

Correlation between Situational Awareness and EEG signals

Jan Luca Kästle^a, Bani Anvari^{a,*}, Jakub Krol^a, Helge A Wurdemann^b

^a Centre for Transport Studies, Department of Civil, Environmental and Geomatic Engineering, University College London, UK

^b Department of Mechanical Engineering, University College London, UK



ARTICLE INFO

Article history:

Received 13 July 2020

Revised 11 October 2020

Accepted 10 December 2020

Communicated by Zidong Wang

Keywords:

Situational Awareness

Electroencephalograph (EEG)

Psychology Experiment Building Language (PEBL) test

Random Forests

Decision trees

Brain computer interface

ABSTRACT

An important aspect in safety-critical domains is Situational Awareness (SA) where operators consolidate data into an understanding of the situation that needs to be updated dynamically as the situation changes over time. Among existing measures of SA, only physiological measures can assess the cognitive processes associated with SA in real-time. Some studies showed promise in detecting cognitive states associated with SA in complex tasks using brain signals (e.g. electroencephalogram/EEG). In this paper, an analytical methodology is proposed to identify EEG signatures associated with SA on various regions of the brain. A new data set from 32 participants completing the SA test in the PEBL is collected using a 32-channel dry-EEG headset. The proposed method is tested on the new data set and a correlation is identified between the frequency bands of β (12 – 30 Hz) and γ (30 – 45 Hz) and SA. Also, activation of neurons in the left and right hemisphere of the parietal and temporal lobe is observed. These regions are responsible for the visuo-spatial ability and memory and reasoning tasks. Among the presented results, the highest achieved accuracy on test data is 67%.

© 2021 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Understanding how people are acquainted with their environment, referred to as Situational Awareness (SA), is crucial in complex tasks that involve processing multiplexed information by a human operator. In order to reduce the cognitive load and undesirable and potentially fatal incidents in safety-critical domains, there is a need to identify the lack of SA in a timely manner [1–3]. Safety risks have emerged, for instance, from drivers handling highly automated vehicles that dramatically increase in numbers on our roads [3]. In this case, drivers are provided with increasingly sophisticated autonomy features (e.g. cruise control, lane keeping ability). Changing from a more automated vehicle state to a lower one requires the driver to regain Situational Awareness in order to safely conduct their vehicle. Identifying a potential lack of SA and subsequently giving suitable feedback to the user can prevent potential collisions.

The most cited approach for defining SA's processing trajectory is established by Endsley [4], where SA is divided into three levels: SA level 1 associated with the subjects' perception of their environment; SA level 2 associated with their comprehension; and SA level 3 associated with projection into a future state. It is noted that sufficient SA at lower levels is required, to reach adequate SA at

subsequent levels. Only if SA level 3 is achieved, the subject can adequately act to meet their objectives [4].

Existing measures of SA are categorised into subjective and objective measures. In the majority of previous studies, SA is assessed through task performance questionnaires where the results are mapped onto a certain score for each level of SA [5–7]. These subjective measures can be inaccurate as the subjects may misstate their responses [8]. On the other hand, behavioural and physiological measures give an objective understanding of the subject's reflexive emotional response which can be more accurate in the assessment and analysis of SA [9]. Behavioural sensors, including eye-tracking devices, can give an insight into the subject's perception of a situation. Data from physiological sensors is used to assess certain domains such as stress (EMG¹ [10], EDA² [11]), sleepiness and fatigue (EOG³ [12], ECG⁴ [13], and PPG⁵ [14]). Nonetheless, each of the mentioned sensors is able to detect only a limited range of emotional responses and consequently none of them is sensitive to all possible mental states [15–17].

One of the most promising physiological signals are brain activity recordings, as they simultaneously exhibit signals for all

¹ Electromyography.

² Electrodermal activity.

³ Electrooculography.

⁴ Electrocardiography.

⁵ Photoplethysmography.

* Corresponding author.

E-mail address: b.anvari@ucl.ac.uk (B. Anvari).

sensory inputs and therefore provide a deep insight into a subject’s physiological response. A widely used sensor to measure brain activity based on voltage fluctuations is the electroencephalograph (EEG) [18]. It is commonly used to acquire physiological measurement data in order to predict subject’s mental state, especially to detect workload [19], mental fatigue [20], sleepiness [21], and drowsiness [22] in different environments [23,24]. Given close relationship between the aforementioned mental states and SA [4], EEG is used in this study to assessment SA.

Few studies have looked into the correlation between SA and data from physiological measures. Among those, the SA of construction workers was monitored in [25] to identify subjects that are more prone to accidents and hazardous situations based on their personality. High SA is hereby was defined based on the frequency of visual scanning of the potential hazardous area. Since the SA metric in [25] was task-specific, they looked into only assessing perception of subjects (SA level 1) using eye-tracking glasses. Another study used fNIRS⁶ to demonstrate the effect of augmented reality on mental workload and SA [26]. Higher SA and less mental workload was linked to higher scores in two secondary tasks (1-back auditory task and questions about the environment) and reduced prefrontal cortex activity, during a real-life walk using either a handheld display or an augmented reality wearable display. It is noted that both studies performed comparative analysis between two subject groups and did not attempt to define an quantitative measure of what is considered to be high SA.

Three studies used EEG to link brain activity to SA. In [27] a 128-channel Geodesic Sensor Net (GSN) with semi-wet electrodes was used to identify the brain region showing high activity during loss of SA. This study also used comparative metric of SA, where loss of SA in subjects was artificially induced and the activity between the two considered cases were compared. The frequency band was not stated in this study. An earlier investigation [28] revealed a correlation between certain brain frequencies and SA, however, their sample size was small (8 participants) and no brain regions were identified. In [29], brain frequency bands are identified during loss of SA. The results are based on data collected using a 14-channel semi-wet EEG headset with a limited number of 10 participants and for a task-specific application.

Given the promising the results achieved to identify the relationship between EEG data and SA, the present study focuses on using brain activity as an objective measure to identify two classes of SA (*high SA* and *low SA*) of a subject. The choice of considering these two categories for SA is motivated by real-world applications such as hand-over tasks between humans and autonomous systems (i.e. shared autonomy). Here, it is of paramount importance to understand if a human is able to take back control of the autonomous system in a timely and safe manner, i.e. if the human has high or low SA. We propose an analytical methodology for identifying both spatial areas and frequency bands associated with SA in the brain (see Fig. 1). This is related to the feature definition employed in this study, where each of them is associated with unique spatial locations and frequency band. EEG was ultimately chosen as the objective measurement device in this study, as compared to other sensors available for measuring brain activity (e.g., MEG, fMRI and fNIRS), EEG is non-invasive [30], easy to apply [31], has a high temporal resolution (up to 1000 Hz [32]), and is portable [33]. However, mapping physiological data to SA is challenging due to high measurement noise and difficulty of disentangling signals into specific stimuli [31,34].

