

Review of eye tracking metrics involved in emotional and cognitive processes

Vasileios Skaramagkas¹, Giorgos Giannakakis^{1,2}, *Member, IEEE*, Emmanouil Ktistakis^{1,3}, Dimitris Manousos¹, Ioannis Karatzanis¹, Nikolaos S. Tachos⁴, *Member, IEEE*, Evanthia Tripoliti⁴, *Member, IEEE*, Kostas Marias^{1,6}, *Member, IEEE*, Dimitrios I. Fotiadis⁵, *Fellow, IEEE*, and Manolis Tsiknakis^{1,6}, *Member, IEEE*

Abstract—Eye behaviour provides valuable information revealing one’s higher cognitive functions and state of affect. Although eye tracking is gaining ground in the research community, it is not yet a popular approach for the detection of emotional and cognitive states. In this paper, we present a review of eye and pupil tracking related metrics (such as gaze, fixations, saccades, blinks, pupil size variation, etc.) utilized towards the detection of emotional and cognitive processes, focusing on visual attention, emotional arousal and cognitive workload. Besides, we investigate their involvement as well as the computational recognition methods employed for the reliable emotional and cognitive assessment. The publicly available datasets employed in relevant research efforts were collected and their specifications and other pertinent details are described. The multimodal approaches which combine eye-tracking features with other modalities (e.g. biosignals), along with artificial intelligence and machine learning techniques were also surveyed in terms of their recognition/classification accuracy. The limitations, current open research problems and prospective future research directions were discussed for the usage of eye-tracking as the primary sensor modality. This study aims to comprehensively present the most robust and significant eye/pupil metrics based on available literature towards the development of a robust emotional or cognitive computational model.

Index Terms—eye tracking, gaze, pupil, fixations, saccades, smooth pursuit, blinks, stress, visual attention, emotional arousal, cognitive workload, emotional arousal datasets, cognitive workload datasets

Vasileios Skaramagkas, Giorgos Giannakakis, Emmanouil Ktistakis, Dimitris Manousos, Ioannis Karatzanis are with the Institute of Computer Science, Foundation for Research and Technology Hellas (FORTH), GR-700 13 Heraklion, Crete, Greece (Email: vskaramag@ics.forth.gr, ggian@ics.forth.gr, mandim@ics.forth.gr, karatzan@ics.forth.gr)

Giorgos Giannakakis is with the Institute of AgriFood and Life Sciences, University Research Centre, Hellenic Mediterranean University, Heraklion, Greece.

Emmanouil Ktistakis is with the Institute of Computer Science, Foundation for Research and Technology Hellas (FORTH) and the Laboratory of Optics and Vision, School of Medicine, University of Crete, Heraklion, Greece (Email: mankti@ics.forth.gr)

Nikolaos S. Tachos and Evanthia E. Tripoliti, are with the Department of Biomedical Research, Institute of Molecular Biology and Biotechnology, FORTH, GR-451 10, Ioannina, Greece (Email: ntachos@gmail.com, etripoliti@gmail.com)

Dimitrios I. Fotiadis is with the Department of Biomedical Research, Institute of Molecular Biology and Biotechnology, FORTH, Ioannina, Greece and the Department of Materials Science and Engineering, Unit of Medical Technology and Intelligent Information Systems, University of Ioannina, GR-451 10, Ioannina, Greece (phone: +302651009006, fax: +302651008889, e-mail: fotiadis@uoi.gr.)

Kostas Marias and Manolis Tsiknakis are with the Institute of Computer Science, Foundation for Research and Technology Hellas (FORTH) and the Department of Electric and Computer Engineering, Hellenic Mediterranean University, GR-710 04 Heraklion, Crete, Greece (Email: kmarias@ics.forth.gr, tsiknaki@ics.forth.gr)

I. INTRODUCTION

The investigation of emotional and cognitive processes that modulate human behaviour requires a comprehensive research approach. Various psychophysiological and psychophysical modalities (such as electroencephalography (EEG), event-related potentials (ERP), electrodermal activity (EDA), electrocardiography (ECG), facial expressions, body posture, etc.) have been employed in the relevant emotion -or attention-related literature. The eye and pupillary response in relation to emotional or cognitive processing provides valuable information for one’s higher cognitive function and state of affect [1], [2], [3]. However, there is not a concrete comprehensive guide of the utilization of eye and pupil behaviour towards this objective.

Over the last years, there is an increasing research effort in the area of emotional and cognitive functions. A significant portion of this research is based on neurophysiological data, investigating the pattern and behaviour of the implicated neural networks. Another approach is the investigation of the human body’s physiological and physical measures (biosignals). Eye and pupil behavioural measures, although been mediated by the autonomic nervous system (ANS) just like biosignals, have remained in the research background [4].

Lately, the evolution of eye tracking hardware/software enables their usage on convenient wearable devices boosted the related eye tracking research. Robust eye trackers, in terms of accuracy, portability and ease of use, have been developed, that are able to unobtrusively monitor eye movements in real-time. Among the different eye tracking systems, the head-mounted type has become the most popular since they are used in daily indoor/outdoor activities. In addition, various computational algorithms have been developed to efficiently extract metrics associated with the behaviour of the eye.

To our knowledge, a comprehensive guide of specific eye patterns during cognitive and emotional processing does not exist. Most related studies apply machine learning techniques to the multivariate ocular feature set in order to categorize the data into predefined user’s states.

In this manuscript, the ocular features are investigated in the context of user’s emotional arousal, visual attention and cognitive workload. Emotional arousal is a fundamental concept in emotion theory, which is involved in the majority of affective experiences. It constitutes one of the two main axes in Russell’s two-dimensional circular space of emotions [5]. In this perspective, the level of arousal level is the structural ele-

ment contributing, to a lesser or greater degree, to all emotions. Along the same line, visual attention and cognitive workload are significant indices of the human cognitive function and performance. Reduced performance may imply a deficiency of information processing. This may be caused by a limited pool of resource (cognitive capacity), or by a more conservative response under higher cognitive workloads (response caution) [6].

According to the literature, among the different eye movement metrics, there can be an initial categorization between those that are most correlated with visual attention, those that are more relevant with emotional arousal and those that are best indicators for the cognitive workload. However, there are no specific or unique eye features that provide a discrimination capability among visual, emotional and cognitive processes. For this reason, the collection of the various eye features and their positive or negative influence on an individual's emotional or cognitive state is vital.

In the present review, we investigate eye and pupil behaviour related metrics that best describe and express the processes of emotional arousal, visual attention and cognitive workload. Firstly, we determine the scope of these three emotional/cognitive processes (section II) and we describe the nature and the corresponding underlying physiological functions of the eye movements and pupil behaviour features (section III). A short presentation of eye-tracker systems focusing on the most widely used, the video-based trackers and the techniques of gaze estimation (Section IV) is being performed. Then, this paper's main scope is described which is the investigation and the association of eye/pupil metrics with the three emotional/cognitive processes in urban environments and daily life activities (section V). Besides, we provide comprehensive information on the emotional/cognitive recognition methods either only using eye metrics or using multimodal approaches (section VI). Finally, publicly available eye behaviour datasets are presented (section VII). The fields of research reviewed relate to eye metrics involved in emotional/cognitive processes (e.g. emotional arousal, visual attention, cognitive workload) including basic research in order for the reader to be able to recognize the most robust features that can be utilized on her/his research. The selection of the studies is based on the scope of our review and the investigation of the correlation of eye metrics with affective and cognitive processes in urban environments and daily life activities such as driving, reading, working on a PC etc.

The main aim of this review is to

- identify the eye metrics that have a significant relation to the investigated emotional arousal, cognitive workload, visual attention processes
- determine efficient recognition methods based on eye metrics for the investigated emotional arousal, cognitive workload, visual attention processes
- specify the most robust and relevant ocular combined feature set in a multimodal approach.

This analysis expects to aid related future experimental design and research towards an efficient selection of eye metrics.

II. EMOTIONAL AND COGNITIVE PROCESSES

In natural vision cognitive and affective factors influence an observer's visual attention. Cognitive and emotional factors play a dominant role in active gaze control and can determine attention allocation in complex scenes [7]. It is known that emotional arousal modifies the allocation of attentional resources [8] but also conversely, how attention is allocated during emotional arousal can significantly alter an emotional state [9]. Both emotion and attention seem to modulate both early and late stages of visual processing [8].

In this section, the emotional and cognitive processes are defined in detail and their specific characteristics are presented. Visual attention is first presented and then, emotional arousal and cognitive workload as factors that affect visual attention.

A. Visual attention

The process by which a user selects a specific element from all the available information in order to further examine is called *visual attention*. In other words, the term "visual attention" refers to the collection of various cognitive operations that isolate the relevant from irrelevant information from cluttered visual scenes [10]. Attention remains a crucial area of investigation within education, psychology, cognitive neuroscience, and neuropsychology [11]. In recent years, active research is being carried out to determine the source of sensory and attention-triggering signals, the effects of these sensory points on the coordination properties of the sensory neurons as well as the relationship between attention and other behavioural and cognitive processes including memory and psychological vigilance.

B. Emotional arousal and stress

Emotional arousal is a state that describes the level of calmness (i.e., low arousal) or excitation (i.e., high arousal) elicited by a stimulus. In [12], arousal is defined as a global feeling of dynamism or lethargy that involves mental activity and physical preparedness to act. The most common manifestation of increased arousal and negative valence is stress. There are various dimensional models of affect, the most known of which is the circumplex model of Russell [5], which maps emotions along predefined axes. The physiological stress response modulates, among other body functions, eye behaviour and function [13]. Previous studies have investigated the relationship between the level of emotional arousal and response inhibition [14].

C. Cognitive workload

The *cognitive workload* is defined as the level of an individual's measurable mental effort in order to cope with one or more cognitively demanding tasks [15]. According to cognitive load theory, there are three types of cognitive load: 1) intrinsic load 2) extraneous load and 3) germane load [16]. The intrinsic load comes from the complexity of the task and its association with the user, extraneous load is caused by the presentation style of the material and germane load refers to the ability of the user to fully understand the material [17].

Although the cognitive workload is considered to be a subjective personality property, cognitive load can be partially demarcated under three quantified measurement types:

- performance
- subjective
- physiological.

Performance measurements assess workload through the ability of a user to perform tasks or functions of a system. Subjective measurements are based on the judgments of the users regarding the workload associated with the execution of a task or a system function. Physiological measurements evaluate the physiological responses of the user during specific task demands [18].

III. FEATURES REPRESENTING EYE AND PUPIL BEHAVIOUR

In this section, some of the most important features that represent the behaviour of eye motion are presented and discussed.

A. Visual fixations

Ocular fixation is defined as the maintenance of eye gaze on a single location [19]. Human beings can fixate only when they possess a fovea in the anatomy of the eye. A fovea is typically located at the center of macula lutea of the retina and dictates the point of clearest vision [20], [21]. In recent years, following the development of eye trackers, the detection of fixation related metrics has become more robust and easier to implement. The most commonly used fixation metrics used are: number of fixations, number of fixations on each area of interest, total number of fixations, fixation duration, total fixation duration i.e. the cumulative duration of all of the fixations on a particular area of interest (AOI), time to first fixation on target, fixation density and repeat fixations [22].

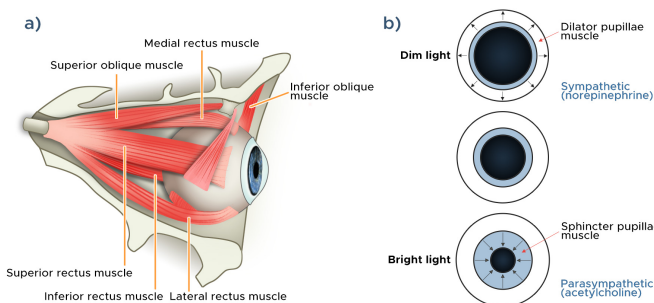


Fig. 1. (a) Eye anatomy responsible for the eye movements. The superior and inferior rectus muscles are responsible for the eye's vertical movements, whereas the lateral and medial rectus muscles control horizontal movements, (b) Function and implicated muscles of pupil dilation and contraction.

B. Saccades

Saccadic movements or *saccades* are called the instantaneous and ballistic changes of the eyes between fixation points [23]. Their amplitude can be very small such as in reading situations or relatively large when, for example, a user is gazing around a room [24]. Saccades, just like fixations, can be discriminated into involuntary or voluntary depending on

their duration. Usually, a saccadic movement appears with a frequency of 2 or 3 times every second. The most commonly used saccadic movement features are: number of saccades, saccadic velocity/amplitude, saccade rate and duration [25].

C. Microsaccades

Microsaccades form one of the three fixational eye movements, along with tremor and drift [26]. They serve as a counteraction to retinal adaptation by generating small random displacements of the retinal image in stationary viewing [27]. They move the retinal image in a distance of some hundreds of cones and they have a relatively constant duration of 25 msec and this is why a linear correlation of their peak velocity with their amplitude is being observed [26]. Microsaccades are most possibly conjugate eye movements.

D. Smooth pursuit eye movements

Smooth pursuit eye movements allow the gaze to be maintained on selected objects regardless of whether the subject or the objects are stationary or moving [28]. Smooth pursuit eye movements have a maximum velocity of about 100/s and latency 100-130 ms. Drugs, fatigue, alcohol, and even distraction degrade the quality of these movements.

E. Pupil

The size of the pupil is controlled by two sets of muscles, the constrictor and dilator pupillae, which are governed by the sympathetic (SNS) and parasympathetic (PNS) divisions of the autonomic nervous system (ANS) [29]. It reflects autonomic involuntary activity and it is associated with emotional, cognitive or sexual arousal [30]. The pupil size variation is a characteristic measure that is affected and is characteristic in the investigation of mental or cognitive processes [13]. However, pupil metrics are susceptible to issues affecting their reliability that need to be taken into consideration. It is known that pupil's sensitivity to illumination conditions may affect the pupil size [31], [32] leading to the constriction of the pupil when the amount of light increases. Besides, pupil size is also affected by environmental conditions such as humidity and temperature [33]. Age is another significant parameter for pupil size variation, as it is referred a marked reduction of pupil size with the ageing [34]. Furthermore, pupil metrics can also be affected by the position of the camera and the angle of recording [35]. Therefore, placing the experimental stimuli at the center of the FOV or using well-established reference points or estimating pupil size in the normal state could be a good practice when investigating pupil metrics.

F. Eye blinks

Eye blinking is a semi-involuntary action of fast closing and reopening of the eyelid. It occurs as a result of the co-inhibition of eyelid's protractors and retractors muscles. Blinking serves the spreading of the corneal tear film across the frontal surface of the cornea [36]. The average duration of a single eye blink is 0.1-0.4 sec [37]. The blink rate (BR), measured in blinks/min, is influenced by environmental factors

(humidity, temperature, brightness), and physical activity [38]. Also, research evidence suggests that eye blink rate may be tied with emotional and cognitive processing, especially with attentional engagement and mental workload [39].

