Journal of Computing in Civil Engineering

Wayfinding Information Cognitive Load Prediction based on Functional Near-Infrared Spectroscopy (fNIRS)

--Manuscript Draft--

Manuscript Number:	
Full Title:	Wayfinding Information Cognitive Load Prediction based on Functional Near-Infrared Spectroscopy (fNIRS)
Manuscript Region of Origin:	UNITED STATES
Article Type:	Technical Paper
Manuscript Classifications:	Artificial intelligence and machine learning; Disaster warning systems; Forecasting; Human-computer interaction; Stochastic processes
Funding Information:	National Science Foundation (1937878) — Dr. Jing Du National Institute of Standards and Technology (60NANB18D152) — Qi Zhu
Abstract:	Amid the rapid development of building information technologies, wayfinding information has become more accessible to building users and first responders. As a result, a realistic risk of cognitive load related to the wayfinding information processing starts to emerge. As cognition-driven adaptive wayfinding information systems become increasingly captivated to overcome challenges of cognition overload due to overwhelming information, a practical and non-invasive method to monitor and predict cognitive loads during the processing of wayfinding information is needed. This paper introduces a Functional Near-Infrared Spectroscopy (fNIRS) based method to identify cognition load related to wayfinding information processing. It provides a holistic fNIRS signal analytical pipeline to extract hemodynamic response features in prefrontal cortex (PFC) for cognitive load classification and prediction. A human-subject experiment (N=15) based on the Sternberg working memory test was performed to model the relationship between fNIRS features and cognitive load. Personalized models were also evaluated to capture individual difference and identify unique contributing features to each person. The results find that fNIRS-based model can predict cognitive load with satisfactory performance (avg. recall rate 73.41±5.49 percent). The findings also demonstrate that personalized models, instead of universal models, are needed for predicting cognitive load based on neuroimaging data. fNIRS has demonstrated comparable advantages over other neuroimaging methods in cognitive load prediction given its robustness to motor artefacts and the satisfactory predictability.
Corresponding Author:	Jing Du, Ph.D. University of Florida Gainesville, FL UNITED STATES
Corresponding Author E-Mail:	eric.du@essie.ufl.edu
Order of Authors:	Qi Zhu Yangming Shi, Ph.D. Jing Du, Ph.D.
Suggested Reviewers:	
Opposed Reviewers:	
Additional Information:	
Question	**Response**
The flat fee for including color figures in print is $800, regardless of the number of color figures. There is no fee for online only color figures. If you decide to not	No

print figures in color, please ensure that the color figures will also make sense when printed in black-and-white, and remove any reference to color in the text. Only one file is accepted for each figure. Do you intend to pay to include color figures in print? If yes, please indicate which figures in the comments box.	
Authors are required to attain permission to re-use content, figures, tables, charts, maps, and photographs for which the authors do not hold copyright. Figures created by the authors but previously published under copyright elsewhere may require permission. For more information see http://ascelibrary.org/doi/abs/10.1061/9780784479018.ch03. All permissions must be uploaded as a permission file in PDF format. Are there any required permissions that have not yet been secured? If yes, please explain in the comment box.	No
ASCE does not review manuscripts that are being considered elsewhere to include other ASCE Journals and all conference proceedings (see next question for expanded conference proceeding requirements). Is the article or parts of it being considered for any other publication? If your answer is yes, please explain in the comments box below.	No
Is this article or parts of it already published in print or online in any language? ASCE does not review content already published (see next questions for conference papers and posted theses/dissertations). If your answer is yes, please explain in the comments box below.	No
Has this paper or parts of it been published as a conference proceeding? A conference proceeding may be reviewed for publication only if it has been significantly revised and contains 50% new content. Any content overlap should be reworded and/or properly referenced. If your answer is yes, please explain in the comments box below and be prepared to provide the conference paper.	No

ASCE allows submissions of papers that are based on theses and dissertations so long as the paper has been modified to fit the journal page limits, format, and tailored for the audience. ASCE will consider such papers even if the thesis or dissertation has been posted online provided that the degree-granting institution requires that the thesis or dissertation be posted. Is this paper a derivative of a thesis or dissertation posted or about to be posted on the Internet? If yes, please provide the URL or DOI permalink in the comment box below.	No
Each submission to ASCE must stand on its own and represent significant new information, which may include disproving the work of others. While it is acceptable to build upon one's own work or replicate other's work, it is not appropriate to fragment the research to maximize the number of manuscripts or to submit papers that represent very small incremental changes. ASCE may use tools such as CrossCheck, Duplicate Submission Checks, and Google Scholar to verify that submissions are novel. Does the manuscript constitute incremental work (i.e. restating raw data, models, or conclusions from a previously published study)?	No
Authors are expected to present their papers within the page limitations described in <u><i>Publishing in ASCE Journals: A Guide for Authors</u></i>. Technical papers and Case Studies must not exceed 30 double-spaced manuscript pages, including all figures and tables. Technical notes must not exceed 7 double-spaced manuscript pages. Papers that exceed the limits must be justified. Grossly over-length papers may be returned without review. Does this paper exceed the ASCE length limitations? If yes, please provide justification in the comments box below.	Yes
If yes, please provide justification in the comments box below. as follow-up to "Authors are expected to present their papers within the page limitations described in <u><i><a	Dear Editor: Our paper reports a fNIRS neuroimaging analysis study for wayfinding cognitive load prediction. We included many visualizations to introduce the result of the neuroimaging analysis. Because of the need to visualize the results, the manuscript would be longer than 30 pages if counting all figures. The main body, however, is

href="http://dx.doi.org/10.1061/9780784479018" target="_blank">Publishing in ASCE Journals: A Guide for Authors</u></i>. Technical papers and Case Studies must not exceed 30 double-spaced manuscript pages, including all figures and tables. Technical notes must not exceed 7 double-spaced manuscript pages. Papers that exceed the limits must be justified. Grossly over-length papers may be returned without review. Does this paper exceed the ASCE length limitations? If yes, please provide justification in the comments box below. "	within the limit of 30 pages. We would like to request an exemption as we believe that these visualizations will be helpful for reviewers to evaluate our work. Please kindly consider our request. Thank you! Authors.
All authors listed on the manuscript must have contributed to the study and must approve the current version of the manuscript. Are there any authors on the paper that do not meet these criteria? If the answer is yes, please explain in the comments.	No
Was this paper previously declined or withdrawn from this or another ASCE journal? If so, please provide the previous manuscript number and explain what you have changed in this current version in the comments box below. You may upload a separate response to reviewers if your comments are extensive.	No
Companion manuscripts are discouraged as all papers published must be able to stand on their own. Justification must be provided to the editor if an author feels as though the work must be presented in two parts and published simultaneously. There is no guarantee that companions will be reviewed by the same reviewers, which complicates the review process, increases the risk for rejection and potentially lengthens the review time. If this is a companion paper, please indicate the part number and provide the title, authors and manuscript number (if available) for the companion papers along with your detailed justification for the editor in the comments box below. If there is no justification provided, or if there is insufficient justification, the papers will be returned without review.	NA
If this manuscript is intended as part of a Special Issue or Collection, please provide the Special Collection title and	NA

name of the guest editor in the comments box below.	
Recognizing that science and engineering are best served when data aremade available during the review and discussion of manuscripts andjournal articles, and to allow others to replicate and build on workpublished in ASCE journals, all reasonable requests by reviewers formaterials, data, and associated protocols must be fulfilled. If you are restricted from sharing your data and materials, please explain below.	NA
Papers published in ASCE Journals must make a contribution to the core body of knowledge and to the advancement of the field. Authors must consider how their new knowledge and/or innovations add value to the state of the art and/or state of the practice. Please outline the specific contributions of this research in the comments box.	This paper reports a neuroimaging approach (fNIRS) for assessing the real time cognitive load related to the wayfinding information processing. It is particularly tests the scenarios faced by emergency responders, such as firefighters, when information is incomplete and time is limited in the spatial memory development. The findings are expected to inspire the design of cognition-driven information systems to reduce cognitive load, facing the increasing complexity of modern buildings.
When submitting a new and revised manuscript, authors are asked to include a Data Availability Statement containing one or more of the following statements, with specific items listed as appropriate. Please select any of the statements below that apply to your manuscript. Also, please include the selected statements in a separate "Data Availability Statement" section in your manuscript, directly before the acknowledgements or references. The statement(s) listed in your manuscript should match those you select in your response to this question.	a. Some or all data, models, or code generated or used during the study are available in a repository online in accordance with funder data retention policies.
If there is anything else you wish to communicate to the editor of the journal, please do so in this box.	

Wayfinding Information Cognitive Load Prediction based on Functional Near-Infrared Spectroscopy (fNIRS)

Qi Zhu[1], Yangming Shi, Ph.D.[2], Jing Du, Ph.D.[3*]

[1]Ph.D. Student, Engineering School of Sustainable Infrastructure & Environment, University of Florida, 1949 Stadium Road 454A Weil Hall, Gainesville, FL 32611; email: qizhu@ufl.edu

[2]Postdoctoral Research Associate, Engineering School of Sustainable Infrastructure & Environment, University of Florida, 1949 Stadium Road 454A Weil Hall, Gainesville, FL 32611; email: shiyangming@ufl.edu

[3]Associate Professor, Engineering School of Sustainable Infrastructure & Environment, University of Florida, 1949 Stadium Road 460F Weil Hall, Gainesville, FL 32611; email: eric.du@essie.ufl.edu;

Corresponding Author

ABSTRACT

Amid the rapid development of building information technologies, wayfinding information has become more accessible to building users and first responders. As a result, a realistic risk of cognitive load related to the wayfinding information processing starts to emerge. As cognition-driven adaptive wayfinding information systems become increasingly captivated to overcome challenges of cognition overload due to overwhelming information, a practical and non-invasive method to monitor and predict cognitive loads during the processing of wayfinding information is needed. This paper introduces a Functional Near-Infrared Spectroscopy (fNIRS) based method to identify cognition load related to wayfinding information processing. It provides a holistic fNIRS signal analytical pipeline to extract hemodynamic response features in prefrontal cortex (PFC) for cognitive load classification and prediction. A human-subject experiment (N=15) based on the Sternberg working memory test was performed to model the relationship between fNIRS features and cognitive load. Personalized models were also evaluated to capture individual difference and identify unique contributing features to each person. The results find that fNIRS-based model can predict cognitive load with satisfactory performance (avg. recall rate 73.41±5.49 percent). The findings also demonstrate that personalized models, instead of universal models, are needed for predicting cognitive load based on neuroimaging data. fNIRS has demonstrated comparable advantages over other neuroimaging methods in cognitive load prediction given its robustness to motor artefacts and the satisfactory predictability.

KEYWORDS: fNIRS; Cognitive Load; Sternberg Test; Wayfinding

INTRODUCTION

As of 2020, an estimated 20% of U.S. buildings can be classified as “smart” with 50 billion connected devices forming clusters of Internet of Things (IoT) (Memoori 2015). These new sensing devices and technologies will generate a large amount of real-time informatics in built environments. The way a building’s spatial information is presented is becoming more diverse, mixing signage (e.g., directional, identification, information, warning), text instructions, maps (2D, 3D, interactive), tactile and haptic guidance, sound, verbal guidance, colors, landmark markers, and lighting, etc. (Calori and Vanden-Eynden 2015). Owing to the development of computing powers, new technologies such as radio-frequency identification (RFID) (Willis and Helal 2005), indoor positioning systems (Randell and Muller 2001), better telecommunications (Aporta et al. 2005) and augmented reality (AR) technologies (Olsson et al. 2013) are increasingly available to interact with the building data for building wayfinding, i.e., orienting and navigating inside a building (Mulloni et al. 2011).

