

Multi-Class Classification of Alzheimer's Disease Using Frontal Cortex Non-invasive fNIRS



Thi Kieu Khanh Ho, Minhee Kim, Harish Garg, Byeong C. Kim, Jang Jae Lee, Kyu Yeong Choi, Jae Gwan Kim, Kun Ho Lee, Jong-In Song, and Jeonghwan Gwak

Abstract Early and accurate diagnosis of Alzheimer's disease (AD) is especially important for neurodegenerative disorders allowing patients, who exhibit different patterns of severity and progression risks, to take prevention and treatments before brain damages are shaped. Functional Near-Infrared Spectroscopy (fNIRS), as a clinical test, can be employed in an ecological setting to support AD diagnosis at early stages. In this study, we aimed to comprehensively study the hemodynamic response patterns among four subject groups from 140 participants by visualizing the 4-class distributions based on t-SNE visualization techniques and showing the fNIRS topographic mapping to have a pictorial representation on the surface of the prefrontal cortex region. We demonstrated that the changes in hemodynamic concentrations of oxygenated hemoglobin (HbO) and deoxygenated hemoglobin (Hb) from healthy

T. K. K. Ho · J.-I. Song

School of Electrical Engineering and Computer Science, Gwangju Institute of Science and Technology, Gwangju 61005, Korea

M. Kim · J. G. Kim

Department of Biomedical Science and Engineering, Institute of Science and Technology, Gwangju 61005, Korea

H. Garg

School of Mathematics, Thapar Institute of Engineering and Technology, Deemed University, Patiala, Punjab 147004, India

B. C. Kim

Department of Neurology, Chonnam National University Medical School, Gwangju 61469, Korea

J. J. Lee · K. Y. Choi · K. H. Lee

Gwangju Alzheimer's Disease and Related Dementias (GARD) Cohort Research Center, Chosun University, Gwangju 61452, Korea

K. H. Lee

Department of Biomedical Science, Chosun University, Gwangju 61452, Korea

Aging Neuroscience Research Group, Korea Brain Research Institute, Daegu 41062, Korea

J. Gwak (✉)

Department of Software, Korea National University of Transportation, Chungju 27469, Korea

© Springer Nature Switzerland AG 2022

V. Van Toi et al. (eds.), *8th International Conference on the Development of Biomedical Engineering in Vietnam*, IFMBE Proceedings 85, https://doi.org/10.1007/978-3-030-75506-5_75

955

controls to three stages of AD did exist during four stages of our experimental protocols. Results also revealed that various machine learning techniques were capable of both analyzing AD patterns and its related features and classifying multiple pathophysiological mechanisms with the highest accuracies of $74.39 \pm 4.7\%$ and $84.91 \pm 4.01\%$ corresponding to raw and SMOTE-based features, respectively. These findings demonstrate the feasibility of fNIRS-based approaches to investigate the stages of AD progression.

Keywords fNIRS · Alzheimer's disease · Machine learning · Multi-class classification

1 Introduction

Alzheimer's disease (AD), the most common cause of dementia accounting for 60–80% cases worldwide [1], is an irreversible and chronic progressive neurodegenerative brain disease that is characterized by progressive impairments in the cognitive abilities and memory [2]. By a significant loss of neurons in the neurotransmitter systems, the present proteins such as amyloid-beta deposits and neurofibrillary tangles may disrupt the communication with nerve cells, damage cells and lead to the development of AD [3]. AD patients typically suffer from several symptoms of functional impairments, including memory, communication, reasoning, and behavior functions which interfere with them in daily life while there are currently no medication therapies or treatments being able to halt the disease's progression [4]. Hence, a significant effort has been made to develop the strategies for early detection, particularly at a presymptomatic stage, with the hope that the intervention can delay or even prevent those clinical symptoms.

