Caitlin McElwee
August 2017
AIR Lab at Drexel University and Distributed Research Experience for Undergraduates

Using Machine Learning on fNIRS Data to Detect Changes in Working Memory Load

Abstract

Working memory is the ability for a person to receive, retain, and act on information gained quickly and used after some delay; it is the brain's ability to juggle information for short intervals. The fNIRS (functional Near-InfraRed Spectroscopy) system can be used to measure brain activity during everyday tasks (e.g using a computer) without keeping the participant from normal behavior. A 16-channel fNIRS system was used to measure the hemodynamics in participants' brains while performing the n-back task, an abstract activity applying varying loads on working memory. Only data from pilot studies were collected; the goal of future data collection is to build a machine learning model that can automatically detect the level of working memory demand an individual is experiencing. This model will enable future interactive systems to adapt behavior to better support the user's changing cognitive state.

Introduction and Theory

Working memory is the ability for a person to receive, retain, and act on information gained quickly; it is the brain's ability to function on information given in real time. Studies have determined certain areas of the brain associated with the use of working memory: prefrontal cortex [1].

It is common among brain studies to use PET, MRI, and fMRI systems to record data. However, a major issue with these systems is their bulk. The size and requirements of such systems prevent the measuring of everyday activities like driving or using a computer. Functional near-infrared spectroscopy, or fNIRS, has been shown to measure the same brain events as aforementioned systems but in a much less intrusive form factor, allowing the fNIRS system to be used to measure everyday tasks like those listed above without preventing the participant from normal behavior [2]. An fNIRS system composes of fiber optic wires sending light at near-infrared wavelengths into a person's forehead which change depending on the hemodynamics in the brain. The light bounces back to detectors which measure that change in the light received, allowing the system to detect areas of high and low activity in the brain. This study utilizes an fNIRS system to record brain data while participants perform the n-back task with a computer program.

The n-back task is an abstract task created to isolate working memory from other brain activity present in more realistic but complex activities [3]. In all versions of the task used in this study, the participant is shown ten single-digit numbers. Referred to as stimuli, these numbers are shown in series, on display for 1.0 seconds followed by a blank screen for 1.5 seconds. During that 2.5s window for each stimulus, the participant must decide if the current stimulus is the same as the relevant stimulus. The relevant stimulus depends on what level of n-back the current task is: for 0-back, it is the current stimulus; for 1-back, it is the previously shown stimulus; for 2-back, it not the previous stimulus but the stimulus before that one. For 1-back tasks the first stimulus does not have an answer, and the same for the first two stimuli for 2-back tasks, since those stimuli would rely on stimuli that do not exist.

Related Work

Machine Learning and fNIRS Brain Sensors Systems

More than one study has explored attempting to analyze fNIRS data with machine learning algorithms [2,4,5]. Some succeeded in performing good analysis, resulting in accuracy ratings of upwards of 80% and 90%. These studies showed machine learning algorithms not only can successfully classify fNIRS data but also they can do so with high degrees of accuracy. My study will extend the work from these papers by generating a model trained on fNIRS data for the purpose of passing it on to other projects.

N-Back Tasks with fNIRS Brain Sensors Systems

Multiple studies have measured people's brains while they perform variations of n-back tasks [2,3]. Izzetoglu et Al. showed evidence a fNIRS system can detect varying working memory load while Solovey et Al. focused on using machine learning classification algorithms to detect the state of a person's brain at a given time. Together the studies show yes, working memory load can be measured with fNIRS reliably. My study will extend the work from these papers by independently verifying fNIRS systems can measure working memory load and by training a model using machine learning to detect working memory load in real time.

Approach

Because only pilot studies with lab-associated personnel were performed, no actual experiments were run. This section described what would have been performed had real participants been recruited. For the pilot study, one person (a colleague in the lab) was measured.

When the participant came into the experiment space, they were instructed on what the n-back task is and how to perform each level of the task they were to perform that day. A paper worksheet and practice program were used to ensure the participant was well-versed in the performance of all three levels of n-back task before any brain data was collected. The practice program appeared and its tasks were performed identically to the actual performance program.

Figure 1: The fNIRS sensors are held to a person's forehead with headbands.

The fNIRS sensors were set up on the participant after the practice program was completed. The sensors themselves are held to the participant's forehead with headbands, layered to keep out as much external light as possible (see Figure 1). When the sensors were ready, a physical space tracker was used to mark specific points on the participant's head and then on the corners of each silicon block housing the fNIRS detectors/emitters so as to later on be able to establish what parts of the brain were measured. After the fNIRS system was fully set up and was properly recording data, the participant was presented with the performance program and were instructed to begin. The participant's experience of the program consisted of a black screen with white text (see Figure 2). All text presented after this screen is read aloud by sound files generated by a Text-To-Speech program. The welcome screen was followed by two pages of instruction, read aloud. When the instructions were finished being read, the program was cued to move on to present the n-back tasks. The tasks, ranging from 0-back to 2-back, were presented in a random order and were generated randomly by the program every run; no two participants received the same tasks nor the same order of tasks. Between tasks a short rest was given to the participant, them being allowed ten seconds before the next task begins. Each task starts with a task's instructions being presented visually and audibly to the participant, then the task is presented (see Figure 3). A "yes, the two items match" answer is given by pressing the 'A' key on the keyboard and a "no, the items do not match" answer is given by pressing the 'L' key on the keyboard. After every six tasks, the participant was given a long rest lasting thirty seconds.

