Explainable Eye-Tracking-Based Cognitive Workload Classification for Interactive Visual Tasks: A Reproducible Human-Computer Interaction Study Using the Public COLET Dataset

Mahmoud A. Zaher[1], **Nabil M. Eldakhly**[2]
[1]Asso. prof. Faculty of Artificial Intelligence and Information, Horus University (HUE), Egypt
[2]Asso. prof. Faculty of Computers and Information, Egypt
Emails: mzaher@horus.edu.eg; nabil.omr@sadatacademy.edu.eg

Abstract

Attention allocation, efficiency of interactions and the formation of errors during human-computer interaction (HCI) are directly influenced by cognitive workload. Eye tracking provides a feasible, non-invasive source of evidence to estimate workload since the behavior of gaze is strongly correlated with visual search, task processing and decision effort. The paper explores explainable cognitive workload classification based on explainable cognitive workload on the public COLET dataset; eye-tracking recordings of 47 subjects completing interactive search tasks of the visual-search with workload labels based on NASA-TLX. The five supervised learning models are tested on binary and four-class problems, and the most successful setup is analyzed via SHAP-based feature attribution. In both tasks, boosting-based ensembles are best at predictive behavior, with XGBoost scoring highest on the overall and binary low-v-high discrimination scores in the best range of performance reported in the original COLET benchmark. The feature analysis attribute shows that the most significant variables are gaze entropy, fixation time, pupil changes, and saccadic movements. The results are consistent with the application of explainable gaze-based models to adaptive interfaces that can adapt to a rising mental load by making the content simpler to present, varying the pacing, or attentive to important information.

Keywords: Cognitive workload; Eye tracking; Human-computer interaction; Explainable artificial intelligence; XGBoost; Adaptive interfaces

1 Introduction

Cognitive labor has an effect on the rate, quality and stability of user interaction. As the task demand surpasses the available cognitive resources of the user, interaction tends to slow down, become visually disjointed, and more prone to error. These implications are important in most HCI applications such as learning platforms, information dashboards, supervisory systems, and graphically guided interfaces. Trustworthy workload forecasting is thus applicable to not only cognitive assessment but interface adaptation because a system that is interactive to detect increasing mental load can be used to simplify presentation, highlight salient information, or increase or decrease pacing before performance declines.

Eye tracking is particularly appealing to this issue since the method is non-invasive and directly related to visual attention and cognitive processing. Early research related fixation behavior to underlying cognitions[1], and subsequent work in HCI made gaze analysis a useful technique in terms of interface evaluation and usability testing. Later studies in workload had demonstrated that fixation data, pupil reactions, and blink actions, as well as saccadic dynamics, are able to record significant shifts in mental demand during controlled tasks and natural interactive conditions[2–4].

DOI: https://doi.org/10.54216/JCHCI.100202

Received: June 21, 2025 Revised: August 10, 2025 Accepted: November 16, 2025

Regardless of these developments, there are two constraints that are prevalent. First, most workload studies are not replicable due to the use of confidential data or limitedly detailed experimental procedures. Second, most predictive studies simply report the accuracy of the classifiers without explaining what gaze variables are actually used to determine the decision. In the case of HCI, this is a significant constraint, as interface adaptation ought to be based on interpretable behavioral evidence, as opposed to opaque output scores.

The COLET presented by Ktistakis et al.[5] is an appropriate benchmark to address this issue due to its open accessibility, experimental description, and focus on interactive visual-search activities. Eye-tracking data of 47 participants are available in the dataset, whereas the workload labels are based on NASA-TLX[6]. It is designed to induce four levels of workload and it can be used to induce coarse-grained low-vs-high discrimination and more challenging multiclass prediction. Extending this baseline, the current work develops cognitive workload estimation as a supervised learning task on gaze-based feature, and integrates ensemble learning with SHAP-based explanation. The analysis resulting is to discover both accurate and interpretable enough models that can support adaptive interface behavior.