Since the diagnosis of AD primarily relies on patients' medical history, neuropsychological examinations or clinical rating scores which require experienced clinicians and exhaustive clinical tests, it has been upgraded by adopting brain imaging techniques such as magnetic resonance imaging (MRI), single-photon emission computed tomography (SPECT), positron emission tomography (PET) and electroencephalography (EEG) introduced in recent years. Functional brain imaging tests are modified from a non-paced stage of administration and free doctor-patient interaction tasks to a paradigm event where an automated stimulation can replace doctors [5–7]. Despite the usefulness of these modifications from a neuroimaging standpoint, they still alter the ecology of the interaction between doctors and patients. With this in mind, Functional Near-Infrared Spectroscopy (fNIRS), which is a scalp-located non-invasive method recording neural activity and oscillation within the oxygenated (HbO) and deoxygenated (Hb) hemoglobin in the brain based on the Blood Oxygen Level Dependent (BOLD) effect [8], is one of the most suitable neuroimaging techniques that are feasible for an outpatient environment.

fNIRS offers various benefits over the aforementioned techniques by its higher temporal resolution, high portability, relatively low cost, lightweight measurement, lower susceptibility to motion artifacts, lack of ionizing radiation, and fewer constraints to the subjects during experiments. It can integrate with a mechanical structure resembling EEG [9] thus being appropriate for measurements during clinical administration tests. Furthermore, several studies have examined the feasibility of fNIRS to differentiate the hemodynamic responses between healthy controls and AD patients [10]. fNIRS is also able to successfully monitor AD treatment in the clinic which was reported by [11] showing a significant difference in cerebral blood flows and the effect of memantine on AD patients. These investigations showed that AD patients had lower levels of activation at specific regions in the brain compared to healthy controls during cognitive experiments. However, although these previous studies denote the potential of fNIRS on measuring the differences between healthy and AD groups, there are still various AD stages remaining unknown. It requires the recruitment of several participants at different stages of AD to comprehensively compare and evaluate the efficiency of fNIRS in therapeutic monitoring and diagnosis. Besides, the pathological mechanism of AD progression has not yet been thoroughly documented and investigated, which is expected to have an intact experimental design dealing with multiple subject groups.

To address the challenges, we comprehensively investigate the hemodynamic responses of healthy subjects and patients with three degrees of AD. In detail, the classification studies will be described, including seven machine learning methods that generally require four main components: feature extraction, feature selection, reduction of dimensionality, and processed feature-based classification algorithms. Due to our imbalanced datasets, a preprocessing step called Synthetic Minority Over-sampling Technique (SMOTE) [12] was used to improve multi-class classification accuracies.

2 Materials and Methods

2.1 Participants and Data Preprocessing

In this study, 140 subjects dwelling in Gwangju city in South Korea and the adjacent cities were recruited from the Chonnam National University Hospital and the National Research Center for Dementia (Gwangju, South Korea). A set of medical examinations—Mini-Mental State Examination (MMSE), MRI, PET, and individual interview—was conducted to adequately analyze and diagnose different AD stages. Four categories were subsequently divided among them: HC class: Normal AD biomarkers, cognitively unimpaired (72.7 ± 5.3 years, 21 M/32F), asymptomatic AD (aAD) class: Very mild MCI with the cognitive decline on memory and executive functions (74.5 ± 4.3 years, 15 M/13F), prodromal AD (pAD) class: Explicit brain dysfunction symptoms (75.8 ± 3.9 years, 33 M/17F), AD Dementia (ADD) class:

Severe deterioration of memory, language, and social abilities (75.4 ± 6.8 years, 4 M/5F). The subject with a mental and behavioral disorder was disregarded in this cohort. All subjects underwent our experimental protocol without any previous experience. No subject had any previous experience with our experimental protocol. Each subject was fully informed of the purpose of research and consent forms prior to conducting experiments. Table 1 summarizes the demographic information of all subjects (SD: Standard Deviation).

