A Panoramic Review of Situational Awareness Monitoring Systems

Aly Elshafei
University College of London
United Kingdom
aly.shafei.18@ucl.ac.uk

Daniela Romano
De Montfort University
United Kingdom
daniela.romano@dmu.ac.uk

ABSTRACT

Autonomous vehicles require a human-machine interaction (HMI) system for safe control transition during takeover requests (TORs). Vital to the HMI system is real-time monitoring of drivers' situational awareness, across diverse scenarios, without pre-calibration. This paper reviews situational awareness metrics, and considers sensor and computational model limitations, highlighting current gaps. We present a classification model categorizing awareness measurement methods—session-specific, online, offline—suitable for various applications, aiding researchers in choosing appropriate models.

CCS CONCEPTS

• **Applied computing** → *Transportation.*

ACM Reference Format:
Aly Elshafei and Daniela Romano. 2023. A Panoramic Review of Situational Awareness Monitoring Systems. In *2023 the 6th International Conference on Robot Systems and Applications (ICRSA) (ICRSA 2023), September 22–24, 2023, Wuhan, China.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3655532.3655539

1 INTRODUCTION

Situational awareness (SA) pertains to perceiving and understanding one's environment, crucial in safety-critical contexts like driving. To prevent errors, monitoring systems generate alerts for inadequate SA levels during tasks [17]. Endsley's model categorizes SA into 3 levels: environmental perception (Level 1), comprehension (Level 2), and future projection (Level 3) [17].

SA influences performance and mental workload, where higher workload often diminishes SA and degrades performance [18, 36]. Thus, this review considers workload experiments alongside SA measurement methods.

SA assessment methods in literature are subjective and objective. Subjective methods employ self-reported questionnaires, yielding scores for SA levels [16, 24, 26, 37, 48, 51]. Objective methods include physiological and real-time performance measures.

Notable subjective SA measures include: i) SA General Assessment Test (SAGAT) [53]; ii) Situation Present Assessment Method (SPAM); and iii) Subjective SA Test (SART) [9, 61]. Administered post-trial, they risk summative responses with potential inaccuracies [44], or during trials, impacting real-world use.

Objective SA can utilize physiological measures like electrocardiogram (ECG) [5, 12], electroencephalogram (EEG) [5, 12, 34], functional near-infrared spectroscopy (fNIRS) [5], galvanic skin response [12], gaze [34, 60], and eyelid closure (PERCLOS) [32]. Although used for cognitive load [6, 7, 25], these aren't widely adopted for SA due to equipment attachment challenges.

Indirect objective assessment involves task accuracy, time [19, 30], reflecting SA level. Driver speed, brake reaction, lateral position are indicators. Despite SA monitoring efforts, current methods lack real-world suitability due to equipment, calibration, and universal application limitations.

Recent interest in driving SA assessment introduced various methods. Evers et al. [10] contrasted self-report, physiological, and performance-based methods, favoring reliable hazard perception and reaction time over self-report. Lee et al. [28] explored eye-tracking, correlating metrics with SA. Multi-method SA assessment, including physiological and eye-tracking measures, emerges as valuable for driving evaluation.

1.1 Purpose of the study

The purpose of this literature review is to identify the different methods used to objectively measure SA with a focus on automated driving, illustrate the currently known limitations for the different measurement methods.

A special focus to identify the gaze tracking methods is also presented, showing the gap in the literature of identifying a suitable gaze tracking method that can measure SA in driving. And conclude by highlighting future possibilities for measuring SA.

As far as the author knows this is the only review on SA that investigates the different computational methods and their accuracies. And the different eye gaze measurement models that can be used and their limitations.

Following we also present the state-of-the-art computational methods that are applied to physiological data that can produce highly accurate measurements of SA that can be applied in real-world driving.

2 METHOD

A panoramic review covered ISI Web of Science, IEEE Xplore, and Google Scholar using keywords like "EEG," "ECG," "gaze tracking," "multi-modal feature extraction," "situational awareness," and "driving" within 2006-2022. Only 13 unique papers assessing Driving SA were found and reviewed.

Another search explored behavior linked to situational awareness.

ICRSA 2023, September 22–24, 2023, Wuhan, China

ACM ISBN 979-8-4007-0803-9/23/09
https://doi.org/10.1145/3655532.3655539

To evaluate computational methods analyzing physiological measures applicable to SA, a separate search sought "EEG," "ECG," "gaze tracking," "multi-modal feature extraction," and "computational methods." While 99 papers addressing cognitive distractions were identified, they didn't strictly classify SA.