The contributions of this paper are as follows:

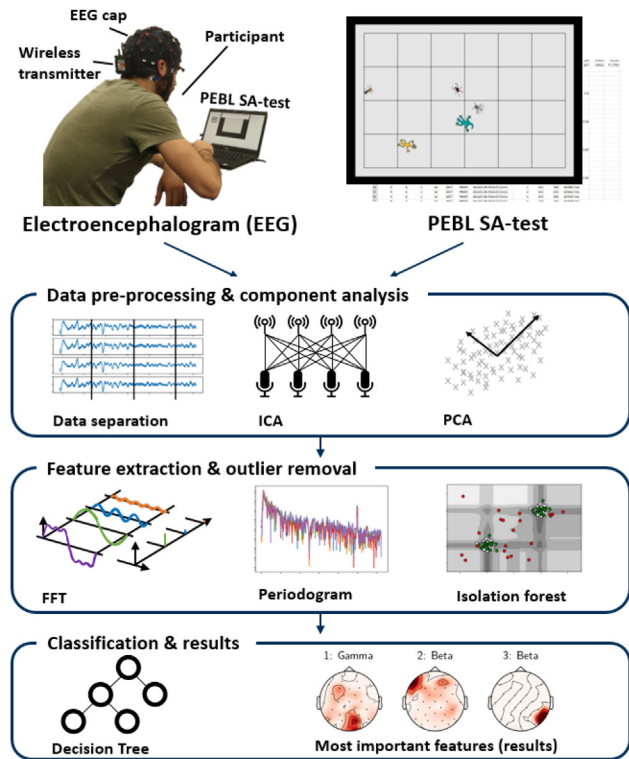


Fig. 1. Overview of the procedure used in this study to identify the spatial locations in a brain and frequency ranges which can be associated with Situation Awareness (SA).

- A new data set was collected using a 32-channel dry-EEG headset from 32 participants completing the well established Psychology Experiment Building Language (PEBL) SA test.
- A new analytical methodology is proposed to identify EEG signatures associated with SA levels (see Fig. 1). In the first stage, labels are defined using the PEBL SA test data. Then, the EEG data is preprocessed. In the third stage, features are extracted using Independent Component Analysis and Principal Component Analysis to determine new temporal signals and transforming them to frequency space with Fast Fourier Transform and periodogram. The obtained frequency characteristics allow classifying experimental runs into two categories of *high* and *low* SA in the fourth stage.
- Using our methodology on the new data set, the most important spatial areas and frequency ranges related to SA are identified.

Section 2 provides technical details of the experimental part of the procedure, and the detailed description of the employed algorithms for data processing are given in Section 3. The results are presented in Section 4, followed by the discussion and concluding remarks in Sections 5 and 6.

2. Experimental procedure

2.1. PEBL’s situational awareness test

PEBL is a framework for psychological assessments [35], which contains a dynamic visual tracking test, referred to as SA test, based on the more general Situational Awareness Global Assessment Technique (SAGAT) [36] proposed by Endsley and Garland [37]. SAGAT is a freeze-probe technique. A simulation is halted after a randomised amount of time and the participant is asked a series of questions linked to their acquaintance with the situation

⁶ Functional Near-Infrared Spectroscopy.

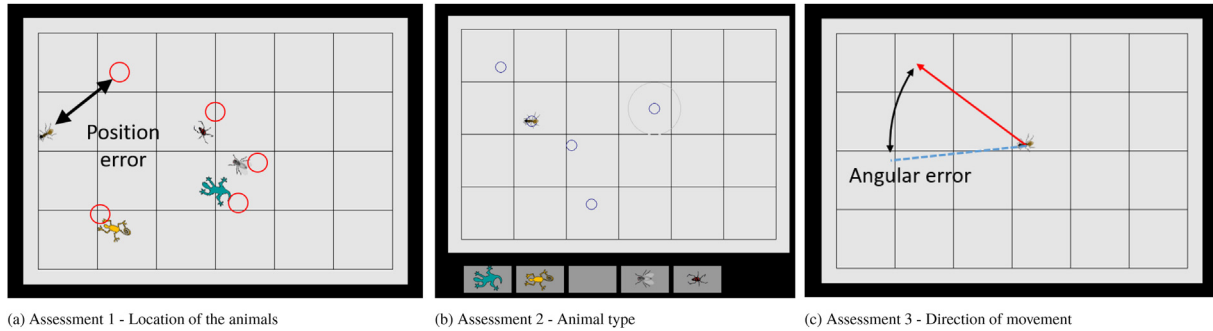


Fig. 2. The assessments of the PEBL SA test for: (a) identifying the most recent location of animals; (b) highlighting the correct animal type for two locations; (c) identifying the most recent direction of movement of the animal. The correct direction of movement is shown in red whilst the selected direction is shown in blue.

in that exact moment. For purposes of statistical significance, this process is repeated several times. Evaluating the participant’s performance towards the ground truth of the simulation gives a measurable SA score [5].

In the PEBL SA test, five animals of different shapes and colours are moving continuously over a grid. The participants are asked to monitor the animals’ positions, types and direction of movement within the grid. Following SAGAT freeze-probe approach, after randomised amount of time the animals disappear from the screen and participants are asked to answer certain questions about most recent situation. In order to evaluate participants’ performance with measurable SA scores, answers to each of the three distinct set of questions are assessed according to corresponding accuracy metric:

1. In order to test participant’s perception (SA level 1), they are asked to mark the last location of all five animals and the average position error is recorded (see Fig. 2a).
2. To test the degree of participant’s comprehension of the situation (SA level 2), they are asked to identify the type of animal located on two given positions (see Fig. 2b). The output includes two binary values specifying if the animals are identified correctly.
3. To evaluate the degree of participant’s projection (SA level 3), they are asked to specify the most recent movement direction of one of the animals, and the angular error is recorded.

A full experiment consists of five blocks and 90 runs that can be used for detailed analysis: Block 0 is a practice session with three sequential runs of the same assessment. Data from this block is not taken into account for further analysis. Blocks 1-3 consist of 15 consecutive runs of the same assessment and the participants are aware of the questions to follow while watching the simulations (45 runs in total). Block 4 features 15 runs of all three assessments in a random order and the participant are unaware of the questions to follow (45 runs in total).

2.2. Participants

Thirty-two subjects (13 female & 19 male) between ages of 18 and 39 participated in the study (Mean = 27.66, SD = 5.46). All participants reported good health and were not sleep-deprived while carrying out the experiment.

2.3. EEG equipment and data acquisition

A 32 channel g.tec g.Nautilus wireless EEG headset with gold-plated dry electrodes was used to measure the brain activity of all participants with a frequency of 250 Hz (see top left image in

Fig. 1). The EEG channels were placed according to the International 10–20 system [18]. A high-pass filter is applied to remove slow frequency drift (> 2 Hz) related to environmental influences. Also, frequencies between 48 Hz and 52 Hz are removed using a notch filter in order to avoid interferences with the power line. The brain activity was recorded throughout the whole duration of the SA test.

3. Methodology

The structure of this section is as follows: the data obtained from the SA test is divided into two categories of *high* and *low* SA in Section 3.1. Subsequently, the SA test labels and EEG data are time synchronised. The data slices with timestamps corresponding to assessment instances are retrieved in Section 3.2. The results are used to obtain spatio-spectral features in Section 3.3. Finally, two algorithms based on decision trees are employed to classify the extracted features into two categories of *high* and *low* SA in Section 3.4.

3.1. Label definition with PEBL SA test data

In the first stage, samples corresponding to each type of assessment (explained in Section 2.1) are labelled as either *high* or *low* SA. The definition of high and low SA categories is motivated by real-world applications such as hand-over tasks between humans and autonomous systems (e.g., shared autonomy). Here, it is of paramount importance to understand if a human is able to take back control of the autonomous system in a timely and safe manner, i.e. if the human has high or low SA. We argue that a human who does not have high SA would not be able to safely take over the control of an autonomous system. This concept is in line with similar approaches by [29, p. 1] stating that “a loss of SA has been associated with poor human performance, which can lead to misjudgement, errors, and life-threatening situations” and distinguishing between high and low SA in safety-critical tasks.