To curtail any environmental disturbance, the experiments were carried out in a confined room. Subjects were requested to seat in a chair and rest by calmly watching a white cross that occurred on the monitor screen during resting periods. Then, they underwent a series of three tasks: (i) Oddball which is a cognitive ability test, (ii) 1-back which is a memory ability test, and (iii) Verbal fluency which is a language ability test. Six fNIRS channels measuring in the frontal cortex area were recorded. Each channel was visually inspected, and channels with large spikes were marked as noises and excluded from our studies. The concentration changes in hemoglobin, HbO, Hb, and total hemoglobin (THb), were calculated based on the Modified Beer-Lambert Law [13]. A low-pass filter with a cut-off frequency of 0.5 Hz was used to remove artifacts. Figure 1 shows the imbalanced datasets and the complexity of distributions

Table 1 Participant Information

	HC	aAD	pAD	ADD
Number of subjects	53	28	50	9
Age \pm SD	72.7 ± 5.3	74.5 ± 4.3	75.8 ± 3.9	75.4 ± 6.8
Gender (M/F)	21/32	15/13	31/17	4/5
MMSE \pm SD	27.0 ± 4.2	26.9 ± 2.5	26.0 ± 3.2	20.2 ± 4.8
Education \pm SD	9.8 ± 4.7	10.2 ± 5.2	10.6 ± 5.2	8.5 ± 5.3

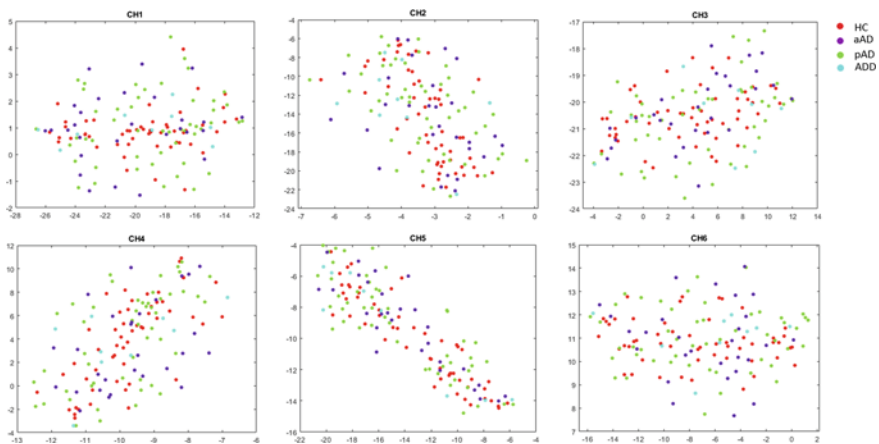


Fig. 1 t-SNE Visualization from 6 channels of HbO signals

among the four classes (i.e., HC, aAD, pAD and ADD) from HbO signals by using the t-Stochastic Neighbor Embedding (t-SNE) visualization technique.

2.2 Experimental Protocol

Figure 2a represents the fNIRS device setup. The transmitter (LED), the detector (photodiode), and the channel (CH) are denoted as the red circle, the black circle, and the green rectangular, respectively. The distance between the transmitter and the detector is 30 mm. Figure 2b displays the experimental protocol, and there are four main stages in our experiments. After finishing recording the fNIRS data, we segmented trial sections corresponding to the experimental stages. Each section was segmented by 60 s. To sum up all four sections, we had a resting section (60 s), the Oddball stage (300 s), the 1-back stage (270 s), and the Verbal stage (390 s). Therefore, the total experimental time for one trial was 1020 s.

