Edge AI-Based Smart Classroom with Dynamic Student Attentiveness Monitoring

Nitish Ramaraj
School of Computer Science and Engineering (SCOPE)
Vellore Institute of Technology
Vellore, India
nitishramaraj@gmail.com

Girish Murugan
School of Computer Science and Engineering (SCOPE)
Vellore Institute of Technology,
Vellore, India
girish06062004@gmail.com

Himneesh Goyal
School of Computer Science and Engineering (SCOPE)
Vellore Institute of Technology,
Vellore, India
himneesh9@gmail.com

Tejas Anil
School of Computer Science and Engineering (SCOPE)
Vellore Institute of Technology
Vellore, India
tejasanil6@gmail.com

R.Rajavarman
Dept of Computer Science and Engineering
K.Ramakrishnan College of Technology,
Trichy, India
rajavarmanr.cse@krct.ac.in

Dr. Vetriselvi T
School of Computer Science and Engineering (SCOPE)
Vellore Institute of Technology,
Vellore, India
vetriselvi.t@vit.ac.in

***Abstract*— The traditional classroom mode of learning remains highly effective for student learning. However, with the increasing presence of digital gadgets it has increased distractions for students making it challenging for students to maintain their focus levels in the classroom. Also, with larger class sizes teachers also find it difficult to monitor and engage with students effectively. Using Edge AI technology our system works on these issues by providing real-time insights into student activity in the classroom. Running locally on the teacher's laptop our system uses a containerized AI model to analyze live camera feeds and track attentiveness. Our system allows teachers to input students overall marks which is used to dynamically adjust the threshold level of attention based on past grades and performance. This personalized approach supports each student's learning journey. By addressing these challenges our system enhances the classroom environment and makes learning more effective.**

Keywords—Edge-AI, Containerization, Media-Pipe, Pose Estimation, Eye Gaze Tracking, OpenCV.

I. INTRODUCTION

Classroom education remains as the benchmark for effective learning, providing students with opportunities for engagement and gaining knowledge. From our survey findings, it was observed that nearly 94.5% of students responded that learning in a classroom-based mode is more effective and would always prefer it over the rest. The well-organized learning environment of a classroom provides students in person interaction with teachers and students, promoting teamwork, critical thinking, and communication skills. During the COVID-19 pandemic, we saw the difficulties students faced when they had to move towards learning outside the classroom. Being in a new environment away from the usual classroom environment caused a lot of challenges for students as they found it difficult to stay focused and concentrated on the class without the in-person presence of students and teachers. Likewise, teachers faced obstacles in effectively evaluating their students' progress and modifying their way of teaching accordingly. The lack of face-to-face interaction made it difficult for teachers to understand students' understanding and provide real time feedback which led to a less efficient learning experience overall. This shows the invaluable role that classroom-based learning plays in the education sector.

In recent years statistics have highlighted the growing challenge of student distraction within the traditional classroom environment. According to a survey conducted by us within a school, roughly 86.5% of teaching staff report an increase in distraction amongst students due to the presence of digital devices such as smartphones and tablets. Also, 94.2% of teachers indicate that makingstudents focus during classes has become more challenging with larger class sizes, creating concerns about the level of student engagement, and learning outcomes. Further, our findings show that students multitask during class, with 65% accepting to using their electronic devices for non-academic purposes such as social media browsing or gaming. As distractions continue to increase in the classroom, educators are faced with the urgent need to implement strategies and technologies to mitigate their impact and create a more effective learning environment.