While real-time data collecting and informatics become easier for building wayfinding, a new challenge has arisen - cognitive overload due to the overwhelming information. Too much information can lead to an increased cognitive load (Eppler and Mengis 2004), causing decreased task performance (Sweller 1988) and/or prejudices in decisions such as stereotyping, i.e., relying on personal experiment instead of facts (Banaji and Greenwald 1994). In emergency wayfinding, for example, first responders often need to build an accurate spatial working memory of unfamiliar spaces in a timely manner (additional time pressure), requiring an intensive retention and processing of received spatial information (such as maps or verbal instructions) (Logie 2014). Yet according to the cognitive load theory (CLT), human cognition has a finite information processing capacity (Sweller 1994). The well-proven Miller's Law shows that in one-dimension judgement tasks (such as temporarily remembering a phone number), an average person can only hold and process 2 to 3 bits of information at one time (Miller 1956). Literature also suggests that mental fatigue due to the prolonged period exposed to excessive information also contributes to increased cognitive load (Tanaka et al. 2015). It further adds to the potential risks of cognitive overload as modern professional works (e.g., engineers and first responders) rely on the processing of spatial information more common and more frequently, given the increasing complexity of modern built environments. Such an apparent gap between the enormous information processing need (both in amount and in time), and the limited processing capacity creates a potentially hazardous situation for certain professionals such as civil engineers and first responders. In the foreseeable future, this gap will be further widened due to the advancement of information technologies and unprecedented accessibility to large amounts of real-time data. Successful information delivery in building wayfinding refers to delivering the right information to the right person at the right time, instead of delivering untailored information to anyone at any time (Fischer 2012). An adaptive information system that is optimized for the real-time and dynamic cognitive load is in an urgent need.

To set the foundation for a cognition-driven adaptive wayfinding information system, this study aims to develop and test a cognitive load monitoring and prediction method triggered by the processing of wayfinding information, such as landmarks, routes and symbolic information related to orientation and target destinations. Specifically, we will test the use of an explicit neurofunctional measurement via Functional Near-Infrared Spectroscopy (fNIRS). Although previous literature has indicated a strong relationship between certain neurophysiological metrics and cognitive load, including pupillary dilation (Matthews et al. 1991; Shi et al. 2020; van der Wel and van Steenbergen 2018) and electroencephalogram (EEG) spectral analysis (e.g., alpha–gamma versus theta–gamma ratio is a golden standard for cognitive load) (Roux and Uhlhaas 2014), fNIRS is more tolerant to motor artefacts and thus is a better candidate for detecting neural activities in wayfinding tasks.

However, two challenges have clouded the feasibility and viability of fNIRS for cognitive load prediction. First, fNIRS is still a relatively new neuroimaging method (Hu and Shealy 2019), and the literature is not clear regarding what features of fNIRS data are strong predictors of cognitive load (Pinti et al. 2018). Second, existing efforts have been concentrated on exploring the universal patterns of fNIRS features related to cognitive load, i.e., knowledge applicable to all people; while in contrast, there might be a significant individual difference pertaining to the fNIRS-cognitive load relationship. This study will narrow the gap by testing a personalized cognitive load prediction method based on fNIRS features extracted from a human-subject experiment. A holistic fNIRS analytical pipeline will be examined to extend the literature of the existing fNIRS analysis method. In addition, we will compare individual differences in terms of fNIRS-based cognitive load prediction and what the unique contributing features are to each person. The remainder of this paper will introduce the theoretical basis and the human-subject experiment led to the cognitive load prediction model.

LITERATURE REVIEW

Spatial Information Types in Wayfinding

Wayfinding in complex buildings requires the acquisition and processing of large amounts of spatial information (Werner et al. 1997). The spatial information people leverage to understand the environment and guild navigation can be hierarchically categorized into three main forms, which are landmarks (L), routes (R), and survey (S, i.e., cognitive maps), i.e., the “LRS” model (Siegel and White 1975). The landmarks are defined as unique and distinctive objects at fixed locations in the environment (Werner et al. 1997). Landmark knowledge pertains to identifying and memorizing landmarks based on their shapes, sizes, colors, and contextual information (Elvins 1997). People can incidentally recognize the landmarks and unconsciously build their landmark knowledge through the navigation in the environment (van Asselen et al. 2006). The process of acquiring landmark knowledge usually does not need much mental efforts. Siegel and White (1975) found that even very young children were able to identify and make use of landmarks in spatial understanding. Route knowledge is a more advanced strategy for spatial knowledge acquisition. Route knowledge is encoded as the knowledge of memorizing the sequences of landmarks or locations from one location to the other location (Siegel and White 1975). People can gain route knowledge either from a map or from a navigation experience (Taylor et al. 1999). The process of developing route knowledge requires a higher mental effort, and thus practices are helpful for the acquisition of route knowledge such as navigating in the environment (Lindberg and Gärling 1983). Lastly, survey knowledge is the process to mentally abstract the configurational information as a map-like representation of the environment (Werner et al. 1997). It is the last and the most mentally demanding spatial knowledge acquisition method. People can gain survey knowledge either from a map study or from real-world exposure to the environment (Elvins 1997). van Asselen et al. (2006) found that route knowledge and survey knowledge coexisted in the process of spatial memory development. They further found that the spatial information representation methods of the environment could significantly affect the process of spatial knowledge acquisition. Subjects who learnt from a map outperformed others in developing the survey knowledge. On the other hand, subjects who learnt by navigating in the environment performed better than others in developing the route knowledge (Taylor et al. 1999; van Asselen et al. 2006).

In summary, evidence has shown that the type of spatial information is a contributing factor to the cognitive load in wayfinding tasks. The LRS model indicates that depending on the phases of spatial knowledge development, landmarks, route and survey information indeed requires different levels of cognitive engagement and leads to different levels of cognitive load. As a result, the experiment design in this study follows the LRS model to test the working memory pertaining to the elements of LRS. But in order to understand how spatial information triggers different levels of cognitive load, it is also necessary to examine the role of different formats of information.

Spatial Information Formats and Cognitive Load

The LRS model specifies that three types of spatial information can lead to different cognitive load levels. For each type of spatial information, it is noted that different ways or formats of display could also play an important role in terms of affecting the cognitive load. For instance, landmarks can be communicated verbally, in written texts, or in visual symbols, with an additional impact on the cognitive load. Cognitive load theory (CLT) (Chandler and Sweller 1991; Kalyuga 2009; Paas et al. 2003; Sweller 1994; Sweller 2010; Sweller et al. 2011) divides the overall cognitive load into three main components: *intrinsic cognitive load* (related to the complexity of tasks), *extraneous cognitive load* (affected by how information is presented) and *germane cognitive load* (devoted to construction of schemas – permanent knowledge about patterns). Among all, the extraneous cognitive load can be easily affected by the formats of information. Baddeley et al. proposed a *dual-coding theory* to describe the internal structure of extraneous cognitive load in relation to the working memory of information processing (Baddeley 1992; Baddeley 2000; Baddeley 2003; Baddeley 2012; Baddeley and Hitch 1974; Miyake and Shah 1999; Moreno and Mayer 2007). They found "dual channels" in human cognition, where different mental activities are activated when people process two distinct categories of information: *phonological information* (i.e., auditory verbal information or visually presented language) and *visuospatial information* (i.e., the visually presented information about objects and space). In addition, a *central executive* cognitive process also takes place to bind information into coherent episodes, shift between tasks or retrieval strategies, and select between attention and inhibition (Sridharan et al. 2008). It is more applicable to explain why there is an increased cognitive load when people process mixed information. The dual channels theory indicates the need to include phonological, visuospatial and mixed information in our experiment design.

In addition, cognitive load has also been found to relate to the mental fatigue levels driven by task duration (Ackerman and Kanfer 2009; Pattyn et al. 2008). Based on a comprehensive literature review, Paas et al. found that the multi-constructs of cognitive load were indeed inter-correlated, including task characteristics (e.g., difficulty), mental efforts and mental fatigue (Paas et al. 2003). Xie and Salvendy proposed that cognitive load should be a physiological measure (the load perceived by the subject) and indicative of the cumulative load over a period of time, rather than instantaneous load, i.e., the mean intensity during the task (Xie and Salvendy 2000). The former one reflects the overall experience of test subjects for the complete task and thus is more aligned with the self-reported mental workload measured by NASA TLX (Xie and Salvendy 2000). Roy et al. found that prolonged task duration could substantially affect the cognitive load in working memory development, and hence, a measurement method that captures the interaction between working memory task difficulty and task duration is needed (Roy et al. 2013). Similar findings have also been reported that task duration is a main contributor to subjective cognitive load (Kessler and Meier 2014; Sandry et al. 2014). As a result, in this study, we consider the holistic cognitive load that is affected by both the information processing needs and the mental fatigue level due to the prolonged task time.

Neurophysiological Measures for Real-Time Cognitive Load Assessment

In wayfinding literature, such as navigation and driving studies, cognitive load related to the wayfinding information processing and decisions is mainly measured as a single entity (Chang and Wang 2010; Haapalainen et al. 2010; Klatzky et al. 2006; Liu et al. 2008). Indicators such as task performance (e.g., errors) and psychometrics (e.g., surveys) are typically used to measure cognitive load in these studies (Bunch and Lloyd 2006; Cabeza and Nyberg 2000; Meilinger et al. 2008; Rossano and Moak 1998; Smith and Hart 2006; Sweller 1988). Although these are important measurements to consider, the lagging nature of the indicators are insufficient to support a real-time monitoring and control mechanism of cognitive load that attempts to distinguish different types of cognitive load (e.g., visual and verbal-driven). A deeper insight into the real-time cognitive load measurement methods that balance the prediction accuracy and mobility required by the wayfinding will enable a fine-tuned mechanism to manipulate and manage real-time cognitive load.

With the recent development of neurophysiological sensing technologies, real-time cognitive load measurement methods have been enriched in the research community (Verney et al. 2001; Wang et al. 2014; Wu et al. 2019). Among the neurophysiological metrics, eye-tracking is the most examined method including visual attention and pupillary data to predict real-time cognitive load (Goldinger and Papesh 2012; Kucewicz et al. 2018; Shi et al. 2020; Shi et al. 2020). de Fockert et al. (2001)'s study found that the ability of visual selected attention is strongly related to the real-time cognitive load level. The higher cognitive load leads to higher visual distractor processing in human brain. Toker et al. (2013) investigated the impact of cognitive load on user's attention patterns when users were exposed to different types of visual graphs (bar graphs and radar graphs). They found that visual attention patterns were changed with task difficulty and visualization methods. At the same time, Pupillometry has a long history in studies of cognitive load and memory (Goldinger and Papesh 2012; Klingner et al. 2011; Kucewicz et al. 2018; Paas et al. 2003). Many studies have suggested that pupil dilation reflects a 'summed index' of neural activity during cognitively demanding tasks (Goldinger and Papesh 2012; Kahneman and Beatty 1966; Papesh et al. 2012). Pupil dilation has been shown to increase under greater cognitive load. As a task becomes progressively more difficult pupil dilation increases.

Despite the wide use of eye tracking for cognitive load measures, evidence recognizes its limitations in a highly dynamic task environment, such as the unexpected impacts of light conditions on pupillary size and an environment with too many distractors (Sharafi et al. 2015). Recently, neurocentral methods have gained popularity in various engineering domains such as engineering decision making (Goucher-Lambert et al. 2017) and hazard detection (Thirunavukkarasu et al. 2016). A representative method is Electroencephalography (EEG) that collects a high-resolution time-domain data for the rapid propagation speed of electric fields of brain activity. Evidence shows that alpha and theta activity is related to cognitive load in a variety of task demands (Antonenko et al. 2010; Gevins and Smith 2003; Sterman et al. 1994). Gevins and Smith (2003) found that alpha and theta oscillations demonstrated different patterns according to the task difficulty and cognitive load. As the cognitive load increases, the alpha activity decreases, and theta activity increases. Antonenko et al. (2010) also found that EEG could provide good resolution data for evaluating individual's cognitive activities during complex learning behaviors. But literature also finds that EEG data can be easily affected by the motor artefacts (Lohani et al. 2019). Recently, another non-invasive optical imaging method i.e., functional near-infrared spectroscopy (fNIRS) has started to show a great potential in cognitive status monitoring and analysis.

fNIRS Method

fNIRS is a novel tool to study and monitor tissue oxygenation changes in the brain non-invasively (Bunce et al. 2006; Du et al. 2020). The attributes of fNIRS like portability, movement tolerability, and safety of use have made it particularly suitable for investigating brain function (Pinti et al. 2018). fNIRS uses the sources (laser or light emitting diode) and detector probes, which are positioned over the scalp surface, to detect the change in optical density caused by the hemodynamic changes mainly expected in the cortical grey matter (Pinti et al. 2018). Then intensity data as a result of the absorption of infrared lights of different waves can be converted into oxy-and deoxygenated signals located in small vessels (<1mm diameter). The activation changes indicate neuronal activity levels as these activities provoke an increasing in oxygen consummation, local cerebral blood flow (CBF) and oxygen delivery (Lloyd-Fox et al. 2010). The measures of cerebral oxygenated and deoxygenated hemoglobin are also correlated with the functional magnetic resonance imaging (fMRI) BOLD signals (Cui et al. 2011). Comparing to other available neuroimaging techniques relying on neurovascular coupling such as fMRI (Ochsner et al. 2002) and positron emission tomography (PET) (Andreasen et al. 1996), and those based on the electromagnetic activity of the brain such as (EEG) (Ray and Cole 1985) and magnetoencephalography (MEG) (Halgren et al. 2000), fNIRS provides a higher temporal resolution. In addition, participants are not restricted to a confined space or are not required to stay in a supine position motionless (Ayaz et al. 2013), thus fNIRS is more robust for monitoring cortical hemodynamics during motor tasks or tasks involving walking given its advantages of low sensitivity to body movements and the systems' portability (Pinti et al. 2018). These features make fNIRS suitable for tracking human brain activities in the complex and real working environment. Previous studies have proven that hemodynamic responses in the prefrontal cortex (PFC) measured by fNIRS can be used to quantify and classify mental workload (Herff et al. 2014). Studies of distinguishing among cognitive states using fNIRS have been applied in various fields such as driving research in the car (Herff et al. 2014; Tsunashima and Yanagisawa 2009) and airplane (Verdière et al. 2018), brain-computer interface (BCI) (Erdoĝan et al. 2019; Hong et al. 2018) and human-robot collaboration (Canning and Scheutz 2013). Given the advantages of fNIRS, this study aims to examine cognitive load monitoring and prediction methods based on fNIRS measures. Especially, various features of fNIRS data are extracted and compared in terms of the predictability pertaining to cognitive load, as well as to the differences among individuals.

