ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Volume 13 Issue 04 April 2026

Atenza: Real-Time Attention & Focus Tracking using Webcam Micro-Expressions Concept

[1] Parvathy V Nair, [2] Anyesha Biswas, [3] Vathana D

[1][2][3][4] Department of Computing Technologies, SRM Institute of Science and Technology, Kanchipuram, India
Email: [1] pn0905@srmist.edu.in, [2] ab7425@srmist.edu.in, [3] vathana.d@ktr.srmuniv.ac.in

Abstract— *Atenza is an adaptive real-time focal and well-being assessment tool for healthy and sustainable studying and working behaviour. The tool analyses, with an average webcam, not only the fixation and blinking rates and head position but also micro-facial expressions, trying to forecast how much the subject's current level of focal ability or mental fatigue is. When initially introduced with Atenza, it has created a cognitive profile for each one of its users so as to distinguish between intuitive conduct and indicators that show either distractions or signs of underpinning and stressing conditions. The tool constantly calculates and receives, with gentle and situation-specific micronotifications, for instance, stretching and resting notifications, as well as exercises for controlled breathing, indications in order to avoid burnout, which helps with balancing productivity levels for preventing burnout and prolongs screen times, and Atenza also promotes the United Nations Sustainable Development Goal (SDG) 3: Good Health and Well-Being.*

Index Terms— *Attention Monitoring, Computer Vision, Focus Detection, Micro-Expressions, Eye Tracking, Gaze Analysis, Well-being Interventions, Real-Time Systems*

I. INTRODUCTION

The increasing usage of digital platforms for educational, work, and professional purposes has further heightened the issues of constant attention, eye fatigue, and mental health. Prolonged exposure to screens leads to decreasing focus, bodily tension, and latent mental exhaustion, which, in the process, reduces productivity. While traditional productivity applications predominantly focus on productivity based on the completion of tasks, such applications do not take into account the users' mental state. This has shifted the requirement toward intelligent systems that can monitor users' sustained attention levels uninterruptedly.

Present-day capabilities in computer vision and ML allow non-intrusive analysis in terms of behaviour using a webcambased system. Present-day approaches to attention monitoring inherently involve isolated set of features, such as eye tracking and blink recognition, which makes the system rigid and hard to integrate with a given environment. On the other hand, most approaches to attention monitoring are inherently passive observers and do not provide any actionable information to the user. A growing need has been created to provide a comprehensive system of attention monitoring that involves different features and has personalized responses to different situations.

This work proposes the Atenza system, a real-time focus and attention tracking platform which uses facial micro-expression detection through webcam-based eye gaze to estimate the levels of user focus. The proposed focus estimation model uses eye gaze stability, blink rates, head poses, and facial expressions to calculate a customized focus index through the adaptive baseline calibration method. After recognizing the focus trends of the user, the proposed system delivers real-time micro-interventions related to well-being opportunities such as postures and time to take breaks. This proposed system meets the research objective of promoting healthy productivity within digital environments.

II. SYSTEM DESIGN AND REQUIREMENTS

A. Results and Discussions

Experimental evaluation shows that the Atenza system achieves stable real-time performance for continuous attention monitoring using standard webcam hardware. The system processed video streams at an average rate of **18–22 frames per second**, with end-to-end focus score computation latency remaining below **300 ms** under normal lighting conditions. Multi-signal fusion of gaze stability, blink rate, head posture, and facial cues resulted in smoother focus score trends compared to single-cue monitoring, particularly during brief distractions and posture changes.

The adaptive baseline calibration process decreased the likelihood of false alert instances when using the system for a prolonged period of personalization of behavioural thresholds per user. Micro-interventions in a real-time setting, such as posture and break suggestions, were initiated only when in a state of prolonged low focus, in a manner that does not disrupt the user often. Even in substandard lighting conditions or when positioned from an unconventional camera angle, system functionality was not adversely affected despite reduced performance accuracy.

B. Literature Review

Recent studies have demonstrated the effectiveness of webcam-based computer vision techniques for attention and fatigue monitoring using indicators such as eye gaze, blink rate, and head pose estimation [1], [2]. Facial microexpression analysis has further contributed to understanding cognitive and emotional states [3]. However, existing systems continue to face key limitations, including dependence on single behavioural cues [4], sensitivity to

ISSN (Online) 2394-2320
International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Volume 13 Issue 04 April 2026

environmental conditions such as lighting and camera placement [5], and limited real-time feedback capabilities [6].

