Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

# Ocular Metrics for Fatigue Assessment: A Survey from Physiology to Machine Learning

**Yaxiong Lei**

yaxiong.lei@essex.ac.uk

University of Essex

**Parastoo Azizinezhad**

University of Essex

**Mohammadreza Jamalifard**

University of Essex

**Sanjay Manohar**

University of Oxford

**Michal Wlodarski**

Solvemed.ai

**Tom Foulsham**

University of Essex

**Javier Andreu-Perez**

University of Essex

**Research Article**

**Keywords:**

**Posted Date:** November 24th, 2026

**DOI:** https://doi.org/10.21203/rs.3.rs-9116014/v1

**License:** This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

**Additional Declarations:** No competing interests reported.

# Ocular Metrics for Fatigue Assessment: A Survey from Physiology to Machine Learning

Yaxiong Lei[a,*], Parastoo Azizinezhad[a], Mohammadreza Jamalifard[a], Sanjay Manohar[b], Michal Wlodarski[c], Tom Foulsham[a], Javier Andreu-Perez[a,*]

[a] *University of Essex, United Kingdom*
[b] *University of Oxford, United Kingdom*
[c] *Solvemed Ltd, United States*

**Abstract**

Fatigue is pervasive and safety-critical, yet objective biomarkers remain limited, resulting in a gap between subjective reports and clinical decision-making. The eye provides a non-invasive readout of central nervous system function: changes in eyelid behaviour, eye movements, and pupillary dynamics track arousal, executive control, and visual strain. This PRISMA-guided, PROSPERO-registered review integrates work in medicine, psychology, and neuroscience with computational methods to map ocular biomarkers to fatigue phenotypes and their detection pipelines. We synthesize evidence linking eyelid metrics, saccadic and pursuit dynamics, fixational stability, and pupillometry to drowsiness, cognitive/mental fatigue, visual fatigue, and clinical fatigue states, then trace how computer vision and machine learning operationalize these signals.

A dominant literature focus on driver drowsiness emerges: eyelid-based indicators are mature and robust, with metrics like increased blink duration and a higher Percentage of Eye Closure (PERCLOS) strongly correlating with performance lapses and microsleeps, while cognitive, visual, and clinical fatigue remain comparatively under-explored in computer science despite richer treatment in clinical domains. Methodologically, the field is shifting from feature engineering to end-to-end and sequence-aware models, but progress is constrained by heterogeneous ground truth labels, scarce public benchmarks, and lab-to-wild generalization gaps. We conclude with a roadmap advocating multimodal, privacy-preserving systems; standardized, phenotype-aware benchmarks; and validation that combines behavioural performance, physiology, and self-report, to enable holistic, ethically grounded fatigue monitoring beyond driving scenarios.

*Keywords:* Fatigue assessment, Ocular biomarkers, Eye movements, Pupillometry, PERCLOS, Drowsiness detection, Cognitive fatigue, Vigilance, Machine learning, Deep learning, Privacy-preserving monitoring

## 1. Introduction

Fatigue represents one of the most common yet poorly understood symptoms in clinical medicine and occupational health, constituting a pervasive complaint that negatively impacts work performance, safety, and quality of life [80]. Despite its ubiquity, affecting an estimated 20% of adults globally [183], it remains a complex phenomenon characterized by a persistent "subjective-objective gap"—a chasm between an individual's profound, disabling experience and the frequent absence of specific, objective biomarkers [41]. The consequences of this gap are significant, not only creating diagnostic challenges but also posing critical threats in high-risk professions like transportation, aviation, and healthcare, where a momentary lapse in alertness can have catastrophic outcomes [230]. The staggering

*Corresponding author

*Email addresses:* yaxiong.lei@essex.ac.uk (Yaxiong Lei), p.azizinezhad@essex.ac.uk (Parastoo Azizinezhad), m.jamalifard@essex.ac.uk (Mohammadreza Jamalifard), sanjay.manohar@ndcn.ox.ac.uk (Sanjay Manohar), michal.wlodarski@solvemed.ai (Michal Wlodarski), foulsham@essex.ac.uk (Tom Foulsham), j.andreu-perez@essex.ac.uk (Javier Andreu-Perez)

economic impact, with fatigue-related productivity losses exceeding $136 billion annually in the United States alone, underscores the pressing need for reliable, non-invasive methods to monitor fatigue in real-time [183].

In response to this challenge, the human eye has emerged as a uniquely accessible window into the functional state of the central nervous system [153]. While the pioneering work of Hess and Polt [91] and Kahneman and Beatty [115] first established pupillometry as a powerful index of cognitive effort, subsequent research has expanded this focus to a full suite of ocular behaviours. The degradation of these functions during fatigue is not merely symptomatic but represents a direct manifestation of the underlying neurophysiological processes. Subtle yet quantifiable changes across the oculomotor system—including the slowing of rapid eye movements (*saccades*), the lengthening of gaze periods (*fixations*), altered *blink* dynamics, and attenuated *pupillary* responses—are directly linked to the brain's arousal systems, executive control networks, and overall cognitive state. Together, these ocular parameters serve as powerful indicators of an individual's level of alertness and engagement [4]. The ability to measure this rich set of parameters non-invasively and continuously makes them ideal for developing real-world monitoring applications [119].

The study of fatigue and its ocular manifestations represents a unique convergence of multiple disciplines. Medicine provides the understanding of fatigue's pathophysiology; psychology contributes insights into its cognitive and emotional dimensions; and computer science and engineering deliver the technological tools and analytical methods necessary to objectively measure and model fatigue states from ocular data [96]. Despite significant progress, the translation of laboratory findings into robust, real-world applications remains a considerable challenge, impeded by confounding variables, individual differences, and the complexities of uncontrolled environments [154].

To address these challenges, this article presents a comprehensive systematic review bridging foundational sciences with modern technological applications. The main contributions of this work are summarized as follows:

- **Interdisciplinary Synthesis:** We integrate neurophysiological evidence with computational frameworks, mapping specific ocular biomarkers (e.g., saccades, task-evoked pupillary response (TEPR)) to distinct fatigue phenotypes (e.g., cognitive, visual, clinical) rather than treating fatigue as a monolithic concept.
- **Bibliometric Analysis:** Analysing 379 studies, we provide a quantitative visualization of research trends, identifying the dominance of drowsiness research and the specific gap in cognitive fatigue modelling.
- **Technological Taxonomy:** We categorize detection frameworks not just by algorithm, but by the *signal* they model (eyelid vs. oculomotor vs. pupillary), explaining the biological rationale for model selection (e.g., why RNNs suit saccadic sequences).
- **Strategic Roadmap:** We propose concrete future directions including multimodal fusion, privacy-preserving federated learning, and phenotype-aware benchmarking to move the field from lab-based drowsiness detection to ubiquitous fatigue monitoring.

The remainder of this paper is organized as follows: Section 3 details the survey methodology. Section 4 and Section 5 provide the definitional and biological foundations. Section 6 outlines experimental protocols. Section 7 surveys computational detection frameworks. Finally, Section 8 discusses future directions and open challenges.

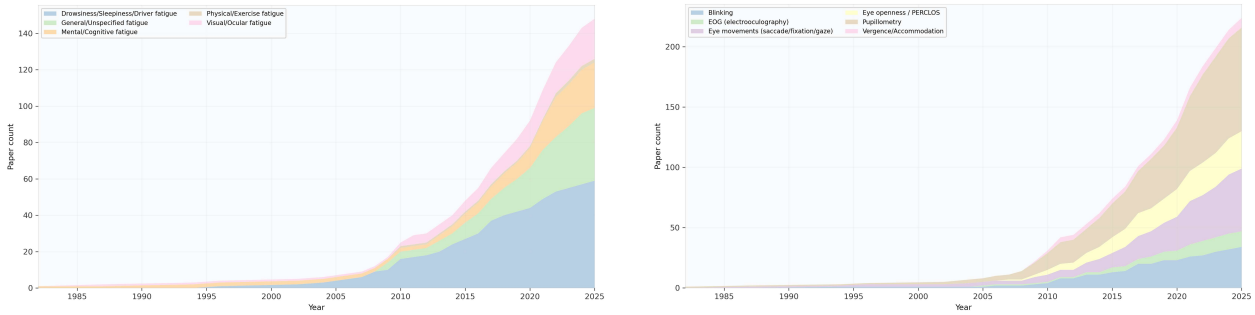
## 2. Related Work

The assessment of fatigue through ocular metrics is a mature and active field of research, spanning disciplines from neuroscience to computer science. To situate the present work and clarify its unique contribution, this section provides a critical analysis of the existing literature. We begin by examining the field's publication trends to demonstrate its growth and shifting focus. We then critique previous review articles, categorizing them by their primary focus to reveal a landscape rich in specialized knowledge but also marked by significant fragmentation. This analysis allows us to precisely identify the research gap that this review aims to fill.

### 2.1. Publication Trends and Research Focus

The scientific interest in ocular fatigue assessment has grown substantially over the past few decades, with a notable acceleration in recent years. This trend is illustrated in Figure 1, which plots the cumulative publication counts from the final corpus of 379 studies identified through the systematic search detailed in Section 3. This steep upward trajectory, beginning around 2010, reflects both an increased recognition of fatigue as a critical safety issue and advancements in the accessibility of eye-tracking technology. An analysis of research focus reveals two key trends.

First, the types of fatigue studied (Figure 1a) show a particular research emphasis on **Drowsiness/Sleepiness**, which constitutes the largest single category, driven primarily by its direct application to driver safety. Second, a similar growth pattern is evident in the specific ocular biomarkers investigated (Figure 1b). While early research was distributed across various metrics, the post-2010 acceleration saw a dramatic rise in studies utilizing `pupillometry` and `eye movements` (saccade/fixation/etc.). Taken together, these trends suggest a paradigm shift within the field, moving beyond established drowsiness-related metrics towards biomarkers that can provide deeper insights into cognitive states and mental workload.




(a) Cumulative publications by fatigue type.

(b) Cumulative publications by ocular biomarker.

Figure 1: Cumulative growth of publications over time, showing a significant increase in research volume starting around 2010. (a) Research categorized by fatigue type reveals a dominant focus on Drowsiness/Sleepiness. (b) Research categorized by ocular biomarker highlights the recent and rapid increase in studies focused on Pupillometry and Eye movements.

### 2.2. Analysis of Existing Reviews

Previous reviews on ocular fatigue detection can be broadly categorized into three main types: those with a domain-specific focus, a biomarker-specific focus, or a technology-specific focus. While providing excellent depth within their niches, this specialization has led to a fragmented literature.

A substantial portion of reviews is dedicated to the domain-specific application of fatigue detection in automotive safety [96, 4, 189, 202, 240, 130, 238, 85, 118, 144, 7, 143]. These works provide comprehensive summaries of methods for detecting driver drowsiness, meticulously documenting the use of eyelid metrics like PERCLOS and facial cues like yawning [96]. While invaluable, this narrow focus often conflates "fatigue" with "drowsiness," leaving biomarkers for cognitive fatigue (e.g., saccadic velocity) under-represented.

A second category focuses in-depth on a specific `biomarker class`. Extensive reviews exist on `pupillometry` as an index of cognitive effort [153, 223] and on `eye movements` (saccades and fixations) as tools to study attention, gaze interaction and protect privacy [182, 137, 146, 33, 139, 51, 74, 134, 136, 135]. By treating these ocular systems as separate fields of study, they often fail to present an integrated view of how they collectively reflect the multifaceted nature of fatigue. This may hinder the development of detection systems which often target multiple indicators.

With the rise of AI, a third category of `technology-specific` reviews has emerged, summarizing the algorithmic evolution from traditional machine learning to deep learning [96, 137, 67]. A frequent limitation is that these are disconnected from the foundational science; ocular metrics are often treated as abstract "input features" with little discussion of the underlying neurophysiology that explains `why` they are predictive. This can lead to a "black box" approach where engineering is divorced from biological understanding.

### 2.3. Identified Research Gap

Our critical analysis reveals a field rich in specialized knowledge but lacking an integrated, interdisciplinary synthesis. The literature is fragmented along the lines of application domains, specific biomarkers, and computational methodologies. This has created a significant research gap: there is no single, comprehensive review that provides a cohesive narrative flowing from the foundational neurophysiology of the full spectrum of ocular biomarkers to the state-of-the-art in their technological application across relevant domains.

Furthermore, while the principle of "task-biomarker specificity"—the observation that certain biomarkers are optimal for different fatigue types—is well-established, the distribution of research effort across these pairings has not been systematically mapped. To address this, we conducted a co-occurrence analysis on the 379 papers identified in our systematic review, coding each for the fatigue types and ocular biomarkers investigated. The results are visualized in Figure 2. This diagram illustrates the frequency of research focus, where the thickness of a ribbon is proportional to the number of studies examining a specific fatigue-biomarker relationship; a single study can contribute to multiple ribbons if it investigates multiple modalities. The visualization makes dominant research patterns clear. For example, the thick ribbon connecting Drowsiness to Blinking and Eye openness/PERCLOS reflects the extensive investigation of these metrics for sleep-related fatigue, clarifying that the figure demonstrates research trends rather than proving biomarker efficacy on its own.

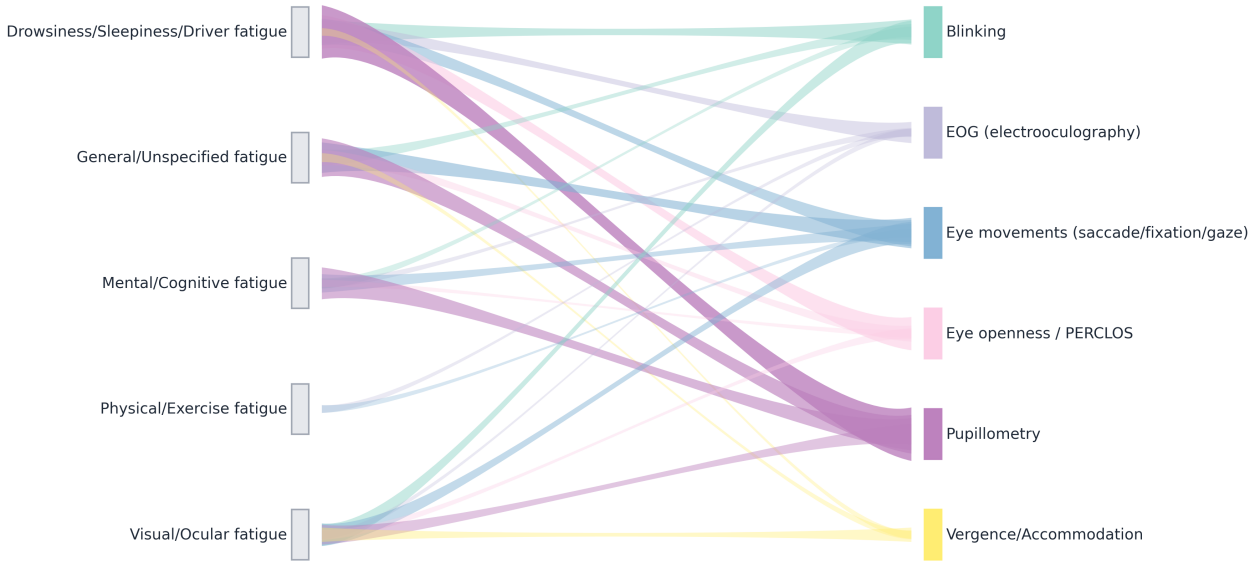



Figure 2: Sankey diagram illustrating the co-occurrence frequency between fatigue types (left) and ocular biomarkers (right) across the 379 studies included in this review. The width of each ribbon is proportional to the number of papers that investigated a given fatigue-biomarker pairing (a single paper can contribute to multiple ribbons). This visualization highlights the dominant research trends in the field, such as the strong focus on eyelid-related metrics for drowsiness.

*Scope and Novelty.* This review therefore integrates *ocular biomarker neurophysiology* with *algorithmic detection frameworks* across fatigue phenotypes. Unlike domain-specific drowsiness or single-biomarker reviews, we (i) map biomarker families to fatigue phenotypes, and (ii) connect neurophysiology to model design/normalization pipelines. Because of the large range of papers and outcome measures, we do not attempt a meta-analysis of effect sizes.

By providing this synthesis, this work aims to serve as a foundational reference for researchers in computer science and related areas fostering a more integrated understanding of the field and charting a clear course for future research.

## 3. Methodology

We conducted this systematic review following the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) 2020 guidelines to ensure transparent reporting of search, screening, and inclusion decisions [165].

Table 1: Search Terms Used in the Systematic Review

| Concept | Search Terms |
|---|---|
| **Ocular Metrics** | pupil*, pupillometr*, pupillar*, pupillograph*, pupil diameter, pupillary unrest, pupillary light reflex, pupil dilation, pupil size, pupil dynamics, pupillary response, pupil fluctuation, pupillary instability, pupil amplitude, blink, fixation, saccade, microsaccade, drift, oculomotor, eye state, eye movement, eye tracking |
| **Fatigue/Vigilance** | fatigue, mental fatigue, physical fatigue, cognitive fatigue, sleepiness, drowsiness, sleep deprivation, vigilance, driver fatigue, task disengagement, cognitive load, workload, attention, sustained attention, alertness, arousal, attention lapse, mind wandering, time on task, memory, performance, performance impairment |
| **Study Design** | detect*, monitor*, measur*, assess*, evaluat*, analys*, model*, algorithm*, experiment*, trial*, survey*, review* |

The review protocol was prospectively registered with PROSPERO (Registration No. CRD420251086741) to document the planned objectives, eligibility criteria, and analysis plan prior to screening, thereby reducing risks of selective reporting and post hoc changes.

### *3.1. Search Strategy and Information Sources*

We searched studies published between 1 January 1970 and 31 May 2025 to capture the modern evolution of ocular-fatigue research. Foundational pupillometry work from the 1960s (e.g., Hess and Polt [91]; Kahneman and Beatty [115]) is cited for historical context but lies outside the formal search window. The search included six major databases spanning medicine, psychology, and engineering: Pubmed/Medline, Web of Science Core Collection, PsycINFO, IEEE Xplore, ACM Digital Library, and Google Scholar. To ensure comprehensive coverage, we also manually searched preprint servers and the reference lists of key review articles.

Our search strategy was structured around three core concepts: (1) ocular metrics, (2) fatigue and vigilance states, and (3) study design and measurement terms. The specific search terms used for each concept are detailed in Table 1. We adapted the search syntax for each database, utilizing controlled vocabularies such as MeSH terms in PubMed and subject headings in PsycINFO where appropriate.

### *3.2. Study Selection and Eligibility Criteria*

Our initial database search yielded 3,846 records. All records were imported into Rayyan, a web-based tool for systematic reviews, to facilitate duplicate removal and screening. After automatically removing 199 duplicates, 3,647 records remained. Subsequently, Rayyan's automated tools further excluded 2,155 records. Concretely, we used Rayyan's automation features (keyword highlighting and rule-based filtering) to identify records that clearly violated the pre-specified eligibility criteria in Table 2 (e.g., non-human studies, non-ocular measurements, non-fatigue phenomena, or non-eligible publication types like conference abstracts lacking full text). To avoid false exclusions, this automation step was applied conservatively: records were auto-excluded only when the title/abstract contained explicit exclusion cues and lacked inclusion cues related to ocular metrics and fatigue/vigilance. To validate this conservative automation step, we manually audited a random subset of auto-excluded records (n=200) and found no eligible studies.

Two reviewers independently screened the titles and abstracts of these 1,492 records against the eligibility criteria detailed in Table 2. This process resulted in the exclusion of 1,066 records, advancing 426 articles to the full-text review stage. The same two reviewers then independently assessed the full texts of these 426 articles for eligibility. Disagreements at any stage were resolved through discussion; if consensus could not be reached, a third senior reviewer (CD) made the final decision. Inter-rater reliability for the screening process was substantial, with a Cohen's kappa coefficient of $k = 0.82$. During the full-text review, 47 reports were excluded for not meeting the population or specific measurement criteria, resulting in a final synthesis of 379 studies. The complete study selection process is illustrated in the PRISMA flow diagram in Figure 3.