2.1 Main themes identified

The literature identifies SA classification using ECG heart rate, EEG, and gaze tracking. However, accuracy remains low for distinguishing specific SA levels, like Level 1.

Given limited SA-focused studies, additional papers on cognitive load were examined. The cognitive load's connection to SA and the search for suitable computational methods prompted this review.

3 RESULTS

3.1 Heart rate measurements

Heart rate variability can be caused by effort, external stress, or medications, and electrocardiograms (ECG) are often utilised to evaluate stress [14]. ECG has also been correlated to SA in driving experiments [2, 3]. In Agrawal et al. [3] 134 participants were involved in a Level 3 automation study, and their heart rate was recorded at a frequency of 256 Hz. Since heart rate (HR) differs for everyone, the signal was normalized for each participant, requiring processing HR offline. They found that HR measurements were higher at the earlier runs in the simulator as drivers were tense, affecting the results. Also, Petersen et al. [43] measured the heart rate in a driving simulator experiment while also measuring the SA. The heart rate at a high level of SA was found to be around 120 BPM compared to 110 BPM at low SA. This conveys there is a slight difference in the heart rate at different SA levels, which could potentially be used as a measurement for SA in the real world. However, limited research has been conducted to investigate the ECG measurement's relation to SA.

The current limitations are that the measurements are subject-specific, include noise, and are currently processed offline. There are difficulties in pattern recognition of ECG data due to noise caused by heart rate variability. ECG has a relationship with both stress and SA which would have to be disentangled. To be usable in the real world, the measuring equipment might need to be attached to the driver (e.g. watch), or embedded in the car (e.g. steering wheel), and the signal processing computations performed in real-time. Also, ECG measurements can't be used to classify if the subject has a high level 1, 2, or 3 level of SA.

3.2 EEG measures

Brain activity can be measured in real-time for example, using time-locked signals of EEG activity, referred to as event-related brain potentials [50]. Jan et. al [33] had shown that there is a correlation between the frequency bands of β (12-30Hz) and γ (30-45Hz) and SA. And neurons in the left and right hemispheres of the parietal and temporal lobe were activated at a high SA. As these regions are responsible for the visuospatial ability and memory and reasoning tasks [33]. Three other studies also used EEG measurements to measure the SA [8, 11, 20]. As for ECG, EEG signals are subject-specific, include noise, and need to be processed in real-time to be usable in a real-world scenario. When the EEG measurements are processed in real time the accuracy decreases considerably. And, again the equipment must be attached to the driver.

3.2.1 *EEG Computations.*

In Jan et. al [33] samples are first labeled in an experiment from their performance results as either having a high or low SA. EEG data is then timestamped with the SA test data. Data must be cleaned from outliers as the machine learning models used for classification would be very sensitive to outliers. A random forest classifier is used which fits several decision tree classifiers and uses averaging to control over-fitting [40], achieving 67% in Jan et. al [33]. In Fernandez et al. [20] subjective scores were taken from the subjects to measure their SA and EEG was then used and achieved 61.5 % accuracy. Catherine et al. [11] SA was measured qualitatively from the performance in a visual search experiment. It was proven that EEG brain activity is directly linked with loss of SA in the experiments. In both experiments the level of SA was measured as either high or low.

Although EEG measurements have shown some success in being linked to two levels of SA however, more computational investigations are needed especially with the wealth of studies in the literature linking EEG measurements with the mental workload [13]. EEG measurements were able to measure high and low mental workload in Hogervorst et al. [29] reaching 90% and in Murata et al. [38] reaching 95% accuracy. This indicates that much higher accuracies could be achieved by EEG measurements in SA classification. These computations were session-specific and taken offline and were classifying two levels of mental workload. When the workload increased to three levels in Zhao et al. [62] and used support vector machine SVM, an accuracy of 81.64 % was achieved.

3.3 Gaze tracking models

Gaze tracking models can be clustered into two types; appearance-based models and featured-based models. Feature-based models either depend on geometric eye models or features obtained from facial landmarks and head orientation, while appearance models are based on the intensities of the eye. Gaze tracking measures have been used in several investigations to measure SA [3, 43]. There are several gaze metrics that can be used in classifying the level of SA of subjects. Gaze fixation is defined by Abbasi et al. [1] as the fixation time on an area of interest for a period between 100 ms and 2000 ms. Remote eye tracker manufacturers such as SmartEye Pro and Tobii consider 200 ms to be the minimum gaze time for a valid fixation. The table 1 shows the definitions of different gaze metrics used in the literature.

3.3.1 *Gaze measurements correlation to SA.*

Louie et al. [35] measured head pose estimation module, pupil tracking module, head movement rate module, and visual focus of attention accurately in a driving context using a 2D camera, however, each of these measurements needs to be correlated to SA which was not completed.