FNIRS-BASED COGNITIVE LOAD ANALYTICAL PIPELINE

It is noted that as a relatively new method, there is no generally agreed analytical pipeline for fNIRS data processing. The hemodynamic response could be affected by the basic physiological activities such as heart beats and therefore, fNIRS data must be cleaned before the analysis (Pinti et al. 2018).Unlike brain electrical activities, there is often a lag between the trigger event and the hemodynamic response, usually ranging from 2 seconds to 8 seconds in term of time-to-peak - calculated as the time to maximum response from the presentation of the stimulus (Huppert et al. 2006) . The unique characteristics of fNIRS data have made it nontrivial for the analysis. Based on the latest fNIRS literature (Verdière et al. 2018), we propose an analytical pipeline for predicting cognitive load status based on the dynamic fNIRS data, as illustrated in ***Signal Acquisition***

The placement design of the number and locations of light emitters and detectors (optodes) should be optimized before data collection. The layout of optodes is determined by the specific purpose of a study concentrating on the areas of interest (AOIs) with different neural functions. In this research, we focused on the prefrontal cortex (PFC). PFC is proven to relate to memory and logical thinking (Herff et al. 2014; Ramnani and Owen 2004). The fNIRS data from PFC has been deemed as indicative of human cognitive load in various studies (Faress and Chau 2013; Hong et al. 2015; Naseer and Hong 2013; Noori et al. 2017; Power and Chau 2013; Verdière et al. 2018). In addition, PFC is not directly affected by motor disabilities and its hair-free feature makes it easier to facilitate probe set-up and to get qualified fNIRS signals (Erdoĝan et al. 2019). As a result, PFC is gaining popularity among fNIRS researchers at a rapid rate in recent years (Faress and Chau 2013; Herff et al. 2014; Pinti et al. 2018; Verdière et al. 2018). Usually, there is a tradeoff between the number of channels and the brain areas required to cover. In some cases, due to the limited number of the channels the device provided, researchers only focus on one or two cortices, which should be including both left and right brain portions. Besides, accurate acquisition of fNIRS signals would require considering the sample rate. Most fNIRS systems nowadays support at least 1 Hz sampling frequency, which is enough for feature extraction. Researchers can also adjust the sampling rate for a more efficient calculation (Lohani et al. 2019). The length of the emitting light also varies across the devices. Usually, most device include two wave lengths which help differentiate CO_2 and O_2 absorption separately.

In this study, an 18-channel fNIRS device – NIRx Sport 2 (NIRx, Berlin) with the sample rate of 7.8125 Hz was used to collect hemodynamic activation levels of PFC, as illuminated in **Fig.2**. Each source emits two near-infrared wavelengths (760 nm and 850 nm) to detect and differentiate between oxygenated and deoxygenated hemoglobin.

Data Pre-processing

The purpose of running a data pre-processing is to remove secondary noises from the raw data and convert the intensity time-series into concentration changes of oxygenated hemoglobin (ΔHbO) and deoxygenated hemoglobin (ΔHbR). The raw intensity data from qualified channels are converted into optical densities (ΔOD). Secondary noises not related to the task-evoked hemodynamic activities should be removed to extract useful information from the OD data. There are two typical kinds of noise identified by previous studies. The first one is physiological noises at specific frequencies associated with heartbeats (~1 Hz), breath (~0.3 Hz), Mayer waves (~0.1 Hz) and very low frequency (< 0.04, VLF) oscillations (Pinti et al. 2018). Digital filtering such as high-pass, low-pass and band-pass filters is a common tool to clean these noises in a specific frequency range. The second noise is related to human motions, such as body movements. Even fNIRS is relatively robust to motion (Pinti et al. 2018), a sufficient data cleaning process to remove artifacts is required. Otherwise, the possible optical decoupling between the optode and the scalp (Cooper et al. 2012) will lead to instantaneous or slow changes in the signal that do not reflect changes in physiology (von Lühmann et al. 2020), and makes the fNIRS data uninterpretable. To tackle this type of noise, motion artifact detection and removal or correction algorithms such as principle component analysis (PCA), spline interpolation, and wavelet analysis are commonly used in fNIRS studies (Brigadoi et al. 2014; Cooper et al. 2012). Both frequency range and artifact detection algorithms should be chosen carefully in order to preserve the stimulus frequency and the task-evoked response (Pinti et al. 2018). In the end, fNIRS optical densities are converted to concentration changes of ΔHbO and ΔHbR by modified Beer-Lambert law with a partial path length factor of 6 (Delpy et al. 1988).

For this study, we used the raw light intensity fNIRS data form six channels covering PFC as shown in **Fig.2** with the sample rate at 7.8125 Hz of each participant. A band-pass filter with the frequency rate between 0.01 Hz and 0.5 Hz was used to remove the influences from physiological changes. After identifying and remove the motion artifacts through hmrMotion Artifact algorithm (Huppert et al. 2009), the processed data were calculated according to Modified Beer-Lambert Law method (Delpy et al. 1988) and output the HbO and HbR concentration variations. For further analysis, we only used ΔHbO2 data of each channel as indicators of changes in regional cerebral blood volume as it is the more sensitive indicator of changes in cerebral blood flow (CBF) (Hoshi et al. 2001).

Feature Processing and Classification

In the processing phase, the pre-processed ΔHbO was used to perform statistical analyses. *mean oxygen consumption* (mean), *standard deviation of oxygen consumption* (std), *peak oxygen consumption* (peak), *area under curve* (AUC), *kurtosis of oxygen consumption distribution* (Kurt) and *Skewness of oxygen consumption* (Skew) are proven to be sensitive to the changes in levels of hemodynamic activities by the existing fNIRS literature, as listed in Table 1. For continuous fNIRS data, literature uses different sliding window sizes to sample training data from a continuous data stream, ranging between 2 seconds and 60 seconds (Bauernfeind et al. 2014; Erdoĝan et al. 2019; Faress and Chau 2013; Hong et al. 2018; Hong et al. 2015; Khan et al. 2019; Verdière et al. 2018; Zhang and Zhu 2019). Extracted features from different channels are usually combined into a limited number of regions of interest (ROI) to reduce the amount of data and the data dimensionality (Verdière et al. 2018). The combination should average data from a similar functional cortex on the same side. Because fNIRS signals usually vary dynamically across different subjects, personalized feature normalization is required to standardize the extracted features for each subject into the same scale and to improve feature interpretability across subjects (Hosseini et al. 2018). Normalized feature sets are then fed into the machine learning model to train the classifier – support vector machine (SVM) in this study. The SVM classifier aims to discriminate different classes by creating an optimum hyperplane. This supervised learning model is widely used in the classification and prediction of neural activity studies due to its simplicity and computational efficiency (Gateau et al. 2015; Naseer and Hong 2015). Although SVM is naturally linear, for datasets with small numbers of features and training examples like the data in this study, a non-linear SVM with a kernel function would be more efficient. We selected the Gaussian radial basis function (RBF), which is efficient by allowing complex separation surface requiring a reduced number of hyper-parameters to tune (Hong et al. 2018), as the kernel to implement our SVM model, as shown in Eq (1):

$$\mathrm{K}(x, x') = \exp(-\gamma\|x - x'\|^2) \ldots \text{Eq (1)}$$

where $||x - x'||^2$ is the squared Euclidean distance between the two feature vectors, γ is the parameter for adjusting the goodness of fit of the RBF model. The one-vs-one scheme (Galar et al. 2011) was used to handle the multi-class classification. It is supposed to be robustness for fNIRS classification studies.

We used the six channels of $\Delta HbO2$ data from PFC to extract features. A sliding window with size 30s with interval 5s was used to split data and obtain block time series for each level trails per subject. The Mean, Peak, Stand Deviation, Kurtosis and Skewness were estimated according to the formulas shown in **Table 1**. respectively. The AUC was calculated as the sum of the absolute values of the signal (Verdière et al. 2018). We defined two regions of interest (ROI)-Left and Right PFC, each including three channels. The features of each channel were aggregated by type in the same ROI. The dimensionality was thus reduced from 18 to 6 per ROI. To bring all feature variables down to a similar scale, the Z-scoring method (Jain et al. 2005) was used to normalize data. As a result, we obtained normalized feature data with six dimensions in each ROI for each of the cognitive load levels per participant. To increase the accuracy of the classifier, we used the grid search method (Bergstra and Bengio 2012) and K-Folds method (k=10) cross-validation to get the best-fit parameters for the model. Then the optimized classifier was applied for training and testing.

In order to test if there are significant individual differences, we will compare model predictability among different subjects. We will also compare how the contributing features vary among people. If it is found that the predictability and/or important features are different among subjects, it suggests fNIRS-cognition relationships are different.

HUMAN-SUBJECT EXPERIMENT

Experiment Setup

A human-subject experiment was conducted based on the Sternberg working memory test, which triggers cognitive load related to the processing of wayfinding related information elements, including the recognition of given information, retention and retrieval phases (Sternberg 1969). The reason we selected the Sternberg test was that it is among the most popular, controlled test to examine mental load related to information processing and working memory development (Ref). This study focuses on the cognitive load prediction in the phase of spatial information processing, and thus, Sternberg test is deemed to be an effective approach for data collection.

Based on the literature review, we designed three categories of information elements in the Sternberg test to reflect different cognitive requirements in processing wayfinding information. **Fig.3** illustrates the specific design of the Sternberg test. For category 1, we used phonological information (letters) to show landmark information. According to the LRS model and the dual channels theory, such information requires the lowest cognitive engagement, and thus it is marked as the "light cognitive load" group. For category 2, we used visuospatial information to show route and survey information. In such a way the cognitive engagement is higher, and we labeled it as the "medium cognitive load" group. Finally, we combined the wayfinding information elements, including mixed phonological information and visuospatial information to indicate landmarks, routes and survey information. It requires participants to process all wayfinding information types and triggers the central executive function according to the cognitive load theory. As a result, we marked it as the "heavy cognitive load" group. Each group includes 40 trials for eventually distributed tests. To be noted, a challenging task to all cognitive load studies is to get the ground truth of cognitive load. Self-reported mental load, such as NASA TLX can be easily affected by the subjective feeling biased toward the newer experience, such as the later phase of an experiment. In addition, it interrupts the continuous flow of information processing which is the real-life experience of most people, such as firefighters in search and rescue tasks and who must process spatial information continuously without any interruptions. As a result, recent literature has started to manipulate the difficulty levels of the task or the information to trigger varying levels of cognitive load, and to use them as the ground truth(Haapalainen et al. 2010) . This study follows the second approach.

Each trial started with the presentation of a list of patterns to recognize (encoding period; 2s in total), followed by a memory retention period (2s) during which the subject must maintain the list of patterns in memory, and a retrial period (2s) in which the subject had to answer whether probe pattern had been displayed before by clicking the corresponding buttons. Between the town sessions was a short fixation phase (0.5s).

As discussed earlier, this study tracks the cognitive load that captures both the information processing needs and the increasing mental fatigue level to better reflect the real-world scenarios (such as predicting a firefighter's increasing risks of cognitive overload over time). **Table 2** lists the use of three levels of task difficulty (related to working memory encoding) as the basis for varying cognitive load levels. The encoding patterns were designed related to wayfinding information, including symbolic information, orientations, words and mixed information.