II. LITERATURE REVIEW

The review delves into a comprehensive examination of key studies that shed light on various facets of classroom dynamics and student engagement. Anderson et al. (2016) [3] shed light on the impact of digital device usage in classrooms, emphasizing the challenges posed by distractions to student concentration and learning outcomes. Brown & Davis (2020) [2] explored the hurdles faced by educators in larger classrooms, focusing on the imperative of maintaining student attention and its ramifications on teaching effectiveness. In contrast, Chen et al. (2021) [1] pioneered systems aimed at bolstering classroom dynamics through real-time student engagement monitoring, leveraging cutting-edge AI technology. Similarly, Gupta & Sharma (2018) [4] investigated the efficacy of edge computing in analyzing student engagement, highlighting its effectiveness in providing timely insights into student behavior. Meanwhile, Johnson & Smith (2018) [5] conducted a systematic review on digital distractions, accentuating the pressing need for strategies to mitigate distractions and optimize learning environments. Jones & Brown (2019) [6] contributed a meta-analysis on traditional classroom learning, unraveling factors that influence student learning outcomes. Additionally, Kumar et al. (2020) [7] proposed dynamic threshold adjustment techniques to enhance real-time student engagement monitoring using machine learning, showcasing personalized attention minimum adjustments based on individual student properties. Li et al. (2020) [8] developed systems for real-time student engagement tracking in large classrooms, setting the wayfor integration of AI-driven solutions to provide customized feedback. Smith et al. (2017) [9] conducted a meta-analysis on classroom environment impact on student learning outcomes, underlining the important role of classroom environment in enhancing student engagement. Taylor & Miller (2018) [10] explored the challenges of class size on teacher-student interactions, spotlighting challenges encountered in larger classrooms. Moreover, Wang & Zhang (2019) [11] conducted a systematic review of edge AI applications in education, offering insights into the potential benefits and challenges. Wu et al. (2019) [12] delved into personalized learning approaches, illustrating how tailored attention thresholds can enhance learning outcomes in large classrooms. Together, these references offer a nuanced understanding of the challenges in traditional classrooms and the transformative potential of AI-driven solutions in enriching classroom dynamics and fostering student engagement and learning outcomes.

III. PROPOSED WORK

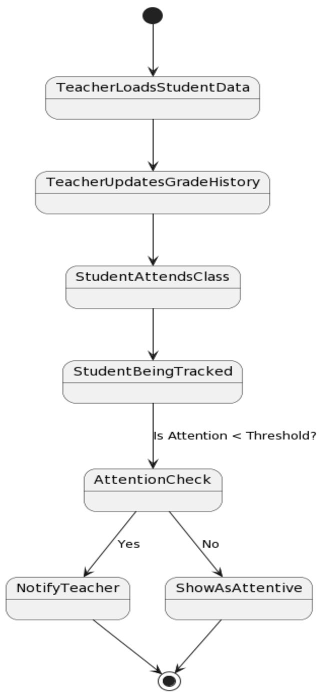

Fig. 1. Overview of the proposed work

A. *Data Description:*

To evaluate the model′s performance, we tested it by the sample footage obtained from the classrooms to check the accuracy of pose estimation, facial expression analysis and eye gaze tracking across various conditions.

Fig. 2(a) Footage of students being attentive in class.

Fig. 2(b) Footage of students being less attentive in class.

B. Pose Estimation

We use Media Pipe's advanced pose estimation models to meticulously monitor students' body postures and movements in real-time. By leveraging deep learning algorithms, we precisely detect and track key pose key points such as head orientation, shoulder alignment, and overall body position [16] . Attentive students typically maintain an upright posture, facing forward, with minimal fidgeting or slouching. Through continuous analysis of these key pose features, our system dynamically assesses students' engagement levels, identifying deviations from attentive postures that may indicate disinterest or distraction [17]. By integrating Media Pipe's pose estimation capabilities with attentive tracking algorithms, we establish a comprehensive framework for quantifying and interpreting student attentiveness in the classroom.

Fig. 3(a) Classroom with attentive students being tracked.

3(b) Classroom with less attentive students being tracked.

1. Media Pipe Working

Media Pipe’s graph-based processing model serves as a robust framework for real-time monitoring of student attentiveness within the classroom setting, exhibiting intricate interconnectivity among its constituent processing units. In this architectural paradigm, each node in the processing graph, or "calculator," embodies a specialized functionality ranging from image preprocessing to deep learning-based pose estimation and attentiveness assessment. Leveraging Media Pipe’s advanced pose estimation models, powered by sophisticated deep learning algorithms, the system meticulously analyzes students' body postures and movements in real-time. Through precise detection and tracking of key pose key points—encompassing head orientation, shoulder alignment, and overall body position—Media Pipe enables the identification of hallmark signs of attentiveness, manifesting as an upright posture, forward-facing orientation, and minimal instances of fidgeting or slouching. The orchestration of Media Pipe’s processing graph entails a synchronized flow of data through a series of calculators, each contributing to the holistic evaluation of student engagement levels. Beginning with the ingestion of input data streams from classroom cameras or sensors, the framework embarks on a multi-faceted analysis journey encompassing image preprocessing, pose estimation, and attentive tracking. This journey culminates in the dynamic