IV. VIDEO-BASED EYE AND PUPIL TRACKING SYSTEMS AND METHODS

The eye tracking technology is being used in a wide variety of disciplines, including psychology, human-computer interactions as well as in commercial and safety applications [40]. In the respective literature, various eye tracking systems have been proposed. Among them, the most common are electrooculography (EOG), photo/videooculography (POG), scleral contact lens/search coil and video-based. The EOG eye tracker was one of the first developed and rely on signal differences recorded from contact electrodes placed around the eyes. This technique is able to capture eye movements dynamics and reveal visual processing information [41], however in modern approaches video-based techniques appear to gain ground in the research community.

A. Video-based eye tracking

Video-based eye-gaze tracking is at the cutting edge of passive eye tracking techniques. It captures the eye movements non-intrusively utilizing a video camera alone [42]. Video-based eye tracking systems that use pupil and corneal reflection have been developed on the advances on hardware and software. The most common designs of video-based eye trackers use infrared/near-infrared light that creates corneal reflections [43], although there are also some webcam-based eye trackers which are much less accurate [44]. The eye tracker associates the corneal reflection and the centre of the pupil in order to compute vectors that relate eye position to locations in the perceived world. With appropriate calibration, these eye trackers are capable of measuring a viewer's point of regard on a planar surface on which calibration points are displayed [45]. There is, however, evidence that video-based eye-trackers produce errors in the measurement of small eye movements [46]. The eye tracking setups are available as head-mounted [47], [48], desktop [49], [50] and mobile devices [51], [52]. A typical mobile eye tracker is presented in Fig. 2.

Head-mounted and mobile devices usually include one (monocular) or two (binocular) eye cameras and a scene/world camera. The eye camera monitors the pupil of the eye, while the scene camera captures the user's field of view.

B. Eye gaze estimation

There are various eye gaze estimation algorithms that utilize Near InfraRed (NIR) illumination of the cornea. Among the most common are the 2D regression, 3D model and cross ratio-based methods. The 2D regression-based methods [53], [54] use the difference between the pupil centre and the corneal glint. Through a mapping function and a transformation matrix derived from a calibration of known gaze points, the gaze coordinates can be extracted. Various studies have evaluated the effect of head movements on system accuracy [55], some

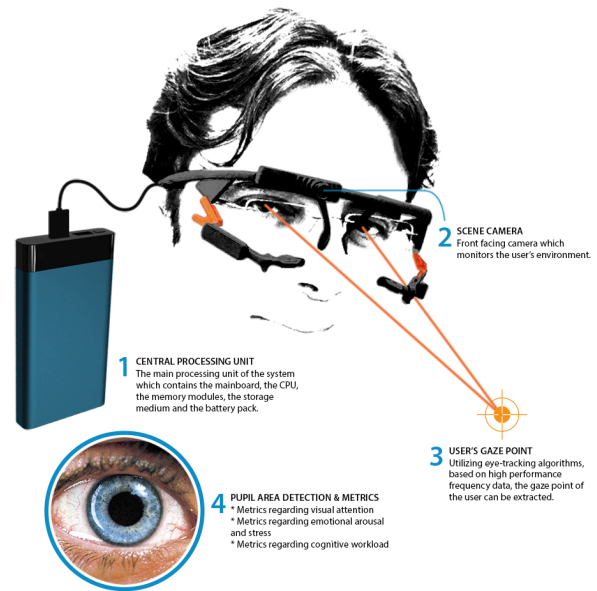


Fig. 2. Example of a mobile eye tracking system

of them using neural networks [56], [57]. The implemented 2D regression methods use various combinations of cameras as well as NIR Light-Emitting Diode (LED). The accuracy of these methods varies according to the setup as well as the movement of the head and ranges from 0.8 to 8 degrees. The 3D model-based methods are divided into two subcategories using one [58], [59] or multiple cameras [60], [61]. These methods use a mathematical model of the human eye to reconstruct the centre of the cornea as well as the optical and visual axes. The accuracy of these methods depends on the number of cameras. Higher accuracy can be produced as compared to the corresponding two dimensional one [62] but requires multiple calibration procedures, including calibration of cameras for 3D measurements, estimation of the position with respect to the cameras and the geometry of the LED. There are however some calibration-free gaze estimation techniques which use either multiple cameras [63], [64] or multiple light sources [65]. Finally, regarding the cross-ratio methods the eye gaze points are calculated by projecting a known pattern of NIR light on a screen (four LEDs on four corners of a computer screen) and then by comparing two perspective projections on the camera [66], [67]. The first projection consists of the virtual images of the corneal reflections of the LEDs (scene plane), while the second projection is the camera projection [62].

C. Estimation of pupil area

Pupil area is an important feature of eye behaviour, thus its estimation is of great significance for many studies. Most of the times, the pupil and iris are darker than the surrounding eye area and therefore thresholds can be applied if the contrast is sufficiently large. Researchers have developed a repetitive threshold algorithm based on a skin-colour model where pupils can be identified by the search of dark areas that satisfy certain anthropometric constraints [68]. Unfortunately, the success

rates of the methods drop abruptly by the presence of dark areas around the eye, such as eyebrows or eyelashes [69]. Another disadvantage of these methods is that they can not model the eyes' closure. In order to overcome this limitation, Tian et al. [70] proposed an eye tracking method that recovers the eye parameters through a dual state model (open/closed eyes). The method requires manual initialization of the eye model and uses a modified Lucas-Kanade tracking algorithm for the tracking of the inner corner of the eye eyelids [71].

Recently, a new method was proposed in [72], which extracts the pupil area through analysis of different eye images intensity levels. In [73], the pupil characteristics are detected using the Cascade Classifier algorithm based on Haar-like features with the help of a histogram equalization method to increase image contrast. Finally, a recent publication presents a method, which initially segments the pupil's region through a convolutional neural network (CNN), and subsequently finds the center of mass of the region [74].

V. EYE AND PUPIL BEHAVIOUR METRICS

In this section, eye and pupil tracking metrics that are correlated with the 3 states investigated (visual attention, emotional arousal/stress and cognitive workload) along with their usage in relevant studies are presented. The most robust of them in terms of their discrimination ability are identified and discussed.

A. Metrics related to visual attention

When a user inspects an image, a video or a real-world scene, several eye inspection patterns and their computational algorithms can provide insights about the user's visual attention and scene perception [75]. Some topics such as center bias, saliency and scan paths are key elements to understand eye movement behaviour related to visual attention.

When viewing an image there is a strong tendency of paying attention more at its center and that is translated to increased fixations at this region. This behavior is known as "center bias" and is well documented [76],[77], [78]. Some attributes or regions of the stimuli are more probable to attract the observer's covert or overt attention, making them salient, such as distinctive color, motion, orientation, or size [79]. Although salience-based schemes are still widely used in computational models, they prove to be poor in predicting eye movement behaviour in natural tasks [80]. Eye movements and especially saccades are being used to investigate several processes, such as visual search. Models of visual search and attention often use scan paths and many studies are trying to quantify scan paths and their nonrandom component [81].

In visual attention studies, fixations form the most used eye movement metric. However, saccades and microsaccades, blinks and pupil size prove also to be helpful. In an attempt to distinguish focal and ambient attention, Krejtz et al. have introduced coefficient K as a metric that takes into account the relationship of fixation duration and its subsequent saccade amplitude. Positive values on the ordinate of coefficient K indicate focal patterns, while negative values suggest ambient visual scanning [82].

A plethora of stimuli and tasks have been used in order to study visual attention. In this section, we will follow a general categorisation of the stimuli (dynamic vs static) and we will present first the studies that used static and then those that used dynamic stimuli.

1) *Fixations*: In a study which compared visual attention in different kinds of videos and static pictures, fixations when viewing stop-motion movies and Hollywood trailers were longer than when viewing natural movies, and the shortest fixations occurred on static images. An explanation for this is that the abrupt image change in movies captures most of the visual attention of the viewer as compared to a simple static image [78].

In [83], participants had to decide on whether or not a specific person (target) appeared on an image. In doing so, they fixated less in the person presented images than in the person absent ones. Moreover, the duration of exploratory fixations (i.e., the fixations until the person was spotted or not) was larger when the person was in the images. On average, the observers spent 428 ms fixating on the target before responding. In [84], it was found that when observing illustrations, most of the fixations appear in saliency regions. When there was only one saliency object in the presented images, saliency rate of the first fixation (SRF) amounted up to 93.8%. When two or more saliency objects are present in a scene, it was found that the observer could not decide at which to focus first resulting in the SRF decrease. It is also reported that SRF and the saliency rate of the longest fixation (SRL), independent of the number of saliency objects present in a scene, were higher than the saliency rate. This may mean that people pay attention firstly on regions of interest which are likely more salient than other areas.

During image viewing, observers tended to pay more attention to emotionally salient regions than on visual ones [7]. In [85], an experiment was conducted with random participants and orthodontists staring at images containing smiling faces. It was reported that people with no expertise in dentistry or orthodontic procedures exhibited larger fixation duration in the eyes than the nose or mouth. In contrast, the orthodontist group spent significantly more time looking at the eyes and mouth than the nose of facial images, determining that past experiences, as well as their education and work background, play a vital role in the visual attention motif.

In [86] sensual photographs of heterosexual couples depicting attractive women and men in sensual situations were presented to heterosexual women and men. Participants, irrespective of their gender, looked longer at the body (vs face) of stimuli. This result suggests that, in tasks related to sexual desire, the locus of spontaneous visual attention is preferentially directed toward the body. Moreover, both men and women looked longer at the bodies of women than of men, suggesting that automatic visual attention associated with sexual desire is prominently oriented toward women's bodies, irrespective of gender.

In a study where images of paintings and faces were presented and participants were asked to grade how much they liked each picture, total fixation duration and number of fixations showed a strong positive correlation with liking ratings,

while no significant difference was found in mean fixation duration. The liking of paintings guided visual attention to the same extent as the liking of faces [87].

In a web task, participants fixated more on the text than on the illustration area. However, using successive web pages with less information seemed to promote more eye fixations on the illustration area [88]. In a similar study, participants made longer fixations during their navigation in the first compared to the second page of news and shopping websites. However, evidence showed that this behaviour was altered while interacting with business websites, in which the duration of fixations remained unchanged irrespective of the number of pages [89]. Furthermore, fixation duration of participants belonging to the millennial generation was shorter than the older participants in a web searching task, meaning that younger people needed less time to process the cognitive load of the image [90]. In the same study, when the main picture of the site was placed at the top of the web page, it received a larger number of fixations. Lastly, according to [91] and an experiment asking participants to browse through a website of products, the number of revisits revealed a tendency for the participants regardless of gender to return to a higher user experience more times, but without a main effect of gender, or interaction between gender and user experience. For fixation count and total fixation duration, there was a significant main effect of user experience, but without the main effect of gender, or interaction between gender and user experience.

Memory is associated with visual attention according to [92]. There is a firm correlation between fixation number and recall and this relationship is stronger for those with lower usage levels of a particular commercial brand when watching commercial videos.

Visual attention has also been linked with pain. In [93] participants were presented pictures showing a natural scene while painful electrical stimuli were applied to their left or right hand. Painful stimulation caused fewer and longer fixations. Moreover, painful stimulation on the right hand induced a rightward bias, i.e. increased initial saccades, total number and duration of fixations to the right hemifield of the screen, while pain applied to the left hand as well as no pain induced a leftward bias that was largest for the direction of first saccades.

During a driving task in [94] participants made two consecutive trips during one of which, they received a phone call on a hands-free device in the vehicle. During the phone call, road signs, other vehicles, and the speedometer were fixated less while no significant differences were observed in fixation duration.

There is an extensive study on visual attention on children. Language and visual cue conditions play a vital role in children's gaze pattern according to [95]. Specifically, both conditions were responsible for the observed increased number of fixations on target landmarks and switches between target objects. Also, it is reported that infants tend to fixate on their mothers before the latter talked to them, they are more likely to look at mothers' hands when mothers were holding objects, and they fixated at their mothers more quickly when mothers were already present in their field of view [96].

2) *Saccades, microsaccades and smooth pursuit*: Viewers looking at printed advertisements made longer saccades on the picture part of the ad compared to the text [97]. Painful stimulation while looking at images of natural scenes caused less and slower saccades suggesting reduced exploratory behavior [92].

It has been reported that when viewing natural movies, observers tend to make both more small and more large saccades (with amplitudes of less than 5 and more than 10 degrees, respectively) on natural movies, whereas saccades of intermediate amplitudes are less frequent than in Hollywood action movie trailers and stop-motion movies [78]. In contrast, saccades on Hollywood trailers show the smallest fraction of large amplitudes. During driving task in [94], saccade duration increased during the hands-free trip compared to the control trip, suggesting longer saccade lengths and thus a more disperse fixation pattern during handsfree phoning.

Microsaccades are widely studied in relation with covert attention and there is evidence of the effect of attentional cue presentation on rate and direction of microsaccades [26], [98], [99]. Microsaccades are shown to follow spatial attention during the cue-target interval in a high degree [100]. Mayberg et al. have proposed that microsaccades are related to both – the overt attentional selection of the task-relevant part of the cue stimulus and the subsequent covert attention shift [101]. According to [102] visual attention provides necessary control mechanisms for triggering smooth pursuit eye movements. After its onset, smooth pursuit also requires non-visual cognitive attention controls in order to achieve and maintain a high accuracy of eye tracking. In a clinical study, the relationship between smooth pursuit eye movements and visual attention in patients with schizophrenia and normal controls was evaluated. The magnitude of the correlation between smooth pursuit gain and visual attention measures was statistically compared to the magnitude of the correlation between smooth pursuit gain and motion perception threshold in the controls [103].

3) *Blinks*: As stated in [39], blink rate can provide useful information about the tendency of the viewer to pay more attention to a specific location in a picture, as the decrease in blink frequency signs the tendency in focusing to this exact location, meaning that the viewer attempts to keep his/her eyes open for a larger time period to observe the desired location.

4) *Pupil*: It should be noted that attention during work is sometimes modulated by the level of expertise. In [104], the percent of change in the pupil size of experienced workers is slightly lower than the novice workers. Expert's eye movements may clearly exhibit a systematic pattern. Moreover, regarding the pupil diameter in [91], in parallel with the fixation count and total fixation duration there was also the important influence of user experience, independently of gender, or interaction between gender and user experience.

As shown on Table I, the metrics that are mostly involved in increased visual attention are the fixation number and duration as well as total fixation time. The reasoning behind this finding lies in the fact that when we focus our attention on a specific object or person, we maintain our gaze more time on a specific AOI, thus more fixations and larger fixation duration on this AOI. Moreover, saccade amplitude, blink rate

TABLE I
SUMMARY OF EYE-RELATED MEASUREMENTS AND THEIR RELATIONSHIP
TO VISUAL ATTENTION

Eye feature	Change	References
Fixation duration	↑↑ ns	[78], [85], [89], [90], [93] [87], [94]
Number of fixations	↑↑	[83], [84], [87], [88] [90], [105], [92], [93], [94] [95]
Total fixation time	↑↑	[83], [85], [86], [87], [105]
Saccade amplitude	↑↑	[78], [94]
Saccade velocity	↑	[92]
Number of saccades	↑↑	[92]
Microsaccade rate	↑	[26], [98], [99]
Smooth pursuit gain	↑	[102], [103]
Blink rate	↓	[39]
Pupil size	↑	[91], [104]

↑↑ / ↓↓ : significant increase/decrease during visual attention at 0.01 level
↑ / ↓ : significant increase/decrease during visual attention at 0.05 level
ns: non-significant difference

and microsaccade rate seem to be useful metrics. The rest of the metrics that are presented are not widely studied, thus no specific conclusion can be drawn.