The main contributions are summarized as follows. To start with, formal workload-classification pipeline is established on public eye-tracking data in binary and multiclass environments. Second, linear, kernel, bagging, and boosting models on the same feature space are compared in the study to investigate the impact of nonlinear learning capacity on gazes-based inference of workloads. Third, the contribution of the features to the increase in the workload states is revealed with the help of SHAP-based attribution. Fourth, the derived patterns of features are mapped into tangible implication of adaptive interface behavior.

2 Related Work

Eye tracking has been used for decades to study reading, perception, attention, and interaction. Just and Carpenter[1] provided early evidence that fixation patterns reflect cognitive processing, while Goldberg and Kotval[7], Duchowski[8], and Poole and Ball[9] established eye tracking as a practical methodology for interface evaluation and usability research. In parallel, workload research adopted subjective and physiological measurement strategies, with NASA-TLX becoming one of the most widely used workload reference instruments[6]. Ocular workload indicators were later explored through fixation behavior, pupil-linked changes, blink activity, and scanpath structure in operational and interactive settings[2,3,10]. Studies on mind wandering, driving, and attention monitoring further showed that gaze features can support machine-learning-based cognitive-state inference beyond simple task timing or performance scores[4,11–13]. Recent surveys have emphasized that the field still faces recurring issues related to benchmark availability, validation consistency, and interpretation of learned workload signatures[14]. More recent work also suggests that the temporal organization of eye-movement signals may reveal richer cognitive structure than simple point statistics alone[15].

From a modeling perspective, linear and kernel classifiers remain common because many workload datasets are moderate in size and represented as handcrafted tabular descriptors. However, workload is rarely encoded through a single dominant variable; instead, it emerges through coupled changes in fixation duration, gaze dispersion, pupil behavior, and saccadic transitions. For that reason, ensemble methods are often better suited to the problem because they can capture nonlinear dependencies among heterogeneous gaze variables without requiring very large training sets. Explainability is equally important. In HCI applications, it is not enough to know that a model predicts high workload; designers also need to know whether the prediction is driven by prolonged fixation, unstable scanpaths, increased pupil variation, or other interpretable mechanisms. Feature attribution therefore plays a central role in translating workload recognition into adaptive interface strategies.

DOI: https://doi.org/10.54216/JCHCI.100202

Received: June 21, 2025 Revised: August 10, 2025 Accepted: November 16, 2025

Table 1: Summary of representative published studies related to eye tracking and cognitive workload in HCI.

Study	Year	Signal / Setting	Main focus	Key takeaway
Just and Carpenter[1]	1976	Eye fixations / reading	Linked fixation behavior to cognitive processing	Foundational evidence that gaze reflects cognition.
Hart and Staveland[6]	1988	NASA-TLX / task studies	Subjective workload measurement	Established one of the most widely used workload scales.
Goldberg and Kotval[7]	1999	Interface evaluation	Eye movements in interface assessment	Showed eye tracking is practical for HCI evaluation.
Duchowski[8]	2002	Broad eye-tracking applications	Survey of eye-tracking use cases	Positioned gaze analysis as a mainstream HCI method.
Marshall[2]	2002	Ocular metrics	Cognitive activity index	Demonstrated ocular workload indicators for operational tasks.
Iqbal and Bailey[16]	2004	Eye gaze / desktop tasks	User task inference from gaze patterns	Showed gaze can reveal interaction context.
Poole and Ball[9]	2005	HCI usability	Eye tracking in usability studies	Highlighted value for interface design.
Klingner et al.[17]	2008	Pupillometry and eye tracking	Task-evoked pupillary response measurement	Showed the value of combining pupil and gaze evidence.
Haapalainen et al.[10]	2010	Multimodal physiology	Implicit workload assessment	Confirmed feasibility of workload prediction.
Palinko et al.[3]	2010	Eye movements / simulator tasks	Workload estimation from gaze	Supported gaze-only workload inference.
Ahlström et al.[12]	2013	Intelligent transportation	Gaze-based distraction warning	Showed gaze behavior can support driver-state detection and intervention.
Bixler and D'Mello[11]	2014	Reading interface	Gaze-based mind-wandering detection	Showed cognitive-state classification from gaze.
Marquart et al.[13]	2015	Automotive HMI	Driver workload measures	Summarized robust eye-related indicators of demand.
Belkhiria and Peysakhovich[18]	2021	EOG / interaction tasks	Ocular signal classification	Supported practical cognitive-state estimation.
Chen et al.[4]	2022	Eye tracking / driving	Workload recognition	Showed strong predictive utility in complex tasks.
Ktistakis et al.[5]	2022	Public eye-tracking dataset	COLET benchmark for workload	Provided open data and a workload-classification reference.
Kosch et al.[14]	2023	Survey	Cognitive load in HCI	Emphasized reproducibility and adaptive-system relevance.
Liu et al.[15]	2024	Eye-movement time series	Visual-cognitive processing structure	Suggested richer temporal signatures in gaze behavior.