For SA levels 1 and 3, the continuous position and angular error values are normally distributed and are related to the accuracy of the subject’s response in the respective assessment. For instance, a higher error corresponds to higher deviation from the ground truth indicating lower SA. We modelled the probability distribution of the error values using a two-component Gaussian mixture model (see Fig. 4). The intersection point of these two Gaussian distributions results in our definition for high and low SA. It is worth mentioning that classifiers such as mixtures of beta distributions and mixtures of uniform and beta distributions resulted in significant lower classifier accuracy of SA.

For SA level 2, the position and angular error values are not continuous and, instead, consists of two binary values, signifying if the

participants chose the correct animal for the two locations. Here, the subject was considered to have high SA if and only if both animals were correctly identified. This stems from the fact that the chance of having at least one animal correct through random selection is 40%. Two correct answers are classified as high SA. All other runs are set to low SA.

3.2. Construction of data samples

Both EEG and SA test data are timestamped using one computer, and the EEG data is mapped onto the results of the SA tests. The maximal offset between the data points recorded from EEG and the SA test is 2ms, which is based on the temporal resolution of the EEG equipment. The EEG data covering the time t prior to the moment when a question was posed and only when the animals were moving on the screen, was isolated. Assuming that participants' SA is independent between the runs, the outcome of each test is treated as an individual sample. In total 2880 labelled samples (32 participants \times 90 runs) are acquired.

The dimension of the retained data y for a single sample is $y \in \mathbb{R}^{p \times N}$, where p corresponds to the number of EEG channels and N is the number of time instances ($N = 250t$), where t is time duration in seconds when the animals move prior to posing a question such that $t \sim U(2, 4)$.

3.3. Feature extraction using ICA and PCA

3.3.1. Component analysis

The cognitive processes are not associated with only one part of the brain as a single stimuli affects multiple spatial locations of a brain. Also, the EEG signal of a single channel consists of a superposition of different stimuli or sources. The goal is to extract features which are related primarily to different SA levels and map them on various regions of the brain. In order to decompose the superposed signal into its constituent stimuli and provide a spatial imprint of each source, Independent Component Analysis (ICA) [38] and Principal Component Analysis (PCA) [39] are used here to diversify the spatial imprints, maximise variability of features and, as a result, add components that encompass more than single channel locations. This allows for identification of more complex signals which impact multiple brain areas.

For p EEG channels and N time instances, both ICA and PCA aim to compute matrix $W \in \mathbb{R}^{p \times p}$ so that the EEG measurements y can be expressed as a linear combination of the underlying source signals $s \in \mathbb{R}^{p \times N}$, such that

$$y = Ws. \tag{1}$$

In (1), the i -th row of y corresponds to a measurement from a single channel, whilst the i -th row of s corresponds to time series associated with single source or stimuli. Eq. (1) approximates a single channel measurement as a weighted summation of *all* identified sources. Also, the imprint the i -th source has on brain locations can be examined using the magnitude of the values in i -th column of W , which, in essence, gives spatial structure associated with a specific stimulus. The higher the magnitude of the coefficient in W , the larger is the impact of a stimuli on the corresponding brain location.

Both ICA and PCA aim to find values of W and s , which best approximate given y . Despite similar linear formulation, the two algorithms use different objectives to identify an appropriate form of s . ICA relies on the assumption that the signals are non-Gaussian and statistically independent [38]. The decorrelation technique of PCA outputs sources such that it minimises the reconstruction error if r ($r < p$) components are used, i.e., $\min_{s,W} \|y - Ws\|_F$, where $W \in \mathbb{R}^{p \times r}$, $s \in \mathbb{R}^{r \times N}$ and $\|\cdot\|_F$ denotes the Frobenius norm [39]. The

principal components can be ranked based on the explained variance, and for EEG measurements, this results in the first principal components showing areas of the brain with largest changes in activation.

Data from all participants is standardised such that:

$$z = \frac{(y - \mu)}{\sigma} \tag{2}$$

Here, μ is the mean and σ is the standard deviation of the samples to approximate a standard normally distributed data set. For both ICA and PCA, the values of W_{ICA} and W_{PCA} are computed using the standardised data from all participants, z , whilst s_{ICA} and s_{PCA} are calculated based on z corresponding to each sample, e.g. $s_{PCA} = W_{PCA}^{-1}z$.

3.3.2. Spectral transformation

The most commonly used EEG features are identified based on the Power Spectral Density (PSD) of different EEG bands which gives information about the amplitude distribution of different frequency band across the brain [40,41]. In line with [42], the PSDs of five EEG bands δ (0–4 Hz), θ (4–8 Hz), α (8–12 Hz), β (12–30 Hz), and γ (30–45 Hz) are analysed. In this paper, two different approaches are used to extract frequency information for different brain regions: a Fast Fourier Transform (FFT) and periodogram with a window function. The FFT is based on Discrete Fourier Transform (DFT) [43] of the original time series which outputs $X_\omega \in \mathbb{C}$ giving information about amplitude and phase of the oscillations at frequency ω . The PSD can be approximated using a magnitude of the output $|X_\omega|^2$. Periodogram estimates PSD by taking the Fourier transform of the auto correlation function of the original time series [44]. In order to reduce the sensitivity of the output to small variations in time series, a sequence is multiplied with window functions to obtain multiple shorter subsequences, and the periodogram is computed for each of them. The final output is computed by calculating the average. Here, the flat top window [45] is used to construct window functions. Although, periodogram preprocessed with a window function is known to be more robust to noise, some useful information might be lost through averaging [46]. Consequently, both datasets obtain using FFT and periodogram are considered in further analysis.

The relationships between different brain areas are included in s_{ICA} and s_{PCA} which are obtained by taking a weighted sum of time series corresponding to multiple channels. Both FFT and periodogram are applied to z, s_{ICA} and s_{PCA} to give $|X_\omega|^2$. ω is assigned with discrete values for $\omega \leq 125$ Hz ($\omega = 0.2k$, where $k \in \mathbb{Z}^+$).

The total magnitude of each band is defined as a sum of $|X_\omega|^2$ corresponding to constituent frequencies, e.g. $|X_\alpha|^2 = \sum_{\omega \in \alpha} |X_\omega|^2$. Additionally, the power spectrum ratio is computed as in Eq. (3) [30]:

$$PSR = \frac{|X_\beta|^2}{(|X_\theta|^2 + |X_\alpha|^2)} \tag{3}$$

This procedure is applied to temporal signals obtained using PCA and ICA, and to original EEG measurements, i.e. z, s_{ICA} and s_{PCA} , resulting in 576 features (6 frequency features \times 32 time series \times 3 types of signals (ICA, PCA and EEG measurements) for both FFT and periodogram.

3.3.3. Outlier detection and removal

The EEG measurements recorded using dry electrodes are prone to noise and artefacts, such as eye blinking and motion artefacts, which create outliers among features. Many machine learning algorithms are sensitive to outliers during the training phase, as a single outlier can affect the classification model. Therefore,

anomalous values have to be detected and subsequently replaced or removed before the data set is fed into the training algorithm.

If the samples with big magnitudes are compared with corresponding time series, as shown in Fig. 3, they usually correspond to high amplitude oscillations characterising artefacts. Instead of discarding samples where at least one of the features has an unusually large value, outlier identification is performed individually for each feature. This is based on the assumption that an artefact does not necessarily impact all frequency ranges and all EEG channels equally. In fact, in the example presented in Fig. 3, the PSD is affected by an artefact only for frequencies smaller than ≈ 30 Hz.