Table 2: Inclusion and Exclusion Criteria

| Domain | Criteria |
|---|---|
| **Population** | **Include:** Human participants of any age or health status.<br>**Exclude:** Animal studies, cell culture studies. |
| **Measurement** | **Include:** Continuous gaze measurements, eye openness, eye movements, and pupillometry using infrared eye-tracking or other sensors.<br>**Exclude:** Static pupil photographs, subjective assessments only. |
| **Phenomenon** | **Include:** Studies measuring fatigue, vigilance, sustained attention, or sleep deprivation.<br>**Exclude:** Studies focused only on emotion, drugs, or clinical diagnosis without a fatigue component. |
| **Outcomes** | **Include:** At least one quantitative ocular metric linked to fatigue or performance.<br>**Exclude:** Studies without statistical analysis of ocular-fatigue relationships. |
| **Publication** | **Include:** Peer-reviewed articles, conference papers, doctoral theses.<br>**Exclude:** Abstracts only, editorials, opinion pieces, patents. |

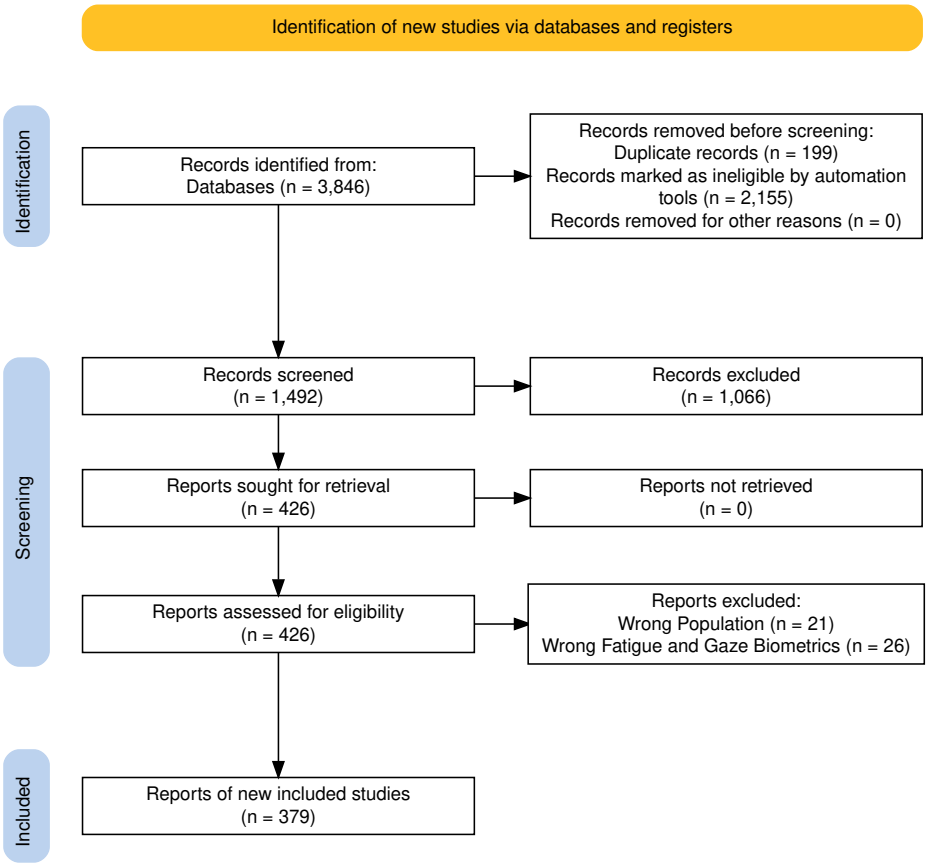



Figure 3: PRISMA 2020 flow diagram illustrating the study identification, screening, and inclusion process.

### *3.3. Data Extraction and Quantitative Synthesis*

Following the final selection of 379 studies, we performed a structured data extraction and quantitative analysis to synthesize the literature and identify key trends and relationships. This went beyond a simple qualitative review to provide a data-driven overview of the field.

First, to capture the dominant themes and terminology across the entire corpus, we extracted keywords from the

titles and abstracts of all included papers. This collection of terms was then used to generate a word cloud, presented in Figure 4, which provides a high-level visual summary of the most prominent topics in ocular fatigue research.




Figure 4: The word cloud generated from the titles and abstracts of the 379 studies included in this review provides a visual summary of the scope covered. The size of each term is proportional to its frequency of occurrence, reflecting the core concepts and dominant terminology within the field of ocular fatigue research.

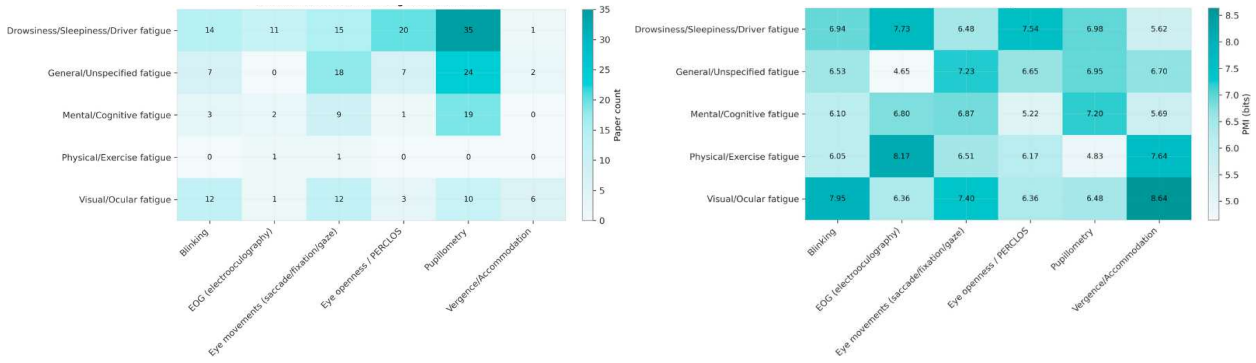
Second, to enable a more granular analysis, two reviewers independently coded each of the 379 included studies for two key pieces of information: (1) the primary type of fatigue being investigated (e.g., Drowsiness, Mental Fatigue, Visual Fatigue), and (2) the primary class of ocular biomarker being measured (e.g., Blinking, Pupillometry, Eye movements).

This structured dataset allowed us to conduct a bibliometric analysis of the relationships between fatigue types and biomarkers. We calculated the raw co-occurrence counts to determine how frequently each fatigue-biomarker pair appeared in the literature. These counts, visualized in the heatmap in Figure 5a, reveal the most heavily researched combinations and highlight the principle of "task-biomarker specificity." To further refine this analysis, we also calculated the Pointwise Mutual Information (PMI) for each pair. PMI is a measure of association that controls for the base frequency of each term, thereby revealing relationships that are stronger than would be expected by chance. The PMI heatmap, shown in Figure 5b, provides a more nuanced view of the informational link between specific fatigue states and their corresponding ocular indicators. These quantitative synthesis methods form the data-driven foundation for the trends and conclusions presented in this review.

## 4. The Multifaceted Nature of Fatigue: A Definitional Framework

Fatigue is a pervasive and complex phenomenon that significantly impacts human health, performance, and well-being. It is a common presenting symptom in primary care and a critical factor in occupational safety, particularly in sectors requiring sustained vigilance [80]. Despite its ubiquity, fatigue is often poorly understood and imprecisely defined, a challenge that complicates its diagnosis, measurement, and management. To build a robust foundation for its objective assessment, it is essential to move beyond simple operational definitions and embrace a multidimensional framework.

Contemporary research makes a foundational distinction between two interdependent components of fatigue: *performance fatigability*, which is an objective and measurable decrease in performance, and *perceived fatigability* (or




(a) Raw co-occurrence counts between fatigue types and ocular biomarkers. Darker cells indicate more papers on that pairing.

(b) Pointwise Mutual Information (PMI) association strength. Higher values indicate stronger-than-expected relationships.

Figure 5: Associations between fatigue types and ocular biomarkers across included studies: **(a)** raw co-occurrence counts; **(b)** PMI scores.

the sensation of fatigue), which is the subjective awareness of a decreased capacity for activity [125, 22, 20]. This distinction is crucial because the perception of fatigue is not merely a symptom of failure but serves an essential **adaptive purpose**. From an evolutionary perspective, the sensation of fatigue functions as a protective, regulatory mechanism. It acts as an early warning signal, alerting the individual to the rising metabolic or cognitive costs of continuing an activity long before catastrophic resource depletion or tissue damage occurs. This subjective experience motivates a shift in behavior—prompting rest, recovery, or a change in strategy—thereby ensuring long-term homeostatic balance and optimizing goal pursuit over time.

A consensus academic definition characterizes fatigue as "the awareness of a decreased capacity for physical and/or mental activity due to an imbalance in the availability, utilization, and/or restoration of resources needed to perform activity" [203, 90]. A key feature that distinguishes pathological fatigue from ordinary, adaptive tiredness is that it is not necessarily or completely relieved by rest and can become disproportionate to the activity performed [185].

This section deconstructs this complex state by first outlining the primary domains in which fatigue manifests — physical, mental, and visual. We then introduce a multi-axial classification framework, presenting four analytical lenses (temporal, etiological, mechanistic, and experiential) that are used to characterize and operationalize these fatigue states. Finally, we address the critical clinical and computational challenge of differentiating fatigue from related states like sleepiness, weakness, and depression.

### *4.1. Primary Domains of Fatigue Manifestation*

While fatigue is a systemic experience, it has been subdivided into different phenomenological dimensions. These map roughly onto its different mental and behavioural effects. This could mean that the various dimensions of fatigue might therefore be measurable in distinct ways.

#### *4.1.1. Physical Fatigue*

Physical fatigue is defined as a transient inability to maintain optimal physical performance, resulting from exertion and leading to a deterioration in the capacity of muscle to produce force or power [70, 69]. Its mechanisms are multifaceted, involving both *peripheral* and *central* factors. `Peripheral fatigue` arises from processes at or distal to the neuromuscular junction, including the depletion of energy substrates (e.g., ATP), the accumulation of metabolic byproducts (e.g., lactate), and dysfunction in calcium handling within muscle fibres [215, 219, 30]. In contrast, central fatigue originates within the central nervous system and is characterized by a reduction in the neural drive from the brain to the muscles, involving a failure of the CNS to send sufficiently strong signals to fully activate muscles [216]. The subjective experience includes symptoms such as muscle pain, weakness, and whole-body tiredness.

#### *4.1.2. Mental and Cognitive Fatigue*

Mental, or cognitive, fatigue is a psychobiological state resulting from prolonged periods of demanding cognitive activity. It is characterized by a subjective sensation of exhaustion and an objective decline in cognitive performance,

and is considered to be independent of sleepiness [167, 103]. It represents a failure of the executive control system to sustain attention and optimize performance over time [192, 206], impacting higher-order functions like concentration, memory, and cognitive control [167].

The neurobiological underpinnings are complex, but previous reports propose disruptions in large-scale brain networks, particularly altered connectivity between frontal and parietal cortices (including the default mode, central executive, and salience networks) [232, 178]. At the neurotransmitter level, mechanisms include glutamate dysfunction and dopamine imbalance, with evidence that glutamine accumulation in the prefrontal cortex may serve as a metabolic marker of cognitive functions [187, 14]. Functional brain imaging studies suggest that as cognitive fatigue develops, activity decreases in the cognitive control network [168], which includes the lateral prefrontal cortex (lPFC), while activity in the default mode network increases. This change, however, appears to be strategic rather than absolute; for instance, one study [24] observed that lPFC activity diminished during low-stakes economic choices but was maintained during the primary cognitive task, suggesting control is disengaged only when the consequences are tolerable. The role of the anterior cingulate cortex (ACC), which is involved in monitoring effort and evaluating the costs and benefits of actions, is also complex; rather than simply remaining engaged, its activity can decrease with fatigue or shift to reflect compensatory effort [126].

#### *4.1.3. Visual Fatigue (Asthenopia)*

Visual fatigue, clinically known as asthenopia, is a collection of symptoms related to discomfort and decreased performance of the visual system following prolonged, visually intensive tasks [71]. It represents a specialized form of fatigue combining accommodative stress (ciliary muscle fatigue from excessive focusing), ocular surface issues (tear film instability from reduced blink rates), and `extraocular muscle strain` (from sustained convergence demands) [120, 177]. In today's digital environment, a key contributor is the vergence-accommodation conflict (VAC), particularly in virtual reality (VR) and 3D displays, where the eyes must converge at one distance (the screen) while focusing at another (the virtual object). Digital eye strain now affects up to 65% of computer users [177] and is increasingly prevalent in children [191].

### *4.2. Analytical Lenses for Classifying Fatigue*

To further deconstruct and operationalize the concept of fatigue for both clinical and computational purposes, several intersecting classification frameworks are used. These are not mutually exclusive categories but rather analytical lenses that provide a structured approach to characterizing a specific fatigue state.

The `temporal lens` categorizes fatigue by its duration and onset. In operational contexts like aviation, this includes *Transient* (acute), *Cumulative* (building up over days), and *Circadian* (tied to the body's internal clock) fatigue. In clinical settings, it is classified by duration as *Recent* (<1 month), *Prolonged* (1-6 months), or *Chronic* (>6 months).

The `etiological lens` organizes fatigue by its underlying cause [170, 163]. This distinguishes between *Primary fatigue*, which is a direct result of a disease process (e.g., in multiple sclerosis) [29], and *Secondary fatigue*, which arises as a consequence of other factors like chronic pain or medication side effects [188]. *Physiological fatigue* describes a normal response to an imbalance in exercise or sleep that is relieved by rest.

The `mechanistic lens` provides a fundamental physiological distinction based on the location of failure. As noted in the description of physical fatigue, this divides the phenomenon into *Central Fatigue*, originating within the CNS, and *Peripheral Fatigue*, arising at or distal to the neuromuscular junction [50, 95]. Cognitive fatigue, as previously described, is a primary example of central fatigue.

Finally, the `psychological and experiential lens` categorizes fatigue by its subjective experience. This includes *Emotional Fatigue* (a core component of burnout) and *Mental Fatigue* (the subjective difficulty in concentrating). A more advanced model, the *Motivational Control Theory*, conceptualizes fatigue not as resource depletion but as an adaptive emotion that signals the increasing costs of a task, prompting a re-evaluation of goal pursuit [93].

### *4.3. Differentiating Fatigue from Clinically and Computationally Related States*

In a clinical context, the nonspecific complaint of "feeling tired" must be carefully differentiated from states with significant symptomatic overlap: sleepiness, weakness, and depression. This process is essential not only for accurate diagnosis but also for robust computational modeling, as the symptomatic overlap creates significant feature ambiguity that can challenge classification algorithms.

#### 4.3.1. Fatigue vs. Sleepiness

The primary difference between fatigue and sleepiness lies in the sensation. *Sleepiness* is specifically the propensity to fall asleep, a direct manifestation of the homeostatic sleep drive [198, 36]. In contrast, *fatigue* is a broader sensation of lacking physical or mental energy. The most reliable differentiator is the response to intervention: sleepiness is effectively relieved by sleep, whereas pathological fatigue is often difficult to relieve, being more intense and prolonged in duration [185, 107].

#### 4.3.2. Fatigue vs. Weakness

This distinction separates a subjective feeling from an objective loss of muscle power. *Weakness* is a quantifiable decrease in muscle strength that can be identified through a physical examination [37, 133]. A patient with *fatigue* but normal strength can generate full power, though they may report that the effort is exhausting or cannot be sustained.

#### 4.3.3. Fatigue vs. Depression

This overlap is substantial and bidirectional [106, 40]. The core distinction lies in motivation. The defining feature of *depression* is often *anhedonia*—a loss of interest or pleasure. The lack of engagement is driven by not "caring" to do things [45, 34]. In contrast, in primary *fatigue* syndromes (like ME/CFS), the core feature is *anergia*—a lack of physical or mental capacity. The patient may still desire to participate but feels unable, and attempting to do so can lead to post-exertional malaise [6, 231].

This symptomatic overlap presents a critical challenge for automated detection systems, which must learn to distinguish these states based on subtle differences in behavioral and physiological feature sets. The comparative characteristics are summarized in Table 3.

Table 3: Comparison of Fatigue to Sleepiness, Weakness, and Depression.

| **Characteristic** | **Fatigue** | **Sleepiness** | **Weakness** | **Depression** |
|---|---|---|---|---|
| **Definition** | A lack of physical and/or mental energy and motivation. | A propensity or urge to fall asleep; drowsiness. | A quantifiable decrease in muscle strength. | A mood disorder characterized by persistent sadness and/or anhedonia. |
| **Primary Sensation** | Feeling exhausted, worn out, drained, or lacking stamina. | Feeling the need to sleep; struggling to stay awake. | Inability to exert force; feeling a lack of power in the muscles. | Feeling sad, empty, hopeless, or worthless; loss of interest or pleasure. |
| **Effect of Rest/Sleep** | Often unrefreshing; may not relieve the underlying feeling of exhaustion [185]. | Relieved by sleep; a nap improves the state [198]. | Rest may alleviate exercise-induced weakness but does not restore strength lost to pathology. | May be associated with insomnia or hypersomnia; sleep is often nonrestorative. |
| **Effect of Activity** | Intensifies the feeling of exhaustion [188]. | Temporarily alleviates the state through increased arousal. | Exacerbates the inability to generate force in the affected muscles. | May improve mood and energy in some cases, but often there is no motivation to be active. |
| **Differentiator** | Lack of energy despite a desire to be active (*anergia*) [6]. | An overwhelming urge to sleep that is relieved by napping. | Objective, measurable loss of muscle power on physical examination [37]. | Pervasive low mood and loss of interest/pleasure (*anhedonia*) [45]. |
| **Locus of Dysfunction** | Central (CNS) or peripheral (muscular) energy/-control systems. | Brain's arousal and sleep-wake regulatory systems [36]. | Neuromuscular pathway (motor neurons, NMJ, muscle fibers) [133]. | Brain circuits involved in mood, motivation, and reward regulation. |

**Lessons Learned**

Fatigue is not a single entity. It must be deconstructed into its components (*perceived* vs. *performance*) and domains (*physical*, *mental*, *visual*). A critical first step in any detection effort is to clearly define the target fatigue phenotype and differentiate it from related states like sleepiness, weakness, or depression, as the optimal biomarkers will differ for each.

## 5. Ocular Biomarkers for Fatigue Assessment

The eyes serve as a particularly accessible window into the functional state of the central nervous system (CNS) [129, 153]. They provide a non-invasive, real-time readout of neural processes that are otherwise hidden from direct observation. Fatigue, in its various forms, induces a cascade of neurobiological changes that manifest as distinct, measurable alterations in pupillary dynamics, oculomotor control, and even the physical appearance of the periorbital region [168, 15, 35]. In some cases, the causes of these changes are well understood; they are direct consequences of the dysregulation of specific neural circuits, including the autonomic nervous system (ANS), brainstem arousal centers like the locus coeruleus, and higher-order control networks involving the basal ganglia and prefrontal cortex [153, 113, 126, 92]. Understanding these neurobiological underpinnings is crucial for developing the objective biomarkers that form the foundation of modern fatigue detection technologies.

### *5.1. Pupillometry: A Direct Line to Central Arousal*

Pupillometry, the measurement of pupil size and reactivity, offers a powerful proxy for the activity of central arousal systems. The pupil's diameter exists in a constant state of dynamic equilibrium, governed by the push-pull antagonism of the two branches of the ANS. The parasympathetic system, via acetylcholine, contracts the iris sphincter muscle to cause constriction (*miosis*), while the sympathetic system, via norepinephrine, contracts the iris dilator muscle to cause dilation (*mydriasis*) [153]. This mechanism is depicted in Figure 6.