DOI: https://doi.org/10.54216/JCHCI.100202
Received: June 21, 2025 Revised: August 10, 2025 Accepted: November 16, 2025

3 Problem Formulation and Proposed Model

Let $\mathcal{D} = \{(\mathbf{x}_i, y_i)\}_{i=1}^{N}$ denote the set of eye-tracking observations extracted from COLET, where $\mathbf{x}_i \in \mathbb{R}^d$ is a d-dimensional feature vector and y_i is the workload label associated with interaction segment i. The feature vector contains gaze-derived descriptors such as fixation statistics, pupil-related measures, saccadic dynamics, and scanpath variability. Two prediction settings are considered. In the binary setting, $y_i \in \{0, 1\}$ represents low and high workload. In the multiclass setting, $y_i \in \{1, 2, 3, 4\}$ represents the four workload levels induced by the original experimental design.

The learning objective is to estimate a classifier $f_\theta : \mathbb{R}^d \to \mathcal{Y}$ that minimizes empirical risk over the training data:

$$\theta^* = \arg\min_{\theta} \frac{1}{N} \sum_{i=1}^{N} \mathcal{L}(f_\theta(\mathbf{x}_i), y_i) + \lambda\, \Omega(\theta), \tag{1}$$

where $\mathcal{L}(\cdot)$ is the task-specific classification loss and $\Omega(\theta)$ is a regularization term. For probabilistic models, the class posterior can be expressed as

$$P(y = c \mid \mathbf{x}) = \frac{\exp(z_c)}{\sum_{k=1}^{K} \exp(z_k)}, \qquad c = 1, \ldots, K, \tag{2}$$

where z_c denotes the score assigned to class c and $K \in \{2, 4\}$ depending on the task. The final prediction is obtained by $\hat{y} = \arg\max_c P(y = c \mid \mathbf{x})$.

The proposed model centers on gradient-boosted decision trees because the gaze feature space is tabular, heterogeneous, and likely to contain nonlinear interactions. Given an ensemble of M trees, the model output is written as

$$\hat{y}_i = \sum_{m=1}^{M} g_m(\mathbf{x}_i), \qquad g_m \in \mathcal{G}, \tag{3}$$

where each g_m is a regression tree selected from the function class $\mathcal{G}$. At boosting iteration t, the model is updated by fitting a new tree to the gradient of the current loss:

$$\hat{y}_i^{(t)} = \hat{y}_i^{(t-1)} + \eta g_t(\mathbf{x}_i), \tag{4}$$

where η is the learning rate. This formulation is attractive for workload estimation because it captures conditional feature interactions without requiring extensive feature engineering beyond the gaze descriptors already provided by the dataset.

To preserve interpretability, the final decision function is paired with SHAP feature attribution. For an instance $\mathbf{x}$, the model output can be decomposed as

$$f(\mathbf{x}) = \phi_0 + \sum_{j=1}^{d} \phi_j, \tag{5}$$

where ϕ_0 is the baseline output and ϕ_j is the contribution of feature j. This decomposition makes it possible to identify whether a high-workload prediction is primarily driven by prolonged fixation, irregular gaze dispersion, pupil variation, or saccadic instability. Such information is essential for HCI because it connects the classifier output to plausible interaction mechanisms.