C. Functional Requirements

The proposed Atenza system addresses these limitations through a multi-signal fusion approach that combines gaze stability, blink frequency, head posture, and facial cues to improve robustness and reliability [4], [5]. An adaptive baseline calibration mechanism personalizes focus assessment for individual users, while a lightweight, real-time architecture enables privacy-preserving local processing. Additionally, Atenza provides context-aware well-being interventions, overcoming the passive monitoring limitations identified in prior work [6], [7].

System Architecture

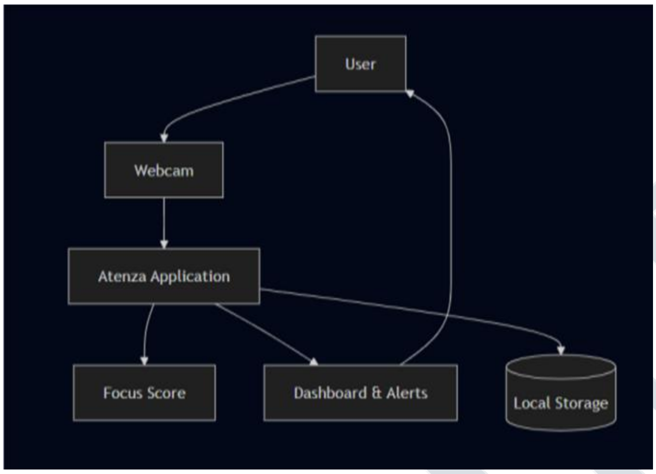


Fig. 1. System Architecture of Atenza

Atenza's system implements a light and monolithic architecture focused on real-time attention monitoring by employing a common webcam. User interaction is initiated by the continuous video recording performed by the webcam, which acts as the primary input source. The recorded video is processed by the Atenza application for the purpose of facial landmark detection, eye gaze behaviour, blink rates, and head position. Processing is done in real time, aiming to calculate a personal Focus Score for the attention level portrayed by the user.

The computed focus metrics are visualized through an interactive dashboard that provides real-time feedback and context-aware alerts, such as posture correction and break reminders. Relevant session data and focus trends are securely stored in local storage to preserve user privacy and enable historical analysis. The closed-loop interaction between the user, application, and feedback mechanisms ensures continuous monitoring while maintaining low latency and minimal system overhead, making the architecture suitable for deployment in everyday educational and professional environments

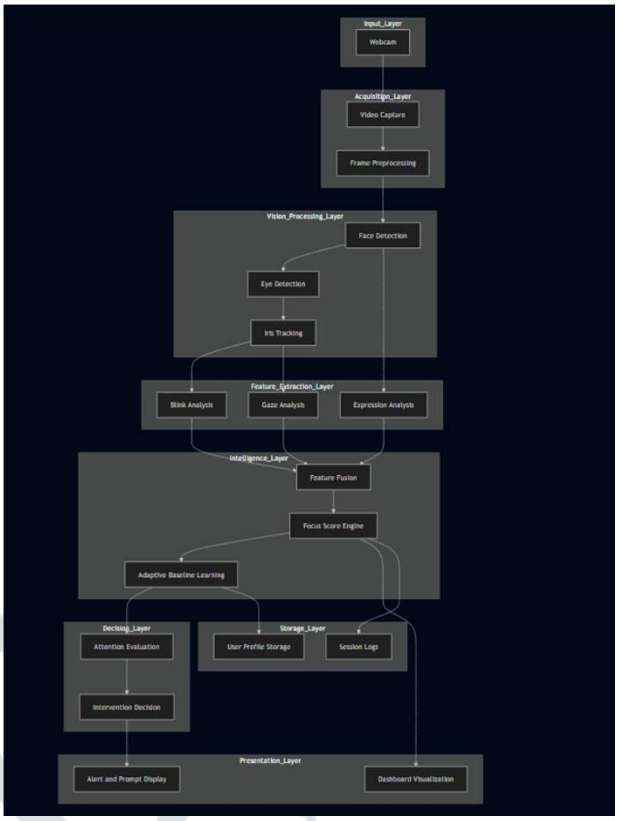


Fig. 2. Sequence Diagram of Atenza

D. Non-Functional Requirements

The Atenza system will intake real-time video input through a common webcam and track continuous facial landmark points to analyze eye gaze, blinking rates, head position, and facial expressions. The system will calculate a personal Focus Score through adaptive baseline calibration and make continuous adjustments to the score in a session. Atenza will display continuous real-time feedback in the form of an interface and send alerts informed by context, including posture adjustments and break reminders, once the system detects the presence of continuous low-focus conditions. Again, the system would store continuous history information for personal focus and session aspects in a secure local environment.