There were earlier studies in the literature that linked SA to gaze metrics in an autonomous vehicle experiment such as Peterson et al. [43] that showed the percentage of time the driver looked at the road had no correlation with a SA. Although, [59] indicated that the lower percentage of time drivers fixated on the road, the higher level of cognitive distraction, which is linked with SA.

Table 1: Eye gaze definitions

Eye gaze metric	Definition
Area of interest (AOI)	The area in each task were the subject needs to focus on to preform the task
Fixation rate	Number of fixations per minute, with a fixation categorized as less than 1° on AOI for 150 msec [49]
Dwell time	Time between fixations on the same AoI added to Total fixation time on an AoI [52]
Nearest neighbor index	Randomness of visual scanning behavior [52]
Percent road Center (PRC)	Percentage of gaze fixations falling in the center of the road [58]
Smooth pursuits	Eyes are in motion tracking a moving target while [4]
Vestibulo-ocular reflex (VOR)	A time slot where the eyes are compensating for head motion and stabilising the foveated area [4]
Optoki-netic nystagmus (OKN)	Sawtooth-like eyemovement patterns [4]
Head pursuit	Pursuit of a moving target using head motion [4]

Gao el al. [23] used Tobii Pro Glasses 3 to track human gaze fixation and used the SAGAT level 1 questionnaire, that focus on perception, in the experiment to measure the level of SA, and the results indicated a weak Pearson correlation coefficient of 0.12 between SA and the fixation time. However, the fixation time was chosen for all cars and pedestrians at the times studied in the experiment, while the area of interest in the situation is the area that needed to be investigated. These results convey there is no consensus in the literature on whether gaze metrics can measure SA in driving or not, as a suitable model of measuring gaze metrics while driving has not been developed yet. The reason there is no consensus in the literature on this topic is due to different gaze patterns in different automation levels not being investigated.

In manual driving, the proportion of looking on the road would be the same at high and low SA, and the proportion of time looking on the road would decrease as the automation level increases, which would affect the relationship between this metric and SA. Rangesh et al. [46] examined the correlation between gaze fixation to different defined regions in automated driving, such as the speedometer, radio, rear mirrors, and the road. In this driving scenario, the right driving mirror was the highest correlated to the observable situational awareness, which is completed by observers watching the drivers.

Zhou et al. [63] conducted the most detailed automated driving experiment as several predictor variables were used including gaze metrics and SA was classified as high and low. The gaze metrics used included fixation on the sky, road, and different mirrors. The gaze metric variables achieved a root mean square error in relation to SA of 0.121 using the lightGBM classifier which is a decision tree classifier.

Zhou et al. [64] used gaze fixation on defined features on the road, while taking into consideration the foveated region, and also took into consideration short term memory of the number of objects that can be remembered by the driver. This data was then processed using SVM achieving a 72.4% accuracy in predicting the SA of the driver which is the highest prediction of SA of a driver in the literature review.

Hofbauer et al. [28] measured the region of interest in driving using the machine learning networks as shown in [27]. Following the identification of the region of interest, it is defined as either undetected, detected, or comprehended in a manual driving experiment. The optimal SA is measured as the weighted sum of situational elements multiplied by their perception level. The perception level was taken as 0 for an undetected situational element, 0.5 for a detected situational element, and 1 for a comprehended situational element. Although this experiment concludes that eye gaze fixation on an ROI is a good indication of the level of SA, however, there was no SAGAT measurement taken in this experiment to validate the results.

3.3.2 Gaze measurements assessing cognitive distraction. Strayer et al. [50] assessed the cognitive distraction in vehicles while drivers were engaged in secondary tasks. Glance data was collected in manual driving and while drivers were engaged in secondary distracting tasks conveying that those glances to the right and left decreased as the cognitive workload increased in a manual driving condition. This is contrary to the results in Yang et al. [59], indicating the differences that occur in the relation of glances and cognitive distraction in manual and autonomous driving.