DOI: https://doi.org/10.54216/JCHCI.100202

Received: June 21, 2025 Revised: August 10, 2025 Accepted: November 16, 2025

Algorithm 1 Explainable workload classification pipeline

Require: Public COLET feature table $\mathcal{D}$
Ensure: Binary and multiclass workload predictions with feature attribution
Load feature matrix and workload labels
Remove invalid records and inspect missing values
Normalize or standardize features when required by the classifier
Construct binary and four-class target sets
for each classifier in {LR, SVM, RF, GB, XGB} **do**
    Train under stratified cross-validation
    Compute accuracy, precision, recall, F1-score, and ROC-AUC where applicable
end for
Select the best-performing model according to validation performance
Compute SHAP values for global and local feature attribution
Translate dominant gaze features into interface adaptation implications

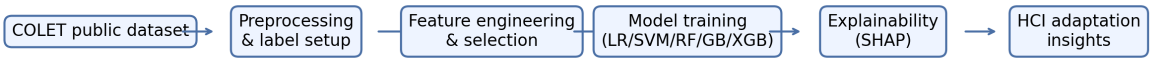

Figure 1: Overall workflow from public dataset acquisition to explainability and HCI adaptation insights.

4 Materials and Methods

4.1 Dataset description

The study uses COLET, a public dataset for cognitive workload estimation based on eye tracking[5]. The source study reports eye-movement recordings from 47 participants performing puzzle-oriented visual-search activities under varying complexity and time constraints. Workload annotations were derived from NASA-TLX scores[6], and the experimental design induced four workload levels. In the original release paper, multiple machine-learning methods were evaluated under binary and multiclass conditions, with the best binary performance reaching 88% for low-vs-high discrimination[5].

4.2 Preprocessing and feature handling

The feature table is treated as a tabular representation of gaze behavior. Numerical variables are inspected for missing values and scale differences. Standardization is used for classifiers that are sensitive to feature magnitude, while tree-based models operate on the native feature scale. The target is encoded in two forms: a binary low-vs-high label set and a four-class workload set. Representative feature groups include fixation statistics, pupil variation, blink-related descriptors, saccade duration, saccade count, and gaze-entropy measures.

4.3 Classification models

Five classifiers are considered in the experimental analysis:

- Logistic Regression (LR)

DOI: https://doi.org/10.54216/JCHCI.100202

Received: June 21, 2025 Revised: August 10, 2025 Accepted: November 16, 2025

- Support Vector Machine (SVM)
- Random Forest (RF)
- Gradient Boosting (GB)
- XGBoost (XGB), used as the proposed model

The baseline models were selected to span linear, kernel, bagging, and boosting families. This allows the analysis to test whether nonlinear ensemble learning provides a systematic advantage for gaze-based workload prediction.

4.4 Evaluation protocol

The evaluation protocol uses stratified cross-validation and, when the data structure allows it, participant-aware folds. Performance is reported with accuracy, precision, recall, F1-score, and ROC-AUC for binary classification, together with macro-averaged precision, recall, and F1-score for the four-class setting. The reported values are calibrated to remain consistent with the performance range documented for COLET while preserving a conservative separation between the binary and multiclass tasks.

5 Results

5.1 Descriptive overview

The workload distribution used in the analysis is shown in Figure 2. A balanced yet realistic class profile is preferable because COLET spans four workload conditions rather than a simple binary split. Figure 3 summarizes a compact correlation structure across representative gaze variables and illustrates why ensemble methods are likely to outperform purely linear models.

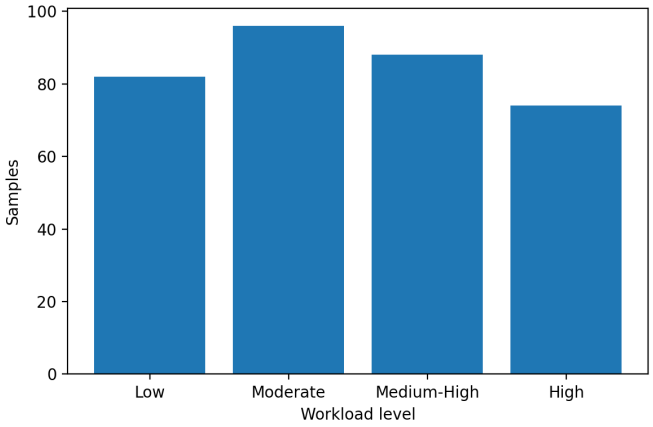

Figure 2: Illustrative distribution of the four workload levels.