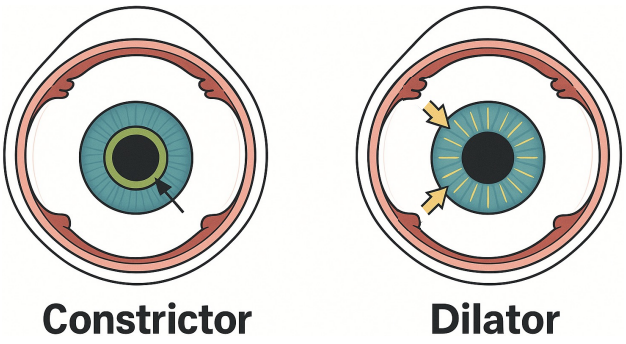



Figure 6: The antagonistic control of pupil size. The pupillary constrictor muscle (green), controlled by the parasympathetic nervous system, contracts to shrink the pupil. The pupillary dilator muscle (yellow), controlled by the sympathetic nervous system, contracts to expand the pupil.

While the ANS provides the final efferent pathway, the command signals for cognitive and arousal-related pupil changes originate deeper within the brain. The principal hub for this control is the *locus coeruleus-norepinephrine (LC-NE) system* [132]. The LC, a small brainstem nucleus, is the primary source of norepinephrine for the brain, regulating arousal, attention, and cognitive effort. A large body of research has established that pupillometry serves as a reliable, real-time index of LC-NE activity [115, 113, 158].

This link provides a powerful neurological framework for understanding fatigue's impact. A critical distinction must be made, however: while low arousal is a defining characteristic of sleepiness, it is not synonymous with all forms of fatigue. The system's *tonic activity*, reflected in baseline pupil diameter, corresponds to overall arousal and is therefore most sensitive to sleep-related states. As sleepiness sets in, the sympathetic tone from the LC wanes, leading to a reduced baseline pupil diameter and increased low-frequency oscillations known as *pupillary unrest* or

*hippus* [149, 35]. In contrast, cognitive fatigue is more closely related to the regulation of mental effort. This is better captured by the system's *phasic activity*—the *Task-Evoked Pupillary Response (TEPR)*—which is the transient dilation of the pupil following a cognitive challenge. This dilation is tightly coupled with the burst-firing of LC neurons and represents the recruitment of neural resources for the task at hand [53, 19].

According to the *adaptive gain theory*, the LC-NE system optimizes performance by modulating "neural gain", enhancing the processing of important information while suppressing noise [81, 68]. From this perspective, cognitive fatigue is a state of suboptimal LC-NE regulation. As a task becomes prolonged, the LC system may shift into a disengaged mode, leading to a progressive *attenuation of the TEPR*. This robust "fatigue effect," where the pupil's response to the same cognitive demand diminishes over time, is a direct ocular marker of mental fatigue and task disengagement [154, 97].

### *5.2. Oculomotor Control: A Behavioural Readout of Cognitive State*

Fatigue impairs the brain's ability to control the intricate network of extraocular muscles that move the eyes (Figure 7). Crucially, these oculomotor deficits do not arise from peripheral muscle fatigue—as the extraocular muscles are known to be highly fatigue-resistant [176]—but are instead a consequence of dysregulation in the high-level neural command centres that plan and execute eye movements [126, 168]. They therefore provide objective, behavioural evidence of a decline in central processing efficiency.

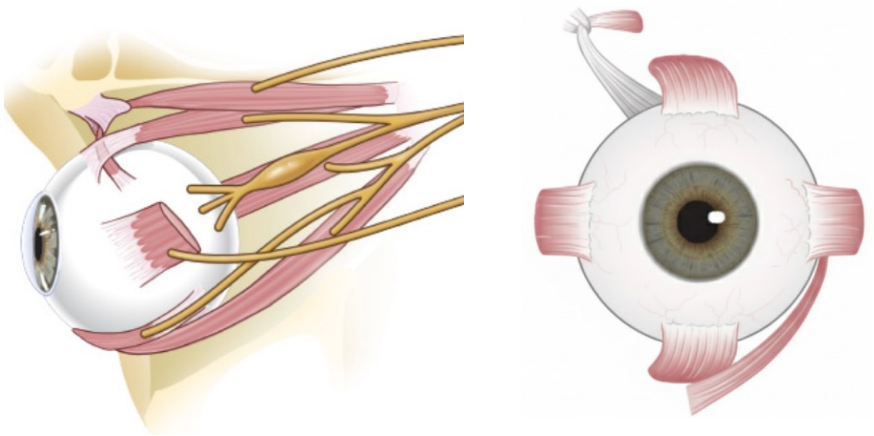


Figure 7: The neuro-muscular basis of oculomotor control. This illustration shows the extraocular muscles responsible for directing the eye's gaze and the cranial nerves that transmit motor commands from the brain [73]. The efficiency of this entire system is highly sensitive to central fatigue, which impairs neural signalling and results in the degradation of eye movements.

This degradation is governed by two key neural systems. The *prefrontal cortex (PFC)*, as the brain's executive hub, is responsible for top-down cognitive control. Since cognitive fatigue is fundamentally an "executive failure" [95], the PFC's ability to direct goal-oriented, non-reflexive eye movements becomes compromised. Concurrently, the *basal ganglia*, critical for the purposive control and selection of movements, also show dysfunction. These subcortical structures suppress unwanted eye movements via inhibitory projections to the superior colliculus [92]. Central fatigue states that disrupt basal ganglia function, for instance through dopamine depletion, manifest directly as impaired eye movement control [201].

This high-level dysregulation produces consistent changes in *saccades*, the rapid, ballistic eye movements used to shift gaze. One of the most reliable findings in fatigue research is a slowing of *saccadic peak velocity* with increased time-on-task and sleep deprivation [35, 194]. This reduction in "oculomotor vigor," which is not under voluntary control, is seen in the fatigue of professionals in high-stakes environments like surgical residents after long shifts [57, 157]. The slowing of saccades is attributed to reduced activity in brainstem burst neurons, decreased noradrenergic drive from the LC, and impaired cortico-collicular signaling. Furthermore, the time taken to initiate a saccade, or *saccade latency*, increases, reflecting slower neural processing and decision-making [246, 52]. Impaired PFC function also leads to a rise in *antisaccade errors*—failures to inhibit a reflexive glance toward a stimulus—which serve as a direct behavioral readout of diminished inhibitory control, a core component of cognitive fatigue [95, 25].

Other oculomotor metrics are similarly affected. *Fixations*, the periods when the eyes are held still to acquire information, tend to lengthen with increasing cognitive and visual fatigue, likely reflecting slower information processing [193, 15]. The stability of fixation also degrades, with an increase in *ocular drift*, signaling a breakdown in fine motor control [56]. Spontaneous *blinking*, a semi-autonomic behavior modulated by dopaminergic pathways in the basal ganglia [235], changes markedly as fatigue progresses into sleepiness. Blink duration increases, and the *Percentage of Eye Closure (PERCLOS)*—the proportion of time the eyelid is ≥80% closed—emerges as one of the most validated biomarkers for drowsiness, strongly correlating with performance lapses and microsleeps [1, 221]. Finally, the ability to track a moving object smoothly (*smooth pursuit*) declines, manifesting as reduced gain and an increase in corrective "catch-up" saccades, reflecting delayed processing in cortical motion areas [75, 246].

### *5.3. Mapping Ocular Changes to Specific Fatigue States*

Different ocular parameters exhibit varying sensitivity to different types of fatigue, allowing for a more nuanced assessment than a single, monolithic score. Mental exhaustion and declining cognitive control are well-captured by metrics such as decreased saccadic peak velocity, increased saccade latency, and a rise in anti-saccade errors [15, 13]. Drowsiness, specifically, is most reliably indexed by eyelid-related metrics like PERCLOS and increased blink duration, which are highly correlated with microsleeps and hold great promise for applications like in-vehicle monitoring systems [1, 193].

In contrast, pupillometry is especially powerful for capturing systemic cognitive load. The attenuation of the TEPR is a hallmark of cognitive fatigue and task disengagement [97]. Moreover, pupillometry can index the systemic stress from intense physical exertion. A study on professional basketball players, for instance, revealed strong correlations between pupillary metrics and markers of physical fatigue like perceived exertion and heart rate variability, suggesting that the pupil provides an integrated measure of an athlete's overall psycho-physiological load [102]. Likewise, visual fatigue from prolonged screen use is indicated by increased fixation duration and decreased saccade velocity [15].

### *5.4. Periorbital Cues as Physical Biomarkers*

Beyond these dynamic functional changes, fatigue—particularly from sleep deprivation—causes visible physical changes in the periorbital region that serve as powerful social cues of an individual's internal state [212, 10]. These include drooping eyelids (*ptosis*) due to reduced neural input to the levator palpebrae muscles, redder eyes from the dilation of conjunctival blood vessels, swollen eyes, and darker under-eye circles. While these cues are readily interpreted by human observers, their symptomatic overlap with expressions of sadness underscores the need for objective, multi-modal assessment [212].

### *5.5. Summary of Ocular Biomarkers*

The various ocular parameters discussed provide a rich, multi-dimensional set of biomarkers for assessing fatigue. As this review has demonstrated, a crucial principle is that of task-biomarker specificity: different ocular metrics show varying sensitivity to different fatigue phenotypes. For instance, *drowsiness* is best captured by eyelid metrics like PERCLOS, *mental fatigue* by a decrease in saccadic peak velocity, and *visual strain* by an increase in fixation duration. While some metrics are more sensitive to sleepiness and arousal (e.g., PERCLOS, blink duration), others are more specific to mental fatigue and executive function (e.g., saccadic velocity, antisaccade errors). Table 4 provides a comprehensive summary of these key metrics, their typical changes with fatigue, the primary states they indicate, and key supporting references. This synthesis is critical for bridging foundational fatigue research with the development of robust, multi-modal fatigue detection systems in computer science.

| Ocular Metric | Description | Fatigue Effect | Associated State | N | Ref |
|---|---|---|---|---|---|

*Continued on next page*

| Ocular Metric | Description | Fatigue Effect | Associated State(s) | N | Ref(s) |
|---|---|---|---|---|---|
| **Eyelid Dynamics** | | | | | |
| PERCLOS | % time eyelid is ≥80% closed over an interval. | *Increases* markedly with drowsiness / microsleeps. | Drowsiness, Sleep Deprivation | 129 | [221, 162, 1] |
| Blink Duration | Mean time from eyelid closing to reopening. | *Increases* (slow/long blinks) as fatigue grows. | Drowsiness | 115 | [193, 189] |
| Blink Rate | Spontaneous blinks per minute. | *Increases* with fatigue/sleepiness; can transiently *decrease* during intense visual demand. | Drowsiness, Inattention | 110 | [193, 236] |
| **Pupillary Dynamics** | | | | | |
| Task-Evoked Pupil Dilation (TEPR) | Phasic pupil increase to cognitive events (effort/engagement). | *Attenuates* with mental fatigue; smaller task-evoked responses. | Mental Fatigue, Disengagement | 44 | [97, 19] |
| Tonic (Baseline) Pupil Diameter | Average pupil size at rest (tonic arousal). | Often *decreases* with time awake/sleepiness, but findings are mixed; interpret with caution. | Arousal Level | 51 | [46, 154] |
| Pupillary Light Reflex (PLR) | Constriction amplitude/latency to light, etc. | *Reduced amplitude* and/or *prolonged latency* after sleep loss; PLR features can index fatigue. | Vigilance / Autonomic Reactivity | 19 | [82, 117] |
| Pupillary Unrest Index (PUI) | Low-frequency spontaneous pupil oscillations in darkness. | *Increases* with daytime sleepiness and following sleep loss. | Sleepiness | 10 | [228, 227, 229] |
| **Saccadic & Pursuit Dynamics** | | | | | |
| Saccadic Peak Velocity | Maximum eye speed during a saccade (oculomotor vigor). | *Decreases* with sleep loss and time-on-task. | Mental Fatigue, Vigilance/Arousal | 68 | [55, 75, 246] |
| Saccade Latency | Reaction time to initiate a saccade after a target appears. | *Increases* with fatigue/sleep deprivation (slower initiation; task-dependent). | Mental Fatigue, Alertness | 55 | [52, 246, 26] |
| Antisaccade Errors | Failures to inhibit a reflexive glance toward a sudden stimulus. | *Increase* with fatigue /sleep loss; inhibitory control degraded. | Executive Dysfunction | 8 | [25, 38] |
| Smooth Pursuit Gain / Accuracy | Ability to match eye velocity to a moving target (gain, error). | *Declines* with fatigue /sleep loss (lower gain; more error; more catch-up saccades). | Low Alertness, Mental Fatigue | 21 | [16, 75, 246] |
| **Fixational Dynamics** | | | | | |
| Fixation Duration | Time the gaze remains on a location before moving. | Tends to *lengthen* with reduced vigilance and mental fatigue (task-dependent distributions). | Attention / Vigilance | 83 | [236, 76] |
| Fixation Stability (Drift/Dispersion) | Stability of fixation (drift velocity /dispersion while fixating). | *Worsens*: drift velocity /instability *increase* with time-on-task. | Mental Fatigue, Inattention | 15 | [56] |
| Microsaccades | Tiny, involuntary saccades during fixation. | *Slower dynamics* with fatigue; changes in rate/velocity with | Alertness / Focus | 9 | [56] |

**Lessons Learned**

The eye provides a rich, multi-dimensional readout of CNS function. A key principle is **task-biomarker specificity**: different ocular metrics are sensitive to different fatigue states. Eyelid dynamics (e.g., PERCLOS) are robust for *drowsiness*, while oculomotor control (e.g., saccadic velocity, vergence) and pupillometry (e.g., TEPR attenuation) are more powerful for tracking *cognitive fatigue* and mental effort.

## 6. Experimental Paradigms for Eliciting and Measuring Fatigue

The objective study of fatigue relies on a structured methodological framework designed to reliably induce a fatigued state and quantify its effects through a combination of behavioural, physiological, and subjective measures. While not an exhaustive survey, this section reviews several of the most common and influential paradigms, instruments, and protocols that form the foundation of experimental fatigue research. We first survey key cognitive tasks used for fatigue induction, then discuss the suite of instruments for its measurement, and conclude by outlining a standard protocol for a robust fatigue experiment. This framework provides the necessary context for understanding how the ocular biomarkers discussed in Section 5 are elicited and validated in a controlled setting.

### *6.1. Fatigue Induction Paradigms*

The effectiveness of a task at inducing cognitive fatigue is largely determined by its ability to place sustained, uninterrupted demands on specific cognitive systems, thereby depleting finite neural resources [60]. Paradigms that feature continuous performance are thus more potent than those with intermittent, trial-based structures that allow for micro-recovery [180]. A comparative summary of major induction tasks is provided in Appendix Appendix A, Table A.6.

#### *6.1.1. Sustained Attention and Vigilance Tasks*

The **Psychomotor Vigilance Task (PVT)** is the gold standard for assessing behavioural alertness [60, 204, 63]. Participants perform a simple reaction-time task where stimuli appear at unpredictable intervals over a 5- to 10-minute block [63, 147, 98]. This random timing forces the maintenance of tonic alertness, and the primary metrics—attentional lapses (reaction times > 500 ms) and reaction time variability—are exceptionally sensitive to sleep loss and circadian disruption [83, 17]. While its cognitive scope is narrow, its high reliability and minimal learning effects make it an ideal probe for changes in vigilance [18].

While the PVT remains the benchmark, other paradigms are also used to assess sustained attention and vigilance. These include the **Sustained Attention to Response Task (SART)**, which requires withholding a response to rare targets amidst a stream of frequent non-targets [89], the **Elevator Counting Task**, where participants must count visual or auditory events over time [222], and more ecologically valid simulations, such as monitoring instruments during a general aviation flight task [226].

#### *6.1.2. Working Memory and Executive Load Tasks*

The **n-Back task** parametrically manipulates cognitive load by requiring participants to monitor a continuous sequence of stimuli and respond if the current stimulus matches the one presented 'n' trials previously [124]. Administered in a continuous block of 20-60 minutes, it places relentless demands on working memory updating and interference control, leading to a reliable build-up of mental fatigue [31]. In contrast, the **Sternberg item recognition task**, which involves encoding and then recalling a small set of items on a trial-by-trial basis, is a poor fatigue inducer due to the intermittent recovery periods between trials [210]. It is better used as a pre-/post-induction probe to measure fatigue's specific impact on short-term memory scanning efficiency [210].

*6.1.3. Conflict and Inhibitory Control Tasks*

The Stroop task is a canonical method for inducing cognitive conflict. It requires participants to name the ink color of a word while inhibiting the more automatic and prepotent process of reading the word itself [108]. The conflict is immediate and inherent to the stimulus properties, demanding top-down inhibition of the dominant reading pathway to allow for the selection of the correct, colour-based response. When administered in a sustained block (e.g., ≥30 minutes), the constant effort required to resolve this interference and maintain top-down executive control serves as a powerful fatigue inducer. This makes the Stroop task a classic measure of interference resolution at the point of response selection.

In contrast to the simultaneous, stimulus-inherent conflict of the Stroop task, the family of Continuous Performance Tasks (CPTs) probes inhibitory control through conflict that arises from context and expectancy over time [186, 61]. In the foundational "go/no-go" design of a CPT, participants monitor a continuous stream of stimuli and must execute a rapid motor response to frequent "go" stimuli while withholding the response to an infrequent "no-go" stimulus. This design allows for the clear quantification of distinct cognitive processes: omission errors (failing to respond to a "go" stimulus) are interpreted as lapses in sustained attention or vigilance, while commission errors (incorrectly responding to a "no-go" stimulus) provide a direct measure of failed inhibitory control.

Advanced CPT variants, such as the AX-CPT, are specifically designed to assess the active maintenance of task goals [205, 197]. In this paradigm, the correct response to a "probe" stimulus (e.g., 'X') is dependent on a "cue" stimulus (e.g., 'A') that was presented moments earlier. Performance is often interpreted through the Dual Mechanisms of Control (DMC) framework, which distinguishes between a resource-intensive proactive control mode (anticipating and preparing for events) and a less demanding reactive mode (engaging control only after conflict is detected). Because proactive control is metabolically expensive, tasks like the AX-CPT that require its sustained engagement are powerful tools for investigating the neurocognitive mechanisms of mental fatigue [39].

*6.1.4. High-Fidelity Ecological Simulations*

Complex simulations, often using virtual reality (VR), such as an **Applied Distributed Cognition Task (ADCT)**, offer high ecological validity by mimicking the multi-faceted demands of real-world operational environments (e.g., a team-based flight simulation) [44, 32, 175]. These paradigms induce a realistic fatigue state by simultaneously taxing executive functions, spatial navigation, and team communication [175]. However, they are resource-intensive, and it is methodologically critical to control for confounds like cybersickness using tools such as the Simulator Sickness Questionnaire (SSQ) [121, 11, 23]

*6.2. Measurement Instruments*

A comprehensive assessment of fatigue requires a multi-modal approach that captures objective physiological changes, observable behavioural decrements, and the subjective experience of the individual.

*6.2.1. Objective Physiological Markers*

As detailed in Section 5, ocular metrics provide a powerful, non-invasive window into the neurophysiological state of a fatigued individual. These measures, including pupillometry (e.g., Task-Evoked Pupillary Response, Pupil Unrest Index) and eyelid kinematics (e.g., PERCLOS, blink duration), are directly linked to the brain's arousal and cognitive control networks [154, 195, 1]. A detailed summary of these objective ocular metrics, their directional changes with fatigue, and their neurophysiological rationale is presented in Appendix Appendix A, Table A.7.