3.3.3 Gaze measurements computational methods. As mentioned, both features based and appearance-based models can be used for gaze tracking measures. Feature based models have been used successfully in [15, 21, 22, 31, 41, 47, 55], the performance of the feature based models was assessed in terms of the accuracy reached, the regions of interest involved, and whether the model is user specific or a global model that can be used with different subjects. Gaze tracking was done using head pose estimation where a regression model is used for each Euler angle producing a mean average error of 5.324 Ruiz et al.[47]. In Fridman, Lee, et al.[22] the feature based model used images from a camera attached to the dashboard of the vehicle achieving an accuracy of 94.6% while taking only confident estimations, and when cross validation is used the accuracy decreased to 65%. The latest feature based model used was in Dari et al., [15] outperforming earlier landmark based gaze estimation techniques by producing an accuracy of 92.3% on 75 video snippets with 20 subjects and 7 region of interest involved. Although, feature based models have been proved to produce great accuracies, however, calibration is always required which is the main drawback for feature-based models and the reason they are not chosen in this research. Appearance based models do not require calibration, the paper Fridman, Langhans, et al.[21] used a random forest classifier to achieve 44% accuracy for 6 classes when used as a global model, and a user based model achieved 91%. In the model used by Fridman, Langhans, et al.[21] a camera was positioned on the dashboard and 50 subjects were involved in driving through a local interstate high way. An infrared camera was used in Naqvi et al.[39] and an Alex net neural network was able to achieve 96.3% accuracy in 17 regions of interest, but the problem with infrared cameras that would be effective in detecting the glint of the eye is that they would be affected by differences that happen in illumination. Park et al.[42] used an appearance model that used a VGG-16 neural network that achieved accuracy within 3.18 degrees. The highest accuracy found in the literature for an appearance based model was in Vora et al. [54] were the camera was attached behind the rear view mirror and the model was trained on 11 drivers achieving an accuracy of 95.2%. for 6 regions of interest. The best gaze tracking model that can be deployed for measuring SA in a driving task would need to achieve high accuracy with at least six regions of interest, require no calibration, and be transferable between different sessions, subjects, and scenarios.

3.3.4 Conclusion section of eye gaze. Yu et al. [60] conveyed scan patterns for subjects in a flight simulation towards defined areas of interest (AOI) were directly linked to their level of SA. This conveys that if the AOI in a driving scenario is identified then the first level of SA (perception) can be classified using gaze metrics. The gaze patterns of drivers differ at different automation levels, which would affect the SA measurements. As in manual driving the driver is most of the time focused on the road and the percentage of time the driver focuses on the road decreases as the automation level increases. An experiment is needed to evaluate gaze metrics measuring SA at different automation levels.

3.4 Multimodal feature extraction

Multimodal feature extraction has been applied in several experiments to measure driver's cognitive load. In Yang et al. [57] ECG, and EEG measurements are taken, and the experimental results show their effectiveness in measuring the driver's fatigue.

Lobo et al. [34] used both EEG data and gaze tracking data to identify the level of cognitive distraction. For the EEG data, the alpha power and theta power are the features extracted, while for the gaze data, the percentage of eyelid closure and the pupil diameter are the features collected. The data is first labeled into the high workload, medium workload, and low workload. The best classification model used was the K-Nearest Neighbor achieving 99% accuracy, however, this approach did not allow transfer learning between different sessions and subjects. Functional near-infrared spectroscopy (fNIRS) conveys the level of oxygen in the blood when brain regions are active [45] and Ahn et al. [5] used EEG, ECG, and fNIRS data to determine mental fatigue. Although, the combined features were not related linearly to the fatigue level, each modality affected the fatigue level and in one subject the multimodal features improved the accuracy 30% compared to EEG only. Chen et al. [12] used a combination of electrocardiogram, galvanic skin response and respiration measurements to detect driver's stress. Data was preprocessed then support vector machine (SVM) and Extreme Learning Machine (ELM) were used to classify the stress level in drivers, with stress representing. The receiver operating characteristic (ROC) curve was used to assess the models used and SVM produced the highest accuracy of 99%.

Heart rate (HR), respiration and eye movement and EEG measures were used to classify the operators state in the form of high and low [56]. An ANN was used achieving 98.5% offline and 82% online.

No studies have yet investigated the use of multi modal feature extraction in driving to assess the SA of drivers. However, multimodal, measurements have shown a higher accuracy and could present a possibility to measure SA with higher accuracies in driving. Table 2 shows cognitive classification achieved using different models and the accuracy to correlating if such relation was investigated.

Table 2: Classification models using EEG and Gaze tracking measurements

Paper	measurement	online	offline	number of classes	model	accuracy
[33]	EEG		X	2	random forest classifier	67%
[20]	EEG		X	2		61.5%
[11]	EEG		X	2		
[43]	Fixation time	X		2		not correlated to SA
[23]	Fixation time	X		2		0.12 pearson correlation
[63]	Gaze metrics	X		2	Light GBM	0.121 RMSE
[64]	Fixation time	X		2	SVM	72.4%
[3]	Percentage on the road gaze fixation					
[28]	Gaze fixation on defined ROI	X		3		

4 DISCUSSION

This literature review focused on experiments measuring the situational awareness of drivers in vehicles. Different physiological measurements in the literature were evaluated such as EEG, eye gaze, ECG, and multi modal measurements.