The experiment consisted of three sessions: (1) preparation session, (2) training session, (3) task performing session. The preparation session (5-10 minutes) was designed to allow participants to familiarize the procedure and potential benefit or risk of the experiment. Participants' demographical information including age, gender, major, degree level was also collected in this session. The participants were guided to familiarize themselves with the fNIRS system. In this session, all the participants were first instructed to wear on the fNIRS device, and the investigators were able to ensure each probe of the fNIRS device accurately collected the neuroimaging data from the target brain regions. The participants were asked to stay claim in a chair and let the fNIRS device to set up the baseline data for each participant. Before starting to collect fNIRS data, we asked subjects to relax themselves with closed eyes and an empty mind for 3 minutes to remove possible hemodynamic responses resulting from previous activities. In the training session, a pre-Sternberg test, which provided the same procedure as the formal test in task performing session with different patterns, was used to train and test subjects for performing the task in the correct way. During the task performing session, each subject was asked to sit in front of a moister screen, which only displayed the test contents without anything else. The environment kept quiet and in the same light condition during the whole experiment to eliminate the additional distractions.

Participants

We recruited 15 subjects (1 Women, mean age = 24 ±2) to participate in the experiment (**Fig.4**). One of the participants' data was excluded from the analysis due to strong noise. All the participants did not have a history of any psychiatric or neurologic abnormalities. We gave the experiment instruction with details to all participants before they signed the informed consent. All the experiments were done at the same location (Francis Hall Room 101 - BIM CAVE at Texas A&M University), with the same devices. The environmental effects can be ruled out as well.

fNIRS Data Pre-Processing Results

After pre-processing, raw intensity signals of each subject was converted into optical density (OD) and then concentration changes of oxyhemoglobin (ΔHbO2). As shown in **Fig.5**, the band-pass filter removed very low frequency systemic fluctuations the respiration and the heart-rate oscillations. Each channel was screened to detect and correct motion artifacts. In this way, the possible noise from global systemic effects originating in superficial layers of scale, dura, and peripheral vasculature but not in the cortex was removed. Then the channels with stability value were retained for further analysis. **Fig.6** illustrates the grand averages of the hemodynamic response under three levels of cognition load of all subjects. From a visual inspection, the mean values of ΔHbO2 increased from level 1 to level 3. Besides, the curve of level 1 is smoother than level 3. These indicates that the hemodynamic responses are different among cognition load levels and therefore can be used for further analysis of classifying and predicting cognition status.

Classification Performance

The first aim of the classification is to detect whether the SVM classifier can distinguish three-level trials for both the individual and group levels. We classified both individual level feature sets (i.e., the training and testing datasets were from each participant data separately) and group level feature sets (i.e., the training and testing datasets were from aggregated data from a total of fourteen participants). **Fig. 7** illustrates the paired feature examples for the individual (participant 4) and group level (total participants). After the feature extraction, we implemented SVM as the classifier to distinguish different cognitive load levels.

Table 3 lists the classification performance, including accuracy, precision and recall rate resulted from the K-Folds method (k=10) cross-validation for each participant. The average classification performance for all participants by the SVM classifier was 74.06±5.01 percent in accuracy, 73.41±5.49 percent in recall rate and 80.93±2.67 percent in precision. It is worth noting that the classification accuracy varied among individual participants. We obtained the best accuracy (85.31%) for participant 10 and the worst one (65.00%) from for 13. It indicates that the fNIRS shows a significant individual difference and the model should be personalized. **Table 4** lists the classification performance at the group level (i.e., aggregating all datasets). Compared to individual level models the aggregated model at the group level shows a relatively low performance. The accuracy, recall rate and precision of group level classification were 62.10%, 60.98%, 62.87% relatively. It further shows that individual difference in fNIRS patterns make it difficult to find a generalizable model for cognitive load prediction. In addition, we further found that the performance of fNIRS-based cognitive load prediction varies at different cognitive load levels. **Fig.8.** shows examples of the individual (Participant 4) and group level classification in the form of confusion matrices. We found that both individual and group level classification showed a higher true positive rate in Level 1 and Level 3 than Level 2. The worse classification performance in level 2 probably resulted from the transition period from level 1 to level 2 and level 2 to level 3. The similarity of the HbO signals between levels made the classifiers easy to false predict the other levels into level 2.

Feature Importance Evaluation

In addition, we further used the wrapper method with a sequential forward feature selection (SFS) method to find the optimal feature subset to achieve better classification performance for both the left hemisphere and the right hemisphere as illustrated in **Fig. 9**. The wrapper method is defined as the feature selection method that employs a search strategy to look through the space of possible feature subsets regarding the quality of the performance of a given algorithm. Comparing to the other feature selection methods including filter methods and embedded methods, the wrapper method can detect the interaction between different variables and find the optimal feature subset for the certain machine learning algorithm. According to our previous study (Shi et al. 2020), the neuro connectivity and interactions can play an important role in reflecting the neuro activities based on the fNIRS data. Thus, we selected the wrapper method in this analysis. At the same time, the wrapper methods usually achieve better predictive accuracy than filter methods. There are four steps of the wrapper method: (1) search for a subset; (2) build a machine learning model; (3) evaluate the prediction performance; (4) repeat the process. In this analysis, we selected step forward feature selection also called sequential forward feature selection (SFS). **Fig. 9** shows one interaction of feature selection according to certain machining learning method. We set the accuracy to evaluate the prediction performance and the K-Folds method (k=10) was also used to cross-validate the feature selection results. **Table 5** lists the ranking orders of these features selected by the wrapper method with sequential forward feature selection during 10-fold. Each participant had different top 5 key features according to their individual prediction model. The most significant top 5 features including (1) Peak_Left Hbo, (2) Peak_Right Hbo, (3) AUC_Left Hbo, (4) Std_Left Hbo, and (5) AUC_Right features. These results further suggested that using a subset of fNIRS features in the review session can be used to predicting the cognitive load with satisfactory accuracy.

DISCUSSION AND CONCLUSIONS

Amid the increasing complexity of modern buildings, wayfinding becomes nontrivial for building users and professionals such as first responders. Although the rapid advancements of building informatics and visualization methods have made wayfinding information more accessible, the processing of excessive spatial information may also lead to potential risks of cognitive overload, leading to erroneous or biased decision making. Adaptive wayfinding information systems that closely monitor a person's information-induced cognitive load, and adapt the information formats and amounts correspondingly, are expected to be a promising solution for certain high-risk professionals who are challenged by wayfinding tasks, such as first responders. A core component of such a system is a real-time cognitive load monitoring and prediction method. This paper examines an explicit cognitive load measure and prediction method based on fNIRS data, i.e., the hemodynamic responses in different brain areas for evaluating the neural functional activation related to information processing. Compared to the existing neurophysiological measurement methods, such as EEG and pupillary dilation, fNIRS data is less affected by motor artefacts and thus is more suited for the real-time cognitive status monitoring for wayfinding tasks. The challenge is that there is still no widely accepted analytical pipeline for fNIRS-based cognitive load measurement. Basic questions, such as what features of fNIRS data can be used to cognitive load, remain unanswered. In addition, many fNIRS studies have focused on revealing the generic patterns hidden in fNIRS data in regard to predicting cognitive load based on aggregate data of multiple human subjects(Erdoĝan et al. 2019; Hong et al. 2015; Verdière et al. 2018). There is an urgent need to test and confirm whether there are inherent differences among individuals.

To narrow the gap, we performed a controlled Sternberg experiment (N=15) to model the relationship between fNIRS data patterns and cognitive load. In the experiment, emergency wayfinding information, including symbolic information, orientations, words and mixed information, were presented to the participants. Participants were required to memorize the presented information and answer questions based on the working memory. Their performance and fNIRS data were recorded for modeling purpose. The information was grouped into three levels of difficulty, based on the information amount and the complexity of the information format. Then fNIRS data was used to classify three levels of cognitive load according to the corresponding difficulty of information memorization tasks.

We applied a feature extraction method and obtained 6 features from the raw fNIRS data (every 40s interval), including *mean oxygen consumption* (mean), *standard deviation of oxygen consumption* (std dev), *peak oxygen consumption* (peak), *area under curve* (AUC), *kurtosis of oxygen consumption* *distribution* (Kurt) and *Skewness of oxygen consumption* (Skew). Then we trained SVM machine learning models to predict cognitive load levels (low, medium, and high) measured by the memorization task difficulty. The results show an average of 74.06±5.01 percent accuracy in predicting cognitive load using fNIRS data only in the individual level. In addition, we found that *peak oxygen consumption* (peak), *area under curve* (AUC) and *standard deviation* (std dev). These features all reflect the volatility of brain oxygen consumption instead of *mean oxygen consumption* (mean). It further justifies the robustness of using fNIRS in tracking cognitive load, as the volatility of hemodynamic response is less affected by body motions compared to the absolute oxygen consumption. In addition, we found significant individual differences in fNIRS-based cognitive load models. The significant features for predicting cognitive load, also varied across the participants. It indicates that each person is unique regarding the patterns of cognitive load due to the processing of wayfinding information, suggesting that personalized models instead of universal models are needed for predicting cognitive load based on neuroimaging data. Based on the findings, we have presented an analytical pipeline for filtering, cleaning and modeling fNIRS data for cognitive load prediction. The findings are expected to add evidence to the fNIRS-based cognitive load prediction literature (Du et al. 2020; Shi et al. 2020), inspire new findings with fNIRS tools, and facilitate the development of future cognition-driven information systems.

One of the limitations to be addressed in the future is to test the cognitive load models in a more realistic wayfinding experiment setting. In this study, we controlled the experiment environments where test subjects were required to sit quietly and to memorize relevant elements of wayfinding information. The findings are more related to the cognitive load for memory development and information encoding. Yet in our previous studies (Shi et al. 2020), we found that cognitive status showed different patterns in information memorization phase (information encoding) and task performance phase (memory retrieval). It suggests the possibilities that the developed cognitive load prediction models may not be applicable for all wayfinding task phases (i.e., reviewing information versus using the working memory in the field). In addition, we will also need to test how extemporaneous wayfinding information, i.e., onsite wayfinding information that is not previously reviewed (e.g., interactive road guidance signage), affects the cognitive load in a simultaneous information encoding and use process.

ACKNOWLEDGMENTS

This material is supported by the National Science Foundation (NSF) under Grants 1937878, as well as the National Institute of Standards and Technology (NIST) under Grant 60NANB18D152. Any opinions, findings, conclusions, or recommendations expressed in this article are those of the authors and do not reflect the views of the NSF and NIST.

DATA AVAILABILITY STATEMENT

All data, models, or code generated or used during the study are available in a repository or online in accordance with funder data retention policies:

https://www.dropbox.com/sh/cnnmkfns4e3rx2l/AABuGrC-BBqiipN5lnHk3UeLa?dl=0

REFERENCES

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

Andreasen, N. C., O'Leary, D. S., Cizadlo, T., Arndt, S., Rezai, K., Ponto, L., Watkins, G. L., and Hichwa, R. D. (1996). "Schizophrenia and cognitive dysmetria: a positron-emission tomography study of dysfunctional prefrontal-thalamic-cerebellar circuitry." *Proceedings of the National Academy of Sciences*, 93(18), 9985-9990.

Antonenko, P., Paas, F., Grabner, R., and Van Gog, T. (2010). "Using electroencephalography to measure cognitive load." *Educational Psychology Review*, 22(4), 425-438.

Aporta, C., Higgs, E., Hakken, D., Palmer, L., Palmer, M., Rundstrom, R., Pfaffenberger, B., Wenzel, G., Widlok, T., and Aporta, C. (2005). "Satellite culture: global positioning systems, Inuit wayfinding, and the need for a new account of technology." *Current anthropology*, 46(5), 729-753.

Ayaz, H., Onaral, B., Izzetoglu, K., Shewokis, P. A., McKendrick, R., and Parasuraman, R. (2013). "Continuous monitoring of brain dynamics with functional near infrared spectroscopy as a tool for neuroergonomic research: empirical examples and a technological development." *Frontiers in human neuroscience*, 7, 871.

Baddeley, A. (1992). "Working memory and conscious awareness." *Theories of memory*, Lawrence Erlbaum Associates, 11-20.

Baddeley, A. (2000). "The episodic buffer: a new component of working memory?" *Trends in cognitive sciences*, 4(11), 417-423.