B. Metrics related to emotional arousal and stress

Emotional arousal and stress are known to affect eye function and behaviour [106]. In this section, their manifestations in eye metrics (such as fixations, saccades, blinks, pupil, etc) are presented.

1) *Fixations*: Researchers employed videos conveying both positive and negative emotions and report that the number and duration of fixations were significantly different between the two emotional states studied [107]. Concerning stress conditions, socially anxious participants did not exhibit initial orienting bias and they had a greater probability to fixate on angry faces (having also greater fixation duration) as compared to non-anxious participants [108]. On the other hand, non-anxious participants showed a higher probability of fixation at happy faces, during the two-second time-interval after stimulus onset [108]. Distribution of gaze points is also considered to be affected by arousal and stress [109], [110]. Gaze features such as gaze direction, gaze congruence and the size of the gaze-cuing effect have been employed in arousal studies [111], [112]. Besides, fixation instability has been associated with trait anxiety in both volitional and stimulus-driven conditions, but it is more pronounced in the presence of threat [110]. People with increased anxiety tend to have their first fixation on the emotional picture, contrary to the neutral one [113].

2) *Saccades*: Arousal has also been associated with saccadic duration, acceleration, and velocity. Saccadic velocity has been considered as an index of arousal/cognitive demand increasing in high arousal states [114]. Involuntary saccades were significantly increased under conditions of arousal with a specific time course for the increase in involuntary movements [115]. Specific inhibitory deficits related to arousal can be revealed through the antisaccade task, where saccadic control is disrupted [116].

3) *Blinks*: The frequency of spontaneous eye blinks increases during stress or other states of emotional arousal [117].

This can be partially attributed to the redirection of blood periorbital eye musculature facilitating rapid eye movements [118]. However, eye blinks decrease during tasks that demand to pay more attention (e.g. reading a difficult text) [117]. According to [119], there is a significant correlation between eye blink frequency and stress level. Artificially triggering of emotional responses by billboards, and more natural emotional responses by simulating car crashes caused a temporary increase in the eye blink frequency.

4) *Pupil*: Pupil size reflects autonomic involuntary activity and it is associated with emotional, cognitive or sexual arousal. Various studies indicate the relation between pupil size variation and affective states [120], [121], [30], [122], [32]. The variation of pupil size has been employed efficiently as an index of stress and arousal [123], [124], [125],[126]. Pupil diameter increases during stress elicited by stressful stimuli in laboratory environment [122], [32], [127], [115].

The pupillary response appears to be greater when the visual stimuli presented are images conveying negative valence information and it tends to be higher among persons reporting higher overall levels of stress [128]. Pupil size may also increase in response to positive, and negative arousing sounds, as compared to emotionally neutral ones [121], [129]. Audience anxiety has been demonstrated to affect the pupil size [123]. The higher the anxiety level, the bigger the pupil becomes, presenting significant differences in expressions of contempting and surprise [130]. Other studies refer that under arousal conditions the pupil dilation increases [109]. In [126], pupil size was positively correlated with HR or GSR and their causal interaction during emotional processing was investigated. An interesting approach is the translation of pupil behaviour in arousal observed due to drug abuse. The pupil dilation among a cocaine-induced paranoia (CIP) group was significantly greater in response to a video image of crack cocaine than a non-CIP group [131] which can be attributed either to the recall of an event of cocaine intake or to the trait anxiety it is caused by it.

Available research data suggests that emotional arousal is a key element in modulating the pupil's response. For instance, as early as 1960 authors [132] reported bi-directional effects of emotion on pupil change, specifically reporting that the pupil constricted when people viewed unpleasant pictures and dilated when they viewed pleasant pictures. Similar results were also reported in other studies [133]. In a recent study, high trait anxiety group's pupil size was increased during tasks with facial expression processing (mainly expressions of condemn and surprise) in relation to low trait anxiety [130].

Table II summarizes the literature findings on the changes in eye and pupil metrics while subjects were under emotional arousal or stress state. As it can be observed, pupil size metric appears as the most common indicator for detecting emotional arousal as its increase clearly reflects the emotional charge in different scenarios. It is important to mention though that the detection or even more the quantification of emotional arousal only through measuring the pupil diameter is questionable. Besides, the limitations referred in Section II should make the researchers cautious on the proper usage of pupil metrics. The majority of the studies stated in Section

TABLE II
SUMMARY OF EYE-RELATED MEASUREMENTS AND THEIR RELATIONSHIP
TO INCREASED EMOTIONAL AROUSAL

Eye feature	Change	References
Fixation duration	↑↑ ns	[108] [107], [113]
Number of fixations	ns	[107]
First fixation probability	↑↑	[30]
Last fixation probability	ns	[113]
Time to first fixation	↑ ns	[107] [128]
Saccadic velocity	↑	[115]
Blink rate	↑↑ ns	[117], [119] [109]
Blink duration	ns	[109]
Pupil size	↑↑	[128], [30], [121], [131], [129], [127], [134], [115]
	↑	[135], [130], [123], [136]
	ns	[107], [109]
Gaze distribution	ns	[109]
Dwell time	ns	[128]

↑ / ↓ : significant increase/decrease of emotional arousal at 0.05 level
 ↑↑ / ↓↓ : significant increase/decrease of emotional arousal at 0.01 level
 ns: non-significant difference

V-B use other physiological signals outside of pupil size in order to determine among emotional arousal states, thus questioning the feasibility of solely using pupil size as an emotions indicator. Furthermore, except pupil dilation, blink duration seems to increase with the increase of arousal level. Besides, an interesting finding is the fact that people with trait anxiety tend to have greater duration of the initial fixation and the duration of the fixations were longer on threat images (conveying negative valence) in relation to neutral images [137]. The other eye metrics demonstrated in Table II do not have a consistent correlation on the arousal.

C. Metrics related to cognitive workload

The study of mental workload, also known as cognitive workload, is a vital aspect in the areas of psychology, ergonomics, and human factors for understanding performance. [138]. Despite the multitudinous and extended research in this area, there is no single definition to describe cognitive workload. Often we refer to cognitive workload as taskload i.e. the effort needed to perform a certain procedure. However, defining workload can be a rather subjective task depending on how different people with different experience and abilities can handle the same task [139], [140]. So, a general definition of mental workload would be the product of factors that contribute to one's workload efficiency for a given task.

Due to the majority of definitions for cognitive workload there is a plethora of ways to measure it. No sensor can give a complete picture of how someone reacts to a task. Therefore, the estimation of multimodal biosensors can assist in the determination of workload levels. In the next paragraphs, we will present the most popular and robust biomarkers that are utilized in estimating cognitive workload. These include measurements through fixations, eye movements, blinks and pupil size. Blink rate, pupil diameter, blink duration and fixation duration seem to be the most frequently used eye-related measures [141].

1) *Fixations*: There is ample evidence that the number and duration of fixations can be an indicator of cognitive effort, especially when associated with the level of experience.

It has been shown that mean fixation duration has a significant negative correlation with the level of cognitive load in simulated flight [142], [143] and driving tasks [144], in video gaming [145] and in arithmetic task [146]. Max fixation duration also shows the same behaviour [146]. Fixation duration failed to reach significant difference with increasing cognitive load in a reading task [147]. In [148], fixation duration was found to be most sensitive to one of the three types of cognitive load; extraneous load.

Number of fixations has been shown to increase with increasing cognitive load in a surgical procedure [149], in chess playing [150], in video gaming [145], in simulated flight task [143] and in a complex task on a website [151].

As said before, the relationship between fixation parameters and cognitive workload in association with the level of expertise has been studied largely. Higher fixation duration has been found in novices compared to experts in several working tasks; surgical environment [149] and chess playing [150], while no significant difference was found between experts and novices map users [152]. No difference was found either in fixation duration between novices and experts users of a training platform in low cognitive load task. There was however increased fixation duration in experts in high cognitive load task [153]. As far as the fixation rate is concerned, novices performed more fixations than experts in a surgical environment [149], but fewer fixations in map using task [152].

2) *Saccades, microsaccades and smooth pursuit*: Although saccades form the type of eye movements that are most evaluated in cognitive workload studies, smooth pursuit eye movements and microsaccades are also studied and they seem to correlate well with cognitive load [154], [155].

The average peak saccadic velocity is found to have a positive relationship with the increasing workload [156] and more specifically, it has been shown to be more sensitive to germane load [148]. The significance of saccadic velocity in determining the amount of cognitive load is demonstrated also through video game scenarios. In [145], the saccadic peak velocity decreased while the speed of the game slowed down and increased rapidly when the game speed raised in proportion to the increase of the difficulty level.

Average and maximum saccadic amplitude also show a moderate positive correlation with difficulty level [146]. On the other hand, [152] have shown that mean saccade amplitude decreases with increasing cognitive load, both in experts and novices map users.

During the performance of driving tasks, the saccade velocity and saccade frequency increased when time pressure got higher, and decreased when the subject was overloaded. The maximum workload was firmly correlated with the maxima of average saccadic velocity and saccade frequency [157]. When performing a secondary task during a driving scenario, a significant increase in drivers' saccade rate was observed as the task difficulty level increased [158].

Saccades can be complementary to fixations (referred in the previous section) and can also indicate the skill level of

the specified clinician as novice surgeons make more saccadic movements compared to the intermediate surgeons [149]. In another study, novice map users made saccades of smaller amplitude than experts [152]. In a low cognitive load task in which participants had to operate a training version of a military land platform, difference in saccade amplitude between novices and experts failed to reach significance [153].

Kosch et al. have shown that certain trajectories (fast and circular) cause increased gaze differences of smooth pursuit eye movements during the presence of cognitive workload [155]. In another study, eye-target synchronization during smooth pursuit eye movement improved under intermediate cognitive load in young normal subjects [159].

Microsaccades are often used to study cognitive processes. Microsaccade rate in mental arithmetic tasks while fixating a central target is found to decrease with increasing task difficulty [160], [161]. In a more recent study, microsaccade rate failed to reach significant change with increasing cognitive load [154]. Microsaccade amplitude seems to increase with increasing task difficulty [160], [154], while the behaviour of microsaccade peak velocity and peak amplitude is not clear [154].

TABLE III

SUMMARY OF EYE-RELATED MEASUREMENTS AND THEIR RELATIONSHIP TO INCREASED COGNITIVE WORKLOAD

Eye feature	Change	References
Fixation duration	↑↑	[150]
	↑	[149]
	ns	[147], [152], [151]
	↓	[144], [146], [153]
	↓↓	[142], [143], [145]
Max fixation duration	↓↓	[146]
Number of fixations/ Fixation frequency	↑↑	[143], [145], [150]
	↑	[151], [149]
Saccadic velocity	↑↑	[145], [143]
	↑	[156], [143]
Saccade amplitude	↑	[146]
	ns	[153]
	↓	[152]
Max saccade duration	↑	[146]
Saccade rate / Number of saccades	↑↑	[158], [162], [149]
Microsaccade rate	↓↓	[160], [161]
	ns	[154]
Microsaccade amplitude	↑↑	[160], [154]
Blink rate	↑	[163]
	↓	[17]
	↓↓	[145], [164], [165]
Blink duration	↓↓	[145]
	↓	[165]
Pupil size	↑↑	[145], [166], [167], [168], [169], [17], [170], [154], [143], [165]
	↑	[163], [171], [132], [172]
	ns	[147]
Pupil size std	↑	[173]

↑ / ↓ : significant increase/decrease during cognitive workload at 0.05 level

↑↑ / ↓↓ : significant increase/decrease during cognitive workload at 0.01 level

ns: non-significant difference

3) *Blinks*: Blink duration is a sensitive indicator of cognitive workload, given the significantly higher presence of short blinks under high visual load conditions [165]. Moreover, a decreasing blink duration tendency as the task became more difficult was observed. In [145], blink duration and frequency

were also shown to maximize at the lowest speed in the video game, and as the workload increased, blink frequency decreased. In [17] blink rate decreased from low to medium cognitive load but no further change was observed in high cognitive load. Borys et al. also showed that maximum blink duration correlated strongly with the number of errors in high cognitive load of an arithmetic task [146].

In another study, the blink rate was shown to drop to low levels compared to a resting-state rate and be sensitive to the phases of microsurgical suture [164]. In a driving scenario in [163], the eye tracker showed increasing blink rates along with the parallel increase in the difficulty of a -parallel with driving- performed secondary task.

4) *Pupil*: The pupil diameter changes have great subject variability leading to the result that the pupil size changed irregularly during various conditions along with the fact that the value of peak points was different under different difficulty tasks [174]. Another similar study in [168], proposed that pupil diameter distinguished differences in workload between task difficulty levels and increased as the task became more demanding.

Pupil area is strongly associated with the user's ongoing task difficulty [166], [132], [172] and mean pupil diameter has been shown to have a positive correlation with the cognitive workload in several different tasks [17], [170], [154]. The standard deviation of the pupil size also increases with cognitive load [173]. In a task where the participants had to watch a multimedia lesson, Zu et al. trying to determine what type of cognitive load affects pupil size, concluded that mean ratio of pupil size change was most sensitive to extraneous and germane load [148]. In accordance with this study, another experimental study showed that pupil size increased when time pressure got higher [143]. This study also showed that pupil size decreased when the subject was overload. Pupil size reached to the maximum where the load was maximum.

In addition, pupil diameter increases analogous to the difficulty of the ongoing secondary task performed by the user [163]. During mentally demanding tasks pupil size appears to increase proportionally to the effort that someone has to make [171]. Moreover, pupil size showed the strongest relationship with changes in workload during the Tetris game [145] as it was positively correlated with the mental workload. In the arithmetic task of [146], max pupil dilation correlated strongly with the number of errors in high cognitive load. The increase of the mental workload in [169] caused the pupils to dilate and as the participants were close to overload the saccade rate increased too. Lastly, according to [147], the topic and difficulty level of a text did not significantly influence pupil size measures.

Average pupil size clearly reflected the effort of multitasking while driving as well as it exhibited the importance of habituation effect. After the IVIS (in-vehicle information systems) task was performed in easy and difficult situations, it was observed that average pupil size was significantly higher each time an IVIS condition was performed for the first time [165]. The pupil size of novice surgeons was larger than that of the intermediate surgeons while performing a surgical task as the inexperience led to increased mental and physical effort [149].

An overview of eye features correlated with cognition difficulty based on our literature review is demonstrated in Table III. As it can be observed, unlike to the case of emotional arousal, there is a greater variety of eye behaviour features that can witness an individual's cognitive effort. Firstly, the fixations number increases when the workload is higher while the fixation duration seems to decrease. Saccade rate, saccade velocity and microsaccade amplitude increase with the increasing workload and so is the pupil size which proves to be a useful indicator of cognitive load conditions. Blink rate, microsaccade rate as well as blink duration seem to decrease with cognitive load.