EEG measurements have been used in [11, 20, 33] but neither of these studies was conducted in a driving simulator, and required post processing of the data, therefore can't be conducted online. EEG achieved the highest accuracy of 67% [33].

ECG has been used to measure SA in driving experiments in [2, 3, 43], however, the presence of noise and post processing makes ECG an unattractive option to measure SA. The most common method used in the literature to measure SA are eye tracking methods. However, the most common gaze tracking method used in driving experiments in the literature is the road monitoring ratio [43] which monitors the percentage of time the driver's gaze fixation is on the road in the journey. This metric is found to be effective in automated vehicles, as an increase in this percentage would be linked to a high SA level. However, in manual driving this metric is not effective in measuring SA as the driver would be looking to the road most of the time. Hofbauer et al. [28] conducted the only experiment that investigated gaze tracking on several regions of interest in a driving scene, and the results showed promise to using gaze tracking to measure driver's situational awareness at real time. However, Hofbauer's method included all the situational elements in the driver's surrounding which would be important in a driving situation for the driver to gain SA and comprehend his/her surroundings however there is an area of interest that the driver looks at in every scenario that has not been investigated and would further improve Hofbauer's method. And although, multi-modal measurements have shown success in measuring the driver's cognitive status they have not yet been used to measure SA which could further improve the results. The highest recorded SA in a driving experiment using physiological measures was using gaze metrics in Zhu et al. [64] reaching 72.4% accuracy.

With the research trend seen in the literature moving towards using gaze tracking to measure SA in driving, which is the right direction. However, the correct AOI is needed to be identified to be able to use the advances in gaze tracking technology to the advantage of measuring SA in driving. And special focus is needed to identify the changes in the gaze patterns in different automation levels.

5 CONCLUSION

In conclusion, SA measurement in driving situations in the literature has been examined through physiological measures. And from the review of the results and the different computational methods used with each physiological measurement, gaze tracking is the most suited to be applied in different situations, subjects, and computations can be done at real time with only a camera as shown in [54] achieving a 95.2% accuracy for 6 regions of interest. Zhu et al. [64] classified SA using gaze metrics using SVM reaching 72.4% accuracy, which is the highest in the literature.

However, further analysis is required to identify the AOI in driving and use that with the gaze fixation measurement to identify the level of SA of drivers in different automation levels, which could lead to a higher accuracy in classifying SA.

REFERENCES

[1] Jibran A Abbasi, Darragh Mullins, Nicolas Ringelstein, Patrice Reilhac, Edward Jones, and Martin Glavin. 2021. An Analysis of Driver Gaze Behaviour at Roundabouts. *IEEE Transactions on Intelligent Transportation Systems* (2021).

[2] Amirhossein S Aghaei, Birsen Donmez, Cheng Chen Liu, Dengbo He, George Liu, Konstantinos N Plataniotis, Huei-Yen Winnie Chen, and Zohreh Sojoudi. 2016. Smart driver monitoring: when signal processing meets human factors: in the driver's seat. *IEEE Signal Processing Magazine* 33, 6 (2016), 35–48.

[3] Shubham Agrawal and Srinivas Peeta. 2021. Evaluating the impacts of situational awareness and mental stress on takeover performance under conditional automation. *Transportation research part F: traffic psychology and behaviour* 83 (2021), 210–225.

[4] Ioannis Agtzidis, Mikhail Startsev, and Michael Dorr. 2019. A ground-truth data set and a classification algorithm for eye movements in 360-degree videos. *arXiv preprint arXiv:1903.06474* (2019).

[5] Sangtae Ahn, Thien Nguyen, Hyojung Jang, Jae G Kim, and Sung C Jun. 2016. Exploring neuro-physiological correlates of drivers' mental fatigue caused by sleep deprivation using simultaneous EEG, ECG, and fNIRS data. *Frontiers in human neuroscience* 10 (2016), 219.

[6] Pavlo Antonenko, Fred Paas, Roland Grabner, and Tamara Van Gog. 2010. Using electroencephalography to measure cognitive load. *Educational psychology review* 22, 4 (2010), 425–438.

[7] Tobias Appel, Natalia Sevcenko, Franz Wortha, Katerina Tsarava, Korbinian Moeller, Manuel Ninaus, Enkelejda Kasneci, and Peter Gerjets. 2019. Predicting cognitive load in an emergency simulation based on behavioral and physiological measures. In *2019 International Conference on Multimodal Interaction*. 154–163.