The outliers can be identified using statistical or algorithmic approaches (e.g. Isolation Forest). In this study both approaches are used. In a statistical method, the threshold Δ is defined based on properties of the sample distribution, such as standard deviation, and the values larger than Δ are denoted as outliers. Three statistical outlier detection methods are used here with Δ set to: 1) $\Delta = \mu + \sigma$, 2) $\Delta = \mu + 2\sigma$ and 3) $\Delta = \tilde{x}$, where μ, σ and \tilde{x} are mean, standard deviation and median of a feature x .

In the Isolation Forest (IF) approach which is used here, a split value between the minimal and maximal value is randomly chosen creating two subsets of the data set. This procedure is repeated until all subsets consist of a single data point or only of identical data points. The maximum amount of splits required to separate a data point is used as a parameter to denote outliers, since anomalous points need fewer splits to be isolated [47].

Once the outliers are identified, they can be replaced with substitute values in the procedure referred to as imputation. In this study, the outliers are imputed with a missing value token (NaN), 0, μ , and $\max x$. Separate classification models are calculated for each of the possible imputations and data with no outlier replacement.

3.4. Classification with Random Forest and Boosted Trees

Two algorithms based on Decision Trees are used to classify EEG data: Random Forest (RF) and Boosted Trees (BT). Decision tree is a flowchart-like model in which a node represents a conditional test, branches depict an outcome of a test, and the leaf (lowest branch of the tree) is associated with one of the classes. The given input is propagated along the tree based on the result of conditional statements at the nodes until it reaches one of the leaves. The decision tree-based algorithms allow highly complex, non-linear decision boundaries between different categories. Also, they are easily interpretable and allow identification of the most statistically significant features where these characteristics are prevalent among conditional splits in the top nodes of the model [48].

When the number of nodes in a decision tree is left unconstrained, the algorithm tends to overfit the training data set. Thus, RF is used here to create multiple decision trees using a subset of the original data set. New data points are classified by all created trees based on a majority vote. The algorithm relies on the assumption that even though constituent trees might overfit to the training data, the generalisation of the model improves the results when the average is taken.

The BT algorithm addresses the underfitting issue of a model, where the classification accuracy is low on the training data. An example of such a model, is a shallow decision tree with small number of nodes. Shallow decision trees are referred to as weak learners and their accuracy is slightly better than classifier which chooses the output randomly. The underlying idea is to construct many weak learners and combine them in order to *boost* the accuracy of the overall model and at the same time prevent overfitting [49].

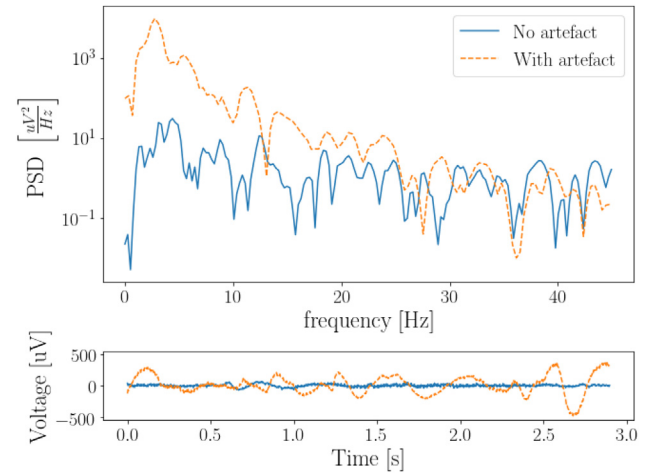


Fig. 3. An example of the effect a motion artefact has on the magnitude of the Power Spectral Density (PSD). For the data sample affected by motion artefacts (dashed orange line) the resulting PSD is up to three magnitudes higher than for the data sample without any artefacts (solid blue line).

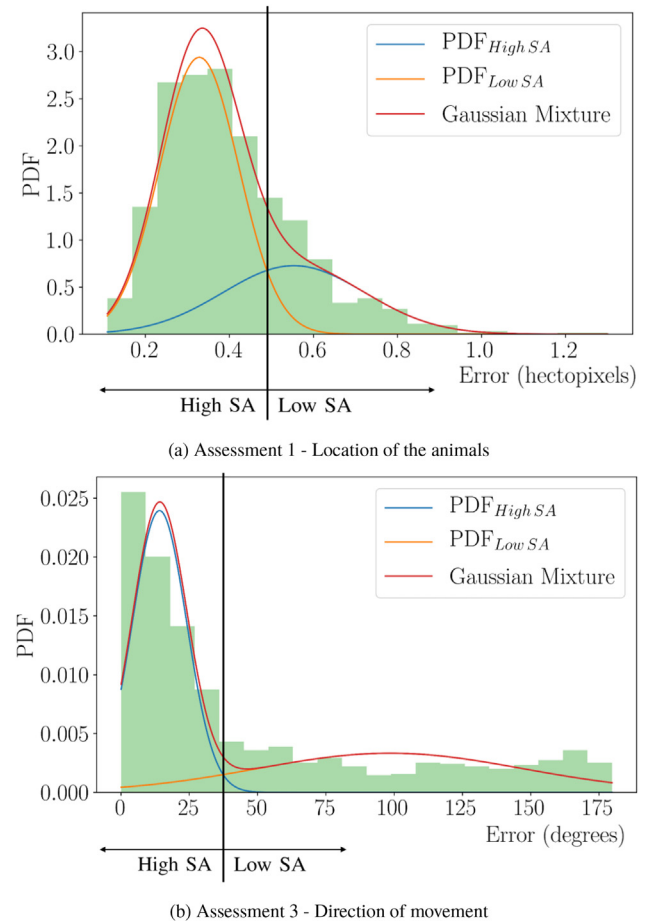


Fig. 4. Two-component Gaussian mixtures are fitted on the position and angular errors in order to identify the threshold for *high* and *low* SA.

In both algorithms, it is possible to determine which features have a higher impact on the classification. One criterion is to calculate how much the impurity of the training data (i.e. ratio of multiple classes in each leaf) is reduced due to split along a specific feature. For example, if a feature separates a set of samples into subsets with no label overlap, this characteristic would be

considered more important. Furthermore, the criterion is proportional to the amount of samples that are segregated, i.e. the feature which well isolates n samples has less impact than the feature which is able to segregate $n + 1$ samples. Thus, features which determine a conditional split in the top nodes have significantly higher impact on the impurity in the subsequent trees and usually are deemed more dominant [48].

4. Results

4.1. PEBL data classification

The histogram of the average error for 1056 runs in the assessment of SA level 1 is shown in Fig. 4(a). The threshold between the two components of the fitted Gaussian mixture is 0.49 hpx which is used to distinguish *high* and *low* SA. Consequently, 76.2% of the data set are labelled as *high* SA and 23.8% are labelled as *low* SA.

In the assessment of SA level 2, 64.6% of the participants provided two correct answers and their data points are labelled as *high* SA, while the remaining 35.4% are labelled as *low* SA.

The results of the assessment on SA level 3 are shown in Fig. 4 (b). The threshold between the two components of the fitted Gaussian mixture is 37.43° which splits the data set into 62.3% *high* SA and 37.7% *low* SA.

4.2. Component analysis

The PCA is conducted using data from all participants and based on the number of EEG channels. 32 principal components are computed. The three most important components for all participants and a single participant are shown in Fig. 5(a) and (b) respectively.

The most dominant PCA structures in Fig. 5 represent the areas of largest changes in the brain activity. It can be seen that the activation patterns for a single participant differ from the average response, which indicates the effect of individual brain plasticity that can encumber classification accuracy.

ICA is conducted using data from all participants and the obtained features are used for classification. Whereas a ranking of components with respect to explained variance can be constructed for PCA, it is worth noting that no equivalent metric is associated with ICA. Hence, no clear criterion of importance can be used to choose a few dominant components.