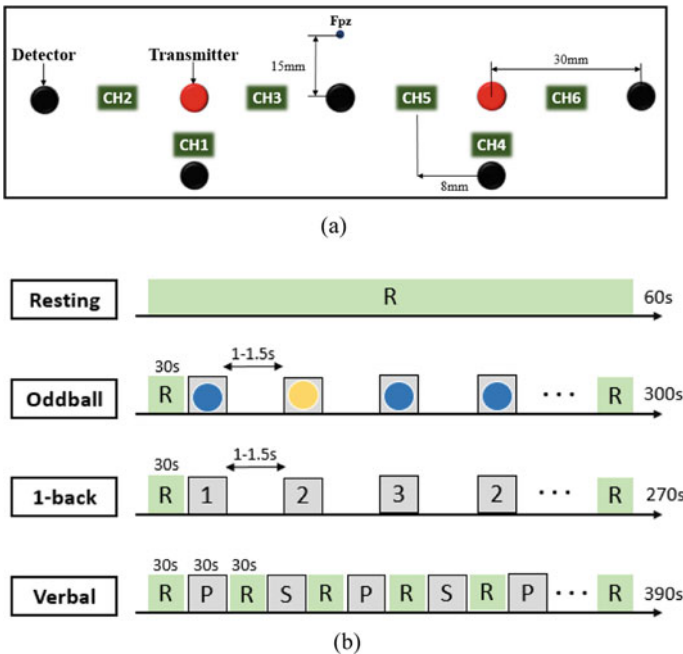


Fig. 2 a fNIRS device setup; b experimental protocol (R: Resting, P: phonemic, S: semantic)

2.3 Classification Algorithms

The following seven supervised machine learning classifiers was selected to classify the four categories:

- Gaussian Discriminant Analysis (GDA) [14]: With a set of samples (each belongs to each class), the intra-class and inter-class matrices were calculated, and the linear transformation was obtained based on solving the generalized eigenvalues. Then, the classification task was performed on the transformed space using Euclidean distance after 200 epochs.
- K-Nearest Neighbor (KNN) [15]: $K = 4$ was chosen as the most optimal value to reduce errors and allow the algorithm to calculate the distance between each data point and the cluster centroid by Euclidean distance. KNN was trained with 200 epochs.
- Gaussian Naïve Bayes (GNB) [16]: The conditional probability for each class was computed based on the prior and the posterior probability. The class showing the highest probability was considered as the final predicted class.
- Support Vector Machine (SVM) [17]: The decision hyperplane was constructed when the margin of the classifier was maximized from support vectors to separate four classes. A sigmoid kernel was specified for the non-linearly separation problem. SVM would stop training whenever the margin error was trivial, and no further optimization is needed.
- Adaptive Boosting (AdaBoost) [18]: A classifier was built by combining 1000 weak classifiers, and the number of estimators is 200. During training, if any misclassified point was found, the weight of that point was boosted. The new weight was updated to the next classifier, and the procedure was repeated.
- Neural Network (NN) [19]: Along with an input layer and an output layer (4 neurons), NN with different settings of hyperparameters was trained until the best optimal set of hidden neurons was obtained (100–80–50). A ReLU activation function, $1e-5$ learning rate, and 300 epochs were manually chosen to achieve the highest accuracy.
- Random Forest (RF) [20]: A decision tree was constructed for every sample and generated output. RF consisted of multiple trees to vote for each predicted output and select the output with the most votes as the final prediction. RF was trained several times with the max depth of the tree as 8.

For a fair comparison, fivefold cross-validation was used for each classifier. The SMOTE step was also applied by over-sampling the minority classes (it drew a line between a set of neighboring samples in the feature space and over-sampled new synthetic points along the line) to improve accuracies in the latter part.

3 Experimental Results

The distributions of four classes based on t-SNE visualization techniques are depicted in Fig. 1. To have an apparently pictorial representation on the surface of the prefrontal cortex, we showed the fNIRS topographic mapping in Fig. 3 taken from the resting and three ability tests. The white digits being shown in Fig. 3 represent the channel numbers and their positions on the prefrontal cortex region. The bar graph on the right side in Fig. 3 denotes the fNIRS signal intensity. It was observed that channels 1, 2, and 3 had a significant impact on HC during Oddball and 1-back and aAD patients during Oddball and pAD during Verbal. Meanwhile, the activation spots located at channels 4, 5, and 6 were observed from pAD during 1-back and ADD patients during Verbal. The local activation regions in the same experimental task were totally opposite. For example, channels 4, 5, and 6 were highlighted in HC whilst not being spotted in aAD during the Oddball stage, or channels 1, 2, and 3 were dominantly punctuated in HC while those channels are inactivated in pAD. It induces that each channel owned peculiar characteristics and behaved differently counting on corresponding experimental tasks; we thereby utilized all channels for the classification tasks.