[8] Chris Berka, Daniel J Levendowski, Gene Davis, Melissa Whitmoyer, Kelly Hale, and K Fuchs. 2006. Objective measures of situational awareness using neurophysiology technology. *Augmented Cognition: Past, Present and Future* (2006), 145–154.

[9] Miriam Bongo and Rosemary Seva. 2022. Effect of Fatigue in Air Traffic Controllers' Workload, Situation Awareness, and Control Strategy. *The International Journal of Aerospace Psychology* 32, 1 (2022), 1–23.

[10] Sarah-Maria Castritius, Patric Schubert, Christoph Dietz, Heiko Hecht, Lynn Huestegge, Magnus Liebherr, and Christian T Haas. 2021. Driver Situation Awareness and Perceived Sleepiness during Truck Platoon Driving–Insights from Eye-tracking Data. *International Journal of Human–Computer Interaction* 37, 15 (2021), 1467–1477.

[11] Di Catherwood, Graham K Edgar, Dritan Nikolla, Chris Alford, David Brookes, Steven Baker, and Sarah White. 2014. Mapping brain activity during loss of situation awareness: an EEG investigation of a basis for top-down influence on perception. *Human factors* 56, 8 (2014), 1428–1452.

[12] Lan-lan Chen, Yu Zhao, Peng-fei Ye, Jian Zhang, and Jun-zhong Zou. 2017. Detecting driving stress in physiological signals based on multimodal feature analysis and kernel classifiers. *Expert Systems with Applications* 85 (2017), 279–291.

[13] Baoquan Cheng, Chaojie Fan, Hanliang Fu, Jianling Huang, Huihua Chen, and Xiaowei Luo. 2022. Measuring and computing cognitive statuses of construction workers based on electroencephalogram: a critical review. *IEEE Transactions on Computational Social Systems* (2022).

[14] Hsiu-Sen Chiang. 2015. Ecg-based mental stress assessment using fuzzy computing and associative petri net. *Journal of Medical and Biological Engineering* 35, 6 (2015), 833–844.

[15] Simone Dari, Nikolay Kadrileev, and Eyke Hüllermeier. 2020. A neural network-based driver gaze classification system with vehicle signals. In *2020 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 1–7.

[16] Amir Dini, Cornelia Murko, Saeed Yahyanejad, Ursula Augsdörfer, Michael Hofbaur, and Lucas Paletta. 2017. Measurement and prediction of situation awareness in human-robot interaction based on a framework of probabilistic attention. In *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 4354–4361.

[17] Mica R Endsley. 1995. Measurement of situation awareness in dynamic systems. *Human factors* 37, 1 (1995), 65–84.

[18] Mica R Endsley and Daniel J Garland. 2000. Pilot situation awareness training in general aviation. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 44. SAGE Publications Sage CA: Los Angeles, CA, 357–360.

[19] Alexander Eriksson and Neville A Stanton. 2017. Takeover time in highly automated vehicles: noncritical transitions to and from manual control. *Human factors* 59, 4 (2017), 689–705.

[20] Raul Fernandez Rojas, Essam Debie, Justin Fidock, Michael Barlow, Kathryn Kasmarik, Sreenatha Anavatti, Matthew Garratt, and Hussein Abbass. 2019. Encephalographic assessment of situation awareness in teleoperation of human-swarm teaming. In *International Conference on Neural Information Processing*. Springer, 530–539.

[21] L Fridman, P Langhans, J Lee, and B Reimer. 2019. Driver gaze region estimation without using eye movement. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*.

[22] Lex Fridman, Joonbum Lee, Bryan Reimer, and Trent Victor. 2016. 'Owl'and 'Lizard': Patterns of head pose and eye pose in driver gaze classification. *IET Computer Vision* 10, 4 (2016), 308–314.

[23] Xiaofeng Gao, Xingwei Wu, Samson Ho, Teruhisa Misu, and Kumar Akash. 2022. Effects of Augmented-Reality-Based Assisting Interfaces on Drivers' Object-wise Situational Awareness in Highly Autonomous Vehicles. *arXiv preprint arXiv:2206.02332* (2022).

[24] Leo J Gugerty. 1997. Situation awareness during driving: Explicit and implicit knowledge in dynamic spatial memory. *Journal of Experimental Psychology: Applied* 3, 1 (1997), 42.

[25] Eija Haapalainen, SeungJun Kim, Jodi F Forlizzi, and Anind K Dey. 2010. Psycho-physiological measures for assessing cognitive load. In *Proceedings of the 12th ACM international conference on Ubiquitous computing*. 301–310.

