ABSTRACT

In recent years, maintaining sustained human attention and concentration has become a critical challenge in scenarios such as online examinations, virtual classrooms, remote interviews, and driving safety systems. The lack of physical supervision in such environments often leads to distractions caused by mobile phone usage, drowsiness, noise disturbances, and loss of visual focus. To address this issue, this project presents a Real-Time AI-Based Human Concentration and Distraction Detection System that integrates computer vision, signal processing, and deep learning techniques to monitor user attention effectively and non-intrusively.

The proposed system utilizes MediaPipe Face Mesh technology for accurate facial landmark detection, enabling real-time tracking of eye movements, iris position, head orientation, and blink patterns. Eye Aspect Ratio (EAR)–based geometric algorithms are employed to detect eye blinks and prolonged eye closure, which are strong indicators of drowsiness and reduced alertness. Gaze direction estimation is achieved through iris-based calibration and heuristic threshold analysis to identify whether the user is focused on the screen or distracted by looking away. Additionally, a classical head alignment method is used to monitor face orientation and determine deviations from the attentive posture.

To enhance distraction detection, the system integrates YOLOv8, a state-of-the-art deep learning object detection model, to identify mobile phone usage in real time. Audio-based noise detection is implemented using Root Mean Square (RMS) signal energy analysis, allowing the system to detect surrounding noise levels that may affect concentration. These multi-modal inputs—visual, behavioral, and audio—are combined using a custom weighted attention scoring algorithm to compute a real-time concentration score ranging from 0 to 100.

The system operates entirely offline, ensuring privacy preservation and low latency, making it suitable for deployment on resource-constrained devices. A real-time visual interface provides continuous feedback, alerts, and warnings through on-screen messages and voice notifications. At the end of each session, a comprehensive performance report is automatically generated, summarizing key metrics such as average concentration level, total eye-closure duration, mobile phone usage time, and noise exposure. These results are presented through a web-based dashboard for easy interpretation and analys