Baddeley, A. (2003). "Working memory: looking back and looking forward." *Nature Reviews. Neuroscience*, 4(10), 829.

Baddeley, A. (2012). "Working memory: theories, models, and controversies." *Annual review of psychology*, 63, 1-29.

Baddeley, A. D., and Hitch, G. (1974). "Working memory." *Psychology of learning and motivation*, 8, 47-89.

Banaji, M. R., and Greenwald, A. G. "Implicit stereotyping and prejudice." *Proc., The psychology of prejudice: The Ontario symposium*, 55-76.

Bauernfeind, G., Steyrl, D., Brunner, C., and Müller-Putz, G. R. "Single trial classification of fnirs-based brain-computer interface mental arithmetic data: a comparison between different classifiers." *Proc., 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, IEEE, 2004-2007.

Bergstra, J., and Bengio, Y. (2012). "Random search for hyper-parameter optimization." *The Journal of Machine Learning Research*, 13(1), 281-305.

Brigadoi, S., Ceccherini, L., Cutini, S., Scarpa, F., Scatturin, P., Selb, J., Gagnon, L., Boas, D. A., and Cooper, R. J. (2014). "Motion artifacts in functional near-infrared spectroscopy: a comparison of motion correction techniques applied to real cognitive data." *Neuroimage*, 85, 181-191.

Bunce, S. C., Izzetoglu, M., Izzetoglu, K., Onaral, B., and Pourrezaei, K. (2006). "Functional near-infrared spectroscopy." *IEEE engineering in medicine and biology magazine*, 25(4), 54-62.

Bunch, R. L., and Lloyd, R. E. (2006). "The cognitive load of geographic information." *The Professional Geographer*, 58(2), 209-220.

Cabeza, R., and Nyberg, L. (2000). "Imaging cognition II: An empirical review of 275 PET and fMRI studies." *Journal of cognitive neuroscience*, 12(1), 1-47.

Calori, C., and Vanden-Eynden, D. (2015). *Signage and wayfinding design: a complete guide to creating environmental graphic design systems*, John Wiley & Sons.

Canning, C., and Scheutz, M. (2013). "Functional near-infrared spectroscopy in human-robot interaction." *Journal of Human-Robot Interaction*, 2(3), 62-84.

Chandler, P., and Sweller, J. (1991). "Cognitive load theory and the format of instruction." *Cognition and instruction*, 8(4), 293-332.

Chang, Y.-J., and Wang, T.-Y. (2010). "Comparing picture and video prompting in autonomous indoor wayfinding for individuals with cognitive impairments." *Personal and Ubiquitous Computing*, 14(8), 737-747.

Cooper, R., Selb, J., Gagnon, L., Phillip, D., Schytz, H. W., Iversen, H. K., Ashina, M., and Boas, D. A. (2012). "A systematic comparison of motion artifact correction techniques for functional near-infrared spectroscopy." *Frontiers in neuroscience*, 6, 147.

Cui, X., Bray, S., Bryant, D. M., Glover, G. H., and Reiss, A. L. (2011). "A quantitative comparison of NIRS and fMRI across multiple cognitive tasks." *Neuroimage*, 54(4), 2808-2821.

Cui, X., Bray, S., and Reiss, A. L. (2010). "Speeded near infrared spectroscopy (NIRS) response detection." *PLoS one*, 5(11), e15474.

de Fockert, J. W., Rees, G., Frith, C. D., and Lavie, N. (2001). "The role of working memory in visual selective attention." *Science*, 291(5509), 1803-1806.

Delpy, D. T., Cope, M., van der Zee, P., Arridge, S., Wray, S., and Wyatt, J. (1988). "Estimation of optical pathlength through tissue from direct time of flight measurement." *Physics in Medicine & Biology*, 33(12), 1433.

Du, J., Zhu, Q., Shi, Y., Wang, Q., Lin, Y., and Zhao, D. (2020). "Cognition digital twins for personalized information systems of smart cities: Proof of concept." *Journal of Management in Engineering*, 36(2), 04019052.

Elvins, T. T. (1997). "VisFiles: virtually lost in virtual worlds—wayfinding without a cognitive map." *ACM SIGGRAPH computer graphics*, 31(3), 15-17.

Eppler, M. J., and Mengis, J. (2004). "The concept of information overload: A review of literature from organization science, accounting, marketing, MIS, and related disciplines." *The information society*, 20(5), 325-344.

Erdoĝan, S. B., Özsarfati, E., Dilek, B., Kadak, K. S., Hanoĝlu, L., and Akın, A. (2019). "Classification of motor imagery and execution signals with population-level feature sets: implications for probe design in fNIRS based BCI." *Journal of neural engineering*, 16(2), 026029.

Faress, A., and Chau, T. (2013). "Towards a multimodal brain–computer interface: combining fNIRS and fTCD measurements to enable higher classification accuracy." *Neuroimage*, 77, 186-194.

Fischer, G. "Context-aware systems: the'right'information, at the'right'time, in the'right'place, in the'right'way, to the'right'person." *Proc., Proceedings of the international working conference on advanced visual interfaces*, ACM, 287-294.

Galar, M., Fernández, A., Barrenechea, E., Bustince, H., and Herrera, F. (2011). "An overview of ensemble methods for binary classifiers in multi-class problems: Experimental study on one-vs-one and one-vs-all schemes." *Pattern Recognition*, 44(8), 1761-1776.

Gateau, T., Durantin, G., Lancelot, F., Scannella, S., and Dehais, F. (2015). "Real-time state estimation in a flight simulator using fNIRS." *PloS one*, 10(3), e0121279.

Gevins, A., and Smith, M. E. (2003). "Neurophysiological measures of cognitive workload during human-computer interaction." *Theoretical Issues in Ergonomics Science*, 4(1-2), 113-131.

Goldinger, S. D., and Papesh, M. H. (2012). "Pupil dilation reflects the creation and retrieval of memories." *Current Directions in Psychological Science*, 21(2), 90-95.

Goucher-Lambert, K., Moss, J., and Cagan, J. (2017). "Inside the mind: using neuroimaging to understand moral product preference judgments involving sustainability." *Journal of Mechanical Design*, 139(4).

Haapalainen, E., Kim, S., Forlizzi, J. F., and Dey, A. K. "Psycho-physiological measures for assessing cognitive load." *Proc., Proceedings of the 12th ACM international conference on Ubiquitous computing*, 301-310.

Halgren, E., Raij, T., Marinkovic, K., Jousmäki, V., and Hari, R. (2000). "Cognitive response profile of the human fusiform face area as determined by MEG." *Cerebral cortex*, 10(1), 69-81.

Herff, C., Heger, D., Fortmann, O., Hennrich, J., Putze, F., and Schultz, T. (2014). "Mental workload during n-back task—quantified in the prefrontal cortex using fNIRS." *Frontiers in human neuroscience*, 7, 935.

Holper, L., and Wolf, M. (2011). "Single-trial classification of motor imagery differing in task complexity: a functional near-infrared spectroscopy study." *Journal of neuroengineering and rehabilitation*, 8(1), 34.

Hong, K.-S., Khan, M. J., and Hong, M. J. (2018). "Feature extraction and classification methods for hybrid fNIRS-EEG brain-computer interfaces." *Frontiers in human neuroscience*, 12, 246.

Hong, K.-S., Naseer, N., and Kim, Y.-H. (2015). "Classification of prefrontal and motor cortex signals for three-class fNIRS–BCI." *Neuroscience letters*, 587, 87-92.

Hoshi, Y., Kobayashi, N., and Tamura, M. (2001). "Interpretation of near-infrared spectroscopy signals: a study with a newly developed perfused rat brain model." *Journal of applied physiology*, 90(5), 1657-1662.

Hosseini, R., Walsh, B., Tian, F., and Wang, S. (2018). "An fNIRS-based feature learning and classification framework to distinguish hemodynamic patterns in children who stutter." *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 26(6), 1254-1263.

Hu, M., and Shealy, T. (2019). "Application of functional near-infrared spectroscopy to measure engineering decision-making and design cognition: Literature review and synthesis of methods." *Journal of Computing in Civil Engineering*, 33(6), 04019034.

Huppert, T. J., Diamond, S. G., Franceschini, M. A., and Boas, D. A. (2009). "HomER: a review of time-series analysis methods for near-infrared spectroscopy of the brain." *Applied optics*, 48(10), D280-D298.

Huppert, T. J., Hoge, R. D., Diamond, S. G., Franceschini, M. A., and Boas, D. A. (2006). "A temporal comparison of BOLD, ASL, and NIRS hemodynamic responses to motor stimuli in adult humans." *Neuroimage*, 29(2), 368-382.

Jain, A., Nandakumar, K., and Ross, A. (2005). "Score normalization in multimodal biometric systems." *Pattern recognition*, 38(12), 2270-2285.

Kahneman, D., and Beatty, J. (1966). "Pupil diameter and load on memory." *Science*, 154(3756), 1583-1585.

Kalyuga, S. (2009). "Cognitive load theory." *Managing Cognitive Load in Adaptive Multimedia Learning*, IGI Global, 34-57.

Kessler, J. B., and Meier, S. (2014). "Learning from (failed) replications: Cognitive load manipulations and charitable giving." *Journal of Economic Behavior & Organization*, 102, 10-13.

Khan, M. A., Zafar, A., and Hong, K.-S. "Detection and Classification of Three-Class Initial Dips Using Vector Phase Analysis with Dual Threshold Circles: An fNIRS Study." *Proc., 2019 19th International Conference on Control, Automation and Systems (ICCAS)*, IEEE, 786-791.

Klatzky, R. L., Marston, J. R., Giudice, N. A., Golledge, R. G., and Loomis, J. M. (2006). "Cognitive load of navigating without vision when guided by virtual sound versus spatial language." *Journal of Experimental Psychology: Applied*, 12(4), 223.

Klingner, J., Tversky, B., and Hanrahan, P. (2011). "Effects of visual and verbal presentation on cognitive load in vigilance, memory, and arithmetic tasks." *Psychophysiology*, 48(3), 323-332.

Kucewicz, M. T., Dolezal, J., Kremen, V., Berry, B. M., Miller, L. R., Magee, A. L., Fabian, V., and Worrell, G. A. (2018). "Pupil size reflects successful encoding and recall of memory in humans." *Scientific reports*, 8(1), 4949.

Lindberg, E., and Gärling, T. (1983). "Acquisition of different types of locational information in cognitive maps: Automatic or effortful processing?" *Psychological Research*, 45(1), 19-38.

Liu, A. L., Hile, H., Kautz, H., Borriello, G., Brown, P. A., Harniss, M., and Johnson, K. (2008). "Indoor wayfinding: Developing a functional interface for individuals with cognitive impairments." *Disability and Rehabilitation: Assistive Technology*, 3(1-2), 69-81.

Lloyd-Fox, S., Blasi, A., and Elwell, C. (2010). "Illuminating the developing brain: the past, present and future of functional near infrared spectroscopy." *Neuroscience & Biobehavioral Reviews*, 34(3), 269-284.

Logie, R. H. (2014). *Visuo-spatial working memory*, Psychology Press.

Lohani, M., Payne, B. R., and Strayer, D. L. (2019). "A review of psychophysiological measures to assess cognitive states in real-world driving." *Frontiers in human neuroscience*, 13, 57.

Matthews, G., Middleton, W., Gilmartin, B., and Bullimore, M. (1991). "Pupillary diameter and cognitive load." *Journal of Psychophysiology*.

Meilinger, T., Knauff, M., and Bülthoff, H. H. (2008). "Working memory in wayfinding—A dual task experiment in a virtual city." *Cognitive Science*, 32(4), 755-770.

Memoori (2015). "Big Data for Smart Buildings 2015 to 2020."

Miller, G. A. (1956). "The magical number seven, plus or minus two: Some limits on our capacity for processing information." *Psychological review*, 63(2), 81.

Miyake, A., and Shah, P. (1999). *Models of working memory: Mechanisms of active maintenance and executive control*, Cambridge University Press.

Moreno, R., and Mayer, R. (2007). "Interactive multimodal learning environments." *Educational psychology review*, 19(3), 309-326.

Mulloni, A., Seichter, H., and Schmalstieg, D. "Handheld augmented reality indoor navigation with activity-based instructions." *Proc., Proceedings of the 13th international conference on human computer interaction with mobile devices and services*, 211-220.

Naseer, N., and Hong, K.-S. (2013). "Classification of functional near-infrared spectroscopy signals corresponding to the right-and left-wrist motor imagery for development of a brain–computer interface." *Neuroscience letters*, 553, 84-89.