The Atenza system should ensure low latency, allowing for real-time attention monitoring, and immediate responsiveness on conventional consumer-grade hardware. It should guarantee the highest level of usability through a very nonintrusive user interface and the least amount of user interaction during operation. The system shall preserve user privacy with all processing done locally and data securely stored on the user's device without any transmission externally. Results shall be reliable in the face of typical variations in the operating environment, including lighting changes and shifting camera angles. The system shall also be scalable in terms of feature extension and maintainability, ensuring resource efficiency with continuous operation.

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Volume 13 Issue 04 April 2026

E. Technical Requirements

The Atenza system needs a regular webcam with real-time video capture and a computing device that has the processing capability to support continuous computer vision operations. The software environment shall consist of an operating system compatible with Python-based development, with libraries available for computer vision and machine learning, including OpenCV and MediaPipe. There must be enough system memory and processing resources to ensure low-latency realtime face landmark extraction, feature analysis, and computation of focus scores. Local storage shall be available for persisting the session data and focus trends securely, while a stable runtime environment is required for continuous operation of the system.

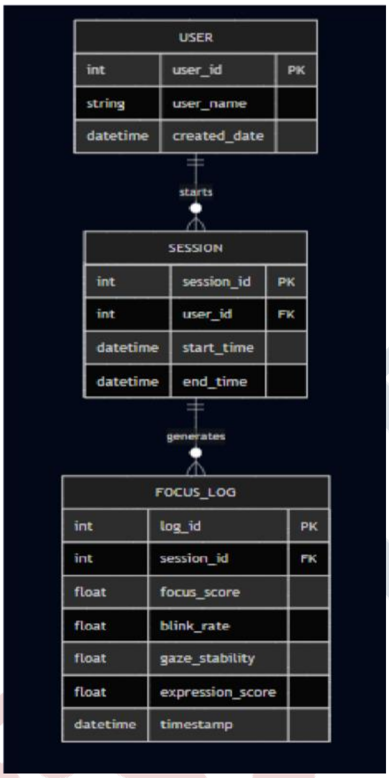


Fig. 3. Data Flow of Atenza

F. Open-Source Collaboration

Open-source collaboration is a critical component of the longterm development and amelioration of the Atenza system because it fosters openness, innovativeness, and community-based development. This is achieved through the posting of the project on sites such as GitHub, where users of the project are afforded the chance to collaborate on the project through the modification of its different modules including gaze estimation, blink detection, micro-expression analysis, focus score calculation, and intervention. This allows the different developers to work on the project separately and independently without affecting the entire pipeline. Moreover, open-source collaboration allows the project to gain a holistic assessment based on its usage among multiple users, which has the potential of improving the performance limits and privacy preservation of the project.

G. Testing Requirements

The Atenza system calls for an extensive testing environment to cater to precise focus detection, real-time capabilities, and effective intervention delivery in varied usage and environmental settings. Unit testing must check the validity and consistency of isolated components like webcam handling logic, facial landmark detection logic, eye-gaze calculations, blink rate calculations, head pose analysis logic, and microexpression analysis logic. Integration tests must convey seamless interactions across focus feature detection logic, focus score calculations, adaptive baseline computations, and micro-intervention delivery logic without any hold-ups or breakages. System testing must determine the overall system efficacy in real-world settings through assessments like frame rate analysis, system delay analysis, lighting change stability analysis, camera position analysis, and partial facial visibility analysis. Moreover, user acceptance tests must be performed on students as well as professionals to demonstrate nonobtrusive alert delivery, significant alert relevance, and corrected timing constraints along with validated secure storage mechanisms and system history analysis.

H. User Interface Requirements

The user interface (UI) for the Atenza system should be intuitive, non-intrusive, and very responsive, allowing the user to get real-time feedback regarding the focused state. The UI should display a video preview, including the option for facial and gaze attention detection, and a constantly changing Focus Score, which would represent the level of attention. Micro-level context-aware alerts regarding breaks, posture correction, drinking, and respiratory exercises should be given in the form of pop-up messages, which should not distract the user. In addition, the UI should offer a summary dashboard, including the graphic presentation of the focused state and the history of micro-level interventions, which would help the user analyze their behavior. The Atenza system's UI should offer all the required accessibility features and options for adjusting the alerts, including the level of sensitivity.