[26] Sogand Hasanzadeh, Behzad Esmaeili, and Michael D Dodd. 2016. Measuring construction workers' real-time situation awareness using mobile eye-tracking. In *Construction Research Congress 2016*. 2894–2904.

[27] Markus Hofbauer, Christopher B Kuhn, Jiaming Meng, Goran Petrovic, and Eckehard Steinbach. 2020. Multi-view region of interest prediction for autonomous driving using semi-supervised labeling. In *2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 1–6.

[28] Markus Hofbauer, Christopher B Kuhn, Lukas Püttner, Goran Petrovic, and Eckehard Steinbach. 2020. Measuring driver situation awareness using region-of-interest prediction and eye tracking. In *2020 IEEE International Symposium on Multimedia (ISM)*. IEEE, 91–95.

[29] Maarten A Hogervorst, Anne-Marie Brouwer, and Jan BF Van Erp. 2014. Combining and comparing EEG, peripheral physiology and eye-related measures for the assessment of mental workload. *Frontiers in neuroscience* 8 (2014), 322.

[30] Mike Howard, Rashmi Sundareswara, Mike Daily, Rajan Bhattacharyya, Sam Kaplan, Nathan Mundhenk, Craig Lee, and Howard Neely. 2013. Using tactile displays to maintain situational awareness during driving. In *2013 IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA)*. IEEE, 228–237.

[31] Yifei Huang, Minjie Cai, Zhenqiang Li, and Yoichi Sato. 2018. Predicting gaze in egocentric video by learning task-dependent attention transition. In *Proceedings of the European conference on computer vision (ECCV)*. 754–769.

[32] Lisheng Jin, Qingning Niu, Haijing Hou, Huacai Xian, Yali Wang, and Dongdong Shi. 2012. Driver cognitive distraction detection using driving performance measures. *Discrete Dynamics in Nature and Society* 2012 (2012).

[33] Jan Luca Kästle, Bani Anvari, Jakub Krol, and Helge A Wurdemann. 2021. Correlation between Situational Awareness and EEG signals. *Neurocomputing* 432 (2021), 70–79.

[34] Jesus L Lobo, Javier Del Ser, Flavia De Simone, Roberta Presta, Simona Collina, and Zdenek Moravek. 2016. Cognitive workload classification using eye-tracking and EEG data. In *Proceedings of the International Conference on Human-Computer Interaction in Aerospace*. 1–8.

[35] Wing-Yue Geoffrey Louie, Iyad Mansour, et al. 2019. Towards a driver monitoring system for estimating driver situational awareness. In *2019 28th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 1–6.

[36] Staffan Magnusson Nählinder. 2009. *Flight simulator training: Assessing the potential*. Ph. D. Dissertation. Linköping University Electronic Press.

[37] Michael D Matthews and Scott A Beal. 2002. *Assessing situation awareness in field training exercises*. Technical Report. Military Academy West Point NY Office of Military Psychology and Leadership.

[38] Atsuo Murata. 2005. An attempt to evaluate mental workload using wavelet transform of EEG. *Human Factors* 47, 3 (2005), 498–508.

[39] Rizwan Ali Naqvi, Muhammad Arsalan, Ganbayar Batchuluun, Hyo Sik Yoon, and Kang Ryoung Park. 2018. Deep learning-based gaze detection system for automobile drivers using a NIR camera sensor. *Sensors* 18, 2 (2018), 456.

[40] Mahesh Pal. 2005. Random forest classifier for remote sensing classification. *International journal of remote sensing* 26, 1 (2005), 217–222.

[41] Cristina Palmero, Javier Selva, Mohammad Ali Bagheri, and Sergio Escalera. 2018. Recurrent cnn for 3d gaze estimation using appearance and shape cues. *arXiv preprint arXiv:1805.03064* (2018).

[42] Seonwook Park, Shalini De Mello, Pavlo Molchanov, Umar Iqbal, Otmar Hilliges, and Jan Kautz. 2019. Few-shot adaptive gaze estimation. In *Proceedings of the IEEE/CVF international conference on computer vision*. 9368–9377.

[43] Luke Petersen, Lionel Robert, X Jessie Yang, and Dawn M Tilbury. 2019. Situational awareness, drivers trust in automated driving systems and secondary task performance. *arXiv preprint arXiv:1903.05251* (2019).

[44] Karl Popper. 2005. *The logic of scientific discovery*. Routledge.

[45] Valentina Quaresima and Marco Ferrari. 2019. A mini-review on functional near-infrared spectroscopy (fNIRS): where do we stand, and where should we go?. In *Photonics*, Vol. 6. MDPI, 87.