Naseer, N., and Hong, K.-S. (2015). "fNIRS-based brain-computer interfaces: a review." *Frontiers in human neuroscience*, 9, 3.

Noori, F. M., Naseer, N., Qureshi, N. K., Nazeer, H., and Khan, R. A. (2017). "Optimal feature selection from fNIRS signals using genetic algorithms for BCI." *Neuroscience letters*, 647, 61-66.

Ochsner, K. N., Bunge, S. A., Gross, J. J., and Gabrieli, J. D. (2002). "Rethinking feelings: an FMRI study of the cognitive regulation of emotion." *Journal of cognitive neuroscience*, 14(8), 1215-1229.

Olsson, T., Lagerstam, E., Kärkkäinen, T., and Väänänen-Vainio-Mattila, K. (2013). "Expected user experience of mobile augmented reality services: a user study in the context of shopping centres." *Personal and ubiquitous computing*, 17(2), 287-304.

Paas, F., Renkl, A., and Sweller, J. (2003). "Cognitive load theory and instructional design: Recent developments." *Educational psychologist*, 38(1), 1-4.

Paas, F., Tuovinen, J. E., Tabbers, H., and Van Gerven, P. W. (2003). "Cognitive load measurement as a means to advance cognitive load theory." *Educational psychologist*, 38(1), 63-71.

Papesh, M. H., Goldinger, S. D., and Hout, M. C. (2012). "Memory strength and specificity revealed by pupillometry." *International Journal of Psychophysiology*, 83(1), 56-64.

Pattyn, N., Neyt, X., Henderickx, D., and Soetens, E. (2008). "Psychophysiological investigation of vigilance decrement: boredom or cognitive fatigue?" *Physiology & behavior*, 93(1-2), 369-378.

Pinti, P., Scholkmann, F., Hamilton, A., Burgess, P., and Tachtsidis, I. (2018). "Current status and issues regarding pre-processing of fNIRS neuroimaging data: An investigation of diverse signal filtering methods within a General Linear Model framework." *Frontiers in human neuroscience*, 12, 505.

Pinti, P., Tachtsidis, I., Hamilton, A., Hirsch, J., Aichelburg, C., Gilbert, S., and Burgess, P. W. (2018). "The present and future use of functional near- infrared spectroscopy (fNIRS) for cognitive neuroscience." *Annals of the New York Academy of Sciences*.

Power, S. D., and Chau, T. (2013). "Automatic single-trial classification of prefrontal hemodynamic activity in an individual with Duchenne muscular dystrophy." *Developmental neurorehabilitation*, 16(1), 67-72.

Ramnani, N., and Owen, A. M. (2004). "Anterior prefrontal cortex: insights into function from anatomy and neuroimaging." *Nature reviews neuroscience*, 5(3), 184-194.

Randell, C., and Muller, H. "Low cost indoor positioning system." *Proc., Ubicomp 2001: Ubiquitous Computing*, Springer, 42-48.

Ray, W. J., and Cole, H. W. (1985). "EEG alpha activity reflects attentional demands, and beta activity reflects emotional and cognitive processes." *Science*, 228(4700), 750-752.

Rossano, M. J., and Moak, J. (1998). "Spatial representations acquired from computer models: Cognitive load, orientation specificity and the acquisition of survey knowledge." *British Journal of Psychology*, 89(3), 481-497.
Roux, F., and Uhlhaas, P. J. (2014). "Working memory and neural oscillations: alpha–gamma versus theta–gamma codes for distinct WM information?" *Trends in cognitive sciences*, 18(1), 16-25.
Roy, R. N., Bonnet, S., Charbonnier, S., and Campagne, A. "Mental fatigue and working memory load estimation: interaction and implications for EEG-based passive BCI." *Proc., 2013 35th annual international conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, IEEE, 6607-6610.
Sandry, J., Genova, H. M., Dobryakova, E., DeLuca, J., and Wylie, G. (2014). "Subjective cognitive fatigue in multiple sclerosis depends on task length." *Frontiers in neurology*, 5, 214.
Schroeter, M. L., Kupka, T., Mildner, T., Uludağ, K., and von Cramon, D. Y. (2006). "Investigating the post-stimulus undershoot of the BOLD signal—a simultaneous fMRI and fNIRS study." *Neuroimage*, 30(2), 349-358.
Sharafi, Z., Soh, Z., and Guéhéneuc, Y.-G. (2015). "A systematic literature review on the usage of eye-tracking in software engineering." *Information and Software Technology*, 67, 79-107.
Shi, Y., Du, J., and Ragan, E. (2020). "Review visual attention and spatial memory in building inspection: Toward a cognition-driven information system." *Advanced Engineering Informatics*, 44, 101061.
Shi, Y., Du, J., and Worthy, D. A. (2020). "The impact of engineering information formats on learning and execution of construction operations: A virtual reality pipe maintenance experiment." *Automation in Construction*, 119, 103367.
Shi, Y., Zhu, Y., Mehta, R. K., and Du, J. (2020). "A neurophysiological approach to assess training outcome under stress: A virtual reality experiment of industrial shutdown maintenance using Functional Near-Infrared Spectroscopy (fNIRS)." *Advanced Engineering Informatics*, 46, 101153.
Siegel, A. W., and White, S. H. (1975). "The development of spatial representations of large-scale environments." *Advances in child development and behavior*, Elsevier, 9-55.
Smith, S. P., and Hart, J. "Evaluating distributed cognitive resources for wayfinding in a desktop virtual environment." *Proc., 3D User Interfaces, 2006. 3DUI 2006. IEEE Symposium on*, IEEE, 3-10.
Sridharan, D., Levitin, D. J., and Menon, V. (2008). "A critical role for the right fronto-insular cortex in switching between central-executive and default-mode networks." *Proceedings of the National Academy of Sciences*, 105(34), 12569-12574.
Sterman, M. B., Mann, C. A., Kaiser, D. A., and Suyenobu, B. Y. (1994). "Multiband topographic EEG analysis of a simulated visuomotor aviation task." *International journal of psychophysiology*, 16(1), 49-56.
Sternberg, S. (1969). "Memory-scanning: Mental processes revealed by reaction-time experiments." *American scientist*, 57(4), 421-457.
Sweller, J. (1988). "Cognitive load during problem solving: Effects on learning." *Cognitive science*, 12(2), 257-285.
Sweller, J. (1994). "Cognitive load theory, learning difficulty, and instructional design." *Learning and instruction*, 4(4), 295-312.
Sweller, J. (2010). "Cognitive load theory: Recent theoretical advances." *Cognitive load theory*, 1, 29-30.
Sweller, J., Ayres, P., and Kalyuga, S. (2011). *Cognitive load theory*, Springer.
Tanaka, M., Ishii, A., and Watanabe, Y. (2015). "Effects of mental fatigue on brain activity and cognitive performance: a magnetoencephalography study." *Anat Physiol*, 4, 1-5.

Taylor, H. A., Naylor, S. J., and Chechile, N. A. (1999). "Goal-specific influences on the representation of spatial perspective." *Memory & cognition*, 27(2), 309-319.

Thirunavukkarasu, G. S., Abdi, H., and Mohajer, N. "A smart HMI for driving safety using emotion prediction of EEG signals." *Proc., 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, IEEE, 004148-004153.

Toker, D., Conati, C., Steichen, B., and Carenini, G. "Individual user characteristics and information visualization: connecting the dots through eye tracking." *Proc., proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 295-304.

Tsunashima, H., and Yanagisawa, K. (2009). "Measurement of brain function of car driver using functional near-infrared spectroscopy (fNIRS)." *Computational intelligence and neuroscience*, 2009.

van Asselen, M., Fritschy, E., and Postma, A. (2006). "The influence of intentional and incidental learning on acquiring spatial knowledge during navigation." *Psychological Research*, 70(2), 151-156.

van der Wel, P., and van Steenbergen, H. (2018). "Pupil dilation as an index of effort in cognitive control tasks: A review." *Psychonomic bulletin & review*, 25(6), 2005-2015.

Verdière, K. J., Roy, R. N., and Dehais, F. (2018). "Detecting pilot's engagement using fNIRS connectivity features in an automated vs. manual landing scenario." *Frontiers in human neuroscience*, 12, 6.

Verney, S. P., Granholm, E., and Dionisio, D. P. (2001). "Pupillary responses and processing resources on the visual backward masking task." *Psychophysiology*, 38(1), 76-83.

von Lühmann, A., Li, X., Müller, K.-R., Boas, D. A., and Yücel, M. A. (2020). "Improved physiological noise regression in fNIRS: A multimodal extension of the General Linear Model using temporally embedded Canonical Correlation Analysis." *NeuroImage*, 208, 116472.

Wang, Q., Yang, S., Liu, M., Cao, Z., and Ma, Q. (2014). "An eye-tracking study of website complexity from cognitive load perspective." *Decision support systems*, 62, 1-10.

Werner, S., Krieg-Brückner, B., Mallot, H. A., Schweizer, K., and Freksa, C. (1997). "Spatial cognition: The role of landmark, route, and survey knowledge in human and robot navigation." *Informatik'97 Informatik als Innovationsmotor*, Springer, 41-50.

Willis, S., and Helal, S. "RFID information grid for blind navigation and wayfinding." *Proc., Wearable Computers, 2005. Proceedings. Ninth IEEE International Symposium on*, IEEE, 34-37.

Wu, C., Cha, J., Sulek, J., Zhou, T., Sundaram, C. P., Wachs, J., and Yu, D. (2019). "Eye-Tracking Metrics Predict Perceived Workload in Robotic Surgical Skills Training." *Human Factors*, 0018720819874544.

Xie, B., and Salvendy, G. (2000). "Prediction of mental workload in single and multiple tasks environments." *International journal of cognitive ergonomics*, 4(3), 213-242.

Zhang, Y., and Zhu, C. (2019). "Assessing Brain Networks by Resting-State Dynamic Functional Connectivity: An fNIRS-EEG Study." *Frontiers in Neuroscience*, 13.

Table 1. fNIRS signal features and related studies

Feature	**Formula**	**References**
Mean (Average)	$Mean(X) = E(X)$…(2)	(Faress and Chau 2013; Hong et al. 2015; Naseer and Hong 2013; Noori et al. 2017; Power and Chau 2013; Verdière et al. 2018)
Peak (Maximum)	$Peak(X) = Max(X)$ …(3)	(Bauernfeind et al. 2014; Cui et al. 2010; Noori et al. 2017; Verdière et al. 2018)
Stand Deviation (Variance)	$Var(X) = E[(X-\mu)^2]$ …(4)	(Holper and Wolf 2011; Noori et al. 2017; Verdière et al. 2018)
Area Under the Curve (AUC)	$AUC(X) = \sum\lvert X\rvert$…(5)	(Schroeter et al. 2006; Verdière et al. 2018)
Skewness	$Skewness(X) = E\left[(\frac{X-\mu}{\sigma})^3\right]$…(6)	(Holper and Wolf 2011; Noori et al. 2017; Verdière et al. 2018)
Kurtosis	$Kurtosis(X) = E\left[(\frac{X-\mu}{\sigma})^4\right]$…(7)	(Holper and Wolf 2011; Noori et al. 2017; Verdière et al. 2018)

Table 2. Three levels of task difficulty corresponding to cognitive load levels

Cognitive Load Levels	Encoding Pattern	Sternberg Trails
1- Light Load	Phonological information for landmarks	Trails 1- 40
2- Medial Load	Visuospatial information for routes and survey	Trails 41 - 80
3- Heavy Load	Combinations	Trails 81 – 120

Table 3. HbO-based classification accuracies of SVM in percentages [%] of individual level feature sets

Subjects	SVM		
	Accuracy	Recall	Precisions
1	73.46	72.94	79.25
2	70.79	71.11	82.79
3	78.16	77.65	78.32
4	76.77	77.05	81.17
5	70.08	71.45	79.33
6	78.63	80.05	83.88
7	72.60	70.10	81.95
8	77 .27	76.29	83.20
9	70.21	68.22	77.14
10	85.31	85.01	86.12
11	75.42	74.61	82.34
12	73.18	71.63	80.71
13	65.00	63.77	76.62
14	70.02	67.93	80.30
Mean	74.06±5.01	73.41±5.49	80.93±2.67