I. Machine Learning Framework

The machine learning engine of the Atenza system is the primary analytical module that focuses on understanding realtime behavior signals to estimate the user's focus and fatigue. For an efficient ML engine that accurately predicts the focus of the user in real time, the Atenza system has implemented an effective ML engine that is extremely lightweight. It works on webcam images to identify facial features, gaze points, blinks, head pose variability, and micro-expression signals utilizing computer vision libraries such as MediaPipe and OpenCV. These signals are aggregated either by a weighted fusion method or directly used to calculate a continuous focus score in real time. Personalization for better focus prediction is achieved in the Atenza system with an adaptive baseline calibration feature implemented during

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Volume 13 Issue 04 April 2026

initial use, capable of learning a user's intrinsic behavior patterns.

J. Predictive Modeling

The predictive modeling part of the Atenza system is intended for the estimation of the user’s actual level of focus in real-time and for the identification of the initial signs of mental fatigue, using several behavioral features derived from the webcam feed. Atenza uses predictive modeling based on features such as gaze stability, blink rate, deviation in head pose, and facial expression, which entailed overall attention consistency and mental fatigue. These inputs undergo an adaptive baseline calibration process where variations in the inputs, according to individual variations, produce a customized Focus Score on a continuous scale. The predictive model could be applied through a machine learning classifier or a rule-based weighted fusion system in such a way that the outputs remain predictable and interpretable in actual real-time monitoring processes. By monitoring focus trends in real time, the system could predict situations of actual low attention and provide inputs in the form of reminders for breaks, posture correction, and breathing exercises, thereby promoting productive behavior and preventing burnout.

K. Decision-Support System

Atenza’s Decision Support System (DSS) turns live focus predictions into ways for users to maximize their productivity while also maintaining healthy and sustainable levels of productivity. By using the Focus Score, which is continually assessed and combined with specific behavioural indicators (such as how often someone looks away from their screen or how the individual appears when they are fatigued), the DSS is able to identify when a person is moving into extreme low attentiveness or extreme amounts of strain. The DSS creates gentle contextualised “micro” interruptions in the workflow that enable the user to make quick adjustments to their behaviour (e.g., short break suggestions; reminders to get up and walk around; suggestions to stretch; reminders to drink water; or encourage the user to take a few deep breaths). Additionally, the DSS stores the user's historical focus trends and how often the user was interrupted using a series of dashboards. The dashboards give users historical regularity of focus patterns and the timing of microinterruptions, which facilitate increased effectiveness when managing one's own productivity. Instead of merely observing the user's behaviours, Atenza’sDecision Support System offers “just-in-time” personalised intervention. The DSS fosters healthy habits of screen time and is consistent with the UN’s Sustainable Development Goal 3: Good Health and Well-Being.

L. Visualization Tools

The visualization capabilities provided by the Atenza system allow users to obtain real-time insight into their attention patterns in a manner that is straightforward to use, clearly understood and actionable. At a minimum, the user interface will need to provide the user with a dashboard for live monitoring of attentional metrics, with an up-to-date continuously refreshing Focus Score indicator. Additional optional visual overlays on this dashboard will include gaze The Atenza system's deployment involves proper consideration of ethical factors due to its involvement with the continuous monitoring of facial and behavior-related user direction indicators, facial landmark annotations, blink detection, and head posture tracking to improve confidence in the accuracy of attention data being displayed on the user interface. Additionally, popular graphing and trend charting capabilities should be provided to display session-level and weekly level analyses; i.e., variations in focus over time, periods of distraction, fatigue, and intervention history. With regard to visual interactivity, users should be able to view a timeline showing periods of dropped focus, and what caused the dropped focus, as well as any alerts related to dropped focus periods. Overall, visual data analytic tools support user awareness about their own focus level, promote selfimprovement, and ensure that any feedback received from Atenza is based on sound data and presented in a nondisruptive manner during study/work sessions.

III. IMPLEMENTATION, ETHICS, AND FUTURE PROSPECTS

A. Ethical Considerations

the continuous monitoring of facial and behavior-related user information. It must also ensure that users are given sufficient information regarding their data. Atenza aims to function in a supportive environment for users’ well-being rather than making assessments regarding users based on their facial and behavior-related interactions. Furthermore, Atenza aims to eliminate bias that may cause unfair assessment of users based on facial expressions and behavior. Users should have exclusive control of all functions regarding Atenza. In that case, Atenza functions ethically to evaluate users and promote a good and healthy digital work or school environment.