VI. MULTIMODAL EMOTIONAL AND COGNITIVE PROCESSES RECOGNITION METHODS

Emotional arousal and cognitive workload detection are two important aspects of human behaviour. Automatic recognition of these processes and their states through computerized techniques could possibly enhance the computers' ability to respond and act more intelligently and drive the user away from negative emotional states or possible health issues without demanding high levels of cognitive effort [175],[176], [177].

Physiological signals are the most commonly used modality combined with eye-related metrics in order to estimate one's emotional and cognitive load level [178]. Changes in physiological arousal during stressful conditions are quantifiable through skin conductance, thermal camera recordings, respiration and breath rate, heart activity, electrodermal activity (EDA), electroencephalogram (EEG), electromyogram (EMG), pupil dilation (PD), photoplethysmography (PPG) and body posture/movements [13], [179], [180], [181]. The cognitive workload is associated to a lesser extend with biosignals than in arousal state. Cardiac measurements [182] as well as blood pressure metrics [183] have shown evidence of quantifying the mental workload. These success rates increase by adding to the above metrics various eye movement data with the PD being the most commonly used eye movement metric followed by blink rate and fixation metrics.

In the following subsections we present studies that make use of eye features and biosignals in order to decide upon the levels of emotional arousal and cognitive workload using machine learning techniques.

A. Emotional arousal recognition

In Table IV, studies combining the above characteristics are presented in order to identify and quantify the emotional arousal and stress level. We can separate the studies presented in Table IV into the ones which make use only of eye features and into the ones which utilize various physiological signals for the recognition of the states under investigation. The nature of the tasks performed in the majority of the studies mainly concerns the viewing of emotion evoking images and videos. The researchers in all of the studies but one, make use of pupil size as a feature to their classifiers, thus confirming that pupil diameter is the most important eye metric for emotion recognition with blink related metrics coming at second place. The rest of the eye features such as fixations and saccades

metrics as observed are complementary to pupil, blinks and the biosignals feature set in order to ensure the efficient discrimination among the emotional states.

For the studies relying solely on eye features (5 in total) the average best accuracy received is 71.5% with the highest being 93%. In order to interpret those results we must first notice that in [184], [185], the researchers attempt to solve multiclass classification problems (4 emotions and levels of arousal) while using one -but the most important- eye feature, the pupil diameter. However, the highest accuracy received lies below 58% for [184] and 54% for [185] respectively. In [186], a classification between the emotional arousal and valence is attempted while exploiting various eye features correlated with fixations, saccades as well as pupil size resulting to the relatively high accuracy of 80.00% taking into account the three class classification problem. The other two studies concern the binary depressed/non-depressed classification problem. In [187], researchers combine pupil diameter with blink rate and percentage of eyelid closure to reach a 75% accuracy while in [188], they reach the high accuracy of 93% with only exploiting blink related metrics pointing out that it is possible to decide upon the existence of depression with high precision.

For the rest of the studies (9 in total) the average accuracy percentage lies at 81.76% and the highest at 90.1%. This approximately 10% difference in average accuracy in contrast to the studies using only eye-related features is due to the use of biosignals and especially the EEG which is proven to be a very useful emotional indicator [13]. The classification problem concerning these studies is shared among multiclass between basic emotions and binary between stressed and non-stressed classes. As it is expected, the binary stressed/non-stressed classification problem studies resulted in higher accuracy near and above 80.00% reaching up to 90.1%. However, the studies aiming to the quantification of basic emotions resulted in relatively high accuracy too with percentages near 80%.

By observing Table IV once again, the SVM classification system is either the one used in the majority of the studies or the one that helps researchers reach the best emotion prediction accuracy as we can see in the respective column. This is possibly due to the various advantages of the SVM classifiers especially when it comes to relatively small datasets and the clear margin of separation between classes [189].

In this subsection, we commented on the various research attempts to quantify emotional arousal and we compared them with respect to the features they used and the accuracy percentages they achieved. Conclusively, we notice that eye features and especially pupil diameter and blink related metrics play a vital role in the estimation of emotions and emotional arousal. However, although eye features help significantly decide in binary classification problems (i.e. stressed/non-stressed), they cannot by themselves help discriminating well among multiclass classification problems (low/medium/high arousal). Therefore, the combination of eye metrics along with biosignals is necessary in order to achieve higher and reliable prediction rates as we observe from the relevant studies at Table IV.

TABLE IV
MACHINE LEARNING TECHNIQUES USED FOR EMOTIONAL AROUSAL AND COGNITIVE WORKLOAD IDENTIFICATION

Study	Population (subjects)	Women/Men	Age	Stimuli or Task	Feature	Classes	Classification system	Best accuracy received (best classifier)
Zheng et al. (2020) [184]	10	0/10	21-28	360° videos presented in VR	PD	angry/happy/sad/relaxed	SVM	57.65%
Tarnowski et al. (2020) [186]	30	0/30	21.25 ±0.74	Emotion evoking video fragments	NoF, FD, SA, SD, PD	high/moderate arousal, valence	SVM	80.00%
Ahmad et al. (2020) [17]	41	20/21	18-37	word game	HR, HRV, PD, BR	low/medium/high workload	RF, NB	91.66% (RF)
Prabhakar et al. (2020) [173]	21	5/16	mean 26	simulated and realtime driving	SR, FF, SV, PD	low/high workload	NN	75.00%
Zhao et al. (2019) [190]	16	10/6	-	Emotion clips	PD, fixation details, saccade details, blink details, EEG	happy/sad/fear/disgust/neutral	SVM	79.71%
Guo et al. (2019) [191]	-	-	-	Images from SEED-V Database	PD, EEG	happy/sad/fear/disgust/neutral	SVM	79.63%
Li et al. (2019) [192]	16	10/6	19-28	Emotion clips	PD, FD, BD, SD, EEG	happy/sad/fear/disgust/neutral	SVM	70.98%
Al-Gawwam and Benaissa (2019) [188]	34	-	18-63	Reading, answering questions	BR, BD, BA	depressed/non-depressed	SVM, AB, NB, BT	93.00% (SVM)
Wu et al. (2019) [168]	8	4/4	26.0 ±1.6	simulated surgical exercises	PD, FD, Gaze entropy, PERCLOS	low/high workload	NB	84.70%
Bozkir et al (2019) [193]	16	-	-	simulated driving tasks	PD, driving performance measures	low/high workload	SVM, DT, RF, KNN	80.70% (SVM)
Chen et al. (2019) [194]	8	-	22-25	simulated flight task	NoF, FD, SD, NoS, saccades, PERCLOS, PD	low/medium/high workload	KNN+SVM, SVN	43.80% (KNN+SVM)
Kosch et al. (2018) [155]	20	9/11	22-34	n-back test	SPEM	low/high workload	SVM	99.50%
Baltaci et al. (2016) [124]	11	2/9	33.0 ±3.46	IAPS pictures	PD, facial temperature	stressed/non-stressed	DT, AB, RF	83.80% (RF)
Liu et al. (2016) [195]	30	15/15	mean 21.4	judgement trials	PD, FF, BF, Fix PD, SV, SD, SPV, SD, SPV	low/high workload	SVM	90.25%
Aracena et al. (2015) [185]	4	0/4	19-27	Image (from IAPS)	PD	low/medium/high arousal	DT	53.60%
Zheng et al. (2014) [196]	5	2/3	22-24	Emotional film clips	PD, EEG	positive/neutral/negative valence	SVM	73.59%
Pedrotti et al. (2014) [32]	33	17/16	41.0 ±11.3	Lane Change Test	PD, EEG, EDA, SKT, Facial cues	stressed/non-stressed	NN, GA, SVM, ANN (Four-Way parallel class.)	79.20%
Ren et al. (2013) [122]	30	16/14	26.8 ±2.56	Stroop CWT	PD, GSR	stressed/non-stressed	NB	85.53%
Algwinem et al. (2013) [187]	60	-	-	Emotion evoking images	PD, BR, PERCLOS	depressed/non-depressed	GMM, SVM	75.00% (SVM)
Zhai et al. (2006) [125]	32	-	21-42	Paced Stroop Test	PD, GSR, BVP, SKTD	stressed/non-stressed	SVM	90.10%
Liao et al. (2005) [109]	5	-	-	Tasks on computer screen	BR, AECS, Saccades, Gaze Distribution, PD, PRV, Head Movement, Mouth Openness, Eyebrow movement	stressed/non-stressed	DBN	86.20%

PD: Pupil Diameter, PRV: Pupil Ratio Variation, PERCLOS: Percentage of eyelid closure, SPEM: Smooth Pursuit Eye Movements, EDA: electrodermal activity, EEG: electroencephalogram, SKT: skin temperature, GSR: Galvanic Skin Response, BVP: Blood Volume Pulse, AECS: Average Eye Closure Speed, PERCLOS: Percentage of eyelid closure, BR: Blink rate, BRV: Blink rate variability, BD: Blink Duration, BA: Blink Amplitude, NoF: Number of fixations, FF: Fixation Frequency, FD: Fixation Duration, SA: Saccade Amplitude, SD: Saccade Duration, NoS: Number of Saccades, SV: Saccade Velocity, SPV: Saccade Peak Velocity, HR: Heart Rate, HRV: Heart Rate Variability
Stroop CWT: Stroop Colour-Word Test, SVM: Support Vector Machines, GA: Genetic Algorithms, RF: Random forests, AB: AdaBoost, GMM: Gaussian Mixture Models, BT: Bagging Trees, DT: Decision Tree

B. Cognitive workload recognition

In Table IV and similar to the above subsection, we present studies with relation to the classification of cognitive load levels. We can again dissociate the studies into the ones which exploit only eye features and into the ones which utilize various physiological signals. The nature of the task varies among the studies, but we can identify three basic patterns: simulated real-life tasks, n-back test and reading/recalling tasks. Regarding the classification problem, all of the studies aim to solve either a binary class problem (low/high) or a three class problem (low/medium/high) with respect to the level of cognition difficulty. The researchers in all of the studies except two, make use of pupil size as a feature to their classifiers, thus confirming that pupil diameter is very vital for the estimation of the cognitive workload with blink related metrics following right after. Aside from smooth pursuit which is used only in [155], the rest of eye features are complementary to pupil, blinks and the biosignals feature set in order to ensure the efficient discrimination among the cognitive states. For the studies relying solely on eye features (5 in total) the average best accuracy received is 78.64% with the highest being 99.5%. In four of these five studies, the researchers aim to binary classify between low and high cognitive load with high success rates. However, in [194], the scientists aim to identify three levels of cognitive load (low/medium/high) but with low accuracy, that of 43.8%. In contrast and in [17], researchers while trying to identify between the same three categories, exploited heart rate related metrics combined with pupil diameter and blink rate and thus managed to achieve over double the accuracy percentage with 91.66%. The other study [193], uses drive measurement features along with pupil size and reaches up to 80.7% accuracy percentage for the binary classification of the level of cognitive load. So, the average accuracy percentage of the studies utilizing as features for the classification a combination of pupil metrics and biosignals is 87.55%.

By looking once again at Table IV, we can see that -likewise the previous subsection- four out of seven concentrated studies achieve their best accuracy percentages based on SVM classification system most probably for the reasons explained at the emotional arousal recognition subsection.

In this subsection, we summarized the various research attempts to classify the level of cognitive load and we compared them with respect to the features they used and the accuracy percentages they achieved. Finally, we conclude that eye features and especially pupil diameter related metrics are significant in order to reliably measure the cognitive effort especially when seeking to solve binary classification problems and determine the existence of workload. However, biosignals are required to the feature equation in order to decide upon more than two levels of cognitive workload and reach greater success rates.

VII. EYE BEHAVIOUR RELATED DATASETS

In this section, we present and discuss the merits and shortcomings of the various datasets related to emotional arousal, cognitive workload and visual attention that contain

eye-tracking features and are publicly available. The relevant data sets identified are listed in Table V, which also presents the number of subjects and their age in each dataset, the stimuli used for invoking the relevant cognitive or affective state and the eye-related features available in each dataset. It should be noted that the various low level eye-related features shown in Table V are those that are provided by the dataset creators.

Regarding the stimuli used, we observe that stimuli used for the creation of the datasets focusing on visual attention were video or image presentations or tasks in which focusing on a specific target was required. For the generation of the emotional arousal datasets, the stimuli used were emotional eliciting images and video clips, whilst for the creation of datasets related to the cognitive workload, the subjects were performing everyday activities or other tasks requiring increased mental effort.

As observed in Table V, the basic eye feature included in the vast majority of the datasets is the eyes' 2D gaze coordinates. These coordinates enable the estimation of fixations and saccades related characteristics using a variety of fixation detection algorithms that are based either on velocity, i.e. Velocity-Threshold identification, Hidden Markov model identification, dispersion, i.e. Dispersion-Threshold identification and Minimum Spanning Trees identification, or the area of interest, i.e. Area of Interest identification [210].

However, as evidenced from the analysis and summary provided in Section V, pupil and blink related features play a vital role in classifying the levels of emotional arousal and in distinguishing the various levels of cognitive load. Unfortunately, only two of the datasets shown in Table V contain information of the pupil diameter, thus making it difficult to accurately estimate the level of emotional arousal or cognitive workload. Finally, only one of the datasets contains blinks related features, which are known to be highly correlated with emotional and cognitive processes.

On the other hand, the Eye Tracking Movies Database (ETMD) [198] provides the research community with a significant number of eye metrics, i.e. 2D gaze points, fixation points, and pupil position and diameter. In addition, as reported in [211], the movie clips used as stimuli were carefully selected in order to evoke different levels of emotional arousal based on the six basic emotions, i.e. happiness, sadness, disgust, fear, surprise, and anger.

Regarding cognitive workload, the EGTEA Gaze+ dataset provides researchers with a relatively large number of subject data and eye features that include 2D gaze points, fixations, saccades, blinks and pupil diameter. However, the EGTEA Gaze+ dataset was originally developed in an effort to study attention and action in first-person vision (FPV), by generating a dataset of meal preparation tasks captured in a naturalistic kitchen environment as reported in [69].