*6.2.2. Subjective Instruments*

Self-report scales are indispensable for capturing the perceived, internal state of fatigue.

- **State Scales:** Used for in-task tracking, these brief scales capture momentary feelings. The **Karolinska Sleepiness Scale (KSS)** is a 9-point scale used for frequent sampling of subjective sleepiness during a task [8, 155]. The **Stanford Sleepiness Scale (SSS)** serves a similar purpose with a 7-point format [94].
- **Trait Inventories:** These longer questionnaires characterize a participant's general disposition. The **Multidimensional Fatigue Inventory (MFI-20)** assesses five dimensions of fatigue (e.g., General, Mental, Physical) and is best used as a baseline or pre-/post-intervention measure [97].

A summary of key subjective instruments, their constructs, and best-use cases is provided in Appendix Appendix A, Table A.8.

### *6.3. A Canonical Experimental Protocol*

A robust cognitive fatigue experiment typically follows a three-phase, within-subjects design to isolate the effects of the fatiguing manipulation. This structure allows each participant to serve as their own control, increasing statistical power and accounting for individual differences. The canonical protocol is detailed in Appendix Appendix A, Table A.9.

1. **Baseline Phase:** The experiment begins by establishing a stable, pre-fatigue reference point. Measures include trait questionnaires (e.g., MFI-20), state scales (e.g., KSS), a brief behavioural probe (e.g., 3-minute PVT), and resting-state physiological recordings (e.g., tonic pupil diameter) under controlled conditions (e.g., constant luminance, fixed position) [119].
2. **Induction Phase:** This is the core manipulation, where a continuous, demanding cognitive task (e.g., a 45-minute n-back task) is performed to systematically deplete cognitive resources [3]. Continuous behavioural and physiological data are collected throughout, with periodic subjective probes (e.g., KSS every 10 minutes) to track the build-up of fatigue.
3. **Post-task Phase:** Immediately following the induction, the full battery of baseline measures is repeated under identical conditions. This allows for the quantification of net change (delta) in all dependent variables. Comparing pre- and post-task performance on a transfer task (e.g., the PVT) is critical for assessing whether fatigue induced by one task impairs performance on another [119].

By integrating demanding induction paradigms with a multi-modal measurement suite within a structured pre-post protocol, researchers can rigorously investigate the causes and consequences of cognitive fatigue.

**Lessons Learned**

The validity of any fatigue detection model depends on the quality of the data used to train it. A robust experimental protocol involves a within-subjects, pre-post design with a demanding, induction task (e.g., n-Back or game) with continuous data collection. Measurement must be multi-modal, combining objective physiological data (like ocular metrics), subjective reports (like KSS), and behavioural performance measures (like PVT lapses) to create a comprehensive ground truth.

## 7. Technological Frameworks for Fatigue Detection

The translation of ocular biomarkers into practical, real-world fatigue detection systems has been accelerated by advances in eye-tracking hardware and the application of sophisticated machine learning algorithms. These technological frameworks are moving beyond simple threshold-based alarms toward nuanced, predictive models capable of classifying different levels and types of fatigue. This section critically analyses these frameworks, organised not by the algorithms themselves, but by the nature of the ocular signals they model. This approach directly connects the computational methods to the underlying neurophysiology discussed previously, evaluating why certain techniques are better suited for specific fatigue-related biomarkers and highlighting the unique challenges each presents.

### *7.1. Modelling Eyelid Dynamics for Drowsiness Detection*

Eyelid dynamics are the most established and widely implemented biomarkers for detecting drowsiness, primarily due to their strong, direct correlation with the physiological state of descending into sleep. Metrics such as *Percentage of Eye Closure (PERCLOS)*, increased *blink duration*, and altered *blink rate* serve as robust indicators of microsleeps and severe lapses in alertness [221, 1].

The relative simplicity of these signals has made them amenable to traditional computer vision and machine learning techniques. Because the core task is often event detection (identifying a blink) or geometric feature extraction (measuring eyelid distance), early and effective systems were built using pipelines that first locate the face and eyes

(e.g., using Viola-Jones or Haar cascades) and then apply a classifier to the extracted features [88, 48, 161, 160, 99]. For example, a linear Support Vector Machine (SVM) can effectively classify eye state to estimate PERCLOS from edge-detected eye images, demonstrating the power of feature-based approaches for this well-defined problem [47].

More recently, deep learning models, particularly Convolutional Neural Networks (CNNs), have reported exceptionally high accuracies in drowsiness detection, often exceeding 95% [159, 200]. A Multitask Cascaded CNN (MTCNN), for instance, can be used to detect not only eye closure but also other behavioural cues like yawning, achieving over 99% accuracy in controlled conditions [105]. The primary advantage of these end-to-end models is not necessarily in modelling the blink signal itself, but in their robustness to confounding real-world variables. By learning features directly from raw pixels, they can better handle variations in lighting, head pose, and occlusions (e.g., glasses) that often cause traditional feature-based pipelines to fail. However, the computational expense of these models raises questions about their necessity for a signal that is already well-characterized, particularly in resource-constrained environments like in-vehicle driver monitoring systems. For such applications, the robustness of eyelid-based metrics is often enhanced by fusing them with other physiological signals, such as heart rate or electroencephalogram (EEG) data, to provide a more holistic and reliable assessment of the user's state [64, 239].

### *7.2. Modelling Saccadic and Pursuit Dynamics for Cognitive Fatigue*

While eyelid dynamics are excellent markers for drowsiness, they are less sensitive to the subtler state of *cognitive fatigue*. As established in Section 5, the dynamics of voluntary eye movements—such as reduced *saccadic peak velocity*, increased *saccade latency*, a rise in *antisaccade errors*, and decreased *smooth pursuit gain*—potentially serve as a behavioural readout of declining efficiency in the brain's executive control circuits, namely the prefrontal cortex and basal ganglia [194, 52, 25, 75, 15].

Modelling these signals presents significant challenges. First, extracting these features requires high-fidelity eye-tracking with sufficient temporal and spatial resolution, as well as robust signal processing to identify saccadic events and filter out noise. Statistical and functional analysis techniques, such as Fourier transforms, are often employed to analyse the temporal and frequency domain characteristics of saccade data [233]. Second, traditional machine learning approaches often treat these metrics as independent data points averaged over a time window (e.g., mean saccadic velocity). This approach, while informative, discards the rich temporal information contained within the *sequence* of eye movements. For instance, a gradual decline in saccadic velocity followed by a compensatory increase in fixation duration may be a more powerful predictor of fatigue than either metric alone.

This temporal dependency suggests that models designed for sequential data, such as Recurrent Neural Networks (RNNs), LSTMs, or Transformers, may be particularly well-suited for this task. By processing eye movements as a sequence, these models can learn the complex interplay between different oculomotor events over time. Emerging research using deep neural embeddings to model gaze sequences represents a step in this direction, showing that such models can learn state-related changes in oculomotor patterns [152]. The future of detecting cognitive fatigue likely lies in these sequence-aware architectures, which can move beyond simple feature averaging to capture the dynamic signature of a brain struggling to maintain executive control.

### *7.3. Modelling Pupillometry for Arousal and Cognitive Load*

Pupillometry offers a direct, non-invasive window into the brain's central arousal and cognitive effort systems, governed primarily by the locus coeruleus-norepinephrine (LC-NE) system (see Section 5) [132, 113]. Key metrics like the *Task-Evoked Pupillary Response (TEPR)*, tonic *baseline pupil diameter*, and the *Pupillary Unrest Index (PUI)* are powerful indicators of cognitive load, task engagement, and arousal [19, 46, 228].

The primary challenge in using pupillometry is that the cognitive signal is a subtle modulation superimposed on a much larger and more dominant signal: the pupillary light reflex (PLR). The pupil's diameter is highly sensitive to ambient and focal luminance, and this light-driven variation can easily overwhelm the smaller changes related to cognition. Much research has focused on statistical and signal processing techniques to disentangle these signals, such as by dynamically modelling and subtracting the influence of luminance or using frequency-domain analysis like Hilbert transforms [171, 100] and applied ambient lighting sensors from devices to correct pupillometry calculation [27].

This is where end-to-end deep learning models hold the most promise. A deep neural network, particularly a temporal architecture, could theoretically learn to decouple the cognitive signal from the light reflex by taking both the raw pupil size (often smoothed with a Kalman filter [173]) and a measure of environmental luminance as inputs.

Table 5: High-level comparison of common model families used in ocular-fatigue detection pipelines. The table summarizes typical strengths, robustness constraints, and best-fit fatigue phenotypes; reported performance varies widely due to dataset and labelling heterogeneity.

| **Model family** | **Typical inputs** | **Strengths** | **Common failure modes / robustness limits** | **Best suited for** |
|---|---|---|---|---|
| Linear / kernel SVM, LR | Hand-crafted eyelid or pupil features; landmark geometry | Data-efficient; interpretable; low-latency deployment | Sensitive to feature extraction errors (lighting, occlusion, head pose); limited temporal modeling | Eyelid-based drowsiness (e.g., PERCLOS), simple state classification |
| Tree models (RF, XGBoost) | Aggregated ocular features + context (time-of-day, task) | Handles non-linearities; robust to mixed feature scales; fast inference | Still depends on stable feature extraction; can overfit small private datasets | Mixed-feature fatigue scoring; rapid prototyping |
| CNNs (frame-based) | Eye-region crops / face video frames | Strong representation learning; improved robustness to appearance variation | Can exploit dataset-specific shortcuts; domain shift under real lighting/camera changes | Drowsiness from eye/face appearance; blink/closure detection |
| RNN /LSTM /GRU | Sequences of features or embeddings | Captures temporal dynamics (fatigue build-up); models event order | Label noise and sequence length sensitivity; harder to interpret; drift over long sessions | Cognitive fatigue signatures in saccade/fixation sequences; temporal pupil dynamics |
| Transformers / SSMs | Long sequences of embeddings; multimodal tokens (video+IMU+PPG) | Long-range dependencies; multimodal fusion; scalable pretraining | Data-hungry; compute/latency constraints; needs careful calibration | Generalization-focused modeling; multimodal, in-the-wild monitoring |

By learning the complex, non-linear function of the PLR, the model could effectively isolate the residual signal related to cognitive state. Unified models like CLERA, which jointly learn to estimate cognitive load and analyse eye-region features from raw video, exemplify this integrated approach [59]. Through multi-frame data fusion and temporal integration, pupillometry calculations have been achieved on mobile devices [213]. The development of such deep, context-aware models that can robustly measure pupillary markers of cognitive effort in uncontrolled lighting environments remains a key frontier for the field.

### *7.4. Challenges in Validation and Benchmarking*

A critical analysis of the literature reveals that while many fatigue detection models report impressive performance, validating and comparing these systems is fraught with systemic challenges. Accuracy is the most commonly reported metric, with many recent deep learning studies claiming values exceeding 90% and even 99% in some cases [200, 159, 105]. However, such claims must be interpreted with caution. Fatigue detection datasets are often imbalanced, with non-fatigued states being far more prevalent than instances of severe fatigue, a context in which metrics like the Area Under the Receiver Operating Characteristic curve (AUROC) provide a more robust evaluation.

A more fundamental issue is the ambiguity of "ground truth." The validity of any supervised learning model depends on the quality of its labels, which in fatigue research are often derived from subjective self-report scales like the Karolinska Sleepiness Scale (KSS) [243, 244]. While indispensable, subjective feelings do not always align perfectly with objective performance impairment. This creates a critical ambiguity: it is often unclear whether a model is learning to detect an objective physiological state or simply learning to predict a person's subjective report of fatigue [151].

This validation challenge is compounded by a lack of standardized datasets. The vast majority of high-performing models are trained and validated on private, internally collected datasets, often recorded under controlled laboratory or simulator conditions [156]. This absence of public, large-scale benchmarks makes it nearly impossible to perform fair, head-to-head comparisons of different algorithms and contributes to a reproducibility problem, where impressive results reported in one paper cannot be easily verified or built upon by other research groups.

Furthermore, there is a significant generalization gap between model performance in a controlled lab and in a real-world environment. A model achieving high accuracy in a driving simulator with fixed lighting may see its performance degrade significantly in a real vehicle, where it must contend with dynamic sunlight, shadows, frequent head movements, and other unconstrained factors [12]. Studies that attempt real-world validation often report more modest performance, highlighting the brittleness of many current approaches. To advance the field, there is a critical need for public, multimodal benchmark datasets that synchronize ocular data with ground truth labels from multiple sources—subjective reports, objective performance measures, and other physiological signals—to foster the development of systems that are not only accurate in the lab but also robust in the real world.

## 8. Discussion

This review has synthesized the extensive research connecting fatigue-related ocular biomarkers to detection applications. We have traced the field's evolution from foundational neurophysiology to modern computational algorithms. Our analysis reveals a field that has established strong connections in certain areas, particularly the pathway from drowsiness research to detection applications in driving scenarios. However, other significant frontiers remain largely unexplored. This section charts a course for future research by discussing existing work and proposing concrete directions for novel inputs (i.e., new ocular biomarkers), models (i.e., new architectures and training methods), and experimental paradigms (i.e., moving from the lab to "in-the-wild" settings). It also addresses the systemic challenges that must be overcome to translate these technologies into robust and ethically responsible real-world applications.

### *8.1. Expanding the Frontier Beyond Drowsiness*

A primary conclusion of this review is that the field's focus on drowsiness, while successful, leaves other fatigue-related states under-researched. Fatigue is a multifaceted construct that encompasses far more than the simple propensity to sleep. The future of this field lies in leveraging the tools of pupillometry and eye movement analysis to build a nuanced understanding of this broader spectrum. We propose a focused effort to develop objective ocular biomarkers for cognitive states, not only for drowsiness but also for different cognitive states.

One avenue is using eye movements as a direct window into mental fatigue and executive function. A decrease in saccadic peak velocity is a reliable, involuntary marker of declining neural efficiency, reflecting a slowing of the brain's processing speed with time-on-task. The antisaccade task is a more specific probe that assesses inhibitory control governed by the prefrontal cortex. An increase in antisaccade errors directly signals impaired executive function, a core component of cognitive fatigue. Future work should develop paradigms that can measure these subtle saccadic metrics in applied settings.

Pupillometry offers a direct line to the brain's arousal systems. The locus coeruleus–norepinephrine system influences pupil diameter, making it a powerful proxy for cognitive effort and engagement. While lab-grade eye-trackers validate this link, the challenge is translation to ubiquitous sensors. Computer vision must deliver robust pupillometry on standard cameras, enabling smartphones and laptops to monitor cognitive states in everyday environments.

#### *8.1.1. Innovating Research Paradigms and Models*

Progress requires a shift in how we collect data and build models. While small, controlled experiments established foundational links, they must be complemented by large-scale, ecological paradigms that move research from the lab into real-life settings. Most existing studies are lab-based, with very few applied in real-world scenarios. Future studies should collect ocular and behavioural data from thousands of individuals during their daily routines. Such large-scale longitudinal datasets, while less controlled, are essential for developing models that can generalize to real-world complexity.

Multimodal methods have proven to have better performance and robustness than single-modality methods for driving fatigue detection. As ocular biomarkers do not exist in isolation, next-generation fatigue detection systems should be multimodal, fusing ocular data with physiological and behavioural streams. For example, models can integrate gaze data with posture, head movement, and activity levels, resulting in a more holistic and accurate assessment of an individual's fatigue state.

*8.1.2. Overcoming Systemic and Ethical Hurdles*

The field must confront several systemic challenges, the first being the ground truth problem. Many models are trained on subjective self-reports, such as the Karolinska Sleepiness Scale, yet subjective feelings do not always align with objective impairment. Future work should focus on validating a composite ground truth that combines performance lapses, physiological signals, and self-reports. This will create a more valid and reliable target for AI models.

Ethical implications are also central. Continuous monitoring of cognitive states raises critical questions about privacy, data ownership, and misuse. Systems must be built with privacy-preserving algorithms and robust security protocols. Researchers, developers, and policymakers must collaborate to establish clear guidelines for deployment, ensuring these tools enhance safety and well-being without enabling unwarranted surveillance. The goal is to evolve from simple drowsiness alarms to sophisticated, ethically grounded systems that accurately reflect the human cognitive state.

*8.2. Domain Gaps and Computer-Science Frontiers Beyond Driving*

*Knowledge Work and Cognitive Fatigue at the Desktop.* Computer science has not yet built robust detectors for cognitive fatigue in various workplaces, such as for software developers or teachers. In desk work, for instance, current systems rarely observe saccadic vigor, antisaccade errors, or pursuit gain while users type, read, or code. Operating systems and browsers expose few hooks for ocular micro-tasks, and office suites do not log fine-grained interaction states that could help separate overload from boredom. The field lacks models that fuse TEPR attenuation with keystroke timing, mouse micro-pauses, and window-focus entropy, as well as datasets that capture week-long work rhythms with controlled luminance and head pose. This gap blocks the development of adaptive interfaces that could throttle notifications or suggest breaks before errors spike.

*Extended Reality (XR), 3D, and Screen-Induced Visual Fatigue.* Extended reality headsets induce vergence-accommodation conflict and visual strain. Researchers have not yet standardized pipelines that can separate this optical burden from mental fatigue. Few models link fixation instability, pursuit breakdown, or accommodative micro-fluctuations to session-level comfort. Furthermore, rendering stacks seldom expose the photometric metadata needed to normalize pupil traces, and eye-tracking APIs provide landmarks about area of interest objectives but not cognitive or visual fatigue indices. This hinders the development of comfort-aware foveated rendering, session scheduling, and motion sickness prevention.

*Safety-Critical Supervision and Remote Operations.* Air-traffic control, radiology, drone piloting, and industrial control rooms demand sustained vigilance. Computer science has yet to build unobtrusive, high-precision fatigue detectors for these stations. Such systems require low false-alarm rates, calibrated uncertainty, and human-in-the-loop controls. Models need to monitor for performance drift across shifts and handovers, and vision pipelines must be robust to glasses, masks, and dim lighting. The lack of benchmark tasks beyond the PVT that reflect real-world monitoring work limits trust in alarms designed to pause image queues, suggest operator rotations, or trigger secondary reviews.

*Mobile, In-the-Wild Longitudinal Monitoring.* While phone cameras can detect blinks and pupil changes, existing pipelines are fragile. Algorithms often fail due to variations in skin tone, motion artifacts, and auto-exposure adjustments. As users move and tilt their devices, ambient light changes constantly [112, 138]. The field needs more research on motion-tolerant gaze estimation and advanced image processing to enable reliable mobile eye-tracking in unconstrained environments [139, 137, 140, 28]. These advancements depend on more representative datasets and higher-precision gaze estimation algorithms, i.e. integrating computer graphics techniques to model facial structures from individual facial images to gaze estimation pipeline. This gap currently blocks population-scale studies that could map circadian patterns and workload to ocular signals.

*Multimodal Fusion Beyond Video.* Eye video tells only part of the story. IMUs provide behaviour sequence insights, rPPG recovers heart rate variability, microphones capture vocal prosody, and keystroke and pointer traces index cognitive tempo. EEG offers direct insight into brain dynamics. Principled fusion of these streams with ocular features remains a key research frontier. While benefited from the scaling law, large, dataset-driven temporal models may learn to mitigate some sensor-specific noise, most systems do not expose per-modality confidence levels. The field also lacks causal models to disentangle the effects of light, arousal, and task demand, which keeps models brittle when one modality fails.

*Ground Truth, Taxonomy, and Labelling Infrastructure.* Labels for fatigue remain inconsistent across domains, with many datasets conflating sleepiness with cognitive fatigue. Few studies provide both performance-based and physiological anchors for their labels. Annotation tools rarely capture crucial metadata like lighting, display luminance, camera settings, and head distance. There is no shared label ontology that maps tasks to biomarker families, which prevents effective transfer learning and meta-learning across corpora and stalls progress toward generalizable fatigue representations.