[46] Akshay Rangesh, Nachiket Deo, Kevan Yuen, Kirill Pirozhenko, Pujitha Gunaratne, Heishiro Toyoda, and Mohan M Trivedi. 2018. Exploring the situational awareness of humans inside autonomous vehicles. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 190–197.

[47] Nataniel Ruiz, Eunji Chong, and James M Rehg. 2018. Fine-grained head pose estimation without keypoints. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*. 2074–2083.

[48] Paul M Salmon, Neville A Stanton, Guy H Walker, Daniel Jenkins, Darshna Ladva, Laura Rafferty, and Mark Young. 2009. Measuring Situation Awareness in complex systems: Comparison of measures study. *International Journal of Industrial Ergonomics* 39, 3 (2009), 490–500.

[49] Tommy Strandvall. 2009. Eye tracking in human-computer interaction and usability research. In *IFIP Conference on Human-Computer Interaction*. Springer, 936–937.

[50] David L Strayer, Jonna Turrill, Joel M Cooper, James R Coleman, Nathan Medeiros-Ward, and Francesco Biondi. 2015. Assessing cognitive distraction in the automobile. *Human factors* 57, 8 (2015), 1300–1324.

[51] Lindsay Sturre, Dan Chiappe, Kim-Phuong L Vu, and Thomas Z Strybel. 2015. Using Eye Movements to Test Assumptions of the Situation Present Assessment Method. In *International Conference on Human Interface and the Management of Information*. Springer, 45–52.

[52] Koen van de Merwe, Henk van Dijk, and Rolf Zon. 2012. Eye movements as an indicator of situation awareness in a flight simulator experiment. *The International Journal of Aviation Psychology* 22, 1 (2012), 78–95.

[53] Arie P Van Den Beukel and Mascha C Van Der Voort. 2013. The influence of time-criticality on situation awareness when retrieving human control after automated driving. In *16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013)*. IEEE, 2000–2005.

[54] Sourabh Vora, Akshay Rangesh, and Mohan Manubhai Trivedi. 2018. Driver gaze zone estimation using convolutional neural networks: A general framework and ablative analysis. *IEEE Transactions on Intelligent Vehicles* 3, 3 (2018), 254–265.

[55] H Wang, Z Chen, and Y Zhou. [n. d.]. Hybrid coarse-fine classification for head pose estimation, arXiv. org. *URL http://search. proquest. com/docview/2170061213* ([n. d.]).

[56] Glenn F Wilson and Christopher A Russell. 2003. Real-time assessment of mental workload using psychophysiological measures and artificial neural networks. *Human factors* 45, 4 (2003), 635–644.

[57] Guosheng Yang, Yingzi Lin, and Prabir Bhattacharya. 2010. A driver fatigue recognition model based on information fusion and dynamic Bayesian network. *Information Sciences* 180, 10 (2010), 1942–1954.

[58] Shiyan Yang, Jonny Kuo, and Michael G Lenné. 2018. Analysis of gaze behavior to measure cognitive distraction in real-world driving. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 62. SAGE Publications Sage CA: Los Angeles, CA, 1944–1948.

[59] Shiyan Yang, Kyle M Wilson, Trey Roady, Jonny Kuo, and Michael G Lenné. 2022. Evaluating driver features for cognitive distraction detection and validation in manual and Level 2 automated driving. *Human factors* 64, 4 (2022), 746–759.

[60] Chung-san Yu, Eric Min-yang Wang, Wen-Chin Li, and Graham Braithwaite. 2014. Pilots' visual scan patterns and situation awareness in flight operations. *Aviation, space, and environmental medicine* 85, 7 (2014), 708–714.

[61] Guangtao Zhang, Katsumi Minakata, and John Paulin Hansen. 2019. Enabling Real-Time measurement of situation awareness in robot teleoperation with a head-mounted display. In *Human Factors Society Conference*. 169.

[62] Chunlin Zhao, Chongxun Zheng, Min Zhao, Yaling Tu, and Jianping Liu. 2011. Multivariate autoregressive models and kernel learning algorithms for classifying driving mental fatigue based on electroencephalographic. *Expert Systems with Applications* 38, 3 (2011), 1859–1865.

[63] Feng Zhou, X Jessie Yang, and Joost CF de Winter. 2021. Using eye-tracking data to predict situation awareness in real time during takeover transitions in conditionally automated driving. *IEEE Transactions on Intelligent Transportation Systems* 23, 3 (2021), 2284–2295.

[64] Haibei Zhu, Teruhisa Misu, Sujitha Martin, Xingwei Wu, and Kumar Akash. 2021. Improving driver situation awareness prediction using human visual sensory and memory mechanism. In *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 6210–6216.