Cognishift: A Multimodal Deep Learning Framework for Real-Time Cognitive Load Detection and Adaptive Human-Computer Interaction Optimization

S.H.C. Indeewaree*, G.K.S. Jayalath

General Sir John Kotelawala Defence University, Ratmalana, 10390, Sri Lanka
**42-adc-0010@kdu.ac.lk*

Digital technology is deeply integrated into daily life, yet it contributes to a growing cognitive performance challenge. Global attention spans have decreased by approximately 47% since 2015, and post-COVID cognitive fatigue has increased by 67%, with digital overload costing the global economy an estimated $1 trillion annually in productivity losses. Despite these trends, most human–computer interfaces remain static and unresponsive to the user's cognitive state, leaving cognitive overload chronically undetected and unmanaged. This study proposes CogniShift, a multimodal deep learning framework that detects real-time cognitive load using only a standard webcam and keyboard – requiring no specialised hardware beyond these standard peripherals. Five passive behavioural signals are analysed: eye movement patterns (blink rate, pupil dilation), keystroke timing dynamics (speed, error rate), mouse movement behaviour (velocity, hesitation), fatigue-related facial microexpressions, and scrolling activity patterns. These signals are processed continually to estimate the cognitive state of the user. A hybrid CNN-LSTM network extracts temporal facial fatigue features; a Transformer encoder fuses multimodal behavioural signals; and an XGBoost classifier maps the combined features to five cognitive states, spanning optimal focus to complete cognitive overload. A reinforcement learning-powered adaptive interface module responds in real time by simplifying UI complexity, adjusting content delivery speed, or suggesting optimally-timed micro-breaks. SHAP-based explainability ensures transparency in all model decisions. The system was trained on four public benchmark datasets (SEED-IV, DROZY, MIT Eye Tracking Corpus, TU Berlin Cognitive Load Dataset) and validated on a newly collected dataset of 50 post-COVID participants. CogniShift achieved a classification accuracy of 89.4% with sub-250 ms inference latency, outperforming single-modality baselines by 22–31%. The adaptive interface reduced task completion errors by 34% and self-reported cognitive fatigue by 41% compared to static interfaces ($p < 0.01$). CogniShift demonstrates that passive behavioural signals captured via standard consumer hardware are sufficient for reliable, real-time cognitive monitoring without specialised physiological equipment. These findings have direct applications in digital health, remote work, online education, and assistive technologies.

Keywords: cognitive load detection, adaptive human-computer interaction, multimodal deep learning, real-time fatigue monitoring, explainable AI