On the other hand the MAMEM Phase I dataset [199] is a dataset that combines multimodal biosignals and eye tracking information gathered under a human-computer interaction framework. The dataset was developed in the vein of the MAMEM project that focused to endow people with motor disabilities with the ability to edit and author multimedia content through mental commands and gaze activity. The

TABLE V
PUBLICLY AVAILABLE DATASETS WITH EYE BEHAVIOUR FEATURES DURING EMOTIONAL AND COGNITIVE PROCESSES

Dataset	Subjects	Age	Gender (f/m)	Stimuli	Nature of task	Eye features	Type of Affect or Cognition
EMOotional attention dataset (EMOd) [197]	16	21-35	-	Emotion-evoking photos selected from IAPS database	Image viewing	Fixation points, fixation duration, fixation maps	Emotional arousal
ETMD [198]	10	-	-	Movie clips from GOG-NIMUS database	Movie watching	2D gaze points {x,y}, Fixation points, pupil/glints position pupil diameter	Emotional arousal
MAMEM Phase I [199]	36	25-73	9/27	Web pages, multimedia content	Computer usage, imaginary movement task	2D gaze points {x,y}	Cognitive workload
ErrP [200]	10	20-45	4/6	Sentences on computer screen	Typewriting procedure	2D gaze points {x,y}	Cognitive workload
GTEA Gaze [201]	14	-	-	Egocentric every day activities	Activities	2D gaze points {x,y}	Cognitive workload
GTEA Gaze+ [201], [202]	26	-	-	Egocentric every day activities	Activities	2D gaze points {x,y}	Cognitive workload
EGTEA Gaze+ [69]	32	-	-	Egocentric every day activities	Activities	2D gaze points {x,y} fixations, saccades, blinks, pupil diameter	Cognitive workload
EyeC3D [203]	21	18-31	5/16	Stereoscopic video sequences from NAMA3DS1, MUSCADE and Poznan multiview video database	Video watching	Fixation points, fixation density maps	Visual attention
SFU_ETDB [204]	15	-	-	Video sequences	Video watching	2D gaze points {x,y} heatmaps	Visual attention
Variability of eye movements when viewing dynamic natural scenes [78]	54	18-34	46/8	Movie shots and stop motion movie scenes	Video watching	2D gaze points {x,y}	Visual attention
DOVES [205]	29	mean 27	18/11	Images from Natural Image Dataset [206]	Image viewing	2D gaze points {x,y}, fixation points	Visual attention
MIT CSAIL [77]	15	18-35	-	Images from Flickr creative commons and LabelMe	Image viewing	2D gaze points {x,y}, fixation points	Visual attention
MIT CVCL [83]	14	18-40	-	Colour pictures of urban environments	Image viewing	2D gaze points {x,y}, fixation points, fixation duration	Visual attention
USC CRCNS [207]	8	23-32	3/5	Video clips	Video watching	2D gaze points {x,y}, fixation number, fixation duration, saccade number, saccade duration	Visual attention
MPIIGaze [208]	15	-	-	Random sequence of on-screen positions	Computer screen observation	2D gaze points {x,y}, head poses	Visual attention
UT Multi-view [209]	50	-	15/35	Visual shrinking target displayed on the monitor	Computer screen observation	2D gaze points {x,y}	Visual attention

dataset includes EEG, eye-tracking, and physiological signals (GSR and Heart rate) collected from 34 individuals, of which 18 were able-bodied and 16 were motor-impaired. The relevant data were collected during the interaction with a specifically designed interface for web browsing and multimedia content manipulation and during imaginary movement tasks. These were tasks that required increased mental effort and often multitasking abilities. Therefore, despite the drawback that the MAMEM Phase I lacks blinks and pupil related features, it may be more suitable that EGTEA Gaze+ for researchers that want to study the changes in mental workload.

Finally, the datasets available for the study of visual attention are very similar both in terms of subjects present, with the exception of the Variability of eye movement when viewing dynamic natural scenes dataset, the features that they

provide, usually 2D gaze points, as well as the nature of the task executed, which is usually video and/or image viewing. Worth noting is the fact that UT Multi-view and Variability of eye movements when viewing dynamic natural scenes datasets contain a relatively larger number of subjects, and that USCCRCNS includes the largest range of eye features, which include 2D gaze points, fixation number, fixation duration, saccade number, and saccade duration.

VIII. DISCUSSION

In recent years, gaze analysis is seen as an interesting area of research regarding emotion and cognition and for revealing attentional focus and other cognitive strategies of an individual. As a result, the robust and consistent estimation of eye-related metrics through eye tracking and their interpretation

for the recognition of emotional or cognitive processes is an important area of current research.

In this paper, a review on the eye features that relate to visual attention, emotional arousal and cognitive workload and their correlation with emotional and cognitive processes was pursued. In the first section, the emotional and cognitive processes are defined. Then, in Section III, the metrics related to the eye movements are presented and explained. Fixations, saccadic movements, pupil size, blinks, microsaccades and smooth pursuit eye movements are the most commonly used eye to describe visual attention, emotional arousal and cognitive workload.

Section IV provides a review of the current state of the art regarding eye and pupil tracking systems with special emphasis on video-based eye trackers and methods for the estimation of eye behaviour metrics. An overview of the various types of eye trackers types is presented together with their corresponding advantages and limitations. It is revealed that a popular trend relates to the use of head-mounted trackers (mainly in the form of glasses) using wearable miniaturized IR eye-cameras attached close to the eye. This setup allows to calculate both eye-related movements and pupil size variation using computer vision algorithms. Various algorithms have been developed for the gaze estimation based on 2D/3D CV models. Calibration is a critical procedure for reliable estimation of eye gaze.

An additional finding of our review is the fact that there are specific metrics of eye and pupil behaviour providing valuable information related to emotional/cognitive processes that may significantly contribute to the recognition of these states. Regarding emotional arousal and stress, the pupil size and blink rate appear to be significantly involved in most of the relevant studies. They both increase during states of increased arousal or stress. On the other hand, a clear pattern regarding all other gaze metrics reviewed did not surface, thus remaining a current research goal. Even if literature related to gaze distribution metrics provides evidence and initial research hypotheses, more studies should be conducted to validate these assumptions. Regarding visual attention, it was found that the number and duration of fixations and total fixation time are closely linked to the tendency of focusing on a specific target, whereas other metrics such as saccade amplitude, microsaccade rate and blink rate seem to form useful tools. As for the identification and quantification of cognitive workload, it was found that pupil size, number of fixations and saccadic velocity are reliable positive indicators of increase in mental workload. Blink duration, saccade rate and microsaccade amplitude are also used for cognitive load identification according to several studies.

Section VI presents studies concerning machine learning approaches for the recognition and classification of the various emotional and cognitive processes as well as makes a comparison between them. The majority of the studies exploits the features derived from pupil diameter and blink characteristics, indicating their significance for the process classification. However, both for the emotional arousal and cognitive workload cases, in order to identify among their levels and states (multiclass problem) with higher accuracy

percentages, it is vital for many biosignals to be included in the features list. Although most of the algorithms -and especially SVM classification system- can recognise and classify the levels or states of each process with a relatively high accuracy rate, none of the studies reported focus on the discrimination among both the emotional and cognitive processes, thus providing a current research gap in the literature and prospective future research direction. Aside from that, another research matter is that the generalizability of the results is questionable and no relevant information is provided by most reported studies.

We also presented the publicly available datasets that are used - or can be used - for the estimation of emotional arousal, cognitive workload and visual attention. A comparative assessment of these the datasets was performed and suggestions were made regarding their suitability for the identification of emotional and cognitive processes based on the eye-related features that they contain. The fact that the available datasets differ not only in the type of affect in which the focus, but also in the number and demographics of participants, the nature of the tasks performed, and the eye-related features included creates non optimal conditions for the computational study of cognitive and affective states based solely on eye-related features.

In addition, it is important to mention that none of the datasets contained data for all - or at least two - of the emotional and cognitive processes investigated in the present manuscript. Therefore, the generation of a dataset that is focused on all of these cognitive and affective states would be extremely beneficiary for the research community.

In conclusion, our review indicates that eye metrics such as fixations, saccades, blinks and pupil size may provide valuable and reliable information for classifying emotional and cognitive processes. For improved classification results and especially for classification between the various cognition levels and emotional states more biosignals except the eye features are required to be used. In addition, more research in this topic needs to be done and more datasets must be created targeting exclusively at tasks related to emotional arousal and cognitive workload leading to better discrimination among the various stages and levels.

This review can be used as a comprehensive guideline for researchers who address issues related to human emotions and cognitive processes and their reflection on eye or pupil related metrics.

ACKNOWLEDGMENT

This work was partially supported by the H2020 specific targeted research project SeeFar: Smart glasses for multifaceted visual loss mitigation and chronic disEase prevention indicator for healthier, saFer, and more productive workplAce foR ageing population. (H2020-SC1-DTH-2018-1, GA No 826429) [212]. This paper reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

REFERENCES

- [1] E. Granholm and S. R. Steinhauser, "Pupillometric measures of cognitive and emotional processes," *International Journal of Psychophysiology*, vol. 52, no. 1, pp. 1–6, 2004, pupillometric Measures of Cognitive and Emotional Processes.
- [2] S. Chen and J. Epps, "Automatic classification of eye activity for cognitive load measurement with emotion interference," *Computer Methods and Programs in Biomedicine*, vol. 110, no. 2, pp. 111–124, 5 2013.
- [3] D. T. Burley, N. S. Gray, and R. J. Snowden, "As Far as the Eye Can See: Relationship between Psychopathic Traits and Pupil Response to Affective Stimuli," *PLOS ONE*, vol. 12, no. 1, p. e0167436, 1 2017.
- [4] M. K. Eckstein, B. Guerra-Carrillo, A. T. Miller Singley, and S. A. Bunge, "Beyond eye gaze: What else can eyetracking reveal about cognition and cognitive development?" *Developmental Cognitive Neuroscience*, vol. 25, pp. 69–91, 6 2017.
- [5] J. A. Russell, "A circumplex model of affect," *Journal of Personality and Social Psychology*, vol. 39, no. 6, pp. 1161–1178, 12 1980.
- [6] S. C. Castro, D. L. Strayer, D. Matzke, and A. Heathcote, "Cognitive workload measurement and modeling under divided attention," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 45, no. 6, pp. 826–839, 2019.
- [7] Y. Niu, R. M. Todd, M. Kyan, and A. K. Anderson, "Visual and emotional salience influence eye movements," in *ACM Transactions on Applied Perception*, vol. 9, no. 3, 7 2012.
- [8] R. D. Lane, P. M.-L. Chua, and R. J. Dolan, "Common effects of emotional valence, arousal and attention on neural activation during visual processing of pictures."
- [9] G. Nix, C. Watson, T. Pyszczynski, and J. Greenberg, "Reducing depressive affect through external focus of attention," *Journal of Social and Clinical Psychology*, vol. 14, pp. 36–52, 3 1995.
- [10] S. A. McMains and S. Kastner, "Visual Attention," in *Encyclopedia of Neuroscience*. Springer Berlin Heidelberg, 11 2008, pp. 4296–4302.
- [11] P. Chavajay and B. Rogoff, "Cultural variation in management of attention by children and their caregivers." *Developmental psychology*, vol. 35, no. 4, pp. 1079–1090, 1999.
- [12] M. Valstar, "Automatic Facial Expression Analysis," in *Understanding Facial Expressions in Communication: Cross-Cultural and Multidisciplinary Perspectives*. Springer India, 1 2015, pp. 143–172.
- [13] G. Giannakakis, D. Grigoriadis, K. Giannakaki, O. Simantiraki, A. Roniotis, and M. Tsiknakis, "Review on psychological stress detection using biosignals," *IEEE Transactions on Affective Computing*, 2019.
- [14] J. Albert, S. López-Martín, and L. Carretié, "Emotional context modulates response inhibition: Neural and behavioral data," *NeuroImage*, vol. 49, no. 1, pp. 914–921, 2010.
- [15] R. McKendrick and A. Harwood, "Cognitive Workload and Workload Transitions Elicit Curvilinear Hemodynamics During Spatial Working Memory," *Frontiers in Human Neuroscience*, vol. 13, pp. 405 – 421, 11 2019.
- [16] J. Sweller, "Cognitive load theory, learning difficulty, and instructional design," *Learning and Instruction*, vol. 4, pp. 295–312, 1 1994.
- [17] M. I. Ahmad, I. Keller, D. A. Robb, and K. S. Lohan, "A framework to estimate cognitive load using physiological data," *Personal and Ubiquitous Computing*, pp. 1–15, 9 2020.
- [18] W. W. Wierwille and F. T. Eggemeier, "Recommendations for mental workload measurement in a test and evaluation environment," *Human Factors*, vol. 35, no. 2, pp. 263–281, 6 1993.
- [19] R. M. Pritchard, W. Heron, and D. O. Hebb, "Visual perception approached by the method of stabilized images." *Canadian journal of psychology*, vol. 14, pp. 67–77, 1960.
- [20] I. Rehman, N. Mahabadi, M. Motlagh, and T. Ali, "Anatomy, Head and Neck, Eye Fovea," in *StatPearls*. Treasure Island (FL): StatPearls Publishing, 2021.
- [21] M. Iwasaki and H. Inomata, "Relation between superficial capillaries and foveal structures in the human retina," *Investigative Ophthalmology and Visual Science*, vol. 27, no. 12, pp. 1698–1705, 1986.
- [22] Z. Bylinskii, M. A. Borkin, N. W. Kim, H. Pfister, and A. Oliva, "Eye fixation metrics for large scale evaluation and comparison of information visualizations," in *Eye Tracking and Visualization. ETVIS 2015. Mathematics and Visualization*. Springer, Cham, 2017, pp. 235–255.
- [23] E. Kowler, "Eye movements: The past 25years," *Vision Research*, vol. 51, pp. 1457–1483, 7 2011.
- [24] D. Purves, G. J. Augustine, D. Fitzpatrick, L. C. Katz, A.-S. LaMantia, J. O. McNamara, and S. M. Williams, "Types of Eye Movements and Their Functions," *Neuroscience. 2nd edition*, 2001.
- [25] M. Rucci and M. Poletti, "Control and Functions of Fixational Eye Movements," *Annual Review of Vision Science*, vol. 1, no. 1, pp. 499–518, 11 2015.
- [26] R. Engbert and R. Kliegl, "Microsaccades uncover the orientation of covert attention," *Vision Research*, vol. 43, pp. 1035–1045, 4 2003.
- [27] L. A. RIGGS, F. RATLIFF, J. C. CORNSWEET, and T. N. CORNSWEET, "The disappearance of steadily fixated visual test objects." *Journal of the Optical Society of America*, vol. 43, pp. 495–501, 1953.
- [28] E. Kowler, J. F. Rubinstein, E. M. Santos, and J. Wang, "Predictive smooth pursuit eye movements," *Annual Review of Vision Science*, vol. 5, pp. 223–246, 9 2019.
- [29] B. Jackson and B. Lucero-Wagoner, "The pupillary system," *Handbook of psychophysiology*, vol. 2, pp. 142–162, 2000.
- [30] M. M. Bradley, L. Miccoli, M. A. Escrig, and P. J. Lang, "The pupil as a measure of emotional arousal and autonomic activation," *Psychophysiology*, vol. 45, no. 4, pp. 602–607, 7 2008.
- [31] A. T. Duchowski, "Eye Tracking Techniques," in *Eye Tracking Methodology*. Cham: Springer International Publishing, 2017, pp. 49–57.
- [32] M. Pedrotti, M. A. Mirzaei, A. Tedescho, J.-R. Chardonnet, F. Merienne, S. Benedetto, and T. Baccino, "Automatic Stress Classification With Pupil Diameter Analysis," *International Journal of Human-Computer Interaction*, vol. 30, no. 3, pp. 220–236, 2014.
- [33] J. C. Eisenach, R. Curry, C. A. Aschenbrenner, R. C. Coghill, and T. T. Houle, "Pupil responses and pain ratings to heat stimuli: Reliability and effects of expectations and a conditioning pain stimulus," *Journal of Neuroscience Methods*, vol. 279, pp. 52–59, 3 2017.
- [34] B. Winn, D. Whitaker, D. B. Elliott, and N. J. Phillips, "Factors affecting light-adapted pupil size in normal human subjects," *Investigative Ophthalmology and Visual Science*, vol. 35, no. 3, pp. 1132–1137, 1994.
- [35] U. Wildenmann and F. Schaeffel, "Variations of pupil centration and their effects on video eye tracking," *Ophthalmic and Physiological Optics*, vol. 33, pp. 634–641, 11 2013. [Online]. Available: <https://pubmed.ncbi.nlm.nih.gov/24102513/>
- [36] A. Bulling, "Eye Movement Analysis for Context Inference and Cognitive-Awareness," Ph.D. dissertation, 2010.
- [37] "Search BioNumbers - The Database of Useful Biological Numbers," 2011. [Online]. Available: <https://bionumbers.hms.harvard.edu/search.aspx>
- [38] G. W. Ousler, K. W. Hagberg, M. Schindelar, D. Welch, and M. B. Abelson, "The ocular protection index," *Cornea*, vol. 27, no. 5, pp. 509–513, 6 2008.
- [39] A. Maffei and A. Angrilli, "Spontaneous blink rate as an index of attention and emotion during film clips viewing," *Physiology and Behavior*, vol. 204, pp. 256–263, 5 2019.
- [40] P. Majaranta and A. Bulling, "Eye Tracking and Eye-Based Human-Computer Interaction," 2014, pp. 39–65.
- [41] H. Kirchner and S. J. Thorpe, "Ultra-rapid object detection with saccadic eye movements: Visual processing speed revisited," *Vision research*, vol. 46, no. 11, pp. 1762–1776, 2006.
- [42] S. Cristina and K. P. Camilleri, "Unobtrusive and pervasive video-based eye-gaze tracking," *Image and Vision Computing*, vol. 74, pp. 21–40, 6 2018.
- [43] H. R. Chennamma and X. Yuan, "A Survey on Eye-Gaze Tracking Techniques," *arXiv preprint arXiv:1312.6410*, 12 2013.
- [44] J. Z. Lim, J. Mountstephens, and J. Teo, "Emotion recognition using eye-tracking: Taxonomy, review and current challenges," *Sensors*, vol. 20, p. 2384, 4 2020. [Online]. Available: <https://www.mdpi.com/1424-8220/20/8/2384>
- [45] D. W. Hansen and Q. Ji, "In the Eye of the Beholder: A Survey of Models for Eyes and Gaze," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 3, pp. 478–500, 2010.
- [46] K. Holmqvist and P. Blißnaut, "Small eye movements cannot be reliably measured by video-based p-cr eye-trackers," *Behavior Research Methods*, vol. 52, pp. 2098–2121, 10 2020. [Online]. Available: <https://doi.org/10.3758/s13428-020-01363-x>
- [47] A. Lanata, A. Greco, G. Valenza, and E. P. Scilingo, "Robust Head Mounted Wearable Eye Tracking System for Dynamical Calibration," *Journal of Eye Movement Research*, vol. 8, no. 5, pp. 1–15, 2015.
- [48] K. Takemura, K. Takahashi, J. Takamatsu, and T. Ogasawara, "Estimating 3-D point-of-regard in a real environment using a head-mounted eye-tracking system," *IEEE Transactions on Human-Machine Systems*, vol. 44, no. 4, pp. 531–536, 2014.
- [49] P. Kasprowski and K. Hareżlak, "Cheap and Easy PIN Entering Using Eye Gaze," *Annales UMCS, Informatica*, vol. 14, no. 1, p. 7584, 2014.