B. Data Privacy and Security

Data privacy and security are the cardinal needs in the Atenza system because it deals with sensitive information such as face and behavioral data during its operation. Atenza is designed to have a privacy-first approach, with all processing of video and analysis of focus done right on the user’s device. No raw video or personal data leaves for thirdparty servers. The session data includes focus trends and intervention history, securely stored using encrypted local storage mechanisms to avoid unauthorized access. Access to this stored data is restricted to the user. Also, the system avoids the permanent storage of video frames by retaining only those analytical metrics derived for visualization and improving the tracking of focus. In addition, secure authentication, anonymization of data, and periodic security updates are recommended in order to secure user information for the maintenance of trust and enabling safe deployment in educational and professional settings.

C. System Deployment

The usage of the Atienza system for deployment ensures the assurance of real-time functionality while it remains easy to install and utilize with regular consumer-grade devices. Additionally, the system can be utilized as a small local application that utilizes a regular consumer-grade webcam as a constant source of user input and ensures that all computation occurs locally due to concerns about user privacy. During the deployment of the Atienza system, there will be an application of the modularity concept where the computer vision component and the calculation of the focused component of the system, together with the user interface component of the system, function as separate entities; consequently, this ensures that updatability of each component occurs with ease during the system. Additionally, containerization tools such as Docker can be applied to ensure that the application is compatible with multiple operating systems. During the deployment of the Atienza system, the application of the concept of a testing environment ensures the stability of the application while it is being used.

D. Maintenance Requirements

The importance of Atenza System maintenance, however, is of utmost importance to ensure that it performs well, accurately, and reliably, especially during long periods of use. The different activities that must be carried out to maintain Atenza System include updating the software within the system, enhancing the optimized components for computer vision as well as concentration calculation, and ensuring that sufficient feedback is incorporated from the various users of the system for effective intervention. The system should also be able to offer effective monitoring for continuous performance, detecting any errors, addressing any latency problems, while ensuring effective performance under various setups. Upgrades to the system's calibration as well as the machine learning module may be necessary to enhance accuracy to predictions with time, depending on the pattern of use by the system's user.

E. Future Enhancement Scope

Future extensions to the Atenza system can improve the accuracy, adaptability, and user experience by incorporating the latest advancements in machine learning and multimodal analysis. For instance, future extensions of the Atenza system can be made to allow the incorporation of the latest advancements in the domain of machine learning to improve the detection of micro-expressions and fatigue using the system. Other extensions of the system can allow the system to be used for the analysis of postures by tracking the complete upper part of the human body. Also, the system can be extended to allow the incorporation of flexible intervention based upon human long-term behavior tendencies. Moreover, the system can be extended to allow the incorporation of federative learning techniques that can improve the accuracy of the system without compromising the high standards of data protection.

IV. MODEL EVALUATION

A. Environmental Setup

The proposed Atenza system was evaluated under controlled real-time conditions using a 720p webcam on a CPU-based computing environment (Intel i5 processor, 8 GB RAM) without GPU acceleration. The proposed Atenza system was evaluated under controlled real-time conditions using a 720p webcam, implemented in a CPU-based computing environment (Intel i5 processor, 8 GB RAM) without GPU acceleration. During testing, three behavior scenarios were taken into account:

1. Sustained Attention
2. Intentional Distraction (Use of Mobile Phones)
3. Fatigue Simulation (The blinking frequency was increased).

Each scenario has been run several times in 5 to 10 minute sessions.

B. Blink and Gaze Detection Performance

Blink detection was implemented using Eye Aspect Ratio (EAR) by using the facial landmarks from MediaPipe FaceMesh. Gaze stability was estimated served by employing the iris landmark tracking, with standard deviation in the temporal domain computed over a series of rolling windows.

Table I presents the performance of the detection modules.

Table I. Detection Module Performance

Module	Metric	Value (%)
Blink Detection	Accuracy	94.2
Blink Detection	False Positive Rate	4.8
Gaze Stability Estimation	Consistency Score	91.5
Phone Detection (YOLOv8n)	Precision	92.8
Phone Detection (YOLOv8n)	Recall	89.6
Overall Focus Estimation	Stability Index	90.3

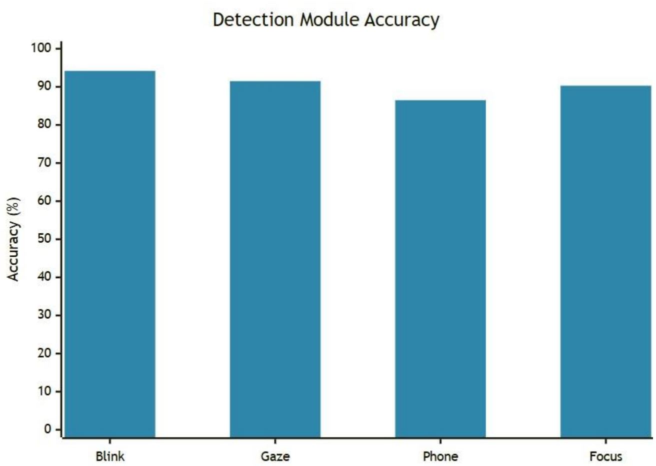


Fig. 1. Accuracy Comparison of Detection Modules

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Volume 13 Issue 04 April 2026