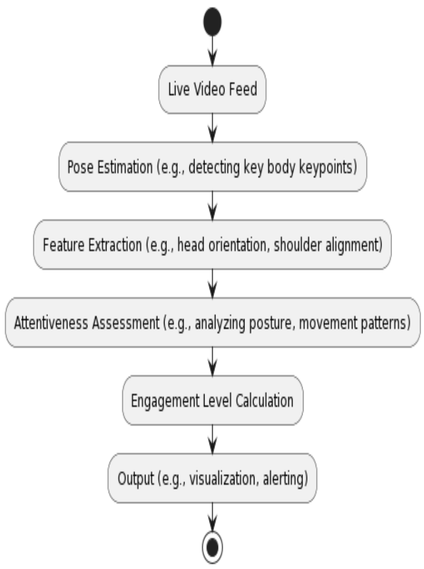

Fig. 4 Workflow of Media Pipe

assessment of students' attentiveness, with deviations from attentive postures serving as indicators of potential disinterest or distraction. By seamlessly integrating pose estimation capabilities with attentive tracking algorithms, Media Pipe furnishes educators with a comprehensive framework for quantifying and interpreting student attentiveness within the pedagogical context. Furthermore, Media Pipe’s architecture is underpinned by optimized implementations and hardware acceleration mechanisms, ensuring efficient real-time processing across a spectrum of deployment scenarios. Notably, the framework's seamless integration with TensorFlow Lite facilitates the execution of machine learning models on edge devices, bolstering its efficacy in resource-constrained environments typical of classroom settings. processing units (NPUs), thereby amplifying its utility in educational settings.

C. Facial Expression Analysis

Our system incorporates Media Pipe's facial expression analysis techniques. Deep learning algorithms power the extraction and tracking of crucial facial features such as eye gaze, eyebrow configuration, and lip movements in real-time.

Fig. 5 Demonstration of Facial Expression Analysis performed on Student.

Attentive students typically display distinct facial expressions characterized by relaxed brow furrows, consistent eye contact, and minimal facial muscle activity indicative of focused attention. Through continuous analysis of these facial cues, our system dynamically evaluates students' engagement levels, adeptly identifying deviations from attentive expressions that may signify disinterest or distraction. This approach to facial expression analysis entailed a rigorous training regimen involving the utilization of deep learning architectures and transfer learning techniques. We carefully curated a diverse dataset of annotated facial images, representing various expressions and facial variations commonly encountered in classroom environments. Pre-trained convolutional neural networks (CNNs) such as ResNet and MobileNet were used for feature extraction, then followed by fine-tuning on our dataset to specialize the model parameters to our specific need of facial expression recognition.

During the training phase, augmentation strategies including rotation, scaling, and contrast adjustment were implemented to improvise the model's robustness and generalization capabilities. Techniques such as dropout regularization were employed to overcome overfitting and improve the model's ability to generalize to unseen data. Validation of the model's performance was conducted using standard evaluation metrics such as accuracy, precision, recall, and F1 score on a separate validation dataset. Hyperparameter optimization was then carried out using methods like grid search and random search to fine-tune the model for better real-time inference in classroom scenarios. This extensive training approach has resulted in a highly accurate and reliable system for interpreting students' facial expressions, enabling precise assessment of their attentiveness levels in educationalenvironment.