*Learning Paradigms for Generalization and Personalization.* Models must generalize across people and devices yet adapt to each individual. Current work underutilizes self-supervised sequence learning on eye-region video, and contrastive learning across modalities remains rare. Test-time adaptation is seldom safeguarded by uncertainty estimation, and continual learning models often suffer from catastrophic forgetting when user patterns (like sleep schedules) shift. The field needs methods that are data-efficient, drift-aware, and energy-efficient for on-device deployment.

*Privacy, Security, and Dual-Use Governance.* Ocular data can reveal identity, health status, and cognitive/emotional states, creating meaningful dual-use risk. We recommend privacy-preserving design as a default rather than an optional add-on. First, systems should prioritize on-device inference and store raw video locally; when possible, only derived ocular events/features (e.g., blink timestamps, TEPR amplitude summaries) should be retained. Second, population-level model improvement should prefer distributed learning systems with secure aggregation so that raw user data never leaves the device. Third, if any telemetry is required, it should use differential privacy (e.g., noise-added feature summaries) and strict data minimization/retention limits. Fourth, sensitive computations can be protected via hardware-backed security (e.g., secure enclaves) and encryption at rest/in transit; access should be gated by explicit consent and purpose limitation.

*Clinical and Pathological Fatigue.* Beyond occupational settings, there is a critical need to apply ocular metrics to clinical populations experiencing pathological fatigue, such as in Myalgic Encephalomyelitis/Chronic Fatigue Syndrome (ME/CFS), multiple sclerosis (MS), or post-viral syndromes. This domain presents unique challenges not found in healthy populations. First, medication confounds, patients often take treatments (e.g., antidepressants, beta-blockers) that directly alter pupillary dynamics and saccadic velocity, requiring baselining protocols that account for pharmacological effects. Second, symptom decoupling, in conditions like MS, subjective fatigue may be high even when motor performance is preserved, or vice-versa, complicating the definition of ground truth. Third, post-exertional malaise (PEM), for ME/CFS patients, standard induction paradigms (like the PVT) may trigger severe, delayed health relapses. Future research must develop "micro-task" assessments—passive monitoring of natural reading or visual search—that can estimate fatigue burden without inducing harm.

### *8.3. Roadmap*

#### *8.3.1. Phenotypes, Tasks, Sensing, and Normalization*

Researchers should define fatigue phenotypes before training models. Cognitive fatigue, for instance, should be mapped to metrics of inhibitory control, sustained attention, and working memory. Protocols should embed micro-tasks into daily use: a one-minute antisaccade block can probe inhibition, a thirty-second smooth pursuit can track cortical timing, and a one-minute TEPR task can measure engagement under fixed luminance. Each task should ship with clear audio-visual guides, randomized timings to reduce anticipation, and logging of ambient light and device exposure settings.

Data collection should favor commodity hardware. A phone's front camera can record the eye region at 30-60 fps, while other sensors (IMU, EEG, etc.) continuously capture body pose, brain activity, and device motion. The rear camera or proximity sensor can estimate face-screen distance, and the microphone can log vocal effort with consent. Each session should store quality indices for each sensor stream. The application should perform on-device segmentation of blinks, saccades, and pupils, with only extracted features and differentially private summaries leaving the device by default.

Pupil features must be robust to changes in light and distance. The pipeline should estimate screen and ambient luminance per trial and invert camera exposure curves to approximate radiance. It should also model corneal reflections to stabilize pupil edges and use head pose to correct for gaze angle and eyelid foreshortening. Blink detection should use both landmark dynamics and intensity profiles. Pursuit gain should be calculated by dividing eye velocity by target velocity in head-centric coordinates. Quality gates must down-weight frames with glare, occlusion, or motion blur.

*8.3.2. Modelling, Fusion, and Learning for Generalization*

Models should be designed to work on sequences. For instance, a vision backbone can convert cropped eye videos into token sequences, an IMU encoder can map head motion into aligned tokens, and a PPG encoder can summarize beat-to-beat intervals with uncertainty. A fusion transformer or a state-space model can then learn the cross-lag relationships between these streams. The model should output fatigue-type logits with calibrated probabilities and per-modality uncertainties. The decoder should also expose a short, textual rationale anchored in specific features, such as a TEPR drop or a loss in pursuit gain.

Pre-training should be self-supervised. A masked-prediction objective can be used to reconstruct eye-region video, while a contrastive objective can align segments from video, IMU, and PPG streams. Fine-tuning should employ a small supervised head with a focal loss for rare events. Test-time adaptation should update normalization layers and low-rank adapters using confidence-weighted steps, freezing global layers to prevent catastrophic drift. Continual [179] or lifelong learning [241, 109] should schedule rehearsal with privacy-preserving prototypes.

*8.3.3. Ground Truth, Labels, and Evaluation*

Labels should combine behaviour and physiology. For example, PVT lapses can provide vigilance anchors, N-back accuracy and response time can provide load anchors, antisaccade error rates can provide inhibition anchors, and TEPR amplitude and latency can provide engagement anchors. Self-report scales should be used to capture fatigue, sleepiness, mood, and pain as separate constructs. Labels must also include context such as hours slept, caffeine intake, and time of day. Each record should include metadata on luminance, camera exposure, and head distance.

Benchmarks and datasets should enforce strict data splits (e.g., subject-independent, device-independent, and illumination-shift) to properly test generalization. Metrics should extend beyond accuracy to include AUROC, AUPRC, calibration error, time-to-detection, and false alarms per hour. Reports must include subgroup performance analysis (by age, sex, skin tone, etc.) and on-device performance metrics like energy per inference and latency. Stress tests should cover corner cases like sudden exposure jumps, motion spikes, and occlusions.

*8.3.4. Personalization, Safety, Privacy, and Governance*

Personalization should be simple and safe. The system should learn a user's baseline for key ocular biomarkers, such as saccadic vigor, during an initial calibration period. Baselines should only be updated when data quality gates pass, and the system should revert to population priors when uncertainty is high. The interface should allow users to view their data trends and to defer or dismiss alerts. Alerting policies should be throttled to avoid alarm fatigue, and high-stakes actions should require confirmation.

Systems should default to edge-computing architectures where inference occurs on-device and raw images are never transmitted. When model updates are necessary, distributed systems should be employed to aggregate gradient updates from users without sharing their personal data. Furthermore, differential privacy protection techniques must be integrated to add noise to the data streams, ensuring that individual re-identification from ocular biometrics is statistically impossible. Synthetic data generation offers another ethical avenue, allowing for the training of robust models on expanding datasets without exposing real human subjects to surveillance risks. Governance frameworks must explicitly prohibit the use of these fatigue metrics for punitive employee monitoring, restricting their use to safety-critical interventions and health awareness.

## 9. Limitations

While this review employed a systematic, data-driven methodology to synthesize a broad and fragmented field, several limitations must be acknowledged. These limitations not only frame the conclusions of this paper but also highlight critical gaps in the primary literature that future research must address.

*9.1. Methodological Heterogeneity of Primary Studies*

A core challenge in synthesizing this body of work is the profound methodological heterogeneity across the included studies. The term "fatigue" is operationalized using a wide array of experimental paradigms, from short-duration cognitive tasks like the n-Back test to prolonged driving simulations and real-world sleep deprivation protocols. Furthermore, the underlying data is captured using a diverse range of hardware, from high-precision, lab-based

eye-trackers to consumer-grade webcams and wearable sensors, each with different sampling rates, accuracy, and noise profiles. While this review does not present a separate quantitative breakdown of every method used in the corpus, this diversity is substantiated throughout our analysis. For instance, Section 6 and Table A.6, B.10 survey the varied experimental tasks used for fatigue induction (e.g., PVT, n-Back, Stroop), while Section 7 and Table 5 document the broad spectrum of analytical pipelines, from classical computer vision with SVMs to modern end-to-end deep learning architectures. This variance in tasks, tools, and participant populations makes a direct quantitative comparison of findings exceptionally difficult. Consequently, this review does not—and cannot—perform a formal meta-analysis of effect sizes, focusing instead on a qualitative synthesis of the research landscape.

### *9.2. Inherent Imbalance in the Research Literature*

The quantitative analysis presented in this review (e.g., Figure 1, Figure 2) clearly demonstrates that the field is heavily skewed towards the study of **drowsiness**, particularly in the context of driver safety. While a primary goal of this paper is to advocate for a broader focus on cognitive, visual, and clinical fatigue, our synthesis is necessarily constrained by the available evidence. The conclusions drawn about drowsiness are supported by a much larger and more mature body of literature than those concerning other fatigue phenotypes. As such, the frameworks and biomarkers for under-researched states like cognitive fatigue in knowledge workers or pathological fatigue in clinical populations are based on a smaller, more nascent evidence base.

### *9.3. Potential Biases in the Review Process*

Like all systematic reviews, this work is susceptible to certain inherent biases. *Publication bias*, the review is based on published articles and conference papers. Research that finds no significant relationship between an ocular metric and a fatigue state (i.e., null results) is less likely to be published. This "file-drawer problem" may lead to an overestimation of the effectiveness and reliability of the biomarkers discussed. Furthermore, the review is subject to potential search and language bias. Although our search strategy was designed to be comprehensive across multiple databases, the specific keywords used may have inadvertently excluded relevant studies from adjacent fields with different terminologies. The restriction to studies published in English may have also omitted valuable research from other regions.

## 10. Conclusion

This review provides a comprehensive synthesis of the use of ocular biomarkers in fatigue research, tracing the field's progression from its neurophysiological roots to contemporary AI-driven detection systems. By deconstructing the multifaceted nature of fatigue and systematically analysing technological frameworks, we have identified a significant gap in the literature: a profound disconnect exists between the mature methods for detecting drowsiness and the under-explored domains of cognitive, visual, and clinical fatigue, highlighting a particular gap between the clinical and computer science fields. We advocate that the future of this research should leverage advanced ocular metrics, such as saccadic dynamics and pupillometry, to break free from the limitations of the driver monitoring paradigm and address these nuanced fatigue states across a broader array of life scenarios. The detailed technical roadmap we have proposed—which emphasizes multimodal fusion, privacy-preserving personalization, and robust "real-world" data collection—offers a clear path forward. Ultimately, this work aims to unify foundational science with computational practice, providing a foundational reference to guide the next generation of research toward creating holistic, ethically-grounded systems that can better understand the complexities of the human fatigue state.

## Acknowledgements

We would like to thank Radek Chrapkiewicz for his insightful discussions and constructive comments that helped shape this work. This research is kindly supported by UKRI BBSRC project EyeWarn (code: APP37953).

## Appendix A. Tables of Experimental Paradigms and Measures

Table A.6: Major fatigue-inducing experimental tasks: procedures, targeted systems, and trade-offs.

| Task (Category) | Procedure & Typical Dose | Primary System(s) | Strengths | Weaknesses / Notes |
|---|---|---|---|---|
| **PVT** (Sustained attention / vigilance) | Simple RT to unpredictable visual onsets; random ISIs; 5–10 min blocks. Lapses (slow/omitted responses) are the primary endpoint. [18] | Tonic alertness; sustained attention. [217] | Gold standard sensitivity to sleep loss/circadian disruption [62, 169]; minimal learning [62]; robust repeatability across settings (lab to operational). [62, 169] | Susceptible to motivation/effort boosts that mask impairment [217]; narrow scope (does not tax WM/executive control). [128] |
| **n-Back** (Working memory load) | Continuous stream; respond when current item matches that 'n' items back; load titrated by 'n' [124]; 20–60 min for induction. [199] | WM updating/-monitoring; interference control; executive control. [20] | Parametrically scalable difficulty [20]; continuous demand yields reliable fatigue build-up [20]; suitable for dose–response designs. | Construct ambiguity at low 'n' (recognition strategies) [124]; can induce frustration; interpretation depends on 'n' level and timing. |
| **Sternberg** (Item recognition) | Encode set of 'k' items, short retention, probe yes/no; manipulate set size; trial-based with microbreaks; 10–20 min per block. [210] | STM scanning/retrieval; serial exhaustive search. [211] | Clean isolation of memory load via set size [114]; well-characterized linear RT vs. set size curve for benchmarking. [211] | Limited fatigue induction due to intermittent recovery [114]; best as a pre/-post probe around a separate inducer. |
| **Stroop Color and Word Test (SCWT)** (Interference/inhibition) | Name ink color while inhibiting word reading; incongruent trials; continuous blocks [150]; ≥30 min recommended for induction. [181] | Inhibitory control; conflict monitoring; executive function. | Strong, reliable interference [150]; sustained mental effort when delivered continuously [181]; objective RT/error decrements. | Short (∼2 min) variants weak as inducers [181]; ecological validity limited outside applied variants. |
| **XR distributed control** (Multi-domain operational) | Team-based VR flight re-entry; distributed roles (e.g., yaw/pitch/thrust); constrained views; continuous coordination under time pressure. [44] | Distributed cognition; teamwork /communication [190]; spatial navigation; multi-domain executive control. | High ecological validity [190]; induces realistic multi-factor fatigue; enables rich multimodal capture (gaze, speech, inputs, physiology). | Resource-intensive; complex analysis (individual vs. team effects) [190]; requires sickness/strain controls (e.g., SSQ). [207] |

Table A.7: Objective ocular metrics relevant to cognitive fatigue: definitions, directional changes, and neurophysiological rationale.

| Metric | Brief Definition | Change w/ Fatigue | Core Neurophysiological Rationale | Notes |
|---|---|---|---|---|
| Tonic pupil size | Baseline pupil diameter at rest or pre-stimulus. | ↓ | Decline in tonic LC–NE activity reduces sympathetic drive to dilator muscle → constricted baseline diameter. [154] | Control luminance and camera exposure; report units and calibration. |

*Continued on next page*

| Metric | Brief Definition | Change w/ Fatigue | Core Neurophysiological Rationale | Notes |
|---|---|---|---|---|
| Task-evoked pupillary response (TEPR) | Transient dilation to cognitive events; amplitude/latency as effort index. [19] | Amplitude ↓ (attenuation) | Lower tonic arousal limits phasic LC–NE gain → reduced event-locked dilation for same task demand. [19] | Use fixed-luminance probes; model blink/occlusion; report baselines. |
| Pupillary light reflex (PLR) | Constriction to light steps; latency/velocity/amplitude. | Amplitude ↓; velocity ↓ | Shift in autonomic balance (PNS/SNS) and reduced responsiveness of iris musculature under drowsiness. [148] | Standardize stimulus photometry; dark adapt where relevant. |
| Pupil unrest index (PUI) | Spontaneous low-frequency oscillations in darkness (mm/min). [195] | ↑ ("fatigue waves") | Instability in central autonomic control as arousal fluctuates (brainstem LC and allied nuclei). [195] | Dark-room recordings; remove respiratory/lighting confounds. |
| Low / High Index of Pupillary Activity (LHIPA) | Low-to-high frequency power ratio of pupil signal (load proxy). [65] | Mixed / context-dependent | High-frequency power tracks rapid phasic processing; low-frequency relates to arousal/luminance; fatigue shifts band balance. [65] | Report band definitions; robust to luminance only if implemented carefully. |
| Blink frequency / duration | Rate and mean closed-time per blink. | ↑ rate and duration [209] | Reduced top-down control (PFC) disinhibits brainstem blink generator; longer blinks approach microsleep. [209, 164] | Track lighting and task state; separate voluntary vs. spontaneous blinks. |
| PERCLOS | % Time eyelid >80% closed in a window. [1] | ↑ | Behavioral manifestation of reduced CNS arousal; linked to DMN/DAN decoupling and vigilance lapses. [1] | Window choice matters; report thresholds; robust in applied settings. |
| Saccade peak velocity (and amplitude) | Ballistic gaze shift speed (and size). | Peak velocity ↓; amplitude may ↓ [15] | Lower excitatory drive to burst neurons (brainstem) from cortical/subcortical arousal systems → slower saccades. [58] | Avoid small-angle measurement bias; exclude blinks/anticipations. |
| Fixation duration / stability | Time eyes are held still; drift/instability indices. | Duration ↑; stability ↓ [131] | Executive disengage deficit under fatigue delays saccade initiation; frontoparietal control less efficient. [131] | Normalize by text/scene structure; report quality/-confidence. |
| VOR cancellation | Voluntary suppression of reflex to keep gaze on head-fixed target. [42] | Impaired (harder to cancel) | Resource-intensive integration (vestibular/visual/proprioceptive; cerebellum & attention networks) degrades with fatigue. [42] | Use head IMU; separate basic VOR gain vs. cancellation. |

Table A.8: Subjective instruments: constructs, formats, best uses, and cautions.

| Instrument | Construct (State vs. Trait) | Format / Range | Best Used For | Notes / Limitations |
|---|---|---|---|---|
| Karolinska Sleepiness Scale (KSS) | State sleepiness (past ∼5 min). [196, 116] | 9-point Likert (1–9) with verbal anchors. [116] | Frequent in-task sampling to track dynamic sleepiness/time-on-task. [116] | Sensitive to momentary context; not a multidimensional fatigue inventory. |
| Stanford Sleepiness Scale (SSS) | State sleepiness (current).[122, 196] | 7-point Likert with descriptive anchors. [122, 196] | Repeated spot checks across protocols; simple single-item administration. [107] | Coarser scale than KSS; similar use-case. [122, 196] |

*Continued on next page*

| Instrument | Construct (State vs. Trait) | Format / Range | Best Used For | Notes / Limitations |
|---|---|---|---|---|
| Multidimensional Fatigue Inventory (MFI-20 / MFI-10) | Trait-like multidomain fatigue (general, physical, activity, motivation, mental). [196] | 20 (or 10) items; 5–7-point Likert per item. [196] | Baseline characterization; pre/post long interventions; clinical cohorts. [196] | Not suited to minute-by-minute tracking within an induction block. [196] |
| Schedule of Fatigue and Anergia (SOFA; SOFA/CFS, SOFA/GP) | Prolonged disabling fatigue; screening severity. [87] | ~10 items; 5-point Likert; versions for specialty and primary care. [123] | Clinical/epidemiological screening for fatigue syndromes. [87] | Limited value in basic cognitive-fatigue lab protocols. [123] |
| Simulator Sickness Questionnaire (SSQ / VRSQ) | Cybersickness (confound). [11] | 16-symptom weighted subscales. [11] | Pre/post XR or simulator exposure to quantify non-fatigue side-effects. [43] | Not a fatigue measure; use to control alternative explanations. [11] |

Table A.9: Canonical cognitive-fatigue experiment: phases, measurements, controls, and outputs.

| Phase | Purpose | Core Measures (examples) | Controls / Setup | Outputs |
|---|---|---|---|---|
| Baseline | Establish stable pre-fatigue reference. [119] | MFI-20 (trait) [108]; KSS/SSS (state) [196]; 3-min PVT [86]; resting ocular baselines (tonic pupil, blink rate). [220] | Constant luminance; fixed camera exposure; head distance/pose logging. | Subject-level normalization; initial calibration and QC indices. |
| Induction (continuous) | Systematically deplete resources. [3] | 45–60 min 2-back *or* long Stroop [174]; continuous eye-tracking; periodic TEPR probes; performance (RT/accuracy). | Quiet/dim room; photometry logged; scheduled brief KSS every 10 min without full recovery. [116] | Time series of behavior, pupil/oculomotor metrics, and subjective sleepiness. |
| Post-task | Quantify net change and transfer. [119] | Final KSS; post PVT; optional post MFI; repeat resting ocular baselines. | Same setup as baseline to ensure comparability. | Pre–post deltas; effect sizes; individual responder profiles. |

## Appendix B. Fatigue Detection Pipelines

Table B.10: Method-family roll-up for fatigue detection: consolidated goals, inputs, typical models, and representative studies.