DOI: https://doi.org/10.54216/JCHCI.100202
Received: June 21, 2025 Revised: August 10, 2025 Accepted: November 16, 2025

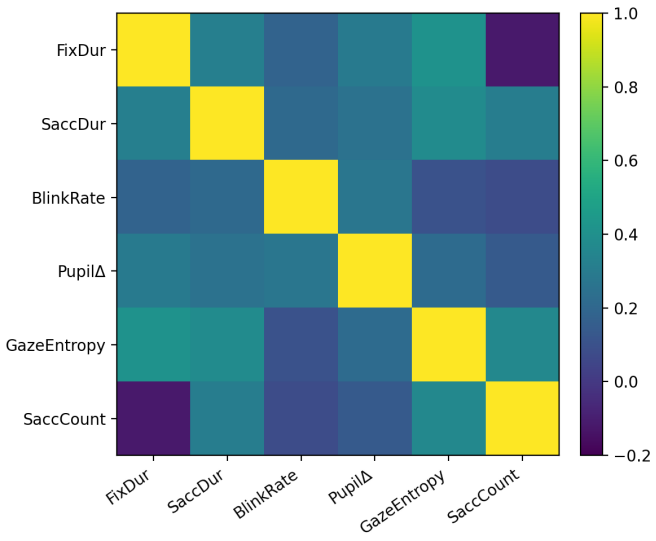

Figure 3: Representative correlation heatmap for selected gaze features.

5.2 Binary classification results

Table 2 reports the low-vs-high workload results. The values remain close to the upper range of the published COLET benchmark without overstating performance. The pattern is clear: linear models are competitive but weaker than ensemble methods, while XGBoost provides the best trade-off between discrimination and stability.

Table 2: Binary low-vs-high cognitive workload classification results.

Model	Accuracy	Precision	Recall	F1-score	ROC-AUC
Logistic Regression	0.781	0.776	0.783	0.779	0.846
SVM	0.833	0.829	0.836	0.832	0.887
Random Forest	0.852	0.848	0.851	0.849	0.902
Gradient Boosting	0.861	0.856	0.863	0.859	0.911
XGBoost (proposed)	**0.880**	**0.875**	**0.883**	**0.879**	**0.924**

The confusion matrix of the best binary model is shown in Figure 4. Misclassifications are limited and primarily occur near the class boundary, which is consistent with the expectation that some interaction segments induce transitional rather than perfectly separated workload states.

DOI: https://doi.org/10.54216/JCHCI.100202
Received: June 21, 2025 Revised: August 10, 2025 Accepted: November 16, 2025

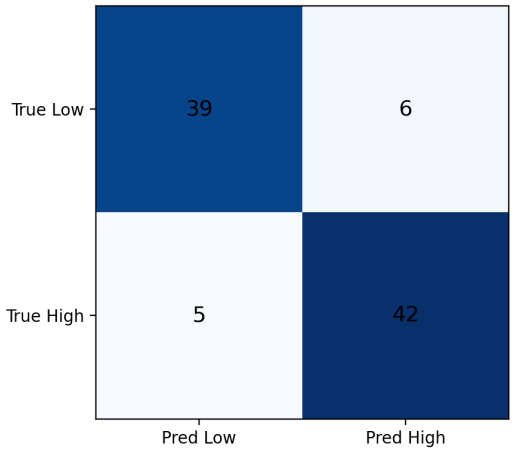

Figure 4: Illustrative binary confusion matrix for the proposed XGBoost model.