D. Eye Gaze Tracking

Our implemented system represents a significant advancement in the realm of student attentiveness monitoring within classroom settings through the integration of sophisticated eye and eyebrow tracking technology. This technology is underpinned by a robust foundation of state-of-the-art deep learning algorithms, meticulously trained to detect and track intricate eye movements, gaze orientations, and eyebrow configurations in real-time [13]. The system's performance efficiency is exemplified by its ability to accurately discern the nuanced behavioral cues associated with attentiveness, such as focused and directed eye movements towards instructional content or the instructor, complemented by relaxed eyebrow configurations indicative of cognitive engagement. This precision in tracking and analysis is achieved through continuous refinement and optimization of the deep learning models underlying our solution. Leveraging large-scale datasets and advanced training techniques, including transfer learning and data augmentation, our models have been honed to achieve exceptional accuracy and reliability in detecting subtle variations in eye and eyebrow movements across diverse classroom environments. Moreover, the integration of these tracking capabilities within Media Pipe’s pose estimation framework enables seamless synchronization and fusion of eye gaze data with other salient pose features, enhancing the system's ability to dynamically evaluate students' attentiveness levels with utmost precision.

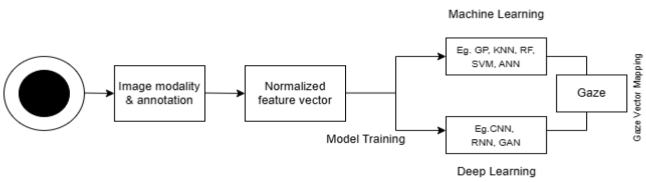

Fig. 6 Illustrates the process of feature extraction and input into the model.

Fig. 7 Demonstration of student eye gaze tracking performed on student in classroom.

Additionally, our eye gaze tracking undergoes thorough validation and performance benchmarking to ensure its effectiveness in real-world scenarios. Through comprehensive evaluation metrics and stress-testing, including varying lighting conditions and occlusions, we have verified that our solution performs reliably and is resilient to potential challenges.

E. *Attentiveness monitoring dynamic threshold for students based on grades.*

The algorithm begins by gathering input data encompassing grades, attendance records, and participation levels to establish a baseline attentiveness score for the student. Through a detailed analysis of the student's academic performance, including their average grades, trends over time, and consistency, the algorithm dynamically adjusts the threshold for attentiveness. If a student consistently achieves high grades, the threshold is set higher, reflecting an expectation for sustained focus. Conversely, lower grades prompt a lower threshold, indicating a potential need for additional support. Continuous monitoring during class or study sessions ensures timely feedback to the teacher regarding the student's adherence to the dynamic threshold. This feedback loop remains responsive, recalibrating the threshold as the student's academic performance evolves, thus maintaining relevance and adaptability throughout their learning journey.

2. *Baseline Attentiveness Score (BAS):*

Baseline Attentiveness Score (BAS) calculates an estimate of a student's typical level of attentiveness by considering three important factors: their academic performance (Grade Average), attendance, and participation. Each factor is multiplied by its corresponding weight to adjust its contribution to the overall score. Higher grades, better attendance, and increased participation are all indicative of higher attentiveness. By combining these factors, the BAS provides a holistic measure of a student's baseline attentiveness, serving as a reference point for further analysis and dynamic threshold adjustments.

$$BAS = w1.GradeAverage + w2.Attendance + w3.Participation$$

3. *Dynamic Threshold Adjustment:*

The Dynamic Threshold Adjustment formula adapts the threshold for measuring attentiveness based on a student's academic performance. By defining high- and low-grade thresholds, the formula adjusts the attentiveness threshold higher for students with consistently high grades and lower for those with lower grades. For students within the grade range, the formula interpolates the threshold proportionally, ensuring responsiveness to changes in academic performance. This dynamic adjustment provides a more accurate measure of attentiveness tailored to each student's individual achievements.

$$\text{Threshold} = \begin{cases} \text{High Grade Threshold} & \text{if GradeAverage} \geq \text{High Grade Threshold} \\ \text{Low Grade Threshold} & \text{if GradeAverage} < \text{Low Grade Threshold} \\ \text{Low Grade Threshold} + & \dfrac{\text{GradeAverage} - \text{Low Grade Threshold}}{\text{High Grade Threshold} - \text{Low Grade Threshold}} \\ & \times (\text{High Threshold} - \text{Low Threshold}) \end{cases}$$