Computing cross-correlation between PCA and ICA features shows that a low correlation between ICA and PCA characteristics (maximum value: 0.146, minimum value: -0.059). This means that the information held by ICA and PCA features is different, and hence both feature sets can convey statistical information which is useful to differentiate between high and low SA.

4.3. Outlier detection

The ratio of values identified as outliers for 576 features using FFT and periodogram are shown in Fig. 6 using boxplots. It can be seen that the smallest proportion of samples is identified as outliers when $\Delta = \mu + 2\sigma$ and the highest when $\Delta = 2\bar{x}$. Also, a smaller percentage of samples are identified as outliers when applying periodogram. Since a more uniform distribution of values are identified as outliers using IF, the output of this algorithm is used for further analysis. This also allows to retain a higher percentage of the original data.

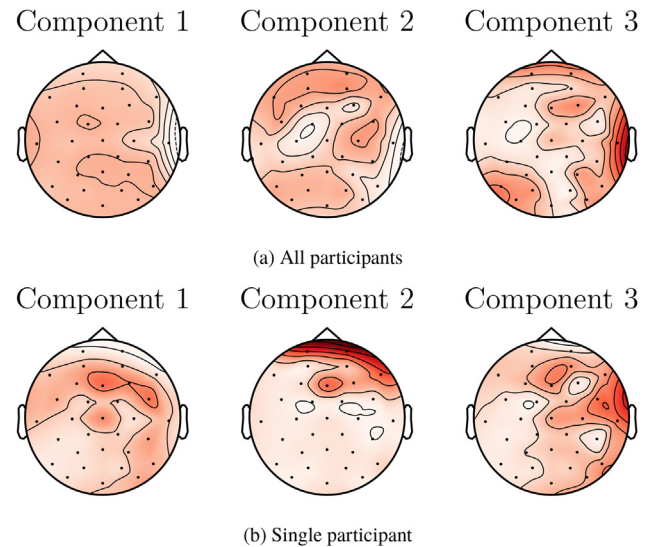


Fig. 5. The magnitude of coefficients in three most important components identified through PCA for (a) all participants and (b) one participant.

4.4. Classification using Random Forest and Boosted Trees

Considering the data containing outliers and the four data sets with imputed outliers, five data sets are used as the input to RF and BT. The data sets are split into a training data set (80%) and a test data set (20%). The training data set is used to create the decision trees for identifying the most significant features. The testing data set is employed to validate the created models.

For BT, three hyperparameters are identified as the most influential: (1) the maximum depth of the decision trees l , (2) the number of estimators M , and (3) the learning rate η , which diminishes the effect of a new decision tree when added to the ensemble model and reduces variance [50]. In order to maximise the test accuracy of the model corresponding to the data set obtained after using periodogram, the parameters are assigned as $M = 100$, $\eta = 0.1$ and $l = 3$.

For RF, (1) the number of trees, (2) the minimum amount of samples in a leaf, (3) the minimum required amount of samples in a leaf to conduct another split and increase the depth, and (4) the number of layers in the trees are examined and only the number of trees is considered as the hyperparameter. The other three parameters are left unconstrained leading to creation of decision trees with high number of nodes. This approach provides a balance between low bias of individual decision trees and variance reduction when the number of decision trees is increased. The number of trees is set to 1000 to maximise the test accuracy.

Table 1 shows the train accuracy, test accuracy, precision and recall for RF and BT, and features identified using FFT and periodogram. Train accuracy and test accuracy describe the percentage of correctly classified samples for the training data set and testing data set respectively. Precision shows the ratio of correctly classified features (true positives, TP) over all retrieved features (including false positives, FP) as in (4):

$$\text{Precision} = \frac{TP}{TP + FP} \quad (4)$$

Recall describes the ratio of correctly classified features (TP) over all relevant items (including false negatives, FN) as in (5):

$$\text{Recall} = \frac{TP}{TP + FN} \quad (5)$$

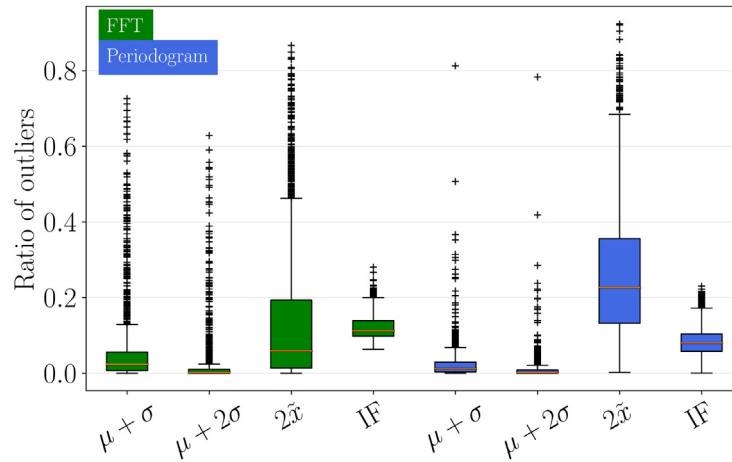


Fig. 6. Percentage of values identified as outliers among 576 features using FFT and periodogram.

Table 1

Train accuracy, test accuracy, precision, and recall for all outlier imputations for BT/RF and periodogram/FFT.

		BT NaN	BT 0	BT μ	BT max x	BT outlier	RF 0	RF μ	RF max x	RF outlier
periodogram	train accuracy	0.901	0.753	0.878	0.772	0.791	1.000	1.000	1.000	1.000
	test accuracy	0.651	0.655	0.670	0.653	0.653	0.635	0.660	0.656	0.663
	precision	0.652	0.644	0.665	0.649	0.646	0.630	0.650	0.656	0.655
	recall	0.807	0.854	0.826	0.826	0.839	0.842	0.848	0.811	0.839
FFT	train accuracy	0.822	0.902	0.771	0.852	0.750	1.000	1.000	1.000	1.000
	test accuracy	0.633	0.627	0.622	0.668	0.653	0.649	0.644	0.649	0.658
	precision	0.629	0.630	0.615	0.652	0.643	0.644	0.638	0.649	0.658
	recall	0.839	0.804	0.863	0.870	0.854	0.832	0.839	0.814	0.807

It can be seen that the highest test accuracy of 67% is achieved for BT with outliers replaced by mean values. This method also features the highest value for precision (66.5%), whereas the highest recall is reported for BT with outliers replaced by zeros (85.4%). The maximum values for test accuracy, precision and recall are slightly lower for features extracted with FFT (66.8%, 65.8%, 87%). The highest reported accuracy is achieved for BT with outliers replaced by a maximal value (66.8%). For RF, all reported test accuracies are lower compared to BT. Additionally, no analysis has been carried out with outliers imputed with NaN values, as they are not supported by the used RF algorithm.

5. Discussion

The division of SA test data into two categories with the threshold set to the intersection of two PDF curves constituting Gaussian mixture resulted in data imbalance where two thirds of samples were classified as *high SA* and one third as *low SA* (see Fig. 4). Since no participants reported any tiredness or fatigue prior to the test, a good overall SA is expected. Fluctuation of SA within a single run can be explained with momentary distractions and shifts in concentration of the subject. For the assessment of SA level 1, the Gaussian mixture curve fits the data set well (error in Cumulative Distribution Function (CDF) is 7.49%), however, the two clusters are not well separated with mean values of $PDF_{High SA}$ and $PDF_{Low SA}$ being close to each other. The Gaussian mixture also fits well the data points corresponding to the assessment for SA level 3 (error in CDF is 5.09%), and the difference between the mean values of the PDFs is large indicating that the data set can be easily separated into two data sets. The approach of assigning labels to the results of the SA level 2's test also show a similar ratio between *high* and *low SA* for the assessment of SA levels 1 and 3.

