**The Potential of Machine Learning in Detecting Cognitive Load**

**Authors:**

\***Q. DANG1**, M. ANORUO1, M. KUCUKOSMANOGLU2, S. CONKLIN2, G. KARGOSHA2, J. BROOKS2;

1UNIV. OF MARYLAND, BALTIMORE COUNTY, UNITED STATE; 2DPRIME.AI, VIRGINIA, UNITED STATE

**Disclosures**:

**Q. Dang**: None. **M. Anoruo:** None, **M. Kucukosmanoglu:** None, **S. Conklin**: None, **G. Kargosha**: None, **J. Brooks**: None

**Abstract:**

Pupil size has been used as an indicator of cognitive load since diameter increases with task difficulty and reflects changes in attention and arousal mediated by the autonomic nervous system. By measuring the dynamic changes in pupil size during different neurocognitive tasks, researchers can detect cognitively loading events and infer the degree of cognitive load. This study investigated the potential of detecting the precise onset of cognitively loading events by analyzing pupil and eye movement responses across a variety of laboratory-based tasks for which stimulus onset was known. Of interest was determining whether there were eye-tracking parameters (i.e. pupillometry and gaze position) that indicated onset of loading across cognitive domains. Participants (n=58) completed the Dot Probe Task (DPT), Mental Arithmetic (MA), Psychomotor Vigilance Task (PVT), and Visual Working Memory (VWM) while pupil size and gaze positions were recorded with a computer-mounted eye tracker. The problem was framed as a binary classification task, labeling 1-second windows as "1" if they occurred immediately after a trial onset, and "0" otherwise. A Convolutional Neural Network (CNN) architecture was employed as the model for training the tasks, and the model's performance was evaluated using the F1 score metric. Among the tasks, the DPT demonstrated the highest F1 score of 0.87, followed by MA (0.70), PVT (0.66), and VWM (0.38). When training all four tasks together, the model achieved an F1 score of 0.54, indicating limited generalization ability. Although the general model underperformed compared to the individual DPT, MA, and PVT tasks, it displayed promising results beyond random guessing. The model utilized both pupil size and gaze position for prediction, with pupil size being the primary driver. The average F1 score using both pupil size and gaze position was 0.72, only slightly dropping to 0.67 when gaze position was excluded. The CNN model effectively detected common patterns in pupil size, where the pupil initially constricted after the onset and followed by a pupil dilation related to cognitive workload. Our preliminary findings suggest that after the onset, the constriction in pupil size indicated participants' attentional shift toward the newly presented problem, with lighting changes making a minor contribution to the constriction. The light changing have a minor contribution to the constriction. The shifting focus took about 0.6 to 0.8 second. Subsequently, the pupil dilation indicated the onset of cognitive workload, with the peak of pupil size and the latency of the peaks strongly influenced by the cognitive workload magnitude and task duration. For short tasks such as PVT, peaks occurred around 0.5 to 1 second after the onset, while medium tasks like VWM and DPT exhibited peaks around 1.5 to 2.5 seconds. Longer tasks like MA typically showed peaks after approximately 3.5 seconds. These findings emphasize the potential of machine learning to accurately predict cognitive load based on pupil and eye movement responses, contributing to advancements in personalized learning and optimizing neurocognitive workload allocation.