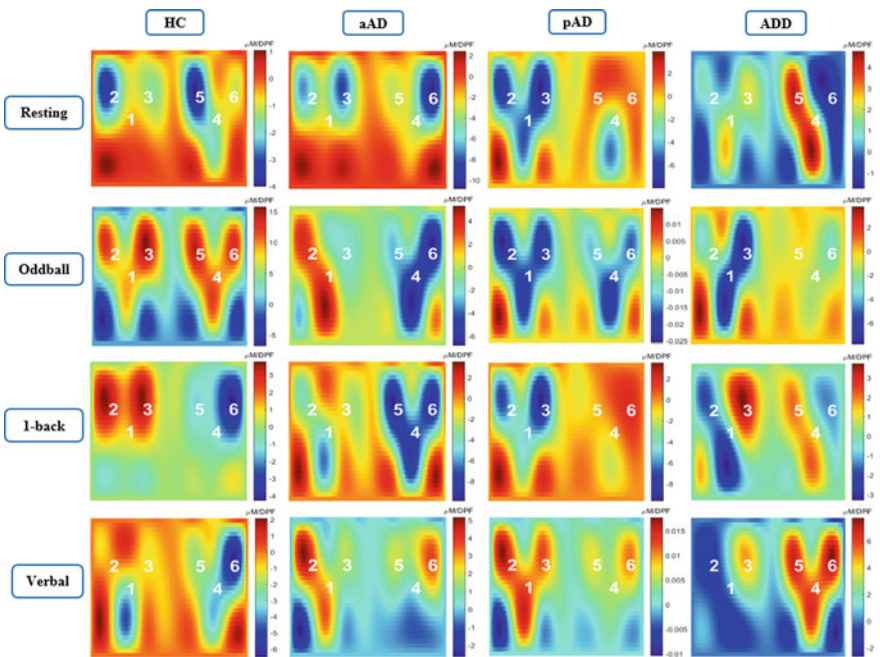


Fig. 3 Comparison of activation maps (topographic mapping) using HbO based on four experimental stages measuring from HC subjects and three AD-degree patients during four experimental stages

Regarding classification tasks, we evaluated the performances of different classifiers on each experimental task dealing with the original features from HbO, Hb, and THb. In general, all classifiers performed better on the THb input compared to the other hemoglobin types. In addition, NN classifiers with the accuracy of $74.39 \pm 4.7\%$ outperformed the others, followed by SVM, GDA, and RF (see Fig. 4). Compared to the other classifiers which require to identify features, break down to different parts, and recombine in the final stages, NN was more flexible since it adopted a sufficient amount of our original data size without any feature engineering steps and learned the high-level features in an end-to-end manner. The remaining classifiers obtained low accuracies. This indicates that the original data was not easily separated using a convoluted decision boundary in KNN, a probabilistic approach from GNB, or boosting algorithms from AdaBoost.

Due to the imbalance of our datasets that patients with three stages of AD were difficult to recruit, SMOTE was applied to magnify the number of small aAD samples and extremely small ADD samples to be balanced with two dominant HC and pAD classes. As expected, the mean classification accuracies considerably improved up to $84.91 \pm 4.01\%$, and the efficacy of each classifier on the resampled data behaved similarly as it applied on the original dataset (see Fig. 5).