The data set containing features extracted using a combination of periodogram with window function mostly achieves a lower ratio of outliers compared to the one employing FFT (see Fig. 6). This result is expected since the preprocessing of data using window function is designed to ensure that the resulting PSD is more robust to slight changes of timeseries and consequently is less sensitive to noise. In the methodology where $\Delta = 2\bar{x}$, for some features over 60% of the samples are excluded, drastically reducing the amount of useful information. This can be explained by examining the distributions of features which in most cases are similar to log-normal distribution, where the median is always smaller (sometimes significantly) than the mean.

Despite artefacts being known to have a high impact on a broad range of applications when EEG measurements are used, the outliers do not impact the accuracy of the classification in this study. This is due to the fact that models based on decision trees are significantly less sensitive to high absolute values both during training and testing, when compared to methods where the model contains terms directly proportional to the input (e.g. logistic regression). In models based on decision trees, the output is decided only based on the condition of an input being greater than a certain threshold value. Thus, for example, replacing a value in the data set with another value at the upper range of the distribution, e.g. the maximum of the retained samples, results in the same output as the original high magnitude input.

It is notable that despite the train accuracy of 100% for classification with RF (see Table 1), the model gives a comparable performance on the test set to BT algorithm for example (train accuracy between 75 – 90%). This is because despite each constituent decision tree overfitting the data, once the average results based on “vote of majority” is used, the output is more generalisable. Furthermore, models with alternative values of hyperparameters do not provide better results.

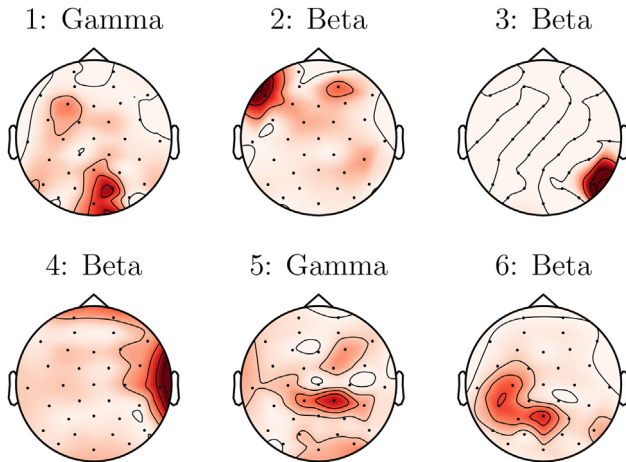


Fig. 7. The magnitude of coefficients in spatial structures corresponding to six statistically most significant features in the classifier with the highest test accuracy.

Comparing our results of the accuracy of the classifier to similar studies such as [29], presenting a classification in the context of teleoperation of human-swarm teaming, might conclude that our accuracy of 67% seems not very high. However, looking closer into the classification, the methodology is fundamentally different: Whereas we have solely included objective measurements such as position and angular error values, the study in [29] combines objective and subjective data (from a SAGAT questionnaire) to identify low and high SA for borderline cases. Hence, we applied the methodology using the Linear Discriminant Analysis (LDA) for classification as proposed in [29] to our dataset using solely objective measurements to be able to compare accuracy results. It should be noted that this involved rejection of all features based on ICA and PCA. We also have used “PO3” and “PO8” channel locations instead of “O1” and “O2” as the latter were not among our measurements. The accuracy achieved following methodology in [29] was 61.5%, therefore the methodology proposed here outperforms similar studies when considering objective data only.

The obtained RF model gives us insights into the most influential features. The identified six most important features are shown in Fig. 7. Four of these features lie in the β band constituting of higher frequencies (12–30 Hz) which indicates an increase of attention and alertness of the subject with high activation [51]. The other presented features lie in γ band (30–45 Hz), responsible for cognitive functioning and information processing [52]. Fig. 7 shows that both left (components 2 and 6) and right hemisphere (components 1, 3, 4, and 5) are responsible for SA.

Also, all four main lobes, frontal lobe (component 2), parietal lobe (components 5 and 6), occipital lobe (components 1 and 3), and temporal lobe (component 4) are activated in the presented features in Fig. 7. All of the identified brain regions are known to carry out tasks needed to acquire SA. The right hemisphere is known to positively correlate with vigilance for simple tasks [53], and intuition [54] whereas logical and linear thinking are attributed to the left hemisphere [55], however, results on brain lateralisation are not universally agreed on [56]. The frontal lobe in the brain is, among others, responsible for concentration, spatial abilities, and short-term memory, while tasks carried out in the parietal lobe include spatial and visual perception, and memory tasks [57]. The occipital lobe interprets vision, e.g. colour, light, and movement, and the temporal lobe sequences and organises visual and auditory input and carries out reasoning tasks [57].

Looking at a similar study [27], loss of SA showed a strong activation of neurons in the orbitofrontal cortex. The frequency band is not stated in this paper. Using a 128-channel GSN EEG sensors, 7

electrodes are placed on the orbitofrontal cortex which are prone to artefacts induced by eye movement or blinking. Also, semi-wet electrodes allow for higher quality signal collection in comparison to dry electrodes. We did not obtain similar results in our study which can be explained by using dry electrodes and having two electrodes covering the areas close to the orbitofrontal cortex (FP1 and FP2).

In [29], activation in θ , α , and β band decreased with loss of SA. This is in line with the results of our study, as we found a correlation between β band and SA. An earlier investigation [28] also revealed a positive correlation between the power spectrum in the θ band and SA, however, their sample size with eight participants is small and no brain regions are identified.

6. Conclusions and future work

A novel analytical methodology for correlating physiological signals using EEG to Situational Awareness was presented in this paper. A new data set from 32 participants completing the SA test in the PEBL framework was collected using a 32-channel dry-EEG headset. A correlation has been found between the β (12–30 Hz) and γ (30–45 Hz) frequency bands and SA. The observed activation of neurons occurred in the four main lobes of the brain (frontal, parietal, occipital and temporal cortex). The combination of these frequencies and brain regions are responsible for concentration and visuo-spatial abilities, which are known to be important in order to build up SA. The highest achieved accuracy of the classifier is 67%.

In future work, a higher number of participants will be invited to conduct the experiment in order to construct separate classifiers for each of SA levels. Thus, the impact of the amount of SA for each level could also be assessed. It was shown in [58–60] that combination of eye-tracking and EEG can increase classification accuracy in multiple contexts similar to SA. Consequently, the alternative data set, which will include measurements from both sensors, will be used in future study. This especially will help to understand the perception of the subject, but is prone to the look-but-failed-to-see phenomenon [61]. Furthermore, the proposed method relies on the assumption that data can be split into *high* and *low* SA using a Gaussian mixture. Other methods incorporating different underlying distributions need to be explored that can be used to divide SA test results into larger number of classes to differentiate between degree of SA levels. Finally, different machine learning algorithms, such as Support Vector Machine and neural networks, could be assessed for improving the classification accuracy.

CRedit authorship contribution statement

Jan Luca Kästle: Methodology, Software, Formal analysis, Investigation, Data curation, Writing - original draft, Visualization. **Bani Anvari:** Conceptualization, Methodology, Investigation, Resources, Writing - review & editing, Visualization, Supervision, Project administration, Funding acquisition. **Jakub Krol:** Methodology, Software, Formal analysis, Data curation, Writing - review & editing, Visualization. **Helge A Wurdemann:** Conceptualization, Methodology, Resources, Supervision, Writing - review & editing, Visualization, Project administration, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This work is supported by the Engineering and Physical Sciences Research Council (Grant Nos. EP/R037795/1 and EP/S01800X/1). The authors also express appreciation to the g.tec's team, especially Emlyn Clay and Matt Pitkin for their fantastic support with the equipment.