- [50] A. Bulling, F. Alt, and A. Schmidt, "Increasing the security of gaze-based cued-recall graphical passwords using saliency masks," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '12. New York, NY, USA: Association for Computing Machinery, 2012, p. 3011–3020.
- [51] C. Holland, A. Garza, E. Kurtova, J. Cruz, and O. Komogortsev, "Usability Evaluation of Eye Tracking on an Unmodified Common Tablet," in *Conference on Human Factors in Computing Systems*, New York, USA, 4 2013, pp. 295–300.
- [52] T. Imabuchi, O. D. A. Prima, H. Kikuchi, Y. Horie, and H. Ito, "Visible-spectrum remote eye tracker for gaze communication," in *Sixth International Conference on Graphic and Image Processing (ICGIP 2014)*, Y. Wang, X. Jiang, and D. Zhang, Eds., vol. 9443. SPIE, 3 2015.
- [53] J. N. Chi, C. Zhang, Y. T. Yan, Y. Liu, and H. Zhang, "Eye gaze calculation based on nonlinear polynomial and generalized regression neural network," in *5th International Conference on Natural Computation, ICNC 2009*, vol. 3, 2009, pp. 617–623.
- [54] C. Ma, K.-A. Choi, B.-D. Choi, and S.-J. Ko, "Robust remote gaze estimation method based on multiple geometric transforms," *Optical Engineering*, vol. 54, no. 8, p. 083103, 8 2015.
- [55] J. J. Cerrolaza, A. Villanueva, and R. Cabeza, "Taxonomic study of polynomial regressions applied to the calibration of video-oculographic systems," in *Eye Tracking Research and Applications Symposium (ETRA)*. New York, USA: ACM Press, 2008, pp. 259–266.
- [56] Z. Zhu and Q. Ji, "Eye and gaze tracking for interactive graphic display," *Machine Vision and Applications*, vol. 15, no. 3, pp. 139–148, 7 2004.
- [57] J. N. Chi, C. Zhang, Y. T. Yan, Y. Liu, and H. Zhang, "Eye gaze calculation based on nonlinear polynomial and generalized regression neural network," in *5th International Conference on Natural Computation, ICNC 2009*, vol. 3, 2009, pp. 617–623.
- [58] E. D. Guestrin and M. Eizenman, "General theory of remote gaze estimation using the pupil center and corneal reflections," *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 6, pp. 1124–1133, 6 2006.
- [59] A. Meyer, M. Böhme, T. Martinetz, and E. Barth, "A single-camera remote eye tracker," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 4021. Springer Verlag, 2006, pp. 208–211.
- [60] C. C. Lai, S. W. Shih, and Y. P. Hung, "Hybrid method for 3-D gaze tracking using glint and contour features," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 25, no. 1, pp. 24–37, 1 2015.
- [61] T. Ohno and N. Mukawa, "A free-head, simple calibration, gaze tracking system that enables gaze-based interaction," in *Eye Tracking Research and Applications Symposium (ETRA)*, 2004, pp. 115–122.
- [62] A. Kar and P. Corcoran, "A review and analysis of eye-gaze estimation systems, algorithms and performance evaluation methods in consumer platforms," pp. 495–519, 8 2017.
- [63] T. Nagamatsu, J. Kamahara, and N. Tanaka, "Calibration-free gaze tracking using a binocular 3d eye model," in *CHI '09 Extended Abstracts on Human Factors in Computing Systems*, ser. CHI EA '09. New York, NY, USA: Association for Computing Machinery, 2009, p. 3613–3618.
- [64] D. Model and M. Eizenman, "User-calibration-free remote eye-gaze tracking system with extended tracking range," in *2011 24th Canadian Conference on Electrical and Computer Engineering (CCECE)*, 2011, pp. 001 268–001 271.
- [65] C. H. Morimoto, A. Amir, and M. Flickner, "Detecting eye position and gaze from a single camera and 2 light sources," in *Object recognition supported by user interaction for service robots*, vol. 4, 2002, pp. 314–317 vol.4.
- [66] J. B. Huang, Q. Cai, Z. Liu, N. Ahuja, and Z. Zhang, "Towards accurate and robust cross-ratio based gaze trackers through learning from simulation," in *Eye Tracking Research and Applications Symposium (ETRA)*. New York, USA: Association for Computing Machinery, 2014, pp. 75–82.
- [67] F. L. Coutinho and C. H. Morimoto, "Free head motion eye gaze tracking using a single camera and multiple light sources," in *Brazilian Symposium of Computer Graphic and Image Processing*, 2006, pp. 171–178.
- [68] D. González-Ortega, F. J. Díaz-Pernas, M. Martínez-Zarzuela, M. Antón-Rodríguez, J. F. Díez-Higuera, and D. Boto-Giralda, "Real-time hands, face and facial features detection and tracking: Application to cognitive rehabilitation tests monitoring," *Journal of Network and Computer Applications*, vol. 33, no. 4, pp. 447–466, 7 2010.
- [69] Y. Li, M. Liu, and J. M. Rehg, "In the Eye of Beholder: Joint Learning of Gaze and Actions in First Person Video," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 11209 LNCS. Springer Verlag, 9 2018, pp. 639–655.
- [70] Y. L. Tian, T. Kanade, and J. F. Cohn, "Dual-state parametric eye tracking," in *4th IEEE International Conference on Automatic Face and Gesture Recognition, FG2000*, 2000, pp. 110–115.
- [71] B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in *Proceedings of the 7th international joint conference on Artificial intelligence - Volume 2*, ser. IJCAI'81. Vancouver, BC, Canada: Morgan Kaufmann Publishers Inc., Aug. 1981, pp. 674–679.
- [72] H. Mohsin and S. H. Abdullah, "Pupil detection algorithm based on feature extraction for eye gaze," in *2017 6th International Conference on Information and Communication Technology and Accessibility, ICTA 2017*, vol. 2017-December. Institute of Electrical and Electronics Engineers Inc., 4 2018, pp. 1–4.
- [73] C. Ionescu, C. Fosalau, D. Petrisor, and C. Zet, "A pupil center detection algorithm based on eye color pixels differences," in *2015 E-Health and Bioengineering Conference (EHB)*, 2015, pp. 1–4.
- [74] S. Y. Han, Y. Kim, S. H. Lee, and N. I. Cho, "Pupil center detection based on the unet for the user interaction in VR and AR environments," in *26th IEEE Conference on Virtual Reality and 3D User Interfaces, VR 2019 - Proceedings*. Institute of Electrical and Electronics Engineers Inc., 3 2019, pp. 958–959.
- [75] G. Underwood, "Cognitive processes in eye guidance: Algorithms for attention in image processing," *Cognitive Computation*, vol. 1, no. 1, pp. 64–76, 2 2009.
- [76] B. W. Tatler, "The central fixation bias in scene viewing: Selecting an optimal viewing position independently of motor biases and image feature distributions," *Journal of Vision*, vol. 7, pp. 4–4, 11 2007.
- [77] T. Judd, K. Ehinger, F. Durand, and A. Torralba, "Learning to predict where humans look," in *IEEE International Conference on Computer Vision (ICCV)*, 2009.
- [78] M. Dorr, T. Martinetz, K. R. Gegenfurtner, and E. Barth, "Variability of eye movements when viewing dynamic natural scenes," *Journal of Vision*, vol. 10, no. 10, pp. 28–28, 08 2010.
- [79] J. M. Wolfe and T. S. Horowitz, "What attributes guide the deployment of visual attention and how do they do it?" *Nature Reviews Neuroscience*, vol. 5, pp. 495–501, 2004.
- [80] B. W. Tatler, M. M. Hayhoe, M. F. Land, and D. H. Ballard, "Eye guidance in natural vision: Reinterpreting salience," *Journal of Vision*, vol. 11, no. 5, pp. 5–5, 05 2011.
- [81] I. Gilchrist and M. Harvey, "Evidence for a systematic component within scan paths in visual search," *Visual Cognition*, vol. 14, pp. 704–715, 8 2006.
- [82] K. Krejtz, A. Duchowski, I. Krejtz, A. Szarkowska, and A. Kopacz, "Discerning ambient/focal attention with coefficient k," *ACM Transactions on Applied Perception*, vol. 13, 5 2016.
- [83] K. A. Ehinger, B. Hidalgo-Sotelo, A. Torralba, and A. Oliva, "Modelling search for people in 900 scenes: A combined source model of eye guidance," *Visual Cognition*, vol. 17, no. 6-7, pp. 945–978, 2009.
- [84] Y. Zheng, X. Lan, J. Li, and F. Meng, "The contrast experiment on visual attention regions and saliency maps based on the eye tracker," in *2014 11th International Conference on Fuzzy Systems and Knowledge Discovery, FSKD 2014*. Institute of Electrical and Electronics Engineers Inc., 12 2014, pp. 830–834.
- [85] Y. Matsumura and K. Arai, "Influence of orthodontic appliances on visual attention to smiling faces by eye-tracking evaluation," *Orthodontic Waves*, 12 2019.
- [86] M. Bolmont, F. Bianchi-Demicheli, M. P. Boisgontier, and B. Cheval, "The Woman's Body (Not the Man's One) Is Used to Evaluate Sexual Desire: An Eye-Tracking Study of Automatic Visual Attention," *Journal of Sexual Medicine*, vol. 16, no. 2, pp. 195–202, 2 2019.
- [87] J. Goller, A. Mitrovic, and H. Leder, "Effects of liking on visual attention in faces and paintings," *Acta Psychologica*, vol. 197, pp. 115–123, 6 2019.
- [88] H. C. Liu, "Eye-tracking viewers' processing of web-based multimedia information," in *2009 Joint Conferences on Pervasive Computing, JCPC 2009*, 2009, pp. 699–704.
- [89] B. Pan, G. Gay, H. A. Hembrooke, G. K. Gay, L. A. Granka, M. K. Feusner, and J. K. Newman, "The Determinants of Web Page Viewing Behavior: An Eye-Tracking Study," *Eye Tracking Research and Applications Symposium (ETRA)*, no. January, pp. 147 – 154, 2004.
- [90] F. Espigares-Jurado, F. Muñoz-Leiva, M. B. Correia, C. M. Sousa, C. M. Ramos, and L. Fáisca, "Visual attention to the main image of a