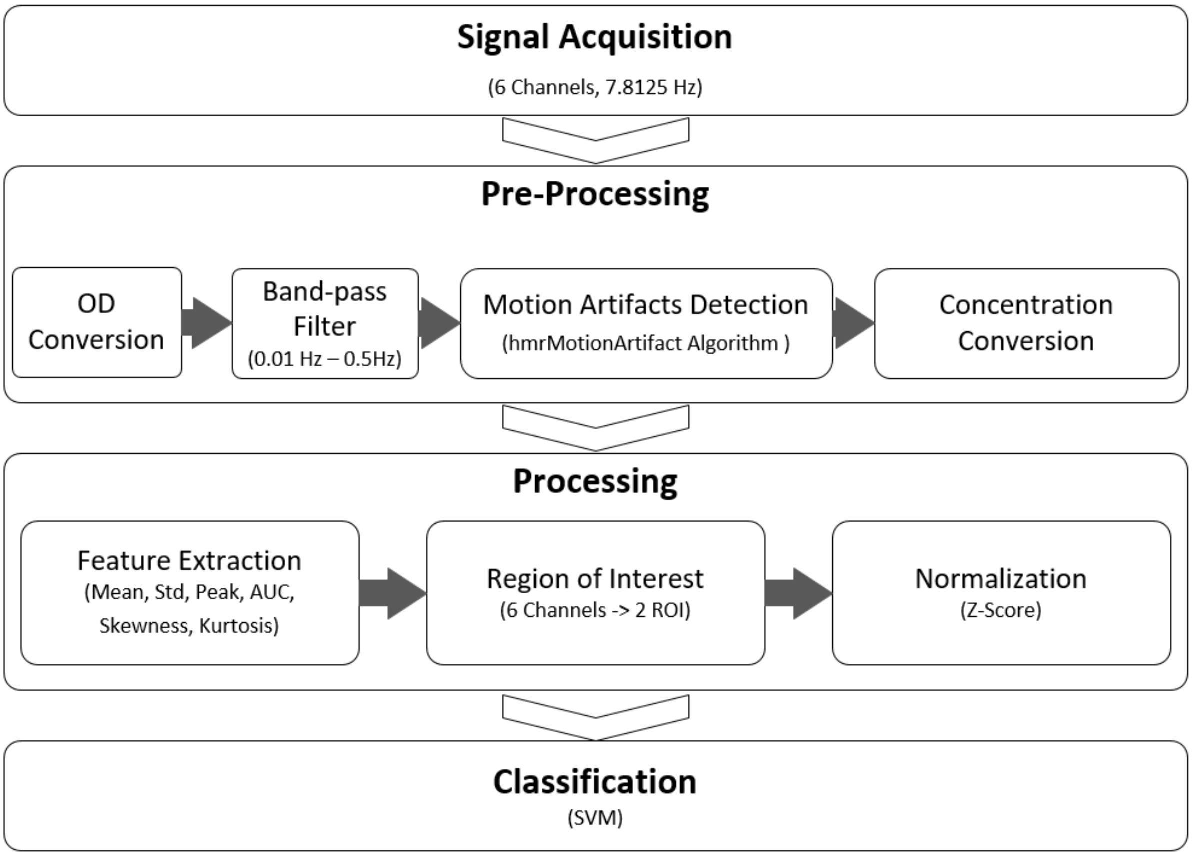
Table 4. HbO-based classification accuracies of SVM in percentages [%]

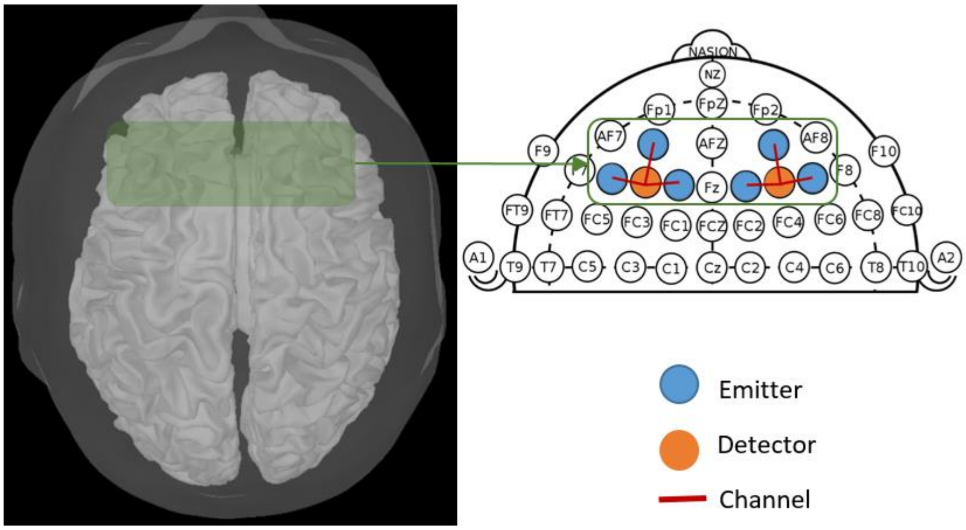
Categories	SVM		
	Accuracy	Recall	Precisions
Individual Level (mean)	74.06	73.41	80.93
Group Level	62.10	60.98	62.87

Table 5. The Top five features of each subject and group level data sets selected by the wrapper method

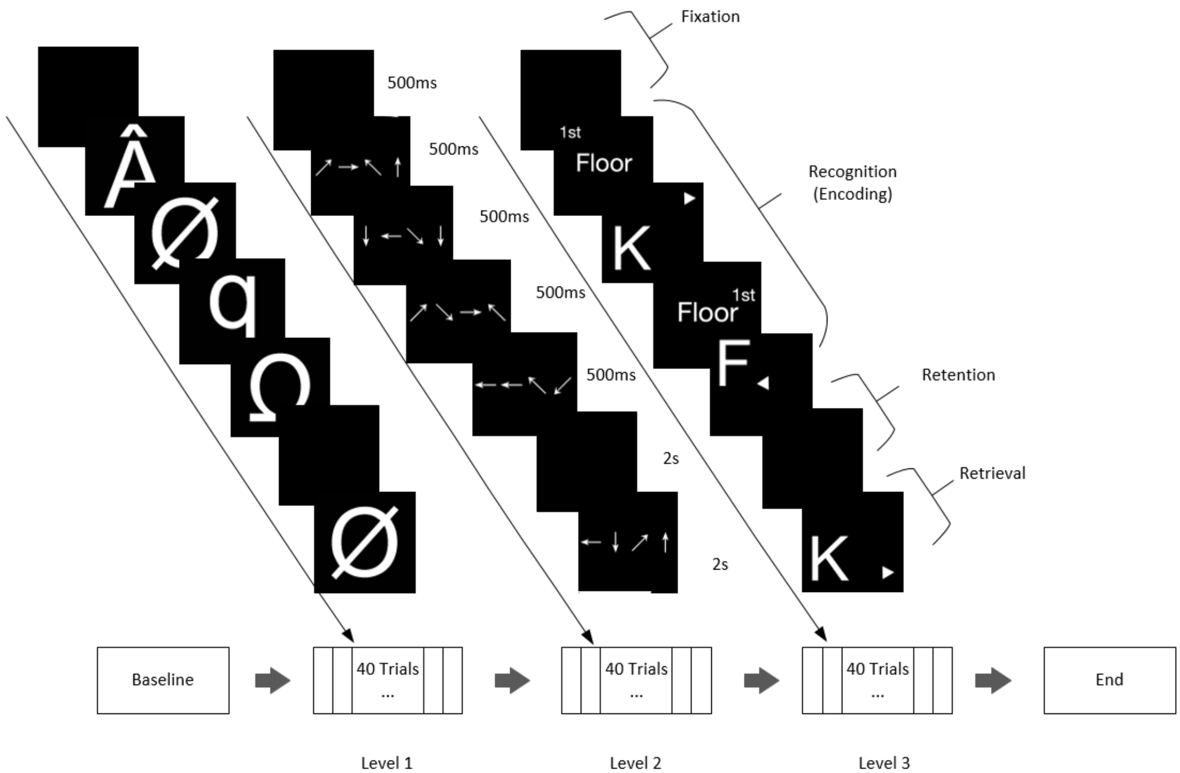
Subjects	Top features	Subjects	Top features	Subjects	Top features
1	1. AUC_Right 2. Peak_Right 3. Std_Right 4. Kurt_Left 5. AUC_Left	6	1. Peak_Righ 2. Std_Right 3. Mean_Right 4. Std_Left 5. Mean_Left	11	1. Peak_Right 2. Std_Right 3. Skew_Left 4. AUC_Left 5. Peak_Left
2	1. Skew_Right 2. Skew_Left 3. Kurt_Left 4. AUC_Left 5. Std_Left	7	1. AUC_Right 2. Peak_Right 3. Std_Right 4. AUC_Left 5. Peak_Left	12	1. Std_Right 2. Kurt_Left 3. AUC_Left 4. Peak_Left 5. Std_Left
3	1. Skew_Right 2. Peak_Right 3. Std_Right 4. Peak_Left 5. Mean_Left	8	1. Kurt_Right 2. AUC_Right 3. Std_Right 4. Kurt_Left 5. Std_Left	13	1. Std_Right 2. Mean_Right 3. Skew_Left 4. Std_Left 5. Mean_Left
4	1. AUC_Right 2. Peak_Right 3. AUC_Left 4. Peak_Left 5. Mean_Left	9	1. Skew_Right 2. AUC_Right 3. Peak_Right 4. Std_Right 5. AUC_Left	14	1. AUC_Right 2. Peak_Right 3. Mean_Right 4. AUC_Left 5. Std_Left
5	1. Kurt_Right 2. Peak_Right 3. Skew_Left 4. Peak_Left 5. Std_Left	10	1. Skew_Right 2. Kurt_Right 3. AUC_Right 4. Std_Right 5. AUC_Left	Group Level	1. Kurt_Right 2. Skew_Left 3. AUC_Left 4. Std_Left 5. Mean_Left

List of Tables

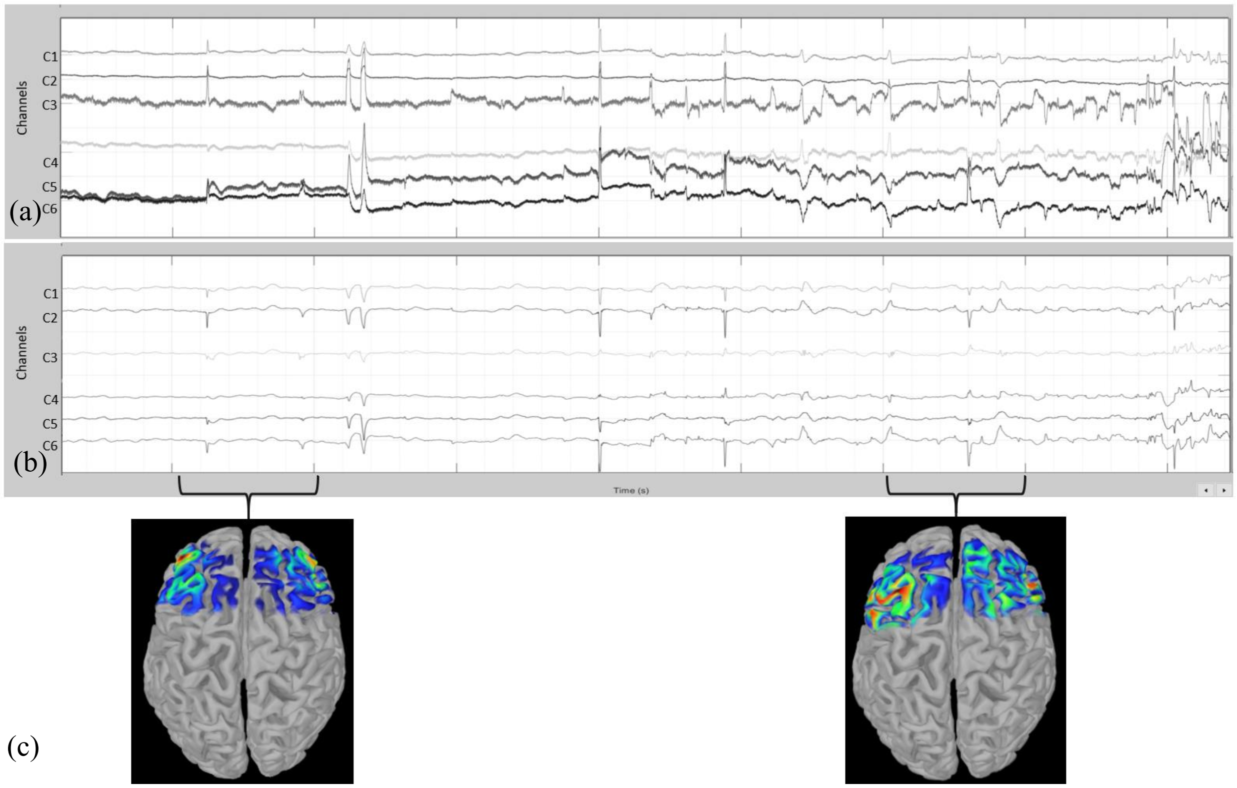

Signal Acquisition
(6 Channels, 7.8125 Hz)
Pre-Processing
OD Conversion
Band-pass Filter
(0.01 Hz – 0.5Hz)
Motion Artifacts Detection
(hmrMotionArtifact Algorithm)
Concentration Conversion
Processing
Feature Extraction
(Mean, Std, Peak, AUC, Skewness, Kurtosis)
Region of Interest
(6 Channels -> 2 ROI)
Normalization
(Z-Score)
Classification
(SVM)