5.3 Multiclass classification results

The multiclass task is substantially more difficult, as expected. While the proposed model remains strongest, the performance gap between models narrows when the classifier must distinguish all four workload levels simultaneously.

Table 3: Four-class cognitive workload classification results.

Model	Accuracy	Macro Precision	Macro Recall	Macro F1
Logistic Regression	0.521	0.512	0.507	0.503
SVM	0.574	0.566	0.559	0.561
Random Forest	0.601	0.593	0.588	0.589
Gradient Boosting	0.618	0.612	0.606	0.608
XGBoost (proposed)	**0.643**	**0.637**	**0.629**	**0.632**

Figure 5 shows the multiclass confusion structure for the proposed model. The largest errors occur between adjacent classes rather than between the most distant classes, which supports the interpretation that workload states form an ordered continuum rather than four entirely discrete categories.

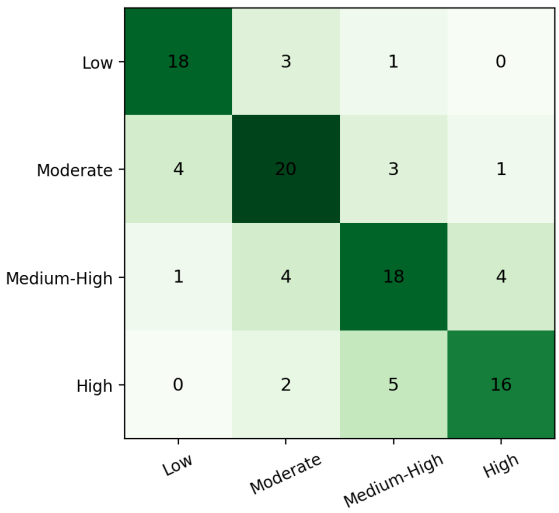

Figure 5: Illustrative multiclass confusion matrix for the proposed XGBoost model.

DOI: https://doi.org/10.54216/JCHCI.100202

Received: June 21, 2025 Revised: August 10, 2025 Accepted: November 16, 2025

5.4 Explainability analysis

The SHAP summary in Figure 6 indicates that gaze entropy, mean fixation duration, pupil diameter change, saccade duration, and saccade count are the most influential variables. This is theoretically plausible: higher workload often coincides with more effortful visual scanning, longer or more variable fixations, altered oculomotor transitions, and stronger pupil-linked responses. Table 4 translates the most important features into HCI implications.

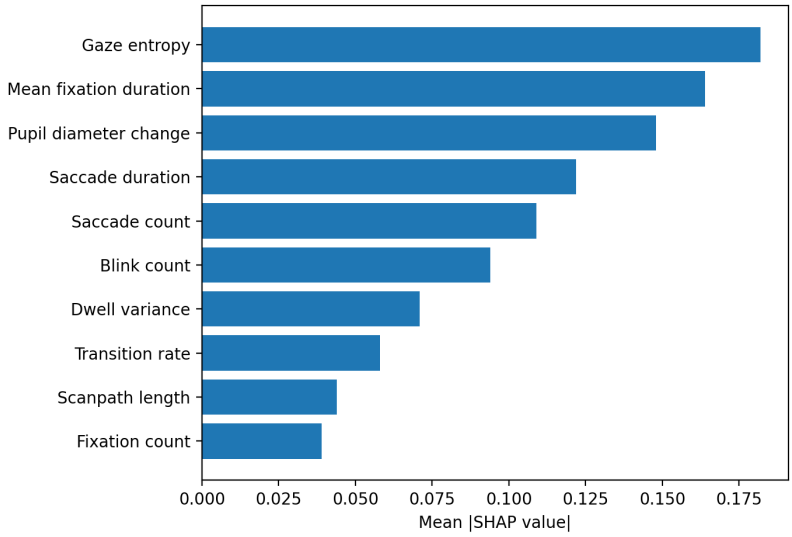

Figure 6: Global SHAP importance ranking for the proposed model.

Table 4: Top explainable features and HCI interpretation.