| Family / Goal | Core Inputs (Signals/Features) | Typical Models / Pipelines | Typical Metrics (reported range) | Representative Studies |
|---|---|---|---|---|
| **Vision-only** eye-state / PERCLOS (driver DMS); real-time alerts | IR/RGB face video; eye closure (PERCLOS), EAR, blink/yawn, head nod | Classical CV + SVM /Haar; CNN /MTCNN; lightweight embedded pipelines | Acc. 80–98%; some field validations; e.g., 83–93–98% | [79, 47, 166, 21, 88, 9, 237, 49, 208, 200, 214, 184, 145, 224] |
| Pupil/iris segmentation & pupil-based workload/fatigue indices | Eye crops; iris/pupil masks; TEPR/tonic pupil; robustness to luminance/emotion | U-Net-like segmentation + decision nets; feature selection + classifiers | Precision/Acc. up to ~97%; robustness under confounds | [245, 225, 110] |

*Continued on next page*

| Family / Goal | Core Inputs (Signals/Features) | Typical Models / Pipelines | Typical Metrics (reported range) | Representative Studies |
|---|---|---|---|---|
| **Multimodal fusion** (EEG/EOG + vision/HR); higher accuracy with fewer sensors | EEG bands; forehead EOG; PERCLOS; HR from video; fusion features | Sparse-DBN; CRF /Neural field; ANN /SVM; dual-indicator logic; 1D-CNN fusion | Acc. ~96–98%; AUC ~0.96; corr. 0.66–0.88 | [239, 142, 242, 64, 101, 151, 78, 234, 111, 218] |
| Cognitive-load /vigilance classification (lab tasks: n-back, Stroop, visual fatigue) | Eye-tracking: pupil /blink /fixation /saccade; task performance; smartphone pupillometry | SVM/linear models; feature ablation/selection; preprocessing pipelines | Binary vigilance 70–80% (best 76.8%); visual-fatigue up to 92.9% | [77, 72, 110, 104, 2] |
| Information-theoretic gaze metrics (entropy-based drowsiness) | Gaze deviation, fixation/transition entropy series | FCE /SGE /GTE indices (no heavy learning) | Strong correlation with KSS; FCE most robust | [243] |
| Operator-behavior action recognition (beyond face/eyes) | Whole-body /upper-body actions (e.g., ATC videos) | CNN action recognition; decision logic | Meets real-time constraints (qualitative) | [54] |
| Data quality, tracking, and preprocessing toolkits (foundational) | Noisy/missing gaze; blink labeling; active IR tracking | Hierarchical interpolation; tracking with Kalman /Gabor; geometric blink algorithms; joint eye region models | Improved extraction accuracy vs. I-VT/U'n'Eye; robust preprocessing | [141, 84, 66, 59] |
| Smartphone/embedded implementations (in-vehicle or naturalistic) | Front camera video; optional audio; on-device features | 3-stage PERCLOS pipelines; lightweight CV; MediaPipe/Dlib stacks | 93% (3-stage) smartphone; MediaPipe ~96% vs. Dlib 87% | [49, 5, 9] |
| Clinical/physiological benchmarks and PST variants | PST features; engineered pupil dynamics; subjective scales (MFI/KSS) | Random Forest; pattern recognition; correlational validation | Acc. ~83% (PST RF); objective–subjective alignment | [127, 172] |
| Oculomotor biometrics under fatigue/inebriation (impact studies) | Micro/macro eye movements under state change | Deep gaze embeddings with state-invariance tuning | ID verification drop under fatigue; mitigation strategies | [152] |

## References


[1] Abe, T. Perclos-based technologies for detecting drowsiness: current evidence and future directions. *Sleep Advances 4*, 1 (2023), zpad006.

[2] Abeysinghe, Y. Evaluating human eye features for objective measure of working memory capacity. In *Proceedings of the 2023 Symposium on Eye Tracking Research and Applications* (2023), pp. 1–3.

[3] Ackerman, P. L., and Kanfer, R. Test length and cognitive fatigue: an empirical examination of effects on performance and test-taker reactions. *Journal of Experimental Psychology: Applied 15*, 2 (2009), 163.

[4] Adão Martins, N. R., Annaheim, S., Spengler, C. M., and Rossi, R. M. Fatigue monitoring through wearables: a state-of-the-art review. *Frontiers in physiology 12* (2021), 790292.

[5] Adhikari, M., Joshi, P., Shrestha, S., and Shaik, S. Vision-based driver drowsiness and distraction detection through behavioral indicators of fatigue. In *SoutheastCon 2025* (2025), pp. 261–266.

[6] Afari, N., and Buchwald, D. Chronic fatigue syndrome: a review. *American Journal of Psychiatry 160*, 2 (2003), 221–236.

[7] Ajayi, O., Kurien, A., Djouani, K., and Dieng, L. A multimodal systematic review of drivers' fatigue detection methodologies, datasets, and models. *IEEE Access* (2025).

[8] Åkerstedt, T., and Gillberg, M. Subjective and objective sleepiness in the active individual. *International journal of neuroscience 52*, 1-2 (1990), 29–37.

[9] Al-Madani, A. M., Gaikwad, A. T., Mahale, V., Ahmed, Z. A., and Shareef, A. A. A. Real-time driver drowsiness detection based on eye movement and yawning using facial landmark. In *2021 International Conference on Computer Communication and Informatics (ICCCI)* (2021), IEEE, pp. 1–4.

[10] Axelsson, J., Sundelin, T., Ingre, M., Van Someren, E. J., Olsson, A., and Lekander, M. Beauty sleep: experimental study on the perceived health and attractiveness of sleep deprived people. *Bmj 341* (2010), c6614.

[11] Aykent, B., Merienne, F., Guillet, C., Paillot, D., and Kemeny, A. Motion sickness evaluation and comparison for a static driving simulator and a dynamic driving simulator. *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering 228*, 7 (2014), 818–829.

[12] Azim, T., Jaffar, M. A., and Mirza, A. M. Automatic fatigue detection of drivers through pupil detection and yawning analysis. In *2009 fourth international conference on innovative computing, information and control (ICICIC)* (2009), IEEE, pp. 441–445.

[13] Badham, S. P., and Hutchinson, C. V. Characterising eye movement dysfunction in myalgic encephalomyelitis/chronic fatigue syndrome. *Graefe's Archive for Clinical and Experimental Ophthalmology 251*, 12 (2013), 2769–2776.

[14] Baek, J. H., Park, H., Kang, H., Kim, R., Kang, J. S., and Kim, H. J. The role of glutamine homeostasis in emotional and cognitive functions. *International Journal of Molecular Sciences 25*, 2 (2024), 1302.

[15] Bafna, T., and Hansen, J. P. Mental fatigue measurement using eye metrics: A systematic literature review. *Psychophysiology 58*, 6 (2021), e13828.

[16] Bafna-Rührer, T., Baekgaard, P., and Hansen, J. P. Smooth-pursuit performance during eye-typing from memory indicates mental fatigue. *Journal of Eye Movement Research 15*, 4 (2022), 10–16910.

[17] Basner, M., and Dinges, D. F. Maximizing sensitivity of the psychomotor vigilance test (pvt) to sleep loss. *Sleep 34*, 5 (2011), 581–591.

[18] Basner, M., Mollicone, D., and Dinges, D. F. Validity and sensitivity of a brief psychomotor vigilance test (pvt-b) to total and partial sleep deprivation. *Acta astronautica 69*, 11-12 (2011), 949–959.

[19] Beatty, J. Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological bulletin 91*, 2 (1982), 276.

[20] Behrens, M., Gube, M., Chaabene, H., Prieske, O., Zenon, A., Broscheid, K.-C., Schega, L., Husmann, F., and Weippert, M. Fatigue and human performance: an updated framework. *Sports medicine 53*, 1 (2023), 7–31.

[21] Bhowmick, B., and Kumar, K. C. Detection and classification of eye state in ir camera for driver drowsiness identification. In *2009 IEEE International Conference on Signal and Image Processing Applications* (2009), IEEE, pp. 340–345.

[22] Billones, R., Liwang, J. K., Butler, K., Graves, L., and Saligan, L. N. Dissecting the fatigue experience: A scoping review of fatigue definitions, dimensions, and measures in non-oncologic medical conditions. *Brain, behavior, & immunity-health 15* (2021), 100266.

[23] Bimberg, P., Weissker, T., and Kulik, A. On the usage of the simulator sickness questionnaire for virtual reality research. In *2020 IEEE conference on virtual reality and 3D user interfaces abstracts and workshops (VRW)* (2020), IEEE, pp. 464–467.

[24] Blain, B., Hollard, G., and Pessiglione, M. Neural mechanisms underlying the impact of daylong cognitive work on economic decisions. *Proceedings of the National Academy of Sciences 113*, 25 (2016), 6967–6972.

[25] Bocca, M., Marie, S., and Chavoix, C. Impaired inhibition after total sleep deprivation using an antisaccade task when controlling for circadian modulation of performance. *Physiology & behavior 124* (2014), 123–128.

[26] Bocca, M.-L., and Denise, P. Total sleep deprivation effect on disengagement of spatial attention as assessed by saccadic eye movements. *Clinical Neurophysiology 117*, 4 (2006), 894–899.

[27] Bogucki, A., John, I., Zinkiewicz, Ł., Jachura, M., Jaworski, D., Suwała, K., Chrost, H., Wlodarski, M., Kałużny, J., Campbell, D., et al. Machine learning approach for ambient-light-corrected parameters and the pupil reactivity (pure) score in smartphone-based pupillometry. *Frontiers in Neurology 15* (2024), 1363190.

[28] Cai, X., Zeng, J., Shan, S., and Chen, X. Source-free adaptive gaze estimation by uncertainty reduction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2023), pp. 22035–22045.

[29] Cantor, F. Central and peripheral fatigue: exemplified by multiple sclerosis and myasthenia gravis. *PM&R 2*, 5 (2010), 399–405.

[30] Carroll, T. J., Taylor, J. L., and Gandevia, S. C. Recovery of central and peripheral neuromuscular fatigue after exercise. *Journal of Applied Physiology 122*, 5 (2017), 1068–1076.

[31] Caseras, X., Mataix-Cols, D., Giampietro, V., Rimes, K. A., Brammer, M., Zelaya, F., Chalder, T., and Godfrey, E. L. Probing the working memory system in chronic fatigue syndrome: a functional magnetic resonance imaging study using the n-back task. *Psychosomatic Medicine 68*, 6 (2006), 947–955.

[32] Castro, S. C., Strayer, D. L., Matzke, D., and Heathcote, A. Cognitive workload measurement and modeling under divided attention. *Journal of experimental psychology: human perception and performance 45*, 6 (2019), 826.

[33] Cedillo, A. P. M., and Foulsham, T. Don't look at it! the effects of instruction on attention to social and non-social elements in complex images. *Journal of Vision 22*, 14 (2022), 3231–3231.

[34] Chand, S. P., and Arif, H. Depression. In *StatPearls [Internet]*. StatPearls Publishing, 2023.

[35] Chen, J.-T., Kuo, Y.-C., Hsu, T.-Y., and Wang, C.-A. Fatigue and arousal modulations revealed by saccade and pupil dynamics. *International journal of environmental research and public health 19*, 15 (2022), 9234.

[36] Chervin, R. D. *Sleepiness versus fatigue, tiredness, and lack of energy*. Cambridge University Press, 2014, p. 7–32.

[37] Clark, V. L., and Kruse, J. A. Clinical methods: the history, physical, and laboratory examinations. *Jama 264*, 21 (1990), 2808–2809.

[38] Collet, J., Ftouni, S., Clough, M., Cain, S. W., Fielding, J., and Anderson, C. Differential impact of sleep deprivation and circadian timing on reflexive versus inhibitory control of attention. *Scientific Reports 10*, 1 (2020), 7270.

[39] Corbin, L., and Marquer, J. Is sternberg's memory scanning task really a short-term memory task? *Swiss Journal of Psychology* (2013).

[40] Corfield, E. C., Martin, N. G., and Nyholt, D. R. Co-occurrence and symptomatology of fatigue and depression. *Comprehensive psychiatry 71* (2016), 1–10.

[41] Cortes Rivera, M., Mastronardi, C., Silva-Aldana, C. T., Arcos-Burgos, M., and Lidbury, B. A. Myalgic encephalomyelitis/chronic fatigue syndrome: a comprehensive review. *Diagnostics 9*, 3 (2019), 91.

[42] Crampton, A., Teel, E., Chevignard, M., and Gagnon, I. Vestibular-ocular reflex dysfunction following mild traumatic brain injury: a narrative review. *Neurochirurgie 67*, 3 (2021), 231–237.

[43] Cross, J., Boag-Hodgson, C., Ryley, T., Mavin, T. J., and Potter, L. E. Using extended reality in flight simulators: a literature review. *IEEE transactions on visualization and computer graphics 29*, 9 (2022), 3961–3975.

[44] Cross, J. I., and Boag-Hodgson, C. C. A collaborative virtual reality flight simulator: Efficacy, challenges and potential. *IEEE Transactions on Learning Technologies* (2025).

[45] Cuijpers, P., Stringaris, A., and Wolpert, M. Treatment outcomes for depression: challenges and opportunities. *The Lancet Psychiatry 7*, 11 (2020), 925–927.

[46] Daguet, I., Bouhassira, D., and Gronfier, C. Baseline pupil diameter is not a reliable biomarker of subjective sleepiness. *Frontiers in Neurology 10* (2019), 108.

[47] Darshana, S., Fernando, D., Jayawardena, S., Wickramanayake, S., and DeSilva, C. Efficient perclos and gaze measurement methodologies to estimate driver attention in real time. In *2014 5th International Conference on Intelligent Systems, Modelling and Simulation* (2014), IEEE, pp. 289–294.

[48] Dasgupta, A., George, A., Happy, S., and Routray, A. A vision-based system for monitoring the loss of attention in automotive drivers. *IEEE Transactions on Intelligent Transportation Systems 14*, 4 (2013), 1825–1838.

[49] Dasgupta, A., Rahman, D., and Routray, A. A smartphone-based drowsiness detection and warning system for automotive drivers. *IEEE transactions on intelligent transportation systems 20*, 11 (2018), 4045–4054.

[50] Davis, M. P., and Walsh, D. Mechanisms of fatigue. *J Support Oncol 8*, 4 (2010), 164–74.

[51] Dawson, J., and Foulsham, T. Your turn to speak? audiovisual social attention in the lab and in the wild. *Visual Cognition 30*, 1-2 (2022), 116–134.

[52] De Gennaro, L., Ferrara, M., Urbani, L., and Bertini, M. Oculomotor impairment after 1 night of total sleep deprivation: a dissociation between measures of speed and accuracy. *Clinical Neurophysiology 111*, 10 (2000), 1771–1778.

[53] de Rodez Benavent, S. A., Nygaard, G. O., Harbo, H. F., Tønnesen, S., Sowa, P., Landrø, N. I., Wendel-Haga, M., Etholm, L., Nilsen, K. B., Drolsum, L., et al. Fatigue and cognition: Pupillary responses to problem-solving in early multiple sclerosis patients. *Brain and Behavior 7*, 7 (2017), e00717.

[54] Deng, Y., and Sun, Y. A method to determine the fatigue of air traffic controller by action recognition. In *2020 IEEE 2nd International Conference on Civil Aviation Safety and Information Technology (ICCASIT* (2020), IEEE, pp. 95–97.

[55] Di Stasi, L. L., Catena, A., Cañas, J. J., Macknik, S. L., and Martinez-Conde, S. Saccadic velocity as an arousal index in naturalistic tasks. *Neuroscience & Biobehavioral Reviews 37*, 5 (2013), 968–975.

[56] Di Stasi, L. L., McCamy, M. B., Catena, A., Macknik, S. L., Canas, J. J., and Martinez-Conde, S. Microsaccade and drift dynamics reflect mental fatigue. *European Journal of Neuroscience 38*, 3 (2013), 2389–2398.

[57] Di Stasi, L. L., McCamy, M. B., Macknik, S. L., Mankin, J. A., Hooft, N., Catena, A., and Martinez-Conde, S. Saccadic eye movement metrics reflect surgical residents' fatigue. *Annals of surgery 259*, 4 (2014), 824–829.

[58] Di Stasi, L. L., Renner, R., Staehr, P., Helmert, J. R., Velichkovsky, B. M., Cañas, J. J., Catena, A., and Pannasch, S. Saccadic peak velocity sensitivity to variations in mental workload. *Aviation, space, and environmental medicine 81*, 4 (2010), 413–417.

[59] Ding, L., Terwilliger, J., Parab, A., Wang, M., Fridman, L., Mehler, B., and Reimer, B. Clera: a unified model for joint cognitive load and eye region analysis in the wild. *ACM Transactions on Computer-Human Interaction 30*, 6 (2023), 1–23.

[60] Dinges, D. F., and Powell, J. W. Microcomputer analyses of performance on a portable, simple visual rt task during sustained operations. *Behavior research methods, instruments, & computers 17*, 6 (1985), 652–655.

[61] Dolan, G. P., Stone, D. H., and Briggs, A. H. A systematic review of continuous performance task research in children prenatally exposed to alcohol. *Alcohol & Alcoholism 45*, 1 (2010), 30–38.

[62] Dorrian, J., Rogers, N. L., and Dinges, D. F. Psychomotor vigilance performance: Neurocognitive assay sensitive to sleep loss. *Sleep deprivation* (2004), 39–70.

[63] Drummond, S. P., Bischoff-Grethe, A., Dinges, D. F., Ayalon, L., Mednick, S. C., and Meloy, M. The neural basis of the psychomotor vigilance task. *Sleep 28*, 9 (2005), 1059–1068.

[64] Du, G., Zhang, L., Su, K., Wang, X., Teng, S., and Liu, P. X. A multimodal fusion fatigue driving detection method based on heart rate and perclos. *IEEE transactions on intelligent transportation systems 23*, 11 (2022), 21810–21820.

[65] Duchowski, A. T., Krejtz, K., Gehrer, N. A., Bafna, T., and Bækgaard, P. The low/high index of pupillary activity. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (2020), pp. 1–12.

[66] Egloff, T., Hall, E., Caporusso, N., and Thaman, B. A blink detection algorithm for eye-tracking applications based on facial geometry. In *2024 47th MIPRO ICT and Electronics Convention (MIPRO)* (2024), IEEE, pp. 1381–1386.

[67] El-Nabi, S. A., El-Shafai, W., El-Rabaie, E.-S. M., Ramadan, K. F., Abd El-Samie, F. E., and Mohsen, S. Machine learning and deep learning techniques for driver fatigue and drowsiness detection: a review. *Multimedia Tools and Applications 83*, 3 (2024), 9441–9477.

[68] Eldar, E., Cohen, J. D., and Niv, Y. The effects of neural gain on attention and learning. *Nature Neuroscience 16*, 8 (Jun 2013), 1146–1153.

[69] Enoka, R. M., and Duchateau, J. Translating fatigue to human performance. *Medicine and science in sports and exercise 48*, 11 (2016), 2228.

[70] Evans, W. J., and Lambert, C. P. Physiological basis of fatigue. *American Journal of Physical Medicine & Rehabilitation 86*, 1 (2007), S29–S46.

[71] Fan, L., Wang, J., Li, Q., Song, Z., Dong, J., Bao, F., and Wang, X. Eye movement characteristics and visual fatigue assessment of virtual reality games with different interaction modes. *Frontiers in neuroscience 17* (2023), 1173127.

[72] Farha, N. A., Al-Shargie, F., Tariq, U., and Al-Nashash, H. Artifact removal of eye tracking data for the assessment of cognitive vigilance levels. In *2021 Sixth International Conference on Advances in Biomedical Engineering (ICABME)* (2021), IEEE, pp. 175–179.