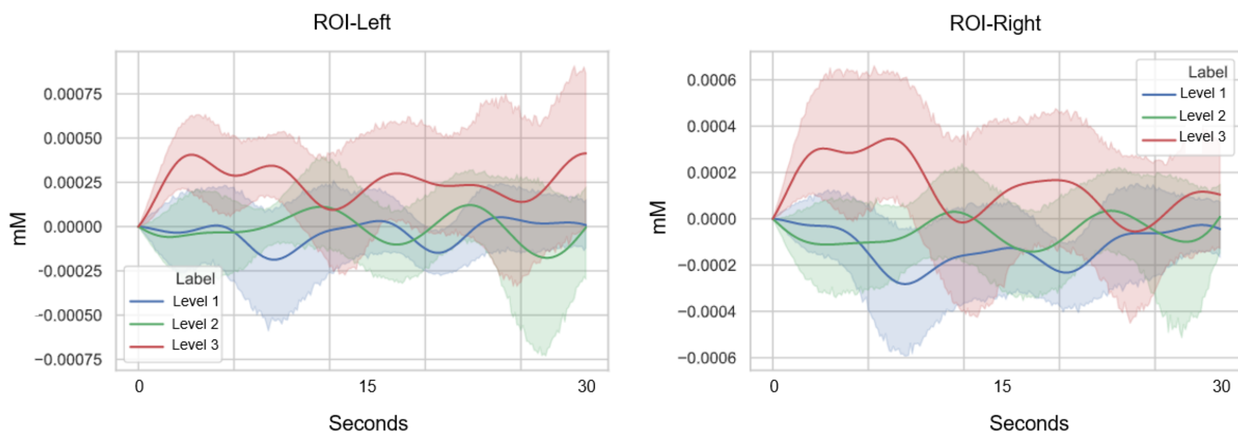

NASION
NZ
FpZ
Fp1
Fp2
AF7
AFZ
AF8
F9
F7
Fz
F8
F10
FT9
FT7
FC5
FC3
FC1
FCZ
FC2
FC4
FC6
FC8
FC10
A1
T9
T7
C5
C3
C1
Cz
C2
C4
C6
T8
T10
A2
Emitter
Detector
Channel

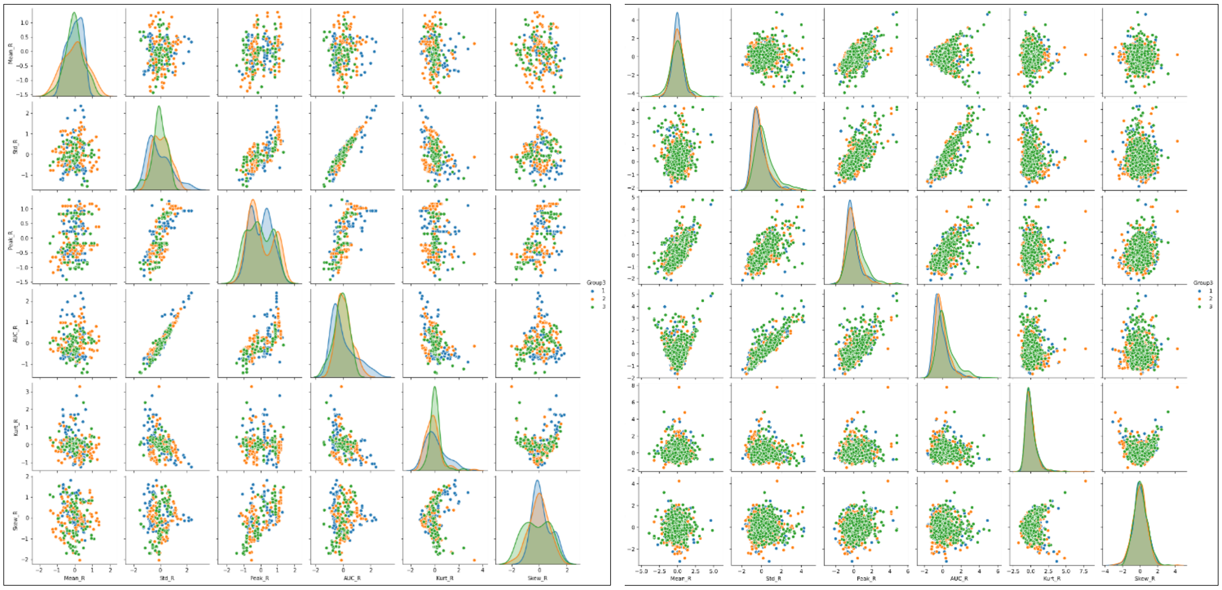
Figure


Fixation
500ms
500ms
500ms
500ms
500ms
2s
2s
1st
Floor
K
1st
Floor
F
K
Recognition
(Encoding)
Retention
Retrieval
Baseline
40 Trials
...
40 Trials
...
40 Trials
...
End
Level 1
Level 2
Level 3

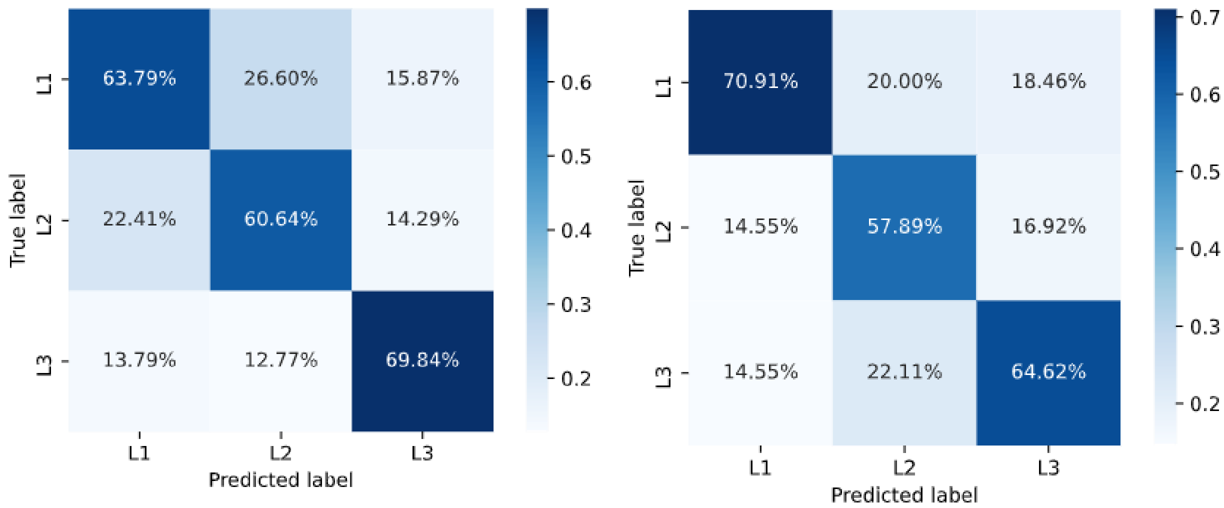
(a) (b)


(a)
Channels
C1
C2
C3
C4
C5
C6
(b)
Channels
C1
C2
C3
C4
C5
C6
Time (s)
(c)

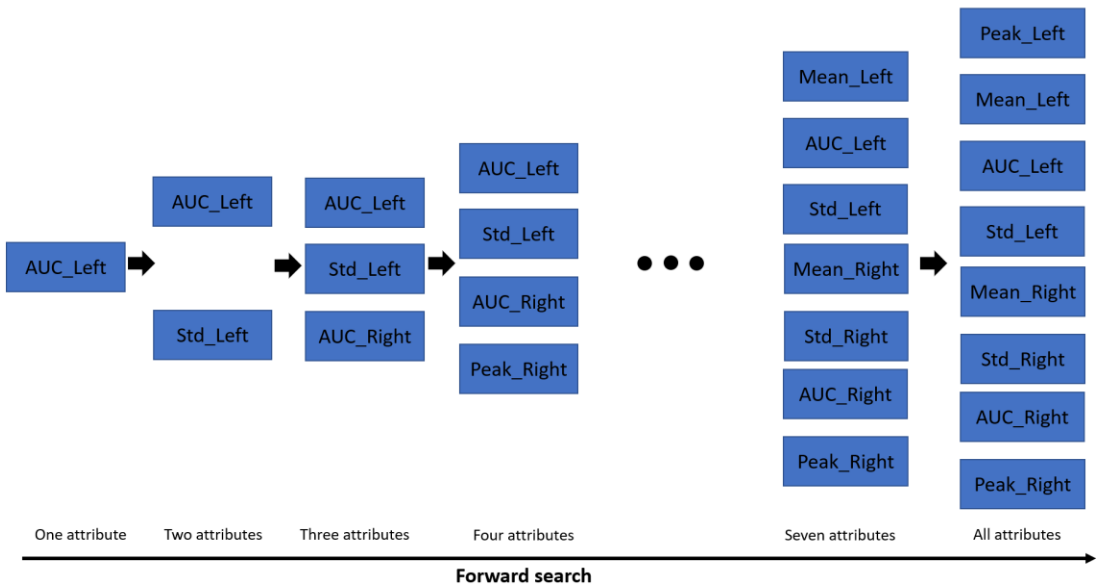

ROI-Left
ROI-Right
mM
Seconds
Label
Level 1
Level 2
Level 3
0.00075
0.00050
0.00025
0.00000
−0.00025
−0.00050
−0.00075
0.0006
0.0004
0.0002
0.0000
−0.0002
−0.0004
−0.0006
0
15
30



(a) (b)


63.79%
26.60%
15.87%
22.41%
60.64%
14.29%
13.79%
12.77%
69.84%
L1
L2
L3
True label
Predicted label
0.6
0.5
0.4
0.3
0.2
70.91%
20.00%
18.46%
14.55%
57.89%
16.92%
14.55%
22.11%
64.62%
L1
L2
L3
True label
Predicted label
0.7
0.6
0.5
0.4
0.3
0.2

(a) (b)


AUC_Left
AUC_Left
Std_Left
AUC_Left
Std_Left
AUC_Right
AUC_Left
Std_Left
AUC_Right
Peak_Right
Mean_Left
AUC_Left
Std_Left
Mean_Right
Std_Right
AUC_Right
Peak_Right
Peak_Left
Mean_Left
AUC_Left
Std_Left
Mean_Right
Std_Right
AUC_Right
Peak_Right
One attribute
Two attributes
Three attributes
Four attributes
Seven attributes
All attributes
Forward search

List of Figures

Figure 1. The analytical pipeline for developing the prediction model of cognitive load with fNIRS.

Figure 2. Placement of emitters and detectors for PFC cognitive load data collection.

Figure 3. Sternberg working memory test: The experiment included three levels of working memory tasks. A trial consisted a Sternberg test with a duration of 2s recognition, 2s retention, 2s retrieval and 500ms fixation. This test was repeated 40 times with similar patterns in each level group.

Figure 4. Participants: a. Setting up fNIRS optodes; b. Performing Sternberg working memory test

Figure 5. fNIRS signal pre-processing example of participant 4: (a) raw intensity signals; (b) ΔHbO2 after pre-processing (noise and motor artefacts removed).(c) summed signals

Figure 6. The averaged hemodynamic responses (ΔHbO2) over subjects and channels under level 1-3. Greyed areas indicate 95% C.I.

Figure 7. Examples of paired features (ROI-right). (a) Individual data (Participant 4); (b) group data.

Figure 8. Examples of Classifier confusion matrices of (a) individual level (participants 4) and (b) group level.

Figure 9. An example of one interaction of feature selection.