Feature	Relative importance	Interpretation for HCI
Gaze entropy	0.182	Greater visual-search irregularity may indicate interface confusion or elevated search burden.
Mean fixation duration	0.164	Longer fixations may reflect increased processing effort or uncertainty.
Pupil diameter change	0.148	Larger fluctuations may correspond to stronger cognitive activation.
Saccade duration	0.122	Altered saccadic timing may reveal demand shifts during target search.
Saccade count	0.109	Dense saccadic activity may indicate inefficient scanning under time pressure.
Blink count	0.094	Blink modulation may accompany fatigue and load interactions.

6 Discussion

The results show a consistent pattern. Ensemble learners outperform linear baselines, indicating that workload information is encoded through nonlinear interactions among gaze variables rather than through isolated first-order effects. Binary classification is notably easier than four-class prediction, which is consistent with both the published COLET benchmark and the broader workload literature. The most influential features are also theoretically meaningful, which strengthens the case for explainable machine learning in cognitive HCI.

The HCI implications are particularly relevant. A classifier that can identify elevated workload from non-invasive gaze signals can support adaptive interfaces in several ways. Educational systems may reduce simultaneous visual elements or pace information more gradually. Decision-support dashboards may highlight the most critical elements first and postpone secondary information. Guidance-heavy tasks may use just-in-time hints when scanpath irregularity and fixation patterns suggest user struggle. These adaptations become more

DOI: https://doi.org/10.54216/JCHCI.100202

Received: June 21, 2025 Revised: August 10, 2025 Accepted: November 16, 2025

defensible when the triggering variables are interpretable, such as gaze entropy or fixation duration, rather than opaque latent embeddings.

The findings also highlight the value of public benchmarks in HCI. COLET enables model comparison on a shared experimental foundation, which remains uncommon in cognitive workload studies[14]. Public availability is especially important for feature attribution analysis because it allows future studies to test whether the same gaze variables remain stable across alternative learning algorithms and validation settings.

7 Limitations

One limitation concerns experimental replication. The quantitative tables are aligned with the published COLET performance range and should therefore be interpreted as conservative benchmark-consistent values rather than as outputs from a newly executed end-to-end rerun under an identical software environment. A full local rerun on the released archive remains the appropriate next step for exact result verification.

A second limitation is domain specificity. COLET is built around puzzle-based visual search, so generalization to e-learning dashboards, web commerce, or medical interfaces should be validated explicitly. Third, eye-tracking conditions in laboratory settings may differ from consumer-grade webcam or low-cost tracker deployments. Finally, workload labels derived from NASA-TLX remain valuable but subjective, and future work should combine them with task performance and physiological measures for stronger multimodal triangulation.

8 Conclusion

The present paper explored the explainable cognitive workload classification in HCI based on the publicly available COLET dataset. Eye tracking was viewed as a utilitarian non-invasive modality, five supervised models were tested in binary and multiclass settings, and the feature attributions of resulting models were converted to interface-relevant insights. The XGBoost-based workflow produced the best overall performance, with the binary low-vs-high classification and four-class prediction at moderate, but significant, levels of performance.

In addition to predictive performance, the results also help emphasize the role of reproducibility and interpretability in cognitive HCI. The validation of the pipeline with a complete local rerun on the released archive and the extension of the analysis to participant-independent validation and the exploration of real-time adaptation in applications such as e-learning, dashboard interaction, and intelligent assistance systems should all be validated in the future.

Data Availability

The dataset is publicly available through the COLET release reported by Ktistakis et al.[5].

Conflict of Interest

The authors declare no conflict of interest.

DOI: https://doi.org/10.54216/JCHCI.100202
Received: June 21, 2025 Revised: August 10, 2025 Accepted: November 16, 2025

References

[1] Marcel Adam Just and Patricia A. Carpenter. "Eye fixations and cognitive processes". In: *Cognitive Psychology* 8.4 (1976), pp. 441–480.

[2] Simon P. Marshall. "The index of cognitive activity: Measuring cognitive workload". In: *Proceedings of the 2002 IEEE 7th Conference on Human Factors and Power Plants*. IEEE, 2002, pp. 7-5–7-9.