- hotel website based on its position, type of navigation and belonging to Millennial generation: An eye tracking study," *Journal of Retailing and Consumer Services*, vol. 52, 1 2020.
- [91] F. Guo, Y. Ding, W. Liu, C. Liu, and X. Zhang, "Can eye-tracking data be measured to assess product design?: Visual attention mechanism should be considered," *International Journal of Industrial Ergonomics*, vol. 53, pp. 229–235, 5 2016.
- [92] L. Simmonds, S. Bellman, R. Kennedy, M. Nenycz-Thiel, and S. Bogomolova, "Moderating effects of prior brand usage on visual attention to video advertising and recall: An eye-tracking investigation," *Journal of Business Research*, 2019.
- [93] K. Schmidt, M. Gamer, K. Forkmann, and U. Bingel, "Pain affects visual orientation: an eye-tracking study," *Journal of Pain*, vol. 19, pp. 135–145, 2 2018. [Online]. Available: <https://pubmed.ncbi.nlm.nih.gov/29030322/>
- [94] C. Desmet and K. Diependaele, "An eye-tracking study on the road examining the effects of handsfree phoning on visual attention," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 60, pp. 549–559, 1 2019.
- [95] H. E. Miller, H. L. Kirkorian, and V. R. Simmering, "Using eye-tracking to understand relations between visual attention and language in children's spatial skills," *Cognitive Psychology*, vol. 117, 3 2020.
- [96] J. M. Franchak, K. S. Kretch, K. C. Soska, and K. E. Adolph, "Head-Mounted Eye Tracking: A New Method to Describe Infant Looking," *Child Development*, vol. 82, no. 6, pp. 1738–1750, 11 2011.
- [97] K. Rayner, C. M. Rotello, A. J. Stewart, J. Keir, and S. A. Duffy, "Integrating text and pictorial information: Eye movements when looking at print advertisements," *Journal of Experimental Psychology: Applied*, vol. 7, no. 3, pp. 219–226, 2001.
- [98] B. D. Corneil, D. P. Munoz, B. B. Chapman, T. Admans, and S. L. Cushing, "Neuromuscular consequences of reflexive covert orienting," *Nature Neuroscience*, vol. 11, pp. 13–15, 1 2008. [Online]. Available: [/record/2008-05455-007](https://doi.org/10.1038/nn2008-05455-007)
- [99] R. Kliegl, M. Rolfs, J. Laubrock, and R. Engbert, "Microsaccadic modulation of response times in spatial attention tasks," *Psychological Research*, vol. 73, pp. 136–146, 3 2009. [Online]. Available: <https://pubmed.ncbi.nlm.nih.gov/19066951/>
- [100] J. Laubrock, R. Kliegl, M. Rolfs, and R. Engbert, "When do microsaccades follow spatial attention?" *Attention, Perception, and Psychophysics*, vol. 72, pp. 683–694, 4 2010. [Online]. Available: <https://link.springer.com/article/10.3758/APP.72.3.683>
- [101] S. Meyberg, P. Sinn, R. Engbert, and W. Sommer, "Revising the link between microsaccades and the spatial cueing of voluntary attention," *Vision Research*, vol. 133, pp. 47–60, 4 2017.
- [102] Y. Chen, P. S. Holzman, and K. Nakayama, "Visual and cognitive control of attention in smooth pursuit," in *The Brain's eye: Neurobiological and clinical aspects of oculomotor research*, ser. Progress in Brain Research, J. Hyona, D. Munoz, W. Heide, and R. Radach, Eds. Elsevier, 2002, vol. 140, pp. 255–265.
- [103] T. A. Stuve, L. Friedman, J. A. Jesberger, G. C. Gilmore, M. E. Strauss, and H. Y. Meltzer, "The relationship between smooth pursuit performance, motion perception and sustained visual attention in patients with schizophrenia and normal controls," *Psychological Medicine*, vol. 27, pp. 143–152, 1 1997. [Online]. Available: <https://pubmed.ncbi.nlm.nih.gov/9122294/>
- [104] F. Ozkan and B. Ulutas, "Use of an eye-tracker to assess workers in ceramic tile surface defect detection," in *International Conference on Control, Decision and Information Technologies, CoDIT 2016*. Institute of Electrical and Electronics Engineers Inc., 10 2016, pp. 88–91.
- [105] F. Guo, Y. Ding, W. Liu, C. Liu, and X. Zhang, "Can eye-tracking data be measured to assess product design?: Visual attention mechanism should be considered," *International Journal of Industrial Ergonomics*, vol. 53, pp. 229–235, 5 2016. [Online]. Available: <http://dx.doi.org/10.1016/j.ergon.2015.12.001>
- [106] J. R. Bardeen and T. A. Daniel, "An Eye-Tracking Examination of Emotion Regulation, Attentional Bias, and Pupillary Response to Threat Stimuli," *Cognitive Therapy and Research*, vol. 41, no. 6, pp. 853–866, 2017.
- [107] M. Alshehri and S. Alghowinem, "An exploratory study of detecting emotion states using eye-tracking technology," in *2013 Science and Information Conference*, 2013, pp. 428–433.
- [108] C. W. Liang, J. L. Tsai, and W. Y. Hsu, "Sustained visual attention for competing emotional stimuli in social anxiety: An eye tracking study," *Journal of Behavior Therapy and Experimental Psychiatry*, vol. 54, pp. 178–185, 3 2017.
- [109] Wenhui Liao, Weihong Zhang, Zhiwei Zhu, and Qiang Ji, "A Real-Time Human Stress Monitoring System Using Dynamic Bayesian Network." Institute of Electrical and Electronics Engineers (IEEE), 1 2006, pp. 70–70.
- [110] G. Laretzaki, S. Plainis, I. Vrettos, A. Chrisoulakis, I. Pallikaris, and P. Bitsios, "Model Threat and trait anxiety affect stability of gaze fixation," *Biological Psychology*, 2011.
- [111] E. Fox, A. Mathews, A. J. Calder, and J. Yiend, "Anxiety and Sensitivity to Gaze Direction in Emotionally Expressive Faces," *Emotion*, vol. 7, no. 3, pp. 478–486, 8 2007.
- [112] J. P. Staab, "The influence of anxiety on ocular motor control and gaze," *Current Opinion in Neurology*, vol. 27, no. 1, pp. 118–124, 2 2014.
- [113] M. G. Calvo and P. Avero, "Time course of attentional bias to emotional scenes in anxiety: Gaze direction and duration," *Cognition and Emotion*, vol. 19, no. 3, pp. 433–451, 4 2005.
- [114] L. L. Di Stasi, A. Catena, J. J. Cañas, S. L. Macknik, and S. Martinez-Conde, "Saccadic velocity as an arousal index in naturalistic tasks," pp. 968–975, 6 2013.
- [115] G. J. DiGirolamo, N. Patel, and C. L. Blaukopf, "Arousal facilitates involuntary eye movements," *Experimental Brain Research*, vol. 234, no. 7, pp. 1967–1976, 7 2016.
- [116] N. Derakshan, T. L. Ansari, M. Hansard, L. Shoker, and M. W. Eysenck, "Anxiety, inhibition, efficiency, and effectiveness: An investigation using the Antisaccade task," *Experimental Psychology*, vol. 56, no. 1, pp. 48–55, 2009.
- [117] G. Giannakakis, M. Pediaditis, D. Manousos, E. Kazantzaki, F. Chiarugi, P. G. Simos, K. Marias, and M. Tsiknakis, "Stress and anxiety detection using facial cues from videos," *Biomedical Signal Processing and Control*, vol. 31, pp. 89–101, 2017.
- [118] I. Pavlidis, J. Levine, and P. Baukol, "Thermal imaging for anxiety detection." Institute of Electrical and Electronics Engineers (IEEE), 11 2002, pp. 104–109.
- [119] M. Haak, S. Bos, S. Panic, and L. Rothkrantz, "Detecting stress using eye blinks and brain activity from EEG signals," in *Proceedings of the 1st driver car interaction and interface (DCII)*, 2009.
- [120] F. Onorati, R. Barbieri, M. Mauri, V. Russo, and L. Mainardi, "Reconstruction and analysis of the pupil dilation signal: Application to a psychophysiological affective protocol," in *International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, 2013, pp. 5–8.
- [121] T. Partala and V. Surakka, "Pupil size variation as an indication of affective processing," *International Journal of Human Computer Studies*, vol. 59, no. 1–2, pp. 185–198, 2003.
- [122] P. Ren, A. Barreto, Y. Gao, and M. Adjouadi, "Affective assessment by digital processing of the pupil diameter," *IEEE Transactions on Affective Computing*, vol. 4, no. 1, pp. 2–14, 2013.
- [123] H. M. Simpson and F. M. Molloy, "Effects of audience anxiety on pupil size," *Psychophysiology*, vol. 8, no. 4, pp. 491–496, Jul. 1971.
- [124] S. Baltaci and D. Gokcay, "Stress Detection in Human-Computer Interaction: Fusion of Pupil Dilation and Facial Temperature Features," *International Journal of Human-Computer Interaction*, vol. 32, no. 12, pp. 956–966, 12 2016.
- [125] J. Zhai and A. Barreto, "Stress detection in computer users based on digital signal processing of noninvasive physiological variables," in *Annual International Conference of the IEEE Engineering in Medicine and Biology - Proceedings*, 2006, pp. 1355–1358.
- [126] C.-A. Wang, T. Baird, J. Huang, J. D. Coutinho, D. C. Brien, and D. P. Munoz, "Arousal effects on pupil size, heart rate, and skin conductance in an emotional face task," *Frontiers in Neurology*, vol. 9, p. 1029, 2018.
- [127] A. O. De Berker, M. Tirole, R. B. Rutledge, G. F. Cross, R. J. Dolan, and S. Bestmann, "Acute stress selectively impairs learning to act," *Scientific Reports*, vol. 6, 7 2016.
- [128] M. O. Kimble, K. Fleming, C. Bandy, J. Kim, and A. Zambetti, "Eye tracking and visual attention to threatening stimuli in veterans of the Iraq war," *Journal of Anxiety Disorders*, vol. 24, no. 3, pp. 293–299, 4 2010.
- [129] V. L. Kinner, L. Kuchinke, A. M. Dierolf, C. J. Merz, T. Otto, and O. T. Wolf, "What our eyes tell us about feelings: Tracking pupillary responses during emotion regulation processes," *Psychophysiology*, vol. 54, no. 4, pp. 508–518, 4 2017.
- [130] L. Wang and Xie Yue, "Attention bias during processing of facial expressions in trait anxiety: An eye-tracking study," in *Proceedings of 2011 International Conference on Electronics and Optoelectronics*, vol. 1, Jul. 2011, pp. 347–350.
- [131] R. B. Rosse, T. N. Alim, S. K. Johri, A. L. Hess, and S. I. Deutch, "Anxiety and pupil reactivity in cocaine dependent subjects endorsing

- cocaine-induced paranoia: preliminary report,” *Addiction*, vol. 90, no. 7, pp. 981–984, 1995.
- [132] E. H. Hess and J. M. Polt, “Pupil size as related to interest value of visual stimuli,” *Science*, pp. 349–350, 1960.
- [133] Z. J. P. S. Steinhauer SR, Boller F, “Pupillary dilation to emotional visual stimuli revisited,” *Psychophysiology*, 1983.
- [134] M. Honma, “Hyper-volume of eye-contact perception and social anxiety traits,” *Consciousness and Cognition*, vol. 22, no. 1, pp. 167–173, 2013.
- [135] A. Babiker, I. Faye, and A. Malik, “Pupillary behavior in positive and negative emotions,” in *IEEE ICSIPA 2013 - IEEE International Conference on Signal and Image Processing Applications*. IEEE Computer Society, 2013.
- [136] S. Ishimaru, S. Jacob, A. Roy, S. S. Bukhari, C. Heisel, N. Grobmann, M. Thees, J. Kuhn, and A. Dengel, “Cognitive State Measurement on Learning Materials by Utilizing Eye Tracker and Thermal Camera,” in *Proceedings of the International Conference on Document Analysis and Recognition, ICDAR*, vol. 8, 1 2018, pp. 32–36.
- [137] L. Quigley, A. L. Nelson, J. Carriere, and C. L. Purdon, “The effects of trait and state anxiety on attention to emotional images: An eye-tracking study,” *Cognition & emotion*, vol. 26, no. 8, pp. 1390–1411, 12 2014.
- [138] M. S. Young, K. A. Brookhuis, C. D. Wickens, and P. A. Hancock, “State of science: mental workload in ergonomics,” pp. 1–17, 1 2015.
- [139] B. Xie and G. Salvendy, “Review and reappraisal of modelling and predicting mental workload in single- and multi-task environments,” *Work and Stress*, vol. 14, no. 1, pp. 74–99, 2000.
- [140] R. L. Charles and J. Nixon, “Measuring mental workload using physiological measures: A systematic review,” *Applied Ergonomics*, vol. 74, pp. 221–232, 1 2019.
- [141] D. Tao, H. Tan, H. Wang, X. Zhang, X. Qu, and T. Zhang, “A systematic review of physiological measures of mental workload,” *International Journal of Environmental Research and Public Health*, vol. 16, 8 2019.
- [142] M. D. Rivecourt, M. N. Kuperus, W. J. Post, and L. J. Mulder, “Cardiovascular and eye activity measures as indices for momentary changes in mental effort during simulated flight,” *Ergonomics*, vol. 51, pp. 1295–1319, 9 2008. [Online]. Available: <https://pubmed.ncbi.nlm.nih.gov/18802817/>
- [143] X. He, L. Wang, X. Gao, and Y. Chen, “The eye activity measurement of mental workload based on basic flight task,” in *IEEE International Conference on Industrial Informatics (INDIN)*, 2012, pp. 502–507.
- [144] H. J. Foy and P. Chapman, “Mental workload is reflected in driver behaviour, physiology, eye movements and prefrontal cortex activation,” *Applied Ergonomics*, vol. 73, pp. 90–99, 11 2018.
- [145] R. Mallick, D. Slayback, J. Touryan, A. J. Ries, and B. J. Lance, “The use of eye metrics to index cognitive workload in video games,” in *Proceedings of the 2nd Workshop on Eye Tracking and Visualization, ETVIS 2016*. Institute of Electrical and Electronics Engineers Inc., 2 2017, pp. 60–64.
- [146] M. Borys, M. Tokovarov, M. Wawrzyk, K. Wesołowska, M. Plechawska-Wójcik, R. Dmytruk, and M. Kaczorowska, “An analysis of eye-tracking and electroencephalography data for cognitive load measurement during arithmetic tasks.” Institute of Electrical and Electronics Engineers Inc., 4 2017, pp. 287–292.
- [147] A. Bayat and M. Pomplun, “The influence of text difficulty level and topic on eye-movement behavior and pupil size during reading,” in *Proceedings - 2016 2nd International Conference of Signal Processing and Intelligent Systems, ICSPIS 2016*. Institute of Electrical and Electronics Engineers Inc., 3 2017.
- [148] T. Zu, J. Hutson, L. C. Loschky, and N. S. Rebello, “Use of eye-tracking technology to investigate cognitive load theory,” *arXiv*, 3 2018. [Online]. Available: <http://arxiv.org/abs/1803.02499>
- [149] G. G. Menekse Dalveren and N. E. Cagiltay, “Insights from surgeons’ eye-movement data in a virtual simulation surgical training environment: effect of experience level and hand conditions,” *Behaviour and Information Technology*, vol. 37, no. 5, pp. 517–537, 5 2018.
- [150] H. Sheridan and E. M. Reingold, “Chess players’ eye movements reveal rapid recognition of complex visual patterns: Evidence from a chess-related visual search task,” *Journal of Vision*, vol. 17, no. 3, pp. 4–4, 3 2017.
- [151] Q. Wang, S. Yang, M. Liu, Z. Cao, and Q. Ma, “An eye-tracking study of website complexity from cognitive load perspective,” *Decision Support Systems*, vol. 62, pp. 1–10, 2014.
- [152] M. Keskin, K. Ooms, A. O. Dogru, and P. De Maeyer, “Exploring the cognitive load of expert and novice map users using eeg and eye tracking,” *ISPRS International Journal of Geo-Information*, vol. 9, pp. 429 – 446.
- [153] E. İşbilir, M. P. Çakır, C. Acartürk, and A. Şimşek Tekerek, “Towards a multimodal model of cognitive workload through synchronous optical brain imaging and eye tracking measures,” *Frontiers in Human Neuroscience*, vol. 13, p. 375, 10 2019.
- [154] K. Krejtz, A. T. Duchowski, A. Niedzielska, C. Biele, and I. Krejtz, “Eye tracking cognitive load using pupil diameter and microsaccades with fixed gaze,” *PLoS ONE*, vol. 13, 9 2018. [Online]. Available: [/pmc/articles/PMC6138399/?report=abstract](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6138399/?report=abstract) <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6138399/>
- [155] T. Kosch, M. Hassib, P. W. Wozniakwozniak, D. Buschek, and F. Alt, “Your eyes tell: Leveraging smooth pursuit for assessing cognitive workload,” 2018. [Online]. Available: <https://doi.org/10.1145/3173574.3174010>
- [156] I. P. Bodala, Y. Ke, H. Mir, N. V. Thakor, and H. Al-Nashash, “Cognitive workload estimation due to vague visual stimuli using saccadic eye movements,” in *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC 2014*. Institute of Electrical and Electronics Engineers Inc., 11 2014, pp. 2993–2996.
- [157] Institute of Electrical and Electronics Engineers, *IEEE 10th International Conference on Industrial Informatics : 25-27 July, 2012, Beijing, China*. IEEE, 2012.
- [158] Y. Yang, M. McDonald, and P. Zheng, “Can drivers’ eye movements be used to monitor their performance? A case study,” *IET Intelligent Transport Systems*, vol. 6, no. 4, pp. 444–452, 12 2012.
- [159] R. Contreras, J. Ghajar, S. Bahar, and M. Suh, “Effect of cognitive load on eye-target synchronization during smooth pursuit eye movement,” *Brain Research*, vol. 1398, pp. 55–63, 6 2011.
- [160] E. Siegenthaler, F. M. Costela, M. B. Mccamy, L. L. D. Stasi, J. Otero-Millan, A. Sonderegger, R. Groner, S. Macknik, and S. Martinez-Conde, “Task difficulty in mental arithmetic affects microsaccadic rates and magnitudes,” *European Journal of Neuroscience*, vol. 39, pp. 287–294, 1 2014.
- [161] X. Gao, H. Yan, and H.-j. Sun, “Modulation of microsaccade rate by task difficulty revealed through between- and within-trial comparisons,” *Journal of Vision*, vol. 15, no. 3, pp. 3–3, 03 2015.
- [162] S. Tokuda and G. Obinata, “Development of an Algorithm to Detect Saccadic Intrusions As an Index of Mental Workload,” in *2012 Proceedings of SICE Annual Conference (SICE)*, Akita, 2012, pp. 1369–1372.
- [163] T. Čegovnik, K. Stojmenova, G. Jakus, and J. Sodnik, “An analysis of the suitability of a low-cost eye tracker for assessing the cognitive load of drivers,” *Applied Ergonomics*, vol. 68, pp. 1–11, 4 2018.
- [164] R. Bednarik, J. Koskinen, H. Vrzakova, P. Bartczak, and A. P. Elomaa, “Blink-Based Estimation of Suturing Task Workload and Expertise in Microsurgery,” in *Proceedings - IEEE Symposium on Computer-Based Medical Systems*, vol. 2018-June. Institute of Electrical and Electronics Engineers Inc., 7 2018, pp. 233–238.
- [165] S. Benedetto, M. Pedrotti, L. Minin, T. Baccino, A. Re, and R. Montanari, “Driver workload and eye blink duration,” *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 14, no. 3, pp. 199–208, 2011.
- [166] M. Pomplun and S. Sunkara, “Pupil Dilation as an Indicator of Cognitive Workload in Human-Computer Interaction,” 2003.
- [167] N. Zhong, M. Li, Y. Wu, and S. Lu, “The impact of different forms of statistical information on reading efficiency, effect, and mental workload: An eye-tracking study,” in *The 2011 IEEE/ICME International Conference on Complex Medical Engineering*, May 2011, pp. 97–102.
- [168] C. Wu, J. Cha, J. Sulek, T. Zhou, C. P. Sundaram, J. Wachs, and D. Yu, “Eye-Tracking Metrics Predict Perceived Workload in Robotic Surgical Skills Training,” *Human Factors*, 2019.
- [169] M. Nakayama and Y. Hayakawa, “Relationships between Oculo-Motor Measures as Task-evoked Mental Workloads during a Manipulation Task,” in *Proceedings of the International Conference on Information Visualisation*, vol. 2019-July. Institute of Electrical and Electronics Engineers Inc., 7 2019, pp. 170–174.
- [170] O. Palinko, A. L. Kun, A. Shyrokov, and P. Heeman, “Estimating cognitive load using remote eye tracking in a driving simulator,” in *Eye Tracking Research and Applications Symposium (ETRA)*, 2010, pp. 141–144.
- [171] W. Soussou, M. Rooksby, C. Forty, J. Weatherhead, and S. Marshall, “EEG and eye-tracking based measures for enhanced training,” in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, 2012, pp. 1623–1626.