[73] Fielding, J., Clough, M., Beh, S., Millist, L., Sears, D., Frohman, A. N., Lizak, N., Lim, J., Kolbe, S., Rennaker, R. L., et al. Ocular motor signatures of cognitive dysfunction in multiple sclerosis. *Nature Reviews Neurology 11*, 11 (2015), 637–645.

[74] Foulsham, T., and Kingstone, A. Covert orienting: the dark matter of social attention. *Trends in Cognitive Sciences* (2025).

[75] Fransson, P.-A., Patel, M., Magnusson, M., Berg, S., Almbladh, P., and Gomez, S. Effects of 24-hour and 36-hour sleep deprivation on smooth pursuit and saccadic eye movements. *Journal of Vestibular Research 18*, 4 (2008), 209–222.

[76] Friedman, L., and Komogortsev, O. V. Evidence for five types of fixation during a random saccade eye tracking task and changes in fixation duration as a function of time-on-task. *Plos one 19*, 9 (2024), e0310436.

[77] Gambiraža, M., Kesedžić, I., Šarlija, M., Popović, S., and Ćosić, K. Classification of cognitive load based on oculometric features. In *2021 44th International Convention on Information, Communication and Electronic Technology (MIPRO)* (2021), IEEE, pp. 377–382.

[78] Gao, X.-Y., Zhang, Y.-F., Zheng, W.-L., and Lu, B.-L. Evaluating driving fatigue detection algorithms using eye tracking glasses. In *2015 7th International IEEE/EMBS Conference on Neural Engineering (NER)* (2015), IEEE, pp. 767–770.

[79] Garcia, I., Bronte, S., Bergasa, L. M., Almazán, J., and Yebes, J. Vision-based drowsiness detector for real driving conditions. In *2012 IEEE Intelligent Vehicles Symposium* (2012), IEEE, pp. 618–623.

[80] Gelfand, J. M., and Douglas, V. C. *Fatigue*. McGraw-Hill Education, New York, NY, 2018.

[81] Gilzenrat, M. S., Nieuwenhuis, S., Jepma, M., and Cohen, J. D. Pupil diameter tracks changes in control state predicted by the adaptive gain theory of locus coeruleus function. *Cognitive, Affective, & Behavioral Neuroscience 10*, 2 (2010), 252–269.

[82] Goldich, Y., Barkana, Y., Pras, E., Zadok, D., Hartstein, M., and Morad, Y. The effects of sleep deprivation on oculomotor responses. *Current eye research 35*, 12 (2010), 1135–1141.

[83] Graw, P., Kräuchi, K., Knoblauch, V., Wirz-Justice, A., and Cajochen, C. Circadian and wake-dependent modulation of fastest and slowest reaction times during the psychomotor vigilance task. *Physiology & behavior 80*, 5 (2004), 695–701.

[84] Gu, H., Ji, Q., and Zhu, Z. Active facial tracking for fatigue detection. In *Sixth IEEE Workshop on Applications of Computer Vision, 2002.(WACV 2002). Proceedings.* (2002), IEEE, pp. 137–142.

[85] Guo, F., Zhou, Y., Wang, X., Li, W., and Cai, J. Literature review of driving fatigue research based on bibliometric analysis. *Journal of Traffic and Transportation Engineering (English Edition)* (2024).

[86] Gupta, C. C., Centofanti, S., Dorrian, J., Coates, A., Stepien, J. M., Kennaway, D., Wittert, G., Heilbronn, L., Catcheside, P., Noakes, M., et al. Altering meal timing to improve cognitive performance during simulated nightshifts. *Chronobiology international 36*, 12 (2019), 1691–1713.

[87] Hadzi-Pavlovic, D., Hickie, I., Wilson, A., Davenport, T., Lloyd, A., and Wakefield, D. Screening for prolonged fatigue syndromes: validation of the sofa scale. *Social psychiatry and psychiatric epidemiology 35*, 10 (2000), 471–479.

[88] Hasan, M. K., Ullah, S. H., Gupta, S. S., and Ahmad, M. Drowsiness detection for the perfection of brain computer interface using viola-jones algorithm. In *2016 3rd International Conference on Electrical Engineering and Information Communication Technology (ICEEICT)* (2016), IEEE, pp. 1–5.

[89] Helton, W. S., Kern, R. P., and Walker, D. R. Conscious thought and the sustained attention to response task. *Consciousness and cognition 18*, 3 (2009), 600–607.

[90] Henningsen, P., Fink, P., Hausteiner-Wiehle, C., and Rief, W. Terminology, classification and concepts. *Medically unexplained symptoms, somatisation and bodily distress: Developing better clinical services* (2011), 43–68.

[91] Hess, E. H., and Polt, J. M. Pupil size as related to interest value of visual stimuli. *Science 132*, 3423 (1960), 349–350.

[92] Hikosaka, O., Takikawa, Y., and Kawagoe, R. Role of the basal ganglia in the control of purposive saccadic eye movements. *Physiological reviews 80*, 3 (2000), 953–978.

[93] Hockey, R. *The psychology of fatigue: Work, effort and control*. Cambridge University Press, 2013.

[94] Hoddes, E., Zarcone, V., Smythe, H., Phillips, R., and Dement, W. C. Quantification of sleepiness: a new approach. *Psychophysiology 10*, 4 (1973), 431–436.

[95] Holtzer, R., Shuman, M., Mahoney, J. R., Lipton, R., and Verghese, J. Cognitive fatigue defined in the context of attention networks. *Aging, Neuropsychology, and Cognition 18*, 1 (2010), 108–128.

[96] Hooda, R., Joshi, V., and Shah, M. A comprehensive review of approaches to detect fatigue using machine learning techniques. *Chronic Diseases and Translational Medicine* (2021).

[97] Hopstaken, J. F., Van Der Linden, D., Bakker, A. B., and Kompier, M. A. A multifaceted investigation of the link between mental fatigue and task disengagement. *Psychophysiology 52*, 3 (2015), 305–315.

[98] Horne, J. A., and Pettitt, A. N. High incentive effects on vigilance performance during 72 hours of total sleep deprivation. *Acta psychologica 58*, 2 (1985), 123–139.

[99] Horng, W.-B., Chen, C.-Y., Chang, Y., and Fan, C.-H. Driver fatigue detection based on eye tracking and dynamic template matching. In *IEEE International Conference on Networking, Sensing and Control, 2004* (2004), vol. 1, IEEE, pp. 7–12.

[100] Hossain, G., and Yeasin, M. Understanding effects of cognitive load from pupillary responses using hilbert analytic phase. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops* (2014), pp. 375–380.

[101] Huang, Y., Deng, Y., and Shi, C. Driving fatigue identification based on multiple non-intrusive indicators. In *2021 6th International Conference on Transportation Information and Safety (ICTIS)* (2021), IEEE, pp. 118–123.

[102] Huyghe, T., Calleja-González, J., Bird, S., and Alcaraz, P. E. Pupillometry as a new window to player fatigue? a glimpse inside the eyes of a euro cup women's basketball team. *Biology of Sport 41*, 1 (2024), 3–15.

[103] Ishii, A., Tanaka, M., and Watanabe, Y. Neural mechanisms of mental fatigue. *Reviews in the Neurosciences 25*, 4 (2014), 469–479.

[104] Islam, R., Zhang, T., Bisen, P. S., and Bae, S. W. Moodpupilar: Predicting mood through smartphone detected pupillary responses in naturalistic settings. In *2024 IEEE 20th International Conference on Body Sensor Networks (BSN)* (2024), IEEE, pp. 1–4.

[105] Jabbar, M. S., Qasim, H. B., Ali, A. H., Al-Salihi, A. K. J., Al-Hilali, A. A., and Majeed, A. H. Based on multitask cascaded convolutional networks mtcnn, implementation of behavioral drowsiness detection system. In *2023 First International Conference on Advances in Electrical, Electronics and Computational Intelligence (ICAEECI)* (2023), IEEE, pp. 1–6.

[106] Jacobsen, P. B., Donovan, K. A., and Weitzner, M. A. Distinguishing fatigue and depression in. In *Seminars in clinical neuropsychiatry* (2003), vol. 8, pp. 229–240.

[107] Jason, L. A., Evans, M., Brown, M., and Porter, N. What is fatigue? pathological and nonpathological fatigue. *PM&R 2*, 5 (2010), 327–331.

[108] Jensen, A. R., and Rohwer Jr, W. D. The stroop color-word test: a review. *Acta psychologica 25* (1966), 36–93.

[109] Jha, S., Schiemer, M., Zambonelli, F., and Ye, J. Continual learning in sensor-based human activity recognition: An empirical benchmark analysis. *Information Sciences 575* (2021), 1–21.

[110] Ji, Z., Li, N., and Lin, Y. Decoding visual fatigue through pupil dynamics. In *2023 20th China International Forum on Solid State Lighting & 2023 9th International Forum on Wide Bandgap Semiconductors (SSLCHINA: IFWS)* (2023), IEEE, pp. 424–428.

[111] Jiao, Y., and Lu, B.-L. Detecting slow eye movement for recognizing driver's sleep onset period with eeg features. In *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (2016), IEEE, pp. 4658–4661.

[112] John, I., Yari, Z., Bogucki, A., Swiatek, M., Chrost, H., Wlodarski, M., Chrapkiewicz, R., and Li, J. Unsupervised deep learning-driven stabilization of smartphone-based quantitative pupillometry for mobile emergency medicine. In *2024 IEEE International Symposium on Biomedical Imaging (ISBI)* (2024), IEEE, pp. 1–5.

[113] Joshi, S., and Gold, J. I. Pupil size as a window on neural substrates of cognition. *Trends in cognitive sciences 24*, 6 (2020), 466–480.

[114] Jou, J. Task-switching cost and repetition priming: Two overlooked confounds in the fixed-set procedure of the sternberg paradigm and how they affect memory set-size effects. *Quarterly Journal of Experimental Psychology 67*, 10 (2014), 1871–1894.

[115] Kahneman, D., and Beatty, J. Pupil diameter and load on memory. *Science 154*, 3756 (1966), 1583–1585.

[116] Kaida, K., Takahashi, M., Åkerstedt, T., Nakata, A., Otsuka, Y., Haratani, T., and Fukasawa, K. Validation of the karolinska sleepiness scale against performance and eeg variables. *Clinical neurophysiology 117*, 7 (2006), 1574–1581.

[117] Kaifie, A., Reugels, M., Kraus, T., and Kursawe, M. The pupillary light reflex (plr) as a marker for the ability to work or drive–a feasibility study. *Journal of Occupational Medicine and Toxicology 16*, 1 (2021), 39.

[118] Kamti, M. K., and Iqbal, R. Evolution of driver fatigue detection techniques—a review from 2007 to 2021. *Transportation research record 2676*, 12 (2022), 485–507.

[119] Karim, E., Pavel, H. R., Nikanfar, S., Hebri, A., Roy, A., Nambiappan, H. R., Jaiswal, A., Wylie, G. R., and Makedon, F. Examining the landscape of cognitive fatigue detection: A comprehensive survey. *Technologies 12*, 3 (2024), 38.

[120] Kaur, K., Gurnani, B., Nayak, S., Deori, N., Kaur, S., Jethani, J., Singh, D., Agarkar, S., Hussaindeen, J. R., Sukhija, J., et al. Digital eye strain-a comprehensive review. *Ophthalmology and therapy 11*, 5 (2022), 1655–1680.

[121] Kennedy, R. S., Lane, N. E., Berbaum, K. S., and Lilienthal, M. G. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The international journal of aviation psychology 3*, 3 (1993), 203–220.

[122] Keramidas, M. E., Siebenmann, C., Norrbrand, L., Gadefors, M., and Eiken, O. A brief pre-exercise nap may alleviate physical performance impairments induced by short-term sustained operations with partial sleep deprivation–a field-based study. *Chronobiology international 35*, 10 (2018), 1464–1470.

[123] Kim, S., Bae, W. K., Kim, J. Y., Jang, M., Kim, J. H., and Noh, H. H. Validation of the korean version of schedule of fatigue and anergia: general physician questionnaire. *Journal of Korean Medical Science 31*, 2 (2016), 159.

[124] Kirchner, W. K. Age differences in short-term retention of rapidly changing information. *Journal of experimental psychology 55*, 4 (1958), 352.

[125] Kluger, B. M., Krupp, L. B., and Enoka, R. M. Fatigue and fatigability in neurologic illnesses: proposal for a unified taxonomy. *Neurology 80*, 4 (2013), 409–416.

[126] Kok, A. Cognitive control, motivation and fatigue: A cognitive neuroscience perspective. *Brain and Cognition 160* (2022), 105880.

[127] Krajewski, J., Schnupp, T., Schnieder, S., Sommer, D., Heinze, C., and Golz, M. Pattern recognition methods: a novel analysis for the pupillographic sleepiness test. In *Proceedings of the 7th International Conference on Methods and Techniques in Behavioral Research* (2010), pp. 1–4.

[128] Kwon, S. J., Kim, Y., and Kwak, Y. Relationship of sleep quality and attention deficit hyperactivity disorder symptoms with quality of life in college students. *Journal of American College Health 68*, 5 (2020), 536–542.

[129] Laeng, B., Sirois, S., and Gredebäck, G. Pupillometry: A window to the preconscious? *Perspectives on psychological science 7*, 1 (2012), 18–27.

[130] Lal, S. K., and Craig, A. A critical review of the psychophysiology of driver fatigue. *Biological psychology 55*, 3 (2001), 173–194.

[131] Lanthier, S., Risko, E., Smilek, D., and Kingstone, A. Measuring the separate effects of practice and fatigue on eye movements during visual search. In *Proceedings of the annual meeting of the cognitive science society* (2013), vol. 35.

[132] Larsen, R. S., and Waters, J. Neuromodulatory correlates of pupil dilation. *Frontiers in neural circuits 12* (2018), 21.

[133] Larson, S. T., and Wilbur, J. Muscle weakness in adults: evaluation and differential diagnosis. *American family physician 101*, 2 (2020), 95–108.

[134] Lei, Y., Cho, H., Buchanan, F., He, S., Gong, X., Wang, Y., and Ye, J. Tinygaze: Lightweight gaze-gesture recognition on commodity mobile devices. In *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems* (2026).

[135] Lei, Y., Gong, X., He, S., Wang, Y., Khamis, M., and Ye, J. The people's gaze: Co-designing and refining gaze gestures with users and experts. In *Proceedings of the 2026 CHI conference on human factors in computing systems* (2026).

[136] Lei, Y., He, S., Feng, H., Zhao, K., Khamis, M., and Ye, J. Protecting privacy in an era of pervasive camera-based devices: Challenges and potential directions. In *Proc. UK Mobile, Wearable and Ubiquitous Systems Research Symposium* (2023).

[137] Lei, Y., He, S., Khamis, M., and Ye, J. An end-to-end review of gaze estimation and its interactive applications on handheld mobile devices. *ACM Computing Surveys 56*, 2 (2023), 1–38.

[138] Lei, Y., Wang, Y., Buchanan, F., Zhao, M., Sugano, Y., He, S., Khamis, M., and Ye, J. Quantifying the impact of motion on 2d gaze estimation in real-world mobile interactions. *arXiv preprint arXiv:2502.10570* (2025).

[139] Lei, Y., Wang, Y., Caslin, T., Wisowaty, A., Zhu, X., Khamis, M., and Ye, J. Dynamicread: Exploring robust gaze interaction methods for reading on handheld mobile devices under dynamic conditions. *Proceedings of the ACM on Human-Computer Interaction 7*, ETRA (2023), 1–17.

[140] Lei, Y., Zhao, M., Wang, Y., He, S., Sugano, Y., Khamis, M., and Ye, J. Mac-gaze: Motion-aware continual calibration for mobile gaze tracking. *arXiv preprint arXiv:2505.22769* (2025).

[141] Li, F., Chen, C.-H., Xu, G., and Khoo, L.-P. Hierarchical eye-tracking data analytics for human fatigue detection at a traffic control center. *IEEE Transactions on Human-Machine Systems 50*, 5 (2020), 465–474.

[142] Li-Hong, Z., Ning, D., and Cai-Kun, Y. Fatigue-driving recognition based on the state of the human eye and electroencephalographic signal. In *2016 Chinese Control and Decision Conference (CCDC)* (2016), IEEE, pp. 1567–1572.

[143] Liu, F., Chen, D., Zhou, J., and Xu, F. A review of driver fatigue detection and its advances on the use of rgb-d camera and deep learning. *Engineering Applications of Artificial Intelligence 116* (2022), 105399.

[144] Liu, F., Li, X., Lv, T., and Xu, F. A review of driver fatigue detection: Progress and prospect. In *2019 IEEE International conference on consumer electronics (ICCE)* (2019), IEEE, pp. 1–6.

[145] Liu, G., Yan, D., and Chen, Z. Research on early warning of driver fatigue status based on image processing. In *2021 33rd Chinese Control and Decision Conference (CCDC)* (2021), IEEE, pp. 2678–2681.

[146] Liversedge, S., Gilchrist, I., and Everling, S. *The Oxford handbook of eye movements*. Oxford University Press, 2011.

[147] Loh, S., Lamond, N., Dorrian, J., Roach, G., and Dawson, D. The validity of psychomotor vigilance tasks of less than 10-minute duration. *Behavior Research Methods, Instruments, & Computers 36*, 2 (2004), 339–346.

[148] Lowenstain, O. Pupillary movements during acute and chronic fatigue. *Invest Ophthalmol Vis Sci 2* (1963), 138–157.

[149] Lowenstein, O., Feinberg, R., and Loewenfeld, I. E. Pupillary movements during acute and chronic fatigue: A new test for the objective evaluation of tiredness. *Investigative Ophthalmology & Visual Science 2*, 2 (1963), 138–157.

[150] MacLeod, C. M. Half a century of research on the stroop effect: an integrative review. *Psychological bulletin 109*, 2 (1991), 163.

[151] Mahmood, H. R., Irshayyid, M. S., Hussien, M. K., et al. Measuring the fatigue of persons based on eog signal processing mmethods. In *2022 5th International Conference on Engineering Technology and its Applications (IICETA)* (2022), IEEE, pp. 283–288.

[152] Makowski, S., Prasse, P., Jäger, L. A., and Scheffer, T. Oculomotoric biometric identification under the influence of alcohol and fatigue. In *2022 IEEE International Joint Conference on Biometrics (IJCB)* (2022), pp. 1–9.

[153] Mathôt, S. Pupillometry: Psychology, physiology, and function. *Journal of cognition 1*, 1 (2018), 16.

[154] McLaughlin, D. J., Zink, M. E., Gaunt, L., Reilly, J., Sommers, M. S., Van Engen, K. J., and Peelle, J. E. Give me a break! unavoidable fatigue effects in cognitive pupillometry. *Psychophysiology 60*, 7 (2023), e14256.

[155] Miley, A. Å., Kecklund, G., and Åkerstedt, T. Comparing two versions of the karolinska sleepiness scale (kss). *Sleep and biological rhythms 14*, 3 (2016), 257–260.

[156] Misra, A., Samuel, S., Cao, S., and Shariatmadari, K. Detection of driver cognitive distraction using machine learning methods. *IEEE Access 11* (2023), 18000–18012.

[157] Muhammed, K., Dalmaijer, E., Manohar, S., and Husain, M. Voluntary modulation of saccadic peak velocity associated with individual differences in motivation. *Cortex 122* (2020), 198–212.

[158] Murphy, P. R., Robertson, I. H., Balsters, J. H., and O'connell, R. G. Pupillometry and p3 index the locus coeruleus–noradrenergic arousal function in humans. *Psychophysiology 48*, 11 (2011), 1532–1543.

[159] Nandyal, S., et al. Driver drowsiness estimation based on hybrid feature extraction and light weighted dense convolutional network. In *2022 IEEE International Conference on Distributed Computing and Electrical Circuits and Electronics (ICDCECE)* (2022), IEEE, pp. 1–6.