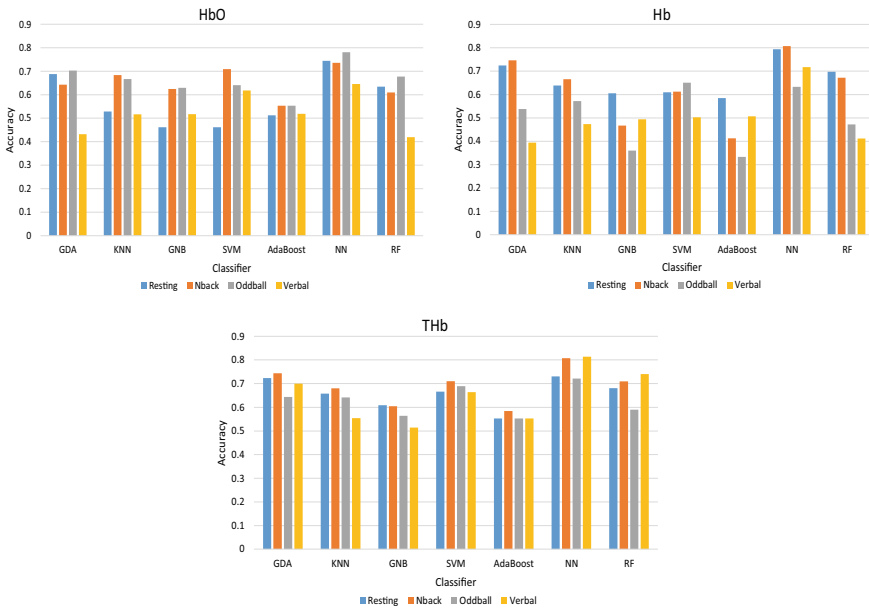


Fig. 4 Classification accuracies of 7 classifiers from three types of hemoglobin when the original data was used

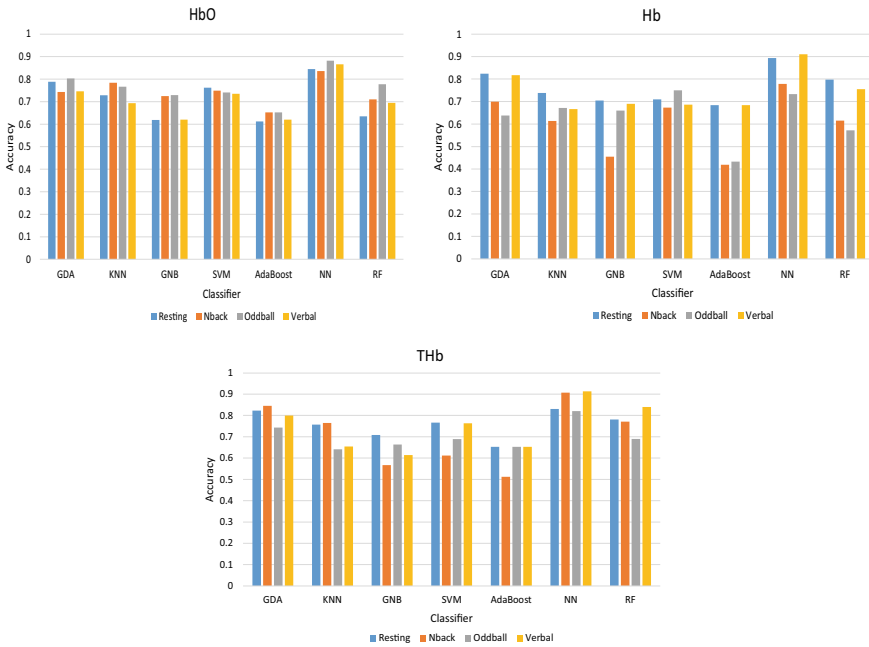


Fig. 5 Classification accuracies of 7 classifiers from three types of hemoglobin when SMOTE was applied

4 Conclusion and Future Works

Emerging evidence yields that exploring alterations of AD’s progress has great importance in understanding and timely treatments of this cognitive deficit disease. In this study, we aimed to access the capability of employing fNIRS as a clinical test being commonly utilized to diagnose AD. We demonstrated that the changes in hemodynamic concentrations from healthy controls to the three stages of AD did exist. We evaluated and optimized the most representative features from three types of hemoglobin signals recorded from four stages of experiments. Thus, a set of machine learning classifiers were able to inexpensively and rapidly classify HC with three stages of AD patients and further supplement the diagnosis of the degree of dementia in AD. Due to the relatively small sample size of AD patients recruited, which could easily result in misclassification performances, a larger cohort should be carried out to validate our present findings. In addition, further advancement of non-linear classifiers, such as deep learning techniques is an indispensable topic to address the challenging fNIRS multi-class classification problems.