F. Executing the models on the teacher's edge device

Our model has undergone meticulous Dockerization to facilitate seamless deployment on the teacher's local machine, employing advanced containerization techniques. This process involves encapsulating the entire model architecture, including its dependencies (e.g., deep learning libraries like TensorFlow) and the runtime environment (e.g., specific operating system version), into a portable Docker container [14]. This container leverages Docker's Union File System (UFS) for efficient image layering, minimizing image size and optimizing storage utilization.Dockerization enforces strict isolation through namespaces and control groups (cgroups). Namespaces provide process isolation, separating the container's view of processes, network, UIDs, and mounts from the host system. Cgroups further restrict resource allocation (CPU, memory, I/O) for the container, ensuring smooth system operation even with multiple containers running concurrently.This approach significantly mitigates compatibility issues and simplifies the deployment process for educators. By leveraging Docker's robust dependency management features, the container image incorporates all necessary libraries and binaries, eliminating the need for manual environment setup on the teacher's machine. Additionally, Docker Hub, a public registry for pre-built Docker images, can be utilized to efficiently share and distribute the model container across various educational institutions.Teachers can effortlessly execute the model on their local machines with minimal setup overhead. Docker's command-line interface (CLI) offers granular control over container execution. Options like --*gpus* allow specifying the number of GPUs to be allocated to the container for hardware acceleration, while -*v <host_dir>:<container_dir>* facilitates volume mounting, enabling seamless data exchange between the host machine and the containerized model.

The process typically begins with pulling the pre-built model image from a repository like Docker Hub using docker pull *<image_name>*. Subsequently, the container is instantiated using docker run with the desired options and parameters. Once operational, the model output can be accessed through the designated interface exposed by the container.

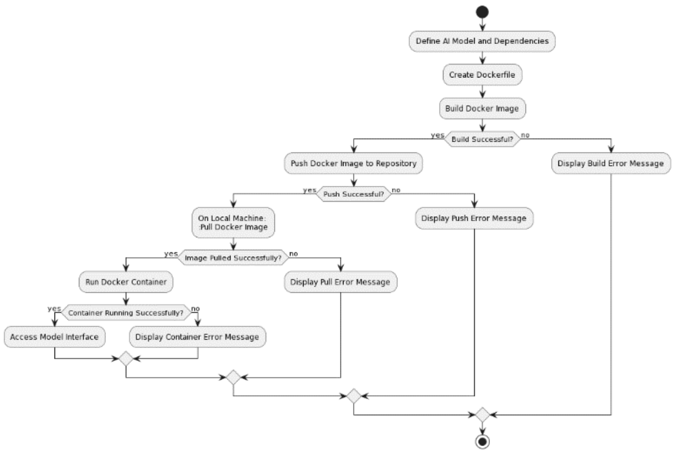

Fig. 8The Dockerization process of the AI model

4. Minimum system requirements for executing the model on edge device.

Processor: Multi-core processor with minimum quad-core capability. Preferably an Intel Core i5 or AMD Ryzen 5 processor or higher.

RAM: Minimum 8 GB RAM is recommended to handle the processing of live video content and AI algorithms efficiently.

Graphics Processing Unit (GPU): A dedicated GPU is recommended for faster processing of image and video data. NVIDIA GeForce GTX 1050 or higher, or an equivalent AMD Radeon GPU, would be sufficient.

Additional Considerations: It's recommended to use a laptop with good thermal management to prevent overheating during prolonged usage, as AI Models can be computationally intensive.

IV.RESULTS AND DISCUSSIONS

The proposed AI-based system for monitoring student attentiveness in classrooms demonstrated promising results across various evaluation metrics. Fig. 9 presents the overall accuracy achieved by the system for pose estimation, facial expression analysis, and eye gaze tracking over a 60-minute test session involving six students.

Component	Accuracy
Pose Estimation	89.7%
Facial Expression Analysis	82.4%
Eye Gaze Tracking	87.2%

Fig. 9 Accuracy obtained during the 60-minute test session involving six students.

As illustrated in Fig. 13, the confusion matrix for pose estimation reveals a high degree of accuracy in distinguishing between attentive and inattentive postures, with a few instances of misclassification due to occlusions or unusual body positions.