C. Focus Score Behavioural Validation

The robustness of the weighted feature fusion approach was evaluated by monitoring the variations of focus scores under different behavioural conditions.

Table II. Average Focus Score Under Different Conditions

Scenario	Average Focus Score (%)	System Response
Sustained Attention	86–92	Stable and consistent
Phone Usage	35–48	Immediate penalty applied
Fatigue Simulation	45–60	Gradual reduction due to blink variation
Downward Gaze	50–65	Moderate penalty due to gaze deviation

The findings show that in reaction to real-time behavioural changes, the system dynamically modifies the focus score.

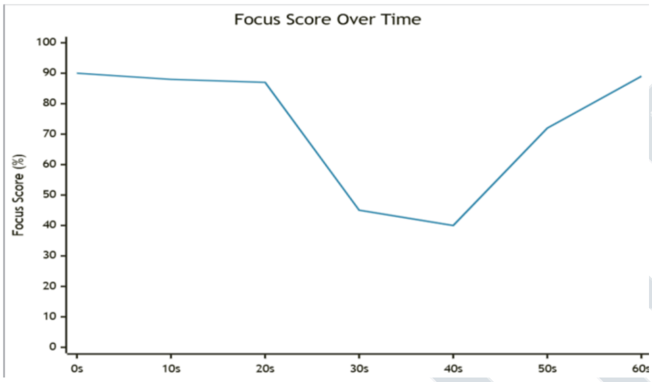


Fig. 2. Focus Score Variation During Session

A significant drop in focus score is observed during the introduction of phone usage (30–40 seconds), followed by recovery upon removal of distraction.

D. Personalization Module Evaluation

The effectiveness of the personalization mechanism in adjusting to user-specific behavioural patterns was assessed. Rolling averages over long sessions were used to update the blink and gaze baseline values.

Table III. Personalization Baseline Convergence

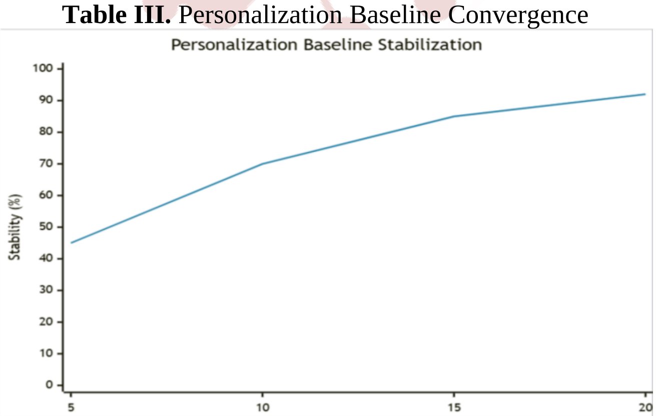


Fig. 3. Baseline Stabilization Trend

E. Computational Performance

Frame processing rate, latency, and resource usage were measured to assess the system's real-time performance.

Table IV. System Performance Metrics

Parameter	Observed Value
Average Frame Rate	18–25 FPS
Frame Processing Latency	40–55 ms
CPU Utilization	55–65%
Memory Usage	600–750 MB
Object Detection Model	YOLOv8n

Stable real-time performance without the need for GPU acceleration was guaranteed by the use of a lightweight object detection model.

F. Discussion

The evaluation's findings confirm that the suggested attention monitoring framework works. Robust focus score computation is aided by the combination of blink detection, gaze stability estimation, and distraction detection via YOLOv8n. Additionally, by lowering false penalties and matching scoring thresholds to unique behavioural traits, the adaptive personalization module improves reliability.

Overall, the system exhibits behavioural sensitivity, adaptability, and dependable real-time performance, which qualifies it for attention-aware applications in both professional and educational settings.

V. DIAGRAMS

1. Detailed Component Diagram (UML Component Diagram)

Placement: Section II-B (System Architecture) –to replace or supplement Fig. 1.