Acknowledgements This research was supported by the Brain Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT & Future Planning (NRF-2016M3C7A1905475, NRF-2019M3C7A1020406), and KBRI basic

research program through Korea Brain Research Institute funded by Ministry of Science and ICT (20-BR-03-02).

Conflicts of Interest The authors have no conflicts of interest to declare.

References

1. Alzheimer's Association (2019) 2019 Alzheimer's disease facts and figures. *Alzheimer's Dementia* 15(3):321–387
2. Veitch DP et al (2019) Understanding disease progression and improving Alzheimer's disease clinical trials: recent highlights from the Alzheimer's disease neuroimaging initiative. *Alzheimer's Dementia* 15(1):106–152
3. Lowe VJ et al (2019) Neuroimaging correlates with neuropathologic schemes in neurodegenerative disease. *Alzheimer's Dementia* 15(7):927–939
4. Ibrahim MM, Gabr MT (2019) Multitarget therapeutic strategies for Alzheimer's disease. *Neural Regen Res* 14(3):437–440
5. Tian F, Yennu A, Smith-Osborne A, Gonzalez-Lima F, North CS, Liu H (2014) Prefrontal responses to digit span memory phases in patients with post-traumatic stress disorder (PTSD): a functional near infrared spectroscopy study. *NeuroImage: Clin* 4:808–819
6. Lancia S, Cofini V, Carrieri M, Ferrari M, Quaresima V (2018) Are ventrolateral and dorsolateral prefrontal cortices involved in the computerized Corsi block-tapping test execution? An fNIRS study. *Neurophotonics* 5(1):11–19
7. Ho TKK, Gwak J, Park CM, Song JI (2019) Discrimination of mental workload levels from multi-channel fNIRS using deep learning-based approaches. *IEEE Access* 7:24392–24403
8. Ferrari M, Quaresima V (2012) A brief review on the history of human functional near-infrared spectroscopy (fNIRS) development and fields of application. *Neuroimage* 63(2):921–935
9. Fazli S, Mehnert J, Steinbrink J, Curio G, Villringer A, Müller KR, Blankertz B (2012) Enhanced performance by a hybrid NIRS–EEG brain computer interface. *Neuroimage* 59:519–529
10. Li R, Nguyen T, Potter T, Zhang T (2019) Dynamic cortical connectivity alterations associated with Alzheimer's disease: An EEG and fNIRS integration study. *NeuroImage: Clin* 21:101622:1–11
11. Araki T et al (2014) The effects of combine treatment of memantine and donepezil on Alzheimer's disease patients and its relationship with cerebral blood flow in the prefrontal area. *Int J Geriatric Psychiatry* 29(9):881–889
12. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP (2002) SMOTE: synthetic minority over-sampling technique. *J Artif Intell Res* 16:321–357
13. Scholkmann F, Kleiser S, Metz AJ, Zimmermann R, Pavia JM, Wolf U, Wolf M (2014) A review on continuous wave functional near-infrared spectroscopy and imaging instrumentation and methodology. *Neuroimage* 85:6–27
14. Ju J, Kolaczyk ED, Gopal S (2003) Gaussian mixture discriminant analysis and sub-pixel land cover characterization in remote sensing. *Remote Sens Environ* 84(4):550–560
15. Weinberger KQ, Saul LK (2009) Distance metric learning for large margin nearest neighbor classification. *J Mach Learn Res* 10(2009):207–244
16. Murphy KP (2006) Naive Bayes Classifiers. University of British Columbia, Department of Computer Science
17. Fung GM, Mangasarian OL (2005) Multicategory proximal support vector machine classifiers. *Mach Learn* 59(1):77–97
18. Sun Y, Liu Z, Todorovic S, Li J (2007) Adaptive boosting for SAR automatic target recognition. *IEEE Trans Aerosp Electron Syst* 43(1):112–125

19. Misra BB, Dehuri S (2007) Functional link artificial neural network for classification task in data mining. *J Comput Sci* 3(12):948–955
20. Pal M (2005) Random forest classifier for remote sensing classification. *Int J Remote Sens* 26(1):217–222