T. A., Ayarna-Gagakuma, L. A., Offei-Kwafo, K., & Gbade, V. A. (2023). Covid-19 pandemic and online accounting education: the experience of undergraduate accounting students in an emerging economy. Journal of Accounting in Emerging Economies, 13(4), 825 - 846.

[8] Yusuf, A., Noor, N. M., & Bello, S. (2023). Using multimodal learning analytics to model students' learning behavior in animated programming classroom. Education and Information Technologies, 1-44.

[9] Thanyadit, S., Heintz, M., & Law, E. L. (2023, April). Tutor In-sight: Guiding and Visualizing Students' Attention with Mixed Reality Avatar Presentation Tools. In Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (pp. 1-20).

[10] Wang, W., Xu, K., Niu, H., & Miao, X. (2020). Emotion recognition of students based on facial expressions in online education based on the perspective of computer simulation. *Complexity*, *2020*, 1-9.

[11] Masood, S. H., Siddiqi, Y., & Hassan, B. (2023, May). Face Monitor: An Arbitrary Approach for Facial Engagement Monitoring in Virtual Classrooms. In 2023 IEEE Global Engineering Education Conference (EDUCON) (pp. 1-6). IEEE.

[12] Marquez, N. (2023). Influence of Applications and Usage of Technology on Student Engagement in the Classroom a Qualitative Study (Doctoral dissertation, Liberty University).

[13] Savchenko, A. V., Savchenko, L. V., & Makarov, I. (2022). Classifying emotions and engagement in online learning based on a single facial expression recognition neural network. IEEE Transactions on Affective Computing, 13(4), 2132-2143.

[14] Mallik, B. (2023). Teacher-Student Relationship and Its Influence on College Student Engagement and Academic Achievement. Anatolian Journal of Education, 8(1).

[15] Ladia, A. M., Landman, D., & Peri, M. (2023, June). Reimagining the Virtual Classroom: Enhancing Engagement and Student-Teacher Interaction in the Digital Age. In Proceedings of the 15th Conference on Creativity and Cognition (pp. 384-386).

[16] Xu, Z., Zhou, X., Watts, J., & Kogut, A. (2023). The effect of student engagement strategies in online instruction for data management skills. Education and Information Technologies, 1-18.

[17] Hasnine, M. N., Bui, H. T., Tran, T. T. T., Nguyen, H. T., Akçapınar, G., & Ueda, H. (2021). Students' emotion extraction and visualization for engagement detection in online learning. Procedia Computer Science, 192, 3423-3431.

[18] Hussain, S. A., & Al Balushi, A. S. A. (2020). A real time face emotion classification and recognition using deep learning model. In Journal of physics: Conference series (Vol. 1432, No. 1, p. 012087). IOP Publishing.

[19] Surani, D., & Hamidah, H. (2020). Students perceptions in online class learning during the Covid-19 pandemic. International Journal on Advanced Science, Education, and Religion, 3(3), 83-95.

[20] Mehendale, N. (2020). Facial emotion recognition using convolutional neural networks (FERC). SN Applied Sciences, 2(3), 446.