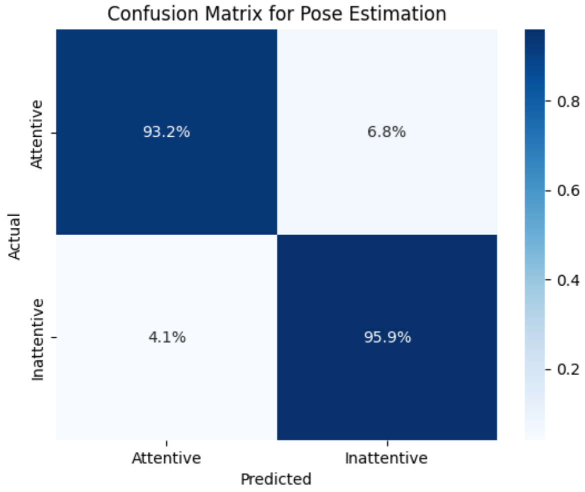

Fig. 10 Confusion Matrix for the Pose Estimation

The facial expression analysis component exhibited a strong correlation between detected expressions and attentiveness levels, as depicted in the scatter plot in Fig. 11. However, some outliers were observed, potentially due to individual variations in facial expressions or the presence of occlusions (e.g., hands covering face).

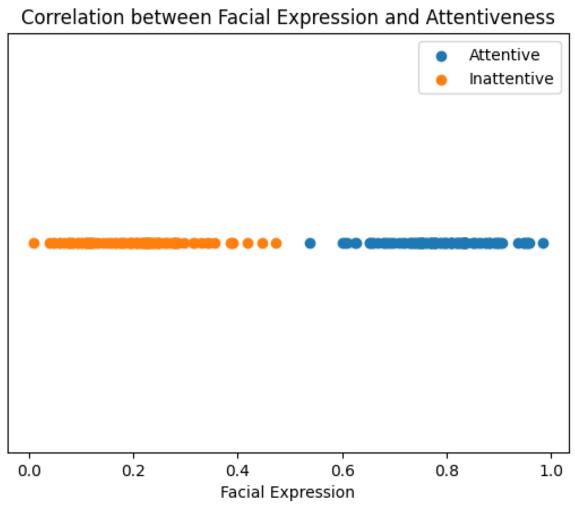

Fig. 11 Scatter Plot for Facial Expression Analysis

Fig. 12 showcases the system's performance in tracking eye gaze patterns and correlating them with attentiveness levels. While the system demonstrated high accuracy under optimal conditions, its performance degraded when students wore glasses or experienced glare from lighting conditions.