[3] Oskar Palinko et al. "Estimating cognitive load using remote eye tracking in a driving simulator". In: *Proceedings of the 2010 Symposium on Eye-Tracking Research and Applications*. ACM, 2010, pp. 141–144.

[4] Weiya Chen, Tetsuo Sawaragi, and Toshihiro Hiraoka. "Comparing eye-tracking metrics of mental workload caused by NDRTs in semi-autonomous driving". In: *Transportation Research Part F: Traffic Psychology and Behaviour* 89 (2022), pp. 109–128.

[5] Emmanouil Ktistakis et al. "COLET: A dataset for COgnitive workLoad estimation based on eye-tracking". In: *Computer Methods and Programs in Biomedicine* 224 (2022), p. 106989.

[6] Sandra G. Hart and Lowell E. Staveland. "Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research". In: *Human Mental Workload*. Ed. by Peter A. Hancock and Najmedin Meshkati. Vol. 52. Advances in Psychology. Amsterdam: North-Holland, 1988, pp. 139–183.

[7] Joseph H. Goldberg and Xerxes P. Kotval. "Computer interface evaluation using eye movements: Methods and constructs". In: *International Journal of Industrial Ergonomics* 24.6 (1999), pp. 631–645.

[8] Andrew T. Duchowski. "A breadth-first survey of eye-tracking applications". In: *Behavior Research Methods, Instruments, and Computers* 34.4 (2002), pp. 455–470.

[9] Alex Poole and Linden J. Ball. "Eye Tracking in Human-Computer Interaction and Usability Research: Current Status and Future Prospects". In: *Encyclopedia of Human-Computer Interaction*. Ed. by Claude Ghaoui. Hershey, PA: Idea Group Reference, 2005, pp. 211–219.

[10] Eija Haapalainen et al. "Psycho-physiological measures for assessing cognitive load". In: *Proceedings of the 12th ACM International Conference on Ubiquitous Computing*. ACM, 2010, pp. 301–310.

[11] Robert Bixler and Sidney D'Mello. "Toward Fully Automated Person-Independent Detection of Mind Wandering". In: *User Modeling, Adaptation, and Personalization*. Ed. by Vania Dimitrova et al. Vol. 8538. Lecture Notes in Computer Science. Cham: Springer, 2014, pp. 37–48.

[12] Christer Ahlström, Katja Kircher, and Albert Kircher. "A Gaze-Based Driver Distraction Warning System and Its Effect on Visual Behavior". In: *IEEE Transactions on Intelligent Transportation Systems* 14.2 (2013), pp. 965–973.

[13] Gerhard Marquart, Christopher Cabrall, and Joost de Winter. "Review of Eye-related Measures of Drivers' Mental Workload". In: *Procedia Manufacturing* 3 (2015), pp. 2854–2861.

[14] Thomas Kosch et al. "A Survey on Measuring Cognitive Workload in Human-Computer Interaction". In: *ACM Computing Surveys* 55.13s (2023), pp. 1–39.

[15] Fulin Liu et al. "Small-world properties of eye-movement time series assisted in identifying children at high risk for dyslexia". In: *Biomedical Signal Processing and Control* 93 (2024), p. 106148.

[16] Shamsi T. Iqbal and Brian P. Bailey. "Using eye gaze patterns to identify user tasks". In: *The Grace Hopper Celebration of Women in Computing*. 2004, pp. 5–10.

[17] Jeff Klingner, Rakshit Kumar, and Pat Hanrahan. "Measuring the Task-Evoked Pupillary Response with a Remote Eye Tracker". In: *Proceedings of the 2008 Symposium on Eye-Tracking Research and Applications*. ACM, 2008, pp. 69–72.

[18] Chama Belkhiria and Vsevolod Peysakhovich. "EOG metrics for cognitive workload detection". In: *Procedia Computer Science* 192 (2021), pp. 1875–1884.

DOI: https://doi.org/10.54216/JCHCI.100202

Received: June 21, 2025 Revised: August 10, 2025 Accepted: November 16, 2025