This publication is using a new dataset which is available at: <https://rdr.ucl.ac.uk/>. Please contact b.anvari@ucl.ac.uk for any further inquiries regarding the data.

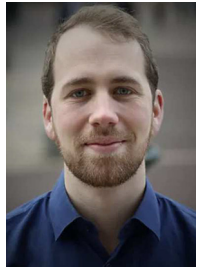
References

- [1] M.R. Endsley, Direct measurement of situation awareness: Validity and use of SAGAT, *Situational Awareness*, Routledge (2017) 129–156.
- [2] V.D. Hopkin, *Human Factors in Air Traffic Control*, CRC Press, 2017.
- [3] S. Samuel, A. Borowsky, S. Zilberstein, D.L. Fisher, Minimum time to situation awareness in scenarios involving transfer of control from an automated driving suite, *Transportation Research Record* 2602 (1) (2016) 115–120.
- [4] M.R. Endsley, Toward a theory of situation awareness in dynamic systems, *Human Factors* 37 (1) (1995) 32–64.
- [5] M.R. Endsley, Situation awareness global assessment technique (SAGAT), in: *Proceedings of the IEEE 1988 National Aerospace and Electronics Conference, IEEE, 1988*, pp. 789–795.
- [6] A.P. Van Den Beukel, M.C. Van Der Voort, 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, 2013, pp. 2000–2005..
- [7] F. Wulf, M. Rimini-Döring, M. Arnon, F. Gauterin, Recommendations supporting situation awareness in partially automated driver assistance systems, *IEEE Transactions on Intelligent Transportation Systems* 16 (4) (2014) 2290–2296.
- [8] K. Popper, *The Logic of Scientific Discovery*, Routledge, 2005.
- [9] M. Feidakis, A review of emotion-aware systems for e-learning in virtual environments, in: *Formative assessment, learning data analytics and gamification*, Elsevier, 2016, pp. 217–242..
- [10] U. Lundberg, R. Kadefors, B. Melin, G. Palmerud, P. Hassmén, M. Engström, I.E. Dohns, Psychophysiological stress and EMG activity of the trapezius muscle, *International Journal of Behavioral Medicine* 1 (4) (1994) 354–370.
- [11] M.V. Villarejo, B.G. Zapirain, A.M. Zorrilla, A stress sensor based on Galvanic Skin Response (GSR) controlled by ZigBee, *Sensors* 12 (5) (2012) 6075–6101.
- [12] T. Åkerstedt, M. Gillberg, Subjective and objective sleepiness in the active individual, *International Journal of Neuroscience* 52 (1–2) (1990) 29–37.
- [13] W. Karlen, C. Mattiussi, D. Floreano, Sleep and wake classification with ecg and respiratory effort signals, *IEEE Transactions on Biomedical Circuits and Systems* 3 (2) (2009) 71–78.
- [14] H.B. Lee, J.S. Kim, Y.S. Kim, H.J. Baek, M.S. Ryu, K.S. Park, The relationship between HRV parameters and stressful driving situation in the real road, in: *2007 6th International Special Topic Conference on Information Technology Applications in Biomedicine, IEEE, 2007*, pp. 198–200.
- [15] F. Akbar, G. Mark, I. Pavlidis, R. Gutierrez-Osuna, An empirical study comparing unobtrusive physiological sensors for stress detection in computer work, *Sensors* 19 (17) (2019) 3766.
- [16] M. Douou, A. Bouabdallah, V. Charfaoui, A light on suitability of physiological sensors for efficient driver drowsiness detection system, *Sensors & Transducers* 224 (8) (2018) 39–50.
- [17] K. Hänsel, R. Poguntke, H. Haddadi, A. Alomainy, A. Schmidt, What to put on the user: Sensing technologies for studies and physiology aware systems, in: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, 2018*, pp. 1–14.
- [18] M. Teplan, Fundamentals of EEG measurement, *Measurement Science Review* 2 (2) (2002) 1–11.
- [19] G. Borghini, L. Astolfi, G. Vecchiato, D. Mattia, F. Babiloni, Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness, *Neuroscience & Biobehavioral Reviews* 44 (2014) 58–75.
- [20] S. Ahn, T. Nguyen, H. Jang, J.G. Kim, S.C. Jun, 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.
- [21] S. Barua, M.U. Ahmed, C. Ahlström, S. Begum, Automatic driver sleepiness detection using EEG, EOG and contextual information, *Expert Systems with Applications* 115 (2019) 121–135.
- [22] S.W. Savage, D.D. Potter, B.W. Tatler, Does preoccupation impair hazard perception? a simultaneous EEG and eye tracking study, *Transportation Research Part F: Traffic Psychology and Behaviour* 17 (2013) 52–62.
- [23] A. Keiper, The age of neuroelectrics, *The New Atlantis* 1 (11) (2006) 4–41.
- [24] K. Kaida, M. Takahashi, T. Åkerstedt, A. Nakata, Y. Otsuka, T. Haratani, K. Fukasawa, Validation of the karolinska sleepiness scale against performance and EEG variables, *Clinical Neurophysiology* 117 (7) (2006) 1574–1581.
- [25] S. Hasanzadeh, B. Esmaeili, M.D. Dodd, Examining the relationship between construction workers' visual attention and situation awareness under fall and tripping hazard conditions: using mobile eye tracking, *Journal of Construction Engineering and Management* 144 (7) (2018) 04018060.
- [26] R. McKendrick, R. Parasuraman, R. Murtza, A. Formwalt, W. Baccus, M. Paczynski, H. Ayaz, Into the wild: neuroergonomic differentiation of hand-held and augmented reality wearable displays during outdoor navigation with functional near infrared spectroscopy, *Frontiers in Human Neuroscience* 10 (2016) 216.
- [27] D. Catherwood, G.K. Edgar, D. Nikolla, C. Alford, D. Brookes, S. Baker, S. White, 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.
- [28] C. Berka, D.J. Levendowski, G. Davis, M. Whitmoyer, K. Hale, S. Fuchs, Objective measures of situational awareness using neurophysiology technology, *Augmented Cognition: Past, Present and Future* (2006) 145–154.
- [29] R.F. Rojas, E. Debie, J. Fidock, M. Barlow, K. Kasmarik, S. Anavatti, M. Garratt, H. Abbass, Encephalographic assessment of situation awareness in teleoperation of human-swarm teaming, *International Conference on Neural Information Processing*, Springer (2019) 530–539.
- [30] C. Zhao, M. Zhao, J. Liu, C. Zheng, Electroencephalogram and electrocardiograph assessment of mental fatigue in a driving simulator, *Accident Analysis & Prevention* 45 (2012) 83–90.
- [31] M.A. Lopez-Gordo, D. Sanchez-Morillo, F.P. Valle, Dry EEG electrodes, *Sensors* 14 (7) (2014) 12847–12870.
- [32] Y.-K. Wang, S.-A. Chen, C.-T. Lin, An EEG-based brain-computer interface for dual task driving detection, *Neurocomputing* 129 (2014) 85–93.
- [33] Y.M. Chi, Y. Wang, Y.-T. Wang, T.-P. Jung, T. Kerth, Y. Cao, A practical mobile dry EEG system for human computer interfaces, *International Conference on Augmented Cognition*, Springer (2013) 649–655.
- [34] F. Wang, H. Wang, R. Fu, Real-time ECG-based detection of fatigue driving using sample entropy, *Entropy* 20 (3) (2018) 196.
- [35] S.T. Mueller, B.J. Piper, The psychology experiment building language (PEBL) and PEBL test battery, *Journal of Neuroscience Methods* 222 (2014) 250–259.
- [36] Y. González-Giraldo, R.E. González-Reyes, S.T. Mueller, B.J. Piper, A. Adan, D.A. Forero, Situation awareness performance in healthy young adults is associated with a serotonin transporter gene polymorphism, *Psychological Reports* 121 (5) (2018) 877–891.
- [37] M.R. Endsley, D.J. Garland, *Situation Awareness Analysis and Measurement*, CRC Press, 2000.
- [38] A. Hyvärinen, Independent component analysis: recent advances, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 371 (1984) (2013) 20110534.
- [39] S. Wold, K. Esbensen, P. Geladi, Principal component analysis, *Chemometrics and Intelligent Laboratory Systems* 2 (1–3) (1987) 37–52.
- [40] S. Sanei, J.A. Chambers, *EEG Signal Processing*, John Wiley & Sons, 2013.
- [41] T.-P. Jung, S. Makeig, M. Stensmo, T.J. Sejnowski, Estimating alertness from the EEG power spectrum, *IEEE Transactions on Biomedical Engineering* 44 (1) (1997) 60–69.
- [42] C. Michel, D. Lehmann, B. Henggeler, D. Brandeis, Localization of the sources of EEG delta, theta, alpha and beta frequency bands using the FFT dipole approximation, *Electroencephalography and Clinical Neurophysiology* 82 (1) (1992) 38–44.
- [43] J.W. Cooley, J.W. Tukey, An algorithm for the machine calculation of complex Fourier series, *Mathematics of Computation* 19 (90) (1965) 297–301.
- [44] P. Stoica, R.L. Moses, *Spectral analysis of signals*, Upper Saddle River, New Jersey, 2005..
- [45] G. Heinzel, A. Rüdiger, R. Schilling, Spectrum and spectral density estimation by the Discrete Fourier transform (DFT), *Tech. rep.*, Max-Planck Institute for Gravitational Physics (2002)..
- [46] A.V. Oppenheim, J.R. Buck, R.W. Schaffer, *Discrete-Time Signal Processing*, vol. 2, Prentice Hall, Upper Saddle River, NJ, 2001.
- [47] F.T. Liu, K.M. Ting, Z.-H. Zhou, Isolation forest, in: *2008 Eighth IEEE International Conference on Data Mining, IEEE, 2008*, pp. 413–422.
- [48] L. Breiman, Random forests, *Machine Learning* 45 (1) (2001) 5–32.
- [49] T. Chen, C. Guestrin, Xgboost: A scalable tree boosting system, in: *Proceedings of the 22nd ACM sigkdd International Conference on Knowledge Discovery and Data Mining, 2016*, pp. 785–794.
- [50] J.H. Friedman, Stochastic gradient boosting, *Computational Statistics & Data Analysis* 38 (4) (2002) 367–378.
- [51] A.T. Pope, E.H. Bogart, D.S. Bartolome, Biocybernetic system evaluates indices of operator engagement in automated task, *Biological Psychology* 40 (1–2) (1995) 187–195.
- [52] P.A. Abhang, B.W. Gawali, S.C. Mehrotra, Technical aspects of brain rhythms and speech parameters, *Introduction to EEG-and Speech-Based Emotion Recognition* (2016) 51–79.
- [53] W.S. Helton, J.S. Warm, L.D. Tripp, G. Matthews, R. Parasuraman, P.A. Hancock, Cerebral lateralization of vigilance: a function of task difficulty, *Neuropsychologia* 48 (6) (2010) 1683–1688.
- [54] S.M. McCrea, Intuition, insight, and the right hemisphere: Emergence of higher sociocognitive functions, *Psychology Research and Behavior Management* 3 (2010) 1.
- [55] N. Kane, M. Kane, Comparison of right & left hemisphere functions, *Gifted Child Quarterly* 23 (1) (1979) 157–167.
- [56] K.M. Mihov, M. Denzler, J. Förster, Hemispheric specialization and creative thinking: A meta-analytic review of lateralization of creativity, *Brain and Cognition* 72 (3) (2010) 442–448.