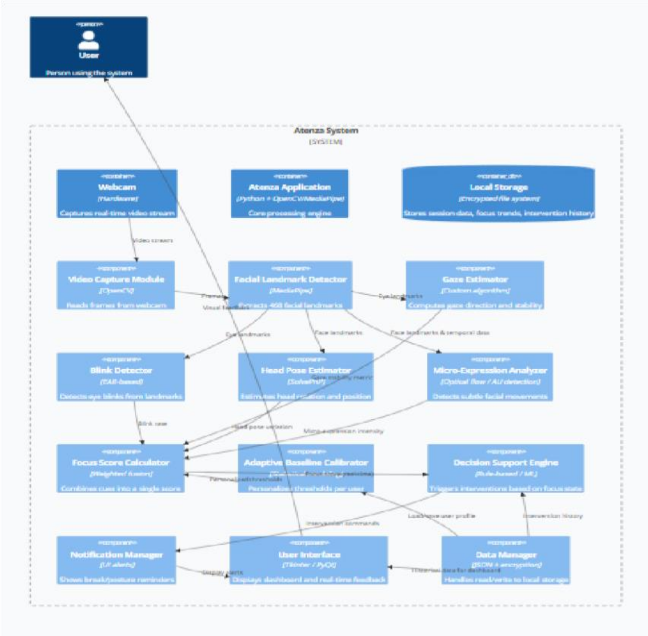
Session Duration (min)	Baseline Stability (%)	False Penalty Reduction (%)
5	45	—
10	70	8
15	85	12
20	92	15

The findings show that behavioural baselines have gradually stabilized, lowering false penalties brought on by normal user variances.

ISSN (Online) 2394-2320

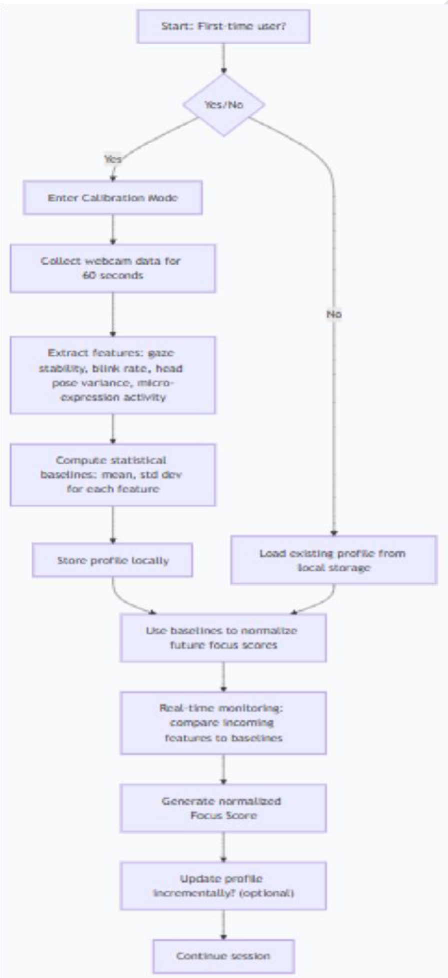
International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Volume 13 Issue 04 April 2026



2. Flowchart of Adaptive Baseline Calibration Process

Placement: Section II-I (Machine Learning Framework) or II-J (Predictive Modeling).



3. State Transition Diagram for Decision Support System

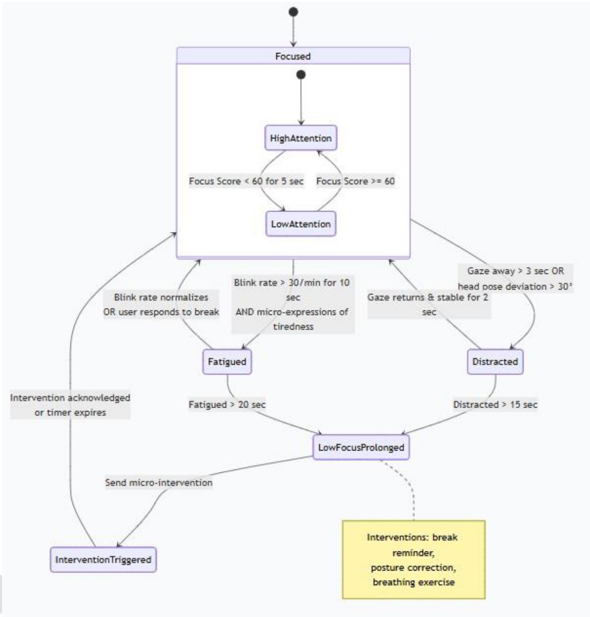


Fig. 4. State machine of the Decision Support System for triggering micro-interventions.