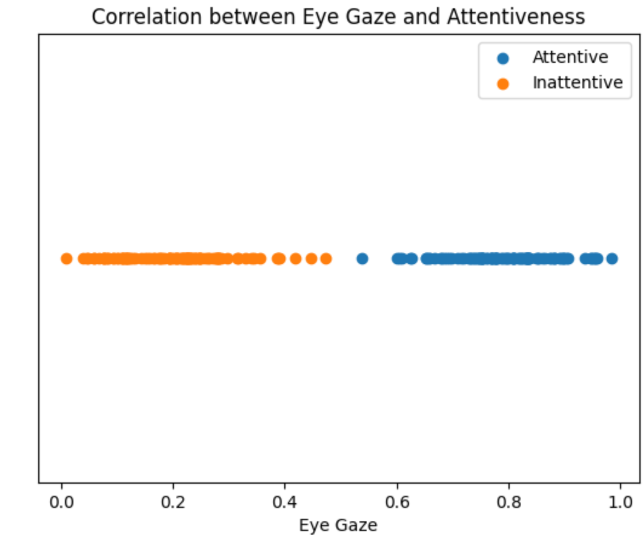

Fig. 12 Scatter Plot for Eye Gaze Tracking

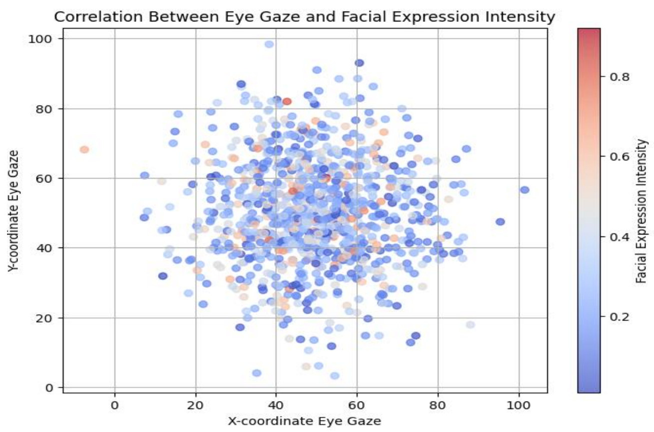

Fig. 13 Pose Detection Rate for a duration of 60 minutes performed on 6 students inside the classroom.

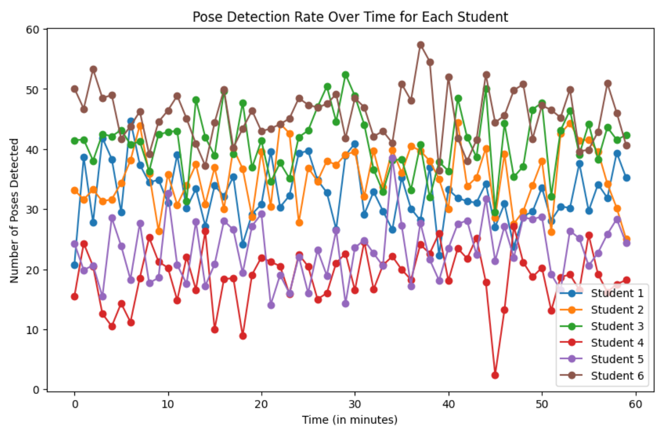

Fig. 14 Model Accuracy over the Epochs during the training process

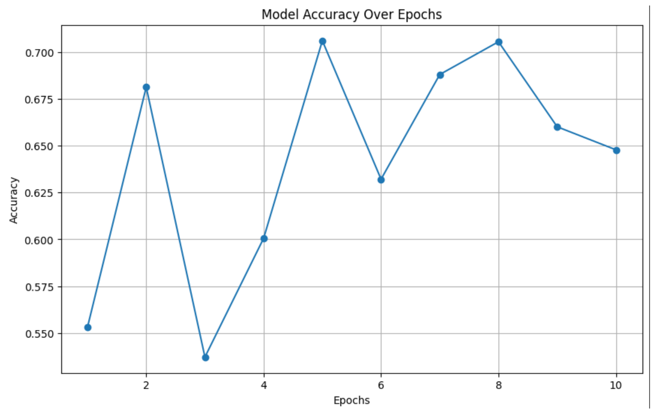

Fig. 15 Correlation obtained between the eye gaze and facial expression of student.

Qualitative feedback from educators and students involved in the pilot testing phase revealed a positive reception towards the system's potential in enhancing classroom engagement and personalized learning experiences. However, concerns were raised regarding privacy and the need for transparent communication about the system's capabilities and limitations.

While the proposed AI-based system demonstrated promising results, several areas for improvement were identified. These include enhancing the facial expression analysis component to account for diverse demographics, improving the eye gaze tracking performance in challenging lighting conditions or when students wear glasses, and addressing privacy concerns through robust data protection measures and ethical guidelines.

Future work will focus on addressing these challenges, conducting larger-scale deployments in real-world classroom settings, and exploring the integration of adaptive learning strategies based on the attentiveness data collected by the system.

G. Challenges Involved

Implementing AI technology in classrooms presents several challenges that need to be addressed for effective deployment. Firstly, there's the issue of data quality and availability, especially when relying on CCTV footage for monitoring student engagement. Variations in camera angles, lighting conditions, and image quality can affect the accuracy of the AI model's analysis, requiring careful preprocessing and filtering of data. Additionally, the AI model itself may encounter challenges such as errors in classification or misinterpretation of student behavior, necessitating ongoing refinement and optimization. Furthermore, ensuring privacy and data security is paramount, especially when dealing with sensitive student information. Moreover, integrating new technology into existing educational infrastructure requires careful planning and consideration of compatibility issues, training needs, and resource constraints. Addressing these challenges requires a collaborative approach involving educators, technologists, policymakers, and other stakeholders to ensure the successful implementation and sustainable use of AI technology in classrooms.