[160] Neshov, N., and Manolova, A. Drowsiness monitoring in real-time based on supervised descent method. In *2017 9th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS)* (2017), vol. 2, IEEE, pp. 660–663.

[161] Nguyen, T., Chew, M.-T., and Demidenko, S. Eye tracking system to detect driver drowsiness. In *2015 6th International conference on automation, robotics and applications (ICARA)* (2015), IEEE, pp. 472–477.

[162] Olson, R. L., Morgan, J. F., Hanowski, R. J., Daily, B., Zimmermann, R. P., Blanco, M., Bocanegra, J. L., Fitch, G. M., and Flintsch, A. M. Assessment of a drowsy driver warning system for heavy vehicle drivers.

[163] Ozuah, P. O., and Sigler, A. Fatigue and weakness. *Primary pediatric care (4th ed., pp. 1079–1084). St. Louis: Mosby* (2001).

[164] Page, C., Liu, C. C., Meltzer, J., and Ghosh Hajra, S. Blink-related oscillations provide naturalistic assessments of brain function and cognitive workload within complex real-world multitasking environments. *Sensors 24*, 4 (2024), 1082.

[165] Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., et al. The prisma 2020 statement: an updated guideline for reporting systematic reviews. *bmj 372* (2021).

[166] Pandey, N. N., and Muppalaneni, N. B. Real-time drowsiness identification based on eye state analysis. In *2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS)* (2021), IEEE, pp. 1182–1187.

[167] Pattyn, N., Van Cutsem, J., Dessy, E., and Mairesse, O. Bridging exercise science, cognitive psychology, and medical practice: Is "cognitive fatigue" a remake of "the emperor's new clothes"? *Frontiers in Psychology 9* (2018), 1246.

[168] Pessiglione, M., Blain, B., Wiehler, A., and Naik, S. Origins and consequences of cognitive fatigue. *Trends in Cognitive Sciences* (2025).

[169] Petrofsky, L. A., Heffernan, C. M., Gregg, B. T., and Smith-Forbes, E. V. Effects of sleep deprivation in military service members on cognitive performance: a systematic review. *Military Behavioral Health 10*, 3 (2022), 202–220.

[170] Phillips, R. O. A review of definitions of fatigue–and a step towards a whole definition. *Transportation research part F: traffic psychology and behaviour 29* (2015), 48–56.

[171] Pignoni, G., Komandur, S., and Volden, F. Accounting for effects of variation in luminance in pupillometry for field measurements of cognitive workload. *IEEE Sensors Journal 21*, 5 (2020), 6393–6400.

[172] Piljugin, O., Badarin, A., Antipov, V., Grubov, V., Yana, B., Tynterova, A., Rafalskiy, V., Shusharina, N., and Hramov, A. Analysis of eye-tracking data during the sternberg working memory task in subjects with asthenic syndrome. In *2022 6th Scientific School Dynamics of Complex Networks and their Applications (DCNA)* (2022), IEEE, pp. 219–222.

[173] Pillai, P., Balasingam, B., and Biondi, F. N. Model-based estimation of mental workload in drivers using pupil size measurements. In *2023 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM)* (2023), IEEE, pp. 815–821.

[174] Pillai, P., Balasingam, B., Jaekel, A., and Biondi, F. N. Comparison of concurrent cognitive load measures during n-back tasks. *Applied Ergonomics 117* (2024), 104244.

[175] Pouliquen-Lardy, L., Milleville-Pennel, I., Guillaume, F., and Mars, F. Remote collaboration in virtual reality: asymmetrical effects of task distribution on spatial processing and mental workload. *Virtual Reality 20*, 4 (2016), 213–220.

[176] Prsa, M., Dicke, P. W., and Thier, P. The absence of eye muscle fatigue indicates that the nervous system compensates for non-motor disturbances of oculomotor function. *Journal of Neuroscience 30*, 47 (2010), 15834–15842.

[177] Pucker, A. D., Kerr, A. M., Sanderson, J., and Lievens, C. Digital eye strain: updated perspectives. *Clinical Optometry* (2024), 233–246.

[178] Qi, P., Ru, H., Gao, L., Zhang, X., Zhou, T., Tian, Y., Thakor, N., Bezerianos, A., Li, J., and Sun, Y. Neural mechanisms of mental fatigue revisited: New insights from the brain connectome. *Engineering 5*, 2 (2019), 276–286.

[179] Rahman, F., Schiemer, M., Sanabria, A. R., and Ye, J. Continual learning in sensor-based human activity recognition with dynamic mixture of experts. *Pervasive and Mobile Computing 110* (2025), 102044.

[180] Ranzijn, R., and Lack, L. The pupillary light reflex cannot be used to measure sleepiness. *Psychophysiology 34*, 1 (1997), 17–22.

[181] Rauch, W., and Schmitt, K. Fatigue of cognitive control in the stroop-task. In *Proceedings of the annual meeting of the cognitive science society* (2009), vol. 31.

[182] Rayner, K., Slattery, T. J., and Bélanger, N. N. Eye movements, the perceptual span, and reading speed. *Psychonomic bulletin & review 17*, 6 (2010), 834–839.

[183] Ricci, J. A., Chee, E., Lorandeau, A. L., and Berger, J. Fatigue in the us workforce: prevalence and implications for lost productive work time. *Journal of occupational and environmental medicine 49*, 1 (2007), 1–10.

[184] Rimal, R., Herceg, M., Vranješ, M., and Grbić, R. Driver monitoring system using an embedded computer platform. In *2023 International Conference on Software, Telecommunications and Computer Networks (SoftCOM)* (2023), IEEE, pp. 1–7.

[185] Rio, C. J., Gehling, G. M., Blumhorst, C., Ross, A., and Saligan, L. N. Defining fatigue from the experiences of patients living with chronic fatigue. *Frontiers in Medicine 11* (2024), 1429275.

[186] Roebuck, H., Freigang, C., and Barry, J. G. Continuous performance tasks: Not just about sustaining attention. *Journal of speech, language, and hearing research 59*, 3 (2016), 501–510.

[187] Rönnbäck, L., and Johansson, B. Long-lasting pathological mental fatigue after brain injury–a dysfunction in glutamate neurotransmission? *Frontiers in behavioral neuroscience 15* (2022), 791984.

[188] Rosenthal, T. C., Majeroni, B. A., Pretorius, R., and Malik, K. Fatigue: an overview. *American family physician 78*, 10 (2008), 1173–1179.

[189] Roy, D., and Nah, F. F.-H. A review on eye-tracking metrics for sleepiness. In *International Conference on Human-Computer Interaction* (2020), Springer, pp. 630–640.

[190] Rybing, J., Nilsson, H., Jonson, C.-O., and Bang, M. Studying distributed cognition of simulation-based team training with dicot. *Ergonomics 59*, 3 (2016), 423–434.

[191] Ryu, H., Ju, U., and Wallraven, C. Decoding visual fatigue in a visual search task selectively manipulated via myopia-correcting lenses. *Frontiers in Neuroscience 18* (2024), 1307688.

[192] Schiphof-Godart, L., Roelands, B., and Hettinga, F. J. Drive in sports: How mental fatigue affects endurance performance. *Frontiers in psychology 9* (2018), 1383.

[193] Schleicher, R., Galley, N., Briest, S., and Galley, L. Blinks and saccades as indicators of fatigue in sleepiness warnings: looking tired? *Ergonomics 51*, 7 (2008), 982–1010.

[194] Schmidt, D., Abel, L., DellOsso, L., and Daroff, R. Saccadic velocity characteristics- intrinsic variability and fatigue. *Aviation, space, and environmental medicine 50*, 4 (1979), 393–395.

[195] Schumann, A., Kietzer, S., Ebel, J., and Bär, K. J. Sympathetic and parasympathetic modulation of pupillary unrest. *Frontiers in Neuroscience 14* (2020), 178.

[196] Shahid, A., Wilkinson, K., Marcu, S., and Shapiro, C. M. *STOP, THAT and one hundred other sleep scales*. Springer, 2012.

[197] Sharifian, F., Schneider, D., Arnau, S., and Wascher, E. Decoding of cognitive processes involved in the continuous performance task. *International Journal of Psychophysiology 167* (2021), 57–68.

[198] Shen, J., Barbera, J., and Shapiro, C. M. Distinguishing sleepiness and fatigue: focus on definition and measurement. *Sleep medicine reviews 10*, 1 (2006), 63–76.

[199] Shigihara, Y., Tanaka, M., Ishii, A., Kanai, E., Funakura, M., and Watanabe, Y. Two types of mental fatigue affect spontaneous oscillatory brain activities in different ways. *Behavioral and Brain Functions 9*, 1 (2013), 2.

[200] Siddik, M. A. D., and Rahman, M. S. Drowsiness and lethargy detection using machine learning techniques. In *2023 International conference on electrical, computer and communication engineering (ECCE)* (2023), IEEE, pp. 1–6.

[201] Sieger, T., Bonnet, C., Serranova, T., Wild, J., Novak, D., Ruuzicka, F., Urgosik, D., Ruzicka, E., Gaymard, B., and Jech, R. Basal ganglia neuronal activity during scanning eye movements in parkinson's disease. *PLoS One 8*, 11 (2013), e78581.

[202] Sikander, G., and Anwar, S. Driver fatigue detection systems: A review. *IEEE Transactions on Intelligent Transportation Systems 20*, 6 (2018), 2339–2352.

[203] Skau, S., Sundberg, K., and Kuhn, H.-G. A proposal for a unifying set of definitions of fatigue. *Frontiers in Psychology 12* (2021), 739764.

[204] Smets, E., Garssen, B., Bonke, B. d., and De Haes, J. The multidimensional fatigue inventory (mfi) psychometric qualities of an instrument to assess fatigue. *Journal of psychosomatic research 39*, 3 (1995), 315–325.

[205] Smid, H., De Witte, M., Homminga, I., and Van Den Bosch, R. Sustained and transient attention in the continuous performance task. *Journal of clinical and experimental neuropsychology 28*, 6 (2006), 859–883.

[206] Solís-Marcos, I., Galvao-Carmona, A., and Kircher, K. Reduced attention allocation during short periods of partially automated driving: An event-related potentials study. *Frontiers in human neuroscience 11* (2017), 537.

[207] Somerville, A., Joiner, K., and Wild, G. Effects of flight experience or simulator exposure on simulator sickness in virtual reality flight simulation. *Multimodal Technologies and Interaction 9*, 3 (2025), 24.

[208] SP, R. M., Reshma, D., et al. A real-time fatigue detection system using multi-task cascaded cnn model. In *2021 10th IEEE International Conference on Communication Systems and Network Technologies (CSNT)* (2021), IEEE, pp. 674–679.

[209] Stern, J. A., Boyer, D., and Schroeder, D. Blink rate: a possible measure of fatigue. *Human factors 36*, 2 (1994), 285–297.

[210] Sternberg, S. High-speed scanning in human memory. *science 153*, 3736 (1966), 652–654.

[211] Sternberg, S. In defence of high-speed memory scanning. *Quarterly Journal of Experimental Psychology 69*, 10 (2016), 2020–2075.

[212] Sundelin, T., Lekander, M., Kecklund, G., Van Someren, E. J., Olsson, A., and Axelsson, J. Cues of fatigue: effects of sleep deprivation on facial appearance. *Sleep 36*, 9 (2013), 1355–1360.

[213] Szczęśniewski, P., Dudzic, M., Badowski, K., Świątek, M., Płatnikow, M., Klawiter, J., Olszewski, W., Kowalow, M., Paryzek, Z., Pawlak, K., et al. Computational pupillometry and the pupil reactivity (pure) score in neurocritical care: Prospects for high-accuracy, lighting-invariant neuromonitoring. *Cureus 17*, 9 (2025).

[214] Tabrizi, P. R., and Zoroofi, R. A. Open/closed eye analysis for drowsiness detection. In *2008 first workshops on image processing theory, tools and applications* (2008), IEEE, pp. 1–7.

[215] Taylor, J. L., Amann, M., Duchateau, J., Meeusen, R., and Rice, C. L. Neural contributions to muscle fatigue: from the brain to the muscle and back again. *Medicine and science in sports and exercise 48*, 11 (2016), 2294.

[216] Theofilidis, G., Bogdanis, G. C., Koutedakis, Y., and Karatzaferi, C. Monitoring exercise-induced muscle fatigue and adaptations: making sense of popular or emerging indices and biomarkers. *Sports 6*, 4 (2018), 153.

[217] Thomann, J., Baumann, C. R., Landolt, H.-P., and Werth, E. Psychomotor vigilance task demonstrates impaired vigilance in disorders with excessive daytime sleepiness. *Journal of Clinical Sleep Medicine 10*, 9 (2014), 1019–1024.

[218] Tokadli, G., Woodruff, K., Wu, P., and Matthews, C. How fatigued could the pilot be? fatigue levels through pvt and physiological indications. In *2024 AIAA DATC/IEEE 43rd Digital Avionics Systems Conference (DASC)* (2024), IEEE, pp. 1–6.

[219] Tornero-Aguilera, J. F., Jimenez-Morcillo, J., Rubio-Zarapuz, A., and Clemente-Suárez, V. J. Central and peripheral fatigue in physical exercise explained: A narrative review. *International journal of environmental research and public health 19*, 7 (2022), 3909.

[220] Tsukahara, J. S., Harrison, T. L., and Engle, R. W. The relationship between baseline pupil size and intelligence. *Cognitive psychology 91* (2016), 109–123.

[221] United States. Federal Motor Carrier Safety Administration. Technology Division. Perclos: A valid psychophysiological measure of alertness as assessed by psychomotor vigilance. Tech. rep., United States Federal Motor Carrier Safety Administration, 1998.

[222] van der Leeuw, G., Leveille, S. G., Jones, R. N., Hausdorff, J. M., McLean, R., Kiely, D. K., Gagnon, M., and Milberg, W. P. Measuring attention in very old adults using the test of everyday attention. *Aging, Neuropsychology, and Cognition 24*, 5 (2017), 543–554.

[223] Van der Wel, P., and Van Steenbergen, H. Pupil dilation as an index of effort in cognitive control tasks: A review. *Psychonomic bulletin & review 25* (2018), 2005–2015.

[224] Wang, P., and Shen, L. A method of detecting driver drowsiness state based on multi-features of face. In *2012 5th International Congress on Image and Signal Processing* (2012), pp. 1171–1175.

[225] Wang, W., Li, Z., Wang, Y., and Chen, F. Indexing cognitive workload based on pupillary response under luminance and emotional changes. In *Proceedings of the 2013 international conference on Intelligent user interfaces* (2013), pp. 247–256.

[226] Wiggins, M. W. Vigilance decrement during a simulated general aviation flight. *Applied Cognitive Psychology 25*, 2 (2011), 229–235.

[227] Wilhelm, B., Giedke, H., Lüdtke, H., Bittner, E., Hofmann, A., and Wilhelm, H. Daytime variations in central nervous system activation measured by a pupillographic sleepiness test. *Journal of sleep research 10*, 1 (2001), 1–7.

[228] Wilhelm, B., Wilhelm, H., Lüdtke, H., Streicher, P., and Adler, M. Pupillographic assessment of sleepiness in sleep-deprived healthy subjects. *Sleep 21*, 3 (1998), 258–265.

[229] Wilhelm, B. J., Widmann, A., Durst, W., Heine, C., and Otto, G. Objective and quantitative analysis of daytime sleepiness in physicians after night duties. *International Journal of Psychophysiology 72*, 3 (2009), 307–313.

[230] Williamson, A., Lombardi, D. A., Folkard, S., Stutts, J., Courtney, T. K., and Connor, J. L. The link between fatigue and safety. *Accident Analysis & Prevention 43*, 2 (2011), 498–515.

[231] Wilshire, C. E., Kindlon, T., Courtney, R., Matthees, A., Tuller, D., Geraghty, K., and Levin, B. Rethinking the treatment of chronic fatigue syndrome—a reanalysis and evaluation of findings from a recent major trial of graded exercise and cbt. *BMC psychology 6*, 1 (2018), 6.

[232] Wylie, G., Yao, B., Genova, H., Chen, M., and DeLuca, J. Using functional connectivity changes associated with cognitive fatigue to delineate a fatigue network. *Scientific reports 10*, 1 (2020), 21927.

[233] Xiao-yan, C., Kai, S., Chun-peng, W., Hong-xuan, Z., and Chun-yuan, L. Analysis of pupil size amplitude signal in field fatigue detection. In *2020 7th International Conference on Information Science and Control Engineering (ICISCE)* (2020), IEEE, pp. 302–305.

[234] Xiao-yu, W., Nai-meng, C., Wan-jun, Y., Zi-chen, W., Huai-lin, Z., and Jia-lan, L. Driver's eeg eye movement fatigue detection based on cmac. In *2019 International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS)* (2019), IEEE, pp. 57–63.

[235] Yang, B., Intoy, J., and Rucci, M. Eye blinks as a visual processing stage. *Proceedings of the National Academy of Sciences 121*, 15 (2024), e2310291121.

[236] Zargari Marandi, R., Madeleine, P., Omland, Ø., Vuillerme, N., and Samani, A. Eye movement characteristics reflected fatigue development in both young and elderly individuals. *Scientific reports 8*, 1 (2018), 13148.

[237] Zeng, S., Li, J., Jiang, L., and Jiang, J. A driving assistant safety method based on human eye fatigue detection. In *2017 29th Chinese Control And Decision Conference (CCDC)* (2017), IEEE, pp. 6370–6377.

[238] Zhang, H., Ni, D., Ding, N., Sun, Y., Zhang, Q., and Li, X. Structural analysis of driver fatigue behavior: A systematic review. *Transportation research interdisciplinary perspectives 21* (2023), 100865.

[239] Zhang, J., and Xiao, N. Eeg and forehead eog-based driver fatigue classification using sparse-deep belief networks. In *2021 IEEE 23rd Int Conf on High Performance Computing & Communications; 7th Int Conf on Data Science & Systems; 19th Int Conf on Smart City; 7th Int Conf on Dependability in Sensor, Cloud & Big Data Systems & Application (HPCC/DSS/SmartCity/DependSys)* (2021), IEEE, pp. 1059–1066.

[240] Zhang, Z., Ning, H., and Zhou, F. A systematic survey of driving fatigue monitoring. *IEEE transactions on intelligent transportation systems 23*, 11 (2022), 19999–20020.

[241] Zheng, J., Qiu, S., Shi, C., and Ma, Q. Towards lifelong learning of large language models: A survey. *ACM Computing Surveys 57*, 8 (2025), 1–35.

[242] Zheng, W.-L., Gao, K., Li, G., Liu, W., Liu, C., Liu, J.-Q., Wang, G., and Lu, B.-L. Vigilance estimation using a wearable eog device in real driving environment. *IEEE transactions on intelligent transportation systems 21*, 1 (2019), 170–184.

[243] Zhu, Y., Xu, Q., Zhen, K., Zhang, R., and Wang, S. Quantitative analysis of eye-tracking data based on information-theoretic tools for measuring driver drowsiness. In *2024 IEEE International Conference on Multimedia and Expo (ICME)* (2024), IEEE, pp. 1–6.

[244] Zhu, Y., Xu, Q., Zhen, K., Zhang, R., and Wang, S. Quantitative analysis of eye-tracking data based on information-theoretic tools for measuring driver drowsiness. In *2024 IEEE International Conference on Multimedia and Expo (ICME)* (2024), pp. 1–6.

[245] Zhuang, Q., Kehua, Z., Wang, J., and Chen, Q. Driver fatigue detection method based on eye states with pupil and iris segmentation. *Ieee Access 8* (2020), 173440–173449.

[246] Zils, E., Sprenger, A., Heide, W., Born, J., and Gais, S. Differential effects of sleep deprivation on saccadic eye movements. *Sleep 28*, 9 (2005), 1109–1115.