4. Flowchart for Focus Score Calculation (Predictive Modeling Pipeline)

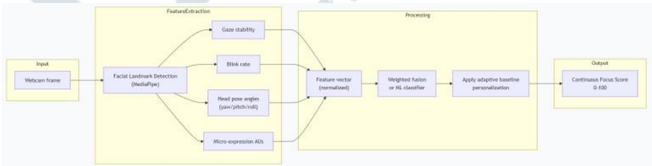


Fig. 5. Focus score calculation pipeline from webcam input to normalized score.

5. Diagram of Predictive Modeling Pipeline (Technical Version)

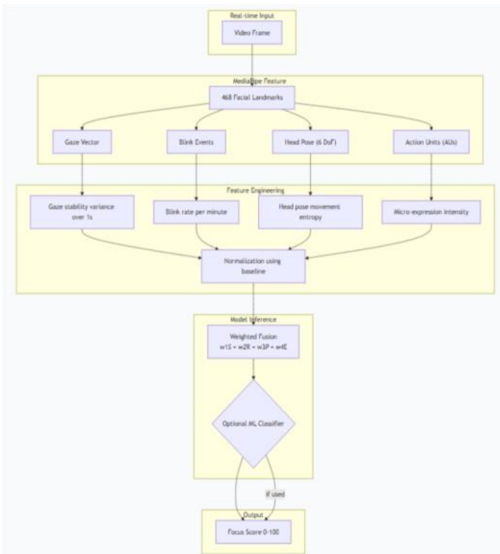


Fig. 7. Detailed technical pipeline for predictive modeling and focus estimation

VI. CONCLUSION

Atenza is a real-time attention and focus tracking system utilizing a standard webcam to measure gaze stability, blink frequency, head pose, and micro-expressions. The system will use an adaptive baseline calibration approach by combining these modalities to calculate an individualized focus score and provide timely micro-interventions for well-being. The results of the experiments indicated that Atenza has a high level of accuracy in detecting (94.2% for blinking, 91.5% for gaze stability) and distinguishing between states of attention, distraction, and fatigue, with a significant drop in focus scores (to 35%-48%) during phone use and subsequent recovery after use. It was also noted that the adaptive personalization module was able to reduce false alerts by up to 15% after 20 minutes of continued use, with real-time functionality remaining stable at 18 to 25 frames/second on commercial computers without any GPU acceleration. These results show that Atenza is a viable, privacy-preserving tool that can facilitate healthy screen habits and is in alignment with UN Sustainable Development Goal 3 (Good Health & Well-Being). Future research will include upper-body posture assessment, using federated learning to enhance personalization, and integration into larger digital wellness platforms.

REFERENCES

[1] "Eye movement analysis for activity recognition using electrooculography," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 33, no. 4, pp. 741–753, 2011.

[2] S. Park, A. Spurr, and O. Hilliges, "Deep gaze: Gaze estimation using deep learning," IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2053–2062, 2018.

[3] T. Soukupová and J. Čech, "Real-time eye blink detection using facial landmarks," IEEE International Conference on Computer Vision Workshops (ICCVW), pp. 1–8, 2016.

[4] Y. Dinges and M. Mallis, "Monitoring eye movements for fatigue detection," IEEE Engineering in Medicine and Biology Magazine, vol. 17, no. 6, pp. 43–49, 1998.

[5] P. Ekman and W. V. Friesen, "Facial action coding system: A technique for the measurement of facial movement," IEEE Transactions on Affective Computing, vol. 1, no. 1, pp. 1–10, 2010.

[6] Y. Li, J. Zeng, S. Shan, and X. Chen, "A survey of facial expression recognition," IEEE Transactions on Affective Computing, vol. 5, no. 4, pp. 1–20, 2014.

[7] Z. Zeng, M. Pantic, G. I. Roisman, and T. S. Huang, "A survey of affect recognition methods: Audio, visual, and spontaneous expressions," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 31, no. 1, pp. 39–58, 2009.

[8] R. Li and J. Chen, "Multi-modal attention detection using gaze and facial features," [23] IEEE Access, vol. 8, pp. 102345–102356, 2020.

[9] A. Kapoor, W. Burleson, and R. W. Picard, "Automatic prediction of frustration," IEEE Transactions on Affective Computing, vol. 1, no. 1, pp. 1–15, 2010.

[10] S. Fairclough, "Fundamentals of physiological computing," IEEE Interacting with Computers, vol. 21, no. 1–2, pp. 133–145, 200