- [172] E. H. Hess and J. M. Polt, "Pupil size in relation to mental activity during simple problem-solving," *Science*, vol. 143, no. 3611, pp. 1190–1192, 1964.
- [173] G. Prabhakar, A. Mukhopadhyay, L. Murthy, M. Modiksha, D. Sachin, and P. Biswas, "Cognitive load estimation using ocular parameters in automotive," *Transportation Engineering*, vol. 2, p. 100008, 12 2020.
- [174] L. Wang, X. He, and Y. Chen, "Distinguishing analysis on workload peak and overload under time pressure with pupil diameter," in *2014 IEEE International Inter-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support, CogSIMA 2014*. IEEE Computer Society, 2014, pp. 151–155.
- [175] R. J. Jacob and K. S. Karn, "Commentary on section 4 - eye tracking in human-computer interaction and usability research: Ready to deliver the promises," in *The Mind's Eye*, J. Hyönä, R. Radach, and H. Deubel, Eds. Amsterdam: North-Holland, 2003, pp. 573–605.
- [176] Y. Fujigaki and K. Mori, "Longitudinal Study of Work Stress among Information System Professionals," *Plastics, Rubber and Composites Processing and Applications*, vol. 9, no. 4, pp. 369–381, 1997.
- [177] S. J. Czaja and J. Sharit, "Age Differences in the Performance of Computer-Based Work," *Psychology and Aging*, vol. 8, no. 1, pp. 59–67, 1993.
- [178] J. Zhai, A. B. Barreto, C. Chin, and C. Li, "Realization of stress detection using psychophysiological signals for improvement of human-computer interactions," in *Conference Proceedings - IEEE SOUTH-EASTCON*, 2005, pp. 415–420.
- [179] G. Giannakakis, K. Marias, and M. Tsiknakis, "A stress recognition system using HRV parameters and machine learning techniques," in *8th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos, (ACIIW 2019)*, 9 2019, pp. 269–272.
- [180] G. Giannakakis, D. Grigoriadis, and M. Tsiknakis, "Detection of stress/anxiety state from EEG features during video watching," in *International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, 11 2015, pp. 34–37.
- [181] G. Giannakakis, D. Manousos, V. Chaniotakis, and M. Tsiknakis, "Evaluation of head pose features for stress detection and classification," in *IEEE EMBS International Conference on Biomedical and Health Informatics (BHI 2018)*, 4 2018, pp. 406–409.
- [182] A. M. Hughes, G. M. Hancock, S. L. Marlow, K. Stowers, and E. Salas, "Cardiac Measures of Cognitive Workload: A Meta-Analysis," *Human Factors*, vol. 61, no. 3, pp. 393–414, 5 2019.
- [183] T. K. Fredericks, S. D. Choi, J. Hart, S. E. Butt, and A. Mital, "An investigation of myocardial aerobic capacity as a measure of both physical and cognitive workloads," *International Journal of Industrial Ergonomics*, vol. 35, no. 12, pp. 1097–1107, 12 2005.
- [184] L. J. Zheng, J. Mountstephens, and J. T. T. Wi, "Multiclass emotion classification using pupil size in vr: Tuning support vector machines to improve performance," *Journal of Physics: Conference Series*, vol. 1529, p. 52062, 2020.
- [185] C. Aracena, S. Basterrech, V. Snáel, and J. Velásquez, "Neural Networks for Emotion Recognition Based on Eye Tracking Data," in *2015 IEEE International Conference on Systems, Man, and Cybernetics*, Oct. 2015, pp. 2632–2637.
- [186] P. Tarnowski, M. Kołodziej, A. Majkowski, and R. J. Rak, "Eye-tracking analysis for emotion recognition," *Computational Intelligence and Neuroscience*, vol. 2020, pp. 1687–5265, Sep 2020.
- [187] S. Alghowinem, R. Goecke, M. Wagner, G. Parker, and M. Breakspear, "Eye movement analysis for depression detection," in *2013 IEEE International Conference on Image Processing, ICIP 2013 - Proceedings*, 2013, pp. 4220–4224.
- [188] S. Al-gawwam and M. Benaissa, "Depression detection from eye blink features," in *2018 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*, 2018, pp. 388–392.
- [189] M. Fernández-Delgado, E. Cernadas, S. Barro, D. Amorim, and A. Fernández-Delgado, "Do we need hundreds of classifiers to solve real world classification problems?" *Journal of Machine Learning Research*, vol. 15, pp. 3133–3181, 2014.
- [190] L. M. Zhao, R. Li, W. L. Zheng, and B. L. Lu, "Classification of Five Emotions from EEG and Eye Movement Signals: Complementary Representation Properties," in *International IEEE/EMBS Conference on Neural Engineering, NER*, vol. 2019-March. IEEE Computer Society, 5 2019, pp. 611–614.
- [191] J. J. Guo, R. Zhou, L. M. Zhao, and B. L. Lu, "Multimodal Emotion Recognition from Eye Image, Eye Movement and EEG Using Deep Neural Networks," in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, 7 2019, pp. 3071–3074.
- [192] T. H. Li, W. Liu, W. L. Zheng, and B. L. Lu, "Classification of Five Emotions from EEG and Eye Movement Signals: Discrimination Ability and Stability over Time," in *International IEEE/EMBS Conference on Neural Engineering, NER*, vol. 2019-March. IEEE Computer Society, 5 2019, pp. 607–610.
- [193] E. Bozkir, D. Geisler, and E. Kasneci, "Person independent, privacy preserving, and real time assessment of cognitive load using eye tracking in a virtual reality setup," in *26th IEEE Conference on Virtual Reality and 3D User Interfaces, VR 2019 - Proceedings*, 3 2019, pp. 1834–1837.
- [194] J. Chen, Q. Zhang, L. Cheng, X. Gao, and L. Ding, "A Cognitive Load Assessment Method Considering Individual Differences in Eye Movement Data," in *IEEE International Conference on Control and Automation, ICCA*, vol. 2019-July. IEEE Computer Society, 7 2019, pp. 295–300.
- [195] X. Liu, T. Chen, G. Xie, and G. Liu, "Contact-Free Cognitive Load Recognition Based on Eye Movement," Nov. 2016.
- [196] W. L. Zheng, B. N. Dong, and B. L. Lu, "Multimodal emotion recognition using EEG and eye tracking data," in *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC 2014*, 11 2014, pp. 5040–5043.
- [197] S. Fan, Z. Shen, M. Jiang, B. L. Koenig, J. Xu, M. S. Kankanalli, and Q. Zhao, "Emotional Attention: A Study of Image Sentiment and Visual Attention," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 12 2018, pp. 7521–7531.
- [198] P. Koutras, A. Katsamanis, and P. Maragos, "Predicting eyes' fixations in movie videos: Visual saliency experiments on a new eye-tracking database," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 8532 LNAI. Springer Verlag, 2014, pp. 183–194.
- [199] S. Nikolopoulos, K. Georgiadis, F. Kalaganis, G. Liaros, I. Lazarou, K. Adam, A. Papazoglou-Chalikias, E. Chatzilari, V. Oikonomou, P. Petrantonakis, I. Kompatsiaris, C. Kumar, R. Menges, S. Staab, D. Müller, K. Sengupta, S. Bostantjopoulou, Z. Katsarou, G. Zeilig, M. Plotnik, A. Gotlieb, S. Fountoukidou, J. Ham, D. Athanasiou, A. Mariakaki, D. Comandici, E. Sabatini, W. Nisticco, and M. Plank, "A Multimodal dataset for authoring and editing multimedia content: The MAMEM project," *Data in Brief*, 2017.
- [200] F. P. Kalaganis, E. Chatzilari, S. Nikolopoulos, I. Kompatsiaris, and N. A. Laskaris, "An error-aware gaze-based keyboard by means of a hybrid BCI system," *Scientific Reports*, vol. 8, no. 1, p. 13176, Sep. 2018.
- [201] A. Fathi, Y. Li, and J. M. Rehg, "Learning to recognize daily actions using gaze," in *Lecture Notes in Computer Science*, vol. 7572. Springer, Berlin, Heidelberg, 2012, pp. 314–327.
- [202] Y. Li, Z. Ye, and J. M. Rehg, "Delving into egocentric actions," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 10 2015, pp. 287–295.
- [203] P. Hanhart and T. Ebrahimi, "EYEC3D: 3D video eye tracking dataset," in *2014 6th International Workshop on Quality of Multimedia Experience, QoMEX 2014*, 12 2014, pp. 55–56.
- [204] H. Hadizadeh, M. J. Enriquez, and I. V. Bajic, "Eye-tracking database for a set of standard video sequences," *IEEE Transactions on Image Processing*, vol. 21, no. 2, pp. 898–903, 2 2012.
- [205] I. Van Der Linde, U. Rajashekar, A. C. Bovik, and L. K. Cormack, "DOVES: A database of visual eye movements," *Spatial Vision*, vol. 22, no. 2, pp. 161–177, 3 2009.
- [206] J. H. Van Hateren and A. Van der Schaaf, "Independent component filters of natural images compared with simple cells in primary visual cortex," *Proceedings of the Royal Society B: Biological Sciences*, vol. 265, no. 1394, pp. 359–366, 3 1998.
- [207] R. Itti, Laurent; Carmi, "Eye-tracking data from human volunteers watching complex video stimuli. CRCNS.org." 2009.
- [208] X. Zhang, Y. Sugano, M. Fritz, and A. Bulling, "Appearance-based gaze estimation in the wild," in *IEEE Conference on Computer Vision and Pattern Recognition*, 10 2015, pp. 4511–4520.
- [209] Y. Sugano, Y. Matsushita, and Y. Sato, "Learning-by-synthesis for appearance-based 3D gaze estimation," in *IEEE Conference on Computer Vision and Pattern Recognition*, 9 2014, pp. 1821–1828.
- [210] D. D. Salvucci and J. H. Goldberg, "Identifying fixations and saccades in eye-tracking protocols," in *Proceedings of the 2000 Symposium on Eye Tracking Research Applications*, ser. ETRA '00. New York, NY, USA: Association for Computing Machinery, 2000, p. 71–78.
- [211] P. Ekman, *Emotions Revealed, Second Edition: Recognizing Faces and Feelings to Improve Communication and Emotional Life*, 2nd ed. New York: Holt Paperbacks, Mar. 2007.
- [212] "SeeFar project." [Online]. Available: <https://www.see-far.eu/>