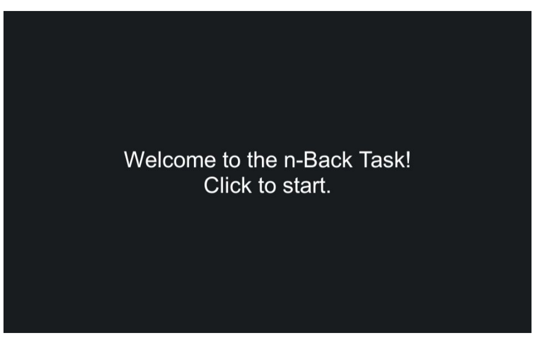

Figure 2: The program opens with a welcome screen, shown here.

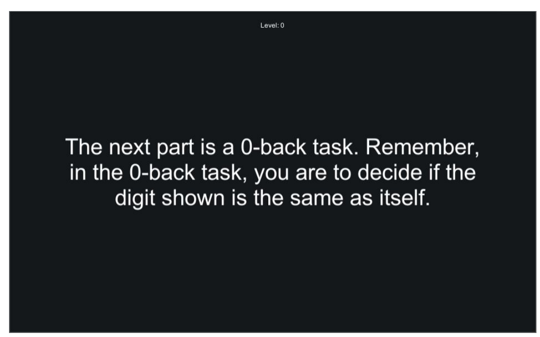

Figure 3: This screen is shown before the program performs a 0-back task.

When the start or end of a task or stimulus was reached, the program sent a flag to the data acquisition software containing information relevant to the event such as what level this new task is, whether the stimulus was answered correctly or not, how many out of how many possible answers were correct, and so on. These flags were used when parsing the data acquisition software's output.

Analysis

Data collected by the acquisition software was exported to a database. The data was then loaded into a Matlab program which, using flags to separate out the tasks as individual events, exported .csv files containing the brain data for each event for each task for each data-collecting channel (six total). The .csv files were manually converted to .arff type files, a file designed for the Weka machine learning program, to be processed by Weka's built-in collection of machine learning algorithms. Weka is an open-source Java program for data mining and machine learning from the University of Waikato[6,7].

Initially it was attempted to use all data points for each channel (195 entries per channel with six channels) but it proved too large for the spreadsheet program used; the features chosen for analysis were the minimum value, the maximum value, the average, the median, and the standard deviation of all data points for each channel for each event (five entries per channel with six channels). These datasets were imported to Weka; they were then to be used to train a machine learning model using Weka's "AttributeSelectedClassifier" algorithm and cross-validation with ten folds.

Results

Due to not receiving Internal Review Board approval, only pilot studies were conducted for this study. Preliminary data collected in pilot studies suggests fNIRS can reliably detect working memory load in users with a rate of ~58% accuracy using the AttributeSelectedClassifier algorithm.

Conclusions

Without having obtained official results, it cannot be determined whether this study reached a solid conclusion. However, the future work based on the ground plan set down by this study can be discussed. This study is a part of a family of similar studies, each researching the effectiveness of fNIRS at detecting changing brain states as a person performs some abstract task designed to single out one specific activity - mind wandering, learning a rule, and so on. These studies collectively are intended to provide foundation for future software referred to as an "intelligent tutoring system". An Intelligent Tutoring System (ITS) is a theoretical program or software package which adjusts to a user to become the perfect teaching tool for that user, like some sort of Digital Aristotle for a modern-day Alexander the Great [8]. If a model trained to accurately detect low, medium, and high levels of working memory was created it would be used by an ITS to monitor the level of engagement and difficulty the user was experiencing. When an ITS knows a user is too challenged or not challenged enough, it will know to change the difficulty to be right for the user. According to Izzetoglu et al. it would also be able to detect when the user gives up (a sharp decline in working memory use) [3]; when an ITS detects such an event, it would know to give the user a break or some help before stepping down the difficulty to make sure they stay engaged.

Future work would include performing experiments as described in this paper, as well as conducting other studies in this family of studies.

References

[1] Smith, E.E. and Jonides, J., 1997. *Working memory: A view from neuroimaging.* Cognitive psychology, 33(1), pp.5-42.

[2] Solovey, E. T., Zec, M., Garcia Perez, E. A., Reimer, B., & Mehler, B. , 2014. *Classifying driver workload using physiological and driving performance data: Two field studies*. Proceedings of the 32nd annual ACM conference on Human factors in computing systems, 4057-4066.

[3] Izzetoglu, M., Bunce, S.C., Izzetoglu, K., Onaral, B. and Pourrezaei, K., 2007. *Functional brain imaging using near-infrared technology.* IEEE Engineering in Medicine and Biology Magazine, 26(4), p.38.

[4] Abibullaev, B. and An, J., 2012. Classification of frontal cortex hemodynamic responses during cognitive tasks using wavelet transforms and machine learning algorithms. Medical engineering & physics, 34(10), pp.1394-1410.

[5] Hennrich, J., Herff, C., Heger, D. and Schultz, T., 2015, August. *Investigating deep learning for fNIRS based BCI*. In Engineering in Medicine and Biology Society (EMBC), 2015 *37th Annual International Conference of the IEEE (pp. 2844-2847)*. IEEE.

[6] Witten, I.H., Frank, E., Hall, M.A. and Pal, C.J., 2016. *Data Mining: Practical machine learning tools and techniques.* Morgan Kaufmann.

[7] Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P. and Witten, I.H., 2009. The WEKA data mining software: an update. *ACM SIGKDD explorations newsletter, 11(1), pp.10-18.*

[8] CGP Grey. (2017). *Digital Aristotle: Thoughts on the Future of Education*. [Online Video]. 5 November 2012. Available from: https://www.youtube.com/watch?v=7vsCAM17O-M..