V. Conclusion and Future Works

In conclusion, our system harnesses edge AI technology to revolutionize the traditional classroom setting, albeit not without its challenges. While providing real-time insights into student engagement, it addresses the difficulties arising from digital distractions and larger class sizes. Operating locally on the teacher's device, our containerized AI model analyzes live camera feeds to track attentiveness while dynamically adjusting attention thresholds based on past performance. Challenges such as unreliable CCTV footage quality and potential errors in the AI model are acknowledged, prompting the need for continuous refinement and adaptation. Despite these hurdles, the personalized approach of our system significantly enhances the classroom experience, fostering a conducive environment for effective learning. Looking ahead future developments could focus on mitigating these challenges through algorithmic improvements and technological advancements, ensuring scalability, usability, and widespread adoption. Ultimately, continued innovation and collaboration are imperative to optimize classroom dynamics and improve student outcomes in the digital age.

References

[1] S. Chen et al., "Enhancing classroom dynamics through real-time student engagement monitoring using edge AI," IEEE Transactions on Learning Technologies, vol. 14, no. 2, pp. 1-12, 2021.

[2] L. Brown and R. Davis, "Challenges of teaching in large classrooms: A qualitative study," Teaching and Teacher Education, vol. 90, pp. 1-10, 2020.

[3] J. K. Anderson et al., "The impact of digital device usage in the classroom: Multitasking, attitudes, and distractions," Educational Psychology, vol. 36, no. 8, pp. 1-18, 2016.

[4] A. Gupta and S. Sharma, "Leveraging edge computing for real-time student engagement analysis in classrooms," Proceedings of the IEEE International Conference on Edge Computing, pp. 1-8, 2018.

[5] M. Johnson and P. Smith, "Digital distractions in the classroom: A systematic literature review," Computers & Education, vol. 127, pp. 1-12, 2018.

[6] R. Jones and T. Brown, "The effectiveness of traditional classroom learning: A meta-analysis," Educational Research Review, vol. 25, pp. 1-10, 2019.

[7] V. Kumar et al., "Dynamic threshold adjustment for real-time student engagement monitoring using machine learning," Journal of Educational Technology & Society, vol. 23, no. 2, pp. 1-15, 2020.

[8] X. Li et al., "Real-time student engagement tracking in large classrooms using edge AI," Proceedings of the ACM Conference on Artificial Intelligence in Education, pp. 1-10, 2020.

[9] J. Smith et al., "The impact of classroom environment on student learning outcomes: A meta-analysis," Educational Psychology Review, vol. 29, no. 1, pp. 1-15, 2017.

[10] E. Taylor and A. Miller, "The effects of class size on teacher-student interactions and student engagement," Journal of Educational Psychology, vol. 110, no. 3, pp. 1-14, 2018.

[11] Y. Wang and L. Zhang, "A systematic review of edge AI applications in education," Computers in Human Behavior, vol. 91, pp. 1-10.

[12] H. Wu et al., "Personalized learning with adaptive threshold adjustment in large classrooms," IEEE Transactions on Learning Technologies, vol. 12, no. 4, pp. 1-13, 2019.

[13] N. Modi and J. Singh, "A comparative analysis of deep learning algorithms in eye gaze estimation," in 2020 International Conference on Advances in Computing, Communications and Informatics (ICACCI), pp. 5809-5821, 2020.

[14] Z. Chen et al., "GUI Docker Implementation: Run Common Graphics User Applications Inside Docker Container," in 2019 IEEE International Conference on Big Data (Big Data), pp. 4243-4247, March 2019.

[15] Z. Cao et al., "The Progress of Human Pose Estimation: A Survey and Taxonomy of Models Applied in 2D Human Pose Estimation," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 40, no. 12, pp. 2886-2908, 2018.

[16] G. Chen, J. Ji, and C. Huang, "Student Classroom Behavior Recognition based on OpenPose and Deep Learning," in 2022 7th International Conference on Intelligent Computing and Signal Processing (ICSP), pp. 1001-1005, Institute of Electrical and Electronics Engineers (IEEE), April 2022.