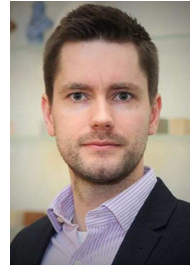
- [57] M.C.H.I. materials, Anatomy of the brain (2018). <https://mayfieldclinic.com/pe-anatbrain.htm>.
- [58] W.-L. Zheng, B.-N. Dong, B.-L. Lu, Multimodal emotion recognition using EEG and eye tracking data, in: 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, IEEE, 2014, pp. 5040–5043.
- [59] W.-L. Zheng, B.-L. Lu, A multimodal approach to estimating vigilance using EEG and forehead EOG, *Journal of Neural Engineering* 14 (2) (2017) 026017.
- [60] F. Putze, J. Hild, R. Kärger, C. Herff, A. Redmann, J. Beyerer, T. Schultz, Locating user attention using eye tracking and EEG for spatio-temporal event selection, in: *Proceedings of the 2013 International Conference on Intelligent User Interfaces*, 2013, pp. 129–136.
- [61] D. Crundall, E. Crundall, D. Clarke, A. Shahar, Why do car drivers fail to give way to motorcycles at t-junctions?, *Accident Analysis & Prevention* 44 (1) (2012) 88–96.



Dr Jakub Krol have received his BSc & MSc degrees in Aeronautical Eng. in 2013 at Imperial College London. Subsequently, he has obtained his PhD degree in 2017 in the same department with the research topics focusing on applications of different dimensional reduction techniques to dynamic estimation & reconstruction of bluff body flows. After brief stint as a quantitative analyst in proprietary trading firm, Jakub has decided to go back to academia and currently works as a Research Fellow in the Intelligent Mobility Group/Lab at UCL (IM@UCL). His primary research interests include data-driven dynamical modelling, estimator design & applications of machine learning to engineering problems.



Jan Luca Kästle received his BSc and MSc in aerospace engineering from University of Stuttgart in 2016 and 2018, specialising in systems engineering and human-machine interaction. He is currently a PhD student in the Intelligent Mobility Group/Lab at UCL (IM@UCL), conducting research on objective quantification of situational awareness in highly automated driving environments. His main research interests are data analytics, computational neuroscience, human factors, and mobility environments including highly automated vehicles.



Helge Wurdemann is Associate Professor of Robotics and leads the UCL Soft Haptics and Robotics Lab at UCL Mechanical Engineering. His research interest focuses on the hardware design and application of innovative stiffness-controllable mechanisms as well as the combination of Artificial Intelligence with control strategies. He has authored over 100 papers published in high-impact journals, leading the field of robotics, and peer-reviewed full-length conference papers, presented at premium robotics conferences.



Dr Bani Anvari is a Lecturer (Assistant Professor) in Intelligent Mobility at UCL. She is the research leader of the Intelligent Mobility Group and Director of IM Lab at UCL (IM@UCL). She is interested in the complexity of interactions and how intuitive, intelligent, human-centric and personalised control systems can influence it. Her research focus has been on Intelligent Mobility, with an emphasis on exploring driver and pedestrian interaction with semi- and fully-autonomous vehicles in different contexts (e.g., shared spaces, CAVs), an area of application that significantly benefits from Robotics and AI.